



**HAL**  
open science

# Exit-problem for Self-interacting and Self-stabilizing Diffusion Processes

Ashot Aleksian

► **To cite this version:**

Ashot Aleksian. Exit-problem for Self-interacting and Self-stabilizing Diffusion Processes. Probability [math.PR]. Université Jean Monnet; ED 488 SIS, 2023. English. NNT : 2023STET039 . tel-04287981v1

**HAL Id: tel-04287981**

**<https://hal.science/tel-04287981v1>**

Submitted on 15 Nov 2023 (v1), last revised 10 Jan 2024 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N° d'ordre NNT : 2023STET039

**THÈSE de DOCTORAT  
DE L'UNIVERSITÉ JEAN MONNET SAINT-ÉTIENNE**

**Membre de l'Université de LYON**

**École Doctorale N° 488  
(SCIENCES INGÉNIERIE SANTÉ SIS)**

**Mathématiques Appliquées**

Soutenue publiquement le 20/11/2023, par :  
**Ashot Aleksian**

---

**Problème de Temps de Sortie pour les Diffusions  
Auto-interagissantes et Auto-stabilisantes**

---

Devant le jury composé de :

Champagnat, Nicolas	Directeur de Recherche / Inria Nancy-Grand Est	Président
Essouabri, Driss	Professeur / UJM Saint-Étienne, ICJ	Examineur
Herrmann, Samuel	Professeur / Université de Bourgogne, IMB	Examineur
Miclo, Laurent	Directeur de Recherche / CNRS	Examineur
Koralov, Leonid	Full professor / University of Maryland	Rapporteur/Examineur
Lelievre, Tony	Professeur / École des Ponts, CERMICS	Rapporteur/Examineur
Tugaut, Julian	MCF HC / UJM Saint-Étienne, ICJ	Directeur de thèse
Kurtzmann, Aline	MCF / Université de Lorraine, IECL	Co-directrice de thèse

# Exit-problem for Self-interacting and Self-stabilizing Diffusion Processes

Ashot Aleksian

A Thesis Presented in Partial Fulfilment of the Requirements for the  
Degree of

DOCTOR OF PHILOSOPHY

APPLIED MATHEMATICS

Supervisors: Aline Kurtzmann – Associate professor, Institut Élie  
Cartan de Lorraine (IECL), University of Lorraine  
Julian Tugaut – Associate professor, Institut Camille  
Jordan (ICJ), Jean Monnet University

Institut Camille Jordan (ICJ)  
JEAN MONNET UNIVERSITY

November 20, 2023



## Résumé

Dans cette thèse, le problème du temps de sortie pour deux types de processus de diffusion non linéaire est étudié. Le premier processus est appelé diffusion auto-interagissante et est défini par l'équation différentielle stochastique suivante, incluant l'interaction du processus avec son propre passé :

$$dX_t = \sigma dW_t - \left( \nabla V(X_t) + \frac{1}{t} \int_0^t \nabla F(X_t - X_s) ds \right) dt.$$

La deuxième équation à laquelle on s'intéresse est un cas particulier du processus de McKean-Vlasov et nous l'appelons la diffusion auto-stabilisante. Elle est définie par l'équation différentielle stochastique suivante, incluant la convolution du processus avec sa loi au temps  $t$ ,  $\mathcal{L}(X_t)$  :

$$dX_t = \sigma dW_t - \left( \nabla V(X_t) + \nabla F * \mathcal{L}(X_t)(X_t) ds \right) dt.$$

Le problème du temps de sortie considéré ici est l'étude du premier instant auquel un processus de diffusion donné sort d'un domaine fixé  $G$ , dans un régime de faible bruit :  $\tau := \inf\{t \geq 0 : X_t \notin G\}$ . En particulier, nous nous intéressons au comportement asymptotique du temps d'atteinte  $\tau$ , lorsque le paramètre de diffusion  $\sigma \rightarrow 0$ , et nous recherchons la loi de type Kramers, c'est-à-dire

$$\tau \approx e^{\frac{2H}{\sigma^2}},$$

où  $H > 0$  est une constante qui contrôle la vitesse de croissance du temps de sortie. De plus, nous décrivons l'emplacement des points où la diffusion peut quitter le domaine  $G$  (ce qui correspond au problème de localisation de sortie).

La thèse se compose de 4 chapitres. Le premier chapitre présente les processus de diffusion auto-interagissants et auto-stabilisants, leurs propriétés importantes et l'état de l'art du problème de sortie.

Les chapitres 2 et 3 sont axés sur le problème du temps de sortie pour les diffusions auto-interagissantes. Dans le chapitre 2, on suppose que les

fonctions  $V$  et  $F$  sont convexes, ce qui nous permet d'utiliser des techniques de couplage standard pour établir la loi de type Kramers et le résultat de localisation de sortie. Le chapitre 3 ne suppose aucune convexité sur  $V$  ou  $F$ , ce qui rend le problème nettement plus difficile et intéressant. Dans ce chapitre, nous démontrons le principe des grandes déviations pour la diffusion auto-interagissante et utilisons des techniques similaires à la théorie de Freidlin-Wentzell, tout en tenant compte du fait que nous sommes dans un cadre non markovien. Avec cette approche, nous établissons la loi de type Kramers ainsi que le résultat de localisation de sortie.

Enfin, le chapitre 4 résout le problème du temps de sortie pour les processus de diffusion auto-stabilisants avec des potentiels généraux  $V$  et  $F$ . Il s'agit d'un problème ouvert depuis plus de 15 ans. Pour l'étudier, nous utilisons des techniques de couplage améliorées afin d'établir la loi de type Kramers et les résultats de localisation de sortie.

### Abstract

In this thesis, exit-time problem for two types of non-linear diffusion processes is considered. The first process is called the Self-interacting diffusion and is defined by the following SDE including interaction of the process with its own path:

$$dX_t = \sigma dW_t - \left( \nabla V(X_t) + \frac{1}{t} \int_0^t \nabla F(X_t - X_s) ds \right) dt.$$

The second equation of interest is a special case of McKean-Vlasov process that we call the Self-stabilizing diffusion. It is defined by the following SDE including the convolution of the process with its law  $\mathcal{L}(X_t)$ :

$$dX_t = \sigma dW_t - \left( \nabla V(X_t) + \nabla F * \mathcal{L}(X_t)(X_t) ds \right) dt.$$

By the exit-time problem, we mean the study of the stopping time  $\tau := \inf\{t \geq 0 : X_t \notin G\}$ , the first time when a given diffusion process exits a fixed domain of attraction  $G$ , in the small noise regime. In particular, we are

interested in the asymptotic behaviour of  $\tau$ , when the diffusion parameter  $\sigma \rightarrow 0$  and we are looking for the so-called Kramers' type law, that is

$$\tau \approx e^{\frac{2H}{\sigma^2}},$$

where  $H > 0$  is a constant that controls the rate of growth of the exit-time. Moreover, we are interested in describing the location of points where the diffusion is likely to leave the domain  $G$  (exit-location problem).

The thesis consists of 4 chapters. The first chapter presents the Self-interacting and the Self-stabilizing diffusion processes, their important properties and the state of the art of the exit-problem.

Chapters 2 and 3 are focused on the exit-time problem for the Self-interacting diffusions. In Chapter 2 functions  $V$  and  $F$  are assumed to be convex, which allows us to use standard coupling method techniques to establish the Kramers' type law and the exit-location results. Chapter 3 does not assume any convexity of  $V$  or  $F$ , which makes the problem significantly more difficult. In this chapter, we prove the large deviations principle for the Self-interacting diffusion and use techniques similar to the Freidlin-Wentzell theory, but taking into account that we are in the non-Markov case. Using this approach, we establish the Kramers' type law and the exit-location results.

Finally, Chapter 4 solves the exit-time problem for the Self-stabilizing diffusion processes for general potentials  $V$  and  $F$ , which was an open problem for more than 15 years. In this case, we use improved coupling techniques to establish the Kramers' type law and the exit-location results.

*Keywords:* Self-interacting diffusion, measure-dependent diffusions, metastability, exit-time, Kramers' law, Freidlin-Wentzell theory, large deviations.

*AMS 2020 Subject Classification:* 60K35, 60J60, 60F10, 60H10.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Self-interacting diffusion . . . . .	1
1.2	Self-stabilizing diffusion . . . . .	6
1.3	Large Deviation Principle . . . . .	8
1.4	Exit-time problem . . . . .	11
1.5	Outline of the thesis . . . . .	14
<b>2</b>	<b>Exit-problem for SID: Convex case</b>	<b>19</b>
2.1	Assumptions and main results . . . . .	21
2.2	Exit-problem . . . . .	27
2.3	Discussions on extension . . . . .	38
<b>3</b>	<b>Exit-problem for SID: General case</b>	<b>41</b>
3.1	Large Deviation Principle . . . . .	45
3.2	Exit-time . . . . .	57
3.3	Generalization . . . . .	90
<b>4</b>	<b>Exit-problem for SSD: General case</b>	<b>93</b>
4.1	Main results . . . . .	101
4.2	Intermediate results . . . . .	110
4.3	Proofs of the main results . . . . .	114
4.4	Proofs of the intermediate results . . . . .	123
	<b>Bibliography</b>	<b>145</b>



# List of Symbols

$ \cdot , \langle \cdot; \cdot \rangle$	Euclidean norm and scalar product in $\mathbb{R}^d$
$\text{Int}(\Gamma)/\overset{\circ}{\Gamma}, \text{Cl}(\Gamma)/\bar{\Gamma}, \partial\Gamma$	Interior, closure and boundary of a set $\Gamma \subset \mathbb{R}^d$
$B_\rho(a) := \{x \in \mathbb{R}^d :  x - a  \leq \rho\}$	A closed ball of radius $\rho$ around $a$ in $\mathbb{R}^d$
$S_\rho(a) := \{x \in \mathbb{R}^d :  x - a  = \rho\}$	A closed sphere of radius $\rho$ around $a$ in $\mathbb{R}^d$
$C([0; T]; \mathbb{R}^d)$	The set of continuous functions on $[0; T]$ with values in $\mathbb{R}^d$ equipped with the norm $\ \phi\ _\infty := \max_{t \in [0; T]}  \phi(t) $
$C_0([0; T]; \mathbb{R}^d)$	Subset of $C([0; T]; \mathbb{R}^d)$ consisting of functions $\psi$ such that $\psi(0) = 0$
$C^2(\mathbb{R}^d; \mathbb{R})$	The set of functions on $\mathbb{R}^d$ with values in $\mathbb{R}$ with continuous partial derivatives up to order 2
$L([0; T]; \mathbb{R}^d)$	The set of integrable w.r.t the Lebesgue measure functions on $[0; T]$ with values in $\mathbb{R}^d$
$L_2([0; T]; \mathbb{R}^d)$	Subset of $L_2([0; T]; \mathbb{R}^d)$ consisting of functions $\psi$ such that $\int_0^T  \psi(s) ^2 ds < \infty$
$\mathcal{M}(\mathbb{R}^d), \mathcal{P}(\mathbb{R}^d)$	The space of signed (bounded) Borel measures on $\mathbb{R}^d$ and its subspace of probability measures
$\mathcal{P}_2(\mathbb{R}^d)$	The set of probability measures on $\mathbb{R}^d$ with finite second moments

$\mathbb{W}_2(\mu, \nu)$	The Wasserstein-2 distance between measures $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ (see e.g. [Vil09, Definition 6.1])
$(W_t)_{t \geq 0}$	The Wiener process
$\mathcal{L}(X)$	The law of a random variable $X$
$\#$	A push-forward symbol (see e.g. [Vil09, p.11])
$\lceil z \rceil, \lfloor z \rfloor$	The least integer greater than or equal to $z \in \mathbb{R}$ and the greatest integer less than or equal to $z \in \mathbb{R}$

The following notation is specific for Chapter 3:

$\mathcal{P}_2(\overline{G})$	The set of probability measures with finite second moments, defined on the closure $\overline{G} \subset \mathbb{R}^d$ , where $G$ is a fixed domain defined in Chapter 3, on page 57
$(\mathfrak{X}, d_{\mathfrak{X}})$	A metric space consisting of triples $\mathbf{x} = (t, \mu, x)$ , $t \in [0, \infty]$ , $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ , and $x \in \mathbb{R}^d$ (see Section 3 of Chapter 3)
T, M, P	Projection maps on the first, second, and third coordinates of an $\mathbf{x} \in \mathfrak{X}$ (see Section 3 of Chapter 3)
$\mathbb{B}_{\rho}(\mu) := \{\nu \in \mathcal{P}_2(\overline{G}) : \mathbb{W}_2(\mu; \nu) \leq \rho\}$	A closed ball of radius $\rho$ around $\mu$ in $\mathcal{P}_2(\overline{G})$ , where $G$ will be the same domain as in Assumption A-2, page 57
$\mathbb{B}_{\rho}(\mathbf{x}) := \{\mathbf{y} \in \mathfrak{X} : d_{\mathfrak{X}}(\mathbf{y}, \mathbf{x}) \leq \rho\}$	A closed ball of radius $\rho$ around $\mathbf{x}$ in $\mathfrak{X}$

In this thesis we work on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , which we assume to be rich enough. Except where otherwise indicated, all random variables and stochastic processes are defined on this probability space and are, in the case of processes,  $\mathcal{F}_t$ -adapted. All the equations in the manuscript are meant to hold a.s., unless stated otherwise.

Note that, because of historical reasons and specifics of the proofs, the notation is adapted to each chapter. Consequently, similar objects may have different names in different chapters. For example, the domain, from which exit-time is established, is called  $G$  in some chapters, but  $\mathcal{D}$  in others. Another example is the main processes of interest that sometimes appear with or without superscripts:  $X^\sigma$ ,  $X$ . Although being different in the entire manuscript, the notation is consistent within each chapter.



# Chapter 1

## Introduction

This thesis is focused on the exit-problem for different non-linear diffusions. In particular, so-called Self-interacting and Self-stabilizing diffusions were considered. The first one is defined by a stochastic differential equation including interaction of each trajectory with its own past. Non-linearity in the second process arises from its dependency on its own law (a particular case of McKean-Vlasov process, see e.g. [HIP08]). By exit-time problem, we mean questions related to some properties of the stopping time defined as the first time when a diffusion leaves the domain of attraction. Formal definitions of the two processes as well as the exit-time problem are provided in the following sections.

### 1.1 Self-interacting diffusion

Systems with path-interaction behaviour have long history of mathematical study. Given the large number of applications for them in game theory, social science, computer science and other related fields, a large number of similar models have been developed during this period. That includes various Pólya urn type and reinforced random walk models in discrete time, and self-interacting diffusion processes in continuous time. In the current work, we are interested in continuous-time models, yet we provide discrete-time case as well.

In 1987 ([Dia88]) the model of reinforced random walks on finite graphs was introduced. Namely, given a graph  $G = (V, E)$ , where  $V$  is the set of

vertices and  $E$  is the set of edges. On the first step of the algorithm all edges of the graph have the same weight  $W(e, 0) = 1$  and, starting from the vertex  $v_0 \in V$ , the probability to go to another vertex  $w \in V$  adjacent to  $v_0$  is equal to:

$$\mathbb{P}(X_1 = w | X_0 = v_0) = \frac{W(\{v_0, w\}, 0)}{\sum_z W(\{v_0, z\}, 0)},$$

where summation is taken over all vertices  $z$  that are adjacent to  $v_0$ . After the first step, the weights are updated by adding one to the edge along which the process ultimately went through, i.e.  $W(\{v_0, w\}, 1) = W(\{v_0, w\}, 0) + \mathbb{1}\{X_1 = w\}$ . This process is repeated inductively. Namely, define  $\mathcal{F}_n := \sigma(X_0, \dots, X_n)$  and for all adjacent vertices:

$$\mathbb{P}(X_{n+1} = w | \mathcal{F}_n) = \frac{W(\{X_n, w\}, n)}{\sum_z W(\{X_n, z\}, n)}.$$

The random walk, defined on graphs the way described above, is called edge-reinforced (ERRW – edge reinforced random walk). Reinforcement in the system is introduced by increasing the probability of edges, through which the process has already passed. In the same work, P. Diaconis proves that for all edges  $e \in E$ , the limit  $\lim_{n \rightarrow \infty} W(e, n)/n$  exists a.s. Moreover, the vector of all weights  $\mathbf{W}(n) = (W(e_1, n), \dots, W(e_N, n))/n$ , where  $N := |E|$ , converges a.s. to a limit on an  $N$ -dimensional simplex, the distribution of which is absolutely continuous with respect to the Lebesgue measure and its density was explicitly obtained by the author in [Dia88]. The work on ERRW on graphs and trees in particular was continued in [Pem88a, Col06a, Col06b].

Another type of reinforced random walks on graphs can be introduced by updating the weights of the vertices instead of the edges. Namely, for all  $v \in V$  define  $W(v, 0) = 1$ . For each consecutive step let us define the probability to go from vertex  $v$  to vertex  $w$  as:

$$\mathbb{P}(X_{n+1} = w | \mathcal{F}_n) = \frac{W(w, n)}{\sum_z W(z, n)},$$

where the summation is taken over all adjacent  $v$  vertices. The process defined this way is called vertex reinforced random walk (VRRW). It turned

out, that behaviour of this process is drastically different from ERRW. Moreover, the form of the graph also affects the outcome result. For example, it was shown in [Pem88b], that if  $G$  is the complete graph on  $N$  vertices, then  $\mathbf{W}(n) \rightarrow (1/N, \dots, 1/N)$  a.s. At the same time, it follows from [Pem88b] and [Pem90] that, in the case of  $G$  being a simple cycle of  $N \geq 5$  nodes, all the limit points of  $\mathbf{W}$  have the form  $(0, \dots, 0, a, 1/2, 1/2 - a, 0, \dots, 0)$  and some convex combinations of these points.

Systems with path-interaction behaviour in continuous time have already been studied by numerous researchers for more than 30 years. In most of the papers the long-time behaviour of the process is considered. One of the first mathematical descriptions of such a process in dimension  $d = 1$  is presented in [NRW87] by J.R. Norris, L.C.G. Rogers, and D. Williams under the name of Self-avoiding random walk:

$$X_t = W_t - \int_0^t g(X_s, L(s, X_s)) ds,$$

where  $g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  is regular enough function and  $\{L(t, x), t \geq 0, x \in \mathbb{R}\}$  is the local time process of  $X$ . Some stochastic properties of the process along with some long-time behaviour results in a particular setting were shown. The main difference between the system considered there and ours is that in [NRW87] there is no renormalization of the interaction term with time.

In [DR92] R.T. Durrett and L.C.G. Rogers introduce a similar (to the previous paper) system that aims to model the shape of a growing polymer, where “newly added units are repelled by existing ones”. Given the physical interpretation of the process, the authors called this model “the Brownian Polymer”:

$$dX_t = dW_t + \left( \int_0^t f(X_t - X_s) ds \right) dt.$$

In this paper asymptotic bounds in dimension  $d = 1$  with some assumptions on interaction function were presented, along with some conjectures on more precise long-time behaviour. Among those, the authors conjectured that, in the symmetrical, *repulsive* case with limited interaction ( $f(-x) = -f(x)$ ,  $xf(x) \geq 0$ , and  $f$  has compact support), one can

show that  $X_t/t \rightarrow 0$  a.s. This repulsive case was later studied in the paper [TTV12], where the conjecture was partially proved (with an additional assumption on smoothness of  $f$ ).

Later, another model of growing polymer was introduced by Benaïm, Ledoux and Raimond [BLR02], for which the drift term depends on its own empirical measure. Namely, they have studied the following process living in a compact smooth connected Riemannian manifold  $M$  without boundary:

$$dX_t = \sum_{i=1}^N G_i(X_t) \circ dW_t^i - \int_M \nabla_x F(X_t, y) \mu_t(dy) dt,$$

where  $F$  is a (smooth) interaction potential,  $(W^1, \dots, W^N)$  is a standard Brownian motion on  $\mathbb{R}^N$ ,  $\mu_t := \frac{1}{t} \int_0^t \delta_{X_s} ds$  and the symbol  $\circ$  stands for the Stratonovich stochastic integration. In the compact setting, they have shown that the asymptotic behaviour of the empirical measure of the process can be related to the analysis of some deterministic dynamical flow. Later, Benaïm and Raimond [BR05] gave sufficient conditions for the almost sure convergence of the empirical measure (again in the compact setting). More recently, Raimond [Rai09] has generalized the previous study and has proved that for the solution of the SDE living on a compact manifold

$$dX_t = dW_t - \frac{g(t)}{t} \int_0^t \nabla_x F(X_t, X_s) ds dt$$

the approximation of the empirical measure by a deterministic flow is no longer valid, unless  $g$  is constant.

