Metastability and Time Scales for Parabolic Equations with Drift 1: The First Time Scale

CLAUDIO LANDIM, JUNGKYOUNG LEE & INSUK SEO

Communicated by M. HAIRER

Abstract

Consider the elliptic operator given by

$$\mathcal{L}_\varepsilon f = \boldsymbol{b} \cdot \nabla f + \varepsilon \Delta f \tag{0.1}$$

for some smooth vector field $\boldsymbol{b} \colon \mathbb{R}^d \to \mathbb{R}^d$ and a small parameter $\varepsilon > 0$. Consider the initial-valued problem

$$\begin{cases} \partial_t u_\varepsilon = \mathcal{L}_\varepsilon u_\varepsilon, \\ u_\varepsilon(0, \cdot) = u_0(\cdot), \end{cases} \tag{0.2}$$

for some bounded continuous function u_0. Denote by $\mathcal{M}_0$ the set of critical points of $\boldsymbol{b}$ which are stable stationary points for the ODE $\dot{\boldsymbol{x}}(t) = \boldsymbol{b}(\boldsymbol{x}(t))$. Under the hypothesis that $\mathcal{M}_0$ is finite and $\boldsymbol{b} = -(\nabla U + \boldsymbol{\ell})$, where $\boldsymbol{\ell}$ is a divergence-free field orthogonal to ∇U, the main result of this article states that there exist a time-scale $\theta_\varepsilon^{(1)}$, $\theta_\varepsilon^{(1)} \to \infty$ as $\varepsilon \to 0$, and a Markov semigroup $\{p_t : t \geq 0\}$ defined on $\mathcal{M}_0$ such that

$$\lim_{\varepsilon \to 0} u_\varepsilon(t\, \theta_\varepsilon^{(1)}, \boldsymbol{x}) = \sum_{\boldsymbol{m}' \in \mathcal{M}_0} p_t(\boldsymbol{m}, \boldsymbol{m}')\, u_0(\boldsymbol{m}')$$

for all $t > 0$ and $\boldsymbol{x}$ in the domain of attraction of $\boldsymbol{m}$ [for the ODE $\dot{\boldsymbol{x}}(t) = \boldsymbol{b}(\boldsymbol{x}(t))$]. The time scale $\theta^{(1)}$ is critical in the sense that, for all time scales ϱ_ε such that $\varrho_\varepsilon \to \infty$, $\varrho_\varepsilon / \theta_\varepsilon^{(1)} \to 0$,

$$\lim_{\varepsilon \to 0} u_\varepsilon(\varrho_\varepsilon, \boldsymbol{x}) = u_0(\boldsymbol{m})$$

for all $\boldsymbol{x} \in \mathcal{D}(\boldsymbol{m})$. Namely, $\theta_\varepsilon^{(1)}$ is the first scale at which the solution to the initial-valued problem starts to change. In a companion paper [20] we extend this result finding all critical time-scales at which the solution of the initial-valued problem (0.2) evolves smoothly in time and we show that the solution u_ε is expressed in terms of the semigroup of some Markov chain taking values in sets formed by unions of critical points of $\boldsymbol{b}$.

1. Introduction

The main concern of the current article is the behavior of the solution u_ε of the equation (0.2) in the regime $\varepsilon \to 0$. This problem is connected to the metastable behavior of the diffusion process induced by the generator $\mathcal{L}_\varepsilon$ given in (0.1), which has been a serious issue in the probability community. Freidlin and Koralov [11,12] found a critical depth $D > 0$ and showed that the the solution $u_\varepsilon(t, x)$ in the interval $t \in [0, e^{(D-\eta)/\varepsilon}]$ and $t \in [e^{(D+\eta)/\varepsilon}, \infty)$ differ significantly for all $\eta > 0$. Therefore, a dramatic phase transition occurs at the scale $\theta_\varepsilon = e^{D/\varepsilon}$. This result has been extended by Koralov and Tcheuko [18] to cases which exhibit multiple metastable time-scales. Ishii and Souganidis [15,16] derived similar results with purely analytical methods.

In this article and the companion paper [20], we characterize the solution u_ε under the assumption that the diffusion process induced by the generator $\mathcal{L}_\varepsilon$ has a Gibbs invariant measure. More precisely, fix a smooth potential $U \colon \mathbb{R}^d \to \mathbb{R}$, and a smooth vector field $\boldsymbol{\ell} \colon \mathbb{R}^d \to \mathbb{R}$. Assume that the vector field $\boldsymbol{\ell}$ is divergence free and is orthogonal to the gradient of U:

$$(\nabla \cdot \boldsymbol{\ell})(\boldsymbol{x}) = 0, \quad (\nabla U)(\boldsymbol{x}) \cdot \boldsymbol{\ell}(\boldsymbol{x}) = 0, \quad \boldsymbol{x} \in \mathbb{R}^d. \tag{1.1}$$

For $\varepsilon > 0$, denote by $\mathcal{L}_\varepsilon$ the elliptic operator given by

$$\mathcal{L}_\varepsilon f = -(\nabla U + \boldsymbol{\ell}) \cdot \nabla f + \varepsilon \Delta f, \quad f \in C^2(\mathbb{R}^d), \tag{1.2}$$

which corresponds to the operator (0.1) when $\boldsymbol{b} = -(\nabla U + \boldsymbol{\ell})$. It has been shown in [28] that the Gibbs measure $\mu_\varepsilon(\mathrm{d}\boldsymbol{x}) = (1/Z_\varepsilon) \exp\{U(\boldsymbol{x})/\varepsilon\} \, \mathrm{d}\boldsymbol{x}$ is an invariant measure for the diffusion induced by the generator $\mathcal{L}_\varepsilon$ for all $\varepsilon > 0$ if and only if $\boldsymbol{\ell}$ satisfies conditions (1.1).

Denote by $\mathcal{L}_\varepsilon$ the generator (1.2), unless otherwise specified. Fix a bounded and continuous function $u_0 \colon \mathbb{R}^d \to \mathbb{R}$ and consider the initial-valued problem

$$\begin{cases} \partial_t u_\varepsilon = \mathcal{L}_\varepsilon u_\varepsilon, \\ u_\varepsilon(0, \cdot) = u_0(\cdot). \end{cases} \tag{1.3}$$

The tools developed in [2,3,21–23] permit to describe the solution of the parabolic equation (1.3), in the domain of attraction of a local minimum $\boldsymbol{m}$, at the time-scale in which the solution is transformed from the value of the initial condition $u_0(\cdot)$ at the local attractor $\boldsymbol{m}$ of the field $\boldsymbol{b} = -(\nabla U + \boldsymbol{\ell})$ to a convex combination of the initial condition calculated at several different local attractors $\boldsymbol{m}'$. A similar result appeared in [5] for the case of sequences of continuous-time Markov chains on a fixed finite state space.

The First Critical Time Scale

Let us now explain our main result in more detail. For two positive sequences $(\alpha_\varepsilon : \varepsilon > 0)$, $(\beta_\varepsilon : \varepsilon > 0)$, we denote by $\alpha_\varepsilon \prec \beta_\varepsilon$, $\beta_\varepsilon \succ \alpha_\varepsilon$ if $\alpha_\varepsilon / \beta_\varepsilon \to 0$ as $\varepsilon \to 0$. The main results of the current article and the companion paper [20]

assert that there exist critical times scales $\theta_\varepsilon^{(1)} \prec \cdots \prec \theta_\varepsilon^{(\mathfrak{q})}$ associated with the potential function U at which the asymptotic behavior of the solution u_ε changes dramatically. We not only characterize these time scales explicitly but also provide precise asymptotics of u_ε along these scales. We also derive the asymptotics of the solution between these time-scales, completely analyzing the behavior of u_ε.

The current article concerns the first time-scale among such a complex multi-scale structure. We explicitly find a time-scale $\theta_\varepsilon^{(1)} \succ 1$, and a Markov semigroup $\{p_t : t \geq 0\}$ defined on the set of local minima $\mathcal{M}_0$ of U such that, for all local minimum $\boldsymbol{m} \in \mathcal{M}_0$,

$$\lim_{\varepsilon \to 0} u_\varepsilon(t\,\theta_\varepsilon^{(1)}, \boldsymbol{x}) \;=\; \sum_{\boldsymbol{m}' \in \mathcal{M}_0} p_t(\boldsymbol{m}, \boldsymbol{m}')\, u_0(\boldsymbol{m}'), \qquad (1.4)$$

for all $t > 0$, $\boldsymbol{x} \in \mathcal{D}(\boldsymbol{m})$. Here, $\mathcal{D}(\boldsymbol{m})$ represents the domain of attraction of $\boldsymbol{m}$ for the ODE $\dot{\boldsymbol{x}}(t) = \boldsymbol{b}(\boldsymbol{x}(t))$, where $\boldsymbol{b} = -(\nabla U + \boldsymbol{\ell})$.

We also show that, for any sequence $1 \prec \varrho_\varepsilon \prec \theta_\varepsilon^{(1)}$,

$$\lim_{\varepsilon \to 0} u_\varepsilon(\varrho_\varepsilon, \boldsymbol{x}) \;=\; u_0(\boldsymbol{m}) \qquad (1.5)$$

for all $\boldsymbol{x} \in \mathcal{D}(\boldsymbol{m})$. Hence, the solution does not change until the time-scale $\theta_\varepsilon^{(1)}$, and it starts to change exactly at $\theta_\varepsilon^{(1)}$ in view of (1.4) and (1.5). The main achievement of the current article is the verification of (1.4) and (1.5). We remark this scale $\theta_\varepsilon^{(1)}$ is the scale $\theta_\varepsilon = e^{D/\varepsilon}$ obtained in [11,12].

To illustrate assertion (1.4), consider a generic potential U, namely a potential for which all critical points are at different heights. Denote by $\boldsymbol{m}_1$ the local minimum associated to the shallowest valley and by $\boldsymbol{m}_2$ the unique local minimum separated from $\boldsymbol{m}_1$ by the shallowest saddle point disconnecting $\boldsymbol{m}_1$ from the other local minima. The local minimum $\boldsymbol{m}_2$ is unique because all saddle points are at different heights. Then $p_t(\boldsymbol{m}, \boldsymbol{m}') = 0$, unless $\boldsymbol{m} = \boldsymbol{m}_1$, $\boldsymbol{m}' = \boldsymbol{m}_2$. In other words, in the time-scale $\theta_\varepsilon^{(1)}$, starting from the basin of attraction of $\boldsymbol{m}_1$, the process waits an exponential time and then jumps to a neighbourhood of $\boldsymbol{m}_2$, where it stays forever.

Multi-scale Structure

The characterization of the remaining scales are the contents of the companion paper [20]. We briefly explain the main result.

Let us start from the second scale which can be inferred from (1.4). The theory of finite-state continuous-time Markov chains asserts that there exist probability measures $\pi_j^{(1)}$, $1 \leq j \leq \mathfrak{n}_1$, on $\mathcal{M}_0$ with disjoint supports, and probability measures $\omega^{(1)}(\boldsymbol{m}, \cdot)$, $\boldsymbol{m} \in \mathcal{M}_0$, on $\{1, \ldots, \mathfrak{n}_1\}$ such that

$$\lim_{t \to \infty} p_t(\boldsymbol{m}, \boldsymbol{m}') \;=\; \sum_{k=1}^{\mathfrak{n}_1} \omega^{(1)}(\boldsymbol{m}, k)\, \pi_k^{(1)}(\boldsymbol{m}') \qquad (1.6)$$

for all $\boldsymbol{m}, \boldsymbol{m}' \in \mathcal{M}_0$. If $\boldsymbol{m}'$ is a transient state all terms in the previous sum vanish. Indeed, the measures $\pi_j^{(1)}$ represent the stationary states of the Markov chain

restricted to the closed irreducible classes, which in turn are the support of the measures $\pi_j^{(1)}$. The weight $\omega^{(1)}(\boldsymbol{m}, k)$ stands for the probability that the Markov chain starting from $\boldsymbol{m}$ is absorbed at the support of the measure $\pi_k^{(1)}$.

If there is only one stationary state or, equivalently, one closed irreducible class, namely $\mathfrak{n}_1 = 1$, then for all time-scales ϱ_ε such that $\theta_\varepsilon^{(1)} \prec \varrho_\varepsilon$, we can readily guess from (1.4) and (1.6) that (note that $\omega^{(1)}(\boldsymbol{m}, 1) = 1$ in this case)

$$\lim_{\varepsilon \to 0} u_\varepsilon(\varrho_\varepsilon, \boldsymbol{x}) = \sum_{\boldsymbol{m}' \in \mathcal{M}_0} \pi_1^{(1)}(\boldsymbol{m}') \, u_0(\boldsymbol{m}')$$

for all local minimum $\boldsymbol{m} \in \mathcal{M}_0$ and $\boldsymbol{x} \in \mathcal{D}(\boldsymbol{m})$. Note that the limit does not depend on $\boldsymbol{m}$ or $\boldsymbol{x}$. This behavior occurs when (a) all the wells associated to local minima of U which are not global minima have the same depth, and (b) either there is only one global minimum or there is more than one and the depth of all wells are the same. In this case of a unique closed irreducible class, the support of the measure $\pi_1^{(1)}$ corresponds to the set of global minima of U. This finishes the description of multi-scale structure for the case $\mathfrak{n}_1 = 1$.

In contrast, if there are more than one closed irreducible classes, the limit of $u_\varepsilon(t\theta_\varepsilon^{(1)}, \boldsymbol{x})$ as $\varepsilon \to 0$ and then $t \to \infty$ depends on the local minimum attracting $\boldsymbol{x}$. In this case, there exists a second and longer time-scale $\theta_\varepsilon^{(2)}$ such that $\theta_\varepsilon^{(1)} \prec \theta_\varepsilon^{(2)}$ and a Markov semigroup $\{p_t^{(2)} : t \geq 0\}$ defined on the set of closed irreducible classes $\{1, \ldots, \mathfrak{n}_1\}$ obtained at the first time-scale such that

$$\lim_{\varepsilon \to 0} u_\varepsilon(t\,\theta_\varepsilon^{(2)}, \boldsymbol{x}) = \sum_{k=1}^{\mathfrak{n}_1} \omega^{(1)}(\boldsymbol{m}, k) \sum_{\ell=1}^{\mathfrak{n}_1} p_t^{(2)}(k, \ell) \sum_{\boldsymbol{m}' \in \mathcal{M}_0} \pi_\ell^{(1)}(\boldsymbol{m}') \, u_0(\boldsymbol{m}')$$

for all $t > 0$ where $\boldsymbol{m} \in \mathcal{M}_0$ is a local minium of U and $\boldsymbol{x} \in \mathcal{D}(\boldsymbol{m})$. Mind that we may restrict the sum over $\boldsymbol{m}'$ to local minima in the support of the measure $\pi_\ell^{(1)}$. We can also verify that, for any sequence ϱ_ε such that $\theta_\varepsilon^{(1)} \prec \varrho_\varepsilon \prec \theta_\varepsilon^{(2)}$, we have

$$\lim_{\varepsilon \to 0} u_\varepsilon(\varrho_\varepsilon, \boldsymbol{x}) = \sum_{k=1}^{\mathfrak{n}_1} \omega^{(1)}(\boldsymbol{m}, k) \sum_{\boldsymbol{m}' \in \mathcal{M}_0} \pi_k^{(1)}(\boldsymbol{m}') \, u_0(\boldsymbol{m}')$$

for all $\boldsymbol{m} \in \mathcal{M}_0$ and $\boldsymbol{x} \in \mathcal{D}(\boldsymbol{m})$. This is exactly the behavior of the solution u_ε in the time scale $t\theta_\varepsilon^{(1)}$ as $\varepsilon \to 0$ and then $t \to \infty$, and the one in the time scale $t\theta_\varepsilon^{(2)}$ as $\varepsilon \to 0$ and then $t \to 0$. This completes the description of the asymptotics of u_ε until the second scale $\theta_\varepsilon^{(2)}$.

More generally, there exist $\mathfrak{q} \geq 1$ and time-scales $\theta_\varepsilon^{(1)} \prec \cdots \prec \theta_\varepsilon^{(\mathfrak{q})}$ such that

$$\lim_{\varepsilon \to 0} u_\varepsilon(t\theta_\varepsilon^{(p)}, \boldsymbol{x}) = \sum_{k=1}^{\mathfrak{n}_{p-1}} \omega^{(p-1)}(\boldsymbol{m}, k) \sum_{\ell=1}^{\mathfrak{n}_{p-1}} p_t^{(p)}(k, \ell) \sum_{\boldsymbol{m}' \in \mathcal{M}_0} \pi_\ell^{(p-1)}(\boldsymbol{m}') \, u_0(\boldsymbol{m}')$$

$$(1.7)$$

for each $1 \leq p \leq \mathfrak{q}, t > 0, \boldsymbol{m} \in \mathcal{M}_0, \boldsymbol{x} \in \mathcal{D}(\boldsymbol{m})$. Furthermore, for each $1 \leq p \leq \mathfrak{q} + 1$, and sequence $(\varrho_\varepsilon : \varepsilon > 0)$ such that $\theta_\varepsilon^{(p-1)} \prec \varrho_\varepsilon \prec \theta_\varepsilon^{(p)}, \boldsymbol{m} \in \mathcal{M}_0$, and $\boldsymbol{x} \in \mathcal{D}(\boldsymbol{m})$,

$$\lim_{\varepsilon \to 0} u_\varepsilon(\varrho_\varepsilon, \boldsymbol{x}) = \sum_{k=1}^{\mathfrak{n}_{p-1}} \omega^{(p-1)}(\boldsymbol{m}, k) \sum_{\boldsymbol{m}' \in \mathcal{M}_0} \pi_k^{(p-1)}(\boldsymbol{m}') u_0(\boldsymbol{m}'). \qquad (1.8)$$

In this formula, $\theta_\varepsilon^{(0)}, \theta_\varepsilon^{(\mathfrak{q}+1)}$ are the constant sequences equal to $1, +\infty$, respectively. Summing up,

- Denote by $\mathfrak{n}_0$ the number of local minima of U so that $\mathfrak{n}_0 > \mathfrak{n}_1 > \cdots > \mathfrak{n}_{\mathfrak{q}} = 1$.
- $p_t^{(p)}, t \geq 0$, is a Markov semigroup on $\{1, \ldots, \mathfrak{n}_{p-1}\}, 1 \leq p \leq \mathfrak{q}$. Here, the semigroup p_t, introduced in (1.4), has been represented by $p_t^{(1)}$.
- For a fixed $1 \leq p \leq \mathfrak{q}, \pi_j^{(p)}, 1 \leq j \leq \mathfrak{n}_p$, are probability measures on $\mathcal{M}_0$ with disjoint supports. They correspond to the extremal invariant probability measures of the Markov chain with transition probability $p_t^{(p)}$.
- $\omega^{(p)}(\boldsymbol{m}, \cdot)$ are probability measures on $\{1, \ldots, \mathfrak{n}_p\}$, where $\omega^{(p)}(\boldsymbol{m}, j)$ stands for the probability that $\boldsymbol{m}$ is absorbed at the support of the probability measure $\pi_j^{(p)}$.

It turns out that all local minima which belong to the support of a measure $\pi_j^{(p)}$ are at the same height: $U(\boldsymbol{m}') = U(\boldsymbol{m}'')$ if $\boldsymbol{m}', \boldsymbol{m}''$ belong to the support of the same measure $\pi_j^{(p)}$. On the other hand, the support of a measure $\pi_j^{(p+1)}$ is formed by the union of the supports of measures $\pi_k^{(p)}, k \in \{1, \ldots, \mathfrak{n}_p\}$. Moreover, $\pi_j^{(p+1)}$ is a convex combination of the corresponding measures $\pi_k^{(p)}$. The rigorous recursive construction of this multi-scale structure is a delicate and complicated task and will be done in the companion paper [20]. Assertions (1.7) and (1.8) will be proven there as well.

Comments on the Proof

The analysis of the asymptotics of the solution $u_\varepsilon(t, \boldsymbol{x})$ of (1.3) is closely related to that of the metastable behavior of the process

$$d\boldsymbol{x}_\varepsilon(t) = -(\nabla U + \boldsymbol{\ell})(\boldsymbol{x}_\varepsilon(t)) \, dt + \sqrt{2\varepsilon} \, dW_t, \qquad (1.9)$$

where $\varepsilon > 0$ denotes a small parameter corresponding to the temperature of the system, and W_t a d-dimension Brownian motion. This relation comes from well-known expression

$$u_\varepsilon(t, \boldsymbol{x}) = \mathbb{E}_{\boldsymbol{x}}^\varepsilon[u_0(\boldsymbol{x}_\varepsilon(t))], \quad t \geq 0, \boldsymbol{x} \in \mathbb{R}^d,$$

where $\mathbb{E}_{\boldsymbol{x}}^\varepsilon$ denotes the expectation with respect to the diffusion process (1.9) starting at $\boldsymbol{x} \in \mathbb{R}^d$.

The proof of the result described above is purely probabilistic and relies on the theory of metastable Markov processes developed in [2,21–24,37]. The metastable

behavior of the process (1.9) has been recently studied in several articles: [31] provided sharp asymptotics on the low-lying spectra which is closely related with the metastability of the process $x_\varepsilon(\cdot)$, [28] established Eyring-Kramers law precisely estimating the mean transition time from a local minimum of U to another one, and finally [29] investigated the metastability among the global minima (i.e., ground states) of U. The last work can be regarded as the analysis of the metastability at the final scale $\theta_\varepsilon^{(\mathfrak{q})}$ described above.

The recursive construction of the multiscale structure presented here appeared before in different contexts. Michel [32] introduced it to study the low-lying eigenvalues of the semiclassical Witten Laplacian associated to a Morse function. We refer to [4,13,26] for the same construction in the context of finite state Markov chains.

The analysis of the multi-scale structure is based on the resolvent approach to metastability developed in [22]. The crucial point consists in showing that the solution of a resolvent equation is asymptotically constant in neighborhoods of local minima. More precisely, denote by $\mathcal{E}(m)$ a small neighborhood of a local minimum m. Fix $\lambda > 0$, $g: \mathcal{M}_0 \to \mathbb{R}$, and let ϕ_ε be the unique solution of the resolvent equation

$$(\lambda - \theta_\varepsilon^{(1)}\mathcal{L}_\varepsilon)\,\phi_\varepsilon \;=\; G := \sum_{m \in \mathcal{M}_0} g(m)\,\chi_{\mathcal{E}(m)},$$

where $\chi_\mathcal{A}$, $\mathcal{A} \subset \mathbb{R}^d$, represents the indicator function of the set $\mathcal{A}$. The function on the right-hand side vanishes at $(\cup_{m \in \mathcal{M}_0} \mathcal{E}(m))^c$ and is constant on each well $\mathcal{E}(m')$. One of the main results of this article asserts that the solution ϕ is asymptotically constant in each well $\mathcal{E}(m)$:

$$\lim_{\varepsilon \to 0} \max_{m \in \mathcal{M}_0} \sup_{x \in \mathcal{E}(m)} |\phi_\varepsilon(x) - f(m)| \;=\; 0, \qquad (1.10)$$

where f is the solution of the reduced resolvent equation

$$(\lambda - \mathfrak{L}_1)\,f \;=\; g, \qquad (1.11)$$

and $\mathfrak{L}_1$ is the generator of the $\mathcal{M}_0$-valued, continuous-time Markov chain whose associated semigroup is the one appearing in (1.4). Property (1.4) of the solution of the initial-valued problem (1.3) is deduced from this property of the resolvent equation.

Background

In a sequence of seminal works, Freidlin and Wentzel, [14] and references therein, investigated the metastable behavior of random perturbations of dynamical systems, and introduced the notion of hierarchy of cycles. Assuming that each cycle has only one subsequent cycle, they described the metastable behavior of the diffusion $dx_\varepsilon(t) = b(x_\varepsilon(t))\,dt + \sqrt{2\varepsilon}dW_t$ at all the time scales other than the critical ones at which the diffusion may jump from one cycle to another. We refer to [11,12,15,16,18] for recent developments of this theory.

With the notation introduced above, the hypothesis that each cycle has only one subsequent cycle means that for each $1 \leq p \leq \mathfrak{q}$, $1 \leq j \leq \mathfrak{n}_{p-1}$, there exists only one k such that $p_t^{(p)}(j, k) > 0$.

In this article, assuming that the drift $\boldsymbol{b}(\cdot)$ can be written as $\boldsymbol{b} = -(\nabla U + \boldsymbol{\ell})$ for a vector field $\boldsymbol{\ell}$ satisfying conditions (1.1), we extend the results in [14] in two directions: describing the behavior of the diffusion (or the one of the solution of the parabolic equation (1.3)) at the critical time-scales $\theta_\varepsilon^{(p)}$, and removing the hypothesis that each cycle has only one subsequent cycle.

As mentioned at the beginning of this introduction, we impose the conditions (1.1) on the vector field $\boldsymbol{\ell}(\cdot)$ for the Gibbs measure $\mu_\varepsilon(\mathrm{d}\boldsymbol{x}) = (1/Z_\varepsilon)\, e^{-U(\boldsymbol{x})/\varepsilon}\, \mathrm{d}\boldsymbol{x}$ to be the stationary state of the diffusion process $\boldsymbol{x}_\varepsilon(\cdot)$. The precise description of the asymptotic behavior of the solution u_ε in the time-scale $\theta_\varepsilon^{(1)}$ presented in (1.4) relies on explicit computations which require an explicit formula for the stationary state and smoothness of its density.

In general (that is, without the hypotheses (1.1)), the quasi-potential, which plays the role of $U(\cdot)$ in the formula for the stationary state, is not smooth and not known explicitly, making it impossible to apply the approach proposed here. We refer to [10,25], for a model (a one-dimensional diffusion on the torus) where the quasi-potential can be computed and the methods presented here applied, despite the lack of smoothness of the quasi-potential.

Estimates of the transition times expectation have been obtained in [6,23]. In the presence of many wells, these expectations may not converge for the following reason. With a very small probability, crossing a saddle point higher than the lowest one, the diffusion may hit a very deep well and remain there a very long time. This contribution might be dominant for the expectation, turning it much larger than predicted.

Uniform estimates, similar to (1.10), for solutions of Dirichlet problems go back at least to Devinatz and Friedman [9], and Day [8]. The convergence to a constant is called in the literature the leveling property of the equation. We refer to Lelièvre, Le Peutrec and Nectoux [30] for a recent account and further references.

Organization

The paper is organized as follows. In Sect. 2, we state the main results. The proof of Theorem 2.2 is divided in two parts. In Sect. 4, we prove that the solution of the resolvent equation is constant on each well, and, in Sect. 9, that the solution of the resolvent equation restricted to the set of local minima of U is asymptotically the solution of the reduced resolvent equation (1.11).

The proof of the local constancy relies on a diffusion mixing time estimate presented in Sect. 3. The proof of the second property of the resolvent equation solution requires an estimate of the time it takes to exit a neighborhood of an unstable equilibrium point, presented in Sect. 5, estimates on the time needed to reach a local minimum of U, the subject of Sect. 6, and test functions which approximate the equilibrium potential between wells, introduced in Sect. 7. In Sect. 8, we add the last piece of the proof, extending the results of Sect. 4 by showing that the solution of the

resolvent equation is actually asymptotically constant in the domain of attractions of a local minimum. In Sect. 9 we prove Theorem 2.2, and Theorem 2.1 in Sect. 10. In the appendices, we present some results needed in the proofs.

2. Model and Main Results

Fix a function $U : \mathbb{R}^d \to \mathbb{R}$ in $C^3(\mathbb{R}^d)$ admitting only a finite number of critical points, all non-degenerate (hence U is a Morse function, cf. [34, Definition 1.7]). Assume that

$$
\lim_{n \to \infty} \inf_{|x| \geq n} \frac{U(x)}{|x|} = \infty, \quad \lim_{|x| \to \infty} \frac{x}{|x|} \cdot \nabla U(x) = \infty,
$$
$$
\lim_{|x| \to \infty} \left\{ |\nabla U(x)| - 2\,\Delta U(x) \right\} = \infty. \tag{2.1}
$$

In this formula and below, $|x|$ represents the Euclidean norm of $x \in \mathbb{R}^d$. Suppose, without loss of generality, that $\min_{x \in \mathbb{R}^d} U(x) = 0$. Consider a vector field $\ell : \mathbb{R}^d \to \mathbb{R}$ in $C^2(\mathbb{R}^d)$, assumed to be divergence free and orthogonal to the gradient of U as stated in (1.1).

Time-Scale

Denote by $\mathcal{M}_0$ the set of local minima of U. For each pair $m' \neq m'' \in \mathcal{M}_0$, denote by $\Theta(m', m'')$ the communication height between m' and m'' as

$$
\Theta(m', m'') := \inf_{z : [0\,1] \to \mathbb{R}^d} \max_{t \in [0,\,1]} U(z(t)), \tag{2.2}
$$

where the minimum is carried over all continuous paths $z(\cdot)$ such that $z(0) = m'$ and $z(1) = m''$. Clearly, $\Theta(m', m'') = \Theta(m'', m')$. Denote by $\Gamma(m)$ the depth of the local minimum $m \in \mathcal{M}_0$:

$$
\Gamma(m) := \min_{m' \neq m} \Theta(m, m') - U(m). \tag{2.3}
$$

Denote by $d^{(1)}$ the depth of the shallowest well, and by $\theta_\varepsilon^{(1)}$ the corresponding time-scale:

$$
d^{(1)} := \min_{m \in \mathcal{M}_0} \Gamma(m), \quad \theta_\varepsilon^{(1)} := e^{d^{(1)}/\varepsilon}.
$$

Gates

Denote by $\Upsilon(m)$ the set of gates of $m \in \mathcal{M}_0$. This is the set of points $x \in \mathbb{R}^d$ for which there exist $m' \in \mathcal{M}_0$, $m' \neq m$, and a continuous path $z : [0, 1] \to \mathbb{R}^d$ such that $z(0) = m$, $z(1) = m'$, $z(1/2) = x$ and $U(z(t)) < U(x) = U(m) + \Gamma(m)$ for all $t \in [0, 1]$, $t \neq 1/2$.

Mind that there might be more than one local minima m' for the same gate $x \in \Upsilon(m)$: there might exist $m_1 \neq m_2$, both different from m, $x \in \Upsilon(m)$, and

continuous paths $z_i \colon [0, 1] \to \mathbb{R}^d$, $i = 1, 2$, such that $z_i(0) = \boldsymbol{m}$, $z_i(1) = \boldsymbol{m}_i$, $z_i(1/2) = \boldsymbol{x}$ and $U(z_i(t)) < U(\boldsymbol{x}) = U(\boldsymbol{m}) + \Gamma(\boldsymbol{m})$ for all $t \in [0, 1]$, $t \neq 1/2$.

Mind also that in the definition of gate, we require $\boldsymbol{m}'$ to be different from $\boldsymbol{m}$. In this way, we exclude from the set of gates points $\boldsymbol{x}$ for which there exists a continuous path $z \colon [0, 1] \to \mathbb{R}^d$ such that $z(0) = \boldsymbol{m}$, $z(1) = \boldsymbol{m}$, $z(1/2) = \boldsymbol{x}$ and $U(z(t)) < U(\boldsymbol{x}) = U(\boldsymbol{m}) + \Gamma(\boldsymbol{m})$ for all $t \in [0, 1]$, $t \neq 1/2$.

Recall that $\boldsymbol{b} = -(\nabla U + \boldsymbol{\ell})$ and that a heteroclinic orbit ϕ from $\boldsymbol{x}$ to $\boldsymbol{y} \in \mathbb{R}^d$ is a solution $\phi \colon \mathbb{R} \to \mathbb{R}^d$ of the ODE

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{b}(\boldsymbol{x}(t)), \tag{2.4}$$

such that

$$\lim_{t \to -\infty} \phi(t) = \boldsymbol{x}, \quad \lim_{t \to \infty} \phi(t) = \boldsymbol{y}.$$

We represent this relation by $\boldsymbol{x} \curvearrowright \boldsymbol{y}$. In other words, $\boldsymbol{x} \curvearrowright \boldsymbol{y}$ indicates the existence of a heteroclinic orbit from $\boldsymbol{x}$ to $\boldsymbol{y}$. We assume also that for all $\boldsymbol{m} \in \mathcal{M}_0$ such that $\Gamma(\boldsymbol{m}) = d^{(1)}$, and $\boldsymbol{\sigma} \in \Upsilon(\boldsymbol{m})$, there exists $\boldsymbol{m}' \in \mathcal{M}_0$, $\boldsymbol{m}' \neq \boldsymbol{m}$, such that

$$\boldsymbol{\sigma} \curvearrowright \boldsymbol{m} \quad \text{and} \quad \boldsymbol{\sigma} \curvearrowright \boldsymbol{m}'. \tag{2.5}$$

The condition (2.5) is crucial. For instance, if there is a heteroclinic orbit from a saddle point $\boldsymbol{\sigma} \in \Upsilon(\boldsymbol{m})$ to a saddle point $\boldsymbol{\sigma}'$, we are not able to determine which is the local minimum visited after $\boldsymbol{m}$ when the process starts from a neighbourhood of $\boldsymbol{m}$.

The assumption (2.5) holds when the dynamical system $\boldsymbol{x}(\cdot)$ defined in (2.4) is a Morse-Smale system. In a Morse-Smale system, for two critical points $\boldsymbol{c}_1$, $\boldsymbol{c}_2$ of $U(\cdot)$, the unstable manifold of $\boldsymbol{c}_1$ and the stable manifold of $\boldsymbol{c}_2$ intersect transversally and thus when $\boldsymbol{c}_1 \curvearrowright \boldsymbol{c}_2$, the index (the number of negative eigenvalue of the Hessian) must strictly decrease along heteroclinic orbits; hence (2.5) follows naturally.

By Proposition A.1, any gate $\boldsymbol{x} \in \Upsilon := \cup_{\boldsymbol{m} \in \mathcal{M}_0} \Upsilon(\boldsymbol{m})$ belongs to the set of critical points of U, denoted by $\mathcal{C}_0$:

$$\mathcal{C}_0 := \{\, \boldsymbol{x} \in \mathbb{R}^d : (\nabla U)(\boldsymbol{x}) = 0 \,\}.$$

By [28, Theorem 2.1], the divergence-free field $\boldsymbol{\ell}$ vanishes at the critical points of U: $\boldsymbol{\ell}(\boldsymbol{x}) = 0$ for all $\boldsymbol{x} \in \mathcal{C}_0$. Denote by $(\nabla^2 U)(\boldsymbol{x})$ the Hessian of U at $\boldsymbol{x}$. Since U is a Morse function, for all $\boldsymbol{\sigma} \in \Upsilon$,

$(\nabla^2 U)(\boldsymbol{\sigma})$ has only one negative eigenvalue, all the others being strictly positive.
$$\tag{2.6}$$

Indeed, by definition, $\boldsymbol{\sigma}$ can not be a local minimum. On the other hand, assume that $\boldsymbol{\sigma}$ is a gate between $\boldsymbol{m}$ and $\boldsymbol{m}'$. If the number of negative eigenvalues is greater than 1, the set $\{\boldsymbol{x} : U(\boldsymbol{x}) < U(\boldsymbol{\sigma})\}$ would be locally connected, and there would be a continuous path from $\boldsymbol{m}$ to $\boldsymbol{m}'$ staying strictly below $U(\boldsymbol{\sigma})$, which is a contradiction.

Denote by $\mathcal{V}(\boldsymbol{m})$ the set of points $\boldsymbol{m}' \in \mathcal{M}_0$, $\boldsymbol{m}' \neq \boldsymbol{m}$, for which (2.5) holds for some $\boldsymbol{\sigma} \in \Upsilon(\boldsymbol{m})$. Hence, $\mathcal{V}(\boldsymbol{m})$ is the set of local minima $\boldsymbol{m}' \in \mathcal{M}_0$, $\boldsymbol{m}' \neq \boldsymbol{m}$, for

which there exist a critical point $\sigma \in \Upsilon(m)$ and heteroclinic orbits from σ to m and σ to m':

$$\mathcal{V}(m) := \left\{ m' \in \mathcal{M}_0 \setminus \{m\} : \exists \sigma \in \Upsilon(m) \text{ such that } \sigma \curvearrowright m, \ \sigma \curvearrowright m' \right\}.$$

Elements of $\mathcal{V}(m)$ are called neighbors of the local minimum m of U. Denote by $\mathcal{S}(m, m'), m' \neq m$, the set of critical points which separate m from m':

$$\mathcal{S}(m, m') := \left\{ \sigma \in \Upsilon(m) : \sigma \curvearrowright m, \ \sigma \curvearrowright m' \right\}. \tag{2.7}$$

Reduced Model

Denote by $(D\ell)(x)$ the Jacobian of ℓ at x. By (2.6), $(\nabla^2 U)(\sigma), \sigma \in \Upsilon$, has only one negative eigenvalue. By [28, Lemma 3.3], $(\nabla^2 U)(\sigma) + (D\ell)(\sigma)$ has also one negative eigenvalue, represented by $-\mu_\sigma < 0$. For $m \in \mathcal{M}_0, \sigma \in \Upsilon(m)$, let the weights $\nu(m), \omega(\sigma)$ be given by

$$\nu(m) := \frac{1}{\sqrt{\det(\nabla^2 U)(m)}}, \quad \omega(\sigma) := \frac{\mu_\sigma}{2\pi \sqrt{-\det \nabla^2 U(\sigma)}} \cdot \tag{2.8}$$

Let $\omega(m, m'), m \neq m' \in \mathcal{M}_0$, be the weight given by

$$\omega(m, m') := \sum_{\sigma \in \mathcal{S}(m,m')} \omega(\sigma). \tag{2.9}$$

Note that $\omega(m, m')$ vanishes if $m' \notin \mathcal{V}(m)$. Moreover, neither $\mathcal{S}(\cdot, \cdot)$ nor $\omega(\cdot, \cdot)$ is symmetric in its arguments. To include the depth of the local minimum m in the definition of the weight $\omega(m, m')$, set

$$\omega_1(m, m') := \omega(m, m') \mathbf{1}\{\Gamma(m) = d^{(1)}\}. \tag{2.10}$$

Denote by $\mathfrak{L}_1$ the generator of the $\mathcal{M}_0$-valued, continuous-time Markov chain given by

$$(\mathfrak{L}_1 h)(m) = \frac{1}{\nu(m)} \sum_{m' \in \mathcal{M}_0} \omega_1(m, m') [h(m') - h(m)]. \tag{2.11}$$

As $\omega(m, m')$ vanishes if m' does not belong to $\mathcal{V}(m)$, the sum can be carried over $\mathcal{V}(m)$.

Theorem 2.1. *Assume that hypotheses (2.1), (2.5) are in force. Fix a bounded and continuous function $u_0 \colon \mathbb{R}^d \to \mathbb{R}$. Denote by u_ε the solution of the parabolic equation (1.3). Then, (1.4) and (1.5) hold for all $t > 0$, where $p_t(\cdot, \cdot)$ is the semigroup associated to the generator $\mathfrak{L}_1$.*

Resolvent Equation

The proof of Theorem 2.1 is based on properties of the resolvent equation presented in this subsection. Denote by $B(x, r)$, $x \in \mathbb{R}^d$, $r > 0$, the open ball of radius r centered at x. Let $\mathcal{W}^r(m)$, $m \in \mathcal{M}_0, r > 0$, be the connected component of the set $\{x \in \mathbb{R}^d : U(x) \leq U(m) + r\}$ containing m.

Fix $m \in \mathcal{M}_0$. All constants r_i below depend on m and $b(\cdot)$, though this does not appear in the notation. Equation (B.1) introduces a positive constant $r_5 > 0$. Choose r_4 small enough for (3.9) to hold with $r_3 = r_5$. By Proposition B.1 and conditions (1), (2) in Sect. 3, $B(m, r_5)$ does not contain critical points of U besides m.

Choose $r_0 > 0$ small enough so that for all $m \in \mathcal{M}_0$,

(a) $\overline{\mathcal{W}^{2r_0}(m)}\backslash\{m\}$ does not contain critical points of U;
(b) $\mathcal{W}^{2r_0}(m)$ is contained in the domain of attraction of m for the ODE (2.4);
(c) $b(x) \cdot n(x) < 0$ for all $x \in \partial\mathcal{W}^{2r_0}(m)$, where $n(\cdot)$ is the exterior normal of the boundary of $\mathcal{W}^{2r_0}(m)$.
(d) $\mathcal{W}^{3r_0}(m) \subset B(m, r_5)$.
(e) $\mathcal{W}^{2r_0}(m) \subset \mathcal{D}_{r_4}(m)$

Since $\mathcal{W}^r(m) = \{x \in \mathbb{R}^d : U(x) \leq U(m)+r\}$, for r small enough $n = \nabla U$ at the boundary of $\mathcal{W}^r(m)$. In particular, as $\ell \cdot \nabla U = 0$, $b(x) \cdot n(x) = -|\nabla U(x)|^2 < 0$ for all $x \in \partial\mathcal{W}^{2r_0}(m)$ and r_0 small enough.

Set

$$\mathcal{E}(m) := \mathcal{W}^{r_0}(m), \quad m \in \mathcal{M}_0. \tag{2.12}$$

For $\lambda > 0$, $g: \mathcal{M}_0 \to \mathbb{R}$, denote by $\phi_\varepsilon = \phi_\varepsilon^{\lambda, g}$ the unique solution of the resolvent equation

$$(\lambda - \theta_\varepsilon^{(1)}\mathcal{L}_\varepsilon)\,\phi_\varepsilon \;=\; G := \sum_{m \in \mathcal{M}_0} g(m)\,\chi_{\mathcal{E}(m)}, \tag{2.13}$$

where $\chi_{\mathcal{A}}$, $\mathcal{A} \subset \mathbb{R}^d$, represents the indicator function of the set $\mathcal{A}$. The function on the right-hand side vanishes at $(\cup_{m \in \mathcal{M}_0}\mathcal{E}(m))^c$ and is constant on each well $\mathcal{E}(m')$. The second main result of this article reads as follows:

Theorem 2.2. *For all $\lambda > 0$ and $g: \mathcal{M}_0 \to \mathbb{R}$,*

$$\lim_{\varepsilon \to 0} \max_{m \in \mathcal{M}_0} \sup_{x \in \mathcal{E}(m)} |\phi_\varepsilon(x) - f(m)| \;=\; 0,$$

where f is the solution of the reduced resolvent equation

$$(\lambda - \mathfrak{L}_1)\,f \;=\; g,$$

and $\mathfrak{L}_1$ is the generator introduced in (2.11).

Comments and Remarks

The proofs of Theorems 2.1 and 2.2 are entirely based on the metastable behavior of the stochastic differential equation

$$\mathrm{d}\boldsymbol{x}_\varepsilon(t) = \boldsymbol{b}(\boldsymbol{x}_\varepsilon(t))\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}W_t, \qquad (2.14)$$

where $\varepsilon > 0$ denotes a small parameter corresponding to the temperature of the system, and W_t a d-dimension Brownian motion.

The proof of Theorem 2.2 is divided in two parts. We first show in Sect. 4 that ϕ_ε is asymptotically constant on each well $\mathcal{E}(\boldsymbol{m})$. Then, we prove that the average of the solution ϕ_ε on a well $\mathcal{E}(\boldsymbol{m})$ converges to $\boldsymbol{f}$.

In Sect. 10, we deduce from Theorem 2.2 and with ideas introduced in [21], the convergence of the finite-dimensional distributions of the process $\boldsymbol{x}_\varepsilon(\cdot)$. A similar result has been obtained by Sugiura in [38] with different ideas in the case $\boldsymbol{\ell} = 0$.

3. Mixing Time of Diffusions

The main result of this section, Theorem 3.1, provides an estimate on the mixing time of a diffusion on $\mathbb{R}^d$. The proof of this result can be skipped in a first reading as the ideas and techniques used to derive the bound on the mixing time will not be used in the next sections.

Fix a function $U_0 \colon \mathbb{R}^d \to \mathbb{R}$ of class C^3 and a vector field $\boldsymbol{\ell}_0 \colon \mathbb{R}^d \to \mathbb{R}^d$ of class C^2 such that

$$(\nabla U_0)(\boldsymbol{x}) \cdot \boldsymbol{\ell}_0(\boldsymbol{x}) = (\nabla \cdot \boldsymbol{\ell}_0)(\boldsymbol{x}) = 0 \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^d. \qquad (3.1)$$

Suppose that U_0 has a local minimum at $\boldsymbol{x} = \boldsymbol{0}$ and that it has no other critical point in a neighborhood of the origin. Furthermore, we assume, for convenience, that $U_0(\boldsymbol{0}) = 0$.

Consider a vector field $\boldsymbol{b}_0 \colon \mathbb{R}^d \to \mathbb{R}^d$ of class C^1 such that

(1) $\boldsymbol{b}_0$ vanishes only at the origin, which is a stable equilibrium point for the dynamical system

$$\dot{\boldsymbol{y}}(t) = \boldsymbol{b}_0(\boldsymbol{y}(t)). \qquad (3.2)$$

(2) There exists $r_3 > 0$ such that

$$\boldsymbol{b}_0(\boldsymbol{x}) = -(\nabla U_0)(\boldsymbol{x}) - \boldsymbol{\ell}_0(\boldsymbol{x}), \quad \boldsymbol{x} \in B(0, r_3).$$

(3) There exist $R > 0$ and a finite constant C_1 such that

$$|\boldsymbol{b}_0(\boldsymbol{x})| \le C_1 |\boldsymbol{x}| \quad \text{and} \quad \|D\boldsymbol{b}_0(\boldsymbol{x})\| \le C_1 |\boldsymbol{x}| \qquad (3.3)$$

for all $|\boldsymbol{x}| > R$, where the matrix norm is defined as

$$\|\mathbb{M}\| = \sup_{|\boldsymbol{y}|=1} |\mathbb{M}\boldsymbol{y}|.$$

(5) Let $\mathbb{H}_0 = (\nabla^2 U_0)(\mathbf{0})$ and $\mathbb{L}_0 = (D\boldsymbol{\ell}_0)(\mathbf{0})$. Assume that

$$- \langle \boldsymbol{b}_0(\boldsymbol{x}), \mathbb{H}_0 \boldsymbol{x} \rangle \geq \frac{1}{2} |\mathbb{H}_0 \boldsymbol{x}|^2 \quad \text{for all } \boldsymbol{x} \in \mathbb{R}. \tag{3.4}$$

where $\langle \cdot, \cdot \rangle$ represents the scalar product in $\mathbb{R}^d$.

The main result of this section requires some notation. Let

$$\mathbb{A}(\boldsymbol{x}) := (D\boldsymbol{b}_0)(\boldsymbol{x}), \quad \mathbb{A} := \mathbb{A}(\mathbf{0}), \quad \text{so that} \quad \mathbb{A} = -(\mathbb{H}_0 + \mathbb{L}_0). \tag{3.5}$$

By [28, Lemmata 4.5 and 4.1], all the eigenvalues of the matrix $\mathbb{A}$ have negative real parts. Therefore, by [19, Theorems 2 and 3, p.414], there exists a positive definite matrix $\mathbb{K}$ such that

$$\mathbb{A}^\dagger \mathbb{K} + \mathbb{K}\mathbb{A} = -\mathbb{I}, \tag{3.6}$$

where $\mathbb{I}$ is the identity.

Let $\mathcal{D}_r \subset \mathbb{R}^d$, $r > 0$, be the set given by

$$\mathcal{D}_r := \{ \boldsymbol{x} \in \mathbb{R}^d : \langle \boldsymbol{x}, \mathbb{H}_0 \boldsymbol{x} \rangle \leq r^2 \}. \tag{3.7}$$

By (3.6), there exists $r_4' > 0$ such that

$$\left\| (\mathbb{A}(\boldsymbol{x}) - \mathbb{A})^\dagger \mathbb{K} + \mathbb{K}(\mathbb{A}(\boldsymbol{x}) - \mathbb{A}) \right\| \leq \frac{1}{2} \tag{3.8}$$

for all $\boldsymbol{x} \in B(0, r_4')$. By (3.7), $\mathcal{D}_{2r_4} \subset B(0, r_4')$ for some $r_4 > 0$. Take r_4 small enough so that

$$\mathcal{D}_{2r_4} \subset B(0, r_3), \tag{3.9}$$

where r_3 has been introduced in condition (2) above.

The main result of this section reads as follows. Denote by $d_{\text{TV}}(\mu, \nu)$ the total variation distance between probability measures μ and ν. Let $\boldsymbol{y}_\varepsilon(\cdot)$ be the diffusion given by

$$d\boldsymbol{y}_\varepsilon(t) = \boldsymbol{b}_0(\boldsymbol{y}_\varepsilon(t)) \, dt + \sqrt{2\varepsilon} \, dW_t. \tag{3.10}$$

The process $\boldsymbol{y}_\varepsilon(\cdot)$ starting at $\boldsymbol{x} \in \mathbb{R}^d$, is represented by $\boldsymbol{y}_\varepsilon(t; \boldsymbol{x})$. Let $t_\varepsilon = \varepsilon^{-\theta}$ for some $\theta \in (0, 1/3)$.

Theorem 3.1. *Denote by π_ε the stationary state of the diffusion $\boldsymbol{y}_\varepsilon(\cdot)$. Then,*

$$\lim_{\varepsilon \to 0} \sup_{\boldsymbol{x} \in \mathcal{D}_{r_4}} d_{\text{TV}}\left(\boldsymbol{y}_\varepsilon(t_\varepsilon; \boldsymbol{x}), \pi_\varepsilon \right) = 0.$$

Remark 3.2. The proof of this result is largely based on [1,27]. Theorem 3.1 follows from [1, Theorem 2.2] when $\mathbb{A}$ is negative definite. As mentioned above, all the eigenvalues of matrix $\mathbb{A}$ have negative real parts, but $\mathbb{A}$ might not be negative definite. The purpose of this section is to extend [1, Theorem 2.2] to this situation.

Proof of Theorem 3.1

The main idea of proof is to approximate the difference $\boldsymbol{y}_\varepsilon(t) - \boldsymbol{y}(t)$ by a Gaussian process. Let

$$\widehat{\boldsymbol{\xi}}(t) \,=\, \frac{1}{\sqrt{2\varepsilon}} \left(\boldsymbol{y}_\varepsilon(t) - \boldsymbol{y}(t)\right).$$

By (3.10) and (3.2),

$$d\widehat{\boldsymbol{\xi}}(t) \,=\, \frac{1}{\sqrt{2\varepsilon}} \left\{ \boldsymbol{b}_0(\boldsymbol{y}_\varepsilon(t)) - \boldsymbol{b}_0(\boldsymbol{y}(t)) \right\} dt \,+\, dW_t \simeq (D\boldsymbol{b}_0)(\boldsymbol{y}(t)) \,\widehat{\boldsymbol{\xi}}(t) dt + dW_t.$$

Hence, it is natural to conjecture that $\widehat{\boldsymbol{\xi}}(t) \simeq \boldsymbol{\xi}(t)$ where $\boldsymbol{\xi}(t)$ is the Gaussian process defined by the SDE

$$d\boldsymbol{\xi}(t) = \mathbb{A}(\boldsymbol{y}(t))\,\boldsymbol{\xi}(t)\,dt \,+\, dW_t, \quad \boldsymbol{\xi}(0) = \boldsymbol{0}. \tag{3.11}$$

Let

$$\boldsymbol{z}_\varepsilon(t) \,:=\, \boldsymbol{y}(t) \,+\, \sqrt{2\varepsilon}\,\boldsymbol{\xi}(t). \tag{3.12}$$

By the previous discussion, we expect that $\boldsymbol{y}_\varepsilon(t) \simeq \boldsymbol{z}_\varepsilon(t)$.

Lemma 3.3. *There exists $r_4 > 0$ such that*

$$\lim_{\varepsilon \to 0} \sup_{\boldsymbol{x} \in \mathcal{D}_{r_4}} d_{\mathrm{TV}}\left(\boldsymbol{y}_\varepsilon(t_\varepsilon; \boldsymbol{x}), \,\boldsymbol{z}_\varepsilon(t_\varepsilon; \boldsymbol{x})\right) = 0. \tag{3.13}$$

Denote by $\mathcal{N}(\boldsymbol{\mu}, \,\Sigma)$ the normal distribution with mean $\boldsymbol{\mu}$ and covariance Σ.

Lemma 3.4. *There exists $r_4 > 0$ such that*

$$\lim_{\varepsilon \to 0} \sup_{\boldsymbol{x} \in \mathcal{D}_{r_4}} d_{\mathrm{TV}}\left(\boldsymbol{z}_\varepsilon(t_\varepsilon; \boldsymbol{x}), \,\mathcal{N}(0, 2\varepsilon\mathbb{H}^{-1})\right) = 0.$$

Proof. The proof is presented in [1, Proposition 3.6], and relies on the fact that $\boldsymbol{z}_\varepsilon(\cdot; \boldsymbol{x})$ is a Gaussian process. In particular, $\boldsymbol{z}_\varepsilon(t; \boldsymbol{x})$ is a normal random variable whose mean and variance can be expressed explicitly. The assertion is thus reduced to a computation of the total variation distance between two normal random variables.

Denote by $\lambda > 0$ the smallest eigenvalue of $\mathbb{H}_0$. The proof starts at [1, display (3.22)], and requires the bound

$$|\boldsymbol{y}(t)|^2 \,\leq\, |\boldsymbol{y}(0)|^2\, e^{-\lambda t},$$

and [1, Lemma B.2]. In the present context, Lemma 3.5 replaces the first estimate, and [1, Lemma B.2] holds because it only needs all the eigenvalues of $(D\boldsymbol{b}_0)(\boldsymbol{0})$ to have a positive real part, a property satisfied by our model as mentioned in Remark 3.2. $\qquad\square$

Proof of Theorem 3.1. Denote by $p_t^\varepsilon(\cdot, \cdot)$ the transition kernel of the process $\mathbf{y}_\varepsilon(\cdot)$ and by $\pi_\varepsilon(\cdot)$ the density of the measure $\pi_\varepsilon(d\mathbf{x})$: $\pi_\varepsilon(d\mathbf{x}) = \pi_\varepsilon(\mathbf{x})d\mathbf{x}$. By definition, and since π_ε is the stationary state of the process $\mathbf{y}_\varepsilon(\cdot)$,

$$
d_{\mathrm{TV}}\left(\mathbf{y}_\varepsilon(t_\varepsilon; \mathbf{x}), \pi_\varepsilon\right) = \frac{1}{2}\int_{\mathbb{R}^d}\left|p_{t^\varepsilon}^\varepsilon(\mathbf{x}, \mathbf{y}) - \pi_\varepsilon(\mathbf{y})\right|d\mathbf{y}
$$
$$
= \frac{1}{2}\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}\left[p_{t^\varepsilon}^\varepsilon(\mathbf{x}, \mathbf{y}) - p_{t^\varepsilon}^\varepsilon(\mathbf{x}', \mathbf{y})\right]\pi_\varepsilon(\mathbf{x}')d\mathbf{x}'\right|d\mathbf{y}.
$$

The previous expression is bounded by

$$
\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left|p_{t^\varepsilon}^\varepsilon(\mathbf{x}, \mathbf{y}) - p_{t^\varepsilon}^\varepsilon(\mathbf{x}', \mathbf{y})\right|\pi_\varepsilon(\mathbf{x}')d\mathbf{x}'d\mathbf{y}
$$
$$
= \int_{\mathbb{R}^d}d_{\mathrm{TV}}(\mathbf{y}_\varepsilon(t_\varepsilon; \mathbf{x}), \mathbf{y}_\varepsilon(t_\varepsilon; \mathbf{x}'))\pi_\varepsilon(\mathbf{x}')d\mathbf{x}'.
$$

By (3.22), the right-hand side is less than or equal to

$$
\int_{\mathcal{D}_{r_4}}d_{\mathrm{TV}}\left(\mathbf{y}_\varepsilon(t_\varepsilon; \mathbf{x}), \mathbf{y}_\varepsilon(t_\varepsilon; \mathbf{x}')\right)\pi_\varepsilon(\mathbf{x}')d\mathbf{x}' + C_0\,\varepsilon
$$

for some finite constant C_0. By Lemma 3.3, and a triangular inequality,

$$
\limsup_{\varepsilon\to 0}\sup_{\mathbf{x}\in\mathcal{D}_{r_4}}\int_{\mathcal{D}_{r_4}}\left|d_{\mathrm{TV}}\left(\mathbf{y}_\varepsilon(t_\varepsilon; \mathbf{x}), \mathbf{y}_\varepsilon(t_\varepsilon; \mathbf{x}')\right)\right.
$$
$$
\left. - d_{\mathrm{TV}}\left(\mathbf{z}_\varepsilon(t_\varepsilon; \mathbf{x}), \mathbf{z}_\varepsilon(t_\varepsilon; \mathbf{x}')\right)\right|\pi_\varepsilon(\mathbf{x}')d\mathbf{x}' = 0.
$$

It remains to show that

$$
\limsup_{\varepsilon\to 0}\sup_{\mathbf{x}\in\mathcal{D}_{r_4}}\int_{\mathcal{D}_{r_4}}d_{\mathrm{TV}}\left(\mathbf{z}_\varepsilon(t_\varepsilon; \mathbf{x}), \mathbf{z}_\varepsilon(t_\varepsilon; \mathbf{x}')\right)\pi_\varepsilon(\mathbf{x}')d\mathbf{x}' = 0. \qquad (3.14)
$$

Since the integrand is bounded by

$$
d_{\mathrm{TV}}\left(\mathbf{z}_\varepsilon(t_\varepsilon; \mathbf{x}), \mathcal{N}(0, 2\varepsilon\mathbb{H}^{-1})\right) + d_{\mathrm{TV}}\left(\mathcal{N}(0, 2\varepsilon\mathbb{H}^{-1}), \mathbf{z}_\varepsilon(t_\varepsilon; \mathbf{x}')\right),
$$

assertion (3.14) follows from Lemma 3.4. $\qquad\square$

Proof of Lemma 3.3

The proof is similar to the one presented in [1, Section 3.3], which is based on conditions (C) or (H) of that article. These conditions, however, are only used in the proof of Lemma 3.3 to derive the estimates presented in Lemmata 3.5, 3.7, 3.10, 3.11, and Proposition 3.8.

Fix $\delta_\varepsilon = \varepsilon^c$ for some $c > 0$. As

$$
\mathbf{y}_\varepsilon(t_\varepsilon; \mathbf{x}) = \mathbf{y}_\varepsilon(\delta_\varepsilon; \mathbf{y}_\varepsilon(t_\varepsilon - \delta_\varepsilon; \mathbf{x})), \quad \mathbf{z}_\varepsilon(t_\varepsilon; \mathbf{x}) = \mathbf{z}_\varepsilon(\delta_\varepsilon; \mathbf{z}_\varepsilon(t_\varepsilon - \delta_\varepsilon; \mathbf{x})),
$$

we have that

$$
\begin{aligned}
d_{\mathrm{TV}} & \left(\boldsymbol{y}_\varepsilon(t_\varepsilon; \boldsymbol{x}),\ \boldsymbol{z}_\varepsilon(t_\varepsilon; \boldsymbol{x}) \right) \\
& \leq\ d_{\mathrm{TV}} \left(\boldsymbol{y}_\varepsilon(\delta_\varepsilon; \boldsymbol{y}_\varepsilon(t_\varepsilon - \delta_\varepsilon; \boldsymbol{x})),\ \boldsymbol{z}_\varepsilon(\delta_\varepsilon; \boldsymbol{y}_\varepsilon(t_\varepsilon - \delta_\varepsilon; \boldsymbol{x})) \right) \\
& \ +\ d_{\mathrm{TV}} \left(\boldsymbol{z}_\varepsilon(\delta_\varepsilon; \boldsymbol{y}_\varepsilon(t_\varepsilon - \delta_\varepsilon; \boldsymbol{x})),\ \boldsymbol{z}_\varepsilon(\delta_\varepsilon; \boldsymbol{z}_\varepsilon(t_\varepsilon - \delta_\varepsilon; \boldsymbol{x})) \right).
\end{aligned} \tag{3.15}
$$

The first term on the right-hand side is bounded in [1, Proposition 3.3] and the second one in [1, Proposition 3.4]. The proof relies on the estimate presented in Proposition 3.8 below.

We sketch the proof of these bounds. For the first one, fix $\boldsymbol{x} \in \mathcal{D}_{r_4}$ and denote by $\mathbb{P}_Y$ and $\mathbb{P}_Z$ the law of process $(\boldsymbol{y}_\varepsilon(s; \boldsymbol{x}))_{s \in [0, \delta_\varepsilon]}$ and $(\boldsymbol{z}_\varepsilon(s; \boldsymbol{x}))_{s \in [0, \delta_\varepsilon]}$, respectively. By Pinsker inequality,

$$
d_{\mathrm{TV}} \left(\boldsymbol{y}_\varepsilon(\delta_\varepsilon; \boldsymbol{x}),\ \boldsymbol{z}_\varepsilon(\delta_\varepsilon; \boldsymbol{x}) \right)^2\ \leq\ -2\, \mathbb{E}_{\mathbb{P}_Y} \left[\log \frac{d\mathbb{P}_Z}{d\mathbb{P}_Y} \right]. \tag{3.16}
$$

The SDE describing the process $\boldsymbol{z}_\varepsilon(\cdot)$ can be written as

$$
d\boldsymbol{z}_\varepsilon(t)\ =\ \left\{ \boldsymbol{b}_0(\boldsymbol{y}(t)) + D\boldsymbol{b}_0(\boldsymbol{y}(t))\, [\, \boldsymbol{z}_\varepsilon(t) - \boldsymbol{y}(t)\,] \right\} dt\ +\ \sqrt{2\varepsilon}\, dW_t. \tag{3.17}
$$

Hence, by Girsanov theorem, (3.10), and (3.17),

$$
\begin{aligned}
\log \frac{d\mathbb{P}_Z}{d\mathbb{P}_Y} =& \frac{1}{\sqrt{2\varepsilon}} \int_0^{\delta_\varepsilon} \big\langle \boldsymbol{b}_0(\boldsymbol{y}_\varepsilon(t)) - \boldsymbol{b}_0(\boldsymbol{y}(t)) + D\boldsymbol{b}_0(\boldsymbol{y}(t))\, [\boldsymbol{y}_\varepsilon(t) - \boldsymbol{y}(t)],\ dW_s \big\rangle \\
& - \frac{1}{4\varepsilon} \int_0^{\delta_\varepsilon} \big| \boldsymbol{b}_0(\boldsymbol{y}_\varepsilon(t)) - \boldsymbol{b}_0(\boldsymbol{y}(t)) + D\boldsymbol{b}_0(\boldsymbol{y}(t))\, [\boldsymbol{y}_\varepsilon(t) - \boldsymbol{y}(t)\,]\big|^2\, ds.
\end{aligned}
$$

Thus, the left-hand side of (3.16) is bounded by

$$
\frac{1}{2\varepsilon} \int_0^{\delta_\varepsilon} \mathbb{E}_{\boldsymbol{x}} \left[\big| \boldsymbol{b}_0(\boldsymbol{y}_\varepsilon(t)) - \boldsymbol{b}_0(\boldsymbol{y}(t)) + D\boldsymbol{b}_0(\boldsymbol{y}(t))\, [\,\boldsymbol{y}_\varepsilon(t) - \boldsymbol{y}(t)\,]\big|^2 \right] ds.
\tag{3.18}
$$

By condition (3.3) on $D\boldsymbol{b}_0$ (which is milder than that of [1]), and the argument presented in [1, Proposition 3.3], we can conclude that $d_{\mathrm{TV}} \left(\boldsymbol{y}_\varepsilon(\delta_\varepsilon; \boldsymbol{x}),\ \boldsymbol{z}_\varepsilon(\delta_\varepsilon; \boldsymbol{x}) \right) \leq \delta_\varepsilon^{1/2}$. We emphasize that in order to control the term $\boldsymbol{b}_0(\boldsymbol{y}_\varepsilon(t)) - \boldsymbol{b}_0(\boldsymbol{y}(t))$, we need the estimate of the fourth moment stated in Proposition 3.8. In all other places, a bound of the second moment suffices.

By Lemma 3.7, the probability that the starting point $\boldsymbol{y}_\varepsilon(t_\varepsilon - \delta_\varepsilon; \boldsymbol{x})$ does not belong to $\mathcal{D}_{r_4}$ vanishes as $\varepsilon \to 0$. This fact together with the bound obtained in the previous paragraph yields that

$$
\limsup_{\varepsilon \to 0} \sup_{\boldsymbol{x} \in \mathcal{D}} d_{\mathrm{TV}} \left(\boldsymbol{y}_\varepsilon(\delta_\varepsilon; \boldsymbol{y}_\varepsilon(t_\varepsilon - \delta_\varepsilon; \boldsymbol{x})),\ \boldsymbol{z}_\varepsilon(\delta_\varepsilon; \boldsymbol{y}_\varepsilon(t_\varepsilon - \delta_\varepsilon; \boldsymbol{x})) \right) = 0.
$$

This completes the estimate of the first term on the right-hand side of (3.15).

We turn to the second term. By Proposition 3.8 the starting points $\boldsymbol{y}_\varepsilon(t_\varepsilon - \delta_\varepsilon; \boldsymbol{x})$ and $\boldsymbol{z}_\varepsilon(t_\varepsilon - \delta_\varepsilon; \boldsymbol{x})$ are close. Since the process $\boldsymbol{z}_\varepsilon(\cdot)$ is Gaussian, the distance

$$
d_{\mathrm{TV}} \left(\boldsymbol{z}_\varepsilon(\delta_\varepsilon; \boldsymbol{w}),\ \boldsymbol{z}_\varepsilon(\delta_\varepsilon; \boldsymbol{w}') \right)
$$

is complete determined by $\boldsymbol{w}$ and $\boldsymbol{w}'$, and one can follow the arguments presented in [1, Proposition 3.3]. All error term appearing in the proof are uniform on the starting point $\boldsymbol{x} \in \mathcal{D}_{r_4}$ because all estimates obtained in the next subsections are uniform. Thus,

$$\lim_{\varepsilon \to 0} \sup_{\boldsymbol{x} \in \mathcal{D}_{r_4}} d_{\mathrm{TV}} \left(\boldsymbol{z}_\varepsilon(\delta_\varepsilon; \boldsymbol{y}_\varepsilon(t_\varepsilon - \delta_\varepsilon; \boldsymbol{x})), \, \boldsymbol{z}_\varepsilon(\delta_\varepsilon; \boldsymbol{z}_\varepsilon(t_\varepsilon - \delta_\varepsilon; \boldsymbol{x})) \right) = 0.$$

This completes the proof of Lemma 3.3.

Exponential Stability

In this subsection and in the next we provide the estimates used in the proofs of Lemmata 3.3 and 3.4. Recall that we denote by $\lambda > 0$ the smallest eigenvalue of $\mathbb{H}_0$. The following lemma substitutes [1, display (2.2)].

Lemma 3.5. *For all $t \geq 0$,*

$$\langle \boldsymbol{y}(t), \, \mathbb{H}_0 \, \boldsymbol{y}(t) \rangle \leqq \mathrm{e}^{-\lambda t} \, \langle \boldsymbol{y}(0), \, \mathbb{H}_0 \, \boldsymbol{y}(0) \rangle. \tag{3.19}$$

Proof. By (3.2) and (3.4),

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle \boldsymbol{y}(t), \, \mathbb{H}_0 \, \boldsymbol{y}(t) \rangle = 2 \langle \boldsymbol{b}_0(\boldsymbol{y}(t)), \, \mathbb{H}_0 \, \boldsymbol{y}(t) \rangle \leq - |\mathbb{H}_0 \, \boldsymbol{y}(t)|^2$$

$$\leq - \lambda \, \langle \boldsymbol{y}(t), \, \mathbb{H}_0 \, \boldsymbol{y}(t) \rangle \tag{3.20}$$

since λ is the smallest eigenvalue of $\mathbb{H}_0$. □

Remark 3.6. Fix $r > 0$. By the previous lemma, $\boldsymbol{y}(t) \in \mathcal{D}_r$ for all $t \geq 0$ provided $\boldsymbol{y}(0) \in \mathcal{D}_r$.

A similar computation for $\boldsymbol{y}_\varepsilon(t)$ instead of $\boldsymbol{y}(t)$ yields the moment estimate stated in the next lemma. This bound plays the role of [1, condition (H)]. Denote by $\mathbb{E}_{\boldsymbol{x}}$ the expectation with respect to $\boldsymbol{y}_\varepsilon(\cdot)$ starting at $\boldsymbol{x}$. Moreover, from now on, all estimates presented hold only for sufficiently small ε.

Lemma 3.7. *Fix $r > 0$. For all $n \geq 1$, there exists a constant $C(n) > 0$ such that*

$$\sup_{t \geq 0} \sup_{\boldsymbol{x} \in \mathcal{D}_r} \mathbb{E}_{\boldsymbol{x}} \langle \boldsymbol{y}_\varepsilon(t), \, \mathbb{H}_0 \, \boldsymbol{y}_\varepsilon(t) \rangle^n \leqq e^{-(n\lambda/4)\,t} \, \langle \boldsymbol{x}, \, \mathbb{H}_0 \, \boldsymbol{x} \rangle + C(n) \, \varepsilon.$$

Proof. By Ito's formula, (3.4), and a similar computation as to (3.20), we get

$$d \langle \boldsymbol{y}_\varepsilon(t), \, \mathbb{H}_0 \, \boldsymbol{y}_\varepsilon(t) \rangle \leq \left[-\lambda \langle \boldsymbol{y}_\varepsilon(t), \, \mathbb{H}_0 \, \boldsymbol{y}_\varepsilon(t) \rangle + 2\mathfrak{h}\varepsilon \right] dt + \sqrt{2\varepsilon} \, \langle 2\mathbb{H}_0 \, \boldsymbol{y}_\varepsilon(t), \, \mathrm{d}W_t \rangle, \tag{3.21}$$

where $\mathfrak{h} = \mathrm{tr}(\mathbb{H}_0)$. Thus, by Ito's formula and (3.21),

$$d \langle \boldsymbol{y}_\varepsilon(t), \, \mathbb{H}_0 \, \boldsymbol{y}_\varepsilon(t) \rangle^n \leq n \langle \boldsymbol{y}_\varepsilon(t), \, \mathbb{H}_0 \, \boldsymbol{y}_\varepsilon(t) \rangle^{n-1} \left[-\lambda \langle \boldsymbol{y}_\varepsilon(t), \, \mathbb{H}_0 \, \boldsymbol{y}_\varepsilon(t) \rangle + 2\mathfrak{h}\varepsilon \right] dt$$

$$+ \sqrt{2\varepsilon} \, \langle 2\mathbb{H}_0 \, \boldsymbol{y}_\varepsilon(t), \, \mathrm{d}W_t \rangle + \frac{n(n-1)}{2} \langle \boldsymbol{y}_\varepsilon(t), \, \mathbb{H}_0 \, \boldsymbol{y}_\varepsilon(t) \rangle^{n-2} \times 8\varepsilon \, |\mathbb{H}_0 \, \boldsymbol{y}_\varepsilon(t)|^2 \, dt.$$

For ε sufficiently small, this expression is bounded by

$$\langle \mathbf{y}_\varepsilon(t),\ \mathbb{H}_0\, \mathbf{y}_\varepsilon(t)\rangle^{n-1} \left[\ -\frac{n\lambda}{2}\, \langle \mathbf{y}_\varepsilon(t),\ \mathbb{H}_0\, \mathbf{y}_\varepsilon(t)\rangle + 2\mathfrak{h} n\varepsilon\ \right] \mathrm{d}t$$
$$+ \sqrt{2\varepsilon}\, \langle 2\mathbb{H}_0\, \mathbf{y}_\varepsilon(t),\ \mathrm{d}W_t\rangle.$$

Since

$$\langle \mathbf{y}_\varepsilon(t),\ \mathbb{H}_0\, \mathbf{y}_\varepsilon(t)\rangle^{n-1}\ \leq\ \frac{n-1}{n}\, \langle \mathbf{y}_\varepsilon(t),\ \mathbb{H}_0\, \mathbf{y}_\varepsilon(t)\rangle^n + \frac{1}{n},$$

for small enough $\varepsilon > 0$,

$$d\, \langle \mathbf{y}_\varepsilon(t),\ \mathbb{H}_0\, \mathbf{y}_\varepsilon(t)\rangle^n\ \leq\ \left[\ -\frac{n\lambda}{4}\, \langle \mathbf{y}_\varepsilon(t),\ \mathbb{H}_0\, \mathbf{y}_\varepsilon(t)\rangle^n + c(n)\varepsilon\ \right] \mathrm{d}t$$
$$+ \sqrt{2\varepsilon}\, \langle 2\mathbb{H}_0\, \mathbf{y}_\varepsilon(t),\ \mathrm{d}W_t\rangle$$

for some finite constant $c(n)$. Hence, by Gronwall's inequality,

$$\mathbb{E}_{\mathbf{x}}\, \langle \mathbf{y}_\varepsilon(t),\ \mathbb{H}_0\, \mathbf{y}_\varepsilon(t)\rangle^n \leq e^{-(n\lambda/4)\,t}\, \langle \mathbf{x},\ \mathbb{H}_0\, \mathbf{x}\rangle + \frac{4c(n)\varepsilon}{n\lambda},$$

as claimed   □

It follows from the estimates derived in the previous lemma, the argument presented in [1, page 1192] (cf. the last line of the proof of [1, Proposition 3.7]) and the dominated and monotone convergence theorems that there exists a finite constant C_0 such that

$$\pi_\varepsilon((\mathcal{D}_{r_4})^c)\ \leqq\ C_0\, \varepsilon \tag{3.22}$$

for all ε sufficiently small.

Gaussian Approximation

Hereafter, we couple the processes $\mathbf{y}_\varepsilon(\cdot)$, $\xi(\cdot)$, and $\mathbf{z}_\varepsilon(\cdot)$ by using the same driving Brownian motion W_t. This coupled probability law and associated expectation will be denoted by $\mathbf{p}_{\mathbf{x}}$ and $\mathbf{E}_{\mathbf{x}}$.

Proposition 3.8. (Gaussian approximation) *There exist constants $\alpha_1,\ \alpha_2 > 0$ such that*

$$\sup_{t \leq t_\varepsilon}\ \sup_{\mathbf{x} \in \mathcal{D}_{r_4}}\ \mathbf{E}_{\mathbf{x}}\left[\ |\, \mathbf{y}_\varepsilon(t; \mathbf{x}) - \mathbf{z}_\varepsilon(t; \mathbf{x})\, |^4\ \right]\ \leqq\ \alpha_1\, \varepsilon^{2+\alpha_2}.$$

This proposition corresponds to [1, display (3.12)] which plays a crucial role in the proof of the main result. Since the proof of [1, display (3.12)] requires conditions (C) and (H) of [1], and these conditions are not assumed here, we develop an alternative approach below, based on [27].

Lemma 3.9. *There exists $c > 0$ such that*

$$\langle \mathbb{K}x, \, \mathbb{A}(y(t))x \rangle \, \leq \, - \, c \, \langle x, \, \mathbb{K}x \rangle.$$

for all $t \geq 0$ and $x \in \mathcal{D}_{2r_4}$.

Proof. By (3.8) and (3.6),

$$2\langle \mathbb{K}x, \, \mathbb{A}(y(t))x \rangle = \left\langle x, \, \left[\mathbb{A}(y(t))^{\dagger}\mathbb{K} + \mathbb{K}\mathbb{A}(y(t)) \right] x \right\rangle$$
$$\leq \left\langle x, \, \left[\mathbb{A}^{\dagger}\mathbb{K} + \mathbb{K}\mathbb{A} \right] x \right\rangle + \frac{1}{2}|x|^2 = -\frac{1}{2}|x|^2 \, .$$

As $\mathbb{K}$ is bounded, the previous term is less than or equal to $- \, c \, \langle x, \, \mathbb{K}x \rangle$ for some positive contant c, as claimed. $\qquad\square$

Lemma 3.10. *For all $n \geq 1$, there exists a finite constant $C(n) > 0$ such that*

$$\sup_{t \geq 0} \, \sup_{x \in \mathcal{D}_{r_4}} \, E_x \left[\langle \boldsymbol{\xi}(t), \, \mathbb{K}\boldsymbol{\xi}(t) \rangle^n \right] \leqq C(n).$$

Proof. By Ito's formula and Lemma 3.9,

$$d\langle \boldsymbol{\xi}(t), \, \mathbb{K}\boldsymbol{\xi}(t) \rangle = \left[2 \langle \mathbb{K}\boldsymbol{\xi}(t), \, \mathbb{A}(y(t))\boldsymbol{\xi}(t) \rangle + \mathfrak{k} \right] dt + 2 \langle \mathbb{K}\boldsymbol{\xi}(t), \, dW_t \rangle$$
$$\leq \left[-2c\langle \boldsymbol{\xi}(t), \, \mathbb{K}\boldsymbol{\xi}(t) \rangle + \mathfrak{k} \right] dt + 2\langle \mathbb{K}\boldsymbol{\xi}(t), \, dW_t \rangle , \qquad (3.23)$$

where $\mathfrak{k} = \text{tr}(\mathbb{K})$. The remainder of the proof is identical to the proof of Lemma 3.7. $\qquad\square$

Lemma 3.11. *For all $n \geq 1$, there exists a finite constant $C(n) > 0$ such that*

$$\sup_{x \in \mathcal{D}_{r_4}} \, E_x \left[\sup_{t \in [0, \, t_{\varepsilon}]} \langle \boldsymbol{\xi}(t), \, \mathbb{K}\boldsymbol{\xi}(t) \rangle^n \right] \leqq C(n) \, t_{\varepsilon}^n.$$

Proof. By Hölder's inequality, it is enough to prove the lemma for n even. Assume that this is the case. Integrating (3.23), as the first term on the right-hand side is negative,

$$\langle \boldsymbol{\xi}(t), \, \mathbb{K}\boldsymbol{\xi}(t) \rangle \leqq \langle x, \, \mathbb{K}x \rangle + 2 \int_0^t \langle \mathbb{K}\boldsymbol{\xi}(t), \, dW_t \rangle ds + \mathfrak{k}t.$$

Therefore,

$$E_x \left[\sup_{t \in [0, \, t_{\varepsilon}]} \langle \boldsymbol{\xi}(t), \, \mathbb{K}\boldsymbol{\xi}(t) \rangle^n \right] \leqq C(n) \left(E_x \left[\sup_{t \in [0, \, t_{\varepsilon}]} \left| \int_0^t \langle \mathbb{K}\boldsymbol{\xi}(t), \, dW_t \rangle \right|^n \right] + t_{\varepsilon}^n \right)$$

$$(3.24)$$

for some finite constant $C(n)$. By the Burkholder–Davis–Gundy inequality and the Hölder inequality, the expectation on the right-hand side is bounded by

$$C(n) \, E_x \left[\left(\int_0^{t_{\varepsilon}} |\mathbb{K}\boldsymbol{\xi}(t)|^2 dt \right)^{n/2} \right] \leq C(n) \, t_{\varepsilon}^{(n/2)-1} \, E_x \left[\int_0^{t_{\varepsilon}} \langle \boldsymbol{\xi}(t), \, \mathbb{K}\boldsymbol{\xi}(t) \rangle^{n/2} \, dt \right].$$

By Fubini's theorem and by Lemma 3.10 [since n is even], this expression is less than or equal to $C(n) \, t_{\varepsilon}^{n/2}$. Inserting this bound is (3.24) completes the proof of the lemma. $\qquad\square$

The proof below is developed in [27] based on ideas of [1].