First step towards studying this process in the non-compact setting, which was  $\mathbb{R}^d$ , was done by A. Kurtzmann in [Kur10]. The author considered the following model with interaction depending on the empirical measure of the process.

$$dX_t = dW_t - \left( \nabla V(X_t) + \frac{1}{t} \int_0^t \nabla_x F(X_t, X_s) ds \right) dt,$$

where  $V$  and  $F$  are regular functions and  $V$  is convex at infinity, which means that there exist constants  $K > 0$  and  $R > 0$  such that for any  $|x| > R$  and for any  $h \in \mathbb{R}^d : |h| = 1$  we have:

$$\langle h; \nabla^2 V(x) h \rangle \geq K.$$



In [Kur10] the ergodic properties of  $X$  were studied as well as certain conditions on  $V$  and  $F$  that guarantee almost sure convergence of  $\mu_t$  in  $*$ -weakly sense. This work generalizes the previous ones of Benaïm and co-authors.

In this thesis we consider SID of the form:

$$dX_t = \sigma dW_t - \left( \nabla V(X_t) + \frac{1}{t} \int_0^t \nabla F(X_t - X_s) ds \right) dt. \quad (1.1)$$

At each point of time the position of the process is driven by the general geometry of the space (introduced by the potential  $V$ ) the interaction with its own path that is defined via the convolution of the empirical measure  $\mu_t := \frac{1}{t} \int_0^t \delta_{X_s} ds$  of the process with the interaction potential  $F$  (see Figure 1.1).

In the paper of V. Kleptsyn and A. Kurtzmann [KK12] dynamic of the form (1.1) with constant  $\sigma = \sqrt{2}$  is considered. The authors proved that, in the case of convex confinement  $V$  with  $\lim_{|x| \rightarrow +\infty} V(x) = +\infty$ , attractive and symmetrical interaction (potential  $F$  is uniformly convex and spherically symmetrical, i.e.  $F(x) = F(|x|)$ ), and some other regularity assumptions, the occupation measure  $\mu_t$  converges (in  $*$ -weakly sense) almost surely to a density  $\rho_\infty$ . It means that there exists a unique density  $\rho_\infty : \mathbb{R}^d \rightarrow \mathbb{R}_+$  such that

$$\mu_t = \frac{1}{t} \int_0^t \delta_{X_s} ds \xrightarrow[t \rightarrow +\infty]{*-weakly} \rho_\infty(x) dx \quad \text{a.s.}$$

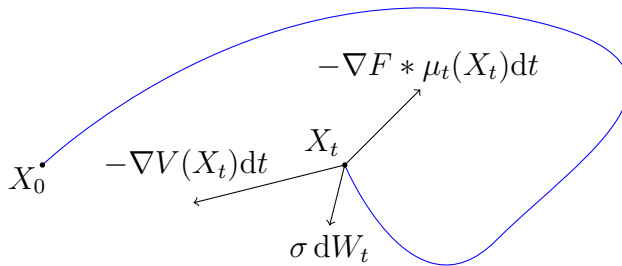


Figure 1.1: Different components of SID and their influence on dynamics of the process.

Moreover,  $\rho_\infty$  is a function satisfying the following functional equation:

$$\rho_\infty = \Pi(\rho_\infty) := \frac{\exp\{-V - F * \rho_\infty\}}{\int_{\mathbb{R}^d} \exp\{-V(y) - F * \rho_\infty(y)\} dy}.$$

In the same paper, the authors also considered the case where the confinement is absent, i.e.  $V \equiv 0$ . It turned out that for this dynamics we can not show that the empirical measure converges a.s. to the same deterministic limit. Yet, V. Kleptsyn and A. Kurtzmann showed that there still exists a limit for almost every  $(\mu_t)_{t \geq 0}$  and these limits are the same up to translation of a deterministic measure. Namely, they proved that there exists a unique (deterministic) symmetric density  $\rho_\infty : \mathbb{R}^d \rightarrow \mathbb{R}_+$ , such that almost surely there exists a random variable  $c_\infty$  such that

$$\mu_t \xrightarrow[t \rightarrow +\infty]{*-weakly} \rho_\infty(x + c_\infty) dx.$$

## 1.2 Self-stabilizing diffusion

Self-stabilizing diffusion is a special case of McKean-Vlasov process, i.e. the process defined by a stochastic differential equation with dependency of the diffusion coefficients on the law of the solution:

$$\begin{cases} dX_t &= a(X_t, \mu_t) dW_t + b(X_t, \mu_t) dt, \\ \mu_t &= \mathcal{L}(X_t). \end{cases}$$

Existence and uniqueness results for this process in the case of Lipschitz continuous  $a$  and  $b$  could be found in the review [CD22a].

This type of processes is a natural generalization of the system of a large number of interacting particles. Historically, these systems were introduced as an attempt to resolve a long-standing theoretical problem, posed by Boltzmann (see e.g. [Bol03]), aiming at studying the thermodynamic limit of elastically colliding particles, when the number of particles tends to infinity. Using physics considerations, Boltzmann formally derived an equation for the probability density function of continuum of particles on the state space of their positions and momenta  $f : (r, p, t) \mapsto \mathbb{R}$  that takes

the following form:

$$\frac{\partial}{\partial t} f(r, p, t) = -\langle \nabla_r f(r, p, t); p \rangle - \langle \nabla V(r); \nabla_p f(r, p, t) \rangle + \left( \frac{\partial f}{\partial t} \right)_{\text{coll}},$$

where  $\nabla V$  describes the force field acting on the particles, and  $\left( \frac{\partial f}{\partial t} \right)_{\text{coll}}$  is a term that aims to describe the change of distribution of the particles due to their collision. It turned out that if we represent particles as balls with fixed radius interacting by colliding with each other, then it is hard to derive thermodynamic limit from only the microscopic behaviour of the system.

Instead, McKean introduced another class of models, in which particles are defined to be point-like and interaction between them is given via the interaction potential. Consider  $N$ -particles system living in  $\mathbb{R}^d$ :  $(X_t^{1,N}, \dots, X_t^{N,N})$ . The dynamics of each particle depends on its position and the position of all other particles at the same time. The latter is introduced in the system by the empirical measure of all particles:  $\mu_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i,N}}$ . A simplified version of the interacting particle system modeling this behaviour is defined as

$$dX_t^{i,N} = a(X_t^{i,N}, \mu_t^N) dW_t^i + b(X_t^{i,N}, \mu_t^N) dt.$$

It was noted that, heuristically, if the number of particles tends to infinity, due to the big number of interacting forces that cancel each other, they should become independent from each other and their empirical measure should converge to the law of one representative particle. This phenomenon is called in the literature the propagation of chaos. The following result is due to McKean. Let diffusion coefficients be defined as

$$a(x, \mu) = f_a(x, K_a * \mu(x)), \quad b(x, \mu) = f_b(x, K_b * \mu(x)),$$

where  $f_a, f_b : \mathbb{R}^d \times \mathbb{R}^m \rightarrow \mathbb{R}^d$  are Lipschitz continuous and  $K_a, K_b : \mathbb{R}^d \rightarrow \mathbb{R}^m$  are two bounded kernels and let

$$d\bar{X}_t^{i,N} = a(\bar{X}_t^{i,N}, \bar{\mu}_t) dW_t^i + b(\bar{X}_t^{i,N}, \bar{\mu}_t) dt,$$

where  $\bar{\mu}_t = \mathcal{L}(\bar{X}_t^i)$  and  $\bar{X}_t^i$  are independent from each other. The result of McKean then states that for any  $T > 0$ :

$$\frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \sup_{t \leq T} \left| X_t^{i,N} - \bar{X}_t^{i,N} \right|^2 \right] \leq \varepsilon(N, T),$$

where  $\varepsilon(N, T) \xrightarrow[N \rightarrow +\infty]{} 0$ . This result was proved for more general drift and diffusion coefficients (see [CD22a, CD22b] for more information).

In the case of Self-stabilizing diffusion that we describe here, the SDE driving the system takes the following form:

$$\begin{cases} dX_t^\sigma = \sigma dW_t - \nabla V(X_t^\sigma) dt - \nabla F * \mu_t^\sigma(X_t^\sigma) dt, \\ \mu_t^\sigma = \mathcal{L}(X_t^\sigma), \end{cases} \quad (1.2)$$

where  $V$ , called confinement potential, represents the environment in which the particle defined by this SDE moves,  $F$ , called interaction potential, controls the way it interacts with its own law that is denoted as  $\mu^\sigma$ .

In the form (1.2) the self-stabilizing diffusion appears in the paper [HIP08]. In this work, the authors proved existence and uniqueness of the process under more relaxed assumptions. In particular,  $\nabla V$  and  $\nabla F$  are assumed to be locally Lipschitz continuous with at most polynomial growth,  $\nabla F$  is assumed to be rotationally invariant and, most importantly,  $V$  is assumed to be uniformly convex at infinity. The last assumption guarantees non-explosiveness of the process, since, under it,  $V$  strongly attracts back to the origin the paths that drifted too far away.

Apart from the obvious parallels that can be drawn between equations (1.1) and (1.2) that define SID and SSD respectively, the following fact makes self-stabilizing case similar to the self-interacting one. As was shown in [CGM08], in the case of convex confinement and interaction and under some other regularity assumptions, the law  $\mu_t^\sigma$  converges to the unique probability measure that is solution of  $\nu = \Pi(\nu)$ , where

$$\Pi(\nu) := \frac{\exp\{-V - F * \nu\}}{\int_{\mathbb{R}^d} \exp\{-V(y) - F * \nu(y)\} dy}$$

which is the same behaviour that was established in [KK12].

### 1.3 Large Deviation Principle

The Large deviation techniques are widely used in this thesis and moreover it was the main tool used to show the exit-time result for Itô diffusion in

the seminal works by M. Freidlin and A. Wentzell (we come back to this problem Section 1.4). In this section we present the definition of the Large Deviation Principle and some results related to it that will be used in this thesis.

Here, by the Large Deviation Principle (LDP) we mean the following asymptotic behaviour of measures.

**Definition 1.1.** Family of measures  $(\nu^\sigma)_{\sigma>0}$  defined on some Banach space  $B$  equipped with Borel sigma-algebra  $\mathcal{B}$  is said to satisfy the Large Deviation Principle with a good rate function  $I$  if for any measurable set  $\Gamma \in \mathcal{B}$ :

$$-\inf_{x \in \Gamma} I(x) \leq \liminf_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \nu^\sigma(\Gamma) \leq \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \nu^\sigma(\Gamma) \leq -\inf_{x \in \bar{\Gamma}} I(x),$$

where  $I : B \rightarrow [0, \infty]$  is a lower-semicontinuous function (this property defines *rate function*) whose level sets  $\{x : I(x) \leq \alpha\}$  are compact subsets of  $B$  for any  $0 \leq \alpha < \infty$  (which means by definition that the rate function is *good*).

Note that our definition of LDP by its appearance deviates from the conventional one (see [DZ10]), but they are equivalent up to multiplication of the “conventional” rate function by  $1/2$ . The reason why we choose the convergence rate of this form is to have the rate  $I$  to correspond to the height of the domain in the case of Itô diffusion process (see Section 1.4).

As can be seen in the definition, the LDP corresponds to a property of exponential decay of improbable events that is controlled by the rate function  $I$ .

The first important notion of Large Deviations theory is exponential equivalence of two families of measures. Consider the following definition.

**Definition 1.2.** Two families of probability measures  $\{\mu_\sigma\}_{\sigma>0}$  and  $\{\tilde{\mu}_\sigma\}_{\sigma>0}$  defined on the same Banach space  $B$  with a norm  $\|\cdot\|_B$  are called *exponentially equivalent* if there exist probability spaces  $\{(\Omega, \mathcal{F}_\sigma, \mathbf{P}_\sigma)\}$  and two families of  $B$ -valued random variables  $\{Z_\sigma\}$  and  $\{\tilde{Z}_\sigma\}$  with joint laws  $\{\mathbf{P}_\sigma\}$  and marginals  $\{\mu_\sigma\}$  and  $\{\tilde{\mu}_\sigma\}$ , respectively, s.t.:

1. For any  $\delta > 0$  the set  $\{\omega : (\tilde{Z}_\sigma, Z_\sigma) \in \Gamma_\delta\}$  is  $\mathcal{F}_\sigma$ -measurable where  $\Gamma_\delta := \{(y, \tilde{y}) \in B \times B : \|y - \tilde{y}\|_B > \delta\}$  - the set of points that are apart of each other at a distance more than  $\delta$ ,
2.  $\limsup_{\sigma \rightarrow 0} \sigma \log \mathbf{P}_\sigma(\Gamma_\delta) = -\infty$ .

As the name suggests, if two families of measures  $\{\mu_\sigma\}_{\sigma>0}$  and  $\{\tilde{\mu}_\sigma\}_{\sigma>0}$  are exponentially equivalent, they should possess same properties related to their exponential convergence to some limit. The following lemma expresses the idea that if two families of measures are exponentially equivalent, LDP for one of them implies LDP for the other (see [DZ10, Theorem 4.2.13] for the proof).

**Lemma 1.3.** *If an LDP with a good rate function  $I$  holds for the family of probability measures  $\{\mu_\sigma\}_{\sigma>0}$ , which are exponentially equivalent to  $\{\tilde{\mu}_\sigma\}_{\sigma>0}$ , then LDP with the same good rate function holds for  $\{\tilde{\mu}_\sigma\}_{\sigma>0}$ .*

The following lemma states that LDP is preserved under continuous mappings (see [DZ10, Theorem 4.2.1] for the proof).

**Lemma 1.4** (Contraction Principle). *Let  $B_1$  and  $B_2$  be two Banach spaces and  $f : B_1 \rightarrow B_2$  be a continuous mapping. Consider a good rate function  $I : B_1 \rightarrow [0, \infty]$ . Then following two statements are true.*

1. *The function  $I'(y) := \inf\{I(x) : x \in B_1, y = f(x)\}$  is a good rate function on the space  $B_2$ .*
2. *If  $I$  controls the LDP associated with a family of probability measures  $\{\mu_\sigma\}_{\sigma>0}$  on  $B_1$  then  $I'$  controls the LDP associated with a family of probability measures  $\{\mu_\sigma \circ f^{-1}\}_{\sigma>0}$  on  $B_2$ .*

Since Brownian motion drives the stochastic behaviour in Itô diffusion process as well as for (1.1) and (1.2), the natural first step is to establish LDP for its path. Consider the process:

$$W_t^\sigma = \sigma W_t$$

where  $W_t$  as previously defined is a  $d$ -dimensional Brownian motion. Let  $\nu_\sigma$  be the probability measure induced by  $W_t^\sigma$  on  $C_0([0, T]; \mathbb{R}^d)$  the space of

all continuous functions  $\phi : [0, T] \rightarrow \mathbb{R}^d$  s.t.  $\phi(0) = 0$ , equipped with usual supremum norm. Consider the following

**Lemma 1.5** (Schilder). *A family of probability measures  $\{\nu_\sigma\}_{\sigma>0}$  induced on  $C_0([0, 1]; \mathbb{R}^d)$  by  $\{W^\sigma\}$  satisfies the LDP with good rate function*

$$I(\phi) = \begin{cases} \frac{1}{4} \int_0^T |\dot{\phi}(t)|^2 dt, & \phi \in H_1 \\ \infty, & \text{otherwise} \end{cases} \quad (1.3)$$

where  $H_1 := \{\int_0^t f(s) ds : f \in L_2([0, T]; \mathbb{R}^d)\}$  is the space of all absolutely continuous functions that possess a square integrable derivative and  $f(0) = 0$ . This space is equipped with the norm  $\|g\|_{H_1} = [\int_0^T |\dot{g}(t)|^2 dt]^{1/2}$ .

This result means that on the time interval  $[0, T]$ , with high probability, for small  $\sigma$ ,  $W_t^\sigma$  stays in a small neighbourhood of 0, while the probability of deviating from this region of space is exponentially small with the rate defined by (1.3).

## 1.4 Exit-time problem

The main goal of this thesis is to obtain the first exit times of self-stabilizing and self-interacting diffusion processes, from some bounded regular domain  $G \subset \mathbb{R}^d$ , i.e., the following stopping time  $\tau_G^\sigma := \inf\{t \geq 0 : X_t^\sigma \notin G\}$ . In particular, we are interested in exits driven by the Brownian motion with small noise  $\sigma > 0$ . Similar problems have been present in physical and chemical literature for a great while, at least since the works by Arrhenius [Arr89b, Arr89a]. The exit-time problem in the case of Itô diffusion was mathematically solved by M. Freidlin and A. Wentzell. The techniques that they used in number of papers, starting with [VF70] and presented in their book [FW98], are known under the name of Freidlin-Wentzell theory (see also [DZ10, §5.6]).

Consider an Itô diffusion in gradient form (also known as reversible case):

$$dX_t = -\nabla V(X_t) dt + \sigma dW_t. \quad (1.4)$$

The Freidlin-Wentzell theory considers a regular domain  $G \subset \mathbb{R}^d$  with only one attraction point of marginal deterministic process (with  $\sigma = 0$ ) that we

denote  $a \in G$ . This property is called stability of  $G$  under  $-\nabla V$  or simply positive invariance. Mathematically, it means that if  $X^0$  is the unique solution of the following ODE

$$X_t^0 = x_0 - \int_0^t \nabla V(X_s^0) ds,$$

then for any  $x_0 \in G$  we have  $\{X_t^0\}_{t \geq 0} \subset G$  and  $X_t^0 \xrightarrow[t \rightarrow \infty]{} a$ .

The Freidlin-Wentzell theory states that, under the stability and some other regularity assumptions on  $G$  and  $V$ , the so-called Kramers' type law for exit-time  $\tau_G^\sigma$  can be established. Namely, if we denote

$$H := \inf_{z \in \partial G} \{V(z) - V(a)\},$$

then for any starting point  $x_0 \in G$  and for any  $\delta > 0$

$$\lim_{\sigma \rightarrow 0} \mathbb{P}_{x_0} \left( e^{\frac{2(H-\delta)}{\sigma^2}} < \tau_G^\sigma < e^{\frac{2(H+\delta)}{\sigma^2}} \right) = 1. \quad (1.5)$$

It means that, with decreasing  $\sigma > 0$ , the exit-time from the stable by  $-\nabla V$  domain  $G$  grows exponentially with rate that depends on the height  $H$  of potential  $V$  inside the domain  $G$  (see fig. 1.2).

The exit-time problem takes central stage in describing metastable behaviour of the process. Namely, if  $V$  is a multi-well potential, with small  $\sigma$ , we expect the process to spend long time around a local minimum, before leaving the domain of attraction around it. Kramers' formula not only allows us to estimate the exit time from the domain of attraction, but also describe the behaviour of the process on the long time scale. If we consider the time  $t(\sigma)$  as a function of parameter  $\sigma$ , increasing with the latter converging to 0. The position  $X_{t(\sigma)}^\sigma$  of the diffusion (1.4) can be estimated by decomposition into cycles techniques that can be found in [FW98].

A similar question was posed by S. Herrmann, P. Imkeller, and D. Peithmann for McKean-Vlasov type diffusion in [HIP08]. In particular, they



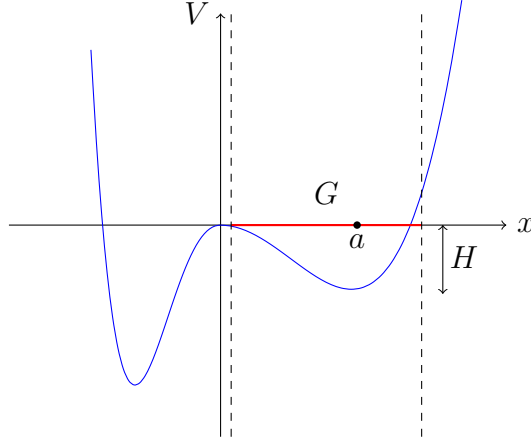


Figure 1.2: A double well potential  $V$  with a domain  $G$  that satisfies assumptions of the Freidlin-Wentzell theory along with  $H$ : the height of  $V$  inside  $G$ .

considered a self-stabilizing diffusion (SSD) of the form

$$\begin{cases} dX_t^\sigma = -\nabla V(X_t^\sigma) dt - \nabla F * \nu_t^\sigma(X_t^\sigma) dt + \sigma dW_t, \\ \nu_t^\sigma = \mathcal{L}(X_t^\sigma), \\ X_0^\sigma = x_0 \in \mathbb{R}^d \text{ a.s.}, \end{cases} \quad (1.6)$$

where  $\mathcal{L}$  denotes the law of random variable  $X_t^\sigma$ . Following the Freidlin-Wentzell theory techniques, in [HIP08] the authors first establish a large deviation principle for SSD with general assumptions on  $V$  and  $F$ . After that, they had to restrict themselves to the case of convex confinement and convex interaction potentials in order to achieve, under some stability and regularity assumptions on  $G$ , the Kramers' type law (1.5). Assuming  $F(0) = 0$ , the rate  $H$  in this case has the form  $H = \inf_{z \in \partial G} \{V(z) - V(a) + F(z - a)\}$ .