Proof of Proposition 3.8. Fix $\boldsymbol{x} \in \mathcal{D}_{r_4}$, and remember from (3.11) that $\boldsymbol{\xi}(t)$ depends on $\boldsymbol{x}$ through the dynamical system $\boldsymbol{y}(\cdot)$ which starts from $\boldsymbol{x}$. Let

$$\boldsymbol{r}_\varepsilon(t) := \frac{\boldsymbol{y}_\varepsilon(t) - \boldsymbol{z}_\varepsilon(t)}{\sqrt{2\varepsilon}} = \frac{\boldsymbol{y}_\varepsilon(t) - \boldsymbol{y}(t)}{\sqrt{2\varepsilon}} - \boldsymbol{\xi}(t). \tag{3.25}$$

We need to prove that there exist positive finite constants α_1, α_2 such that

$$\sup_{t \leq t_\varepsilon} \sup_{\boldsymbol{x} \in \mathcal{D}_{r_4}} \boldsymbol{E}_{\boldsymbol{x}} \left[\langle \boldsymbol{r}_\varepsilon(t), \mathbb{K}\boldsymbol{r}_\varepsilon(t) \rangle^2 \right] \leqq \alpha_1 \varepsilon^{\alpha_2}. \tag{3.26}$$

Since $\boldsymbol{y}_\varepsilon(t)$ and $\boldsymbol{\xi}(t)$ share the same driving Brownian motion, by (3.10), (3.2), and (3.11) that

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle \boldsymbol{r}_\varepsilon(t), \mathbb{K}\boldsymbol{r}_\varepsilon(t) \rangle = 2 \langle \mathbb{K}\boldsymbol{r}_\varepsilon(t), \mathbb{A}(\boldsymbol{y}(t)) \boldsymbol{r}_\varepsilon(t) \rangle - 2 \langle \mathbb{K}\boldsymbol{r}_\varepsilon(t), \boldsymbol{q}_\varepsilon(t) \rangle, \tag{3.27}$$

where

$$\boldsymbol{q}_\varepsilon(t) = \frac{1}{\sqrt{2\varepsilon}} \left\{ \boldsymbol{b}_0(\boldsymbol{y}_\varepsilon(t)) - \boldsymbol{b}_0(\boldsymbol{y}(t)) - (D\boldsymbol{b}_0)(\boldsymbol{y}(t)) (\boldsymbol{y}_\varepsilon(t) - \boldsymbol{y}(t)) \right\}.$$

Let

$$\mathcal{A}_\varepsilon = \mathcal{A}_\varepsilon(\boldsymbol{x}) := \left\{ \boldsymbol{y}_\varepsilon(t) \in \mathcal{D}_{2r_4} \text{ for all } t \in [0, t_\varepsilon] \right\}.$$

By Lemma 3.9 and since $\mathbb{K}$ is positive-definite and bounded, on the event $\mathcal{A}_\varepsilon$ the right-hand side of (3.27) is bounded by

$$-c \langle \boldsymbol{r}_\varepsilon(t), \mathbb{K}\boldsymbol{r}_\varepsilon(t) \rangle - 2 \langle \mathbb{K}\boldsymbol{r}_\varepsilon(t), \boldsymbol{q}_\varepsilon(t) \rangle \leq -c_1 \langle \boldsymbol{r}_\varepsilon(t), \mathbb{K}\boldsymbol{r}_\varepsilon(t) \rangle + C_2 |\boldsymbol{q}_\varepsilon(t)|^2 \tag{3.28}$$

for some finite positive constants c_1, C_2.

Fix $t \in [0, t_\varepsilon]$. Since $\boldsymbol{b}_0 \in C^2(B(\boldsymbol{0}, r_3), \mathbb{R}^d)$, by (3.9) on the event $\mathcal{A}_\varepsilon$,

$$|\boldsymbol{q}_\varepsilon(t)| \leqq \frac{C_0}{\sqrt{\varepsilon}} |\boldsymbol{y}_\varepsilon(t) - \boldsymbol{y}(t)|^2 \leq C_0 \sqrt{\varepsilon} \left\{ |\boldsymbol{r}_\varepsilon(t)|^2 + |\boldsymbol{\xi}(t)|^2 \right\},$$

for some finite constant C_0, whose value may change from line to line. The second inequality follows from (3.25). Therefore, by (3.28),

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle \boldsymbol{r}_\varepsilon(t), \mathbb{K}\boldsymbol{r}_\varepsilon(t) \rangle \leqq -c_1 \langle \boldsymbol{r}_\varepsilon(t), \mathbb{K}\boldsymbol{r}_\varepsilon(t) \rangle$$
$$+ C_3 \varepsilon \left[\langle \boldsymbol{r}_\varepsilon(t), \mathbb{K}\boldsymbol{r}_\varepsilon(t) \rangle^2 + \langle \boldsymbol{\xi}(t), \mathbb{K}\boldsymbol{\xi}(t) \rangle^2 \right].$$

Let $\mathcal{B}_\varepsilon = \mathcal{B}_\varepsilon(\boldsymbol{x})$ be the event defined by

$$\mathcal{B}_\varepsilon := \left\{ \frac{C_3 \varepsilon}{c_1} \left(C_3 \varepsilon t_\varepsilon \sup_{s \in [0, t_\varepsilon]} \langle \boldsymbol{\xi}(s), \mathbb{K}\boldsymbol{\xi}(s) \rangle^2 \right) \leqq \frac{1}{2} \right\}.$$

By Perov's inequality [39, Theorem 3.1], as $r_\varepsilon(0) = \mathbf{0}$, it follows from the previous inequality that on the event $\mathcal{B}_\varepsilon$,

$$\langle r_\varepsilon(t),\, \mathbb{K} r_\varepsilon(t)\rangle \;\leqq\; 2\, C_3\, \varepsilon\, t_\varepsilon\, e^{-c_1 t} \sup_{s \in [0,\, t_\varepsilon]} \langle \xi(s),\, \mathbb{K}\xi(s)\rangle^2\,.$$

for all $t \in [0,\, t_\varepsilon]$. Hence, by Lemma 3.11,

$$\sup_{t \in [0,\, t_\varepsilon]} \sup_{x \in \mathcal{D}_{r_4}} E_x\Big[\langle r_\varepsilon(t),\, \mathbb{K} r_\varepsilon(t)\rangle^2\, \mathbf{1}_{\mathcal{A}_\varepsilon \cap \mathcal{B}_\varepsilon}\Big] \;\leqq\; C_0\, \varepsilon^2\, t_\varepsilon^6 \;=\; C_0\, \varepsilon^{2-6\theta}.$$

Since $\theta < 1/3$, this proves (3.26) on the event $\mathcal{A}_\varepsilon \cap \mathcal{B}_\varepsilon$.

We turn to the event $(\mathcal{A}_\varepsilon \cap \mathcal{B}_\varepsilon)^c$. By the Cauchy–Schwarz inequality,

$$E_x\Big[\langle r_\varepsilon(t),\, \mathbb{K} r_\varepsilon(t)\rangle^2\, \mathbf{1}_{(\mathcal{A}_\varepsilon \cap \mathcal{B}_\varepsilon)^c}\Big]^2 \;\leqq\; E_x\Big[\langle r_\varepsilon(t),\, \mathbb{K} r_\varepsilon(t)\rangle^4\Big]\,\big\{\, P_x(\mathcal{A}_\varepsilon^c) + P_x(\mathcal{B}_\varepsilon^c)\,\big\}.$$

By (3.25) and the Cauchy–Schwarz inequality,

$$E_x\Big[\langle r_\varepsilon(t),\, \mathbb{K} r_\varepsilon(t)\rangle^4\Big]$$
$$\leqq\; \frac{C_0}{\varepsilon^4}\,\Big(E_x\Big[\langle y_\varepsilon(t),\, \mathbb{K} y_\varepsilon(t)\rangle^4 + \langle y(t),\, \mathbb{K} y(t)\rangle^4 + \varepsilon^4\, \langle \xi(t),\, \mathbb{K}\xi(t)\rangle^4\Big]\Big).$$

Hence, by Lemmata 3.5, 3.7 and 3.10,

$$E_x\Big[\langle r_\varepsilon(t),\, \mathbb{K} r_\varepsilon(t)\rangle^4\Big] \;\leq\; \frac{C_0}{\varepsilon^4}$$

for all $t \geq 0$, $x \in \mathcal{D}_{r_4}$.

It remains to show that there exist $c_0 > 0$ and $C_0 < \infty$ such that

$$\sup_{x \in \mathcal{D}_{r_4}} P_x(\mathcal{A}_\varepsilon^c) \;\leq\; C_0\, \varepsilon^{4+c_0} \quad \text{and} \quad \sup_{x \in \mathcal{D}_{r_4}} P_x(\mathcal{B}_\varepsilon^c) \;\leq\; C_0\, \varepsilon^{4+c_0}. \qquad (3.29)$$

Consider the event $\mathcal{A}_\varepsilon$. On the complement of this set,

$$\sup_{t \leq t_\varepsilon} \langle y_\varepsilon(t),\, \mathbb{H}_0 y_\varepsilon(t)\rangle \;\geq\; (2 r_4)^2.$$

By (3.21),

$$\langle y_\varepsilon(t),\, \mathbb{H}_0 y_\varepsilon(t)\rangle \;\leq\; 2\, \mathfrak{h}\, \varepsilon\, t \,+\, 2\sqrt{2\varepsilon} \int_0^t \langle \mathbb{H}_0 y_\varepsilon(s),\, \mathrm{d}W_s\rangle.$$

Thus, as $\varepsilon\, t_\varepsilon \to 0$, for ε small enough, by Markov inequality,

$$P_x(\mathcal{A}_\varepsilon^c) \;\leqq\; P_x\Big[\sup_{t \leq t_\varepsilon} \int_0^t \langle \mathbb{H}_0\, y_\varepsilon(s),\, \mathrm{d}W_s\rangle > \frac{r_1^2}{\sqrt{\varepsilon}}\Big]$$
$$\leq\; C_0\, \varepsilon^8\, E_x\Big[\sup_{t \leq t_\varepsilon} \Big|\int_0^t \langle \mathbb{H}_0\, y_\varepsilon(s),\, \mathrm{d}W_s\rangle\Big|^{16}\Big].$$

By the Burkholder–Davis–Gundy and Hölder inequalities, the right-hand side is bounded by

$$C_0 \, \varepsilon^8 \, \boldsymbol{E_x}\Big[\Big(\int_0^{t_\varepsilon} |\, \mathbb{H}_0 \, \boldsymbol{y}_\varepsilon(s) \,|^2 \, \mathrm{d}s \Big)^8 \Big] \leq C_0 \, \varepsilon^8 \, t_\varepsilon^7 \, \boldsymbol{E_x}\Big[\int_0^{t_\varepsilon} |\, \mathbb{H}_0 \, \boldsymbol{y}_\varepsilon(s) \,|^{16} \, \mathrm{d}s \Big].$$

Hence, by Lemma 3.7,

$$\sup_{\boldsymbol{x} \in \mathcal{D}_{r_4}} \boldsymbol{P_x}(\mathcal{A}_\varepsilon^c) \leq C_0 \, \varepsilon^8 \, t_\varepsilon^8 = C_0 \, \varepsilon^{8(1-\theta)}.$$

As $\theta < 1/3$, the first assertion of (3.29) holds.

We turn to the second assertion. By definition, there exists a positive constant c_0 such that

$$\mathcal{B}_\varepsilon^c = \Big\{ \sup_{s \in [0, t_\varepsilon]} \langle \boldsymbol{\xi}(s), \, \mathbb{K}\boldsymbol{\xi}(s) \rangle \geq \frac{c_0}{\varepsilon \sqrt{t_\varepsilon}} \Big\}.$$

By the Markov inequality and Lemma 3.11

$$\sup_{\boldsymbol{x} \in \mathcal{D}_{r_4}} \boldsymbol{P_x}(\mathcal{B}_\varepsilon^c) \leq C_0 \, \varepsilon^8 \, t_\varepsilon^4 \sup_{\boldsymbol{x} \in \mathcal{D}_{r_4}} \boldsymbol{E_x}\Big[\sup_{s \in [0, t_\varepsilon]} \langle \boldsymbol{\xi}(s), \, \mathbb{K}\boldsymbol{\xi}(s) \rangle^8 \Big] \leq C_0 \, \varepsilon^{8-12\theta}.$$

This proves the second assertion in (3.29) since $\theta < 1/3$.

4. Local Ergodicity

Fix $\lambda > 0$, $\boldsymbol{g} \colon \mathcal{M}_0 \to \mathbb{R}$, and recall that we denote by $\phi_\varepsilon = \phi_\varepsilon^{\lambda, \boldsymbol{g}}$ the unique solution of the resolvent equation (2.13). The main result of this section states that the solution ϕ_ε is asymptotically constant on each well $\mathcal{E}(\boldsymbol{m})$.

Theorem 4.1. *Fix $\lambda > 0$ and $\boldsymbol{g} \colon \mathcal{M}_0 \to \mathbb{R}$. For all $\boldsymbol{m} \in \mathcal{M}_0$,*

$$\lim_{\varepsilon \to 0} \sup_{x, y \in \mathcal{E}(\boldsymbol{m})} |\, \phi_\varepsilon(x) - \phi_\varepsilon(y) \,| = 0.$$

Recall from (2.14) that we represent by $\boldsymbol{x}_\varepsilon(\cdot)$ the diffusion process induced by the generator $\mathcal{L}_\varepsilon$. The proof of Theorem 4.1 is based on mixing properties of $\boldsymbol{x}_\varepsilon(\cdot)$ obtained in [1,14,27]. Denote by $\mathbb{P}_z^\varepsilon$, $z \in \mathbb{R}^d$, the law of $\boldsymbol{x}_\varepsilon(\cdot)$ starting from z. Expectation with respect to $\mathbb{P}_z^\varepsilon$, is represented by $\mathbb{E}_z^\varepsilon$.

We start with elementary facts. By equation (1.3) in [6], conditions (2.1) guarantees that the partition function Z_ε, defined by

$$Z_\varepsilon := \int_{\mathbb{R}^d} e^{-U(\boldsymbol{x})/\varepsilon} \, \mathrm{d}\boldsymbol{x} \tag{4.1}$$

is finite. In particular, the Gibbs measure

$$\mu_\varepsilon(\mathrm{d}\boldsymbol{x}) := Z_\varepsilon^{-1} \, e^{-U(\boldsymbol{x})/\varepsilon} \, \mathrm{d}\boldsymbol{x} := \mu_\varepsilon(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$$

is well defined. Moreover, by Theorem 2.2 and 2.3 in [28], the diffusion $\boldsymbol{x}_\varepsilon(\cdot)$ is positive recurrent and μ_ε is its unique invariant measure. On the other hand, as

we assumed that $\min_{x \in \mathbb{R}^d} U(x) = 0$, by [29, Proposition 3.2] or a straightforward computation, if $\mathcal{M}_\star$ representes the set of absolute minima of U,

$$Z_\varepsilon = [1 + o_\varepsilon(1)] (2\pi\varepsilon)^{d/2} \nu_\star, \quad \text{where} \quad \nu_\star = \sum_{m \in \mathcal{M}_\star} \frac{1}{\sqrt{\det \nabla^2 U(m)}}, \quad (4.2)$$

and, for a local minimum $m \in \mathcal{M}_0$,

$$\mu_\varepsilon(\mathcal{E}(m)) e^{U(m)/\varepsilon} = [1 + o_\varepsilon(1)] \frac{\nu(m)}{\nu_\star}. \quad (4.3)$$

In this formula, and throughout the article, $o_\varepsilon(1)$ represents a remainder which vanishes as $\varepsilon \to 0$, and $\nu(m)$ has been introduced in (2.8).

Denote by $\tau_\mathcal{A}$, $\mathcal{A} \subset \mathbb{R}^d$, the hitting time of the set $\mathcal{A}$:

$$\tau_\mathcal{A} := \inf\{t \geq 0 : x_\varepsilon(t) \in \mathcal{A}\}. \quad (4.4)$$

Recall from (2.12) the definition of $\mathcal{W}^r(m)$. Conditions (b) and (c) in the definition of $\mathcal{E}(m)$ guarantee that the hypotheses of Theorem 6.2 in Chapter 6 of [14] are fulfilled. This results leads to

Proposition 4.2. *Fix $h < H$, and denote by $\mathcal{A}$, $\mathcal{B}$ a connected component of the set $\{x : U(x) < h\}$, $\{x : U(x) < H\}$, respectively. Assume that $\mathcal{A} \subset \mathcal{B}$. Suppose that all critical points c of U in $\mathcal{A}$ are such that $U(c) \leq h_0$ for some $h_0 < h$. Then, for all $\eta > 0$,*

$$\limsup_{\varepsilon \to 0} \sup_{x \in \mathcal{A}} \mathbb{P}_x^\varepsilon \left[\tau_{\partial\mathcal{B}} < e^{(H - h_0 - \eta)/\varepsilon} \right] = 0. \quad (4.5)$$

In particular, for all $m \in \mathcal{M}_0$, $\eta > 0$,

$$\limsup_{\varepsilon \to 0} \sup_{x \in \mathcal{E}(m)} \mathbb{P}_x^\varepsilon \left[\tau_{\partial\mathcal{W}^{2r_0}(m)} < e^{(r_0 - \eta)/\varepsilon} \right] = 0.$$

The estimate in [14, Theorem 6.6.2] is uniform over initial points z belonging to neighborhoods of a critical points. We claim that it holds uniformly over initial points $x \in \mathcal{A}$. Indeed, by [14, Theorem 2.1.2], since the set $\mathcal{A}$ is bounded, if we denote by $\mathcal{N}$ the union of neighborhoods of all critical points of U in $\mathcal{A}$, there exist $T_0 < \infty$, such that

$$\liminf_{\varepsilon \to 0} \inf_{x \in \mathcal{A}} \mathbb{P}_x^\varepsilon [\tau_{\partial\mathcal{N}} < T_0, \tau_{\partial\mathcal{N}} < \tau_{\partial\mathcal{B}}] = 1. \quad (4.6)$$

Assertion (4.5) follows from (4.6), the strong Markov property and [14, Theorem 6.6.2]. Moreover, we could replace h_0 by the minimal value of U on $\mathcal{A}$, but that will not be needed below.

Mixing Times

Fix $\boldsymbol{m} \in \mathcal{M}_0$. All constants, functions, processes which appear in this subsection depend on $\boldsymbol{m}$, but this dependence is omitted in the notation.

Let $\boldsymbol{b}_0 \colon \mathbb{R}^d \to \mathbb{R}^d$ be the field of class C^1 defined in Appendix Appendix B. By (B.3) and condition (d) in the definition of r_0, $\boldsymbol{b}_0(\boldsymbol{x}) = \boldsymbol{b}(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathcal{W}^{3r_0}(\boldsymbol{m})$. By Proposition B.1, the vector field $\boldsymbol{b}_0$ satisfies the hypotheses of Sect. 3.

Denote by $\boldsymbol{x}_\varepsilon^F(\cdot)$ the diffusion process (2.14) with the vector field $\boldsymbol{b}_0$ replacing $\boldsymbol{b}$. Let $\mathbb{P}_{\boldsymbol{z}}^{\varepsilon, F}$, $\boldsymbol{z} \in \mathbb{R}^d$, be the law of $\boldsymbol{x}_\varepsilon^F(\cdot)$ starting from $\boldsymbol{z}$, and $p_\varepsilon^F(\boldsymbol{z}, \cdot\,; t)$ its transition kernel:

$$p_\varepsilon^F(\boldsymbol{z}, B; t) \;=\; \mathbb{P}_{\boldsymbol{z}}^{\varepsilon, F}\left[\, \boldsymbol{x}_\varepsilon^F(t) \in \mathcal{B} \,\right], \quad \boldsymbol{z} \in \mathbb{R}^d, \;\; \mathcal{B} \subseteq \mathbb{R}^d.$$

Denote by μ_ε^F the stationary state of the process $\boldsymbol{x}_\varepsilon^F(\cdot)$.

Proof of Theorem 4.1. Fix $\boldsymbol{m} \in \mathcal{M}_0$. Let

$$\boldsymbol{f}_\varepsilon(\boldsymbol{m}) := \int_{\mathbb{R}^d} \phi_\varepsilon(\boldsymbol{x})\, \mu_\varepsilon^F(\mathrm{d}\boldsymbol{x}) . \tag{4.7}$$

It is enough to prove that for all $\boldsymbol{m} \in \mathcal{M}_0$,

$$\lim_{\varepsilon \to 0}\; \sup_{\boldsymbol{x} \in \mathcal{E}(\boldsymbol{m})}\; |\phi_\varepsilon(\boldsymbol{x}) - \boldsymbol{f}_\varepsilon(\boldsymbol{m})| \;=\; 0.$$

Recall from (2.13) the definition of the function $G \colon \mathbb{R}^d \to \mathbb{R}$. By the stochastic representation of the resolvent equation,

$$\phi_\varepsilon(\boldsymbol{x}) \;=\; \mathbb{E}_{\boldsymbol{x}}^\varepsilon\!\left[\, \int_0^\infty e^{-\lambda s}\, G(\boldsymbol{x}_\varepsilon(\theta_\varepsilon^{(1)} s))\, \mathrm{d}s \,\right]. \tag{4.8}$$

Fix $0 < a < 1/3$, $0 < \eta < r_0/2$, and let $\varrho_\varepsilon = \varepsilon^{-a}$. By definition $\theta_\varepsilon^{(1)}$,

$$\varrho_\varepsilon \prec e^{(r_0 - \eta)/\varepsilon} \prec \theta_\varepsilon^{(1)}. \tag{4.9}$$

Since $\varrho_\varepsilon \prec \theta_\varepsilon^{(1)}$ and G is bounded,

$$\phi_\varepsilon(\boldsymbol{x}) \;=\; \mathbb{E}_{\boldsymbol{x}}^\varepsilon\!\left[\, \int_{\varrho_\varepsilon/\theta_\varepsilon^{(1)}}^\infty e^{-\lambda s}\, G(\boldsymbol{x}_\varepsilon(\theta_\varepsilon^{(1)} s))\, \mathrm{d}s \,\right] + \; R_\varepsilon(\boldsymbol{x}),$$

where, here and below, $R_\varepsilon(\boldsymbol{x})$ represents an error whose value may change from line to line and such that

$$\limsup_{\varepsilon \to 0}\; \sup_{\boldsymbol{y} \in \mathcal{E}(\boldsymbol{m})}\; |R_\varepsilon(\boldsymbol{y})| \;=\; 0.$$

By the Markov property,

$$\phi_\varepsilon(\boldsymbol{x}) \;=\; [1 + R_\varepsilon(\boldsymbol{x})]\, \mathbb{E}_{\boldsymbol{x}}^\varepsilon\!\left[\, \mathbb{E}_{\boldsymbol{x}_\varepsilon(\varrho_\varepsilon)}\!\left[\, \int_0^\infty e^{-\lambda s}\, G(\boldsymbol{x}_\varepsilon(\theta_\varepsilon^{(1)} s))\, \mathrm{d}s \,\right] \right] + \; R_\varepsilon(\boldsymbol{x})$$

$$= \; \mathbb{E}_{\boldsymbol{x}}^\varepsilon\!\left[\, \phi_\varepsilon(\boldsymbol{x}_\varepsilon(\varrho_\varepsilon)) \,\right] + \; R_\varepsilon(\boldsymbol{x})$$

because G is bounded. As $\varrho_\varepsilon \prec e^{(r_0-\eta)/\varepsilon}$, by Proposition 4.2 and since ϕ_ε is uniformly bounded by $(1/\lambda)\,\|g\|_\infty$,

$$\mathbb{E}^\varepsilon_x\big[\phi_\varepsilon(x_\varepsilon(\varrho_\varepsilon))\big] \;=\; \mathbb{E}^\varepsilon_x\big[\phi_\varepsilon(x_\varepsilon(\varrho_\varepsilon))\,\mathbf{1}\{\varrho_\varepsilon < \tau_{(\mathcal{W}^{2r_0}(m))^c}\}\big] \;+\; R_\varepsilon(x)$$

Recall that b and b_0 coincide on $\mathcal{W}^{3r_0}(m)$. By coupling the diffusions $x_\varepsilon(\cdot)$, $x_\varepsilon^F(\cdot)$, and in view of Proposition 4.2, the previous expectation is equal to

$$\mathbb{E}^{\varepsilon,F}_x\big[\phi_\varepsilon(x_\varepsilon^F(\varrho_\varepsilon))\,\mathbf{1}\{\varrho_\varepsilon < \tau_{(\mathcal{W}^{2r_0}(m))^c}\}\big] \;=\; \mathbb{E}^{\varepsilon,F}_x\big[\phi_\varepsilon(x_\varepsilon^F(\varrho_\varepsilon))\big] \;+\; R_\varepsilon(x)\,.$$

Mind that we changed the measure. By condition (e) in the definition of r_0, $\mathcal{E}(m) \subset \mathcal{W}^{2r_0}(m) \subset \mathcal{D}_{r_4}(m)$. Hence, by Theorem 3.1 and since ϕ_ε is uniformly bounded,

$$\mathbb{E}^{\varepsilon,F}_x\big[\phi_\varepsilon(x_\varepsilon^F(\varrho_\varepsilon))\big] \;=\; \int_{\mathbb{R}^d} \phi_\varepsilon(y)\, p_\varepsilon^F(x,\,dy;\varrho_\varepsilon) \;=\; \int_{\mathbb{R}^d} \phi_\varepsilon(y)\, \mu_\varepsilon^F(dy) \;+\; R_\varepsilon(x).$$

As the right-hand side is equal to $f_\varepsilon(m) + R_\varepsilon(x)$, the theorem is proved. □

Recall the definition of the sequence ϱ_ε introduced in (4.9). The proof of Theorem 4.1 yields

Lemma 4.3. *Fix $m \in \mathcal{M}_0$, $b > 0$. Then, for all $\mathcal{A} \subset \mathbb{R}^d$.*

$$\limsup_{\varepsilon\to0}\ \sup_{t\in[2b,\,4b]}\ \sup_{x\in\mathcal{E}(m)} \Big| \mathbb{P}^\varepsilon_x\big[x_\varepsilon(t\theta_\varepsilon^{(1)}) \in \mathcal{A}\big] - \mathbb{P}^\varepsilon_{\mu_\varepsilon^F}\big[x_\varepsilon(t\theta_\varepsilon^{(1)} - \varrho_\varepsilon) \in \mathcal{A}\big] \Big| = 0.$$

Denote by $x_\varepsilon^R(\cdot)$ the diffusion $x_\varepsilon(\cdot)$ reflected at the boundary of $\mathcal{W}^{2r_0}(m)$. Note that we omitted the dependence of $x_\varepsilon^R(\cdot)$ on m. Denote by μ_ε^R the measure μ_ε conditioned to $\mathcal{W}^{2r_0}(m)$, which is the invariant measure of the diffusion $x_\varepsilon^R(\cdot)$. Let finally $\mathbb{P}^{\varepsilon,R}_z$, $z \in \mathcal{W}^{2r_0}(m)$, be the law of $x_\varepsilon^R(\cdot)$ starting from z.

Recall that we denote by $d_{TV}(\mu,\nu)$ the total variation distance between two probability measures ν, μ defined on $\mathbb{R}^d$. Let $\mu_\varepsilon^{\mathcal{E}(m)}$ be the measure μ_ε conditioned to $\mathcal{E}(m)$. We claim that

$$\limsup_{\varepsilon\to0} d_{TV}(\mu_\varepsilon^R, \mu_\varepsilon^F) \;=\; 0. \tag{4.10}$$

Indeed, fix $\mathcal{A} \subset \mathbb{R}^d$, $x \in \mathcal{E}(m)$, and the sequence ϱ_ε introduced in (4.9). By stationarity and Theorem 3.1,

$$\mu_\varepsilon^F(\mathcal{A}) \;=\; \mathbb{P}^{\varepsilon,F}_{\mu_\varepsilon^F}\big[x_\varepsilon^F(\varrho_\varepsilon) \in \mathcal{A}\big] \;=\; \mathbb{P}^{\varepsilon,F}_x\big[x_\varepsilon^F(\varrho_\varepsilon) \in \mathcal{A}\big] \;+\; R_\varepsilon(x),$$

where adopted the convention established in the proof of Theorem 4.1 for the remainder $R_\varepsilon(x)$.

As in the proof of Theorem 4.1, introduce the event $\{\tau_{\partial\mathcal{W}^{2r_0}(m)} \le \varrho_\varepsilon\}$ and its complement. On the event $\{\tau_{\partial\mathcal{W}^{2r_0}(m)} > \varrho_\varepsilon\}$ we may replace the set $\mathcal{A}$ by $\mathcal{A} \cap \mathcal{W}^{2r_0}(m)$, and couple the process $x_\varepsilon^F(\cdot)$, $x_\varepsilon^R(\cdot)$ up to time ϱ_ε. Therefore, the probability on the right-hand side of the previous displayed equation is equal to

$$\mathbb{P}^{\varepsilon,R}_x\big[x_\varepsilon^R(\varrho_\varepsilon) \in \mathcal{A} \cap \mathcal{W}^{2r_0}(m)\big] \;+\; R_\varepsilon^{(2)} \;=\; \mathbb{P}^{\varepsilon,R}_x\big[x_\varepsilon^R(\varrho_\varepsilon) \in \mathcal{A}\big] \;+\; R_\varepsilon^{(2)},$$

where $|R_\varepsilon^{(2)}| \leq 2 \sup_{z \in \mathcal{E}(m)} \mathbb{P}_z^\varepsilon[\tau_{\partial \mathcal{W}^{2r_0}(m)} \leq \varrho_\varepsilon]$. Here, we removed the set $\mathcal{W}^{2r_0}(m)$ because x^R takes value on this set. By Proposition 4.2, $R_\varepsilon^{(2)} \to 0$.

Since the previous estimates are uniform over $x \in \mathcal{E}(m)$, we may average the probability appearing on the right-hand side of the previous displayed equation with respect to the measure $\mu_\varepsilon^{\mathcal{E}(m)}$ to get that

$$\mu_\varepsilon^F(\mathcal{A}) = \mathbb{P}^{\varepsilon,R}_{\mu_\varepsilon^{\mathcal{E}(m)}}\left[x_\varepsilon^R(\varrho_\varepsilon) \in \mathcal{A}\right] + o_\varepsilon(1).$$

Rewrite the previous probability as

$$\mathbb{P}^{\varepsilon,R}_{\mu_\varepsilon^{\mathcal{E}(m)}}\left[x_\varepsilon^R(\varrho_\varepsilon) \in \mathcal{A}\right] = \frac{1}{\mu_\varepsilon(\mathcal{E}(m))} \int_{\mathcal{E}(m)} \mathbb{P}_y^{\varepsilon,R}\left[x_\varepsilon^R(\varrho_\varepsilon) \in \mathcal{A}\right]\mu_\varepsilon(dy).$$

The measure $\mu_\varepsilon^{\mathcal{E}(m)}$ is also the measure μ_ε^R conditioned to $\mathcal{E}(m)$. Since $\mu_\varepsilon(\mathcal{W}^{2r_0}(m) \setminus \mathcal{E}(m))/\mu_\varepsilon(\mathcal{E}(m)) \to 0$, the previous expression is equal to

$$\mathbb{P}^{\varepsilon,R}_{\mu_\varepsilon^R}\left[x_\varepsilon^R(\varrho_\varepsilon) \in \mathcal{A}\right] + R_\varepsilon^{(3)},$$

where $R_\varepsilon^{(3)} \to 0$. Since μ_ε^R is the stationary state, the previous probability is equal to $\mu_\varepsilon^R(\mathcal{A})$.

Putting together the previous estimates yields that

$$\limsup_{\varepsilon \to 0} \sup_{\mathcal{A} \subset \mathbb{R}^d} \left|\mu_\varepsilon^F(\mathcal{A}) - \mu_\varepsilon^R(\mathcal{A})\right| = 0.$$

as claimed in (4.10).

Next result follows from Lemma 4.3 and (4.10), Note that the measure μ_ε^F has been replaced by μ_ε^R.

Corollary 4.4. *Fix $m \in \mathcal{M}_0$, $b > 0$, $\mathcal{A} \subset \mathbb{R}^d$. Then,*

$$\limsup_{\varepsilon \to 0} \sup_{t \in [2b, 4b]} \sup_{x \in \mathcal{E}(m)} \left| \mathbb{P}_x^\varepsilon\left[x_\varepsilon(t\theta_\varepsilon^{(1)}) \in \mathcal{A}\right] - \mathbb{P}_{\mu_\varepsilon^R}^\varepsilon\left[x_\varepsilon(t\theta_\varepsilon^{(1)} - \varrho_\varepsilon) \in \mathcal{A}\right] \right| = 0.$$

5. Exiting Neighborhoods of Unstable Critical Points

The main result of this section, Proposition 5.5, asserts that the time necessary for the diffusion $x_\varepsilon(\cdot)$ to leave neighborhoods of unstable critical points is bounded by ε^{-1}. It also characterizes the exiting sets.

Recall that $\mathcal{C}_0$ denotes the set of critical points of U and set

$$\mathcal{Y}_0 := \mathcal{C}_0 \setminus \mathcal{M}_0,$$

so that $\mathcal{Y}_0$ stands for the collection of critical points of U with index larger than 0.

By [28, Theorem 2.1], $\mathcal{M}_0$ and $\mathcal{Y}_0$ are the set of stable and unstable equilibria of the dynamical system (2.4), respectively. Let $\mathbb{H}^c = (\nabla^2 U)(c)$, $\mathbb{L}^c = (\nabla \cdot \ell)(c)$, $c \in \mathcal{C}_0$, so that $\mathbb{H}^c + \mathbb{L}^c$ denotes the Jacobian of the drift b at the critical point c. Next result asserts that critical points in $\mathcal{Y}_0$ are hyperbolic.

Lemma 5.1. *Fix $c \in \mathcal{Y}_0$. Then, the matrix $\mathbb{H}^c + \mathbb{L}^c$ is invertible and does not have a pure imaginary eigenvalue.*

Proof. Suppose, by contradiction, that ai, $a \in \mathbb{R}$, is an eigenvalue of $\mathbb{H}^c + \mathbb{L}^c$. Denote by v the unit eigenvector corresponding to ai so that $(\mathbb{H}^c + \mathbb{L}^c)v = aiv$. Thus, if $\mathbb{A}^\dagger$ represents the transpose of the matrix $\mathbb{A}$,

$$
\begin{aligned}
aiv \cdot aiv &= v \cdot (\mathbb{H}^c + \mathbb{L}^c)^\dagger (\mathbb{H}^c + \mathbb{L}^c)v \\
&= v \cdot \left\{ (\mathbb{H}^c)^\dagger \mathbb{H}^c + (\mathbb{H}^c)^\dagger \mathbb{L}^c + (\mathbb{L}^c)^\dagger \mathbb{H}^c + (\mathbb{L}^c)^\dagger \mathbb{L}^c \right\} v \,.
\end{aligned}
$$

By [28, Lemma 4.5], the matrix $\mathbb{H}^c \mathbb{L}^c$ is skew-symmetric, so that

$$
-a^2 \|v\|^2 = \|\mathbb{H}^c v\|^2 + \|\mathbb{L}^c v\|^2 \,,
$$

which is a contradiction if $a \neq 0$. If $a = 0$, $\mathbb{H}^c v = 0$ which implies that $v = 0$ since $\mathbb{H}^c$ is invertible. This is also a contradiction to the fact that v is a unit vector. $\square$

The Hartman–Grobman Theorem

Fix from now on a critical point $c \in \mathcal{Y}_0$ of index $k \geq 1$. In this subsection, we use Hartman–Grobman theorem [7, Theorem 1.47], [36, Section 2.8], to define a neighborhood of c.

Denote by $v_x(t)$, $x \in \mathbb{R}^d$, $t \geq 0$, the solution of the ODE (2.4) starting from x, and by $v_{L,x}(t) = v_{L,x}^c(t)$ the solution of the linear ODE

$$
\dot{x}(t) = -(\mathbb{H}^c + \mathbb{L}^c)(x(t) - c) \tag{5.1}
$$

starting from x. The Hartman–Grobman theorem, which can be applied in view of Lemma 5.1, reads as follows:

Theorem 5.2. *Fix $c \in \mathcal{Y}_0$. There exist open neighborhoods $\mathcal{U}_c$, $\mathcal{U}_c^L$ of c and a homeomorphism $\Xi : \mathcal{U}_c \to \mathcal{U}_c^L$ such that $\Xi(c) = c$ and $\Xi(v_x(t)) = v_{L,\Xi(x)}(t)$ for all (x, t) such that $v_x(t) \in \mathcal{U}_c$. In particular, c is the unique critical point of U in $\mathcal{U}_c$.*

Denote by $\mathcal{M}_s = \mathcal{M}_s(c)$, $\mathcal{M}_u = \mathcal{M}_u(c)$ the stable, unstable manifold of c for the dynamical system (2.4), respectively. Hence, for all $x \in \mathcal{M}_s$, $\lim_{t \to \infty} v_x(t) = c$. In contrast, for all $y \in \mathcal{M}_u$ there exists a solution $x(t)$, $t \leq 0$ of (2.4) such that

$$
x(0) = y, \quad \lim_{t \to -\infty} x(t) = c.
$$

Let $\mathcal{M}_{L,s}$, $\mathcal{M}_{L,u}$ be the stable, unstable manifold of c for the linear ODE (5.1). By Theorem 5.2, on the set $\mathcal{U}_c^L$, $\mathcal{M}_{L,s} = \Xi(\mathcal{M}_s)$, $\mathcal{M}_{L,u} = \Xi(\mathcal{M}_u)$.

Choose $r_1 > 0$ so that $B(c, r_1) \subset \mathcal{U}_c^L$. Let $\widehat{\mathcal{N}} = \widehat{\mathcal{N}}(c) := \Xi^{-1}(B(c, r_1))$. For each $y \in \widehat{\mathcal{N}} \setminus \mathcal{M}_s$, let $t(y) = t_c(y)$ be the exit time from $\widehat{\mathcal{N}}$:

$$
t(y) := \inf\{t \geq 0 : v_y(t) \notin \widehat{\mathcal{N}}\}. \tag{5.2}
$$

Clearly, $t(y) = t_L(\Xi(y))$ if $t_L(z)$ represents the exit time from $B(c, r_1)$ for the linear ODE (5.1) starting from z. Denote by $e(y) = e_c(y)$ the exit location of the dynamical systems (2.4) from the set $\widehat{\mathcal{N}}$: $e(y) := v_y(t(y))$. Here, again,

$$\Xi(e(y)) = e_L(\Xi(y)), \tag{5.3}$$

provided that $e_L(z)$ stands for the exit location from the set $B(c, r_1)$ of the linear dynamical systems (5.1) starting from z.

Let $\mathcal{J}_L^a = \mathcal{J}_L^a(c)$ be the elements of $\partial B(c, r_1)$ at distance less than a from $\mathcal{M}_{L,u} \cap \partial \mathcal{B}(c, r_1)$:

$$\mathcal{J}_L^a := \{x \in \partial B(c, r_1) : \exists\, y \in \mathcal{M}_{L,u} \cap \partial \mathcal{B}(c, r_1) \text{ such that } \|x - y\| < a\}.$$

Next result is an assertion about the linear ODE (5.1); its proof is presented in Appendix Appendix D.