In [Tug12] and [Tug16], J. Tugaut studied the exit-time problem for SSD in convex landscape with convex interaction as it was presented in [HIP08]. In these papers he established the Kramers' type law avoiding using the large deviation principle. Instead, the author used various analytical and coupling methods to deal with the problem, thus simplifying calculations.

This work was continued in [Tug18a] and [Tug19] where the same techniques were used to establish the Kramers' type law in the case of confinement potential  $V$  that is of the double-well form. However, exit-time for the SSD with general assumptions (as in Freidlin-Wentzell theory for Itô diffusion) was an open problem, which we resolve in Chapter 4 of this thesis.

Exit-time problem was also studied for the case of self-interacting diffusion. In [AdMKT23], A. Aleksian, P. Del Moral, A. Kurtzmann, and J. Tugaut prove Kramers' type law for SID in which both interaction and confinement potentials  $V$  and  $F$  are convex. This nice property of potentials was used by the authors in order to prove that the occupation measure  $\mu_t^\sigma$  is close to  $\delta_a$  in some sense for  $\sigma$  small enough and for big enough  $t$ , where  $a$  is the unique point of attraction of the dynamical system (because of the convexity assumption, not only inside  $G$ , but globally). The exit-time result was obtained by applying analytical and coupling techniques similar to those used in [Tug16]. Chapter 2 is based on this paper.

For general confinement and interaction potentials, the exit-time problem was solved in [AKT23]. Without the convexity assumption neither control of the occupation measure can be easily established, nor coupling techniques of [Tug16, AdMKT23] can be applied. The authors had to use a different approach. In this case, it was proving the Large Deviation Principle and restoring the logic of Freidlin-Wentzell theory for SID. Chapter 3 is based on this work.

## 1.5 Outline of the thesis

This thesis combines results done by the author in collaboration with Pierre del Moral, Aline Kurtzmann and Julian Tugaut in [AdMKT23, AKT23, AT23]. Each following chapter contains results of corresponding paper.

Chapter 2 considers exit-problem for self-interacting diffusion in the case of convex confinement in interaction potentials. This particular problem is close to the one considered in [Tug16] for self-stabilizing diffusion, thus similar techniques were used in this chapter. First, we prove existence and uniqueness for the system that we consider there. After that, we recall results of [KK12] concerning convergence of the occupation measure  $\mu_t$  towards a Gibbs measure. We use this result in order to prove stabilization of

the occupation measure around  $\delta_a$  in finite time. This fact along with convexity of potentials  $V$  and  $F$  is used to obtain the coupling result between our system and associated Itô diffusion until exit time. This coupling is used to prove Kramers' type law for exit time for self-interacting diffusion that corresponds to the one of the associated Itô process. We also prove exit-location result for this diffusion. As a result, under assumptions of Chapter 2, we prove the following theorem.

**Theorem 1.6.** *Let  $\tau_G^\sigma := \inf \{t \geq 0 : X_t^\sigma \notin G\}$  be the exit-time and let exit-cost be defined as following:*

$$H := \inf_{x \in \partial G} (V(x) + F(x - a) - V(a)) .$$

*Then  $\mathbb{P}\text{-}\lim_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log(\tau_G^\sigma) = H$  that is for any  $\delta > 0$ , we have the Kramers' type law:*

$$\lim_{\sigma \rightarrow 0} \mathbb{P} \left( \exp \left\{ \frac{2}{\sigma^2} (H - \delta) \right\} \leq \tau \leq \exp \left\{ \frac{2}{\sigma^2} (H + \delta) \right\} \right) = 1 .$$

*Moreover, if  $\mathcal{N} \subseteq \partial G$  such that  $\inf_{z \in \mathcal{N}} (V(z) + F(z - a) - V(a)) > H$ , then*

$$\lim_{\sigma \rightarrow 0} \mathbb{P} (X_{\tau_G^\sigma} \in \mathcal{N}) = 0 .$$

Chapter 3 significantly improves this result by assuming confinement and interaction to be more general. In particular, we assume  $V$  and  $F$  to be regular enough with Lipschitz continuous gradients. This poses a lot of complications to the problem, since, in this case, we can not obtain the control of the occupation measure or even coupling with the associated Itô diffusion the same way. Instead, we prove the exit-time result by adapting the Freidlin-Wentzell theory for this system.

Our analysis starts with a “markovization” of the self-interacting diffusion. Even though the SID is not a Markov process, knowing only the past trajectory of the process, we can uniquely continue its subsequent path. The past trajectory can be described by the following triple  $(t_0, \mu_0, x_0)$ , where  $t_0$  is its time-length,  $\mu_0$  is the occupation measure and  $x_0$  is the last

point, while the subsequent path will be defined by the SDE:

$$\begin{cases} dX_t^\sigma &= -\nabla V(X_t^\sigma) dt - \nabla F * \mu_t^\sigma(X_t^\sigma) dt + \sigma dW_t, \\ \mu_t^\sigma &= \frac{t_0}{t_0+t} \mu_0 + \frac{1}{t_0+t} \int_0^t \delta_{X_s^\sigma} ds, \\ X_0^\sigma &= x_0 \text{ a.s.} \end{cases} \quad (1.7)$$

We approach this problem by first proving the large deviation principle for the generalized system above. We use the LDP in order to control the occupation measure until the exit time and restore the logic of Freidlin-Wentzell principle to prove the exit-time result. Namely, we consider when the diffusion comes close to the point of attraction  $a$  and when it deviates significantly away. We prove that the time that the process spends around the point  $a$  is significantly greater than the time spent far away from it, which gives the control of the occupation measure. In order to get the exit-time result, following the Freidlin-Wentzell approach, we once again consider when the process comes close to  $a$  and when it deviates from it. Given that we control the empirical measure of the process, every such deviation can be viewed as an independent exit attempt of the processes of the type (1.7) with  $\mu_0$  close to  $\delta_a$ . Just like in the case of Itô diffusion, since we consider the limit  $\sigma \rightarrow 0$ , the process will exit the domain  $G$  along the least improbable path, that is the one that minimises the quasi-potential. Since, as we pointed out above,  $\mu_t^\sigma \xrightarrow{\sigma \rightarrow 0} \delta_a$  for any  $t$  at least until the exit-time, with  $\sigma \rightarrow 0$ , the quasi-potential will become closer to the one of the following Itô process:

$$dX_t^\sigma = -\left(\nabla V(X_t^\sigma) + \nabla F(X_t^\sigma - a)\right) dt + \sigma dW_t,$$

which helps us to establish the Kramers' type law of the form

$$\lim_{\sigma \rightarrow 0} \mathbb{P} \left( \exp \left\{ \frac{2}{\sigma^2} (H - \delta) \right\} \leq \tau_G^\sigma \leq \exp \left\{ \frac{2}{\sigma^2} (H + \delta) \right\} \right) = 1,$$

where  $H := \inf_{x \in \partial G} (V(x) + F(x - a) - V(a))$ . Moreover, the exit-location result similar to the one of Theorem 1.6 was established.

Chapter 4 considers the exit-problem for the self-stabilizing diffusion with general confinement and interaction potentials. This process is very

similar to the self-interacting diffusion and, after controlling the law of the process, we could use the same technique as in Chapter 3 to establish the exit-time result. Instead, we succeeded in developing techniques to generalize the coupling method presented in [Tug16] and apply them to our non-convex case.

Similarly to the previous chapter, we first show the convergence in finite time of the process  $X^\sigma$  towards the point of attraction  $a$ . After this convergence, we introduce a synchronous coupling of  $X^\sigma$  with an Itô diffusion:

$$dY_t^\sigma = \sigma dW_t - (\nabla V(Y_t^\sigma) + \nabla F(Y_t^\sigma - a)) dt .$$

The idea of the method is based on the fact that, since the processes  $X^\sigma$  and  $Y^\sigma$  are coupled by the same Brownian motion, whenever both  $X^\sigma$  and  $Y^\sigma$  are close to the stable point of attraction  $a$  (in the small neighbourhood of which  $V + F(\cdot - a)$  is assumed to be convex), the distance between them decreases a.s. At the same time, whenever the two processes deviate from  $a$  but still stay inside the domain  $G$ , their maximum scatter can be controlled in terms of the time spent in the annulus confined between  $\partial G$  and a small neighbourhood around  $a$ . After that, we prove that the processes  $X^\sigma$  and  $Y^\sigma$  spend long enough time close to  $a$ , compared to the total time spent far from it, to the point that the attracting effect surpasses the scattering one. We use these facts to prove that for any  $\kappa > 0$ :

$$\lim_{\sigma \rightarrow 0} \mathbb{P}(\sup_t |X_t^\sigma - Y_t^\sigma| > \kappa) = 0,$$

where supremum is taken for  $t$  in a sufficiently wide range to prove the exit-time result.

This coupling allows us to control the law  $\mathcal{L}(X_t^\sigma)$  at least until the exit-time. Moreover, using the same logic as in [Tug16], we prove the Kramers' type law as well as the exit-location result for SSD.



# Chapter 2

## Exit-problem for self-interacting diffusion: Convex case

In this chapter, we are interested in a time-inhomogeneous diffusion. More precisely, we study the following specific diffusion, driven by the Stochastic Differential Equation (SDE):

$$dX_t = \sigma dW_t - \left( \nabla V(X_t) + \frac{1}{t} \int_0^t \nabla F(X_t - X_s) ds \right) dt, X_0 = x_0 \in \mathbb{R}^d \quad (2.1)$$

where  $V, F$  are two potentials on  $\mathbb{R}^d$  and  $\sigma > 0$ . The precise assumptions on the potentials will be given later in Section 2.1. We can already notice that the current position of the process  $X_t$  depends on the whole past trajectory of the process  $(X_s)_{0 \leq s \leq t}$  through the interaction potential  $F$  appearing in the drift term. We call this kind of process a path-interaction process.

### The questions that we study here

A large family of path-dependent processes has been studied by Saporito, see for instance [JS17]. He proves, with his co-authors, existence and uniqueness of such processes. The difference with the process studied here is that we normalize the occupation measure. In the current chapter, we also prove the existence and uniqueness result for general potentials  $V$  and  $F$ , which are not necessarily convex.

A second result that we are obtaining is related to the asymptotic behaviour. Indeed, after proving the existence and the uniqueness of the solution to Diffusion (2.1), we are studying the convergence in long-time of the probability measure  $\mu_t$ . The idea is similar to the one in [KK12].

The main question of this chapter is the exit-time problem of (2.1). We prove that the first exit-time  $\tau$  of the diffusion from some domain satisfies a Kramers' type law in the following sense:

$$\lim_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log(\tau) = H > 0,$$

where the convergence holds in probability.

We could adapt the techniques introduced by Herrmann, Imkeller and Peithmann but only in the case of a convex potential  $V$ . Our aim is to generalize the study also to non-convex potentials. In the present work, we will solve the exit-problem (time and location) for Diffusion (2.1). As will be shown later, the exit-location can be easily obtained once we know the asymptotics of the first exit-time.

## Outline

This chapter is divided into three parts. First, Section 2.1 is devoted to the explanation of the precise assumptions and the statement of the main results. After that, in Section 2.2, we prove Theorem 2.15 and Corollary 2.16, that is we establish the Kramers' type law and the exit-location result. To do so, we first provide some intermediate results. Finally, some possible extensions are discussed in the Appendix.

## Notations

The parameters  $\kappa$ ,  $\xi$ ,  $\epsilon$  and  $\delta$  are arbitrarily small. Other generic constants are denoted by  $C$ . As usual, we denote by  $\mathcal{M}(\mathbb{R}^d)$  the space of signed (bounded) Borel measures on  $\mathbb{R}^d$  and by  $\mathcal{P}(\mathbb{R}^d)$  its subspace of probability measures. In the following,  $\langle \cdot; \cdot \rangle$  stands for the Euclidean scalar product and  $|\cdot|$  is the associated norm.

We first introduce the notion of positively invariant domain.



**Definition 2.1.** Let  $\mathcal{G}$  be a subset of  $\mathbb{R}^d$  and let  $U : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a vector field satisfying some “good assumptions”. For all  $x \in \mathbb{R}^d$ , we consider the dynamical system  $\rho_t(x) = x + \int_0^t U(\rho_s(x)) ds$ . We say that the domain  $\mathcal{G}$  is positively invariant for the flow generated by  $U$  if the orbit  $\{\rho_t(x); t \in \mathbb{R}_+\}$  is included in  $\mathcal{G}$  for all  $x \in \mathcal{G}$ .

Let us recall the definition of the Wasserstein distance.

**Definition 2.2.** For  $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^d)$ , the quadratic Wasserstein distance is defined as

$$\mathbb{W}_2(\mu_1, \mu_2) := \left( \inf \{ \mathbb{E}(|Z_1 - Z_2|^2) \} \right)^{1/2},$$

where the infimum is taken over all the random variables such that  $\mathcal{L}(Z_1) = \mu_1$  and  $\mathcal{L}(Z_2) = \mu_2$ . This corresponds to the minimal  $L^2$ -distance taken over all the couplings between  $\mu_1$  and  $\mu_2$ .

Similarly, the Wasserstein distance  $\mathbb{W}_{2k}$  is defined as

$$\mathbb{W}_{2k}(\mu_1, \mu_2) := \left( \inf \{ \mathbb{E}(|Z_1 - Z_2|^{2k}) \} \right)^{1/(2k)}.$$

In the following, for readability issue, we will omit the  $\sigma$  in the superscript for the process  $X$  as well as the occupation measure  $\mu_t$ . Nevertheless, the reader has to keep in mind that the process  $(X_t, t \geq 0)$  and  $\mu_t$  do depend on  $\sigma$ .

**Definition 2.3.** The minimizer of  $V$  is denoted as  $a$ .

We also introduce the following mapping on the probability measures:

$$\Pi_\sigma(\mu)(dx) := \frac{e^{-\frac{2}{\sigma^2}(V(x)+F*\mu(x))}}{\int_{\mathbb{R}^n} e^{-\frac{2}{\sigma^2}(V(y)+F*\mu(y))} dy} dx.$$

## 2.1 Assumptions and main results

In this section we introduce the assumptions considered in the chapter.

**Assumption 2.4.** We assume some regularity for the potentials  $V$  and  $F$ :  $V \in \mathcal{C}^2(\mathbb{R}^d; \mathbb{R})$ ,  $F \in \mathcal{C}^2(\mathbb{R}^d; \mathbb{R})$ . Also, without loss of generality, we consider only potentials such that  $V \geq 0$ ,  $F \geq 0$ .

Assumption 2.4 is usual in SDE. Since we use Itô calculus techniques, the differentiability assumption above is necessary.

**Assumption 2.5.**  $V$  and  $F$  (and their first two derivatives) have at most a polynomial growth. In other words, there exists a polynomial function  $P$  of degree  $2k$  such that  $P(|x|) \geq 1$  for any  $x \in \mathbb{R}^n$ , and

$$|V(x)| + |F(x)| \leq P(|x|), \quad (2.2)$$

$$|\nabla V(x)| + |\nabla F(x)| \leq P(|x|), \quad (2.3)$$

$$\text{and } \|\nabla^2 V(x)\| + \|\nabla^2 F(x)\| \leq P(|x|). \quad (2.4)$$

This assumption is used in the paper [KK12] to establish the rate of convergence towards the invariant probability measure. We come back to this question on page 25.

**Remark 2.6.** Note, that without any loss of generality, we can choose polynomial  $P$  to be such that  $P(|x|) = C(1 + |x|^{2k})$ . Then, the following property holds: there exists a constant  $\gamma > 0$  such that  $P(|x + y|) \leq \gamma(P(|x|) + P(|y|))$ .

We also need the following assumption to establish the exit-time result:

**Assumption 2.7.** There exist  $\rho, \alpha > 0$  such that for any  $x \in \mathbb{R}^n$ , we have  $\nabla^2 V(x) \geq \rho \text{Id}$  and  $\nabla^2 F(x) \geq \alpha \text{Id}$ . The unique minimizer of  $V$  is denoted as  $a$  and the unique minimizer of  $F$  is 0. Moreover,  $F(0) = 0$

To ensure the existence of the process, we will also use the following assumption, that will help us to exhibit a Lyapunov function:

**Assumption 2.8.**  $\lim_{|x| \rightarrow \infty} V(x) = +\infty$ ,  $\lim_{|x| \rightarrow \infty} \frac{|\nabla V(x)|^2}{V(x)} = +\infty$  and there exists  $m > 0$  such that  $\Delta V(x) \leq mV(x)$  and  $\Delta F(x) \leq mF(x)$ .

**Remark 2.9.** By the latter growth condition,  $|\nabla V|^2 - \Delta V$  is bounded from below.

Finally, we will eventually assume the following, meaning that  $F$  is rotationally invariant:

**Assumption 2.10.** There exists a function  $G$  from  $\mathbb{R}_+$  to  $\mathbb{R}$  such that  $F(x) = G(|x|)$ .

The assumptions on the domain from which the diffusion  $X$  exits are the following.

**Assumption 2.11.** The domain  $\mathcal{D}$  is open and satisfies the following hypotheses:

1. Let  $\varphi$  be the solution to the following equation

$$\varphi_t = x_0 - \int_0^t \nabla V(\varphi_s) ds - \int_0^t \frac{1}{s} \int_0^s \nabla F(\varphi_s - \varphi_r) dr ds.$$

Then, for any  $t \in \mathbb{R}_+$ , we have  $\varphi_t \in \mathcal{D}$  and  $\lim_{t \rightarrow \infty} \varphi_t = a \in \mathcal{D}$ .

2. The domain  $\mathcal{D}$  is positively invariant for the flow generated by the vector field  $x \mapsto -\nabla V(x) - \nabla F * \delta_a(x)$ .
3. For any  $x \in \partial \mathcal{D}$ , define the flow  $\rho(x)$  as the solution to the equation

$$\rho_t(x) = x - \int_0^t \nabla V(\rho_s(x)) ds - \int_0^t \nabla F(\rho_s(x) - a) ds.$$

Moreover, assume that for any  $x \in \mathcal{D}$  the following limit holds:

$$\lim_{t \rightarrow \infty} \rho_t(x) = a.$$

The assumption 1) guarantees that, starting from fixed point  $x_0$ , the deterministic process defined by (2.1) with  $\sigma = 0$  converges towards the point of attraction  $a$ . Of course, we expect  $X_t$  to follow this path with high probability for small enough  $\sigma > 0$ . In fact, as will be shown later, its empirical measure  $\mu_t$  will also converge towards  $\delta_a$  and, after some deterministic time, with high probability, will stay inside a defined-in-advance neighbourhood of  $\delta_a$ . We call this effect: stabilisation of the empirical measure.

That leads to assumptions 2) and 3) above. After the “stabilisation time” we expect our drift term  $V + F * \mu_t$  to have a similar effect as  $V + F * \delta_a$ . Thus, the last two assumptions guarantee that the process  $X_t$  will forever tend to stay inside the domain  $\mathcal{D}$  and be attracted towards the point  $a$ . This is a necessary assumption when considering the exit from a stable domain of attraction solely under the influence of small noise.

## Existence and uniqueness

The first results that we will provide are about the existence and the uniqueness of the solution to the SDE (2.1).

**Theorem 2.1.** *Under the Assumptions 2.4 and 2.8, for any  $x_0 \in \mathbb{R}^n$ , there exists a unique global strong solution  $(X_t, t \geq 0)$  to Equation (2.1).*

*Proof.* Local existence and uniqueness of the solution to (2.1) is standard under the locally Lipschitz assumptions on the vector fields (see for instance [RW00, Theorem 13.1]). We only need to prove here that  $X$  does not explode in a finite time. Let us introduce the following increasing sequence of stopping times. Let  $\tau_0 = 0$  and

$$\tau_n := \inf \left\{ t \geq \tau_{n-1}; \mathcal{E}_t(X_t) + \int_0^t |\nabla \mathcal{E}_s(X_s)|^2 ds > n \right\},$$

where  $\mathcal{E}_t(X_t) := V(X_t) + \frac{1}{t} \int_0^t F(X_t - X_s) ds$ . In order to show that the solution never explodes, we use the Lyapunov functional  $(x, t) \mapsto \mathcal{E}_t(x)$ . As the process  $(t, x) \mapsto \mathcal{E}_t(x)$  is of class  $\mathcal{C}^2$  (in the space variable) and is a  $\mathcal{C}^1$ -semi-martingale (in the time variable), Itô-Ventzell formula applied to  $(x, t) \mapsto \mathcal{E}_{t \wedge \tau_n}(x)$  implies

$$\begin{aligned} \mathcal{E}_{t \wedge \tau_n}(X_{t \wedge \tau_n}) &= V(x_0) + \int_0^{t \wedge \tau_n} \langle \nabla \mathcal{E}_s(X_s), dW_s \rangle - \int_0^{t \wedge \tau_n} |\nabla \mathcal{E}_s(X_s)|^2 ds \\ &\quad + \frac{\sigma^2}{2} \int_0^{t \wedge \tau_n} \Delta \mathcal{E}_s(X_s) ds - \int_0^{t \wedge \tau_n} \int_0^s F(X_s - X_u) du \frac{ds}{s^2}. \end{aligned} \quad (2.5)$$

We note that  $\int_0^{t \wedge \tau_n} \langle \nabla \mathcal{E}_s(X_s), dW_s \rangle$  is a true martingale. By removing the negative terms, using the fact that  $F(0) = 0$  (Assumption 2.7) and by using Assumption 2.8, we get the following bound on the expectation of the Lyapunov functional:

$$\mathbb{E} \mathcal{E}_{t \wedge \tau_n}(X_{t \wedge \tau_n}) \leq V(x_0) + \frac{m\sigma^2}{2} \int_0^t \mathbb{E} \mathcal{E}_{s \wedge \tau_n}(X_{s \wedge \tau_n}) ds.$$

So, Gronwall's Lemma leads to:

$$\mathbb{E} V(X_{t \wedge \tau_n}) \leq \mathbb{E} \mathcal{E}_{t \wedge \tau_n}(X_{t \wedge \tau_n}) \leq V(x_0) e^{m\sigma^2 t/2}.$$

Since, by Assumption 2.8,  $\lim_{|x| \rightarrow \infty} V(x) = \infty$ , the process  $(X_t, t \geq 0)$  can not explode in a finite time, or else it leads to a contradiction with the inequality above. That proves existence of a unique global strong solution.  $\square$

## Convergence rate

In this section we show the long time behaviour for the empirical measure of the self-interacting diffusion in the convex landscape. This framework was considered in the paper of Kleptsyn and Kurtzmann [KK12]. More precisely, they have proved the following

**Theorem 2.12.** [KK12, Theorem 1.6, Theorem 1.12, Proposition 2.5] *Let  $X$  be the solution to the equation (2.1) with  $\sigma = \sqrt{2}$ . Suppose that  $V$  and  $F$  satisfy Assumptions 2.4–2.8. Then:*

1. *There exist  $\alpha, C > 0$  such that for any  $t \geq 0$  big enough:  $\mu_t \in K_{\alpha, C}$  a.s., where*

$$K_{\alpha, C} := \{ \mu \in \mathcal{P}(\mathbb{R}^d); \quad \forall R > 0, \mu(\{y; |y| > R\}) < Ce^{-\alpha R} \}.$$

2. *There exists a unique density  $\rho_\infty : \mathbb{R}^d \rightarrow \mathbb{R}_+$ , such that almost surely*

$$\mu_t = \frac{1}{t} \int_0^t \delta_{X_s} ds \xrightarrow[t \rightarrow +\infty]{*-weakly} \rho_\infty(x) dx.$$

3. *There exists a constant  $m > 0$  such that almost surely, for  $t \rightarrow 0$ , one has the following asymptotic behaviour*

$$\mathbb{W}_2(\mu_t, \rho_\infty) = O(\exp\{-m^{2k+1}\sqrt{\log t}\}),$$

where  $2k$  is the degree of the polynomial  $P$  and  $\mathbb{W}_2$  is the quadratic Wasserstein distance.

Moreover, it was proved that, if  $V$  is symmetric with respect to some point  $q$ , then the corresponding density  $\rho_\infty$  is also symmetric with respect to the same point  $q$ . The authors showed that the density  $\rho_\infty$  is the same limit density as in the result of [CMV03], uniquely defined by the following property:  $\rho_\infty$  is a positive function, proportional to  $e^{-(V+F*\rho_\infty)}$ .

We note that the results of [KK12] were established for the case  $\sigma = \sqrt{2}$ . Nevertheless, one can check that each step of the proof can be reformulated with  $\sigma$ . Moreover, all the asymptotic (with respect to time  $t$ ) results for small  $\sigma$  can be upper bounded without loss of generality by the case of

a constant  $\sigma$ , that one can take to be equal to  $\sqrt{2}$ . That means that for smaller  $\sigma > 0$  we have faster convergence towards the invariant probability measure  $\rho_\infty$ .

We put together the observations above in the form of the following two results.

**Proposition 2.13.** *Under Assumptions 2.4–2.8, there exist  $\alpha, C > 0$ , such that for any  $t \geq 0$  and for any  $\sigma > 0$  small enough  $\mu_t \in K_{\alpha, C}$  a.s.*

**Proposition 2.14.** *Under Assumptions 2.4–2.8, there exists a constant  $m > 0$  such that for any  $x \in \mathbb{R}^n$ , almost surely, the following asymptotics holds:*

$$\mathbb{W}_{2k}(\mu_t, \rho_\infty) = O\left(\exp\{-m(\log t)^{\frac{1}{2k+1}}\}\right),$$

where  $\rho_\infty$  is the unique probability measure such that  $\rho_\infty = \Pi_\sigma(\rho_\infty)$ .

We stress that the convergence rate that we establish in Proposition 2.14 does not depend on  $\sigma$ .

## Main result on exit-problem

The main goal of this chapter is to find some precise upper and lower bounds for the exit-time from some positively invariant domain.

**Theorem 2.15.** *We assume that the potentials  $V, F$  and an open domain  $\mathcal{D}$  satisfy the Assumptions 2.4–2.11. By  $\tau := \inf\{t \geq 0 : X_t \notin \mathcal{D}\}$ , we denote the first time the process  $X$  exits the domain  $\mathcal{D}$ . We introduce the so-called exit-cost:*

$$H := \inf_{x \in \partial\mathcal{D}} (V(x) + F(x - a) - V(a)). \quad (2.6)$$

Then  $\mathbb{P} - \lim_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log(\tau) = H$  that is for any  $\delta > 0$ , we have

$$\lim_{\sigma \rightarrow 0} \mathbb{P} \left( \exp\left\{\frac{2}{\sigma^2}(H - \delta)\right\} \leq \tau \leq \exp\left\{\frac{2}{\sigma^2}(H + \delta)\right\} \right) = 1. \quad (2.7)$$

This statement about the exit-time corresponds to what we denote as the Kramers' type law.

From Theorem 2.15, we immediately obtain the classical statement on the exit-location.

**Corollary 2.16.** Under the same assumptions as the ones of Theorem 2.15, if  $\mathcal{N}$  is a subset of  $\partial\mathcal{D}$  such that  $\inf_{z \in \mathcal{N}} (V(z) + F(z - a) - V(a)) > H$ , then

$$\lim_{\sigma \rightarrow 0} \mathbb{P}(X_\tau \in \mathcal{N}) = 0. \quad (2.8)$$

This means that the diffusion avoids to exit from a part of the boundary where the cost of exiting exceeds the exit-cost of  $\mathcal{D}$ .

## 2.2 Exit-problem

In this section, we prove our main result. First, we give the necessary intermediate results on page 27. More precisely, we show that there exists a time of stabilisation around  $\delta_a$  for the occupation measure, in terms of Wasserstein distance. Then, we show that the process  $X$  solution to (2.1) is close to the solution of the deterministic flow  $(\varphi_t)_{t \geq 0}$ . Using that, we prove in Corollary 2.20 that the probability of leaving a positively invariant domain before the occupation measure remains stuck in the ball of center  $\delta_a$  and radius  $\kappa$  for  $\mathbb{W}_{2k}$  vanishes as  $\sigma$  goes to zero. Then, we consider the coupling between the studied diffusion and the one where the occupation measure is frozen to  $\delta_a$  and we show that these diffusions are close.

After, we provide the proof of Theorem 2.15 then we give the proofs of the intermediate results. Finally, we apply Theorem 2.15 to level sets so that we are in position to prove the exit-location result on page 38.

### Intermediate results

We first introduce a deterministic time, representing the time of stabilisation of the occupation measure, if it occurs, around its supposed limit  $\delta_a$ :

**Definition 2.17.** For any  $\sigma > 0$  and for any  $\kappa > 0$ , we introduce:

$$T_\kappa(\sigma) := \inf \left\{ t_0 \geq 0 : \forall t \geq t_0, \mathbb{E}(\mathbb{W}_{2k}(\mu_t; \delta_a)) \leq \kappa \right\}, \quad (2.9)$$

where  $2k$  is the degree of the polynomial function  $P$ , introduced in Assumption 2.5.

**Proposition 2.18.** For any  $\sigma, \kappa > 0$ , the time  $T_\kappa(\sigma)$  is finite. Moreover, for any  $\kappa > 0$ , there exists  $T_\kappa > 0$  such that

$$\sup_{0 < \sigma < 1} T_\kappa(\sigma) \leq T_\kappa.$$

The proof of Proposition 2.18 is postponed to page 32.

Next, we show that the probability for the process  $X$  to exit from  $\mathcal{D}$  before the time  $T_\kappa(\sigma)$  tends to 0 as  $\sigma$  goes to 0.

We remind the reader that in this work, the noise vanishes. Consequently, it is natural to introduce the deterministic flow  $(\varphi_t)_{t \geq 0}$  defined by the following zero-noise process

$$\dot{\varphi}_t = -\nabla V(\varphi_t) - \frac{1}{t} \int_0^t \nabla F(\varphi_t - \varphi_s) ds, \quad \varphi_0 = x_0. \quad (2.10)$$

We will state that for any  $T > 0$ ,  $(X_t, 0 \leq t \leq T)$  and  $(\varphi_t, 0 \leq t \leq T)$  are uniformly close while the noise goes to zero. Namely,

**Proposition 2.19.** *We assume that the potentials  $V$ ,  $F$  and an open domain  $\mathcal{D}$  satisfy the Assumptions 2.4–2.11. Then, for any  $\xi > 0$  and for any  $T > 0$ , we have:*

$$\lim_{\sigma \rightarrow 0} \mathbb{P} \left( \sup_{t \in [0; T]} |X_t - \varphi_t(x_0)|^2 > \xi \right) = 0. \quad (2.11)$$

The proof of Proposition 2.19 is postponed to page 33. We deduce immediately the following.

**Corollary 2.20.** We assume that the potentials  $V$ ,  $F$  and an open domain  $\mathcal{D}$  satisfy the Assumptions 2.4–2.11. Then:

$$\lim_{\sigma \rightarrow 0} \mathbb{P}(\tau \leq T_\kappa(\sigma)) = 0. \quad (2.12)$$



*Proof.* It is sufficient to consider  $T := T_\kappa$  and  $\xi := \inf_{t \in [0; T_\kappa]} d(\overline{\mathcal{D}}; \varphi_t) > 0$  in Proposition 2.19. We thus have:

$$\begin{aligned} \mathbb{P}(\tau \leq T_\kappa(\sigma)) &= \mathbb{P}\left(\inf \left\{ d(\overline{\mathcal{D}}; X_t) : t \in [0; T_\kappa(\sigma)] \right\} = 0\right) \\ &\leq \mathbb{P}\left(\sup_{t \in [0; T_\kappa(\sigma)]} |X_t - \varphi_t| > \xi\right), \end{aligned}$$

which converges towards 0 as noise vanishes.  $\square$

In [Tug16, Tug18a], Tugaut has proved the Kramers' type law for the exit-time. He has used a coupling between the diffusion of interest ( $X$  here) and another diffusion that is expected to be close to  $X$  if the time is sufficiently large. The main difficulty with the considered self-stabilizing diffusion is in fact that we do not have a uniform (with respect to time) control of the law.

Here, we have proved that the nonlinear quantity appearing in (2.1) (that is  $\frac{1}{t} \int_0^t \delta_{X_s} ds$ ) remains stuck - with high probability - in a small ball (for the  $\mathbb{W}_{2k}$ -distance) of center  $\delta_a$  and radius  $\kappa$  for any  $t \geq T_\kappa(\sigma)$ . The idea is thus to substitute  $\frac{1}{t} \int_0^t \delta_{X_s} ds$  by  $\delta_a$  and to compare the new diffusion with the initial one.

We introduce the diffusion  $(Y_t)_{t \geq 0}$  such that  $Y_t = X_t$  if  $t \leq T_\kappa(\sigma)$  and for any  $t \geq T_\kappa(\sigma)$

$$dY_t = \sigma dW_t - \nabla V(Y_t) dt - \nabla F(Y_t - a) dt. \quad (2.13)$$

**Proposition 2.21.** *We assume that the potentials  $V$ ,  $F$  and an open domain  $\mathcal{D}$  satisfy the Assumptions 2.4–2.11. Then, if  $\kappa$  is small enough, we have for any  $\xi > 0$*

$$\limsup_{\sigma \rightarrow 0} \mathbb{P}\left(\sup_{T_\kappa(\sigma) \leq t \leq \exp\left[\frac{2H+10}{\sigma^2}\right]} |X_t - Y_t| \geq \xi\right) \leq \kappa^k, \quad (2.14)$$

where we remind that  $2k$  is the degree of the polynomial function  $P$  introduced in Assumption 2.5.

The proof of Proposition 2.21 is postponed to page 36.

### Idea of the proof concerning the exit-time

The idea of the proof is to use the fact that diffusions  $Y$  and  $X$  are close to each other at least after the deterministic stabilisation time  $T_\kappa(\sigma)$  and until some fixed deterministic time  $\exp\left\{\frac{2(H+5)}{\sigma^2}\right\}$ . We choose this time to be sufficiently big for our line of reasoning. We can control the proximity of these two diffusions by parameter  $\kappa$ , which represents how close the empirical measure  $\mu_t$  and  $\delta_a$  are. It was already shown in Corollary 2.20, that, with  $\sigma \rightarrow 0$ , the probability of exiting before time  $T_\kappa(\sigma)$  tends to zero. That means that we can focus on the dynamics after the stabilisation of the occupation measure happens. For the upper bound, we show that the event  $\left\{\tau > \exp\left[\frac{2(H+\delta)}{\sigma^2}\right]\right\}$  is unlikely due to the fact that, for small  $\sigma > 0$ , the diffusion  $Y$  leaves a bigger (than  $\mathcal{D}$ ) domain before the time  $\exp\left[\frac{2(H+\delta)}{\sigma^2}\right]$ , which, given the closeness of  $X$  and  $Y$ , gives  $\left\{\tau \leq \exp\left[\frac{2(H+\delta)}{\sigma^2}\right]\right\}$ . Same type of reasoning is used to prove the lower bound  $\exp\left\{\frac{2(H-\delta)}{\sigma^2}\right\}$ .

The proof of the intermediate lemmas are given in the respective section on page 32.

### Proof for the exit-time result

Fix some  $\delta, \kappa > 0$ , decrease it if necessary to be  $\delta < 5$ . For the upper bound, consider the following inequality:

$$\mathbb{P}\left(\tau > e^{\frac{2(H+\delta)}{\sigma^2}}\right) \leq \mathbb{P}\left(\tau > e^{\frac{2(H+\delta)}{\sigma^2}}, \tau_{\mathcal{D}^e}^Y \leq e^{\frac{2(H+\delta)}{\sigma^2}}\right) + \mathbb{P}\left(\tau_{\mathcal{D}^e}^Y > e^{\frac{2(H+\delta)}{\sigma^2}}\right), \quad (2.15)$$

where  $\mathcal{D}^e$  is some enlargement of domain  $\mathcal{D}$  such that its exit-cost is equal to  $H + \frac{\delta}{2}$ , i.e.:

$$\mathcal{D}^e := \left\{x \in \mathbb{R}^d : V(x) + F(x - a) - V(a) < H + \frac{\delta}{2}\right\};$$

and  $\tau_{\mathcal{D}^e}^Y$  is the first exit-time of diffusion  $Y$  from this domain, i.e.:

$$\tau_{\mathcal{D}^e}^Y := \inf\{t : Y_t \notin \mathcal{D}^e\}.$$

Note, that domain  $\mathcal{D}^e$  (since both  $V$  and  $F$  are continuous and convex) satisfies the usual assumptions (see [DZ10]) and  $d_e := d(\mathcal{D}, \partial\mathcal{D}^e) > 0$ . By

classical result of Freidlin-Wentzell theory,

$$\mathbb{P} \left( \tau_{\mathcal{D}^e}^Y > \exp \left[ \frac{2((H + \delta/2) + \delta/2)}{\sigma^2} \right] \right) \xrightarrow{\sigma \rightarrow 0} 0.$$

Let us decrease  $\sigma_\kappa$  if necessary, such that the quantity above will be less than  $\sqrt{\kappa}$  for any  $\sigma < \sigma_\kappa$ . Moreover, the first probability in (2.15) can be bounded by:

$$\mathbb{P}(\tau_{\mathcal{D}^e}^Y \leq e^{\frac{2(H+\delta)}{\sigma^2}} < \tau) \leq \mathbb{P}(|X_{\tau_{\mathcal{D}^e}^Y} - Y_{\tau_{\mathcal{D}^e}^Y}| \geq d_e) \leq 2\kappa^k,$$

where we use Proposition 2.21 and decrease  $\kappa$  and  $\sigma_\kappa$  if necessary.

We approach the lower bound similarly and introduce the contraction of the domain  $\mathcal{D}$ :

$$\mathcal{D}^c := \left\{ x \in \mathbb{R}^d : V(x) + F(x - a) - V(a) < H - \frac{\delta}{2} \right\} \subset \mathcal{D}.$$

If  $\mathcal{D}^c$  turns out to be empty, decrease  $\delta$ . As previously, the domain  $\mathcal{D}^c$  satisfies usual properties and has positive distance with the boundary of the initial domain, that is  $d_c := d(\mathcal{D}^c, \partial\mathcal{D}) > 0$ . We introduce the exit-time from the contracted domain for diffusion  $Y$ :

$$\tau_{\mathcal{D}^c}^Y := \inf \{ t : Y_t \notin \mathcal{D}^c \},$$

and have the following estimate:

$$\begin{aligned} \mathbb{P}(\tau < e^{\frac{2(H-\delta)}{\sigma^2}}) &\leq \mathbb{P}(T_\kappa(\sigma) < \tau < e^{\frac{2(H-\delta)}{\sigma^2}} \leq \tau_{\mathcal{D}^c}^Y) \\ &\quad + \mathbb{P}(\tau \leq T_\kappa(\sigma)) + \mathbb{P}(\tau_{\mathcal{D}^c}^Y \geq e^{\frac{2((H-\delta/2)-\delta/2)}{\sigma^2}}) \\ &\leq \mathbb{P}(|X_\tau - Y_\tau| \geq d_c) + \mathbb{P}(\tau \leq T_\kappa(\sigma)) + 2\kappa^k \\ &\leq 3\kappa^k + \mathbb{P}(\tau \leq T_\kappa(\sigma)) \\ &\leq 4\kappa^k, \end{aligned}$$

by Corollary 2.20, with  $\kappa$  and  $\sigma_\kappa$  small enough. This leads to:

$$\mathbb{P}(e^{\frac{2(H-\delta)}{\sigma^2}} \leq \tau \leq e^{\frac{2(H+\delta)}{\sigma^2}}) \geq 1 - 7\kappa^k,$$

which proves the theorem if we consider  $\kappa \rightarrow 0$ , parameter that uniformly controls the convergence of  $\sigma$  towards 0.

## Proof of the intermediate results

Several propositions are proven here.

### Proof of Proposition 2.18

In the following, we remind the reader that we do not emphasize the dependence on  $\sigma$ , but it will appear everywhere in the computations.