Lemma 5.3. *Fix $c \in \mathcal{Y}_0$ and $a > 0$. Then, there exists $0 < r(a) < r_1$ such that $e_L(z) \in \mathcal{J}_L^a$ for all $z \in \mathcal{B}(c, r(a)) \backslash \mathcal{M}_{L,s}$.*

We turn to the construction of a second neighborhood $\mathcal{N} \subset \widehat{\mathcal{N}}$. Since $\nabla U \cdot \ell \equiv 0$, $(d/dt)U(v_x(t)) = -|\nabla U(v_x(t))|^2 < 0$ for all $x \notin \mathcal{C}_0, t > 0$. Therefore, if x is not a critical point, $U(v_x(t))$ is strictly decreasing in t, and there exists $\eta_0 = \eta_0(r_1) > 0$ such that

$$\max_{x \in \mathcal{M}_u \cap \partial \widehat{\mathcal{N}}} U(x) < U(c) - 3\eta_0. \tag{5.4}$$

Take η_0 small enough so that there is no critical point $c' \in \mathcal{C}_0$ such that

$$U(c') \in [U(c) - \eta_0, U(c)). \tag{5.5}$$

Lemma 5.4. *For all $c \in \mathcal{Y}_0$, there exists $r_2 = r_2(c) > 0$ such that,*

$$\sup_{y \in \Xi^{-1}(B(c, r_2)) \backslash \mathcal{M}_s} U(e(y)) \leq U(c) - 2\eta_0.$$

Proof. For $a > 0$, let

$$\mathcal{J}^a = \{x \in \partial \widehat{\mathcal{N}} : \exists\, y \in \mathcal{M}_u \cap \partial \widehat{\mathcal{N}} \text{ such that } \|x - y\| < a\}.$$

By (5.4) and the fact that $|\nabla U|$ is bounded on compact sets, there exists $a_0 > 0$ such that

$$\sup_{x \in \mathcal{J}^{a_0}} U(x) \leq U(c) - 2\eta_0. \tag{5.6}$$

Since $\Xi^{-1}: \mathcal{U}_c^L \to \mathcal{U}_c$ is continuous, it is uniformly continuous on the compact set $\overline{B(c, r_1)}$. Therefore, there exists $b_0 > 0$ such that $\|\Xi^{-1}(x) - \Xi^{-1}(y)\| \leq a_0$ for all $x, y \in B(c, r_1)$ satisfying $\|x - y\| \leq b_0$, Therefore,

$$\Xi^{-1}(\mathcal{J}_L^{b_0}) \subset \mathcal{J}^{a_0}. \tag{5.7}$$

Let $r(b_0) > 0$ be the positive constant whose existence is asserted in Lemma 5.3. Set $r_2 = r(b_0) \wedge r_1$. By Lemma 5.3, $e_L(\Xi(y)) \in \mathcal{J}_L^{b_0}$ for $\Xi(y) \in B(c, r_2) \backslash \Xi(\mathcal{M}_s)$. Therefore, by (5.3) and (5.7), for $y \in \Xi^{-1}(B(c, r_2)) \setminus \mathcal{M}_s$,

$$e(y) = \Xi^{-1}(e_L(\Xi(y))) \in \mathcal{J}^{a_0}.$$

This along with (5.6) imply that

$$\sup_{y \in \Xi^{-1}(B(c, r_2)) \backslash \mathcal{M}_s} U(e(y)) \leq U(c) - 2\eta_0,$$

which completes the proof of the lemma. $\qquad\square$

Exit problem from $\widehat{\mathcal{N}}$

Denote by $\mathcal{N} = \mathcal{N}(c)$ the closure of the set $\Xi^{-1}(B(c, r_2))$, where r_2 has been introduced in Lemma 5.4. As the set $\widehat{\mathcal{N}}$ contains an unstable equilibrium c, the exit problem from $\widehat{\mathcal{N}}$ does not follow from the Friedlin-Wentzell theory, but has been investigated in [17].

Proposition 5.5. *Fix $c \in \mathcal{Y}_0$. Then,*

$$\limsup_{\varepsilon \to 0} \sup_{z \in \mathcal{N}} \mathbb{P}_z^\varepsilon \big[U(x_\varepsilon(\tau_{\partial\widehat{\mathcal{N}}})) > U(c) - \eta_0 \big] = 0.$$

Moreover, for all $C > 0$,

$$\limsup_{\varepsilon \to 0} \sup_{z \in \mathcal{N}} \mathbb{P}_z^\varepsilon \Big[\tau_{\partial\widehat{\mathcal{N}}} > \frac{C}{\varepsilon} \Big] = 0.$$

Proof. Since the set $\widehat{\mathcal{N}}$ contains only one unstable equilibrium, the second assertion of the proposition follow from [17, Theorem 2.1], which presents an estimate for a fixed starting point in the interior of $\widehat{\mathcal{N}}$. However, a careful reading of the proof reveals that all estimates hold uniformly on compact subsets of the interior of $\widehat{\mathcal{N}}$, such as $\mathcal{N}$.

We turn to the first assertion of the proposition. Let $\mathcal{Q} \subset \partial\widehat{\mathcal{N}}$ be given by

$$\mathcal{Q} = \{e(y) : y \in \mathcal{N} \setminus \mathcal{M}_s\} \cup (\mathcal{M}_u \cap \partial\widehat{\mathcal{N}}).$$

By [17, Theorem 2.3], for any open neighborhood $\mathcal{U} \subset \partial\widehat{\mathcal{N}}$ of $\mathcal{Q}$ in $\partial\widehat{\mathcal{N}}$,

$$\limsup_{\varepsilon \to 0} \sup_{z \in \mathcal{N}} \mathbb{P}_z^\varepsilon \big[x_\varepsilon(\tau_{\partial\widehat{\mathcal{N}}}) \notin \mathcal{U} \big] = 0.$$

Note that [17, Theorem 2.3] is stated for a fixed starting point in the interior of $\widehat{\mathcal{N}}$, but as in the first part of the proof, all estimates in the proof of this result hold uniformly on compact subsets of the interior of $\widehat{\mathcal{N}}$.

By (5.4) and Lemma 5.4,

$$\sup_{x \in \mathcal{Q}} U(x) \leq U(c) - 2\eta_0.$$

To complete the proof, it remains to choose a neighborhood $\mathcal{U}$ small enough so that

$$\sup_{x \in \mathcal{U}} U(x) \leq U(c) - \eta_0.$$

$\qquad\square$

6. Hitting Wells

The main result of this section, Theorem 6.1 below, asserts that starting from a compact set, the diffusion $x_\varepsilon(\cdot)$ hits some well $\mathcal{E}(m)$ in a time bounded by ε^{-1}.

Denote by $\mathcal{E}(\mathcal{A})$, $\mathcal{A} \subset \mathbb{R}^d$, the union of the wells in $\mathcal{A}$:

$$\mathcal{E}(\mathcal{A}) = \bigcup_{m \in \mathcal{M}_0 \cap \mathcal{A}} \mathcal{E}(m).$$

Let

$$\Lambda_H = \{x \in \mathbb{R}^d : U(x) \le H\}, \quad H \in \mathbb{R}.$$

Theorem 6.1. *Fix $H > \min_{x \in \mathbb{R}^d} U(x)$. Suppose that there is no critical point $c \in \mathcal{C}_0$ such that $U(c) = H$. Then, for all $C > 0$,*

$$\limsup_{\varepsilon \to 0} \sup_{z \in \Lambda_H} \mathbb{P}_z^\varepsilon \left[\tau_{\mathcal{E}(\Lambda_H)} > \frac{C}{\varepsilon} \right] = 0.$$

Fix $h_0 < h_1$, and denote by $\mathcal{A}$, $\mathcal{B}$ a connected component of the set $\{x \in \mathbb{R}^d : U(x) < h_0\}$, $\{x \in \mathbb{R}^d : U(x) < h_1\}$, respectively. Assume that $\mathcal{A} \subset \mathcal{B}$, and that there are no critical points c of U in $\mathcal{B} \setminus \mathcal{A}$. Then, for all $C > 0$,

$$\limsup_{\varepsilon \to 0} \sup_{z \in \mathcal{A}} \mathbb{P}_z^\varepsilon \left[\tau_{\mathcal{E}(\mathcal{B})} > \frac{C}{\varepsilon} \right] = 0.$$

Corollary 6.2. *Fix $R > 0$ large enough for Λ_R to contain all the local minima of U. For all constant $C > 0$,*

$$\limsup_{\varepsilon \to 0} \sup_{x \in \Lambda_R} \mathbb{P}_x^\varepsilon \left[\tau_{\mathcal{E}(\mathcal{M}_0)} > \frac{C}{\varepsilon} \right] = 0.$$

Next result follows from Proposition 4.2 and Theorem 6.1.

Corollary 6.3. *Fix $h_0 < h_1$, and denote by $\mathcal{A}$, $\mathcal{B}$ a connected component of the set $\{x \in \mathbb{R}^d : U(x) < h_0\}$, $\{x \in \mathbb{R}^d : U(x) < h_1\}$, respectively. Assume that $\mathcal{A} \subset \mathcal{B}$, and that there are no critical points c of U in $\mathcal{B} \setminus \mathcal{A}$. Then,*

$$\limsup_{\varepsilon \to 0} \sup_{x \in \mathcal{A}} \mathbb{P}_x^\varepsilon \left[\tau_{\partial \mathcal{B}} < \tau_{\mathcal{E}(\mathcal{B})} \right] = 0.$$

Remark 6.4. We expect the optimal time scale to be $O(\log \varepsilon^{-1})$ instead of $O(\varepsilon^{-1})$.

Denote by $\mathcal{N}(\mathcal{A})$, $\mathcal{A} \subset \mathbb{R}^d$, the union, carried over all critical points c in $\mathcal{Y}_0 \cap \mathcal{A}$, of the neighborhoods $\mathcal{N}(c)$ introduced in the previous section:

$$\mathcal{N}(\mathcal{A}) = \bigcup_{c \in \mathcal{Y}_0 \cap \mathcal{A}} \mathcal{N}(c).$$

Lemma 6.5. *Under the hypotheses of Theorem 6.1, for all $C > 0$,*

$$\limsup_{\varepsilon \to 0} \sup_{z \in \Lambda_H} \mathbb{P}_z^\varepsilon \left[\tau_{\mathcal{N}(\Lambda_H) \cup \mathcal{E}(\Lambda_H)} > \frac{C}{\varepsilon} \right] = 0,$$

$$\limsup_{\varepsilon \to 0} \sup_{z \in \mathcal{A}} \mathbb{P}_z^\varepsilon \left[\tau_{\mathcal{N}(\mathcal{B}) \cup \mathcal{E}(\mathcal{B})} > \frac{C}{\varepsilon} \right] = 0.$$

Proof. For each $z \in \Lambda_H$, $\upsilon_z(t)$ reaches the set $\mathcal{N}(\Lambda_H) \cup \mathcal{E}(\Lambda_H)$ in finite time. Therefore, the assertion of the lemma follows from [14, Theorem 2.1.2]. □

Proof of Theorem 6.1

We prove the first assertion of the theorem. The arguments for the second one are similar. Recall the definition of $\eta_0 = \eta_0(c)$, $c \in \mathcal{Y}_0$, introduced at (5.5), and let

$$\mathcal{H}(c) = \{x : U(x) \le U(c) - \eta_0\}.$$

By definition of η, there is no critical point c' of U such that $U(c') = U(c) - \eta_0$.

The proof is carried out by induction on $|\mathcal{Y}_0 \cap \Lambda_H|$, the number of critical point $c \in \mathcal{Y}_0$ which belong to Λ_H. If there are no such critical points in Λ_H the assertion of the theorem follows from Lemma 6.5.

Consider the general case. Decompose the probability $\mathbb{P}_z^\varepsilon[\tau_{\mathcal{E}(\Lambda_H)} > C\varepsilon^{-1}]$ into

$$\mathbb{P}_z^\varepsilon \left[\tau_{\mathcal{E}(\Lambda_H)} > \frac{C}{\varepsilon}, \ \tau_{\mathcal{E}(\Lambda_H)} = \tau_{\mathcal{E}(\Lambda_H) \cup \mathcal{N}(\Lambda_H)} \right]$$

$$+ \mathbb{P}_z^\varepsilon \left[\tau_{\mathcal{E}(\Lambda_H)} > \frac{C}{\varepsilon}, \ \tau_{\mathcal{E}(\Lambda_H)} > \tau_{\mathcal{E}(\Lambda_H) \cup \mathcal{N}(\Lambda_H)} \right].$$

The first probability is bounded above by $\mathbb{P}_z^\varepsilon[\tau_{\mathcal{E}(\Lambda_H) \cup \mathcal{N}(\Lambda_H)} > C\varepsilon^{-1}]$. By Lemma 6.5 this expression vanishes as $\varepsilon \to 0$. In view of this result, it remains to show that

$$\limsup_{\varepsilon \to 0} \sup_{z \in \Lambda_H} \mathbb{P}_z^\varepsilon \left[\tau_{\mathcal{E}(\Lambda_H)} > \frac{C}{\varepsilon}, \ \tau_{\mathcal{E}(\Lambda_H)} > \tau_{\mathcal{E}(\Lambda_H) \cup \mathcal{N}(\Lambda_H)}, \ \tau_{\mathcal{E}(\Lambda_H) \cup \mathcal{N}(\Lambda_H)} \le \frac{C}{2\varepsilon} \right] = 0.$$

By the strong Markov property, the last display is bounded by

$$\limsup_{\varepsilon \to 0} \sup_{z \in \mathcal{E}(\Lambda_H) \cup \mathcal{N}(\Lambda_H)} \mathbb{P}_z^\varepsilon \left[\tau_{\mathcal{E}(\Lambda_H)} > \frac{C}{2\varepsilon} \right].$$

Since $\mathbb{P}_z^\varepsilon[\tau_{\mathcal{E}(\Lambda_H)} > \frac{C}{2\varepsilon}] = 0$ if $z \in \mathcal{E}(\Lambda_H)$, it suffices show that, for each $c \in \mathcal{Y}_0 \cap \Lambda_H$,

$$\limsup_{\varepsilon \to 0} \sup_{z \in \mathcal{N}(c)} \mathbb{P}_z^\varepsilon \left[\tau_{\mathcal{E}(\Lambda_H)} > \frac{C}{2\varepsilon} \right] = 0.$$

By Proposition 5.5, it is enough to prove that

$$\limsup_{\varepsilon \to 0} \sup_{z \in \mathcal{N}(c)} \mathbb{P}_z^\varepsilon \left[\tau_{\mathcal{E}(\Lambda_H)} > \frac{C}{2\varepsilon}, \ \tau_{\partial \widehat{\mathcal{N}}(c)} \le \frac{C}{4\varepsilon}, \ x_\varepsilon(\tau_{\partial \widehat{\mathcal{N}}(c)}) \in \mathcal{H}(c) \right] = 0.$$

By the strong Markov property, the left-hand side is bounded from above by

$$\limsup_{\varepsilon \to 0} \sup_{z \in \mathcal{H}(c)} \mathbb{P}_z^\varepsilon \Big[\tau_{\mathcal{E}(\Lambda_H)} > \frac{C}{4\varepsilon} \Big] = 0.$$

As c belongs to Λ_H, $U(c) \leq H$ and $\mathcal{H}(c) \subset \Lambda_H$. Thus, $\tau_{\mathcal{E}(\Lambda_H)} \leq \tau_{\mathcal{E}(\mathcal{H}(c))}$, and it is enough to prove that

$$\limsup_{\varepsilon \to 0} \sup_{z \in \mathcal{H}(c)} \mathbb{P}_z^\varepsilon \Big[\tau_{\mathcal{E}(\mathcal{H}(c))} > \frac{C}{4\varepsilon} \Big] = 0.$$

This identity follows from the induction hypothesis. Indeed, as the critical point c belongs to Λ_H and not to $\mathcal{H}(c)$, the number of critical points in $\mathcal{Y}_0 \cap \Lambda_H$ is strictly greater than the one in $\mathcal{Y}_0 \cap \mathcal{H}(c)$.   □

We conclude this section with two results on hitting times of wells. The first one follows from Theorem 2.2 and [22, Lemma 4.2]. It will be used in Sect. 10 in the proof of Theorem 2.1. We state it here, before the proof of Theorem 2.2, to have all hitting time estimates of wells in the same section.

Lemma 6.6. *For all* $m \in \mathcal{M}_0$,

$$\limsup_{a \to 0} \limsup_{\varepsilon \to 0} \sup_{x \in \mathcal{E}(m)} \mathbb{P}_x^\varepsilon \big[\tau_{\mathcal{E}(\mathcal{M}_0) \backslash \mathcal{E}(m)} \leq a\, \theta_\varepsilon^{(1)} \big] = 0.$$

The last result asserts that starting from the domain of attraction of a local minima the well associated to this local minimum is attained before the other ones. Recall that we denote by $v_x(t)$, $x \in \mathbb{R}^d$, $t \geq 0$, the solution of the ODE (2.4) starting from x. Denote by $\mathcal{D}(m)$, $m \in \mathcal{M}_0$, the domain of attraction of m:

$$\mathcal{D}(m) = \big\{ x \in \mathbb{R}^d : \lim_{t \to \infty} v_x(t) = m \big\}.$$

Lemma 6.7. *Let* $m \in \mathcal{M}_0$ *and* $\mathcal{K}$ *be a compact subset of* $\mathcal{D}(m)$. *Then*,

$$\liminf_{\varepsilon \to 0} \inf_{x \in \mathcal{K}} \mathbb{P}_x^\varepsilon \big[\tau_{\mathcal{E}(\mathcal{M}_0)} = \tau_{\mathcal{E}(m)} \big] = 1.$$

Proof. Let $\mathcal{F}(m) := \mathcal{D}(m) \backslash \mathcal{E}(m)$ so that $\partial \mathcal{F}(m) = \partial \mathcal{D}(m) \cup \partial \mathcal{E}(m)$. Then,

$$\mathbb{P}_x^\varepsilon \big[\tau_{\mathcal{E}(\mathcal{M}_0)} = \tau_{\mathcal{E}(m)} \big] \geq \mathbb{P}_x^\varepsilon \big[\tau_{\partial \mathcal{F}(m)} = \tau_{\partial \mathcal{E}(m)} \big].$$

Therefore, it suffices to show that

$$\liminf_{\varepsilon \to 0} \inf_{x \in \mathcal{K}} \mathbb{P}_x^\varepsilon \big[\tau_{\partial \mathcal{F}(m)} = \tau_{\partial \mathcal{E}(m)} \big] = 1. \tag{6.1}$$

Since $\mathcal{K}$ is contained in the domain of attraction of m, the solution $v_x(t)$ of the ODE (2.4) starting from $x \in \mathcal{K}$ exits the domain $\mathcal{F}(m)$ at $\partial \mathcal{E}(m)$. Thus, by [14, Chapter 2, Theorem 1.2], (6.1) holds.

The estimate in [14, Chapter 2, Theorem 1.2] is not uniform in x, and just asserts that

$$\liminf_{\varepsilon \to 0} \mathbb{P}_x^\varepsilon \big[\tau_{\partial \mathcal{F}(m)} = \tau_{\partial \mathcal{E}(m)} \big] = 1$$

for all $x \in \mathcal{D}(m)$. However, the bound [14, Chapter 2, Theorem 1.2] holds uniformly over $x \in \mathcal{K}$ (The variable $a(t)$ in the proof depends on x but can be bounded uniformly on any compact set of $\mathcal{D}(m)$, see the displayed equation above (1.6)). This completes the proof of the lemma.   □

7. Test Functions

In this section, we construct the test functions used in Sect. 9 to estimate the solution of the resolvent equation. This test function appeared before in [6,23,28]. For this reason we just present its definition and main properties, and refer the reader to [29] for proofs.

Fix a local minimum $\boldsymbol{m}$. The test function proposed below and denoted by Q_ε is an approximation of the equilibrium potential $h\colon \mathbb{R}^d \to [0, 1]$ defined by $h(\boldsymbol{x}) = \mathbb{P}^\varepsilon_{\boldsymbol{x}}[\tau_{\mathcal{E}(\boldsymbol{m})} < \tau_{\mathcal{E}(\mathcal{M}_0)\setminus\mathcal{E}(\boldsymbol{m})}]$. In particular, inside the wells the test function will be either very close to 1 or very close to 0. In contrast, in very small neighborhoods of saddle points it will change from 0 to 1. To capture this behavior, we linearize the generator at the saddle point and set Q_ε to be close to the equilibrium potential for the linearized generator.

Around a Saddle Point

Fix a saddle point $\boldsymbol{\sigma}$ of U such that $\boldsymbol{m} \curvearrowleft \boldsymbol{\sigma} \curvearrowright \boldsymbol{m}'$ for distinct local minima $\boldsymbol{m}, \boldsymbol{m}'$ of U. Let $\mathbb{H}^\sigma = \nabla^2 U(\boldsymbol{\sigma})$, $\mathbb{L}^\sigma = (D\boldsymbol{\ell})(\boldsymbol{\sigma})$. By (2.6), $\mathbb{H}^\sigma$ has a unique negative eigenvalue. Denote by $-\lambda_1, \lambda_2, \ldots, \lambda_d$ the eigenvalues of $\mathbb{H}^\sigma$, where $-\lambda_1$ represents the unique negative eigenvalue. Mind that we omit the dependence on $\boldsymbol{\sigma}$ which is fixed. Let $\boldsymbol{e}_1, \boldsymbol{e}_k, k \geq 2$, be the unit eigenvector associated with the eigenvalue $-\lambda_1, \lambda_k$, respectively. Choose $\boldsymbol{e}_1$ pointing towards $\boldsymbol{m}$: for all sufficiently small $a > 0$, $\boldsymbol{\sigma} + a\boldsymbol{e}_1$ belongs to the domain of attraction of $\boldsymbol{m}$. For $\boldsymbol{x} \in \mathbb{R}^d$ and $k = 1, \ldots, d$, write $\hat{x}_k = (\boldsymbol{x} - \boldsymbol{\sigma}) \cdot \boldsymbol{e}_k$, so that $\boldsymbol{x} = \boldsymbol{\sigma} + \sum_{m=1}^d \hat{x}_m \boldsymbol{e}_m$.

Let

$$\delta = \delta(\varepsilon) := (\varepsilon \log \frac{1}{\varepsilon})^{1/2} .$$

Fix a large constant $J > 0$ to be chosen later, and denote by $\mathcal{A}_\varepsilon^\pm$, $\mathcal{C}_\varepsilon$ the d-dimensional rectangles defined by

$$\mathcal{A}_\varepsilon^- := \left\{ \boldsymbol{x} \in \mathbb{R}^d : \hat{x}_1 \in \left[-\frac{J\delta}{\sqrt{\lambda_1}} - \varepsilon^2, -\frac{J\delta}{\sqrt{\lambda_1}} \right], \hat{x}_k \in \left[-\frac{2J\delta}{\sqrt{\lambda_k}}, \frac{2J\delta}{\sqrt{\lambda_k}} \right], 2 \leqq k \leqq d \right\}$$

$$\mathcal{C}_\varepsilon := \left\{ \boldsymbol{x} \in \mathbb{R}^d : \hat{x}_1 \in \left[-\frac{J\delta}{\sqrt{\lambda_1}}, \frac{J\delta}{\sqrt{\lambda_1}} \right], \hat{x}_k \in \left[-\frac{2J\delta}{\sqrt{\lambda_k}}, \frac{2J\delta}{\sqrt{\lambda_k}} \right], 2 \leqq k \leqq d \right\}$$

$$\mathcal{A}_\varepsilon^+ := \left\{ \boldsymbol{x} \in \mathbb{R}^d : \hat{x}_1 \in \left[\frac{J\delta}{\sqrt{\lambda_1}}, \frac{J\delta}{\sqrt{\lambda_1}} + \varepsilon^2 \right], \hat{x}_k \in \left[-\frac{2J\delta}{\sqrt{\lambda_k}}, \frac{2J\delta}{\sqrt{\lambda_k}} \right], 2 \leqq k \leqq d \right\} .$$

Fig. 1 illustrates these definitions and the next ones.

Recall from (2.8) that $\mathbb{H}^\sigma + \mathbb{L}^\sigma$ has a unique negative eigenvalue, denoted by $-\mu$. Denote by $\mathbb{A}^\dagger$ the transpose of a matrix $\mathbb{A}$. By [28, display (8.1)], the matrix $\mathbb{H}^\sigma - (\mathbb{L}^\sigma)^\dagger$ also has a unique negative eigenvalue equal to $-\mu$. Denote by $\boldsymbol{v}$ the unit eigenvector of $\mathbb{H}^\sigma - (\mathbb{L}^\sigma)^\dagger$ associated with $-\mu$. By [28, Lemma 8.1], $\boldsymbol{v} \cdot \boldsymbol{e}_1 \neq 0$. We assume that $\boldsymbol{v} \cdot \boldsymbol{e}_1 > 0$, as we can take $-\boldsymbol{v}$ instead of $\boldsymbol{v}$ if this inner product is negative.

Let $p_\varepsilon\colon \mathcal{C}_\varepsilon \to \mathbb{R}$ be given by

$$p_\varepsilon(\boldsymbol{x}) := \frac{1}{M_\varepsilon} \int_{-\infty}^{(\boldsymbol{x}-\boldsymbol{\sigma})\cdot\boldsymbol{v}} e^{-\frac{\mu}{2\varepsilon}t^2} \, dt, \tag{7.1}$$
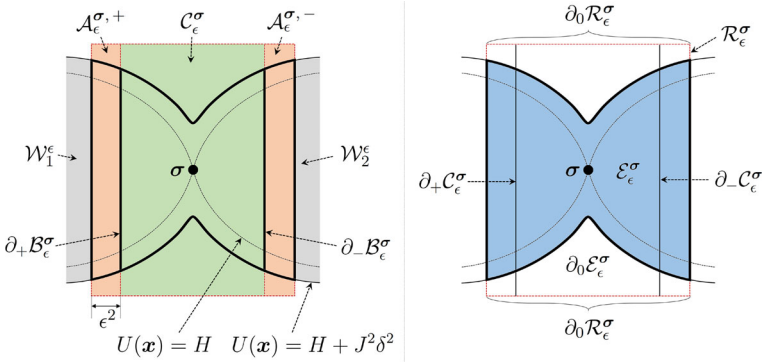
Fig. 1. The sets around a saddle point σ appearing in the definition of the test function

where the normalizing constant M_ε is given by

$$M_\varepsilon = \int_{-\infty}^{\infty} e^{-\frac{\mu}{2\varepsilon}t^2}\,\mathrm{d}t = \sqrt{\frac{2\pi\varepsilon}{\mu}}. \qquad (7.2)$$

We extend continuously the function p_ε to the d-dimensional rectangle $\mathcal{R}_\varepsilon = \mathcal{A}_\varepsilon^- \cup \mathcal{C}_\varepsilon \cup \mathcal{A}_\varepsilon^+$ as follows. For $\widehat{x} = \sigma + \sum_{k=1}^{d} \widehat{x}_k e_k \in \mathcal{A}_\varepsilon^+$, let

$$\overline{x}_r = \sigma + \frac{J\delta}{\sqrt{\lambda_1}}e_1 + \sum_{k=2}^{d}\widehat{x}_k\,e_k. \qquad (7.3)$$

We define $\overline{x}_l$ similarly for $x \in \mathcal{A}_\varepsilon^-$, replacing on the right-hand side of the previous formula the first plus sign by a minus sign. Clearly, $\overline{x}_r$ and $\overline{x}_l$ belong to $\mathcal{C}_\varepsilon$. We extend the definition of p_ε to $\mathcal{R}_\varepsilon$ by setting $p_\varepsilon : \mathcal{A}_\varepsilon^- \cup \mathcal{A}_\varepsilon^+ \to \mathbb{R}$ as

$$
\begin{aligned}
p_\varepsilon(x) &= 1 + \varepsilon^{-2}\Big[\hat{x}_1 - \frac{J\delta}{\sqrt{\lambda_1}} - \varepsilon^2\Big][1 - p_\varepsilon(\overline{x}_r)], \quad x \in \mathcal{A}_\varepsilon^+, \\
p_\varepsilon(x) &= \varepsilon^{-2}\Big[\hat{x}_1 + \frac{J\delta}{\sqrt{\lambda_1}} + \varepsilon^2\Big]p_\varepsilon(\overline{x}_l), \quad x \in \mathcal{A}_\varepsilon^-.
\end{aligned}
\qquad (7.4)
$$

The function p_ε is an approximating solution of the Dirichlet problem $\mathcal{L}_\varepsilon^\dagger f = 0$ in $\mathcal{R}_\varepsilon$ with boundary conditions $f = 1$ on the points of $\mathcal{R}_\varepsilon$ where $\hat{x}_1 = J\delta/\sqrt{\lambda_1} + \varepsilon^2$ and $f = 0$ on the ones such that $\hat{x}_1 = -J\delta/\sqrt{\lambda_1} - \varepsilon^2$. This is the content of [29, Proposition 6.2], which states that the integral of $\theta_\varepsilon^{(1)}\,|\mathcal{L}_\varepsilon^\dagger f|$ on a set slightly smaller than $\mathcal{R}_\varepsilon$ vanishes as $\varepsilon \to 0$. This result also justifies the definition of the test function p_ε.

The test function $p_\varepsilon(\cdot)$ constructed above depends on σ and m. To stress this fact, it is sometimes represented by $p_\varepsilon^{\sigma, m}(\cdot)$.
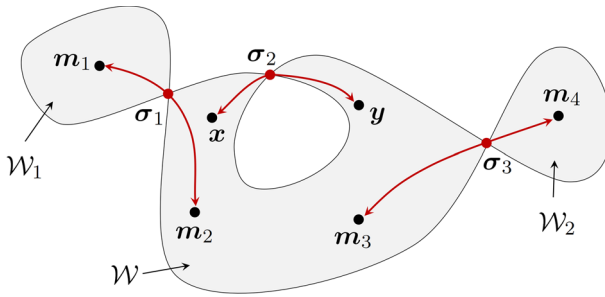
Fig. 2. The saddle points $\boldsymbol{\sigma}_2$ in $\partial \mathcal{W}$ does not belong to $\mathcal{S}_H(\mathcal{W})$ because it leads to critical points in $\mathcal{W}$

Inside a Well

In this subsection we define a test function Q_ε on $\mathbb{R}^d$ with the help of the test functions $p_\varepsilon^{\boldsymbol{\sigma},\boldsymbol{m}}$ introduced in the previous subsection. Recall that we denote by $B(\boldsymbol{x}, r)$ the open ball of radius r centered at $\boldsymbol{x}$.

Fix a height H such that $U(\boldsymbol{\sigma}) = H$ for some saddle point $\boldsymbol{\sigma}$ of U. Denote by $\mathcal{W}$ a connected component of the set $\{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) < H\}$. Assume that there exists a saddle point $\boldsymbol{\sigma}' \in \partial \mathcal{W}$ satisfying condition (a) below and that condition (b) is fulfilled for all saddle points $\boldsymbol{\sigma}' \in \partial \mathcal{W}$ satisfying (a). Here,

(a) There exists $\delta_0 > 0$ such that $B(\boldsymbol{\sigma}', \delta) \cap \{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) < H\}$ is not contained in $\mathcal{W}$ for all $0 < \delta < \delta_0$;
(b) There exists $\boldsymbol{m}, \boldsymbol{m}' \in \mathcal{M}_0$ such that

$$\boldsymbol{\sigma}' \curvearrowright \boldsymbol{m} \quad \text{and} \quad \boldsymbol{\sigma}' \curvearrowright \boldsymbol{m}'. \tag{7.5}$$

Condition (b) prevents the existence of a heteroclinic orbit from $\boldsymbol{\sigma}$ to a critical point which is not a local minimum. Clearly, if conditions (a) and (b) hold, $\boldsymbol{m} \in \mathcal{W}$ and $\boldsymbol{m}' \notin \mathcal{W}$ or the contrary.

Let $\mathcal{S}_H(\mathcal{W}) = \{\boldsymbol{\sigma}_1, \ldots \boldsymbol{\sigma}_p\}$ be the (non-empty) set of saddle points $\boldsymbol{\sigma} \in \partial \mathcal{W}$ satisfying (a) and (b). Note that there might be saddle points $\boldsymbol{\sigma}$ in $\partial \mathcal{W}$ which do not belong to $\mathcal{S}_H(\mathcal{W})$ because they lead to critical points in $\mathcal{W}$: $\boldsymbol{\sigma} \curvearrowright \boldsymbol{x}, \boldsymbol{\sigma} \curvearrowright \boldsymbol{y}$ for $\boldsymbol{x}$, $\boldsymbol{y} \in \mathcal{W} \cap \mathcal{C}_0$. Figure 2 illustrates this possibility. On the other hand, as $\boldsymbol{\sigma} \in \partial \mathcal{W}$, $U(\boldsymbol{\sigma}) = H$ for all $\boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W})$.

Write, whenever needed to clarify, $\mathcal{W}$ as $\mathcal{W}_1$, and denote by $\mathcal{W}_j$, $2 \leq j \leq m$, the connected components of the set $\{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) < H\}$ which share with $\mathcal{W}$ a common saddle point in $\mathcal{S}_H(\mathcal{W})$. Hence, for each $\mathcal{W}_j$, $2 \leq j \leq m$, there exist $\boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W}) \cap \overline{\mathcal{W}_j}$, $\boldsymbol{m} \in \mathcal{W} \cap \mathcal{M}_0$, $\boldsymbol{m}' \in \mathcal{W}_j \cap \mathcal{M}_0$ such that $\boldsymbol{\sigma} \curvearrowright \boldsymbol{m}, \boldsymbol{\sigma} \curvearrowright \boldsymbol{m}'$.

Fix $\eta > 0$ small enough so that there is no critical point $\boldsymbol{x}$ with height in the interval $(H, H + 2\eta)$. Let Ω be the connected component of the set $\{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) \leq H + \eta\}$ which contains $\mathcal{W}$ (and thus all connected components $\mathcal{W}_j$), and set

$$\mathcal{K}_\varepsilon := \left\{ \boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) \leq H + J^2 \delta^2 \right\} \cap \Omega.$$

Denote by $\partial_0 \mathcal{R}_\varepsilon^\sigma$, $\boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W})$, the boundary of the set $\mathcal{R}_\varepsilon^\sigma$, introduced in the previous subsection, given by

$$\partial_0 \mathcal{R}_\varepsilon^\sigma = \left\{ \boldsymbol{x} \in \mathcal{R}_\varepsilon^\sigma : \hat{x}_k = \pm \frac{2J\delta}{\sqrt{\lambda_k}} \text{ for some } 2 \leqq k \leqq d \right\}.$$

By the proof of [28, Lemma 8.3],

$$U(\boldsymbol{x}) \geqq U(\boldsymbol{\sigma}) + \frac{3}{2} J^2 \delta^2 [1 + o_\varepsilon(1)] \tag{7.6}$$

for all $\boldsymbol{x} \in \partial_0 \mathcal{R}_\varepsilon$. In particular, $\partial_0 \mathcal{R}_\varepsilon$ is contained in the complement of $\mathcal{K}_\varepsilon$ provided that ε is sufficiently small.

Let $\mathcal{E}_\varepsilon^\sigma := \mathcal{R}_\varepsilon^\sigma \cap \mathcal{K}_\varepsilon$, $\boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W})$. Denote by $\mathcal{W}_1^\varepsilon$ the connected component of $\mathcal{K}_\varepsilon \backslash (\bigcup_{\boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W})} \mathcal{E}_\varepsilon^\sigma)$ which intersects $\mathcal{W}_1$, and let $\mathcal{W}_2^\varepsilon = \mathcal{K}_\varepsilon \backslash (\mathcal{W}_1^\varepsilon \cup \bigcup_{\boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W})} \mathcal{E}_\varepsilon^\sigma)$. With this notation,

$$\Omega = \bigcup_{\boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W})} \mathcal{E}_\varepsilon^\sigma \cup \mathcal{W}_1^\varepsilon \cup \mathcal{W}_2^\varepsilon \cup (\Omega \backslash \mathcal{K}_\varepsilon). \tag{7.7}$$

For each $\boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W})$, denote by $\boldsymbol{m}_\sigma$ the local minimum $\boldsymbol{m}$ in $\mathcal{W}$ such that $\boldsymbol{\sigma} \curvearrowright \boldsymbol{m}$. Recall the notation introduced at the end of the previous subsection, and let $q^\sigma = p^{\boldsymbol{\sigma}, \boldsymbol{m}_\sigma}$. Consider the test function $Q_\varepsilon : \mathcal{K}_\varepsilon \to \mathbb{R}$ given by

$$Q_\varepsilon(\boldsymbol{x}) = 1, \ \boldsymbol{x} \in \mathcal{W}_1^\varepsilon; \quad Q_\varepsilon(\boldsymbol{y}) = 0, \ \boldsymbol{y} \in \mathcal{W}_2^\varepsilon;$$
$$Q_\varepsilon(\boldsymbol{x}) = q_\varepsilon^\sigma(\boldsymbol{x}), \ \boldsymbol{x} \in \mathcal{E}_\varepsilon^\sigma, \ \boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W}). \tag{7.8}$$

By (7.4), the function Q_ε is continuous on $\mathcal{K}_\varepsilon$. Moreover, if $\mathcal{G}_\varepsilon$ represents the open set formed by the union of the interiors of the set $\mathcal{E}_\varepsilon^\sigma$, $\boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W})$, and the interior of the sets $\mathcal{W}_i^\varepsilon$, $i = 1, 2$,

$$\|\nabla Q_\varepsilon\|_{L^\infty(\mathcal{G}_\varepsilon)} = O(\varepsilon^{-1/2}) \text{ and } \|\Delta Q_\varepsilon\|_{L^\infty(\mathcal{G}_\varepsilon)} = O(\varepsilon^{-3/2}).$$

We can extend Q_ε to Ω keeping these bounds out of a $(d-1)$ dimensional manifold:

$$\|Q_\varepsilon\|_{L^\infty(\Omega_0)} \leq 1, \ \|\nabla Q_\varepsilon\|_{L^\infty(\Omega_0)} = O(\varepsilon^{-1/2}), \text{ and } \|\Delta Q_\varepsilon\|_{L^\infty(\Omega_0)} = O(\varepsilon^{-3/2}) \tag{7.9}$$

where $\Omega_0 = \Omega \backslash \mathfrak{M}$, and $\mathfrak{M}$ is $(d-1)$ dimensional manifold at which the gradient of Q_ε is discontinuous. We further impose the condition that Q_ε vanishes away from Ω:

$$Q_\varepsilon \equiv 0 \text{ on } \{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) > H + \frac{\eta}{2}\}, \tag{7.10}$$

respecting the previous bounds. The function Q_ε is the test function associated to the well $\mathcal{W}$ and height H.