1. As was mentioned above, the invariant probability measure of self-interacting diffusion (and, at the same time, the weak-\* limit of its empirical measure a.s.) is the unique solution to the equation

$$\mu_\infty = \Pi_\sigma(\mu_\infty),$$

where  $\Pi_\sigma$  is defined as:

$$\Pi_\sigma(\mu)(x) = e^{-2(V+F*\mu)(x)/\sigma^2} / \int e^{-2(V+F*\mu)(z)/\sigma^2} dz.$$

The same invariant probability measure appears in the self-stabilizing diffusion, small-noise limit of which was studied in [HT10]. There, authors studied the case of double-wells potentials which is more general than our diffusion. In this paper the result, that can be transformed in our context as following, was proved. If the moments of invariant probability measures  $\mu_\infty$  are uniformly bounded with respect to  $\sigma$ , then  $\delta_a$  is the weak-\* limit of  $\mu_\infty$  with  $\sigma \rightarrow 0$  a.s. Note, that indeed, moments of  $\mu_t$  are uniformly bounded for any  $t > 0$ . Indeed, this is due to the fact that  $\mu_t \in K_{\alpha,C}$  for any  $t > 0$  and for some  $\alpha, C$  that do not depend on  $\sigma$  (Proposition 2.13). It proves that

$$\mu_\infty \xrightarrow[\sigma \rightarrow 0]{\text{weak-*}} \delta_a \text{ a.s.}$$

2. Let us consider the expectation  $\mathbb{E}(\mathbb{W}_{2k}(\mu_t; \delta_a))$ . First, let us show its existence. To do that, we use the fact that for any  $t > 0$ ,  $\mu_t \in K_{\alpha,C}$  almost surely and get

$$\mathbb{W}_{2k}(\mu_t; \delta_a) \leq \left( 2^{2k-1} \int |x|^{2k} \mu_t(dx) + 2^{2k-1} |a|^{2k} \right)^{1/(2k)} \leq \text{Const},$$

where the last constant depends only on  $\alpha, C, a$  and  $k$ . Therefore, since the random variable is bounded by a constant almost surely, expectation exists.

3. Now, we can finish the proof by separating the expectation of the distance between  $\mu_t$  and  $\delta_a$  into two parts and find the limit:

$$\mathbb{E}(\mathbb{W}_{2k}(\mu_t; \delta_a)) \leq \mathbb{E}(\mathbb{W}_{2k}(\mu_t; \mu_\infty)) + \mathbb{E}(\mathbb{W}_{2k}(\mu_\infty; \delta_a)) \xrightarrow[\sigma \rightarrow 0]{t \rightarrow \infty} 0,$$

where the limit is not just iterated, but holds for the pair  $(t, \sigma)$ , since the rate of convergence of  $\mu_t$  towards  $\mu_\infty$  in time does not depend on  $\sigma$ , which was shown in Proposition 2.14. Therefore, for any  $\kappa > 0$  we can find  $\sigma_0$  small enough and  $t_0$  big enough such that  $T_\kappa(\sigma) < T_\kappa < \infty$  for any  $\sigma < \sigma_0$ , which does not only prove existence and finiteness of  $T_\kappa(\sigma)$ , but also its uniformness with respect to  $\sigma$ .

### Proof of Proposition 2.19

First of all, we fix some  $\xi$  and introduce the following stopping time  $\mathcal{T} := \inf\{t : |X_t^\sigma - \varphi_t|^2 \geq \xi\}$ . We apply Itô formula and get the following result, for  $\omega \in \{\mathcal{T} > t\}$  (the choice of this event will be clear further) :

$$\begin{aligned} |X_t - \varphi_t|^2 &= 2 \int_0^t \langle X_s - \varphi_s, dX_s - d\varphi_s \rangle + d\sigma^2 t \\ &\leq d\sigma^2 t - 2 \int_0^t \langle X_s - \varphi_s; \nabla V(X_s) - \nabla V(\varphi_s) \rangle ds \\ &\quad - \int_0^t \frac{2}{s} \int_0^s \langle X_s - \varphi_s, \nabla F(X_s - X_z) - \nabla F(\varphi_s - \varphi_z) \rangle dz ds \\ &\quad + 2\sigma \int_0^t \langle X_s - \varphi_s, dW_s \rangle. \end{aligned}$$

Let  $\text{Lip}_{\nabla F}^{K'}$  be a Lipschitz constant of  $\nabla F$  inside the following compact

$$K' := \{x : |x - \varphi_t|^2 \leq \xi, \text{ for some } t > 0\}.$$

Due to our assumptions, this set is indeed compact at least for small  $\xi$ , which we can decrease without any loss of generality. We remind that  $\rho$  is

the convexity constant of  $V$ . We thus have

$$\begin{aligned}
|X_t - \varphi_t|^2 &\leq d\sigma^2 t - 2\rho \int_0^t |X_s - \varphi_s|^2 ds + 2\sigma \int_0^t \langle X_s - \varphi_s, dW_s \rangle \\
&\quad + \text{Lip}_{\nabla W}^{K'} \int_0^t \frac{2}{s} \int_0^s \langle |X_s - \varphi_s|^2 + |X_s - \varphi_s| \cdot |X_z - \varphi_z| \rangle dz \\
&\leq d\sigma^2 t - 2\rho \int_0^t |X_s - \varphi_s|^2 ds + 2\sigma \int_0^t \langle X_s - \varphi_s, dW_s \rangle \quad (2.16) \\
&\quad + \text{Lip}_{\nabla W}^{K'} \int_0^t \frac{1}{s} \int_0^s (3|X_s - \varphi_s|^2 + |X_z - \varphi_z|^2) dz.
\end{aligned}$$

Note then that by the Burkholder-Davis-Gundy inequality, we get for some constant  $C > 0$ :

$$\begin{aligned}
\mathbb{E} \left( \sup_{[0, t \wedge \mathcal{T}]} \left| 2\sigma \int_0^s \langle X_z - \varphi_z, dW_z \rangle \right| \right) &\leq C\sigma^2 \mathbb{E} \sqrt{\int_0^{t \wedge \mathcal{T}} |X_s - \varphi_s|^2 ds} \\
&\leq C\sigma^2 \sqrt{\int_0^t \mathbb{E} \left( \sup_{z \in [0, s \wedge \mathcal{T}]} (|X_z - \varphi_z|^2) \right) ds}.
\end{aligned}$$

Let us consider the following random variable:  $\sup_{s \in [0; t \wedge \mathcal{T}]} |X_s - \varphi_s|^2$ . The fact that we consider the supremum before time  $t \wedge \mathcal{T}$  gives us that for any  $\omega$  we consider only such  $s$ , that  $s \leq \mathcal{T}(\omega)$ , which in turn means that we can apply estimation (2.16) for any  $s \in [0, t \wedge \mathcal{T}]$ . We also recall that  $t \leq T$  and derive:

$$\begin{aligned}
\mathbb{E} \left( \sup_{s \in [0; t \wedge \mathcal{T}]} |X_s - \varphi_s|^2 \right) &\leq d\sigma^2 T + C\sigma^2 \sqrt{\int_0^t \mathbb{E} \left( \sup_{z \in [0, s \wedge \mathcal{T}]} (|X_z - \varphi_z|^2) \right) dz} \\
&\quad + 4\text{Lip}_{\nabla W}^{K'} \int_0^t \mathbb{E} \left( \sup_{z \in [0, s \wedge \mathcal{T}]} |X_z - \varphi_z|^2 \right) ds,
\end{aligned}$$

which is bounded by

$$\begin{aligned}
\mathbb{E} \left( \sup_{s \in [0; t \wedge \mathcal{T}]} |X_s - \varphi_s|^2 \right) &\leq d\sigma^2 T + \frac{C\sigma^2}{2} \left[ 1 + T \mathbb{E} \left( \sup_{s \in [0, t \wedge \mathcal{T}]} (|X_s - \varphi_s|^2) \right) \right] \\
&\quad + 4\text{Lip}_{\nabla W}^{K'} \int_0^t \mathbb{E} \left( \sup_{z \in [0, s \wedge \mathcal{T}]} |X_z - \varphi_z|^2 \right) ds,
\end{aligned}$$

where in the last inequality we used  $\sqrt{x} \leq (1+x)/2$ . Now, if we denote  $u_t := \mathbb{E} \left( \sup_{s \in [0; t \wedge \mathcal{T}]} |X_s - \varphi_s|^2 \right)$ , we have

$$u_t \leq \frac{1}{1 - CT\sigma^2/2} \left( \frac{2Td + C}{2} \sigma^2 + 4\text{Lip}_{\nabla W}^{K'} \int_0^t u_s ds \right),$$

for small enough  $\sigma$  (such that  $1 - CT\sigma^2/2 > 0$ ). Thus, using Grönwall lemma, we get

$$u_t \leq \frac{(2Td + C)\sigma^2}{2(1 - CT\sigma^2/2)} \exp \left\{ \frac{4\text{Lip}_{\nabla W}^{K'}}{1 - CT\sigma^2/2} T \right\} = O(\sigma^2). \quad (2.17)$$

This in particular means, that  $\mathbb{E} \left( \sup_{s \in [0; T \wedge \mathcal{T}]} |X_s - \varphi_s|^2 \right) \leq O(\sigma^2)$ . Nevertheless, to show the necessary result, we have to get rid of the stopping time  $\mathcal{T}$  in the previous equation. It is sufficient to show, that  $\mathbb{P}(\mathcal{T} \leq T) \xrightarrow{\sigma \rightarrow 0} 0$ .

Indeed, by its definition,  $\mathcal{T}$  is the first time when the difference  $|X_t - \varphi_t|^2$  reaches  $\xi$ . But under the assumption  $\mathcal{T} \leq T$  and due to (2.17), by decreasing  $\sigma$  we can control  $|X_t - \varphi_t|^2$  and make it small enough, such that  $|X_{\mathcal{T}} - \varphi_{\mathcal{T}}|^2 < \xi$  (in some sense), which contradicts the definition of  $\mathcal{T}$ . Rigorously,

$$\mathcal{T} < T \Rightarrow \sup_{[0, \mathcal{T} \wedge T]} |X_s - \varphi_s|^2 = \sup_{[0, \mathcal{T}]} |X_s - \varphi_s|^2 \geq \xi.$$

Thereby,

$$\mathbb{P}(\mathcal{T} < T) \leq \mathbb{P} \left( \sup_{[0, \mathcal{T} \wedge T]} |X_s - \varphi_s|^2 \geq \xi \right) \leq O(\sigma^2),$$

by Markov inequality.

To conclude the proof of Proposition 2.19, we consider

$$\begin{aligned} \mathbb{P} \left( \sup_{t \in [0; T]} |X_t - \varphi_t(x_0)|^2 > \xi \right) &\leq \mathbb{P} \left( \sup_{t \in [0; T]} |X_t - \varphi_t(x_0)|^2 > \xi, \mathcal{T} > T \right) \\ &\quad + \mathbb{P}(\mathcal{T} \leq T) \\ &\leq \mathbb{P} \left( \sup_{t \in [0; T \wedge \mathcal{T}]} |X_t - \varphi_t(x_0)|^2 > \xi \right) + O(\sigma^2) \\ &\leq O(\sigma^2), \end{aligned}$$

by Markov inequality and (2.17), which completes the proof.

**Proof of Proposition 2.21**

Let us define

$$W_a(x) := V(x) + F(x - a) \quad \text{and} \quad W_{\mu_t}(x) := V(x) + F * \mu_t(x),$$

with the occupation measure  $\mu_t := \frac{1}{t} \int_0^t \delta_{X_s} ds$ . For any  $t \geq T_\kappa(\sigma)$ , we have

$$d|X_t - Y_t|^2 = -2 \langle X_t - Y_t; \nabla W_{\mu_t}(X_t) - \nabla W_a(Y_t) \rangle dt.$$

We thus have

$$\begin{aligned} \frac{d}{dt} |X_t - Y_t|^2 &= -2 \langle X_t - Y_t; \nabla W_{\mu_t}(X_t) - \nabla W_{\mu_t}(Y_t) \rangle \\ &\quad + 2 \langle X_t - Y_t; \nabla F(Y_t - a) - \nabla F * \mu_t(Y_t) \rangle. \end{aligned}$$

However,  $\nabla^2 W_{\mu_t} = \nabla^2 V + \nabla^2 F * \mu_t \geq (\rho + \alpha) \text{Id}$  with  $\rho + \alpha > 0$ . So, putting  $\gamma(t) := |X_t - Y_t|^2$ , Cauchy-Schwarz inequality leads to

$$\gamma'(t) \leq -2(\alpha + \rho)\gamma(t) + 2\sqrt{\gamma(t)} |\nabla F(Y_t - a) - \nabla F * \mu_t(Y_t)|.$$

However, by the growth condition (2.2) on  $F$ , we have for any probability measures  $\mu, \nu$  the following control

$$|\nabla F * \mu(x) - \nabla F * \nu(x)| \leq C(1 + |x|^{2k}) \mathbb{W}_{2k}^{2k}(\mu, \nu)$$

where  $2k$  is the degree of the polynomial  $P$  introduced in Assumption 2.5. We introduce the set

$$\mathcal{A}_\kappa := \{\omega \in \Omega : \mathbb{W}_{2k}^{2k}(\mu_t, \delta_a) \leq \kappa^k\}.$$

By Markov inequality, we have  $\mathbb{P}(\mathcal{A}_\kappa^c) \leq 7\kappa^k$  then  $\mathbb{P}(\mathcal{A}_\kappa) \geq 1 - \kappa^k$ . This implies for any  $t \geq T_\kappa(\sigma)$  and for any  $\omega \in \mathcal{A}_\kappa$ :

$$\gamma'(t) \leq -2(\alpha + \rho)\gamma(t) + 2C\kappa^k \sqrt{\gamma(t)} (1 + |Y_t|^{2k}).$$

However,  $\gamma(t) = 0$  for any  $t \leq T_\kappa(\sigma)$ . This means that

$$\left\{ t \geq 0 : \gamma(t) > \frac{C^2 \kappa^{2k} (1 + |Y_t|^{2k})^2}{(\alpha + \rho)^2} \right\} \subset \{t \geq 0 : \gamma'(t) < 0\}.$$



By [BRTV98, Lemma 3.7], we deduce that

$$\sup_{T_\kappa(\sigma) \leq t \leq \exp\left[\frac{2H+10}{\sigma^2}\right]} \gamma(t) \leq \frac{C^2 \kappa^{2k} \left(1 + \sup_{T_\kappa(\sigma) \leq t \leq \exp\left[\frac{2H+10}{\sigma^2}\right]} |Y_t|^{2k}\right)^2}{(\alpha + \rho)^2},$$

if  $\omega \in \mathcal{A}_\kappa$ . We now consider  $R > 0$  such that the exit-cost of the diffusion  $Y$  from the ball of center  $a$  and radius  $R$  is at least  $H + 6$ , meaning that

$$\inf\{V(x) + F(x - a) - V(a) : x \in B(a, R)\} \geq H + 6.$$

Then, by Freidlin-Wentzell theory, we deduce that

$$\lim_{\sigma \rightarrow 0} \mathbb{P} \left( \sup_{T_\kappa(\sigma) \leq t \leq \exp\left[\frac{2H+10}{\sigma^2}\right]} |Y_t - a| \geq R \right) = 0.$$

However, we have

$$\begin{aligned} & \mathbb{P} \left( \sup_{T_\kappa(\sigma) \leq t \leq \exp\left[\frac{2H+10}{\sigma^2}\right]} |X_t - Y_t| \geq \xi \right) \\ & \leq \mathbb{P} \left( \sup_{T_\kappa(\sigma) \leq t \leq \exp\left[\frac{2H+10}{\sigma^2}\right]} |Y_t - a| \geq R \right) \\ & + \mathbb{P} \left( \sup_{T_\kappa(\sigma) \leq t \leq \exp\left[\frac{2H+10}{\sigma^2}\right]} \gamma(t) \geq \xi^2, \sup_{T_\kappa(\sigma) \leq t \leq \exp\left[\frac{2H+10}{\sigma^2}\right]} |Y_t - a| < R \right) \\ & + \mathbb{P}(\mathcal{A}_\kappa^c). \end{aligned}$$

The first term tends to 0 as  $\sigma$  goes to 0. The second term is equal to 0 provided that

$$\xi > \frac{C \kappa^k (1 + 2^{2k-1} (R + |a|^{2k}))}{\alpha + \rho}.$$

In other words, if  $\kappa$  is small enough, the second term is equal to 0 uniformly with respect to  $\sigma$ . The third term is less than  $\sqrt{\kappa}$ . This concludes the proof.

### Proof for the exit-location result

We can apply Theorem 2.15 to the level sets of the potential  $W_a := V + F * \delta_a$ .

By definition of  $\mathcal{N}$  in Corollary 2.16, there exists a constant  $\xi > 0$  such that

$$\inf_{z \in \mathcal{N}} (V(z) + F(z - a) - V(a)) = H + 3\xi .$$

We introduce the set

$$\mathcal{K}_{H+2\xi} := \{x \in \mathbb{R}^d : V(x) + F(x - a) - V(a) < H + 2\xi\} .$$

If we denote by  $\tau_\xi$  the first exit-time of  $X$  from  $\mathcal{K}_{H+2\xi}$ , then we obtain

$$\lim_{\sigma \rightarrow 0} \mathbb{P} \left\{ \exp \left[ \frac{2}{\sigma^2} (H + 2\xi - \eta) \right] < \tau_\xi < \exp \left[ \frac{2}{\sigma^2} (H + 2\xi + \eta) \right] \right\} = 1 \quad (2.18)$$

for any  $\eta > 0$ . By construction of  $\mathcal{K}_{H+2\xi}$ ,  $\mathcal{N} \subset \mathcal{K}_{H+2\xi}^c$ , which implies

$$\begin{aligned} \mathbb{P} \{X_\tau \in \mathcal{N}\} &\leq \mathbb{P} \{X_\tau \notin \mathcal{K}_{H+2\xi}\} \\ &\leq \mathbb{P} \{\tau_\xi \leq \tau\} \\ &\leq \mathbb{P} \left\{ \tau_\xi \leq \exp \left[ \frac{2(H + 3\xi)}{\sigma^2} \right] \right\} + \mathbb{P} \left\{ \exp \left[ \frac{2H + \xi}{\sigma^2} \right] \leq \tau \right\} . \end{aligned}$$

Applying (2.18) with  $\eta := \xi$  to the first term and Theorem 2.15 to the second one, we obtain the result.

## 2.3 Discussions on extension

In this Section, we provide some ideas on how the results of the chapter can be extended.

First, we can modify our equation by adding a reflection at the boundary as it was done for McKean-Vlasov case, for example in [AdRR<sup>+</sup>22]. Note, that if the boundary of reflection contains the closure of the domain from which we want to exit, the result for exit-time does not change and is immediate, since, unlike in McKean-Vlasov case, there is no interaction with the law of the process for self-interacting diffusion.

Second, in this chapter we take a diffusion coefficient which is proportional to the identity matrix. However, it could be relevant for some problems related to optimization to consider a more general diffusion coefficient.

Third, in the current work we do not derive an Arrhenius law, that is to say the convergence of  $\frac{\sigma^2}{2} \log(\mathbb{E}[\tau])$  towards  $H$ . To obtain such a result, it requires to use the large deviations techniques instead of the coupling method that is used here.

Fourth, we point out that the potentials  $V$  and  $F$  are both assumed to be uniformly convex. The techniques used in this chapter are not adapted for a more general case. One shall use the techniques close to [DZ10, FW98] to relax the convexity assumptions.

Finally, we could also study SDEs where the nonlinear part of the drift is more general. that is to say:

$$X_t = x_0 + \sigma W_t - \int_0^t \nabla V(X_s) ds - \int_0^t b(X_s, \mu_s) ds, \quad (2.19)$$

where  $\mu_s := \frac{1}{t} \int_0^t \delta_{X_s} ds$  and  $x \mapsto b(x, \mu)$  is differentiable whereas  $\mu \mapsto b(x, \mu)$  is  $L$ -differentiable, see [RW19] and references therein. However, this will require some adaptations of our methods.



# Chapter 3

## Exit-problem for self-interacting diffusion: General case

In this chapter we consider a stochastic process  $(X_t^\sigma, t \geq 0)$  living in  $\mathbb{R}^d$  defined by a stochastic differential equation that includes interaction of the process with its own passed trajectory in the drift term. Consider the following *Self-interacting Diffusion* (SID):

$$\begin{cases} dX_t^\sigma = -\nabla V(X_t^\sigma) dt - \nabla F * \mu_t^\sigma(X_t^\sigma) dt + \sigma dW_t, \\ \mu_t^\sigma = \frac{1}{t} \int_0^t \delta_{X_s^\sigma} ds, \\ X_0^\sigma = x_0 \in \mathbb{R}^d \text{ a.s.} \end{cases} \quad (3.1)$$

$V : \mathbb{R}^d \rightarrow \mathbb{R}$  is called *confinement potential* and represents general geometry of the space.  $F : \mathbb{R}^d \rightarrow \mathbb{R}$  is called *interaction potential*. Its convolution with  $\mu_t^\sigma$ , which is called *empirical measure*, represents interaction term  $\nabla F * \mu_t^\sigma = \frac{1}{t} \int_0^t \nabla F(\cdot - X_s^\sigma) ds$ . Note also that  $(W_t, t \geq 0)$  denotes the standard  $d$ -dimensional Brownian motion,  $\delta_x$  is the Dirac measure on  $\mathbb{R}^d$  concentrated at point  $x$  and  $\sigma > 0$  is a parameter controlling the noise of the system.

In this work we establish the Kramers' type law for the exit-time from an open bounded domain of attraction  $G \subset \mathbb{R}^d$  of the Self-interacting diffusion

(3.1) with general assumptions on confinement and interaction. Moreover, we assume that inside the domain  $G$  there is a unique point of attraction  $a$ . The precise assumptions on potentials  $V$  and  $F$ , and domain  $G$  that we need for the main theorem will be presented and discussed later (the baseline assumptions are given on pages 46 and 57, whereas their generalised version is presented in Section 3.3). The following theorem is the main result of this chapter.

**Theorem 3.1.** *Let Assumptions A-1 and A-2 be fulfilled. Let the process  $X^\sigma$  be the unique strong solution of the system (3.1) with initial condition  $x_0 \in G$ . Let  $\tau_G^\sigma := \inf\{t : X_t^\sigma \notin G\}$  denote the first time when  $X^\sigma$  exits the domain  $G$ . Let  $H := \inf_{x \in \partial G} \{V(x) + F(x - a) - V(a)\}$  be the height of the effective potential. Then, the following two results hold:*

1. *Kramers' law: for any  $\delta > 0$*

$$\lim_{\sigma \rightarrow 0} \mathbb{P}_{x_0} \left( e^{\frac{2(H-\delta)}{\sigma^2}} < \tau_G^\sigma < e^{\frac{2(H+\delta)}{\sigma^2}} \right) = 1; \quad (3.2)$$

2. *Exit-location: for any closed set  $N \subset \partial G$  such that  $\inf_{z \in N} \{V(z) + F(z - a) - V(a)\} > H$  the following limit holds:*

$$\lim_{\sigma \rightarrow 0} \mathbb{P}_{x_0} (X_{\tau_G^\sigma}^\sigma \in N) = 0. \quad (3.3)$$

## Some remarks on dynamics and initial condition

Let us introduce some notation and notions that will be useful to describe laconically later derivations in this chapter.

First, let us define the marginal deterministic system corresponding to (3.1). By  $X^0$  we will denote the solution of the following deterministic system

$$\begin{cases} \frac{d}{dt} X_t^0 = -\nabla V(X_t^0) - \frac{1}{t} \int_0^t \nabla F(X_t^0 - X_s^0) ds, \\ X_0^0 = x_0 \in \mathbb{R}^d. \end{cases} \quad (3.4)$$

Note that, moreover, we can generalize the initial conditions of diffusion (3.1). This process is non-Markov, but, at the same time, knowing only the

past trajectory of the process, we can construct its subsequent path. We will describe the notion of previous path using three parameters.  $t_0 \geq 0$  will represent its time-length,  $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$  will represent the occupation measure of the past trajectory and  $x_0 \in \mathbb{R}^d$  its end point starting from which we continue the dynamic. We contract initial conditions into a vector  $\mathbf{x} = (t_0, \mu_0, x_0) \in \mathfrak{X} = [0, \infty] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$ , and, in order to be able to operate with the three components of  $\mathbf{x}$  separately, we introduce the following projection mappings. Let  $T : \mathfrak{X} \rightarrow [0; \infty]$  be the projection to the first coordinate returning initial time  $t_0$ . For the initial empirical measure, let  $M : \mathfrak{X} \rightarrow \mathcal{P}_2(\mathbb{R}^d)$  be the mapping that returns the second coordinate of  $\mathbf{x} \in \mathfrak{X}$ . Finally, for the the starting point of the process, respectively, define the mapping  $P : \mathfrak{X} \rightarrow \mathbb{R}^d$  that returns the third coordinate of  $\mathbf{x} \in \mathfrak{X}$ .

Following the classical notation for diffusions, we introduce the following system of equations:

$$\begin{cases} dX_t^{\mathbf{x},\sigma} = -\nabla V(X_t^{\mathbf{x},\sigma}) dt - \nabla F * \mu_t^{\mathbf{x},\sigma}(X_t^{\mathbf{x},\sigma}) dt + \sigma dW_t, \\ \mu_t^{\mathbf{x},\sigma} = \frac{T\mathbf{x}}{T\mathbf{x}+t} M\mathbf{x} + \frac{1}{T\mathbf{x}+t} \int_0^t \delta_{X_s^{\mathbf{x},\sigma}} ds, \\ X_0^{\mathbf{x},\sigma} = P\mathbf{x} \text{ a.s.} \end{cases} \quad (3.5)$$

As before, we define marginal deterministic system that corresponds to the equation (3.5) as

$$\begin{cases} \frac{d}{dt} X_t^{\mathbf{x},0} = -\nabla V(X_t^{\mathbf{x},0}) - \frac{T\mathbf{x}}{T\mathbf{x}+t} \int_{\mathbb{R}^d} \nabla F(X_t^{\mathbf{x},0} - s) M\mathbf{x}(ds) \\ \quad - \frac{1}{T\mathbf{x}+t} \int_0^t \nabla F(X_t^{\mathbf{x},0} - X_s^{\mathbf{x},0}) ds, \\ X_0^{\mathbf{x},0} = P\mathbf{x}. \end{cases} \quad (3.6)$$

Hereafter, we will drop one part of the subscript or both (initial conditions and  $\sigma$ ) if it is clear which one of the equations is meant. Though, to avoid confusion we will always use the subscript 0 to emphasize that we consider specifically the deterministic system. For  $\mathbf{x}$  such that  $T\mathbf{x} = \infty$  we

naturally extend the definition of the processes (3.5) and (3.6) as

$$\begin{cases} dX_t^{\mathbf{x},\sigma} = -\nabla V(X_t^{\mathbf{x},\sigma}) dt - \nabla F * M_{\mathbf{x}}(X_t^{\mathbf{x},\sigma}) dt + \sigma dW_t, \\ X_0^{\mathbf{x},\sigma} = P_{\mathbf{x}} \text{ a.s.} \end{cases} \quad (3.7)$$

Despite that introduction of  $\mathbf{x}$  aims to describe a notion of “previous path” of the process, the set  $\mathfrak{X}$  is defined to be more general. Not only we allow  $t_0$  to be equal to infinity, but also  $\mu_0$  belongs to  $\mathcal{P}_2(\mathbb{R}^d)$ , which, in general, does not restrict it to be an empirical measure of any path.

Note that topology naturally defined on  $[0; \infty]$  by open intervals in usual sense and intervals of the form  $[0, x)$ ,  $(x; \infty]$  is metrizable. Let us denote  $d_{[0,\infty]}$  some metric on this space. We also equip the set  $\mathcal{P}_2(\mathbb{R}^d)$  with Wasserstein-2 metric (see e.g. [Vil09, Definition 6.1]). Thus, the Cartesian product  $\mathfrak{X}$  is a complete separable metric space with the following metric

$$d_{\mathfrak{X}}(\mathbf{x}_1, \mathbf{x}_2) = \max(d_{[0,\infty]}(T\mathbf{x}_1, T\mathbf{x}_2); \mathbb{W}_2(M_{\mathbf{x}_1}, M_{\mathbf{x}_2}); |P_{\mathbf{x}_1} - P_{\mathbf{x}_2}|),$$

where  $\mathbb{W}_2$  is the Wasserstein distance.

Existence and uniqueness result for stochastic differential equations that depend on their path is standard for the case of Lipschitz continuous drift term (see e.g. [RW00, Theorem 11.2]). It is worth noting that introducing  $\mathbf{x} \in \mathfrak{X}$  does not complicate matters in any way, as long as  $\nabla V$  and  $\nabla F$  remain Lipschitz continuous. Assumption A-1 ensures that we are working in this framework.

## Outline of the chapter

Section 3.1 is dedicated to the Large deviation principle. First, we define what we mean by the large deviation principle (LDP), then we establish, under the set of assumptions A-1, LDP for self-interacting diffusion in Theorem 3.3. Next, some useful (for exit-time problem) results related to LDP are presented.

Section 3.2 deals with the exit-time problem. First, we present the set of assumptions A-2 that contains assumptions on domain  $G$ , its boundary and



stability properties. We discuss the role that each of these assumptions play in later proof and present some examples of confinement and interaction potentials along with possible domains  $G$  that one can choose. Then, we present the main result of the chapter, that is Theorem 3.1. This result is followed on page 59 by a section where we present auxiliary lemmas that are later used to prove the main Theorem 3.1 (page 63). These lemmas are later proved on page 68.

Section 3.3 is focused on the generalization of Assumptions A-1 and A-2. In this section, we examine the scenario of a locally Lipschitz continuous drift term ( $\nabla V$  and  $\nabla F$ ) and an unbounded domain  $G$  from which we want to exit.

### 3.1 Large Deviation Principle

In this chapter Large Deviations techniques are widely used. Recall that in this thesis by the Large Deviation Principle (LDP) we mean the following asymptotic behaviour of measures.

**Definition 3.2.** Family of measures  $(\nu^\sigma)_{\sigma>0}$  defined on some Banach space  $B$  equipped with Borel sigma-algebra  $\mathcal{B}$  is said to satisfy the Large Deviation Principle with a good rate function  $I$  if for any measurable set  $\Gamma \in \mathcal{B}$ :

$$-\inf_{x \in \Gamma} I(x) \leq \liminf_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \nu^\sigma(\Gamma) \leq \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \nu^\sigma(\Gamma) \leq -\inf_{x \in \Gamma} I(x),$$

where  $I : B \rightarrow [0, \infty]$  is a lower-semicontinuous function (this property defines *rate function*) whose level sets  $\{x : I(x) \leq \alpha\}$  are compact subsets of  $B$  for any  $0 \leq \alpha < \infty$  (which means by definition that the rate function is *good*).

For more information about LDP and the results related to it that we use in this chapter, see Chapter 1.

#### Establishing the LDP for the SID

In this section we prove LDP for Self-Interacting diffusions of the type (3.5). First, we introduce the following group of main assumptions.

**Assumptions A-1.**

1. (regularity) Potentials  $V$  and  $F$  belong to the space  $C^2(\mathbb{R}^d; \mathbb{R})$ .
2. (Lipschitz continuity) There exist  $\text{Lip}_{\nabla V}, \text{Lip}_{\nabla F} > 0$  such that we have  $|\nabla V(x) - \nabla V(y)| \leq \text{Lip}_{\nabla V}|x - y|$  and  $|\nabla F(x) - \nabla F(y)| \leq \text{Lip}_{\nabla F}|x - y|$  for any  $x, y \in \mathbb{R}^d$ .
3. (boundedness of  $\nabla F$ ) There exists  $C_{\nabla F} > 0$  such that  $|\nabla F(x)| \leq C_{\nabla F}$  for any  $x \in \mathbb{R}^d$ .

Not only the first two assumptions above help to establish the global existence and uniqueness result for (3.1) and (3.5) easily, but they also play an important role in proving the Large Deviation Principle. The third assumption A-1.3 is only used in Lemma 3.5 and can be substituted, for example, with a condition that diffusion  $X^{\mathbf{x}, \sigma}$  does not leave some bounded domain in  $\mathbb{R}^d$ . Since it is indeed the case in Section 3.2, Assumption A-1.3 is not as restrictive as it could seem. By the same logic, Assumption A-1.2 can be relaxed to locally Lipschitz case whenever the existence and uniqueness problem is resolved. We will stress again this observations in Section 3.3.

Consider the following theorem that is the main result of the section.

**Theorem 3.3.** *Under Assumptions A-1.1, A-1.2 for any  $\mathbf{x} \in \mathfrak{X}$  the probability measures  $(\nu^{\mathbf{x}, \sigma})_{\sigma > 0}$  induced on  $C([0, T]; \mathbb{R}^d)$  by the process  $(X_t^{\mathbf{x}, \sigma})_{t \leq T}$ , which is the unique solution of the system (3.5), satisfy LDP with the following good rate function:*

$$\begin{aligned}
I_T^{\mathbf{x}}(f) &:= \frac{1}{4} \int_0^T \left| \dot{f}(t) + \nabla V(f(t)) + \frac{\mathbf{T}\mathbf{x}}{\mathbf{T}\mathbf{x} + t} \nabla F * \mathbf{M}\mathbf{x}(f(t)) \right. \\
&\quad \left. + \frac{1}{\mathbf{T}\mathbf{x} + t} \int_0^t \nabla F(f(t) - f(s)) ds \right|^2 dt, \text{ for } f \in H_1^{\mathbf{P}\mathbf{x}} \quad (3.8)
\end{aligned}$$

and  $I_T^{\mathbf{x}}(f) := \infty$  otherwise. Here,  $H_1^{\mathbf{x}}$  is the space of absolutely continuous functions with square-integrable derivatives that start at point  $x = \mathbf{P}\mathbf{x}$ ;  $H_1^{\mathbf{P}\mathbf{x}} := \{\mathbf{P}\mathbf{x} + \int_0^{\cdot} g(s) ds : g \in L_2([0, T]; \mathbb{R}^d)\}$ .

*Proof.* Let us define function  $G : C_0([0, T]; \mathbb{R}^d) \rightarrow C([0, T]; \mathbb{R}^d)$  that maps every function  $g$  to the unique solution of the following equation:

$$\begin{aligned} f(t) = & P_{\mathbf{x}} - \int_0^t \nabla V(f(s)) ds - \int_0^t \frac{T_{\mathbf{x}}}{T_{\mathbf{x}} + s} \nabla F * M_{\mathbf{x}}(f(s)) ds \\ & - \int_0^t \frac{1}{T_{\mathbf{x}} + s} \int_0^s \nabla F(f(s) - f(u)) du ds + g(t). \end{aligned}$$

Let us show that  $G$  is continuous. Fix some  $\delta > 0$  and two functions  $g_1, g_2 \in C_0([0, T]; \mathbb{R}^d)$  such that  $\|g_1 - g_2\|_{\infty} \leq \delta$ . Let  $f_1 = G(g_1)$ ,  $f_2 = G(g_2)$ . Then,

$$\begin{aligned} |f_1(t) - f_2(t)| \leq & \int_0^t |\nabla V(f_1(s)) - \nabla V(f_2(s))| ds \\ & + \int_0^t \frac{T_{\mathbf{x}}}{T_{\mathbf{x}} + s} \int_{\mathbb{R}^d} |\nabla F(f_1(s) - u) - \nabla F(f_2(s) - u)| M_{\mathbf{x}}(du) ds \\ & + \int_0^t \frac{1}{T_{\mathbf{x}} + s} \int_0^s |\nabla F(f_1(s) - f_1(u)) - \nabla F(f_2(s) - f_2(u))| du ds \\ & + |g_1(t) - g_2(t)|. \end{aligned}$$

Use Lipschitz continuity of  $\nabla V$ ,  $\nabla F$  (Assumption A-1.2) and get:

$$\begin{aligned} |f_1(t) - f_2(t)| \leq & (\text{Lip}_{\nabla V} + \text{Lip}_{\nabla F}) \int_0^t |f_1(s) - f_2(s)| ds \\ & + \int_0^t \frac{\text{Lip}_{\nabla F}}{T_{\mathbf{x}} + s} \int_0^s |f_1(u) - f_2(u)| du ds + \delta \\ \leq & \left( \text{Lip}_{\nabla V} + \left(1 + \frac{T}{T_{\mathbf{x}}}\right) \text{Lip}_{\nabla F} \right) \int_0^t |f_1(s) - f_2(s)| ds + \delta. \end{aligned}$$

Therefore, by Grönwall's inequality

$$\|f_1 - f_2\|_{\infty} \leq \delta \exp \left\{ \left( \text{Lip}_{\nabla V} + \left(1 + \frac{T}{T_{\mathbf{x}}}\right) \text{Lip}_{\nabla F} \right) T \right\};$$

it means the continuity of the map  $G$  for any possible  $\mathbf{x} \in \mathfrak{X}$ .

By the uniqueness of the solution to (3.5), we can express  $\nu^{\mathbf{x}, \sigma}$  as  $\nu^{\mathbf{x}, \sigma} = G_{\#} \eta^{\sigma}$ , where  $\eta^{\sigma}$  is the probability measure induced on  $C_0([0, T]; \mathbb{R}^d)$  by the path of Brownian motion  $W^{\sigma} = \sigma W$ . LDP for the path of Brownian

motion with vanishing noise is known under the name of Schilder theorem ([DZ10, Theorem 5.2.3]). Since  $G$  is continuous, we can apply the Contraction principle (see [DZ10, Theorem 4.2.1]) and conclude that the family of measures  $(\nu^{\mathbf{x},\sigma})_{\sigma \geq 0}$  satisfies an LDP with the good rate function

$$I_T^{\mathbf{x}}(f) = \inf_{g \in G^{-1}(f)} \frac{1}{4} \int_0^T |\dot{g}(t)|^2 dt$$

$$= \begin{cases} \frac{1}{4} \int_0^T |\dot{f}(t) + \nabla V(f(t)) + \frac{\mathbb{T}\mathbf{x}}{\mathbb{T}\mathbf{x}+t} \int_{\mathbb{R}^d} \nabla F(f(t) - u) M\mathbf{x}(du) \\ \quad + \frac{1}{\mathbb{T}\mathbf{x}+t} \int_0^t \nabla F(f(t) - f(u)) du|^2 dt, & \text{for } f \in H_1^{\text{Px}}, \\ \infty, & \text{otherwise.} \end{cases}$$

□

**Remark 3.4.** For simplicity of the notation, we define the rate function corresponding to the system (3.1) as  $I_T^{x_0}$ , i.e.

$$I_T^{x_0}(f) := \begin{cases} \frac{1}{4} \int_0^T |\dot{f}(t) + \nabla V(f(t)) \\ \quad + \frac{1}{t} \int_0^t \nabla F(f(t) - f(s)) ds|^2 dt, & \text{for } f \in H_1^{x_0}, \\ \infty, & \text{otherwise.} \end{cases}$$

## Results related to the LDP

The following lemma generalizes the large deviation principle for the case of converging initial conditions.

**Lemma 3.5.** *Under assumptions A-1, for any sequence  $\{\mathbf{x}_n\}_1^\infty$  such that  $\mathbf{x}_n \xrightarrow[n \rightarrow \infty]{d_x} \mathbf{x}$  the following inequalities hold:*

1. For any closed  $\Phi \subset C([0, T]; \mathbb{R}^d)$

$$\limsup_{\substack{n \rightarrow \infty \\ \sigma \rightarrow 0}} \frac{\sigma^2}{2} \log \mathbb{P}_{\mathbf{x}_n}(X^\sigma \in \Phi) \leq - \inf_{\phi \in \Phi} I_T^{\mathbf{x}}(\phi).$$

2. For any open  $\Psi \subset C([0, T]; \mathbb{R}^d)$

$$\liminf_{\substack{n \rightarrow \infty \\ \sigma \rightarrow 0}} \frac{\sigma^2}{2} \log \mathbb{P}_{\mathbf{x}_n}(X^\sigma \in \Psi) \geq - \inf_{\phi \in \Psi} I_T^{\mathbf{x}}(\phi).$$

*Proof.* We will show that the families of measures  $(\nu^{\mathbf{x}_\sigma, \sigma})_{\sigma > 0}$  and  $(\nu^{\mathbf{x}, \sigma})_{\sigma > 0}$ , which are families of probability measures induced on  $C([0, T]; \mathbb{R}^d)$  by  $X^{\mathbf{x}_\sigma, \sigma}$  and  $X^{\mathbf{x}, \sigma}$  respectively, are exponentially equivalent (see for example [DZ10, Definition 4.2.10]) for any  $\{\mathbf{x}_\sigma\}_{\sigma > 0}$  such that  $\mathbf{x}_\sigma \xrightarrow[\sigma \rightarrow 0]{d_{\mathbf{x}}} \mathbf{x}$ .