Main Estimate

Next lemma is a crucial step in the proof of Theorem 2.2. To stress below the dependence of the set $\mathcal{C}_\varepsilon$, $\mathcal{A}_\varepsilon^\pm$ on a saddle point σ, we add a superscript σ in the notation. Denote by $\partial_\pm \mathcal{C}_\varepsilon^\sigma$ the boundary of the set $\mathcal{C}_\varepsilon^\sigma$ given by

$$\partial_\pm \mathcal{C}_\varepsilon^\sigma \;=\; \Big\{ x \in \mathcal{C}_\varepsilon^\sigma \,:\, \hat{x}_1 = \pm \frac{J\delta}{\sqrt{\lambda_1}} \Big\}, \quad \text{and let}$$

$$\mathcal{B}_\varepsilon^\sigma \;:=\; \mathcal{C}_\varepsilon^\sigma \cap \mathcal{K}_\varepsilon, \quad \partial_\pm \mathcal{B}_\varepsilon^\sigma := \partial_\pm \mathcal{C}_\varepsilon^\sigma \cap \mathcal{K}_\varepsilon.$$

Recall from (4.2) the definition of $\nu_\star$, and from (7.3) the definition of $\overline{x}_r$, $\overline{x}_l$. In the statement of Lemma 7.1, the vectors v and e_1 depend on σ, but this dependence is omitted from the notation. For $c > 0$, let

$$\Lambda_{c,\varepsilon} := \big\{ x \in \mathbb{R}^d \,:\, U(x) \leq H - c\,J^2\delta^2 \big\}.$$

Lemma 7.1. *There exists $c_0 > 0$, such that for all $0 < c < c_0$, $\mathbf{g}\colon \mathcal{M}_0 \to \mathbb{R}$,*

$$e^{H/\varepsilon} \int_\Omega Q_\varepsilon\,(-\mathcal{L}_\varepsilon\,\phi_\varepsilon)\,d\mu_\varepsilon \;=\; \sum_{\sigma \in \mathcal{S}_H(\mathcal{W})} J(\sigma) \,+\, o_\varepsilon(1), \qquad (7.11)$$

where $J(\sigma) = J_+(\sigma) - J_-(\sigma)$, and

$$J_+(\sigma) \;=\; -\,[\,1 + o_\varepsilon(1)\,]\,\frac{\varepsilon\,\sqrt{\mu^\sigma}\,(v \cdot e_1)}{(2\pi\varepsilon)^{(d+1)/2}\,\nu_\star} \int_{\partial_+ \mathcal{B}_\varepsilon^\sigma \cap \Lambda_{c,\varepsilon}}$$

$$e^{-\frac{1}{2\varepsilon}x \cdot (\mathbb{H}^\sigma + \mu^\sigma\,v \otimes v)\,x}\,\phi_\varepsilon(x)\,\mathrm{S}(dx)$$

$$-\,[\,1 + o_\varepsilon(1)\,]\,\frac{1}{(2\pi)^{(d+1)/2}\,\nu_\star\,\sqrt{\mu^\sigma}\,\varepsilon^{(d+3)/2}} \int_{\mathcal{A}_\varepsilon^{\sigma,+} \cap \Lambda_{c,\varepsilon}} \phi_\varepsilon(x)\,\frac{\mathbb{L}^\sigma \overline{x}_r \cdot e_1}{\overline{x}_r \cdot v}$$

$$e^{-\frac{1}{2\varepsilon}\overline{x}_r \cdot (\mathbb{H}^\sigma + \mu^\sigma\,v \otimes v)\,\overline{x}_r}\,dx.$$

In this formula, $\mathrm{S}(dx)$ represents the surface measure on the $(d-1)$-dimensional manifold $\partial_+ \mathcal{B}_\varepsilon^\sigma \cap \Lambda_{c,\varepsilon}$, and $J_-(\sigma)$ is obtained from $J_+(\sigma)$ by removing the minus sign and replacing $\partial_+ \mathcal{B}_\varepsilon^\sigma$, $\mathcal{A}_{a,\varepsilon}^{\sigma,+}$ by $\partial_- \mathcal{B}_\varepsilon^\sigma$, $\mathcal{A}_{a,\varepsilon}^{\sigma,-}$, respectively.

The proof of this result is omitted as it is the content of [29, Section 7].

8. Domain of Attraction

Fix $\sigma \in \mathcal{S}_H(\mathcal{W})$. Denote by n_σ the local minimum m of U which does not belong to $\mathcal{W}$ and such that $\sigma \curvearrowright m$. The main result of this section asserts that we may replace $\phi_\varepsilon(x)$ in the formula for $J_+(\sigma)$, $J_-(\sigma)$ by $\phi_\varepsilon(m_\sigma)$, $\phi_\varepsilon(n_\sigma)$, respectively.

Proposition 8.1. *There exists $c_0 > 0$, such that for all $0 < c < c_0$,*

$$\lim_{\varepsilon \to 0} \sup_{x \in \partial_+ \mathcal{B}_\varepsilon^\sigma \cap \Lambda_{c,\varepsilon}} |\phi_\varepsilon(x) - \phi_\varepsilon(m_\sigma)| = 0, \quad \lim_{\varepsilon \to 0} \sup_{x \in \partial_- \mathcal{B}_\varepsilon^\sigma \cap \Lambda_{c,\varepsilon}} |\phi_\varepsilon(x) - \phi_\varepsilon(n_\sigma)| = 0$$

for all $\sigma \in \mathcal{S}_H(\mathcal{W})$. A similar result holds if we replace $\partial_+ \mathcal{B}_\varepsilon^\sigma$, $\partial_- \mathcal{B}_\varepsilon^\sigma$ by $\mathcal{A}_\varepsilon^{\sigma,+}$, $\mathcal{A}_\varepsilon^{\sigma,-}$, respectively.

The proof of Proposition 8.1 is based on the following general result. Recall that we denote by $\mathcal{D}(\boldsymbol{m})$, $\boldsymbol{m} \in \mathcal{M}_0$, the domain of attraction of $\boldsymbol{m}$.

Proposition 8.2. *Fix* $\boldsymbol{m} \in \mathcal{M}_0$, *and a sequence* $(\mathcal{K}_\varepsilon)_{\varepsilon>0}$ *of subsets of* $\mathcal{D}(\boldsymbol{m})$. *Assume that* $\bigcup_{\varepsilon>0} \mathcal{K}_\varepsilon$ *is a bounded set, and*

$$\liminf_{\varepsilon \to 0} \inf_{\boldsymbol{x} \in \mathcal{K}_\varepsilon} \mathbb{P}^\varepsilon_{\boldsymbol{x}} \left[\tau_{\mathcal{E}(\mathcal{M}_0)} = \tau_{\mathcal{E}(\boldsymbol{m})} \right] = 1 \tag{8.1}$$

Then,

$$\limsup_{\varepsilon \to 0} \sup_{\boldsymbol{x} \in \mathcal{K}_\varepsilon} \left| \phi_\varepsilon(\boldsymbol{x}) - \phi_\varepsilon(\boldsymbol{m}) \right| = 0.$$

Proof. Recall the definition of the function $G \colon \mathbb{R}^d \to \mathbb{R}$ introduced in (2.13). By the stochastic representation of the solution of the resolvent equation,

$$\phi_\varepsilon(\boldsymbol{x}) = \mathbb{E}^\varepsilon_{\boldsymbol{x}} \left[\int_0^\infty e^{-\lambda s} G(\boldsymbol{x}_\varepsilon(\theta^{(1)}_\varepsilon s)) \, \mathrm{d}s \right]. \tag{8.2}$$

As G is bounded, the absolute value of the time integral is bounded by $\lambda^{-1} \|\boldsymbol{g}\|_\infty$. Therefore, as $\bigcup_{\varepsilon>0} \mathcal{K}_\varepsilon$ is a bounded set and $U(\boldsymbol{x}) \to \infty$ as $|\boldsymbol{x}| \to \infty$, taking R sufficiently large in Corollary 6.2,

$$\phi_\varepsilon(\boldsymbol{x}) = \mathbb{E}^\varepsilon_{\boldsymbol{x}} \left[\int_0^\infty e^{-\lambda s} G(\boldsymbol{x}_\varepsilon(\theta^{(1)}_\varepsilon s)) \, \mathrm{d}s \, \mathbf{1}\left\{ \tau_{\mathcal{E}(\mathcal{M}_0)} \leq \frac{C_0}{\varepsilon} \right\} \right] + R_\varepsilon(\boldsymbol{x}), \tag{8.3}$$

where, here and below, $R_\varepsilon(\boldsymbol{x})$ is an error term such that

$$\lim_{\varepsilon \to 0} \sup_{\boldsymbol{x} \in \bigcup_{\varepsilon>0} \mathcal{K}_\varepsilon} |R_\varepsilon(\boldsymbol{x})| = 0.$$

Consider the time integral in the interval $[0, \tau_{\mathcal{E}(\mathcal{M}_0)}/\theta^{(1)}_\varepsilon]$. As G is bounded and $\varepsilon \theta^{(1)}_\varepsilon \to \infty$, the expectation of this piece is bounded by $R_\varepsilon(\boldsymbol{x})$. By the strong Markov property, the second piece is equal to

$$\mathbb{E}^\varepsilon_{\boldsymbol{x}} \left[\mathbb{E}^\varepsilon_{\boldsymbol{x}_\varepsilon(\tau_{\mathcal{E}(\mathcal{M}_0)})} \left[\int_0^\infty e^{-\lambda s} G(\boldsymbol{x}_\varepsilon(\theta^{(1)}_\varepsilon s)) \, \mathrm{d}s \right] e^{-\lambda \tau_{\mathcal{E}(\mathcal{M}_0)}/\theta^{(1)}_\varepsilon} \mathbf{1}\left\{ \tau_{\mathcal{E}(\mathcal{M}_0)} \leq \frac{C_0}{\varepsilon} \right\} \right].$$

By the same reasons invoked above, this expression is equal to

$$\mathbb{E}^\varepsilon_{\boldsymbol{x}} \left[\mathbb{E}^\varepsilon_{\boldsymbol{x}_\varepsilon(\tau_{\mathcal{E}(\mathcal{M}_0)})} \left[\int_0^\infty e^{-\lambda s} G(\boldsymbol{x}_\varepsilon(\theta^{(1)}_\varepsilon s)) \, \mathrm{d}s \right] \mathbf{1}\left\{ \tau_{\mathcal{E}(\mathcal{M}_0)} \leq \frac{C_0}{\varepsilon} \right\} \right] + R_\varepsilon(\boldsymbol{x}).$$

In conclusion,

$$\phi_\varepsilon(\boldsymbol{x}) = \mathbb{E}^\varepsilon_{\boldsymbol{x}} \left[\phi_\varepsilon(\boldsymbol{x}_\varepsilon(\tau_{\mathcal{E}(\mathcal{M}_0)})) \mathbf{1}\left\{ \tau_{\mathcal{E}(\mathcal{M}_0)} \leq \frac{C_0}{\varepsilon} \right\} \right] + R_\varepsilon(\boldsymbol{x}).$$

Applying Corollary 6.2 once more, as ϕ_ε is uniformly bounded, the right-hand side is equal to

$$\mathbb{E}^\varepsilon_{\boldsymbol{x}} \left[\phi_\varepsilon(\boldsymbol{x}_\varepsilon(\tau_{\mathcal{E}(\mathcal{M}_0)})) \right] + R_\varepsilon(\boldsymbol{x}) = \mathbb{E}^\varepsilon_{\boldsymbol{x}} \left[\phi_\varepsilon(\boldsymbol{x}_\varepsilon(\tau_{\mathcal{E}(\boldsymbol{m})})) \right] + R_\varepsilon(\boldsymbol{x}),$$

where we used hypothesis (8.1) and the uniform boundedness of ϕ_ε in the last step. To complete the proof, it remains to recall the assertion of Theorem 4.1. $\qquad \square$
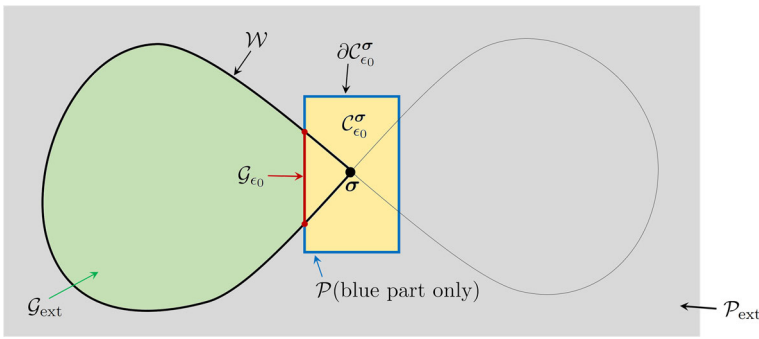
Fig. 3. The sets $\mathcal{G}_{\text{ext}}$ and $\mathcal{P}_{\text{ext}}$ introduced in the proof of Proposition 8.1

Recall that $\boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W})$ is fixed and that $\boldsymbol{\sigma} \curvearrowright \boldsymbol{m_\sigma}, \boldsymbol{\sigma} \curvearrowright \boldsymbol{n_\sigma}$. Denote by $B[\boldsymbol{x}, r]$ the closed ball of radius r centered at $\boldsymbol{x}$, and by $\mathcal{W}'$ the connected component of the set $\{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) < U(\boldsymbol{\sigma})\}$ whose closure contains $\boldsymbol{\sigma}$ and $\boldsymbol{n_\sigma}$. Next lemma is a consequence of Theorem 5.2.

Lemma 8.3. *There exist $\delta > 0$ such that $(B[\boldsymbol{\sigma}, \delta] \cap \overline{\mathcal{W}}) \backslash \{\boldsymbol{\sigma}\}$ is contained in the domain of attraction $\mathcal{D}(\boldsymbol{m_\sigma})$ of $\boldsymbol{m_\sigma}$, and $(B[\boldsymbol{\sigma}, \delta] \cap \overline{\mathcal{W}'}) \backslash \{\boldsymbol{\sigma}\}$ is contained in the domain of attraction $\mathcal{D}(\boldsymbol{n_\sigma})$ of $\boldsymbol{n_\sigma}$.*

Proof of Proposition 8.1. We prove the first assertion, as the second is similar. By Lemma 8.3, there exists $\varepsilon_1 > 0$ such that $\partial_+ \mathcal{B}_\varepsilon^{\boldsymbol{\sigma}} \cap \Lambda_{c,\varepsilon} \subset \mathcal{D}(\boldsymbol{m_\sigma})$ for all $\varepsilon < \varepsilon_1$. Therefore, by Proposition 8.2, it suffices to show that

$$\liminf_{\varepsilon \to 0} \inf_{\boldsymbol{x} \in \partial_+ \mathcal{B}_\varepsilon^{\boldsymbol{\sigma}} \cap \Lambda_{c,\varepsilon}} \mathbb{P}_{\boldsymbol{x}}^\varepsilon [\, \tau_{\mathcal{E}(\boldsymbol{m_\sigma})} = \tau_{\mathcal{E}(\mathcal{M}_0)} \,] = 1. \tag{8.4}$$

Recall that $\mathcal{W}$ represents the well that contains $\boldsymbol{m_\sigma}$. Let $\mathcal{G}_\varepsilon = \partial_+ \mathcal{B}_\varepsilon^{\boldsymbol{\sigma}} \cap \overline{\mathcal{W}}$. By Lemma 8.3, there exists $\varepsilon_0 > 0$ such that $\mathcal{G}_\varepsilon \subset \mathcal{D}(\boldsymbol{m_\sigma})$ for all $\varepsilon \leq \varepsilon_0$. We claim that

$$\inf_{\boldsymbol{x} \in \partial_+ \mathcal{B}_\varepsilon^{\boldsymbol{\sigma}} \cap \Lambda_{c,\varepsilon}} \mathbb{P}_{\boldsymbol{x}}^\varepsilon [\tau_{\partial \mathcal{C}_{\varepsilon_0}^{\boldsymbol{\sigma}}} = \tau_{\mathcal{G}_{\varepsilon_0}}] = 1 - o_\varepsilon(1) \tag{8.5}$$

where $\partial \mathcal{C}_\varepsilon^{\boldsymbol{\sigma}}$ represents the boundary of $\mathcal{C}_\varepsilon^{\boldsymbol{\sigma}}$.

To prove (8.5), let $\mathcal{P} := \partial \mathcal{C}_{\varepsilon_0}^{\boldsymbol{\sigma}} \backslash \mathcal{G}_{\varepsilon_0}$, $\mathcal{G}_{\text{ext}} = \overline{\mathcal{W} \backslash \mathcal{C}_{\varepsilon_0}^{\boldsymbol{\sigma}}}$, $\mathcal{P}_{\text{ext}} = \overline{\mathbb{R}^d \backslash [\mathcal{G}_{\text{ext}} \cup \mathcal{C}_{\varepsilon_0}^{\boldsymbol{\sigma}}]}$. Figure 3 illustrates these sets. By definition,

$$\{\, \tau_{\partial \mathcal{C}_{\varepsilon_0}^{\boldsymbol{\sigma}}} = \tau_{\mathcal{G}_{\varepsilon_0}} \,\} = \{\, \tau_{\mathcal{G}_{\varepsilon_0}} < \tau_{\mathcal{P}} \,\} = \{\, \tau_{\mathcal{G}_{\text{ext}}} < \tau_{\mathcal{P}_{\text{ext}}} \,\}$$

for all $\boldsymbol{x} \in \partial_+ \mathcal{B}_\varepsilon^{\boldsymbol{\sigma}} \cap \Lambda_{c,\varepsilon}, \varepsilon < \varepsilon_0$. Therefore, by definition of the set $\mathcal{P}_{\text{ext}}$, $\Big\{ \tau_{\mathcal{E}(\boldsymbol{m_\sigma})} < \tau_{\mathcal{P}_{\text{ext}}} \Big\} \subset \{\tau_{\partial \mathcal{C}_{\varepsilon_0}^{\boldsymbol{\sigma}}} = \tau_{\mathcal{G}_{\varepsilon_0}}\}$.

Fix $\varepsilon < \varepsilon_0$. Recall that we denote by $B(\boldsymbol{x}, r)$ the open ball of radius r centered at $\boldsymbol{x}$. By [28, Lemma 9.2], there exists a finite constant C_0, whose value may change from line to line, such that

$$\text{cap}_\varepsilon(B(\boldsymbol{y}, \varepsilon), \mathcal{E}(\boldsymbol{m_\sigma})) \geqq C_0 \, \varepsilon^d \, Z_\varepsilon^{-1} \, e^{-U(\boldsymbol{y})/\varepsilon}$$

for all $y \in \partial_+ \mathcal{B}_\varepsilon^\sigma \cap \Lambda_{c,\varepsilon}$. In this formula, $\mathrm{cap}_\varepsilon(\mathcal{A}, \mathcal{B})$ stands for the capacity between the sets $\mathcal{A}$, $\mathcal{B}$ for the diffusion $x_\varepsilon(\cdot)$ and is defined in Appendix Appendix C. On the other hand, by the proof of [28, Lemma 9.3], there exists a finite constant C_0 such that

$$\mathrm{cap}_\varepsilon(B(y, \varepsilon), \mathcal{P}_{\mathrm{ext}}) \leq C_0 \, Z_\varepsilon^{-1} \, e^{-U(\sigma)/\varepsilon}$$

for all $y \in \partial_+ \mathcal{B}_\varepsilon^\sigma \cap \Lambda_{c,\varepsilon}$. By [23, Proposition 7.2], there exists a finite constant C_0 such that

$$\mathbb{P}_x^\varepsilon[\,\tau_{\mathcal{P}_{\mathrm{ext}}} < \tau_{\mathcal{E}(m_\sigma)}\,] \;\leq\; C_0 \, \frac{\mathrm{cap}_\varepsilon(B(x, \varepsilon), \, \mathcal{P}_{\mathrm{ext}})}{\mathrm{cap}_\varepsilon(B(x, \varepsilon), \, \mathcal{E}(m_\sigma))}$$

for all $x \in \mathbb{R}^d$. Combining the previous estimates yields that for all $x \in \partial_+ \mathcal{B}_\varepsilon^\sigma \cap \Lambda_{c,\varepsilon}$

$$\mathbb{P}_x^\varepsilon[\,\tau_{\mathcal{P}_{\mathrm{ext}}} < \tau_{\mathcal{E}(m_\sigma)}\,] \;\leq\; C_0 \, \varepsilon^{-d} \, e^{(U(x)-U(\sigma))/\varepsilon} \;\leq\; C_0 \, \varepsilon^{-d} \, e^{-cJ^2\delta^2/\varepsilon} \;=\; C_0 \, \varepsilon^{cJ^2-d}.$$

This expression vanishes as $\varepsilon \to 0$ for large enough J. To complete the proof of assertion (8.5), it remains to recall that $\{\tau_{\mathcal{E}(m_\sigma)} < \tau_{\mathcal{P}_{\mathrm{ext}}}\} \subset \{\tau_{\partial \mathcal{C}_{\varepsilon_0}^\sigma} = \tau_{\mathcal{G}_{\varepsilon_0}}\}$.

We turn to the proof of (8.4). For $x \in \partial_+ \mathcal{B}_\varepsilon^\sigma \cap \Lambda_{c,\varepsilon}$, by the strong Markov property and (8.5),

$$\mathbb{P}_x^\varepsilon[\tau_{\mathcal{E}(m_\sigma)} = \tau_{\mathcal{E}(\mathcal{M}_0)}] \geq \mathbb{P}_x^\varepsilon[\tau_{\mathcal{E}(m_\sigma)} = \tau_{\mathcal{E}(\mathcal{M}_0)}, \, \tau_{\partial \mathcal{C}_{\varepsilon_0}^\sigma} = \tau_{\mathcal{G}_{\varepsilon_0}}]$$

$$\geq \inf_{y \in \mathcal{G}_{\varepsilon_0}} \mathbb{P}_y^\varepsilon[\tau_{\mathcal{E}(m_\sigma)} = \tau_{\mathcal{E}(\mathcal{M}_0)}] \, \mathbb{P}_x^\varepsilon[\tau_{\partial \mathcal{C}_{\varepsilon_0}^\sigma} = \tau_{\mathcal{G}_{\varepsilon_0}}]$$

$$= (1 - o_\varepsilon(1)) \inf_{y \in \mathcal{G}_{\varepsilon_0}} \mathbb{P}_y^\varepsilon[\tau_{\mathcal{E}(m_o)} = \tau_{\mathcal{E}(\mathcal{M}_0)}] \, .$$

The last infimum is $1 - o_\varepsilon(1)$ by Lemma 6.7 because $\mathcal{G}_{\varepsilon_0} \subset \mathcal{D}(m_0)$. $\qquad\square$

Proposition 8.1 provides a simple formula for the quantities $J_\pm(\sigma)$ introduced in Lemma 7.1.

Lemma 8.4. *For all* $\mathbf{g} \colon \mathcal{M}_0 \to \mathbb{R}$,

$$e^{H/\varepsilon} \int_\Omega Q_\varepsilon \, (-\mathcal{L}_\varepsilon \, \phi_\varepsilon) \, d\mu_\varepsilon \;=\; \frac{1}{2\pi \nu_\star} \sum_{\sigma \in \mathcal{S}_H(\mathcal{W})} [\phi_\varepsilon(m_\sigma) - \phi_\varepsilon(n_\sigma)]$$

$$\frac{\mu^\sigma}{\sqrt{-\det \mathbb{H}^\sigma}} + o_\varepsilon(1). \tag{8.6}$$

Proof. By Proposition 8.1, in the formula for $J_+(\sigma)$ presented in Lemma 7.1, we may replace $\phi_\varepsilon(x)$ by $\phi_\varepsilon(m_\sigma)$ at a cost $o_\varepsilon(1)$, and we are left with a Gaussian type integral. A straightforward computation, presented in the proof of [29, Lemma 7.5], together with [29, Lemma 7.3] yields that

$$\frac{\varepsilon \, \sqrt{\mu^\sigma} \, (v \cdot e_1)}{(2\pi \varepsilon)^{(d+1)/2} \, \nu_\star} \int_{\partial_+ \mathcal{B}_\varepsilon^\sigma \cap \Lambda_{c,\varepsilon}} e^{-\frac{1}{2\varepsilon} x \cdot (\mathbb{H}^\sigma + \mu^\sigma \, v \otimes v)} \, \mathrm{S}(dx)$$

$$= \frac{1}{2\pi \, \nu_\star} \frac{\lambda_1^\sigma}{\sqrt{-\det \mathbb{H}^\sigma}} + o_\varepsilon(1).$$

Similarly, by the proof of [29, Lemma 7.7],

$$\frac{1}{(2\pi)^{(d+1)/2} \, v_\star \, \sqrt{\mu^\sigma} \, \varepsilon^{(d+3)/2}} \int_{\mathcal{A}_\varepsilon^{\sigma,+} \cap \Lambda_{c,\varepsilon}} \frac{\mathbb{L}^\sigma \overline{x} \cdot e_1}{\overline{x} \cdot v} \, e^{-\frac{1}{2\varepsilon} \overline{x} \cdot (\mathbb{H}^\sigma + \mu^\sigma \, v \otimes v) \, \overline{x}} \, dx$$

$$= \frac{1}{2\pi v_\star} \frac{\lambda_1^\sigma}{\sqrt{-\det \mathbb{H}^\sigma}} \frac{(\mathbb{L}^\sigma (\mathbb{H}^\sigma)^{-1} v) \cdot e_1}{v \cdot e_1} + o_\varepsilon(1).$$

By the proof of [29, Proposition 5.7],

$$v \cdot e_1 + (\mathbb{L}^\sigma (\mathbb{H}^\sigma)^{-1} v) \cdot e_1 = \frac{\mu^\sigma}{\lambda_1^\sigma} v \cdot e_1.$$

Combining the previous estimates yields that

$$J_+(\sigma) = -\frac{1}{2\pi \, v_\star} \frac{\mu_1^\sigma}{\sqrt{-\det \mathbb{H}^\sigma}} \phi_\varepsilon(m_\sigma) + o_\varepsilon(1).$$

The same argument leads to the same formula for $J_-(\sigma)$ with a plus sign and $\phi_\varepsilon(n_\sigma)$ instead of $\phi_\varepsilon(m_\sigma)$. This completes the proof of the lemma. □

9. Proof of Theorem 2.2

Recall from (2.11), (4.7) the definitions of the generator $\mathfrak{L}_1$, and the function $f_\varepsilon \colon \mathcal{M}_0 \to \mathbb{R}$, respectively. The main result of this section is

Theorem 9.1. *For all $\lambda > 0$, $g \colon \mathcal{M}_0 \to \mathbb{R}$,*

$$(\lambda - \mathfrak{L}_1) f_\varepsilon = g + o_\varepsilon(1).$$

Proof of Theorem 2.2. The assertion follows from two observations. The sequence f_ε is uniformly bounded and the equation $(\lambda - \mathfrak{L}_1) f = g$ has a unique solution. □

The remainder of this section is devoted to the proof of Theorem 9.1. Fix $m \in \mathcal{M}_0$. Let $\mathcal{W}$ be the connected component of the set $\{x \in \mathbb{R}^d : U(x) < U(m) + \Gamma(m)\}$ which contains m. By definition, $\mathcal{W}$ does not contain any other local minimum of U (in particular, the present situation is different from the one represented in Fig. 2, where $\mathcal{W}$ contains more than one local minimum). Recall from (7.5) the definition of the set $\mathcal{S}_H(\mathcal{W})$.

Lemma 9.2. *There exists a saddle point $\sigma \in \partial \mathcal{W}$ satisfying condition (a) in (7.5). Condition (b) is fulfilled for all saddle points $\sigma' \in \partial \mathcal{W}$ satisfying (a). Moreover, $\mathcal{S}_H(\mathcal{W}) = \Upsilon(m)$, where $H = U(m) + \Gamma(m)$.*

Proof. By Proposition A.1, there exist a local minimum m' of U different from m and a continuous path $z \colon [0, 1] \to \mathbb{R}^d$ such that $U(m) + \Gamma(m) = \Theta(m, m')$, $z(0) = m$, $z(1) = m'$, and

$$\max_{t \in [0, 1]} U(z(t)) = U(z(1/2)) = \Theta(m, m'), \quad U(z(s))$$

$$< U(z(1/2)), \quad s \in [0, 1] \setminus \{1/2\}, \tag{9.1}$$

and $\boldsymbol{\sigma} := z(1/2)$ is a saddle point of U. In particular, $\boldsymbol{\sigma} \in \partial \mathcal{W}$. Condition (a) is satisfied because $\boldsymbol{m}' \neq \boldsymbol{m}$ and $\mathcal{W}$ contains only the local minimum $\boldsymbol{m}$.

We turn to condition (b). Let $\boldsymbol{\sigma}$ be a saddle point in $\partial \mathcal{W}$ satisfying (a). By definition of $\mathcal{W}$ and with the help of the solution of the ODE (2.4), it is possible to construct a continuous path $z \colon [0, 1] \rightarrow \mathbb{R}^d$ such that $z(0) = \boldsymbol{m}' \in \mathcal{M}_0$, $z(1/2) = \boldsymbol{\sigma}, z(1) = \boldsymbol{m}'' \in \mathcal{M}_0$, and

$$U(z(s)) < U(\boldsymbol{\sigma}), \quad s \neq 1/2,$$

for some $\boldsymbol{m}', \boldsymbol{m}'' \in \mathcal{M}_0$. As $\boldsymbol{\sigma}$ satisfies (a), we may assume without loss of generality that $\boldsymbol{m}' \in \mathcal{W}$, $\boldsymbol{m}'' \in \overline{\mathcal{W}}^c$. Since $\mathcal{W}$ contains a unique local minimum, $\boldsymbol{m}' = \boldsymbol{m}$. Therefore, since $U(\boldsymbol{\sigma}) = U(\boldsymbol{m}) + \Gamma(\boldsymbol{m})$, by definition of $\Upsilon(\boldsymbol{m})$, $\boldsymbol{\sigma} \in \Upsilon(\boldsymbol{m})$. Hence, by condition (2.5), there exists $\boldsymbol{m}''' \in \mathcal{M}_0, \boldsymbol{m}''' \notin \mathcal{W}$, such that $\boldsymbol{\sigma} \frown \boldsymbol{m}, \boldsymbol{\sigma} \frown \boldsymbol{m}'''$, which is condition (b).

Assume that $\boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W})$. By definition, it satisfies (a). Thus, by the previous paragraph, $\boldsymbol{\sigma} \in \Upsilon(\boldsymbol{m})$. Conversely, suppose that $\boldsymbol{\sigma} \in \Upsilon(\boldsymbol{m})$. By definition, there exists a local minimum $\boldsymbol{m}' \neq \boldsymbol{m}$ and a continuous path $z \colon [0, 1] \rightarrow \mathbb{R}^d$ such that $z(0) = \boldsymbol{m}, z(1) = \boldsymbol{m}'$ for which (9.1) holds. By Proposition A.1, $\boldsymbol{\sigma} := z(1/2)$ is a saddle point of U. Since $\mathcal{W}$ has a unique local minimum, $\boldsymbol{m}' \notin \mathcal{W}$. Thus, condition (a) holds for $\boldsymbol{\sigma}$. By (2.5), condition (b) also holds, so that $\boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W})$. $\qquad \square$

Proof of Theorem 9.1. Fix $\boldsymbol{m} \in \mathcal{M}_0$. Let $\mathcal{W}$ be the connected component of the set $\{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) < U(\boldsymbol{m}) + \Gamma(\boldsymbol{m})\}$ which contains $\boldsymbol{m}$. By Lemma 9.2, there exists a saddle point $\boldsymbol{\sigma} \in \partial \mathcal{W}$ satisfying condition (a) in (7.5), and condition (b) is fulfilled for all saddle points $\boldsymbol{\sigma}' \in \partial \mathcal{W}$ satisfying (a). We may therefore apply Lemma 7.1.

Let Q_ε be the test function constructed in Sect. 7 associated to the well $\mathcal{W}$, and recall that $H = U(\boldsymbol{m}) + \Gamma(\boldsymbol{m})$. Multiply both sides of (2.13) by the test function Q_ε and integrate over $\mathbb{R}^d$ to deduce that

$$\int_\Omega Q_\varepsilon \, (\lambda - \theta_\varepsilon^{(1)} \, \mathcal{L}_\varepsilon) \, \phi_\varepsilon \, d\mu_\varepsilon \; = \; \boldsymbol{g}(\boldsymbol{m}) \int_{\mathcal{E}(\boldsymbol{m})} Q_\varepsilon \, d\mu_\varepsilon, \tag{9.2}$$

where Ω is given by (7.7).

By definition of Q_ε, the right-hand side is equal to $\boldsymbol{g}(\boldsymbol{m}) \, \mu_\varepsilon(\mathcal{E}(\boldsymbol{m}))$. Similarly, as ϕ_ε is uniformly bounded and

$$\mu_\varepsilon \Big(\bigcup_{\boldsymbol{\sigma} \in \mathcal{S}_H(\mathcal{W})} \mathcal{E}_\varepsilon^{\boldsymbol{\sigma}} \, \cup \, (\Omega \setminus \mathcal{K}_\varepsilon) \Big) \; = \; o_\varepsilon(1) \, \mu_\varepsilon(\mathcal{E}(\boldsymbol{m})),$$

as Q_ε vanishes on $\mathcal{W}_2$ and is equal to 1 on $\mathcal{W}_1$, by definition of $\mathbf{f}_\varepsilon$,

$$\lambda \int_\Omega Q_\varepsilon \, \phi_\varepsilon \, d\mu_\varepsilon \; = \; \lambda \, \mathbf{f}_\varepsilon(\boldsymbol{m}) \, \mu_\varepsilon(\mathcal{E}(\boldsymbol{m})) \; + \; o_\varepsilon(1) \, \mu_\varepsilon(\mathcal{E}(\boldsymbol{m})) \,. \tag{9.3}$$

It remains to consider the term in (9.2) involving the generator $\mathcal{L}_\varepsilon$. We examine two cases separately.

Case 1: Assume that $\Gamma(\boldsymbol{m}) > d^{(1)}$. As $\Gamma(\boldsymbol{m}) > d^{(1)}$ and $e^{-U(\boldsymbol{m})/\varepsilon}/\mu_\varepsilon(\mathcal{E}(\boldsymbol{m}))$ $\leq C_0$ for some finite constant independent of ε, $\theta_\varepsilon^{(1)} = e^{d^{(1)}/\varepsilon} \prec e^{\Gamma(\boldsymbol{m})/\varepsilon}$ $\leq C_0\, e^{[\Gamma(\boldsymbol{m})+U(\boldsymbol{m})]/\varepsilon}\, \mu_\varepsilon(\mathcal{E}(\boldsymbol{m}))$. Hence, by Lemma 8.4, as the right-hand side of (8.6) is bounded and $H = \Gamma(\boldsymbol{m}) + U(\boldsymbol{m})$,

$$\theta_\varepsilon^{(1)} \int_\Omega Q_\varepsilon\,(-\mathcal{L}_\varepsilon)\,\phi_\varepsilon\,\mathrm{d}\mu_\varepsilon \;=\; o_\varepsilon(1)\,\mu_\varepsilon(\mathcal{E}(\boldsymbol{m})).$$

Combining the previous estimates yields that

$$\mathbf{f}_\varepsilon(\boldsymbol{m}) = \frac{1}{\lambda}\,\mathbf{g}(\boldsymbol{m}) \;+\; o_\varepsilon(1).$$

By (2.10) and the definition of $\mathfrak{L}_1$, as $\Gamma(\boldsymbol{m}) > d^{(1)}$, $(\mathfrak{L}_1 \boldsymbol{f}_\varepsilon)(\boldsymbol{m}) = 0$, which completes the proof of the theorem in Case 1.

Case 2: Assume that $\Gamma(\boldsymbol{m}) = d^{(1)}$. Multiply both sides of (8.6) by $e^{-U(\boldsymbol{m})/\varepsilon}$. Since $\theta_\varepsilon^{(1)} = e^{d^{(1)}/\varepsilon} = e^{\Gamma(\boldsymbol{m})/\varepsilon} = e^{[H-U(\boldsymbol{m})]/\varepsilon}$, by Lemma 8.4,

$$\theta_\varepsilon^{(1)} \int_\Omega Q_\varepsilon\,(-\mathcal{L}_\varepsilon)\,\phi_\varepsilon\,d\mu_\varepsilon$$
$$= \left\{ \frac{1}{2\pi\,\nu_\star} \sum_{\sigma \in \mathcal{S}_H(\mathcal{W})} [\phi_\varepsilon(\boldsymbol{m}) - \phi_\varepsilon(\boldsymbol{n}_\sigma)]\, \frac{\mu^\sigma}{\sqrt{-\det \mathbb{H}^\sigma}} \;+\; o_\varepsilon(1) \right\} e^{-U(\boldsymbol{m})/\varepsilon}$$

because $\boldsymbol{m}_\sigma = \boldsymbol{m}$ for all $\sigma \in \mathcal{S}_H(\mathcal{W})$, as $\boldsymbol{m}$ is the only local minima of U in $\mathcal{W}$.

Since $e^{-U(\boldsymbol{m})/\varepsilon}/\mu_\varepsilon(\mathcal{E}(\boldsymbol{m})) \leq C_0$ for some finite constant independent of ε, we may replace in the previous formula, $o_\varepsilon(1)\, e^{-U(\boldsymbol{m})/\varepsilon}$ by $o_\varepsilon(1)\, \mu_\varepsilon(\mathcal{E}(\boldsymbol{m}))$. On the other hand, by (4.3), $e^{-U(\boldsymbol{m})/\varepsilon}/\nu_\star = [1 + o_\varepsilon(1)]\, \mu_\varepsilon(\mathcal{E}(\boldsymbol{m}))/\nu(\boldsymbol{m})$. We may therefore rewrite the right-hand side of the previous equation as

$$\left\{ \frac{1}{2\pi\,\nu(\boldsymbol{m})} \sum_{\sigma \in \mathcal{S}_H(\mathcal{W})} [\phi_\varepsilon(\boldsymbol{m}) - \phi_\varepsilon(\boldsymbol{n}_\sigma)]\, \frac{\mu^\sigma}{\sqrt{-\det \mathbb{H}^\sigma}} \;+\; o_\varepsilon(1) \right\} \mu_\varepsilon(\mathcal{E}(\boldsymbol{m})).$$

By Lemma 9.2, $\mathcal{S}_H(\mathcal{W}) = \Upsilon(\boldsymbol{m})$. Thus, by (2.7) and by definition of $\boldsymbol{n}_\sigma$, introduced at the beginning of Sect. 8, $\{\boldsymbol{n}_\sigma : \sigma \in \mathcal{S}_H(\mathcal{W})\} = \{\boldsymbol{n}_\sigma : \sigma \in \Upsilon(\boldsymbol{m})\} = \mathcal{V}(\boldsymbol{m})$. Hence, by Theorem 4.1, the previous expression can rewritten as

$$\left\{ \frac{1}{2\pi\,\nu(\boldsymbol{m})} \sum_{\boldsymbol{m}' \in \mathcal{V}(\boldsymbol{m})} [\boldsymbol{f}_\varepsilon(\boldsymbol{m}) - \boldsymbol{f}_\varepsilon(\boldsymbol{m}')] \sum_{\sigma \in \mathcal{S}(\boldsymbol{m},\boldsymbol{m}')} \frac{\mu^\sigma}{\sqrt{-\det \mathbb{H}^\sigma}} \;+\; o_\varepsilon(1) \right\} \mu_\varepsilon(\mathcal{E}(\boldsymbol{m})).$$

By (2.8), (2.9), (2.10) and (2.11), the previous expression is equal to

$$\left\{ (-\mathfrak{L}_1 \boldsymbol{f}_\varepsilon)(\boldsymbol{m}) \;+\; o_\varepsilon(1) \right\} \mu_\varepsilon(\mathcal{E}(\boldsymbol{m})).$$

To complete the proof of the theorem, it remains to combine the estimates obtained at the beginning of the proof with this last one. $\qquad\square$

Trace Processes

Let $\boldsymbol{y}_\varepsilon(t)$ be the process $\boldsymbol{x}_\varepsilon(t)$ speeded-up by $\theta_\varepsilon^{(1)}$: $\boldsymbol{y}_\varepsilon(t) = \boldsymbol{x}_\varepsilon(t\theta_\varepsilon^{(1)})$. Denote by $\mathbb{Q}_{\boldsymbol{x}}^\varepsilon$ the probability measure on $C(\mathbb{R}_+, \mathbb{R}^d)$ induced by the process $\boldsymbol{y}_\varepsilon(t)$ starting from $\boldsymbol{x}$. We use the same symbol $\mathbb{Q}_{\boldsymbol{x}}^\varepsilon$ to represent the expectation with respect to the measure $\mathbb{Q}_{\boldsymbol{x}}^\varepsilon$.

Denote by $T_\varepsilon(t)$ the time spent by $\boldsymbol{y}_\varepsilon(\cdot)$ on $\mathcal{E}(\mathcal{M}_0)$ up to time $t > 0$:

$$T_\varepsilon(t) \ := \ \int_0^t \chi_{\mathcal{E}(\mathcal{M}_0)}(\boldsymbol{y}_\varepsilon(s))\, \mathrm{d}s.$$

Let $S_\varepsilon(\cdot)$ be the generalized inverse of the non-decreasing process $T_\varepsilon(\cdot)$:

$$S_\varepsilon(t) \ := \ \sup\{s \geq 0 : T_\varepsilon(s) \leq t\}, \quad t \geq 0.$$

Define the trace process of $\boldsymbol{y}_\varepsilon(\cdot)$ on $\mathcal{E}(\mathcal{M}_0)$ by

$$\boldsymbol{y}_\varepsilon^{\mathrm{T}}(t) \ := \ \boldsymbol{y}_\varepsilon(S_\varepsilon(t)), \quad t \geq 0, \tag{9.4}$$

which is an $\mathcal{E}(\mathcal{M}_0)$-valued Markov process. Let $\Phi : \mathcal{E}(\mathcal{M}_0) \to \mathcal{M}_0$ be the projection given by $\Phi = \sum_{\boldsymbol{m}\in\mathcal{M}_0} \boldsymbol{m}\, \chi_{\mathcal{E}(\boldsymbol{m})}$. The next result is a consequence of Theorem 2.2 and [22, Theorem 2.3].

Theorem 9.3. *Fix $\boldsymbol{m} \in \mathcal{M}_0$, and a sequence $\boldsymbol{x}_\varepsilon \in \mathcal{E}(\boldsymbol{m})$. Starting from $\boldsymbol{x}_\varepsilon$, the process $\Phi(\boldsymbol{y}_\varepsilon^{\mathrm{T}}(t))$ converges in the Skorohod topology to the $\mathcal{M}_0$-valued continuous-time Markov chain induced by the generator $\mathfrak{L}_1$ starting from $\boldsymbol{m}$. Moreover, for all $T > 0$,*

$$\lim_{\varepsilon\to 0} \sup_{\boldsymbol{x}\in\mathcal{E}(\mathcal{M}_0)} \mathbb{Q}_{\boldsymbol{x}}^\varepsilon\left[\int_0^T \chi_{\mathbb{R}^d\setminus\mathcal{E}(\mathcal{M}_0)}(\boldsymbol{y}_\varepsilon(t))\, dt \right] = 0. \tag{9.5}$$

The next assertion is a consequence of (9.5). We refer to [21, display (3.2)] for a proof.

Lemma 9.4. *For all $t \geq 0$ and $\delta > 0$,*

$$\limsup_{\varepsilon\to 0} \sup_{\boldsymbol{x}\in\mathcal{E}(\mathcal{M}_0)} \mathbb{Q}_{\boldsymbol{x}}^\varepsilon[\, S_\varepsilon(t) > t + \delta\,] = 0.$$

10. Proof of Theorem 2.1: Finite-Dimensional Distributions

The main result of this section, Theorem 10.1, states that in the time-scale $\theta_\varepsilon^{(1)}$ the finite-dimensional distributions of the diffusion $\boldsymbol{x}_\varepsilon(t)$ converge to those of the $\mathcal{M}_0$-valued Markov chain whose generator is given by $\mathfrak{L}_1$ introduced in (2.11).

Denote by $D(\mathbb{R}_+, \mathcal{M}_0)$ the space of right-continuous functions $\boldsymbol{y} : \mathbb{R}_+ \to \mathcal{M}_0$ with left limits endowed with the Skorohod topology. Let $\mathcal{Q}_{\boldsymbol{m}}$, $\boldsymbol{m} \in \mathcal{M}_0$, be the measure on $D(\mathbb{R}_+, \mathcal{M}_0)$ induced by the continuous-time $\mathcal{M}_0$-valued Markov chain associated to the generator $\mathfrak{L}_1$ starting from $\boldsymbol{m}$. Expectation with respect to $\mathcal{Q}_{\boldsymbol{m}}$ is also represented by $\mathcal{Q}_{\boldsymbol{m}}$.

Theorem 10.1. *Fix $m \in \mathcal{M}_0$, and $x \in \mathcal{D}(m)$. Then,*

$$\lim_{\varepsilon \to 0} \mathbb{E}_x^\varepsilon \Big[\prod_{j=1}^{\mathfrak{n}} F_j(x_\varepsilon(\theta_\varepsilon^{(1)} t_j)) \Big] = \mathcal{Q}_m \Big[\prod_{j=1}^{\mathfrak{n}} F_j(y(t_j)) \Big]$$

for all $\mathfrak{n} \geq 1$, $0 < t_1 < \cdots < t_\mathfrak{n}$ and bounded continuous functions $F_j \colon \mathbb{R}^d \to \mathbb{R}$, $1 \leq j \leq \mathfrak{n}$.

The proof of Theorem 10.1 is based on the next result.

Proposition 10.2. *Fix $r_0 > 0$ small enough to fulfill the conditions above equation (2.12), and recall from this equation the definition of the wells $\mathcal{E}(m')$, $m' \in \mathcal{M}_0$. Fix $m \in \mathcal{M}_0$. Then,*

$$\lim_{\varepsilon \to 0} \mathbb{P}_{x_\varepsilon}^\varepsilon \Big[\bigcap_{j=1}^{\mathfrak{n}} \{x_\varepsilon(\theta_\varepsilon^{(1)} t_{j,\varepsilon})) \in \mathcal{E}(m_j)\} \Big] = \mathcal{Q}_m \Big[\bigcap_{j=1}^{\mathfrak{n}} \{y(t_j)) = m_j\} \Big]$$

for all $\mathfrak{n} \geq 1$, $0 < t_1 < \cdots < t_\mathfrak{n}$, $m_1, \ldots, m_\mathfrak{n} \in \mathcal{M}_0$, and sequences $x_\varepsilon \in \mathcal{E}(m)$, $t_{j,\varepsilon} \to t_j$.

It follows from this result that

$$\lim_{\varepsilon \to 0} \mathbb{P}_{x_\varepsilon}^\varepsilon \Big[\bigcap_{j=1}^{\mathfrak{n}} \{x_\varepsilon(\theta_\varepsilon^{(1)} t_{j,\varepsilon})) \in \mathcal{E}(\mathcal{M}_0)\} \Big] = 1 \tag{10.1}$$

for all $m \in \mathcal{M}_0$, $\mathfrak{n} \geq 1$, $0 < t_1 < \cdots < t_\mathfrak{n}$, $m_1, \ldots, m_\mathfrak{n} \in \mathcal{M}_0$, and sequences $x_\varepsilon \in \mathcal{E}(m)$, $t_{j,\varepsilon} \to t_j$.

Proof of Theorem 10.1

We prove the result for $\mathfrak{n} = 1$ as the arguments in the general case are identical. Fix $t > 0$, $\eta > 0$ and a bounded continuous function $F \colon \mathbb{R}^d \to \mathbb{R}$. By continuity, there exists $\delta_0 > 0$ such that

$$\max_{m \in \mathcal{M}_0} \sup_{x \in \mathcal{W}^{2\delta_0}(m)} |F(x) - F(m)| \leq \eta. \tag{10.2}$$

Fix $r_0 < \delta_0$ small enough to fulfill the conditions of Proposition 10.2. Consider the wells $\mathcal{E}(m)$ defined by (2.12).

Recall from (4.4) that we represent by $\tau_{\mathcal{A}}$ the hitting time of the set $\mathcal{A}$, and let $\tau = \tau_{\mathcal{E}(m)}$. By Lemma 6.7, Corollary 6.2, and the strong Markov property, as $x \in \mathcal{D}(m)$ and F is bounded,

$$\mathbb{E}_x^\varepsilon \big[F(x_\varepsilon(\theta_\varepsilon^{(1)} t)) \big] = \mathbb{E}_x^\varepsilon \Big[\mathbb{E}_{x_\varepsilon(\tau)}^\varepsilon \big[F(x_\varepsilon(\theta_\varepsilon^{(1)} t - \tau)) \big] \chi_{\tau \leq \varepsilon^{-1}} \Big] + R_\varepsilon^{(1)},$$

where $|R_\varepsilon^{(1)}| \to 0$. The expectation on the right-hand side has to be understood as the expectation of $\mathbb{E}_{x_\varepsilon(\tau)}^\varepsilon \big[F(x_\varepsilon(\theta_\varepsilon^{(1)} t - s)) \big]$ for $s = \tau$. By definition of r_0, the

wells $\mathcal{E}(m')$, $m' \in \mathcal{M}_0$, and (10.2), the right-hand side of the previous equation is equal to

$$\sum_{m' \in \mathcal{M}_0} F(m') \, \mathbb{E}_x^\varepsilon \Big[\mathbb{P}_{x_\varepsilon(\tau)}^\varepsilon \big[x_\varepsilon(\theta_\varepsilon^{(1)} t - \tau) \in \mathcal{E}(m') \big] \chi_{\tau \le \varepsilon^{-1}} \Big] + R_\varepsilon^{(1)} + R_\varepsilon^{(2)} + R_\eta,$$

where $|R_\eta| \le \eta$ and

$$|R_\varepsilon^{(2)}| \le \|F\|_\infty \sup_{y \in \mathcal{E}(m)} \sup_{t - (\varepsilon \theta_\varepsilon^{(1)})^{-1} \le s \le t} \mathbb{P}_y^\varepsilon \big[x_\varepsilon(\theta_\varepsilon^{(1)} s) \notin \mathcal{E}(\mathcal{M}_0) \big].$$

By (10.1), $R_\varepsilon^{(2)} \to 0$. By Proposition 10.2, Lemma 6.7 and Corollary 6.2, as $\varepsilon \to 0$, the sum converges to

$$\sum_{m' \in \mathcal{M}_0} F(m') \, \mathcal{Q}_m \big[y(t) = m' \big] = \mathcal{Q}_m \big[F(y(t)) \big],$$

which completes the proof of the theorem. □

Proof of Proposition 10.2

The proof relies on a lemma, which appeared before in [21, Lemma 3.1] for discrete-valued Markov processes.

Lemma 10.3. *Fix $t > 0$ and m, $m' \in \mathcal{M}_0$. Then, for all $x \in \mathcal{E}(m)$, $b \in (0, t/3)$ and sequence $t_\varepsilon \to t$,*

$$\mathbb{Q}_x^\varepsilon \big[y_\varepsilon^{\mathrm{T}}(t - 3b) \in \mathcal{E}(m') \big] \le \mathbb{Q}_x^\varepsilon \big[y_\varepsilon(t_\varepsilon) \in \mathcal{E}(m') \big] + R_\varepsilon(x, t, b),$$

where,

$$\lim_{b \to 0} \limsup_{\varepsilon \to 0} \sup_{x \in \mathcal{E}(m)} R_\varepsilon(x, t, b) = 0.$$

Proof. Fix $t > 0$, m, $m' \in \mathcal{M}_0$, $x \in \mathcal{E}(m)$, a sequence $t_\varepsilon \to t$, and $2 < \alpha < 3$. By (9.4) and the trivial fact that $S_\varepsilon(t) \ge t$, for $b \in (0, t/3)$

$$\mathbb{Q}_x^\varepsilon [y_\varepsilon^{\mathrm{T}}(t - 3b) \in \mathcal{E}(m')] = \mathbb{Q}_x^\varepsilon [y_\varepsilon(S_\varepsilon(t - 3b)) \in \mathcal{E}(m')]$$
$$\le \mathbb{Q}_x^\varepsilon [A_\varepsilon(t, b)] + \mathbb{Q}_x^\varepsilon [B_\varepsilon(t, b)],$$

where

$$A_\varepsilon(t, b) = \{ S_\varepsilon(t - 3b) > t - \alpha b \},$$
$$B_\varepsilon(t, b) = \{ y_\varepsilon(s) \in \mathcal{E}(m') \text{ for some } s \in [t - 3b, t - \alpha b] \}.$$

By Lemma 9.4, as $\alpha < 3$,

$$\limsup_{\varepsilon \to 0} \sup_{x \in \mathcal{E}(\mathcal{M}_0)} \mathbb{Q}_x^\varepsilon [A_\varepsilon(t, b)] = 0.$$

On the other hand,

$$\mathbb{Q}_{\boldsymbol{x}}^{\varepsilon}[B_{\varepsilon}(t,\,b)] \leq \mathbb{Q}_{\boldsymbol{x}}^{\varepsilon}[\,\boldsymbol{y}_{\varepsilon}(t_{\varepsilon}) \in \mathcal{E}(\boldsymbol{m}')\,] + \mathbb{Q}_{\boldsymbol{x}}^{\varepsilon}[\,B_{\varepsilon}(t,\,b),\;\boldsymbol{y}_{\varepsilon}(t_{\varepsilon}) \notin \mathcal{E}(\boldsymbol{m}')\,].$$

It remains to prove that

$$\limsup_{b \to 0}\;\limsup_{\varepsilon \to 0}\;\sup_{\boldsymbol{x} \in \mathcal{E}(\boldsymbol{m})}\;\mathbb{Q}_{\boldsymbol{x}}^{\varepsilon}[\,B_{\varepsilon}(t,\,b),\;\boldsymbol{y}_{\varepsilon}(t_{\varepsilon}) \notin \mathcal{E}(\boldsymbol{m}')\,] \;=\; 0.$$

By Lemma 6.6, the definition of $B_{\varepsilon}(t,\,b)$, and the strong Markov property,

$$\limsup_{b \to 0}\;\limsup_{\varepsilon \to 0}\;\sup_{\boldsymbol{x} \in \mathcal{E}(\boldsymbol{m})}\;\mathbb{Q}_{\boldsymbol{x}}^{\varepsilon}[\,B_{\varepsilon}(t,\,b),\;\boldsymbol{y}_{\varepsilon}(t_{\varepsilon}) \in \mathcal{E}(\mathcal{M}_0) \setminus \mathcal{E}(\boldsymbol{m}')\,] \;=\; 0.$$

On the other hand, as $\alpha > 2$, for ε sufficiently small, $t_{\varepsilon} - s \in [2b,\,4b]$ for all $s \in [t - 3b,\,t - \alpha b]$. Hence, by the strong Markov property and Proposition 10.4,

$$\limsup_{b \to 0}\;\limsup_{\varepsilon \to 0}\;\sup_{\boldsymbol{x} \in \mathcal{E}(\boldsymbol{m})}\;\mathbb{Q}_{\boldsymbol{x}}^{\varepsilon}[\,B_{\varepsilon}(t,\,b),\;\boldsymbol{y}_{\varepsilon}(t_{\varepsilon}) \notin \mathcal{E}(\mathcal{M}_0)\,] \;=\; 0.$$

The assertion of the lemma follows from the previous estimates. □

Proof of Proposition 10.2. The proof is similar to the one of [21, Proposition 2.1]. We consider the case $\mathfrak{n} = 1$, the general one being similar.

Fix $t > 0$, $\boldsymbol{m}$, $\boldsymbol{m}' \in \mathcal{M}_0$, and sequences $\boldsymbol{x}_{\varepsilon} \in \mathcal{E}(\boldsymbol{m})$, $t_{\varepsilon} \to t$. By Theorem 9.3,

$$\mathcal{Q}_{\boldsymbol{m}}[\boldsymbol{y}(t) = \boldsymbol{m}'] \;=\; \lim_{\delta \to 0} \mathcal{Q}_{\boldsymbol{m}}[\boldsymbol{y}(t - 3\delta) = \boldsymbol{m}'] \;=\; \lim_{\delta \to 0} \lim_{\varepsilon \to 0} \mathbb{Q}_{\boldsymbol{x}_{\varepsilon}}^{\varepsilon}[\,\boldsymbol{y}_{\varepsilon}^{\mathrm{T}}(t - 3\delta) \in \mathcal{E}(\boldsymbol{m}')\,].$$

Thus, by Lemma 10.3,

$$\mathcal{Q}_{\boldsymbol{m}}[\boldsymbol{y}(t) = \boldsymbol{m}'] \;\leq\; \liminf_{\varepsilon \to 0} \mathbb{Q}_{\boldsymbol{x}_{\varepsilon}}^{\varepsilon}[\,\boldsymbol{y}_{\varepsilon}(t_{\varepsilon}) \in \mathcal{E}(\boldsymbol{m}')\,] \;\leq\; \limsup_{\varepsilon \to 0} \mathbb{Q}_{\boldsymbol{x}_{\varepsilon}}^{\varepsilon}[\,\boldsymbol{y}_{\varepsilon}(t_{\varepsilon}) \in \mathcal{E}(\boldsymbol{m}')\,].$$

Since

$$1 \;=\; \sum_{\boldsymbol{m}' \in \mathcal{M}_0} \mathcal{Q}_{\boldsymbol{m}}[\boldsymbol{y}(t) = \boldsymbol{m}'] \quad \text{and} \quad \sum_{\boldsymbol{m}' \in \mathcal{M}_0} \mathbb{Q}_{\boldsymbol{x}_{\varepsilon}}^{\varepsilon}[\,\boldsymbol{y}_{\varepsilon}(t_{\varepsilon}) \in \mathcal{E}(\boldsymbol{m}')\,] \;\leq\; 1,$$

the inequalities in the penultimate formula must be identities for each $\boldsymbol{m} \in \mathcal{M}_0$, which completes the proof of the proposition. □

Avoiding Wells

We complete the proof Proposition 10.2 by showing that the probability that the process is not in a well when it starts from a well is very small. This is the content of Proposition 10.4 below, the main result of this subsection.

Proposition 10.4. *For all* $\boldsymbol{m} \in \mathcal{M}_0$,

$$\limsup_{b \to 0}\;\limsup_{\varepsilon \to 0}\;\sup_{\boldsymbol{x} \in \mathcal{E}(\boldsymbol{m})}\;\sup_{t \in [2b,\,4b]}\;\mathbb{Q}_{\boldsymbol{x}}^{\varepsilon}[\,\boldsymbol{y}_{\varepsilon}(t) \notin \mathcal{E}(\mathcal{M}_0)\,] \;=\; 0.$$

The proof of this proposition requires some preliminary estimates. Fix $\eta \in (0, r_0/2)$ so that there is no critical point $c \in \mathcal{C}_0$ such that $U(c) \in (U(m'), U(m') + \eta)$ for some $m' \in \mathcal{M}_0$. Fix $m \in \mathcal{M}_0$, and let

$$\mathcal{R} = \mathcal{R}(m) = (\mathbb{R}^d \setminus \mathcal{E}(\mathcal{M}_0)) \cap \left\{ x \in \mathbb{R}^d : U(x) < U(m) + \eta/2 \right\}.$$

Denote by $\mathcal{W}$ the connected component of the set $\{x \in \mathbb{R}^d : U(x) < U(m) + d^{(1)}\}$ which contains m. We claim that

$$\mathcal{W} \cap \mathcal{R} = \varnothing. \tag{10.3}$$

Indeed, if $y \in \mathcal{R}$, $U(y) < U(m) + \eta/2$. By definition of $d^{(1)}$ and $\mathcal{E}(m)$, all points $z \in \mathcal{W}$ such that $U(z) \leq U(m) + \eta/2$ are contained in $\mathcal{E}(m)$. Hence, $\mathcal{W} \cap \mathcal{R} = \mathcal{E}(m) \cap \mathcal{R}$. By definition of $\mathcal{R}$, $\mathcal{E}(m) \cap \mathcal{R} = \varnothing$, which completes the proof of the claim.

Lemma 10.5. *Fix $m \in \mathcal{M}_0$. Then,*

$$\limsup_{a \to 0} \limsup_{\varepsilon \to 0} \sup_{x \in \mathcal{E}(m)} \mathbb{Q}_x^\varepsilon \left[\tau_\mathcal{R} \leq a \right] = 0.$$

Proof. By Lemma 6.6, it suffices to show that

$$\mathbb{Q}_x^\varepsilon \left[\tau_\mathcal{R} \leq a \right] \leq 2 \max_{m' \in \mathcal{M}} \sup_{z \in \mathcal{E}(m')} \mathbb{Q}_z^\varepsilon \left[\tau_{\mathcal{E}(\mathcal{M}_0) \setminus \mathcal{E}(m')} < 2a \right] + R_\varepsilon(x),$$

where $\sup_{x \in \mathcal{E}(m)} |R_\varepsilon(x)| \to 0$.

To prove the previous bound, first observe that

$$\mathbb{Q}_x^\varepsilon [\tau_\mathcal{R} \leq a] = \mathbb{Q}_x^\varepsilon [\tau_\mathcal{R} \leq a, \sigma_{\mathcal{E}(\mathcal{M}_0)} \leq \iota_\varepsilon] + \mathbb{Q}_x^\varepsilon [\tau_\mathcal{R} \leq a, \sigma_{\mathcal{E}(\mathcal{M}_0)} > \iota_\varepsilon], \tag{10.4}$$

where $\sigma_\mathcal{A}$, $\mathcal{A} \subset \mathbb{R}^d$, is the first time after $\tau_\mathcal{R}$ that the process visits $\mathcal{A}$:

$$\sigma_\mathcal{A} := \inf\{t > \tau_\mathcal{R} : x_\varepsilon(t) \in \mathcal{A}\},$$

$\iota_\varepsilon = a + \varepsilon^{-1}/\theta_\varepsilon^{(1)}$. By the strong Markov property, the second term on the right-hand is bounded by

$$\sup_{z \in \mathcal{R}} \mathbb{P}_z^\varepsilon \left[\tau_{\mathcal{E}(\mathcal{M}_0)} \geq \varepsilon^{-1} \right].$$

To keep notation simple, we replaced the measure $\mathbb{Q}_z^\varepsilon$ by $\mathbb{P}_z^\varepsilon$. By Corollary 6.2, this expression is bounded by a remainder $R_\varepsilon(x)$ such that $\sup_{x \in \mathcal{E}(m)} |R_\varepsilon(x)| \to 0$.

We turn to the first term on the right-hand side of (10.4). It can be written as

$$\mathbb{Q}_x^\varepsilon [\tau_\mathcal{R} \leq a, \sigma_{\mathcal{E}(m)} \leq \iota_\varepsilon] + \mathbb{Q}_x^\varepsilon [\tau_\mathcal{R} \leq a, \sigma_{\mathcal{E}(\mathcal{M}_0) \setminus \mathcal{E}(m)} \leq \iota_\varepsilon] \tag{10.5}$$

By the strong Markov property, the first term is bounded by

$$\sup_{z \in \mathcal{R}} \mathbb{P}_z^\varepsilon \left[\tau_{\mathcal{E}(m)} < 2a\theta_\varepsilon^{(1)} \right] = \max_\mathcal{A} \sup_{z \in \mathcal{A}} \mathbb{P}_z^\varepsilon \left[\tau_{\mathcal{E}(m)} < 2a\theta_\varepsilon^{(1)} \right],$$

where the maximum is carried over all connected components of $\mathcal{R}$. The number of connected component is finite because $U(x) \to \infty$ as $|x| \to \infty$. Fix a connected component $\mathcal{A}$ of $\mathcal{R}$, and let $\mathcal{B}$ be the connected component of $\{x \in \mathbb{R}^d : U(x) < U(m) + \eta\}$ containing $\mathcal{A}$. Since there are no critical points $c \in \mathcal{C}_0$ such that $U(c) \in (U(m), U(m) + \eta)$, by Corollary 6.3,

$$\lim_{\varepsilon \to 0} \sup_{z \in \mathcal{A}} \mathbb{P}^\varepsilon_z \big[\tau_{\partial \mathcal{B}} < \tau_{\mathcal{E}(\mathcal{B})} \big] = 0.$$

On the other hand, by (10.3), $\mathcal{E}(m) \subset \mathbb{R}^d \backslash \mathcal{B}$, so that $\tau_{\partial \mathcal{B}} < \tau_{\mathcal{E}(m)}$. Hence,

$$\sup_{z \in \mathcal{A}} \mathbb{P}^\varepsilon_z \big[\tau_{\mathcal{E}(m)} < 2a\theta^{(1)}_\varepsilon \big] \leq \sup_{z \in \mathcal{A}} \mathbb{P}^\varepsilon_z \big[\tau_{\mathcal{E}(m)} < 2a\theta^{(1)}_\varepsilon, \ \tau_{\mathcal{E}(\mathcal{B})} < \tau_{\mathcal{E}(m)} \big] + o_\varepsilon(1).$$

By the strong Markov property, this expression is bounded by

$$\sup_{z \in \mathcal{E}(\mathcal{B})} \mathbb{P}^\varepsilon_z \big[\tau_{\mathcal{E}(m)} < 2a\theta^{(1)}_\varepsilon \big] + o_\varepsilon(1).$$

Since $\mathcal{B}$ and $\mathcal{E}(m)$ are disjoint this expression is less than or equal to

$$\max_{m' \in \mathcal{M}} \sup_{z \in \mathcal{E}(m')} \mathbb{P}^\varepsilon_z \big[\tau_{\mathcal{E}(\mathcal{M}_0) \backslash \mathcal{E}(m')} < 2a\theta^{(1)}_\varepsilon \big] + o_\varepsilon(1).$$

We turn to the second term of (10.5). Since $\varepsilon^{-1} \prec \theta^{(1)}_\varepsilon$, it is bounded by

$$\mathbb{P}^\varepsilon_x \big[\tau_{\mathcal{E}(\mathcal{M}_0) \backslash \mathcal{E}(m)} < 2a\theta_\varepsilon \big],$$

which completes the proof of the lemma. $\qquad\square$

Proof of Proposition 10.4. Recall the definition of the set $\mathcal{R}$ introduced just before Lemma 10.5. Denote by $\mathcal{W}$ the connected component of the set $\{x \in \mathbb{R}^d : U(x) < U(m) + d^{(1)}\}$ which contains m. By (10.3), $\mathcal{R} \cap \mathcal{W} = \varnothing$. Clearly,

$$\mathbb{Q}^\varepsilon_x \big[y_\varepsilon(t) \in \mathbb{R}^d \setminus \mathcal{E}(\mathcal{M}_0) \big] \leq \mathbb{Q}^\varepsilon_x \big[y_\varepsilon(t) \in \mathbb{R}^d \setminus \{\mathcal{E}(\mathcal{M}_0) \cup \mathcal{R}\} \big] + \mathbb{Q}^\varepsilon_x \big[\tau_{\mathcal{R}} \leq t \big].$$

Recall that $\eta < r_0/2$, choose a time-scale ϱ_ε satisfying (4.9), and let $\kappa_\varepsilon = \varrho_\varepsilon/\theta^{(1)}_\varepsilon$. By Corollary 4.4, the first term on the right-hand side is bounded by

$$\mathbb{Q}^\varepsilon_{\mu^R_\varepsilon} \big[y_\varepsilon(t - \kappa_\varepsilon) \in \mathbb{R}^d \setminus \{\mathcal{E}(\mathcal{M}_0) \cup \mathcal{R}\} \big] + o_\varepsilon(1),$$

where the error is uniform over $t \in [2b, 4b]$, $x \in \mathcal{E}(m)$. As μ_ε is the stationary state and μ^R_ε the measure μ_ε conditioned to $\mathcal{W}^{2r_0}(m)$, the previous expression is equal to

$$\frac{\mu_\varepsilon(\mathbb{R}^d \setminus \{\mathcal{E}(\mathcal{M}_0) \cup \mathcal{R}\})}{\mu_\varepsilon(\mathcal{W}^{2r_0}(m))} + o_\varepsilon(1) = o_\varepsilon(1),$$

where the error terms are uniform on $x \in \mathcal{E}(m)$ and $t \in [2b, 4b]$.

It remains to show that

$$\limsup_{b \to 0} \limsup_{\varepsilon \to 0} \sup_{x \in \mathcal{E}(m)} \sup_{t \in [2b, 4b]} \mathbb{Q}^\varepsilon_x [\tau_{\mathcal{R}} \leq t] = 0.$$

This is a direct consequence of Lemma 10.5 since $\mathbb{Q}^\varepsilon_x[\tau_{\mathcal{R}} \leq t] \leq \mathbb{Q}^\varepsilon_x[\tau_{\mathcal{R}} \leq 4b]$ for all $t \leq 4b$. $\qquad\square$

Proof of Theorem 2.1

The assertion of Theorem 2.1 in the time scale $\theta_\varepsilon^{(1)}$ is a particular case of Theorem 10.1. We turn to the second claim.

Fix a time-scale ϱ_ε such that $1 \prec \varrho_\varepsilon \prec \theta_\varepsilon^{(1)}$, $\boldsymbol{m} \in \mathcal{M}_0$, $\boldsymbol{x} \in \mathcal{D}(\boldsymbol{m})$, $\eta > 0$, and a bounded continuous function F. Define the wells $\mathcal{E}(\boldsymbol{m}')$, $\boldsymbol{m}' \in \mathcal{M}_0$, as in the proof of Proposition 10.2, to fulfill (10.2). First, assume that there exists ε_0 such that $\varrho_\varepsilon \geq \varepsilon^{-2}$ for all $\varepsilon < \varepsilon_0$.

By [14, Theorem 2.1.2], there exists $T > 0$ such that

$$\mathbb{P}_x^\varepsilon[\boldsymbol{x}_\varepsilon(T) \notin \mathcal{E}(\boldsymbol{m})] = o_\varepsilon(1).$$

Hence, by the Markov property,

$$\mathbb{E}_x^\varepsilon[F(\boldsymbol{x}_\varepsilon(\varrho_\varepsilon))] = \mathbb{E}_x^\varepsilon\Big[\mathbb{E}_{\boldsymbol{x}_\varepsilon(T)}^\varepsilon\big[F(\boldsymbol{x}_\varepsilon(\varrho_\varepsilon - T))\big]\mathbf{1}\{\boldsymbol{x}_\varepsilon(T) \in \mathcal{E}(\boldsymbol{m})\}\Big] + o_\varepsilon(1).$$

As $\boldsymbol{x}_\varepsilon(T)$ belongs to $\mathcal{E}(\boldsymbol{m})$ and $\varrho_\varepsilon \prec \theta_\varepsilon^{(1)}$, by Lemma 6.6, inside the second expectation on the right-hand side we may insert the indicator of the set $\mathcal{A}_1 = \{\boldsymbol{x}_\varepsilon(\varrho_\varepsilon - T) \notin \mathcal{E}(\mathcal{M}_0)\backslash\mathcal{E}(\boldsymbol{m})\}$ at a cost $o_\varepsilon(1)$. By Proposition 4.2, we may also insert the indicator of the set $\mathcal{A}_2 = \{U(\boldsymbol{x}_\varepsilon(\varrho_\varepsilon - T - \varepsilon^{-1})) \leq U(\boldsymbol{m}) + d^{(1)} + 2r_0\}$ at the same cost. Hence, the left-hand side of the previous displayed equation is equal to

$$\mathbb{E}_x^\varepsilon\Big[\mathbb{E}_{\boldsymbol{x}_\varepsilon(T)}^\varepsilon\big[F(\boldsymbol{x}_\varepsilon(\varrho_\varepsilon - T))\,\mathbf{1}\{\mathcal{A}_1 \cap \mathcal{A}_2\}\big]\mathbf{1}\{\boldsymbol{x}_\varepsilon(T) \in \mathcal{E}(\boldsymbol{m})\}\Big] + o_\varepsilon(1).$$

By the Markov property the previous expectation is equal to

$$\mathbb{E}_x^\varepsilon\Big[\mathbb{E}_{\boldsymbol{x}_\varepsilon(T)}^\varepsilon\Big[\mathbf{1}\{\mathcal{A}_2\}\mathbb{E}_{\boldsymbol{x}_\varepsilon(\varrho_\varepsilon - T - (1/\varepsilon))}^\varepsilon\big[F(\boldsymbol{x}_\varepsilon(1/\varepsilon))\,\mathbf{1}\{\mathcal{A}_1'\}\big]\Big]\mathbf{1}\{\boldsymbol{x}_\varepsilon(T) \in \mathcal{E}(\boldsymbol{m})\}\Big],$$

where $\mathcal{A}_1' = \{\boldsymbol{x}_\varepsilon(\varepsilon^{-1}) \notin \mathcal{E}(\mathcal{M}_0)\backslash\mathcal{E}(\boldsymbol{m})\}$. Since $U(\boldsymbol{x}_\varepsilon(\varrho_\varepsilon - T - \varepsilon^{-1})) \leq U(\boldsymbol{m}) + d^{(1)} + 2r_0$, by Theorem 6.1 and Proposition 4.2, in the third expectation, we may insert the indicator of the set $\mathcal{A}_3 = \{\boldsymbol{x}_\varepsilon(1/\varepsilon) \in \mathcal{E}(\mathcal{M}_0)\}$ at a cost $o_\varepsilon(1)$. If by bad luck, there are critical points $\boldsymbol{c}$ such that $U(\boldsymbol{c}) = U(\boldsymbol{m}) + d^{(1)} + 2r_0$, we add to this constant a positive value to make sure that this does not happen. As $\mathcal{A}_4 = \mathcal{A}_1' \cap \mathcal{A}_3 = \{\boldsymbol{x}_\varepsilon(\varepsilon^{-1}) \in \mathcal{E}(\boldsymbol{m})\}$, by (10.2), the previous expression is equal to

$$F(\boldsymbol{m})\,\mathbb{E}_x^\varepsilon\Big[\mathbb{E}_{\boldsymbol{x}_\varepsilon(T)}^\varepsilon\Big[\mathbf{1}\{\mathcal{A}_2\}\mathbb{P}_{\boldsymbol{x}_\varepsilon(\varrho_\varepsilon - T - (1/\varepsilon))}^\varepsilon\big[\mathcal{A}_4\big]\Big]\mathbf{1}\{\boldsymbol{x}_\varepsilon(T) \in \mathcal{E}(\boldsymbol{m})\}\Big] + R(\varepsilon, \eta),$$

where $|R(\varepsilon, \eta)| \leq \eta + o_\varepsilon(1)$. We may now go backward in the argument to conclude that the previous expression is equal to $F(\boldsymbol{m}) + R(\varepsilon, \eta)$, which completes the proof of the theorem in the case where $\varrho_\varepsilon \geq \varepsilon^{-2}$ for all ε small.

Assume that this is not the case. We may suppose that $\varrho_\varepsilon \leq \varepsilon^{-2}$ for all ε small enough. If there is a subsequence which does not satisfy this condition, it is treated as in the first part of the proof.

By [14, Theorem 2.1.2], there exists $T > 0$ such that

$$\mathbb{P}_x^\varepsilon[\boldsymbol{x}_\varepsilon(T) \notin \mathcal{W}^{r_0/2}(\boldsymbol{m})] = o_\varepsilon(1).$$

Hence, by the Markov property,

$$\mathbb{E}_x^\varepsilon[\, F(\boldsymbol{x}_\varepsilon(\varrho_\varepsilon))\,] \;=\; \mathbb{E}_x^\varepsilon\Big[\mathbb{E}_{\boldsymbol{x}_\varepsilon(T)}^\varepsilon\big[\, F(\boldsymbol{x}_\varepsilon(\varrho_\varepsilon - T))\,\big]\,\mathbf{1}\{\boldsymbol{x}_\varepsilon(T) \in \mathcal{W}^{r_0/2}(\boldsymbol{m})\,\}\Big] \;+\; o_\varepsilon(1).$$

As $\boldsymbol{x}_\varepsilon(T) \in \mathcal{W}^{r_0/2}(\boldsymbol{m})$, by Proposition 4.2, in the second expectation, we may insert the indicator of the set $\mathcal{A} = \{\boldsymbol{x}_\varepsilon(\varrho_\varepsilon - T) \in \mathcal{E}(\boldsymbol{m})\}$ at a cost $o_\varepsilon(1)$. At this point, we may repeat the arguments presented at the end of the first part of the proof to conclude. □

Data Availibility Statement Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Appendix A. The Potential U

We present in this section elementary properties of the potential U and the dynamical system (2.4). The main result establishes the existence of a path which perfects the infimum in (2.2).

Proposition A.1. *Fix a local minimum $\boldsymbol{m} \in \mathcal{M}_0$. Then, there exist a local minimum $\boldsymbol{m}' \in \mathcal{M}_0$, $\boldsymbol{m}' \neq \boldsymbol{m}$ and a continuous path $\boldsymbol{z} \colon [0, 1] \to \mathbb{R}^d$ such that $\boldsymbol{z}(0) = \boldsymbol{m}$,*

$z(1) = m'$, and

$$\max_{t \in [0, 1]} U(z(t)) = U(z(1/2)) = U(m) + \Gamma(m) = \Theta(m, m'),$$

$$U(z(s)) < U(z(1/2)), \quad s \in [0, 1] \setminus \{1/2\}.$$

Moreover, if $z(\cdot)$ is such a path, then $z(1/2)$ is a saddle point of U.

Proof of Proposition A.1

The proof is based on three lemmata. Fix $m \in \mathcal{M}_0$, and let $\mathcal{W}$ be the connected component of $\{x \in \mathbb{R}^d : U(x) < U(m) + \Gamma(m)\}$ containing m. By definition of $\Gamma(m)$,

$$\mathcal{M}_0 \cap \mathcal{W} = \{m\}. \tag{A.1}$$

Lemma A.2. *Fix $m \in \mathcal{M}_0$ There is a connected component $\mathcal{W}'$ of $\{x \in \mathbb{R}^d : U(x) < U(m) + \Gamma(m)\}$ such that $\mathcal{W} \cap \mathcal{W}' = \varnothing$ and $\overline{\mathcal{W}} \cap \overline{\mathcal{W}'} \neq \varnothing$.*

The proof of this result is given in a subsection below. Recall that we denote by $B(x, r)$ the open ball of radius r centered at x. Let

$$\mathcal{A}(x, r) = (B(x, r) \setminus \{x\}) \cap \{y \in \mathbb{R}^d : U(y) < U(x)\}.$$

Lemma A.3. *Fix $H \in \mathbb{R}$, and let $\mathcal{W}_1$ and $\mathcal{W}_2$ be two disjoint connected components of $\{x \in \mathbb{R}^d : U(x) < H\}$. If $\overline{\mathcal{W}_1} \cap \overline{\mathcal{W}_2} \neq \varnothing$, then $\overline{\mathcal{W}_1} \cap \overline{\mathcal{W}_2} = \partial \mathcal{W}_1 \cap \partial \mathcal{W}_2$ and any element σ of $\overline{\mathcal{W}_1} \cap \overline{\mathcal{W}_2}$ is a saddle point such that $U(\sigma) = H$. Moreover, for all $r > 0$ small enough, $\mathcal{A}(\sigma, r)$ has two connected components: $\mathcal{A}(\sigma, r) \cap \mathcal{W}_1$ and $\mathcal{A}(\sigma, r) \cap \mathcal{W}_2$*

The proof of this lemma is presented in a later subsection below.

Lemma A.4. *Let $m', m'' \in \mathcal{M}_0$ and let $z : [0, 1] \to \mathbb{R}^d$ be a continuous path such that*

$$z(0) = m', \quad z(1) = m'', \quad U(z(1/2)) = \Theta(m', m'')$$
$$U(z(t)) < U(z(1/2) \text{ for } t \in [0, 1] \setminus \{1/2\}. \tag{A.2}$$

Then, $z(1/2)$ is a saddle point.