Indeed, if we define  $Z_t^\sigma = X_t^{\mathbf{x}_\sigma, \sigma} - X_t^{\mathbf{x}, \sigma}$ , where  $X_t^{\mathbf{x}_\sigma, \sigma}$  and  $X_t^{\mathbf{x}, \sigma}$  are driven by the same Brownian motion, then

$$\begin{aligned} |Z_t^\sigma| &\leq |\mathbb{P}_{\mathbf{x}_\sigma} - \mathbb{P}_{\mathbf{x}}| + \int_0^t |\nabla V(X_s^{\mathbf{x}_\sigma, \sigma}) - \nabla V(X_s^{\mathbf{x}, \sigma})| ds \\ &+ \int_0^t \left| \frac{\mathbb{T}_{\mathbf{x}_\sigma}}{\mathbb{T}_{\mathbf{x}_\sigma + s}} \nabla F * \mathbb{M}_{\mathbf{x}_\sigma}(X_s^{\mathbf{x}_\sigma, \sigma}) - \frac{\mathbb{T}_{\mathbf{x}}}{\mathbb{T}_{\mathbf{x} + s}} \nabla F * \mathbb{M}_{\mathbf{x}}(X_s^{\mathbf{x}, \sigma}) \right| ds \\ &+ \int_0^t \int_0^s \left| \frac{1}{\mathbb{T}_{\mathbf{x}_\sigma + s}} \nabla F(X_s^{\mathbf{x}_\sigma, \sigma} - X_z^{\mathbf{x}_\sigma, \sigma}) - \frac{1}{\mathbb{T}_{\mathbf{x} + s}} \nabla F(X_s^{\mathbf{x}, \sigma} - X_z^{\mathbf{x}, \sigma}) \right| dz ds. \end{aligned}$$

In order to separate the effect of closeness of  $\mathbb{T}_{\mathbf{x}_\sigma}$  to  $\mathbb{T}_{\mathbf{x}}$  from  $Z_t^\sigma$  in the last two integrals, we add and subtract expressions  $\frac{\mathbb{T}_{\mathbf{x}_\sigma}}{\mathbb{T}_{\mathbf{x}_\sigma + s}} \nabla F * \mathbb{M}_{\mathbf{x}}(X_s^{\mathbf{x}, \sigma})$  and  $\frac{1}{\mathbb{T}_{\mathbf{x}_\sigma + s}} \nabla F(X_s^{\mathbf{x}, \sigma} - X_z^{\mathbf{x}, \sigma})$  in the corresponding integrals. Since  $\nabla V$  and  $\nabla F$  are Lipschitz continuous, we get

$$\begin{aligned} |Z_t^\sigma| &\leq (1 + \text{Lip}_{\nabla F}) d_{\mathbf{x}}(\mathbf{x}_\sigma, \mathbf{x}) + (\text{Lip}_{\nabla V} + \text{Lip}_{\nabla F}) \int_0^t |Z_s^\sigma| ds \\ &+ \int_0^t \int_{\mathbb{R}^d} \left| \left( \frac{\mathbb{T}_{\mathbf{x}_\sigma}}{\mathbb{T}_{\mathbf{x}_\sigma + s}} - \frac{\mathbb{T}_{\mathbf{x}}}{\mathbb{T}_{\mathbf{x} + s}} \right) \nabla F(X_s^{\mathbf{x}, \sigma} - z) \right| \mathbb{M}_{\mathbf{x}}(dz) ds \\ &+ \text{Lip}_{\nabla F} \int_0^t \frac{1}{\mathbb{T}_{\mathbf{x}_\sigma + s}} \int_0^s (|Z_s^\sigma| + |Z_z^\sigma|) dz ds \\ &\leq (1 + \text{Lip}_{\nabla F}) d_{\mathbf{x}}(\mathbf{x}_\sigma, \mathbf{x}) + \left( \text{Lip}_{\nabla V} + 2\text{Lip}_{\nabla F} + \frac{T}{\mathbb{T}_{\mathbf{x}_\sigma}} \right) \int_0^t |Z_s^\sigma| ds \\ &+ C_{\nabla F} \int_0^t \left| \frac{\mathbb{T}_{\mathbf{x}_\sigma}}{\mathbb{T}_{\mathbf{x}_\sigma + s}} - \frac{\mathbb{T}_{\mathbf{x}}}{\mathbb{T}_{\mathbf{x} + s}} \right| ds, \end{aligned}$$

where we get the last inequality by applying assumption A-1.3 to

$$\int_0^t \int_{\mathbb{R}^d} \left| \left( \frac{\mathbb{T}_{\mathbf{x}_\sigma}}{\mathbb{T}_{\mathbf{x}_\sigma + s}} - \frac{\mathbb{T}_{\mathbf{x}}}{\mathbb{T}_{\mathbf{x} + s}} \right) \nabla F(X_s^{\mathbf{x}, \sigma} - z) \right| \mathbb{M}_{\mathbf{x}}(dz) ds.$$

Since the expression inside the integral is bounded, we have the following equality:  $\int_0^t \left| \frac{T\mathbf{x}_\sigma}{T\mathbf{x}_\sigma+s} - \frac{T\mathbf{x}}{T\mathbf{x}+s} \right| ds = O(|T\mathbf{x}_\sigma - T\mathbf{x}|) = O(d_{\mathfrak{X}}(\mathbf{x}_\sigma, \mathbf{x}))$ . Thus, by Grönwall's inequality,

$$|Z_t^\sigma| \leq (1 + \text{Lip}_{\nabla F} + C_{\nabla F}) O(d_{\mathfrak{X}}(\mathbf{x}_\sigma; \mathbf{x})) \\ \times \exp \left\{ \left( \text{Lip}_{\nabla V} + 2\text{Lip}_{\nabla F} + \frac{T}{T\mathbf{x}_\sigma} \right) T \right\}.$$

It means that  $\mathbb{P}(|Z_t^\sigma| \geq \delta) = 0$  for any  $\delta$  if we choose  $\sigma$  to be small enough. That proves exponential equivalence of  $(\nu^{\mathbf{x}_\sigma, \sigma})_{\sigma>0}$  and  $(\nu^{\mathbf{x}, \sigma})_{\sigma>0}$ , and, by contraction principle (see [DZ10, Theorem 4.2.13]), the lemma itself.  $\square$

The following lemma elaborates on the idea of the LDP and provides a tool to study the asymptotic behaviour ( $\sigma \rightarrow 0$ ) of the process  $X^\sigma$  with respect to its initial conditions.

**Lemma 3.6.** *For any compact subset  $\mathcal{C} \subset \mathfrak{X}$  the following inequalities hold:*

1. *For any closed  $\Phi \subset C([0, T]; \mathbb{R}^d)$*

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup_{\mathbf{x} \in \mathcal{C}} \mathbb{P}_{\mathbf{x}}(X^\sigma \in \Phi) \leq - \inf_{\mathbf{x} \in \mathcal{C}} \inf_{\phi \in \Phi} I_T^{\mathbf{x}}(\phi).$$

2. *Similarly, for any open  $\Psi \subset C([0, T]; \mathbb{R}^d)$*

$$\liminf_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \inf_{\mathbf{x} \in \mathcal{C}} \mathbb{P}_{\mathbf{x}}(X^\sigma \in \Psi) \geq - \sup_{\mathbf{x} \in \mathcal{C}} \inf_{\phi \in \Psi} I_T^{\mathbf{x}}(\phi).$$

*Proof.* For the first inequality, for each fixed  $\varepsilon > 0$  let us define

$$I^\varepsilon := \min \left\{ \inf_{\mathbf{x} \in \mathcal{C}} \inf_{\phi \in \Phi} I_T^{\mathbf{x}}(\phi) - \varepsilon; \frac{1}{\varepsilon} \right\}.$$

By Lemma 3.5 for any  $\mathbf{x} \in \mathcal{C}$  there exists small enough  $\sigma_{\mathbf{x}} > 0$  such that for any  $\sigma < \sigma_{\mathbf{x}}$

$$\frac{\sigma^2}{2} \log \sup_{\mathbf{y} \in B_{\sigma_{\mathbf{x}}}(\mathbf{x})} \mathbb{P}_{\mathbf{y}}(X^\sigma \in \Phi) \leq -I^\varepsilon.$$

Since  $\mathcal{C}$  is compact, we can cover it by finite amount of  $B_{\sigma_{\mathbf{x}_i}}(\mathbf{x}_i)$  for some  $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathcal{C}$ . Then, for any  $\sigma < \min_{1 \leq i \leq m} \sigma_{\mathbf{x}_i}$

$$\frac{\sigma^2}{2} \log \sup_{\mathbf{x} \in \mathcal{C}} \mathbb{P}_{\mathbf{x}}(X^\sigma \in \Phi) \leq -I^\varepsilon,$$

and that proves the first inequality. One can prove the second inequality the same way.  $\square$

By definition, rate functions are lower semicontinuous, i.e. for any  $\mathbf{x} \in \mathfrak{X}$ ,  $\liminf_{\phi \rightarrow \phi_0} I_T^{\mathbf{x}}(\phi) \geq I_T^{\mathbf{x}}(\phi_0)$ , or, equivalently, all level sets  $L_\alpha := \{\phi \in C([0, T]; \mathbb{R}^d) : I_T^{\mathbf{x}}(\phi) \leq \alpha\}$  are closed for  $0 \leq \alpha < \infty$ . An immediate consequence of this property is that infima of  $I_T^{\mathbf{x}}$  are achieved over compact sets. The following lemma extends this lower semicontinuity property of our rate functions to the case of converging initial conditions besides the argument of  $I_T$ . In other words,  $I_T$  as a function of two arguments  $(\mathbf{x}, \phi) \in \mathfrak{X} \times C([0, T]; \mathbb{R}^d)$  still possesses lower semicontinuity property.

**Lemma 3.7.** *For any  $T > 0$ ,  $\mathbf{x} \in \mathfrak{X}$ ,  $\phi \in C([0, T]; \mathbb{R}^d)$ , and any sequences  $\{\mathbf{x}_n\}$ ,  $\{\phi_n\}$ , where  $\mathbf{x}_n \in \mathfrak{X}$ ,  $\phi_n \in C([0, T]; \mathbb{R}^d)$ , such that  $\mathbf{x}_n \xrightarrow[n \rightarrow \infty]{d_{\mathfrak{X}}} \mathbf{x}$  and  $\phi_n \xrightarrow[n \rightarrow \infty]{} \phi$  the following inequality holds:*

$$\liminf_{n \rightarrow \infty} I_T^{\mathbf{x}_n}(\phi_n) \geq I_T^{\mathbf{x}}(\phi).$$

*Proof.* Without loss of generality, let us assume that all  $\phi_n \in H_1^{\mathbf{P}\mathbf{x}}$  (those  $\phi_n$  for which it is not true do not influence the  $\liminf$  since  $I_T^{\mathbf{x}_n}(\phi_n) = \infty$ ). Then

$$\begin{aligned} 4I_T^{\mathbf{x}_n}(\phi_n) &= \int_0^T |\dot{\phi}_n(t) + \nabla V(\phi_n(t)) \\ &\quad + \frac{\mathbf{T}\mathbf{x}_n}{\mathbf{T}\mathbf{x}_n + t} \int_{\mathbb{R}^d} \nabla F(\phi_n(t) - u) \mathbf{M}\mathbf{x}_n(du) \\ &\quad + \frac{1}{\mathbf{T}\mathbf{x}_n + t} \int_0^t |\nabla F(\phi_n(t) - \phi_n(u)) du|^2 dt. \end{aligned} \quad (3.9)$$

Let us add and subtract two terms of the form  $A := \frac{\mathbf{T}\mathbf{x}}{\mathbf{T}\mathbf{x}+t} \int_{\mathbb{R}^d} \nabla F(\phi_n(t) - u) \mathbf{M}\mathbf{x}(du)$  and  $B := \frac{1}{\mathbf{T}\mathbf{x}+t} \int_0^t \nabla F(\phi_n(t) - \phi_n(u)) du$ . Note that for any  $c \in (0, 1)$  and for any  $a, b \in \mathbb{R}^d$  the following inequality holds:  $|a + b|^2 \geq (1 - c)|a|^2 + (1 - \frac{1}{c})|b|^2$ . Using two previous statements, let us split (3.9) into two parts, one of which does not depend on  $\mathbf{x}_n$ , but depends on  $\mathbf{x}$  instead. We then have

$$\begin{aligned} 4I_T^{\mathbf{x}_n}(\phi_n) &\geq (1 - c)4I_T^{\mathbf{x}}(\phi_n) \\ &\quad + \left(1 - \frac{1}{c}\right) \int_0^T \left| \frac{\mathbf{T}\mathbf{x}_n}{\mathbf{T}\mathbf{x}_n + t} \int_{\mathbb{R}^d} \nabla F(\phi_n(t) - u) \mathbf{M}\mathbf{x}_n(du) - A \right. \\ &\quad \left. + \frac{1}{\mathbf{T}\mathbf{x}_n + t} \int_0^t \nabla F(\phi_n(t) - \phi_n(u)) du - B \right|^2 dt \\ &=: (1 - c)4I_T^{\mathbf{x}}(\phi_n) + \left(1 - \frac{1}{c}\right) \mathbf{I}. \end{aligned}$$

Since  $I_T^{\mathbf{x}}(\phi_n)$  is a good rate function, by taking  $\liminf$  from both sides of the inequality above we get

$$\liminf_{n \rightarrow \infty} I_T^{\mathbf{x}_n}(\phi_n) \geq (1 - c)I_T^{\mathbf{x}}(\phi) + \frac{1}{4} \left(1 - \frac{1}{c}\right) \liminf_{n \rightarrow \infty} \mathbf{I}, \quad (3.10)$$

for all  $c \in (0, 1)$ . Therefore, in order to prove the lemma, it is enough to show that  $\liminf_{n \rightarrow \infty} \mathbf{I} = 0$ . Then we could take the limit  $c \rightarrow 0$  from both sides of (3.10) and obtain the necessary result.

Let us consider  $\mathbf{I}$ . We use inequality of the form  $|a + b|^2 \leq 2|a|^2 + 2|b|^2$  for  $\mathbf{I}$  and get an upper bound of the form:

$$\begin{aligned} \mathbf{I} &\leq 2 \int_0^T \left| \frac{\mathbf{T}\mathbf{x}_n}{\mathbf{T}\mathbf{x}_n + t} \int_{\mathbb{R}^d} \nabla F(\phi_n(t) - u) \mathbf{M}\mathbf{x}_n(du) \right. \\ &\quad \left. - \frac{\mathbf{T}\mathbf{x}}{\mathbf{T}\mathbf{x} + t} \int_{\mathbb{R}^d} \nabla F(\phi_n(t) - u) \mathbf{M}\mathbf{x}(du) \right|^2 dt \\ &\quad + 2 \int_0^T \left| \frac{1}{\mathbf{T}\mathbf{x}_n + t} \int_0^t \nabla F(\phi_n(t) - \phi_n(u)) du \right. \\ &\quad \left. - \frac{1}{\mathbf{T}\mathbf{x} + t} \int_0^t \nabla F(\phi_n(t) - \phi_n(u)) du \right|^2 dt \\ &=: 2\mathbf{II}_1 + 2\mathbf{II}_2. \end{aligned}$$



Let us apply dominated convergence theorem to both  $\text{II}_1$  and  $\text{II}_2$ . In order to do so, we prove that integrands are uniformly bounded.

First, consider  $\text{I}_1$ . Of course, fractions  $\frac{T\mathbf{x}_n}{T\mathbf{x}_n+t}$  and  $\frac{T\mathbf{x}}{T\mathbf{x}+t}$  are bounded by 1. As for  $\int_{\mathbb{R}^d} |\nabla F(\phi_n(t) - u)| \text{M}\mathbf{x}_n(du)$ , we introduce the following decomposition

$$\begin{aligned} & \int_{\mathbb{R}^d} |\nabla F(\phi_n(t) - u)| \text{M}\mathbf{x}_n(du) \\ &= \int_{\mathbb{R}^d} |\nabla F(\phi_n(t) - u) - \nabla F(\phi(t) - u) + \nabla F(\phi(t) - u)| \text{M}\mathbf{x}_n(du) \quad (3.11) \\ &\leq \text{Lip}_{\nabla F} |\phi_n(t) - \phi(t)| + \int_{\mathbb{R}^d} |\nabla F(\phi(t) - u)| \text{M}\mathbf{x}_n. \end{aligned}$$

By Assumption A-1.2, the expression above is bounded by:

$$\text{Lip}_{\nabla F} \|\phi_n - \phi\|_\infty + C_{\nabla F} \left( \max_{t \in [0, T]} |\phi(t)| + \left( \int_{\mathbb{R}^d} |u|^2 \text{M}\mathbf{x}_n \right)^{1/2} \right),$$

and since  $\int |u|^2 \text{M}\mathbf{x}_n \xrightarrow[n \rightarrow \infty]{} \int |u|^2 \text{M}\mathbf{x}$  (see [Vil09, Theorem 6.9]), the integral  $\int_{\mathbb{R}^d} |\nabla F(\phi_n(t) - u)| \text{M}\mathbf{x}_n(du)$  and, as a consequence,  $\int_{\mathbb{R}^d} |\nabla F(\phi_n(t) - u)| \text{M}\mathbf{x}(du)$  are uniformly (in  $t$ ) bounded.

For  $\text{II}_2$ , we also use lipschitzness of  $\nabla F$  when needed, convergence of  $\phi_n$  in uniform norm topology towards  $\phi$ , as well as bounds of the form  $|\phi(t)| \leq \max_{t \in [0, T]} |\phi(t)|$ . That easily gives us uniform (in  $t$ ) boundedness of the integrand of the integral  $\text{II}_2$ .

Thus, to calculate  $\lim_{n \rightarrow \infty} \text{I}$ , we can use the dominated convergence theorem and pass the limit inside both of the integrals  $\text{II}_1$  and  $\text{II}_2$ .

Since  $\nabla F$  is continuous and since the uniform boundedness in  $t$  and in  $u$  on finite time interval  $[0, T]$  of  $\nabla F(\phi_n(t) - \phi_n(u))$  can be easily established,  $\lim_{n \rightarrow \infty} \text{II}_2$  is clearly equal to 0. Limits of components of  $\text{II}_1$  are also obvious, except for  $\int_{\mathbb{R}^d} |\nabla F(\phi_n(t) - u)| \text{M}\mathbf{x}_n(du)$  that needs some attention. Let us show that expression  $\int_{\mathbb{R}^d} |\nabla F(\phi(t) - u)| \text{M}\mathbf{x}(du)$  is actually its limit with

$n \rightarrow \infty$ . Similarly to computations in (3.11), we get

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} \nabla F(\phi_n(t) - u) M_{\mathbf{x}_n}(du) - \int_{\mathbb{R}^d} \nabla F(\phi(t) - u) M_{\mathbf{x}}(du) \right| \\ & \leq \left| \int_{\mathbb{R}^d} \nabla F(\phi(t) - u) M_{\mathbf{x}_n}(du) - \int_{\mathbb{R}^d} \nabla F(\phi(t) - u) M_{\mathbf{x}}(du) \right| \\ & \quad + \text{Lip}_{\nabla F} \|\phi_n - \phi\|_{\infty}. \end{aligned}$$

As was pointed out before, convergence of measures in Wasserstein distance gives convergence of respective integrals, since  $\nabla F$  is Lipschitz continuous [Vil09, Theorem 6.9].

This is the last remark needed to observe that  $\lim_{n \rightarrow \infty} \mathbf{I} = 0$ . Thus, the lemma is proved by (3.10).  $\square$

As was pointed out before, lower semicontinuity guarantees that infima of a function are achieved over compact sets. We summarise this property by the following corollary.

**Corollary 3.8.** *For any  $T > 0$ , any compact set  $\Phi \subset C([0, T]; \mathbb{R}^d)$ , and any compact set  $\mathcal{C} \subset \mathfrak{X}$  there exist  $\phi^* \in \Phi$  and  $\mathbf{x}^* \in \mathcal{C}$  such that*

$$\inf_{\mathbf{x} \in \mathcal{C}} \inf_{\phi \in \Phi} I_T^{\mathbf{x}}(\phi) = I_T^{\mathbf{x}^*}(\phi^*).$$

## Compactness results

In this chapter compact subsets of  $\mathfrak{X} \times C([0, T]; \mathbb{R}^d)$  of particular form are considered. Let us present two results about compactness of some sets that will be used later in the proof.