Proof. Recall that we denote by $\upsilon_x(t)$ the solution of the ODE (2.4) starting from x. For $s \geq 0$, let $\psi_s : [0, 1] \to \mathbb{R}^d$ be the continuous path defined by

$$\psi_s(t) = \upsilon_{z(t)}(s).$$

As U decreases along the solutions of the ODE,

$$U(\psi_s(t)) = U(\upsilon_{z(t)}(s)) \leq U(z(t)).$$

We claim that

$$U(\psi_s(1/2)) = U(z(1/2)) = U(\psi_0(1/2)) \quad \text{for all } s > 0. \tag{A.3}$$

Suppose, by contradiction, that there exists $s_0 > 0$ such that $U(\psi_{s_0}(1/2)) < U(z(1/2))$. By (A.2), for all $t \neq 1/2$,

$$U(\psi_{s_0}(t)) \leq U(z(t)) < U(z(1/2)).$$

Thus, since by hypothesis, $U(\psi_{s_0}(1/2)) < U(z(1/2))$,

$$\max_{t \in [0,1]} U(\psi_{s_0}(t)) < U(z(1/2)) = \Theta(\boldsymbol{m}', \boldsymbol{m}'') . \tag{A.4}$$

As $\boldsymbol{m}'$, $\boldsymbol{m}''$ are critical points, $\upsilon_{\boldsymbol{n}}(s) = \boldsymbol{n}$ for $\boldsymbol{n} = \boldsymbol{m}'$, $\boldsymbol{m}''$, $s > 0$, so that $\psi_{s_0}(0) = \upsilon_{z(0)}(s_0) = \upsilon_{\boldsymbol{m}'}(s_0) = \boldsymbol{m}'$, $\psi_{s_0}(1) = \boldsymbol{m}''$. Therefore, the continuous path $\psi_{s_0} \colon [0, 1] \to \mathbb{R}^d$ satisfies $\psi_{s_0}(0) = \boldsymbol{m}'$, $\psi_{s_0}(1) = \boldsymbol{m}''$ and (A.4). This contradicts to the definition of $\Theta(\boldsymbol{m}', \boldsymbol{m}'')$, and completes the proof of claim (A.3).

It follows from (A.3) and from the fact U strictly decreases along trajectories which do not start from critical points that $z(1/2)$ is a critical point of U.

It remains to show that $z(1/2)$ is a saddle point. Clearly, $z(1/2)$ is not a local minimum. Suppose, by contradiction, that $z(1/2)$ is not a saddle point. Then, by Lemma A.8 below, the set $\mathcal{A}(z(1/2), r)$ is connected for sufficiently small $r > 0$. Since z is continuous, there is $\eta_0 = \eta_0(r) > 0$ such that $z(t) \in B(z(1/2), r)$ for all $t \in [1/2 - \eta_0, 1/2 + \eta_0]$. Therefore, $z(t) \in \mathcal{A}(z(1/2), r)$ for all $t \in [1/2 - \eta_0, 1/2 + \eta_0] \setminus \{1/2\}$. Since $\mathcal{A}(z(1/2), r)$ is connected and open, it is path connected. Therefore, there is a continuous path $z_1 \colon [1/2 - \eta_0, 1/2 + \eta_0] \to \mathcal{A}(z(1/2), r)$ such that $z_1(1/2 \pm \eta_0) = z(1/2 \pm \eta_0)$. Define a path $z_2 \colon [0, 1] \to \mathbb{R}^d$ as

$$z_2(t) = \begin{cases} z(t) & t \in [0, 1/2 - \eta_0) \cup (1/2 + \eta_0, 1] \\ z_1(t) & t \in [1/2 - \eta_0, 1/2 + \eta_0] . \end{cases}$$

Thus, z_2 is a continuous trajectory from $\boldsymbol{m}'$ to $\boldsymbol{m}''$ such that $U(z(t)) < U(z(1/2))$ for all $t \in [0, 1]$. This contradicts the definition of $\Theta(\boldsymbol{m}', \boldsymbol{m}'')$, and completes the proof of the lemma. $\qquad\square$

Proof of Proposition A.1. Fix $\boldsymbol{m} \in \mathcal{M}_0$. Let $\mathcal{W}'$ be given by Lemma A.2, and denote by $\boldsymbol{\sigma}$ an element of $\overline{\mathcal{W}} \cap \overline{\mathcal{W}}'$. By Lemma A.3, $\boldsymbol{\sigma}$ is a saddle point, $\boldsymbol{\sigma} \in \Upsilon(\boldsymbol{m})$, and, for sufficiently small $r > 0$, $\mathcal{A}(\boldsymbol{\sigma}, r)$ has two connected components $\mathcal{A}(\boldsymbol{\sigma}, r) \cap \mathcal{W}$ and $\mathcal{A}(\boldsymbol{\sigma}, r) \cap \mathcal{W}'$. By Hartman–Grobman theorem, there are two continuous path $\phi_1, \phi_2 \colon (-\infty, 0] \to \mathbb{R}^d$ such that

$$\lim_{t \to -\infty} \phi_j(t) = \boldsymbol{\sigma} \quad \phi_1(s) \in \mathcal{A}(\boldsymbol{\sigma}, r) \cap \mathcal{W} \quad \phi_2(s) \in \mathcal{A}(\boldsymbol{\sigma}, r) \cap \mathcal{W}'$$

for all $s \leq 0$. Since $\mathcal{W}$, $\mathcal{W}'$ are connected, we may extend continuously these trajectories to $s > 0$ in such a way that $\phi_1(s) \in \mathcal{W}$, $\phi_2(s) \in \mathcal{W}'$ for all $s \geq 0$. As $\boldsymbol{\sigma} \in \Upsilon(\boldsymbol{m})$, by (2.5),

$$\lim_{s \to \infty} \phi_1(s) = \boldsymbol{m}, \quad \lim_{s \to \infty} \phi_2(s) = \boldsymbol{m}',$$

where $\boldsymbol{m}'$ is a local minimum of U one in $\mathcal{W}'$.

Concatenating the paths ϕ_1, ϕ_2 and reparametrizing it, we obtain a continuous path $z : [0, 1] \to \mathbb{R}^d$ from $\boldsymbol{m}$ to $\boldsymbol{m}'$ such that $z(1/2) = \boldsymbol{\sigma}$. By Lemma A.3, $U(\boldsymbol{\sigma}) = U(\boldsymbol{m}) + \Gamma(\boldsymbol{m})$. Therefore, by construction, $z(\cdot)$ fulfills all conditions required in Proposition A.1.

It remains to check the final assertion of the proposition, which follows from Lemma A.4. $\qquad\qquad\square$

Proof of Lemma A.3

Throughout this subsection, we will use the fact that an open connected subset of $\mathbb{R}^d$ is path-connected.

Lemma A.5. *Homeomorphisms preserve the number of open connected components.*

Proof. Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be open sets, and let $\varphi : \mathcal{U}_1 \to \mathcal{U}_2$ be a homeomorphism. Denote by $\mathcal{U}_{j,1}, \ldots, \mathcal{U}_{j,n_j}$ the connected components of $\mathcal{U}_j$, $j = 1, 2$. Since φ is continuous, $\varphi(\mathcal{U}_{1,k})$ is connected. As φ is surjective, $\mathcal{U}_2 = \cup_{1 \leq k \leq n_1} \varphi(\mathcal{U}_{1,k})$, so that $n_2 \leq n_1$. Since φ^{-1} is continuous, the same argument yields the reverse inequality. $\square$

Lemma A.6. *Let $\boldsymbol{p}$ be non-critical point of U. Then, for sufficiently small $r > 0$, the manifold $\{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) = U(\boldsymbol{p})\}$ divides $B(\boldsymbol{p}, r)$ into two connected components which are $B(\boldsymbol{p}, r) \cap \{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) < U(\boldsymbol{p})\}$ and $B(\boldsymbol{p}, r) \cap \{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) > U(\boldsymbol{p})\}$. In particular, $\mathcal{A}(\boldsymbol{p}, r)$ is connected. Furthermore, there is a continuous path $z : [0, 1] \to B(\boldsymbol{p}, r)$ such that*

$$z(0) = \boldsymbol{p} \quad z((0, 1]) \subset B(\boldsymbol{p}, r) \cap \{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) < U(\boldsymbol{p})\} .$$

Proof. Fix $\boldsymbol{p} = (p_1, \ldots, p_d) \in \mathbb{R}^d$ be a non-critical point. Then, $\nabla U(\boldsymbol{p}) \neq 0$ so that there is $1 \leq j \leq d$ such that

$$\frac{\partial U}{\partial x_j}(\boldsymbol{p}) \neq 0 .$$

Assume, without loss of generality, that $j = d$. For $\boldsymbol{x} \in \mathbb{R}^d$, let

$$\widetilde{\boldsymbol{x}} = (x_1, \ldots, x_{d-1}) .$$

By the implicit function theorem, there exist $r > 0$ and a C^1-function $g : \mathbb{R}^{d-1} \to \mathbb{R}$ such that

$$g(\widetilde{\boldsymbol{p}}) = p_d, \quad U(\widetilde{\boldsymbol{x}}, g(\widetilde{\boldsymbol{x}})) = U(\boldsymbol{p}) \text{ for all } \widetilde{\boldsymbol{x}} \in B_{d-1}(\widetilde{\boldsymbol{p}}, r)$$

where $B_{d-1}(\widetilde{\boldsymbol{p}}, r)$ is a $(d - 1)$-dimensional ball with radius $r > 0$ centered at $\widetilde{\boldsymbol{p}}$. Decompose the set $B(\boldsymbol{p}, r)$ into three parts:

$$\mathcal{P}_1 = B(\boldsymbol{p}, r) \cap \{(\widetilde{\boldsymbol{x}}, y) \in \mathbb{R}^d : y > g(\widetilde{\boldsymbol{x}})\}$$
$$\mathcal{P}_2 = B(\boldsymbol{p}, r) \cap \{(\widetilde{\boldsymbol{x}}, y) \in \mathbb{R}^d : y < g(\widetilde{\boldsymbol{x}})\}$$

$$\mathcal{P}_3 = B(\boldsymbol{p}, r) \cap \{(\widetilde{\boldsymbol{x}}, \boldsymbol{y}) \in \mathbb{R}^d : \boldsymbol{y} = g(\widetilde{\boldsymbol{x}})\} .$$

By definition of g, $\mathcal{P}_3 = B(\boldsymbol{p}, r) \cap \{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) = U(\boldsymbol{p})\}$ and

$$U(\boldsymbol{x}) \neq U(\boldsymbol{p}) \text{ for all } \boldsymbol{x} \in \mathcal{P}_1 \cup \mathcal{P}_2 . \tag{A.5}$$

Suppose that there is $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{P}_1$ such that $U(\boldsymbol{x}) < U(\boldsymbol{p}) < U(\boldsymbol{y})$. As $\mathcal{P}_1$ is path-connected, there is a path in $\mathcal{P}_1$ connecting $\boldsymbol{x}$ to $\boldsymbol{y}$. Since U is continuous, this path must pass through a point $\boldsymbol{z} \in \mathcal{P}_1$ such that $U(\boldsymbol{z}) = U(\boldsymbol{p})$, and this contradicts (A.5). Therefore, $U(\boldsymbol{x}) > U(\boldsymbol{p})$ for all $\boldsymbol{x} \in \mathcal{P}_1$ or $U(\boldsymbol{x}) < U(\boldsymbol{p})$ for all $\boldsymbol{x} \in \mathcal{P}_1$. Let $\boldsymbol{v} = \nabla U(\boldsymbol{p})$. For sufficiently small $\eta > 0$, $U(\boldsymbol{p}+\eta\boldsymbol{v}) > U(\boldsymbol{p})$ and $U(\boldsymbol{p}-\eta\boldsymbol{v}) < U(\boldsymbol{p})$. Thus, there is $\boldsymbol{x}, \boldsymbol{y} \in B(\boldsymbol{p}, r)$ such that $U(\boldsymbol{x}) < U(\boldsymbol{p}) < U(\boldsymbol{y})$. Therefore, one of the sets $\mathcal{P}_1, \mathcal{P}_2$ is $B(\boldsymbol{p}, r) \cap \{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) < U(\boldsymbol{p})\}$ and the other one is $B(\boldsymbol{p}, r) \cap \{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) > U(\boldsymbol{p})\}$.

Finally, since $\mathcal{P}_3$ is the graph of a C^1 function, there are paths $z_i : [0, 1] \to B(\boldsymbol{p}, r)$ such that

$$z_i(0) = \boldsymbol{p} \ \ z_i((0, 1]) \subset \mathcal{P}_i .$$

This completes the proof of the lemma. □

Critical points. The next two lemmata provide the number of connected componentes of the set $\mathcal{A}(\boldsymbol{c}, r)$, $\boldsymbol{c} \in \mathcal{C}_0$, in terms of the index of the critical points $\boldsymbol{c}$.

Lemma A.7. *Let $\boldsymbol{\sigma}$ be a saddle point of U. Then, for sufficiently small $r > 0$, the set $\mathcal{A}(\boldsymbol{\sigma}, r) = (B(\boldsymbol{\sigma}, r) \backslash \{\boldsymbol{\sigma}\}) \cap \{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) < U(\boldsymbol{\sigma})\}$ has exactly two connected components.*

Proof. By [33, Lemma 2.2], since U is nondegenerate at $\boldsymbol{\sigma}$, the image of U near $\boldsymbol{\sigma}$ is locally diffeomorphic to the quadratic function $F : \mathbb{R}^d \to \mathbb{R}$ given by

$$F(\boldsymbol{x}) = -x_1^2 + \sum_{i=2}^{d} x_i^2 .$$

Therefore, for sufficiently small $r > 0$, $\mathcal{A}(\boldsymbol{\sigma}, r)$ is diffeomorphic to the set

$$[\, B(\boldsymbol{0}, r) \backslash \{\boldsymbol{0}\} \,] \cap F^{-1}((-\infty, 0)) = [\, B(\boldsymbol{0}, r) \backslash \{\boldsymbol{0}\} \,] \cap \{\boldsymbol{x} \in \mathbb{R}^d : -x_1^2$$
$$+ \sum_{i=2}^{d} x_i^2 < 0\} .$$

Since the set on the right-hand side has two connected components, by Lemma A.5, $\mathcal{A}(\boldsymbol{\sigma}, r)$ has also two connected components. □

Lemma A.8. *Let $\boldsymbol{c}$ be a critical point of U with index greater or equal to 2. Then, for sufficiently small $r > 0$, $\mathcal{A}(\boldsymbol{c}, r)$ is path-connected.*

Proof. By [33, Lemma 2.2], since U is nondegenerate at c, the image of U near c is locally diffeomorphic to to the quadratic function $F : \mathbb{R}^d \to \mathbb{R}$ given by

$$F(x) \;=\; - \sum_{i=1}^{k} x_i^2 \;+\; \sum_{i=k+1}^{d} x_i^2$$

where $k \geq 2$ is the index of c. Therefore, for sufficiently small $r > 0$, $\mathcal{A}(c, r)$ is diffeomorphic to

$$[\, B(\mathbf{0}, r) \setminus \{\mathbf{0}\}\,] \cap \Big\{ x \in \mathbb{R}^d : - \sum_{i=1}^{k} x_i^2 \;+\; \sum_{i=k+1}^{d} x_i^2 < 0 \Big\}.$$

Since this set is connected, by Lemma A.5, $\mathcal{A}(c, r)$ is also connected, and therefore path-connected. □

Level sets. In this subsection, we examine the connected components of the level sets of U.

Lemma A.9. *Fix $H \in \mathbb{R}$. Let $\mathcal{H}$ be a connected component of $\{x \in \mathbb{R}^d : U(x) < H\}$. Let $\mathcal{G} \subset \{x \in \mathbb{R}^d : U(x) < H\}$ be a connected set satisfying $\mathcal{G} \cap \mathcal{H} \neq \varnothing$. Then, $\mathcal{G} \subset \mathcal{H}$. The same assertion holds if we replace all strict inequalities by inequalities.*

Proof. Let $x_0 \in \mathcal{H}$. Then, $\mathcal{H}$ is the largest connected set $\mathcal{F}$ satisfying

$$x_0 \in \mathcal{F} \subset \{x \in \mathbb{R}^d : U(x) < H\}\,.$$

As $\mathcal{G} \cap \mathcal{H} \neq \varnothing$, there exists $x_0 \in \mathcal{G} \cap \mathcal{H}$. As $\mathcal{G}$ belongs to the previous class, $\mathcal{G} \subset \mathcal{H}$.

The same proof yields the second assertion of the lemma. □

Lemma A.10. *Fix $H \in \mathbb{R}$. Let $\mathcal{H}$ be a connected component of the set $\{x \in \mathbb{R}^d : U(x) < H\}$ or one of the set $\{x \in \mathbb{R}^d : U(x) \leq H\}$. Then, $U(x_0) = H$ for all $x_0 \in \partial \mathcal{H}$. Moreover,*

 (1) *If $\mathcal{H}$ is an open set, then x_0 is not a local minimum*
 (2) *If $\mathcal{H}$ is a closed set, then x_0 is not a local maximum*

Proof. Fix $x_0 \in \partial \mathcal{H}$. Since U is continuous, $U(x_0) \leq H$. Assume by contradiction that $U(x_0) < H$. Let $\mathcal{G}$ be the connected component of the set $\{x \in \mathbb{R}^d : U(x) < H\}$ containing x_0. Since U is smooth, there exists $r > 0$ such that

$$\max_{y \in B(x_0, r)} U(y) < H$$

As $x_0 \in \partial \mathcal{H}$, there exists $z \in B(x_0, r) \cap \mathcal{H}$. Hence, by the previous displayed equation, $B(x_0, r) \subset \mathcal{H}$, so that $x_0 \in \mathcal{H}$, in contradiction to the fact that $x_0 \in \partial \mathcal{H}$. This completes the proof of the first assertion.

Suppose that $\mathcal{H}$ is a connected component of the set $\{x \in \mathbb{R}^d : U(x) < H\}$, and fix $x_0 \in \partial \mathcal{H}$. By the first assertion of the lemma, $U(x_0) = H$. Suppose by contradiction

that x_0 is a local minimum. Then, there exists $r > 0$ such that $U(y) \geq U(x_0)$ for all $y \in B(x_0, r)$. Therefore, $B(x_0, r) \cap \{x \in \mathbb{R}^d : U(x) < H\} = \varnothing$, so that $B(x_0, r) \cap \mathcal{H} = \varnothing$. This contradicts the fact that $x_0 \in \partial\mathcal{H}$.

Suppose that $\mathcal{H}$ is a connected component of the set $\{x \in \mathbb{R}^d : U(x) \leq H\}$, and fix $x_0 \in \partial\mathcal{H}$. By the first assertion of the lemma, $U(x_0) = H$. Suppose by contradiction that x_0 is a local maximum. Then, there exists $r > 0$ such that $U(y) < U(x_0)$ for all $y \in B(x_0, r)\backslash\{x_0\}$. Therefore, $B(x_0, r)\backslash\{x_0\} \subset \{x \in \mathbb{R}^d : U(x) < H\}$.

Since $x_0 \in \partial\mathcal{H}$, $B(x_0, r) \cap \mathcal{H}^o \neq \varnothing$, where $\mathcal{H}^o$ is the interior of $\mathcal{H}$. Fix $x_1 \in B(x_0, r) \cap \mathcal{H}^o$ and let $\mathcal{G}$ be the connected component of $\{x \in \mathbb{R}^d : U(x) < H\}$ containing x_1. As $x_1 \in \mathcal{H}^o \subset \mathcal{H}$, by definition of $\mathcal{H}$, $\mathcal{G} \subset \mathcal{H}$. On the other hand, by Lemma A.9, $B(x_0, r) \setminus \{x_0\} \subset \mathcal{G}$, so that $B(x_0, r) \setminus \{x_0\} \subset \mathcal{H}$. As $x_0 \in \mathcal{H}$, $B(x_0, r) \subset \mathcal{H}$. This contradicts the fact that $x_0 \in \partial\mathcal{H}$, and completes the proof of the lemma. □

For the next lemma, we extend the definition of $\Theta(m, m')$ to subsets of $\mathcal{M}_0$. For two disjoint non-empty subsets $\mathcal{M}'$ and $\mathcal{M}''$ of $\mathcal{M}_0$, define

$$\Theta(\mathcal{M}', \mathcal{M}'') = \min_{m' \in \mathcal{M}', m'' \in \mathcal{M}''} \Theta(m', m''). \tag{A.6}$$

Lemma A.11. *Let $\mathcal{H} \subset \mathbb{R}^d$ be a connected component of level set $\{x \in \mathbb{R}^d : U(x) < c_0\}$ for some $c_0 \in \mathbb{R}$. Let $\mathcal{M}$, $\mathcal{M}'$ be disjoint non-empty subsets of $\mathcal{M}_0$.*

(1) *If $\mathcal{M}$, $\mathcal{M}' \subset \mathcal{H}$, then $\Theta(\mathcal{M}, \mathcal{M}') < c_0$.*
(2) *If $\mathcal{M} \subset \mathcal{H}$ and $\mathcal{M}' \subset \mathbb{R}^d \backslash \mathcal{H}$, then $\Theta(\mathcal{M}, \mathcal{M}') \geq c_0$.*

Proof. Let $\mathcal{M}$, $\mathcal{M}'$ be disjoint non-empty subsets of $\mathcal{M}_0$ contained in $\mathcal{H}$. Since $\mathcal{H}$ is open connected set, it is a path connected set. Thus, there exists a connected path $z : [0, 1] \to \mathcal{H}$ such that $z(0) \in \mathcal{M}$ and $z(1) \in \mathcal{M}'$. Since $z(t) \in \mathcal{H}$ for all $t \in [0, 1]$, we have $\max_{t \in [0, 1]} U(z(t)) < c_0$ and thus by (A.6), $\Theta(\mathcal{M}, \mathcal{M}') < c_0$. This proves the first assertion.

To prove the second assertion note that any path connecting $\mathcal{M}$ and $\mathcal{M}'$ must pass through $\partial\mathcal{H}$ on which the value of U is c_0. □

Proof of Lemma A.3. Let $\sigma \in \overline{\mathcal{W}_1} \cap \overline{\mathcal{W}_2}$. We claim that $\sigma \in \partial\mathcal{W}_1 \cap \partial\mathcal{W}_2$. Indeed, by definition $\sigma \in \overline{\mathcal{W}_1}$. It remains to show that $\sigma \notin \mathcal{W}_1$. Assume, by contradiction, that $\sigma \in \mathcal{W}_1$. Then, there exists $r > 0$ such that $B(\sigma, r) \subset \mathcal{W}_1$. Since $\mathcal{W}_1 \cap \mathcal{W}_2 = \varnothing$, $B(\sigma, r) \cap \mathcal{W}_2 = \varnothing$, which contradicts the fact that $\sigma \in \overline{\mathcal{W}_2}$. Thus, $\sigma \in \partial\mathcal{W}_1$. The same argument shows that $\sigma \in \partial\mathcal{W}_2$, proving the claim. By Lemma A.10, $U(\sigma) = H$, and σ is not a local minimum.

By definition, there exists $r > 0$ such that $B(\sigma, r) \cap \mathcal{W}_1 \neq \varnothing$ and $B(\sigma, r) \cap \mathcal{W}_2 \neq \varnothing$. Since $\sigma \in \partial\mathcal{W}_1 \cap \partial\mathcal{W}_2$, $\sigma \notin \mathcal{W}_1 \cup \mathcal{W}_2$, so that $(B(\sigma, r)\backslash\{\sigma\}) \cap \mathcal{W}_1 \neq \varnothing$ and $(B(\sigma, r)\backslash\{\sigma\}) \cap \mathcal{W}_2 \neq \varnothing$. Hence, by definition of $\mathcal{W}_1$ and $\mathcal{W}_2$, $\mathcal{A}(\sigma, r)$ is not empty.

We claim that $\mathcal{A}(\sigma, r)$ is not connected. Suppose, by contradiction, that $\mathcal{A}(\sigma, r)$ is connected. Let $x_1 \in B(\sigma, r) \cap \mathcal{W}_1$, $x_2 \in B(\sigma, r) \cap \mathcal{W}_2$. Since $U(x_j) < U(\sigma)$, $x_1, x_2 \in \mathcal{A}(\sigma, r)$. Since $\mathcal{A}(\sigma, r)$ is open, there exists a continuous path $z : [0, 1] \to \mathbb{R}^d$ connecting x_1 to x_2 in $\mathcal{A}(\sigma, r)$. In particular, $\sup_{0 \leq t \leq 1} U(z(t)) < U(\sigma) = H$. Since $x_1 \in \mathcal{W}_1$ and $\mathcal{W}_1$ is a connected component of the set $\{x :$

$U(x) < H\}$, all points in this path, including x_2, belong to $\mathcal{W}_1$. As $x_2 \in \mathcal{W}_2$ and $\mathcal{W}_1 \cap \mathcal{W}_2 = \varnothing$, this is a contradiction, which proves the claim.

Since σ is not a local minimum, and $\mathcal{A}(\sigma, r)$ is not empty and not connected, by Lemmata A.6, A.8, σ is a saddle point. By Lemma A.7, $\mathcal{A}(\sigma, r)$ has exactly two components. Let $\mathcal{A}_1, \mathcal{A}_2$, be the connected component which intersects with $\mathcal{W}_1$, $\mathcal{W}_2$, respectively. Since $\mathcal{A}_j$ is a connected set contained in $\{x \in \mathbb{R}^d : U(x) < U(\sigma)\}$, by Lemma A.9, $\mathcal{A}_1 \subset \mathcal{W}_1$ and $\mathcal{A}_2 \subset \mathcal{W}_2$. Hence, $\mathcal{A}_1 \neq \mathcal{A}_2$ and $\mathcal{A}_1 = \mathcal{A}(\sigma, r) \cap \mathcal{W}_1$, $\mathcal{A}_2 = \mathcal{A}(\sigma, r) \cap \mathcal{W}_2$. $\qquad\square$

Proof of Lemma A.2

The proof relies on several lemmata.

Lemma A.12. *Let $\mathcal{K}_n$ be a decreasing sequence of compact connected sets and let $\mathcal{K} := \bigcap_{n=1}^{\infty} \mathcal{K}_n$. Then, $\mathcal{K}$ is connected.*

Proof. Suppose, by contradiction, that $\mathcal{K}$ is not connected. In consequence, there are two disjoint open sets $\mathcal{U}$ and $\mathcal{V}$ such that $\mathcal{K} \cap \mathcal{U} \neq \varnothing, \mathcal{K} \cap \mathcal{V} \neq \varnothing$, and $\mathcal{K} \subset \mathcal{U} \cup \mathcal{V}$. Since $\mathcal{K}_n \cap \mathcal{V} \neq \varnothing$ and $\mathcal{U} \cap \mathcal{V} = \varnothing$, $\mathcal{K}_n \setminus \mathcal{U} \neq \varnothing$.

We claim that $\mathcal{K}_n \cap \partial \mathcal{U} \neq \varnothing$. Suppose by contradiction that $\mathcal{K}_n \cap \partial \mathcal{U} = \varnothing$. In this case, $\mathbb{R}^d = [\mathcal{K}_n \cap \partial \mathcal{U}]^c = \mathcal{K}_n^c \cup (\partial \mathcal{U})^c$, so that $\mathcal{K}_n = \mathcal{K}_n \cap (\partial \mathcal{U})^c$. Hence, $\mathcal{K}_n \setminus \mathcal{U} = \mathcal{K}_n \cap \mathcal{U}^c = \mathcal{K}_n \cap (\partial \mathcal{U})^c \cap \mathcal{U}^c = \mathcal{K}_n \cap [(\partial \mathcal{U}) \cup \mathcal{U}]^c = \mathcal{K}_n \cap \overline{\mathcal{U}}^c \subset \overline{\mathcal{U}}^c$. Therefore, as $\overline{\mathcal{U}}^c$ is an open set, for all $x \in \mathcal{K}_n \setminus \mathcal{U}$, there exists $r(x) > 0$ such that $B(x, r(x)) \subset \overline{\mathcal{U}}^c$. Since $\mathcal{K}_n$ is compact, $\mathcal{K}_n \setminus \mathcal{U}$ is compact so that there are finitely many $x_1, \ldots, x_k \in \mathcal{K}_n \setminus \mathcal{U}$ such that

$$\mathcal{K}_n \setminus \mathcal{U} \subset \bigcup_{j=1}^{k} B(x_j, r(x_j)) .$$

Therefore, $\mathcal{K}_n \subset \mathcal{U} \cup \bigcup_{j=1}^{k} B(x_j, r(x_j))$. However, since $B(x_j, r(x_j)) \subset \overline{\mathcal{U}}^c$ for all $j, \mathcal{U} \cap \bigcup_{j=1}^{k} B(x_j, r(x_j)) = \varnothing$, in contradiction with the connectedness of $\mathcal{K}_n$. This proves the claim.

As $(\mathcal{K}_n \cap \partial \mathcal{U})$ is a decreasing sequence of compact sets, $\mathcal{K} \cap \partial \mathcal{U} = \bigcap_{n=1}^{\infty} (\mathcal{K}_n \cap \partial \mathcal{U}) \neq \varnothing$ by Cantor's intersection theorem. Let $x_0 \in \mathcal{K} \cap \partial \mathcal{U}$. Since $\mathcal{U}$ is open, $x_0 \notin \mathcal{U}$ so that $x_0 \in \mathcal{V}$. Since $\mathcal{V}$ is open, there exists $r_0 > 0$ such that $B(x_0, r_0) \subset \mathcal{V}$ so that $B(x_0, r_0) \cap \mathcal{U} = \varnothing$, which contradicts the fact that $x_0 \in \partial \mathcal{U}$. This completes the proof of the lemma. $\qquad\square$

Lemma A.13. *Let $\mathcal{K}_n$ be a decreasing sequence of compact sets. Suppose that $\mathcal{K} := \bigcap_{n=1}^{\infty} \mathcal{K}_n$ is contained in an open set $\mathcal{U}$. Then, there exists $N \in \mathbb{N}$ such that $\mathcal{K}_N \subset \mathcal{U}$.*

Proof. Suppose, by contradiction, that each $\mathcal{K}_n$ is not contained in $\mathcal{U}$. Then, for each $n \in \mathbb{N}$, there exists $x_n \in \mathcal{K}_n \setminus \mathcal{U} \subset \mathcal{K}_1 \setminus \mathcal{U}$. As $\mathcal{K}_1 \setminus \mathcal{U}$ is compact, there is a

subsequence $(x'_n)_{n \geq 1}$ which converges to a point $x_0 \in \mathcal{K}_1 \setminus \mathcal{U}$. Since $x'_j \in \mathcal{K}_m \setminus \mathcal{U}$ for all $j \geq m$, $x_0 \in \mathcal{K}_m \setminus \mathcal{U}$. Therefore,

$$x_0 \in \bigcap_{n=1}^{\infty} (\mathcal{K}_n \setminus \mathcal{U}) = \mathcal{K} \setminus \mathcal{U} = \varnothing$$

which is a contradiction. $\qquad \square$

Lemma A.14. *Let $\mathcal{H}$ be a connected component of the set $\{x \in \mathbb{R}^d : U(x) \leq H\}$, and let $\mathcal{U}$ be an open set containing $\mathcal{H}$. Then, there is $N \in \mathbb{N}$ such that the connected component of*

$$\left\{ x \in \mathbb{R}^d : U(x) \leq H + \frac{1}{N} \right\}$$

containing $\mathcal{H}$ is contained in $\mathcal{U}$

Proof. Let $\mathcal{R}_n$ be the connected component of

$$\left\{ x \in \mathbb{R}^d : U(x) \leq H + \frac{1}{n} \right\}$$

containing $\mathcal{H}$ and let $\mathcal{R} := \bigcap_{n=1}^{\infty} \mathcal{R}_n$. Since $\mathcal{H} \subset \mathcal{R}_n$ for all $n \geq 1$, $\mathcal{H} \subset \mathcal{R}$. On the other hand, as $(\mathcal{R}_n)_{n \geq 1}$ is a decreasing sequence of compact connected sets, by Lemma A.12, $\mathcal{R}$ is connected. Since $\mathcal{R} \cap \mathcal{H} \neq \varnothing$, $\mathcal{R} \subset \{x \in \mathbb{R}^d : U(x) \leq H\}$ and $\mathcal{H}$ is a connected component of $\{x \in \mathbb{R}^d : U(x) \leq H\}$, by Lemma A.9, $\mathcal{R} \subset \mathcal{H}$. By the previous two estimates, $\mathcal{H} = \mathcal{R}$.

As $\mathcal{R} \subset \mathcal{U}$ by Lemma A.13, there exists $N \in \mathbb{N}$ such that $\mathcal{R}_N \subset \mathcal{U}$. $\qquad \square$

Lemma A.15. *Fix $H \in \mathbb{R}$. Let $\mathcal{H}$ be a connected component of $\{x \in \mathbb{R}^d : U(x) < H\}$. Then, $\overline{\mathcal{H}}$ is path connected.*

Proof. As $\mathcal{H}$ is open and connected, it is path connected. It remains to show that the boundary $\partial \mathcal{H}$ is path connected to $\mathcal{H}$. Fix $x_0 \in \partial \mathcal{H}$. By Lemma A.10, $U(x_0) = H$. Assume that x_0 is not a critical point of U. By Lemma A.6, there exists $r > 0$ such that the manifold $\{x \in \mathbb{R}^d : U(x) = U(x_0)\}$ divides $B(x_0, r)$ into two parts:

$$B(x_0, r) \cap \{x \in \mathbb{R}^d : U(x) < U(x_0)\},$$
$$B(x_0, r) \cap \{x \in \mathbb{R}^d : U(x) > U(x_0)\}.$$

Since $x_0 \in \partial \mathcal{H}$, $B(x_0, r) \cap \mathcal{H} \neq \varnothing$ so that $B(x_0, r) \cap \{x \in \mathbb{R}^d : U(x) < U(x_0)\} \neq \varnothing$. By Lemma A.9, $B(x_0, r) \cap \{x \in \mathbb{R}^d : U(x) < U(x_0)\} \subset \mathcal{H}$. By Lemma A.6, there is a path $z : [0, 1] \to B(x_0, r)$ such that $z(0) = x_0$ and $z((0, 1]) \subset B(x_0, r) \cap \{x \in \mathbb{R}^d : U(x) < U(x_0)\} \subset \mathcal{H}$. Hence, x_0 is path-connected to $\mathcal{H}$.

Suppose that x_0 is a critical point. By Lemma A.10, x_0 is not a local minimum. By the Hartman–Grobman Theorem, there is $T > 0$ and a continuous path $z : [0, T] \to \mathbb{R}^d$ in the unstable manifold of x_0 such that $z(0) = x_0$ and $z((0, T]) \subset \mathcal{H}$. This completes the proof of the lemma. $\qquad \square$

Lemma A.16. *Let $\mathcal{H}$ be a connected component of $\{x \in \mathbb{R}^d : U(x) \leq H\}$, and $\mathcal{H}^o$ the interior of $\mathcal{H}$. Denote by $\mathcal{W}_i$, $i \geq 1$, the connected components of $\mathcal{H}^o$. Then, the number of connected components is finite. Moreover,*

(1) Let $\mathcal{W}'_j$ be a connected component of $\{x \in \mathbb{R}^d : U(x) < H\}$ which intersects with $\mathcal{W}_j$. Then, $\overline{\mathcal{W}_j} = \overline{\mathcal{W}'_j}$.

(2) $\mathcal{H}$ is path connected. In particular, for each i, there is j such that $\overline{\mathcal{W}_i} \cap \overline{\mathcal{W}_j} \neq \varnothing$.

Proof. Consider the open set $\mathcal{W}_1$. Since critical points of U are not isolated points, it is not possible to have $U(x) = H$ for all $x \in \mathcal{W}_1$. Hence, $\mathcal{W}'_1$ well defined and $\mathcal{W}_1 \cap \mathcal{W}'_1 \neq \varnothing$. Let $x_0 \in \mathcal{W}_1 \cap \mathcal{W}'_1$.

Claim 1: Let $x_1 \in \mathcal{W}_1$ such that $U(x_1) = H$. Then, x_1 is a local maximum.

Fix $x_1 \in \mathcal{W}_1$ such that $U(x_1) = H$. Since $\mathcal{W}_1$ is open, there exists $r_1 > 0$ such that $B(x_1, r_1) \subset \mathcal{W}_1$. Let r_1 be small enough so that there is no critical point in $B(x_1, r_1) \backslash \{x_1\}$. Let $y \in B(x_1, r_1)$ such that $U(y) = H$. Since $U(x) \leq H$ for all $x \in B(x_1, r_1)$, y is a critical point (since $\nabla U(y) = 0$). Hence, $U(x) < H$ for all $x \in B(x_1, r_1) \setminus \{x_1\}$. Therefore, x_1 is a local maximum, as claimed.

Let

$$\widehat{\mathcal{W}}_1 := \{x \in \mathcal{W}_1 : x \text{ is not a local maximum}\} \subset \mathcal{W}_1 .$$

Since there are finitely many local maximum in $\mathcal{W}_1$, $\widehat{\mathcal{W}}_1$ is open and connected. By Claim 1, $U(x) < H$ for all $x \in \widehat{\mathcal{W}}_1$.

By construction, $x_0 \in \widehat{\mathcal{W}}_1$.

Claim 2: $\widehat{\mathcal{W}}_1 \subset \mathcal{W}'_1 \subset \mathcal{W}_1$.

Since $\mathcal{W}'_1$ is a connected component of $\{x \in \mathbb{R}^d : U(x) < H\}$ intersecting with $\widehat{\mathcal{W}}_1$ and since $U(x) < H$ for all $x \in \widehat{\mathcal{W}}_1$, by Lemma A.9, $\widehat{\mathcal{W}}_1 \subset \mathcal{W}'_1$. Let $x_2 \in \mathcal{W}'_1$. Since $x_0 \in \mathcal{W}'_1$, there is a continuous path $z \colon [0, 1] \to \mathcal{W}'_1$ in $\mathcal{W}'_1$ from x_0 to x_2 such that $U(z(t)) < H$ for all $0 \leq t \leq 1$. Since $\mathcal{H}$ is a connected component of $\{x \in \mathbb{R}^d : U(x) \leq H\}$ containing x_0, this path is contained in $\mathcal{H}$: $z(t) \in \mathcal{H}$ for all $0 \leq t \leq 1$. As $U(z(t)) < H$, by Lemma A.10, $z(t) \in \mathcal{H}^o$ for all $0 \leq t \leq 1$. As $\mathcal{W}_1$ is a connected component of $\mathcal{H}^o$ and $x_0 \in \mathcal{W}_1$, $x_2 \in \mathcal{W}_1$, as claimed.

By definition, the set $\mathcal{W}'_1$ contains a local minimum. By Claim 2. so does $\mathcal{W}_1$. Since the connected components are disjoints, each one contains at least one local minimum of U and there are a finite number of critical points, the set $\mathcal{H}^o$ has a finite number of connected components. This is the first assertion of the lemma.

By Claim 2, $\overline{\widehat{\mathcal{W}}_1} \subset \overline{\mathcal{W}'_1} \subset \overline{\mathcal{W}_1}$. Since local maxima $y \in \mathcal{W}_1$ are accumulation points, $\overline{\widehat{\mathcal{W}}_1} = \overline{\mathcal{W}_1}$ so that $\overline{\mathcal{W}'_1} = \overline{\mathcal{W}_1}$. This proves the second assertion of the lemma.

Denote by n the number of connected components of $\mathcal{H}^o$, so that $\mathcal{H} = \overline{\mathcal{H}^o} = \overline{\bigcup_{i=1}^n \mathcal{W}_i} = \bigcup_{i=1}^n \overline{\mathcal{W}_i}$. By Lemma A.15, $\overline{\mathcal{W}_i} = \overline{\mathcal{W}'_1}$ is path-connected.

Claim 3: for all $i \neq j \in \{1, \ldots, n\}$, there exists $i = i_0, \ldots, i_k = j$ such that

$$\overline{\mathcal{W}_{i_m}} \cap \overline{\mathcal{W}_{i_{m+1}}} \neq \varnothing, \quad 0 \leq m < k. \tag{A.7}$$

Suppose this property does not hold. Then, there exists $i \neq j \in \{1, \ldots, n\}$ which are not connected in the sense (A.7). Let A be the set of indices in $\{1, \ldots, n\}$ which are

connected to i in the sense (A.7). The sets $\cup_{k \in A} \overline{\mathcal{W}_k}$, $\cup_{k \notin A} \overline{\mathcal{W}_k}$ are compact, disjoint and non-empty. Thus, there exist disjoint opens sets $\mathcal{U}$, $\mathcal{V}$ such that $\cup_{k \in A} \overline{\mathcal{W}_k} \subset \mathcal{U}$, $\cup_{k \notin A} \overline{\mathcal{W}_k} \subset \mathcal{V}$. This contradicts the fact that $\mathcal{H} = \bigcup_{i=1}^{n} \overline{\mathcal{W}_i}$ is connected, and proves Claim 3.

Since each set $\overline{\mathcal{W}_i}$ is path-connected, by property (A.7), the set $\mathcal{H}$ is also path-connected. $\qquad\square$

Lemma A.17. *The connected component of the set $\{x \in \mathbb{R}^d : U(x) \le \Theta(m', m'')\}$ containing m' also contains m''.*

Proof. Let $\mathcal{H}$ be the connected component of the set $\{x \in \mathbb{R}^d : U(x) \le \Theta(m', m'')\}$ containing m'. Suppose, by contradiction, that $m'' \notin \mathcal{H}$. Since $\mathcal{H}$ and $\{m''\}$ are compact sets, there is an open set $\mathcal{U}$ such that $\mathcal{H} \subset \mathcal{U}$ and $m'' \notin \mathcal{U}$. By Lemma A.14, there is $n \in \mathbb{N}$ such that the connected component of

$$\{x \in \mathbb{R}^d : U(x) \le \Theta(m', m'') + \frac{1}{n}\}$$

containing $\mathcal{H}$ is contained in $\mathcal{U}$. This connected component does not contain m''. Thus, by Lemma A.11, $\Theta(m', m'') \ge \Theta(m', m'') + \frac{1}{n}$, which is a contradiction. $\square$

Lemma A.18. *Let m, $m' \in \mathcal{M}_0$ be two different local minima. Then,*

$$U(m), \ U(m') < \Theta(m, m') .$$

Proof. We only prove for m because $\Theta(\cdot, \cdot)$ is symmetric. Since m is a local minimum, there exists $\eta > 0$ such that m is a unique local minimum of a connected component of $\{x \in \mathbb{R}^d : U(x) < U(m) + 2\eta\}$ containing m. Therefore, by Lemma A.11, for all continuous path z connecting local minimum m and any other local minimum, we have

$$\max_{t \in [0,1]} U(z(t)) > U(m) + \eta$$

which implies

$$\Theta(m, m') \ge U(m) + \eta > U(m) .$$

$\qquad\square$

Proof of Lemma A.2. Fix $m \in \mathcal{M}_0$. Let $\mathcal{H}$ be the connected component of the set $\{x \in \mathbb{R}^d : U(x) \le U(m) + \Gamma(m)\}$ containing m. By definition, there exists $m' \in \mathcal{M}_0$ such that

$$\Theta(m, m') = \min_{m'' \in \mathcal{M}_0 \setminus \{m\}} \Theta(m, m'') = U(m) + \Gamma(m).$$

By Lemma A.17, $m' \in \mathcal{H}$.

As in Lemma A.16, denote by $\mathcal{H}^o$ the interior of $\mathcal{H}$. Let $\mathcal{W}_j$, $1 \le i \le n$, the open connected components of $\mathcal{H}^o$. Assume that $m \in \mathcal{W}_1$.

We assert that m is the unique local minimum in $\mathcal{W}_1$. Indeed, as in Lemma A.16, let $\mathcal{W}_1'$ be a connected component of $\{x \in \mathbb{R}^d : U(x) < U(m) + \Gamma(m)\}$ which

intersects with $\mathcal{W}_1$. By (A.1), $\mathcal{W}_1'$ contains one and only one local minimum of U. On the other hand, by Claim 2 in Lemma A.16, $\mathcal{W}_1' \subset \mathcal{W}_1$ and all elements of $\mathcal{W}_1 \backslash \mathcal{W}_1'$ are local maxima. This proves the assertion.

By Lemma A.18, $U(\boldsymbol{m}') < \Theta(\boldsymbol{m}, \boldsymbol{m}')$. As $\boldsymbol{m}' \in \mathcal{H}$, by Lemma A.10, $\boldsymbol{m}' \in \mathcal{H}^o$ so that $n \geq 2$.

By Lemma A.16-(2), there is $1 < k \leq n$ such that $\overline{\mathcal{W}_1} \cap \overline{\mathcal{W}_k} \neq \varnothing$. Let $\mathcal{W}, \mathcal{W}'$ be connected components of $\{\boldsymbol{x} \in \mathbb{R}^d : U(\boldsymbol{x}) < \Theta(\boldsymbol{m}, \boldsymbol{m}')\}$ intersecting with $\mathcal{W}_1$, $\mathcal{W}_k$, respectively. By Lemma A.16, $\overline{\mathcal{W}} = \overline{\mathcal{W}_1}$ and $\overline{\mathcal{W}'} = \overline{\mathcal{W}_k}$ so that $\overline{\mathcal{W}} \cap \overline{\mathcal{W}'} \neq \varnothing$. □

Appendix B. Extension of the Vector Field b

Fix $\boldsymbol{m} \in \mathcal{M}_0$. In this subsection, we define a new vector field $\boldsymbol{b}_0 : \mathbb{R}^d \to \mathbb{R}^d$ which coincides with $\boldsymbol{b}$ in a neighborhood of $\boldsymbol{m}$ and satisfies the hypotheses of Sect. 3. Assume that $\boldsymbol{m} = \boldsymbol{0}$, and let

$$\mathbb{H} = (\nabla^2 U)(\boldsymbol{0}), \quad \mathbb{L} = (D\boldsymbol{\ell})(\boldsymbol{0}).$$

By Taylor expansion, for $\boldsymbol{x} \simeq \boldsymbol{0}$,

$$- \boldsymbol{b}(\boldsymbol{x}) \cdot \mathbb{H}\boldsymbol{x} = \left[(\mathbb{H} + \mathbb{L})\boldsymbol{x} + O(|\boldsymbol{x}|^2) \right] \cdot \mathbb{H}\boldsymbol{x} = |\mathbb{H}\boldsymbol{x}|^2 + O(|\boldsymbol{x}|^3),$$

where the second equality comes from the fact that $\mathbb{H}\mathbb{L}$ is skew-symmetric. Thus, there exists $r_5 > 0$ such that

$$= - \boldsymbol{b}(\boldsymbol{x}) \cdot \mathbb{H}\boldsymbol{x} \geq \frac{1}{2} |\mathbb{H}\boldsymbol{x}|^2 \quad \text{for all } \boldsymbol{x} \in B(\boldsymbol{0}, 2r_5). \tag{B.1}$$

If needed, modify the definition of $r_5 > 0$ for

$$|\mathbb{K}_{\boldsymbol{x}} \boldsymbol{y}| \leq \frac{1}{2} |\mathbb{H}\boldsymbol{y}| \quad \text{for all } \boldsymbol{x} \in B(\boldsymbol{0}, 2r_5), \quad \boldsymbol{y} \in \mathbb{R}^d. \tag{B.2}$$

where $\mathbb{K}_{\boldsymbol{x}} = (\nabla^2 U + D\boldsymbol{\ell})(\boldsymbol{x}) - (\mathbb{H} + \mathbb{L})$.

For $\boldsymbol{x} \notin B(\boldsymbol{0}, r_5)$, let

$$\boldsymbol{r}(\boldsymbol{x}) = \frac{r_5}{|\boldsymbol{x}|}\boldsymbol{x} \in \partial B(\boldsymbol{0}, r_5)$$

and let $\boldsymbol{b}_0 : \mathbb{R}^d \to \mathbb{R}^d$ be given by

$$\boldsymbol{b}_0(\boldsymbol{x}) = \begin{cases} \boldsymbol{b}(\boldsymbol{x}), & \boldsymbol{x} \in B(\boldsymbol{0}, r_5) \\ \boldsymbol{b}(\boldsymbol{r}(\boldsymbol{x})) + (D\boldsymbol{b})(\boldsymbol{r}(\boldsymbol{x}))(\boldsymbol{x} - \boldsymbol{r}(\boldsymbol{x})), & \boldsymbol{x} \in B(\boldsymbol{0}, r_5)^c. \end{cases} \tag{B.3}$$

The main result of this section reads as follows.

Proposition B.1. *The vector field $\boldsymbol{b}_0$ fullfils all conditions of Sect. 3. Condition (2) holds for $r_3 = r_5$.*

The proof relies on two lemmata.

Lemma B.2. *The vector field $\boldsymbol{b}_0$ belongs to $C^1(\mathbb{R}^d, \mathbb{R}^d)$. Moreover, there exists a finite constant C_1 such that*

$$|\boldsymbol{b}_0(\boldsymbol{x})| \leq C_1|\boldsymbol{x}| \quad and \quad \|D\boldsymbol{b}_0(\boldsymbol{x})\| \leq C_1|\boldsymbol{x}| \tag{B.4}$$

for all $\boldsymbol{x} \in B(\boldsymbol{0}, r_5)^c$.

Proof. By a straightforward computation, for $|\boldsymbol{x}| > r_5$,

$$(\partial_{x_k}\boldsymbol{b}_0)(\boldsymbol{x}) = \partial_{x_k}\big\{ [D\boldsymbol{b}(\boldsymbol{r}(\boldsymbol{x}))] (\boldsymbol{x} - \boldsymbol{r}(\boldsymbol{x})) \big\} + \partial_{x_k}\{ \boldsymbol{b}(\boldsymbol{r}(\boldsymbol{x})) \}. \tag{B.5}$$

Since

$$\partial_{x_k}\boldsymbol{r}(\boldsymbol{x}) = r_5 \frac{\boldsymbol{e}_k}{|\boldsymbol{x}|} - \frac{x_k}{|\boldsymbol{x}|^3} \boldsymbol{x},$$

the matrix $\partial_{x_k} [D\boldsymbol{b}(\boldsymbol{r}(\boldsymbol{x}))]$ is uniformly bounded on $B(\boldsymbol{0}, r_5)^c$.

Since $\boldsymbol{x} \to \boldsymbol{r}(\boldsymbol{x})$ as $\boldsymbol{x}$ approaches $\partial B(\boldsymbol{0}, r_5)$, the boundedness of $\partial_{x_k} [D\boldsymbol{b}(\boldsymbol{r}(\boldsymbol{x}))]$ yields that $\partial_{x_k}\boldsymbol{b}_0(\boldsymbol{x}) \to \partial_{x_k}\boldsymbol{b}(\boldsymbol{x})$ as $\boldsymbol{x}$ approaches to $\partial B(\boldsymbol{0}, r_5)$. This proves that $\boldsymbol{b}_0 \in C^1(\mathbb{R}^d, \mathbb{R}^d)$. The first assertion of (B.4) follows from the definition of $\boldsymbol{b}_0$. The second one from (B.5) and the boundedness of $\partial_k [D\boldsymbol{b}(\boldsymbol{r}(\boldsymbol{x}))]$ on $B(0, r_5)^c$. □

Lemma B.3. *For all $\boldsymbol{x} \in \mathbb{R}^d$,*

$$- \boldsymbol{b}_0(\boldsymbol{x}) \cdot \mathbb{H}\boldsymbol{x} \geq \frac{1}{2}|\mathbb{H}\boldsymbol{x}|^2$$

Proof. By (B.1) the condition is satisfied for $\boldsymbol{x} \in B(\boldsymbol{0}, 2r_5)$. Fix $\boldsymbol{x} \notin B(\boldsymbol{0}, r_5)$ so that

$$- \boldsymbol{b}_0(\boldsymbol{x}) \cdot \mathbb{H}\boldsymbol{x} = - \boldsymbol{b}(\boldsymbol{r}(\boldsymbol{x})) \cdot \mathbb{H}\boldsymbol{x} - (D\boldsymbol{b})(\boldsymbol{r}(\boldsymbol{x}))(\boldsymbol{x} - \boldsymbol{r}(\boldsymbol{x})) \cdot \mathbb{H}\boldsymbol{x} \tag{B.6}$$

Since $\boldsymbol{x} = (|\boldsymbol{x}|/r_5) \boldsymbol{r}(\boldsymbol{x})$ and since $\boldsymbol{r}(\boldsymbol{x}) \in B(\boldsymbol{0}, 2r_5)$, by (B.1), the first term on the right-hand side can be estimated by

$$- \frac{|\boldsymbol{x}|}{r_5} \boldsymbol{b}(\boldsymbol{r}(\boldsymbol{x})) \cdot \mathbb{H}\boldsymbol{r}(\boldsymbol{x}) \geq \frac{|\boldsymbol{x}|}{2r_5} | \mathbb{H}\boldsymbol{r}(\boldsymbol{x}) |^2 = \frac{r_5}{2|\boldsymbol{x}|} |\mathbb{H}\boldsymbol{x}|^2.$$

For the second term, write

$$- (D\boldsymbol{b})(\boldsymbol{r}(\boldsymbol{x})) = \mathbb{H} + \mathbb{L} + \mathbb{K}_{\boldsymbol{x}} \quad and \quad \boldsymbol{x} - \boldsymbol{r}(\boldsymbol{x}) = \Big(1 - \frac{r_5}{|\boldsymbol{x}|} \Big) \boldsymbol{x}.$$

Since $\mathbb{H}\mathbb{L}$ is skew-symmetry, the second term of (B.6) is equal to

$$\Big(1 - \frac{r_5}{|\boldsymbol{x}|} \Big) (\mathbb{H} + \mathbb{L} + \mathbb{K}_{\boldsymbol{x}}) \boldsymbol{x} \cdot \mathbb{H}\boldsymbol{x}$$

$$= \Big(1 - \frac{r_5}{|\boldsymbol{x}|} \Big) \Big(|\mathbb{H}\boldsymbol{x}|^2 + \mathbb{K}_{\boldsymbol{x}}\boldsymbol{x} \cdot \mathbb{H}\boldsymbol{x} \Big) \geq \frac{1}{2} \Big(1 - \frac{r_5}{|\boldsymbol{x}|} \Big) |\mathbb{H}\boldsymbol{x}|^2 .$$

The last inequality comes from (B.2). Adding the previous estimates completes the proof of the lemma. □

Proof of Proposition B.1. To check the first condition, suppose that $\boldsymbol{b}_0(\boldsymbol{x}) = 0$ for some $\boldsymbol{x} \in \mathbb{R}^d$. Lemma B.3 implies that $\boldsymbol{x} = \boldsymbol{0}$. Thus $\boldsymbol{0}$ is the only equilibrium of the dynamical system (3.2). Since the behavior of this ODE near $\boldsymbol{0}$ is identical to that of $\boldsymbol{x}(\cdot)$, the origin is a stable equilibrium. Condition (2) in Sect. 3 for $r_3 = r_5$ follows from the definition of $\boldsymbol{b}_0$. The third and fourth conditions have been derived in Lemmata B.2 and B.3, respectively. □

Appendix C. Potential Theory

In sake of completeness we introduce in this section the capacity between sets. Fix two disjoint non-empty bounded domains $\mathcal{A}$ and $\mathcal{B}$ of $\mathbb{R}^d$ with $C^{2,\alpha}$-boundaries for some $\alpha \in (0, 1)$. Assume that the perimeters of $\mathcal{A}$, $\mathcal{B}$ are finite and that the distance between the sets is positive. Let $\Omega = (\overline{\mathcal{A}} \cup \overline{\mathcal{B}})^c$ so that $\partial\Omega = \partial\mathcal{A} \cup \partial\mathcal{B}$. The equilibrium potentials $h^\varepsilon_{\mathcal{A}, \mathcal{B}}$ between $\mathcal{A}$ and $\mathcal{B}$ with respect to the processes $x_\varepsilon(\cdot)$ is given by

$$h^\varepsilon_{\mathcal{A}, \mathcal{B}}(x) = \mathbb{P}^\varepsilon_x[\tau_\mathcal{A} < \tau_\mathcal{B}], \quad x \in \mathbb{R}^d,$$

and the capacity by

$$\mathrm{cap}_\varepsilon(\mathcal{A}, \mathcal{B}) = \varepsilon \int_\Omega |\nabla h^\varepsilon_{\mathcal{A}, \mathcal{B}}|^2 \, d\mu_\varepsilon.$$

We refer to [22] for equivalent formulations and properties of the capacity.

Appendix D. Analysis of a Linear ODE

In this section, we prove Lemma 5.3. To simplify notation, we fix $c \in \mathcal{Y}_0$, assumed to be equal to 0, $c = 0$, and write $\mathbb{A} = -(\mathbb{H}^c + \mathbb{L}^c)$. By Lemma 5.1, the matrix $\mathbb{A}$ is invertible and does not have a pure imaginary eigenvalue. All the results given in this section holds for such a matrix $\mathbb{A}$.

Real Jordan Canonical Form

Suppose that a matrix $\mathbb{K}$ can be written as a block matrix of the form

$$\mathbb{K} = \begin{bmatrix} \mathbb{K}_1 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \ddots & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{K}_n \end{bmatrix},$$

where $\mathbb{K}_1, \ldots, \mathbb{K}_n$ are matrices of possibly different sizes and $\mathbb{O}$ denotes the zero matrix of suitable size. We represent such a matrix as $\mathbb{K} = \mathrm{diag}(\mathbb{K}_1, \ldots, \mathbb{K}_n)$. We start by a review of a real Jordan canonical form of $\mathbb{A}$. By [35, Theorem 2.5], there exists an invertible matrix $\mathbb{U}$ such that

$$\mathbb{A} = \mathbb{U} \mathbb{J} \mathbb{U}^{-1} \tag{D.1}$$

where $\mathbb{J}$ is of the form

$$\mathbb{J} = \mathrm{diag}(\mathbb{E}^-_1, \ldots, \mathbb{E}^-_{u_1}, \mathbb{F}^-_1, \ldots, \mathbb{F}^-_{u_2}, \mathbb{E}^+_1, \ldots, \mathbb{E}^+_{s_1}, \mathbb{F}^+_1, \ldots, \mathbb{F}^+_{s_2})$$

$$\mathbb{E}^\pm_k = \begin{bmatrix} \lambda^\pm_k & 1 & 0 \\ 0 & \ddots & 1 \\ 0 & 0 & \lambda^\pm_k \end{bmatrix}, \quad \mathbb{F}^\pm_k = \begin{bmatrix} \mathbb{B}^\pm_k & \mathbb{I}_2 & 0 \\ 0 & \ddots & \mathbb{I}_2 \\ 0 & 0 & \mathbb{B}^\pm_k \end{bmatrix},$$

and

$$\mathbb{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbb{B}_k^{\pm} = \begin{bmatrix} \alpha_k^{\pm} & \beta_k^{\pm} \\ -\beta_k^{\pm} & \alpha_k^{\pm} \end{bmatrix}.$$

In this formula, λ_k^+ and α_k^+ are positive while λ_k^- and α_k^- are negative real numbers. The eigenvalues of $\mathbb{A}$ are $\lambda_k^{\pm}$ and $\alpha_k^{\pm} + i\beta_k^{\pm}$. Thus, by Lemma 5.1, $\lambda_k^{\pm}$ and $\alpha_k^{\pm}$ cannot be 0. (Note that the real numbers $\beta_k^{\pm}$ are also different from 0 because the eigenvalues $\alpha_k^{\pm} + i\beta_k^{\pm}$ are complex numbers, but this will not be used below). By (D.1)

$$e^{t\mathbb{A}} = \mathbb{U}e^{t\mathbb{J}}\mathbb{U}^{-1}$$

where

$$e^{t\mathbb{J}} = \mathrm{diag}(e^{t\mathbb{E}_1^-}, \ldots, e^{t\mathbb{E}_{u_1}^-}, e^{t\mathbb{F}_1^-}, \ldots, e^{t\mathbb{F}_{u_2}^-}, e^{t\mathbb{E}_1^+}, \ldots, e^{t\mathbb{E}_{s_1}^+}, e^{t\mathbb{F}_1^+}, \ldots, e^{t\mathbb{F}_{s_2}^+}).$$
(D.2)

Suppose that $\mathbb{E}_k^{\pm}$ is a $j \times j$ matrix. An elementary computation yields that

$$e^{t\mathbb{E}_k^{\pm}} = e^{t\lambda_k^{\pm}} \begin{bmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{j-1}}{(j-1)!} & \frac{t^j}{j!} \\ 0 & 1 & t & \cdots & \frac{t^{j-2}}{(j-2)!} & \frac{t^{j-1}}{(j-1)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & t \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$
(D.3)

Similarly, if $\mathbb{F}_k^{\pm}$ is a $2j \times 2j$ matrix,

$$e^{t\mathbb{F}_k^{\pm}} = e^{t\alpha_k^{\pm}} \begin{bmatrix} \mathbb{S} & t\mathbb{S} & \frac{t^2}{2}\mathbb{S} & \cdots & \frac{t^{j-1}}{(j-1)!}\mathbb{S} & \frac{t^j}{j!}\mathbb{S} \\ \mathbb{O} & \mathbb{S} & t\mathbb{S} & \cdots & \frac{t^{j-2}}{(j-2)!}\mathbb{S} & \frac{t^{j-1}}{(j-1)!}\mathbb{S} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \cdots & \mathbb{S} & t\mathbb{S} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} & \cdots & \mathbb{O} & \mathbb{S} \end{bmatrix},$$
(D.4)

where $\mathbb{O}$ denotes the 2×2 zero matrix and

$$\mathbb{S} = \begin{bmatrix} \cos(t\beta_k^{\pm}) & \sin(t\beta_k^{\pm}) \\ -\sin(t\beta_k^{\pm}) & \cos(t\beta_k^{\pm}) \end{bmatrix}.$$

Stable and Unstable Manifolds

Recall from (5.1) that we represent by $\upsilon_{L,\boldsymbol{x}}(t)$ the solution of the linear ODE (5.1). With the notation of this section, it can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\upsilon_{L,\boldsymbol{x}}(t) = \mathbb{A}\,\upsilon_{L,\boldsymbol{x}}(t), \quad \upsilon_{L,\boldsymbol{x}}(t) = e^{t\mathbb{A}}\boldsymbol{x} = \mathbb{U}e^{t\mathbb{J}}\mathbb{U}^{-1}\boldsymbol{x}. \tag{D.5}$$

Recall that $\mathcal{M}_{L,s}$, $\mathcal{M}_{L,u}$ represent the stable, unstable manifold of $\boldsymbol{c} = \boldsymbol{0}$ for the linear ODE (5.1). By (D.5),

$$\mathcal{M}_{L,u} : = \left\{\boldsymbol{y} \in \mathbb{R}^d : \lim_{t \to -\infty} e^{t\mathbb{J}}\mathbb{U}^{-1}\boldsymbol{y} = \boldsymbol{0}\right\} \quad \mathcal{M}_{L,s} :$$
$$= \left\{\boldsymbol{y} \in \mathbb{R}^d : \lim_{t \to +\infty} e^{t\mathbb{J}}\mathbb{U}^{-1}\boldsymbol{y} = \boldsymbol{0}\right\}.$$

Denote by $m \in \mathbb{N}$ the size of the matrix $\mathrm{diag}(e^{t\mathbb{E}_1^-}, \ldots, e^{t\mathbb{E}_{u_1}^-}, e^{t\mathbb{F}_1^-}, \ldots, e^{t\mathbb{F}_{u_2}^-})$, and by $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_d\}$ the column vectors of $\mathbb{U}$, $(\boldsymbol{u}_i = \mathbb{U}\boldsymbol{e}_i$, where $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d\}$ stands for the canonical basis of $\mathbb{R}^d$). By (D.2), (D.3), and (D.4),

$$\mathcal{M}_{L,u} = \langle\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\rangle \quad \text{and} \quad \mathcal{M}_{L,s} = \langle\boldsymbol{u}_{m+1}, \ldots, \boldsymbol{u}_d\rangle, \tag{D.6}$$

where $\langle S\rangle$ denotes the vector space spanned by S. The following lemma is a direct consequence of the discussion above.

Lemma D.1. *There exists $C_0 > 0$ such that for all $\boldsymbol{y} \in \mathcal{M}_{L,s}$ and $t \geq 0$,*

$$\|\upsilon_{L,\boldsymbol{y}}(t)\| \leq C_0\|\boldsymbol{y}\|.$$

Proof. Write $\gamma = \min\{|\lambda_1^-|, \ldots, |\lambda_{u_1}^-|, |\alpha_1^-|, \ldots, |\alpha_{u_2}^-|\} > 0$. Then, by (D.2), (D.3), (D.4), (D.5), and (D.6), it is clear that there exists a polynomial $P(t)$ depending only on d such that

$$\|\upsilon_{L,\boldsymbol{y}}(t)\| \leq e^{-\gamma t}P(t)\|\boldsymbol{y}\|.$$

The conclusion of lemma follows immediately. $\qquad\qquad\square$

By (D.6), $\mathbb{R}^d = \mathcal{M}_{L,u} \oplus \mathcal{M}_{L,s}$. Hence, for each $\boldsymbol{y} \in \mathbb{R}^d$, there exists a unique decomposition

$$\boldsymbol{y} = \boldsymbol{v}^u(\boldsymbol{y}) + \boldsymbol{v}^s(\boldsymbol{y}) \tag{D.7}$$

such that $\boldsymbol{v}^u(\boldsymbol{y}) \in \mathcal{M}_{L,u}$ and $\boldsymbol{v}^s(\boldsymbol{y}) \in \mathcal{M}_{L,s}$. Next lemma provides the basic property of this decomposition.

Lemma D.2. *There is $c_0 < \infty$ such that*

$$\|\boldsymbol{v}^s(\boldsymbol{y})\| \leq c_0\|\boldsymbol{y}\| \text{ for all } \boldsymbol{y} \in \mathbb{R}^d.$$

The proof is based on the following elementary result.

Lemma D.3. *Let V and W be subspaces of $\mathbb{R}^d$ such that $V \cap W = \{\mathbf{0}\}$. Then, there exists $\zeta = \zeta(V, W) > 0$ such that*

$$\sup_{v \in V \setminus \{\mathbf{0}\}, \, w \in W \setminus \{\mathbf{0}\}} \frac{|\langle v, w \rangle|}{\|v\| \|w\|} = 1 - \zeta$$

where sup is taken over all non-zero vectors.

Proof. Let us define $F : (V \setminus \{\mathbf{0}\}) \times (W \setminus \{\mathbf{0}\}) \to \mathbb{R}$ as

$$F(v, w) = \frac{|\langle v, w \rangle|}{\|v\| \|w\|}.$$

Since $F(cv, c'w) = F(v, w)$ for all $c, c' \neq 0$, we have

$$\sup_{v \in V \setminus \{\mathbf{0}\}, \, w \in W \setminus \{\mathbf{0}\}} F(v, w) = \sup_{v \in V, \, w \in W: \|v\| = \|w\| = 1} F(v, w).$$

Since the set $S_0 = \{(v, w) : v \in V, w \in W : \|v\| = \|w\| = 1\}$ is compact and $F(\cdot, \cdot)$ is continuous, the function F achieve the maximum at certain $(v^*, w^*) \in S_0$. Then

$$\sup_{v \in V \setminus \{\mathbf{0}\}, \, w \in W \setminus \{\mathbf{0}\}} F(v, w) = F(v^*, w^*).$$

Note that $F(v^*, w^*) < 1$ by the Cauchy–Schwarz inequality (the equality cannot hold because of the assumption $V \cap W = \{\mathbf{0}\}$). This completes the proof. □

Proof of Lemma D.2. Since $\mathcal{M}_{L,u} \cap \mathcal{M}_{L,s} = \{\mathbf{0}\}$, by Lemma D.3, there exists a constant $c > 1$ such that, for all $y \in \mathbb{R}^d$,

$$
\begin{aligned}
|\langle v^u(y), v^s(y) \rangle| &\leq \sqrt{(c-1)/c} \|v^u(y)\| \|v^s(y)\| \\
&\leq \frac{1}{2} \|v^u(y)\|^2 + \frac{c-1}{2c} \|v^s(y)\|^2,
\end{aligned}
$$

where we applied Young's inequality in the last step. Therefore,

$$-2c \langle v^u(y), v^s(y) \rangle \leq c \|v^u(y)\|^2 + (c-1) \|v^s(y)\|^2.$$

Reorganizing, we obtain

$$\|v^s(y)\|^2 \leq c \left(\|v^u(y)\|^2 + 2 \langle v^u(y), v^s(y) \rangle + \|v^s(y)\|^2 \right) = c \|y\|^2.$$

This completes the proof. □

Proof of Lemma 5.3. Recall the constant C_0 and c_0 from Lemmata D.1 and D.2, respectively, and define $r > 0$ as

$$r = \frac{a}{3 C_0 c_0}. \tag{D.8}$$

Suppose that $y \in \mathcal{B}(\mathbf{0}, r)$. As in (D.7), decompose $y \in \mathbb{S}^d$ into

$$y = v^u(y) + v^s(y)$$

so that

$$\upsilon_{L,\boldsymbol{y}}(t) = \upsilon_{L,\boldsymbol{v}^u(\boldsymbol{y})}(t) + \upsilon_{L,\boldsymbol{v}^s(\boldsymbol{y})}(t).$$

Note that $\upsilon_{L,\boldsymbol{v}^u(\boldsymbol{y})}(t) \in \mathcal{M}_{L,u}$ and $\upsilon_{L,\boldsymbol{v}^s(\boldsymbol{y})}(t) \in \mathcal{M}_{L,s}$ for all $t \geq 0$ since $\mathcal{M}_{L,u}$ and $\mathcal{M}_{L,s}$ are invariant under the dynamical system (5.1). Recall the definition (5.2) of $t_L(\cdot)$ and write

$$\boldsymbol{w}^u(\boldsymbol{y}) := \upsilon_{L,\boldsymbol{v}^u(\boldsymbol{y})}(t_L(\boldsymbol{y})) \in \mathcal{M}_{L,u} \quad \text{and} \quad \boldsymbol{w}^s(\boldsymbol{y}) := \upsilon_{L,\boldsymbol{v}^s(\boldsymbol{y})}(t_L(\boldsymbol{y})) \in \mathcal{M}_{L,s},$$

so that

$$\boldsymbol{e}_L(\boldsymbol{y}) = \boldsymbol{w}^u(\boldsymbol{y}) + \boldsymbol{w}^s(\boldsymbol{y}) \tag{D.9}$$

By (5.2), $\|\boldsymbol{e}_L(\boldsymbol{y})\| = r_1$. Moreover, by Lemmata D.1, D.2, and by the definition (D.8) of r,

$$\|\boldsymbol{w}^s(\boldsymbol{y})\| \leq C_0\|\boldsymbol{v}^s(\boldsymbol{y})\| \leq C_0 c_0\|\boldsymbol{y}\| \leq C_0 c_0 r = \frac{a}{3}. \tag{D.10}$$

Therefore, by the triangle inequality,

$$\left\| \boldsymbol{e}_L(\boldsymbol{y}) - \frac{r_1}{\|\boldsymbol{w}^u(\boldsymbol{y})\|}\boldsymbol{w}^u(\boldsymbol{y}) \right\| \leq \left\| \boldsymbol{w}^u(\boldsymbol{y}) - \frac{r_1}{\|\boldsymbol{w}^u(\boldsymbol{y})\|}\boldsymbol{w}^u(\boldsymbol{y}) \right\| + \|\boldsymbol{w}^s(\boldsymbol{y})\|$$
$$= \left| 1 - \frac{r_1}{\|\boldsymbol{w}^u(\boldsymbol{y})\|} \right| \|\boldsymbol{w}^u(\boldsymbol{y})\| + \|\boldsymbol{w}^s(\boldsymbol{y})\|.$$

Since $\|\boldsymbol{e}_L(\boldsymbol{y})\| = r_1$, this expression is equal to

$$\left| \|\boldsymbol{w}^u(\boldsymbol{y})\| - r_1 \right| + \|\boldsymbol{w}^s(\boldsymbol{y})\| = \left| \|\boldsymbol{w}^u(\boldsymbol{y})\| - \|\boldsymbol{e}_L(\boldsymbol{y})\| \right| + \|\boldsymbol{w}^s(\boldsymbol{y})\|.$$

By (D.9) and (D.10), this expression is bounded by $2\|\boldsymbol{w}^s(\boldsymbol{y})\| < (2/3)\,a$. This completes the proof of the lemma since

$$\frac{r_1}{\|\boldsymbol{w}^u(\boldsymbol{y})\|}\,\boldsymbol{w}^u(\boldsymbol{y}) \in \mathcal{M}_{L,u} \cap \partial\mathcal{B}(\boldsymbol{0},\,r_1).$$

References

1. BARRERA, G., JARA, M.: Thermalisation for small random perturbations of dynamical systems. *Ann. Appl. Probab.* **30**, 1164–1208, 2020
2. BELTRÁN, J., LANDIM, C.: Tunneling and metastability of continuous time Markov chains. *J. Stat. Phys.* **140**, 1065–1114, 2010
3. BELTRÁN, J., LANDIM, C.: Tunneling and metastability of continuous time Markov chains II. *J. Stat. Phys.* **149**, 598–618, 2012
4. BELTRÁN, J., LANDIM, C.: Metastability of reversible condensed zero range processes on a finite set. *Probab. Theory Relat. Fields* **152**, 781–807, 2012
5. BERTINI, L., GABRIELLI, D., LANDIM, C.: Metastable Γ-expansion of finite state Markov chains level two large deviations rate functions. *Ann. Appl. Probab.* **34**, 3820–3869, 2024
6. BOVIER, A., ECKHOFF, M., GAYRARD, V., KLEIN, M.: Metastability in reversible diffusion process. I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc.* **6**, 399–424, 2004

7. CHICONE, C.: Ordinary Differential Equations with Applications, vol. 31. Text in Applied Mathematics. Springer, New York (2010)
8. DAY, M.V.: Exponential leveling for stochastically perturbed dynamical systems. *SIAM J. Math. Anal.* **13**, 532–540, 1982
9. DEVINATZ, A., FRIEDMAN, A.: Asymptotic behavior of the principal eigenfunction for a singularly perturbed Dirichlet problem. *Indiana Univ. Math. J.* **27**, 143–157, 1978
10. FAGGIONATO, A., GABRIELLI, D.: A representation formula for large deviations rate functionals of invariant measures on the one dimensional torus. *Ann. Inst. Henri Poincaré Probab. Stat.* **48**, 212–234, 2012
11. FREIDLIN, M.I., KORALOV, L.: Nonlinear stochastic perturbations of dynamical systems and quasi-linear parabolic PDE's with a small parameter Probab. *Theory Relat. Fields* **147**, 273–301, 2010
12. FREIDLIN, M.I., KORALOV, L.: Metastability for nonlinear random perturbations of dynamical systems. *Stoch. Proc. Appl.* **120**, 1194–1214, 2010
13. FREIDLIN, M., KORALOV, L.: Metastable distributions of Markov Chains with rare transitions. *J. Stat. Phys.* **167**, 1355–1375, 2017
14. FREIDLIN, M.I., WENTZELL, A.D.: Random Perturbations of Dynamical Systems, vol. 260. Second edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, New York (1998)
15. ISHII, H., SOUGANIDIS, P.E.: Metastability for parabolic equations with drift: Part I. *Indiana Univ. Math. J.* **64**, 875–913, 2015
16. ISHII, H., SOUGANIDIS, P.E.: Metastability for parabolic equations with drift: part II. The quasilinear case. *Indiana Univ. Math. J.* **66**, 315–360, 2017
17. KIFER, Y.: The exit problem for small random perturbations of dynamical systems with a hyperbolic fixed point. *Isr. J. Math.* **40**, 74–96, 1981
18. KORALOV, L., TCHEUKO, L.: Quasi-linear equations with a small diffusion Term and the evolution of hierarchies of cycles. *J. Theor. Probab.* **29**, 867–895, 2016
19. LANCASTER, P., TISMENETSKY, M.: The Theory of Matrices with Applications, 2nd edn. Academic press. Inc., San Diego (1985)
20. LANDIM, C., LEE, J., SEO, I.: Metastability and time scales for parabolic equations with drift II: the longer time scales. arXiv:2402.07695, 2024
21. LANDIM, C., LOULAKIS, M., MOURRAGUI, M.: Metastable Markov chains: from the convergence of the trace to the convergence of the finite-dimensional distributions. *Electron. J. Probab.* **23**, 1–34, 2018
22. LANDIM, C., MARCONDES, D., SEO, I.: A Resolvent Approach to Metastability. To appear in J. Eur. Math. Soc. arXiv:2102.00998, 2023
23. LANDIM, C., MARIANI, M., SEO, I.: A Dirichlet and a Thomson principle for non-selfadjoint elliptic operators, Metastability in non-reversible diffusion processes. *Arch. Rational Mech. Anal.* **231**, 887–938, 2017
24. LANDIM, C., SEO, I.: Metastability of non-reversible random walks in a potential field, the Eyring–Kramers transition rate formula. *Commun. Pure Appl. Math.* **71**, 203–266, 2018
25. LANDIM, C., SEO, I.: Metastability of one-dimensional, non-reversible diffusions with periodic boundary conditions. *Ann. Inst. H. Poincaré Probab. Stat.* **55**, 1850–1889, 2019
26. LANDIM, C., XU, T.: Metastability of finite state Markov chains: a recursive procedure to identify slow variables for model reduction. *ALEA Lat. Am. J. Probab. Math. Stat.* **13**, 725–751, 2016
27. LEE, S., RAMIL, M., SEO, I.: Asymptotic stability and cut-off phenomenon for underdamped Langevin dynamics. In preparation, 2023
28. LEE, J., SEO, I.: Non-reversible metastable diffusions with Gibbs invariant measure I: Eyring–Kramers Formula. *Probab. Theory Relat. Fields* **182**, 849–903, 2022
29. LEE, J., SEO, I.: Non-reversible metastable diffusions with Gibbs invariant measure II: Markov Chain Convergence. *J. Stat. Phys.* **189**, 25, 2022

30. Lelièvre, T., Le Peutrec, D., Nectoux, B.: The exit from a metastable state: concentration of the exit point distribution on the low energy saddle points, part 2. *Stoch. PDE: Anal. Comput.* **10**, 317–357, 2022
31. Le Peutrec, D., Michel, L.: Sharp spectral asymptotics for nonreversible metastable diffusion processes. *Probab. Math. Phys.* **1**, 3–53, 2019
32. Michel, L.: About small eigenvalues of the Witten Laplacian. *Pure Appl. Anal.* **1**, 149–206, 2019
33. Milnor, J.: Morse Theory, vol. 51. Annals of Mathematics Studies. Princeton University Press, New Jersey (1969)
34. Nicolaescu, L.I.: An Invitation to Morse Theory, 2nd edn. Universitext. Springer, New York (2011)
35. Palis, J., de Melo, W.: Geometric Theory of Dynamical Systems: An Introduction. Springer, New York (1982)
36. Perko, L.: Differential Equations and Dynamical Systems, vol. 7, 3rd edn. Texts in Applied Mathematics. Springer Verlag, New York (2001)
37. Rezakhanlou, F., Seo, I.: Scaling limit of small random perturbation of dynamical systems. *Ann. Inst. H. Poincaré Probab. Stat.* **59**, 867–903, 2023
38. Sugiura, M.: Metastable behaviors of diffusion processes with small parameter. *J. Math. Soc. Jpn.* **47**, 755–788, 1995
39. Webb, J.R.L.: Extensions of Gronwall's inequality with quadratic growth terms and applications. *Electron. J. Qual. Theor. Differ.* **61**, 1–12, 2018

Claudio Landim
IMPA,
Estrada Dona Castorina 110, J. Botanico,
Rio de Janeiro
22460 Brazil.
e-mail: landim@impa.br

and

Univ. Rouen Normandie, CNRS, LMRS UMR 6085,
76000 Rouen
France.

and

Jungkyoung Lee
School of Mathematics,
Korea Institute for Advanced Study,
Seoul
Republic of Korea.
e-mail: jklee@kias.re.kr

and

Insuk Seo
Department of Mathematical Sciences,
Seoul National University and Research Institute of Mathematics,
Seoul
Republic of Korea.
e-mail: insuk.seo@snu.ac.kr