**Lemma 3.9.** *For any  $T > 0$  and any compact subsets  $C_1, C_2 \subset \mathbb{R}^d$  the following set*

$$\mathcal{C} := \{\mathbf{x} \in \mathfrak{X} : T \leq T_{\mathbf{x}} \leq \infty, P_{\mathbf{x}} \in C_1, M_{\mathbf{x}} \in \mathcal{P}_2(C_2)\}.$$

*is a compact subset of  $\mathfrak{X}$ .*

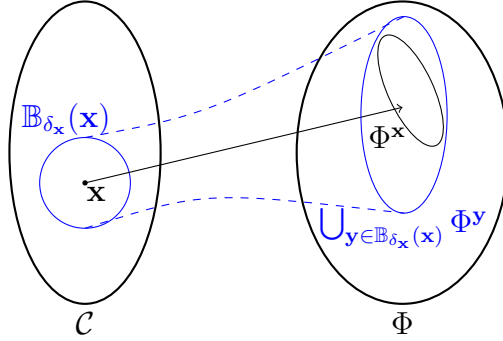
*Proof.* Projections of  $\mathcal{C}$  on first two axes are obviously compact subsets in  $[0, \infty]$  and  $\mathbb{R}^d$ . For compactness of the projection on the third axis, note that by Prokhorov's theorem this set is compact in weak topology which is metrizable by Wasserstein-2 distance (see [Vil09, Theorem 6.9]).  $\square$

**Lemma 3.10.** *Let  $\Phi^{\mathbf{x}} := \{\phi \in C([0, T]; \mathbb{R}^d) : I_T^{\mathbf{x}}(\phi) \leq 1\}$ , where  $\mathbf{x} \in \mathcal{C}$ ,  $T > 0$ , and  $\mathcal{C}$  is a compact subset of  $\mathfrak{X}$ . Then*

$$\Phi = \bigcup_{\mathbf{x} \in \mathcal{C}} \Phi^{\mathbf{x}}$$

*is a relatively compact subset of  $C([0, T]; \mathbb{R}^d)$ .*

*Proof.* In the case of complete metric spaces, the notion of relative compactness is equivalent to totally boundedness. By definition, the set  $\Phi$  is totally bounded if for any  $\varepsilon > 0$  there exists a finite cover of  $\Phi$  with open balls of radius  $\varepsilon$ . The strategy of the proof is the following. We prove that for any  $\varepsilon > 0$  and any  $\mathbf{x} \in \mathcal{C}$  there exists  $\delta_{\mathbf{x}}^{\varepsilon} > 0$  small enough such that  $\bigcup_{\mathbf{y} \in \mathbb{B}_{\delta_{\mathbf{x}}^{\varepsilon}}(\mathbf{x})} \Phi^{\mathbf{y}}$  is covered by a finite number of balls of radius  $\varepsilon$ . Since  $\mathcal{C}$  is itself a compact subset of  $\mathfrak{X}$  and  $\{\mathbb{B}_{\delta_{\mathbf{x}}^{\varepsilon}}(\mathbf{x})\}_{\mathbf{x} \in \mathcal{C}}$  is its cover by open sets, we can extract finite subcover  $\{\mathbb{B}_{\delta_{\mathbf{x}_i}^{\varepsilon}}(\mathbf{x}_i)\}_1^n$  and thus prove the lemma, since  $\Phi \subset \bigcup_{i=1, \dots, n} \bigcup_{\mathbf{y} \in \mathbb{B}_{\delta_{\mathbf{x}_i}^{\varepsilon}}(\mathbf{x}_i)} \Phi^{\mathbf{y}}$  is covered by the finite amount of open balls of radius  $\varepsilon$ .



As a result, all we have to prove is that for any  $\varepsilon > 0$  and any  $\mathbf{x} \in \mathcal{C}$  there exists  $\delta_{\mathbf{x}}^{\varepsilon} > 0$  small enough such that  $\bigcup_{\mathbf{y} \in \mathbb{B}_{\delta_{\mathbf{x}}^{\varepsilon}}(\mathbf{x})} \Phi^{\mathbf{y}}$  is a totally bounded subset of  $C([0, T]; \mathbb{R}^d)$ . In the following, we provide the calculations for the

case where  $0 < T_{\mathbf{x}} < \infty$ . These derivations can be easily adapted to cover the general case.

First of all, for any function  $\phi \in \bigcup_{\mathbf{y} \in \mathbb{B}_{\delta_{\mathbf{x}}}(\mathbf{x})} \Phi^{\mathbf{y}}$  if  $\delta_{\mathbf{x}}^{\varepsilon}$  is small enough, the following integral is bounded by a positive constant  $C_1$ :

$$\int_0^T |\dot{\phi}(s)|^2 ds \leq C_1.$$

Indeed, by adding and subtracting  $\nabla V(\phi(s)) + \frac{T_{\mathbf{y}}}{T_{\mathbf{y}+s}} \nabla F * M_{\mathbf{y}}(\phi(s)) + \frac{1}{T_{\mathbf{y}+s}} \int_0^s \nabla F(\phi(s) - \phi(z)) dz$  inside the absolute value, we get:

$$\begin{aligned} \int_0^t |\dot{\phi}(s)|^2 ds &\leq 4 \int_0^t \left| \dot{\phi}(s) + \nabla V(\phi(s)) + \frac{T_{\mathbf{y}}}{T_{\mathbf{y}+s}} \nabla F * M_{\mathbf{y}}(\phi(s)) \right. \\ &\quad \left. + \frac{1}{T_{\mathbf{y}+s}} \int_0^s \nabla F(\phi(s) - \phi(z)) dz \right|^2 ds + 4 \int_0^t |\nabla V(\phi(s))|^2 ds \\ &\quad + 4 \int_0^t \left| \frac{T_{\mathbf{y}}}{T_{\mathbf{y}+s}} \nabla F * M_{\mathbf{y}}(\phi(s)) \right|^2 ds \\ &\quad + 4 \int_0^t \left| \frac{1}{T_{\mathbf{y}+s}} \int_0^s \nabla F(\phi(s) - \phi(z)) dz \right|^2 ds. \end{aligned}$$

Let  $C > 0$  be a generic positive constant. Since  $I_T^{\mathbf{y}}(\phi) \leq 1$  for respective  $\mathbf{y} \in \mathbb{B}_{\delta_{\mathbf{x}}}(\mathbf{x})$  and using Lipschitz continuity of  $\nabla V$  and  $\nabla F$  (Assumption A-1.2), we get

$$\begin{aligned} \int_0^t |\dot{\phi}(s)|^2 ds &\leq 4 + 8\text{Lip}_{\nabla V}^2 \int_0^t |\phi(s)|^2 ds + 8T |\nabla V(0)|^2 \\ &\quad + 4\text{Lip}_{\nabla F}^2 \int_0^t \int_{\mathbb{R}^d} |\phi(s) - z|^2 M_{\mathbf{y}}(dz) ds \\ &\quad + \frac{4t\text{Lip}_{\nabla F}^2}{T_{\mathbf{y}}^2} \int_0^t \int_0^s |\phi(s) - \phi(z)|^2 dz ds \\ &\leq C + C \left( 1 + \frac{C}{T_{\mathbf{y}}^2} \right) \int_0^t |\phi(s)|^2 ds + C \int_{\mathbb{R}^d} |z|^2 M_{\mathbf{y}}(dz). \end{aligned}$$

Since all  $\mathbf{y}$  belong to a ball of radius  $\delta_{\mathbf{x}}^{\varepsilon}$  of  $\mathbf{x}$ , then the following inequalities hold:  $|T_{\mathbf{y}} - T_{\mathbf{x}}| \leq \delta_{\mathbf{x}}^{\varepsilon}$ ,  $\left| \int |z|^2 M_{\mathbf{y}}(dz) - \int |z|^2 M_{\mathbf{x}}(dz) \right| \leq \delta_{\mathbf{x}}^{\varepsilon}$  and

$|\mathbf{P}\mathbf{y} - \mathbf{P}\mathbf{x}| \leq \delta_{\mathbf{x}}^\varepsilon$ . Moreover, without loss of generality, we can choose  $\delta_{\mathbf{x}}^\varepsilon$  to be small enough such that  $\mathbf{T}\mathbf{y} > \mathbf{T}\mathbf{x} - \delta_{\mathbf{x}}^\varepsilon > 0$ .

In addition, since  $\phi \in H_1^{\mathbf{P}\mathbf{x}}$  and by Hölder's inequality, we can bound  $\int_0^t |\dot{\phi}(s)|^2 ds \leq 2T(|\mathbf{P}\mathbf{x}|^2 + \int_0^t \int_0^s |\dot{\phi}(u)|^2 du ds)$ . As a result, there exist constants  $C_2, C_3 > 0$  such that

$$\int_0^t |\dot{\phi}(s)|^2 ds \leq C_2 + C_3 \int_0^t \int_0^s |\dot{\phi}(u)|^2 du ds.$$

So, by Grönwall's inequality, we get

$$\int_0^T |\dot{\phi}(s)|^2 ds \leq C_2 e^{C_3 T} =: C_1. \quad (3.12)$$

Finally, we use the Arzelà–Ascoli theorem to prove that the set

$$\bigcup_{\mathbf{y} \in \mathbb{B}_{\delta_{\mathbf{x}}^\varepsilon}(\mathbf{x})} \Phi^{\mathbf{y}}$$

is totally bounded. We use the bound (3.12) and get the following inequalities showing uniform equicontinuity and pointwise boundedness of functions that belong to  $\bigcup_{\mathbf{y} \in \mathbb{B}_{\delta_{\mathbf{x}}^\varepsilon}(\mathbf{x})} \Phi^{\mathbf{y}}$ .

$$|\phi(t_2) - \phi(t_1)|^2 \leq (t_2 - t_1) \int_{t_1}^{t_2} |\dot{\phi}(s)|^2 ds \leq C_1(t_2 - t_1).$$

This completes the proof. □

## 3.2 Exit-time

First, we present the following assumptions on domain  $G$  and geometry of the potentials that drive the system inside the domain. We remind that the point  $x_0 \in \mathbb{R}^d$  is the (deterministic) initial condition of the main process (3.1) and  $X^0$  is its deterministic limit defined in (3.4). Consider:

**Assumptions A-2.**

1. (domain  $G$ )  $G \subset \mathbb{R}^d$  is an open, bounded, connected set such that  $\partial G = \partial \overline{G}$ . The boundary  $\partial G$  is a smooth  $(d - 1)$ -dimensional hypersurface. The point  $x_0 \in G$ .
2. (point of attraction  $a$ ) There exists  $a \in G$  such that  $\{X_t^0\}_{t \geq 0} \subset G$  and  $X_t^0 \xrightarrow[t \rightarrow \infty]{} a$ .
3. (stability of  $\overline{G}$  under the effective potential) Let  $\phi$  be defined as  $\phi_t^x = x - \int_0^t \nabla V(\phi_s) ds - \int_0^t \nabla F(\phi_s - a) ds$ . For any  $x \in \overline{G}$ ,  $\{\phi_t^x\}_{t > 0} \subset G$  and  $\phi_t^x \xrightarrow[t \rightarrow \infty]{} a$ . Moreover,  $\nabla V(a) = \nabla F(0) = 0$ .
4. (strong attraction around  $a$ ) There exist  $\Delta_\mu, \Delta_x > 0$  small enough and a constant  $K > 0$ , such that for any  $\mu \in \mathbb{B}_{\Delta_\mu}(\delta_a)$  and for any  $x \in B_{\Delta_x}(a)$ ,  $\langle \nabla V(x) + \nabla F * \mu(x); x - a \rangle \geq K|x - a|^2$ .

Assumption A-2.1 defines domain  $G$  and its regularity properties. Assumption A-2.2 defines point  $a$  as a unique attractor of the deterministic process  $X^0$  starting at fixed  $x_0 \in G$ . Since, as was proved in Theorem 3.3, the process  $X^\sigma$  satisfies LDP for any finite time interval  $[0; T]$ , we expect it to be close to its deterministic limit and also to converge towards  $a$  with high probability. As a consequence, we also expect  $\mu_t^\sigma$  to converge towards  $\delta_a$  in a finite time (this is shown in Lemma 3.11). Assumption A-2.3 ensures that for  $\mathbf{x} \in \mathfrak{X}$  such that  $M\mathbf{x}$  is close to  $\delta_a$ , starting from any  $P\mathbf{x} \in G$ , deterministic process  $X^{\mathbf{x},0}$  will be still attracted to  $a$ . Assumption A-2.4 elaborates on this idea and suggests that the attracting forces around  $a$  will be stronger than interaction forces. These facts combined produce Lemma 3.12. These effects also suggest that  $\mu^\sigma$  for small  $\sigma$  should stay close to  $\delta_a$  with high probability at least until exit-time, since inside  $G$  there will always be a force that pushes  $X^\sigma$  towards  $a$ . That is shown in Lemma 3.15.

Now, let us recall the main theorem before proving it. Under the sets of Assumptions A-1 and A-2 the following theorem holds:

**Theorem 3.1.** Let the process  $X^\sigma$  be the unique strong solution of the system (3.1) with initial condition  $x_0 \in G$ . Let  $\tau_G^\sigma := \inf\{t : X_t^\sigma \notin G\}$  denote the first time when  $X^\sigma$  exits the domain  $G$ . Let  $H := \inf_{x \in \partial G} \{V(x) + F(x - a) - V(a)\}$  be the height of the effective potential. Then, the following two results hold:

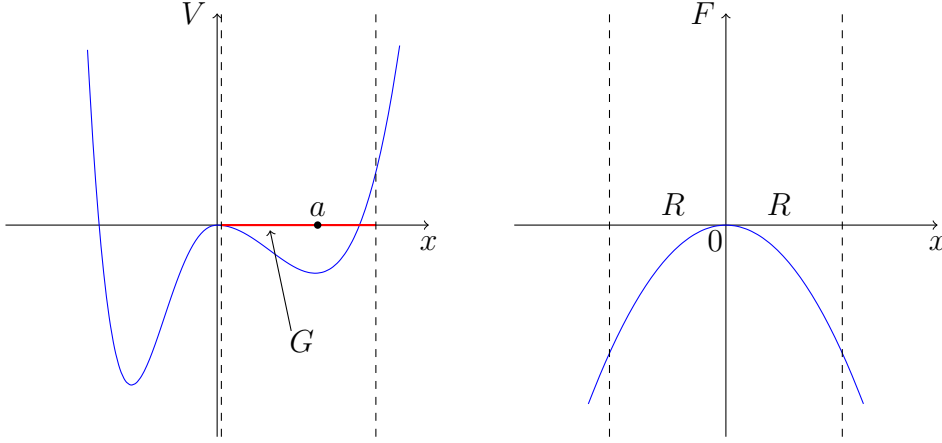


Figure 3.1: Examples of possible  $V$ ,  $F$ , and  $G$  in dimension  $d = 1$ . Despite Assumptions A-1.2 and A-1.3 are not satisfied globally, they are satisfied inside  $G$  and for  $R := \sup\{|z - y| : y, z \in G\}$ , which is enough for exit-time result.

1. Kramers' law: for any  $\delta > 0$

$$\lim_{\sigma \rightarrow 0} \mathbb{P}_{x_0} \left( e^{\frac{2(H-\delta)}{\sigma^2}} < \tau_G^\sigma < e^{\frac{2(H+\delta)}{\sigma^2}} \right) = 1; \quad (3.13)$$

2. Exit-location: for any closed set  $N \subset \partial G$  such that  $\inf_{z \in N} \{V(z) + F(z - a) - V(a)\} > H$  the following limit holds:

$$\lim_{\sigma \rightarrow 0} \mathbb{P}_{x_0} (X_{\tau_G^\sigma}^\sigma \in N) = 0. \quad (3.14)$$

## Auxiliary results

### Initial descent to the point of attraction

The first result that we will use for proving the main theorem describes convergence of the solution to the system (3.1) with vanishing noise towards point  $a$  in a constant time. Moreover, we show that its occupation measure also converges towards  $\delta_a$  at some constant time with high probability for small  $\sigma$ . As a matter of fact, as will be shown in Lemma 3.15, with  $\sigma \rightarrow 0$ , not only  $\mu^\sigma$  converges towards  $\delta_a$  in some constant time, but it also stays in the neighbourhood of  $\delta_a$  at least until exit-time of  $X^\sigma$  from the domain  $G$ . We call this “effect of stabilization of the occupation measure”.

Consider the following lemma:

**Lemma 3.11.** *For any  $\rho > 0$  small enough there exists big enough time  $T_{st}^\rho > 0$  such that*

$$\lim_{\sigma \rightarrow 0} \mathbb{P}_{x_0} \left( X_{T_{st}^\rho}^\sigma \notin B_\rho(a) \right) = 0 \quad (3.15)$$

and

$$\lim_{\sigma \rightarrow 0} \mathbb{P}_{x_0} \left( \mu_{T_{st}^\rho}^\sigma \notin \mathbb{B}_\rho(\delta_a) \right) = 0. \quad (3.16)$$

### Convergence of deterministic process towards $a$

The following lemma claims that the deterministic process  $X^{\mathbf{x},0}$  driven by (3.6) with some suitable initial conditions converges towards point  $a$  at most at some constant time (that depends on how close should the process get to the point  $a$ ). This result is shown for any starting point  $\mathbf{P}\mathbf{x} \in G$ , any  $\mathbf{M}\mathbf{x}$  close enough to  $\delta_a$  ( $\mathbf{M}\mathbf{x} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)$ ), and any  $\mathbf{T}\mathbf{x}$  big enough ( $\mathbf{T}\mathbf{x} \geq T_{st}^\rho$ ). We also show that the occupation measure  $\mu^0$  will not move far from  $\delta_a$  in the process of convergence.

Let  $X^{\mathbf{x},0}$  be the solution of (3.6) and define

$$\mathcal{C}_1 := \{ \mathbf{x} \in \mathfrak{X} : T_{st}^\rho \leq \mathbf{T}\mathbf{x} \leq \infty, \mathbf{M}\mathbf{x} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a), \text{ and } \mathbf{P}\mathbf{x} \in G \}.$$

**Lemma 3.12.** *For any  $\rho > 0$  small enough there exists  $T_1^\rho > 0$  such that for any  $\varepsilon, \vartheta > 0$  small enough and any  $\mathbf{x} \in \mathcal{C}_1$  the following estimations hold:*

1.  $X_s^{\mathbf{x},0} \in G$  and  $\mu_s^{\mathbf{x},0} \in \mathbb{B}_{(1+2\varepsilon)\rho}(\delta_a)$  for any  $0 \leq s < T_1^\rho$ .
2.  $X_s^{\mathbf{x},0} \in B_{(1-\vartheta)\rho}(a)$  and  $\mu_s^{\mathbf{x},0} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)$  for any  $s \geq T_1^\rho$ .

### Attraction of stochastic process towards $a$

The following lemma claims that there exists time  $T_1^\rho > 0$  such that at most at that time with high probability the following event happens. Starting at any point  $\mathbf{P}\mathbf{x}$  inside  $G$  with some suitable initial conditions  $\mathbf{M}\mathbf{x}$  and  $\mathbf{T}\mathbf{x}$ , stochastic process  $X^{\mathbf{x},\sigma}$  either comes sufficiently close to point  $a$  or leaves domain  $G$  entirely.



Define the first time when the process either comes close enough to point  $a$  or leaves the domain  $G$  as the stopping time  $\tau_0^{\mathbf{x}} := \inf\{t \geq 0 : X_t^{\mathbf{x},\sigma} \in B_\rho(a) \cup \partial G\}$ . Let  $\mathcal{C}_1 := \{\mathbf{x} \in \mathfrak{X} : T_{\text{st}}^\rho \leq T\mathbf{x} \leq \infty; M\mathbf{x} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a), \text{ and } P\mathbf{x} \in G\}$ . Then the following lemma holds.

**Lemma 3.13.** *For any  $\rho, \varepsilon > 0$  small enough there exist time  $T_1^\rho > 0$  big enough such that*

$$\limsup_{\sigma \rightarrow 0} \sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}}(\tau_0 > T_1^\rho) = 0. \quad (3.17)$$

### Behaviour in the annulus between $B_\rho(a)$ and $\partial G$

The following lemma claims that the probability that, with  $t$  big enough, the process  $X^{\mathbf{x},\sigma}$  stays in between  $B_\rho(a)$  and  $\mathbb{R}^d \setminus G$  without touching any of those sets, decays exponentially with  $\sigma \rightarrow 0$ . Moreover, with  $t \rightarrow +\infty$  the rate of this decay tends to  $-\infty$ .

Introduce the following stopping time:

$$\gamma := \inf\{t \geq T_{\text{st}}^\rho : \mu_t^{\mathbf{x},\sigma} \notin \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)\},$$

where  $\mu^\sigma$  is part of the solution to equation (3.1). In other words,  $\gamma$  is a random time that indicates when the occupation measure  $\mu^\sigma$  moves significantly far away from  $\delta_a$  after first descending to it by Lemma 3.11. We also recall that  $\tau_0^{\mathbf{x}} := \inf\{t \geq 0 : X_t^\sigma \in B_\rho(a) \cup \partial G\}$ , and  $\mathcal{C}_1 = \{\mathbf{x} \in \mathfrak{X} : T_{\text{st}}^\rho \leq T\mathbf{x} \leq \infty; M\mathbf{x} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a); \text{ and } P\mathbf{x} \in G\}$ .

**Lemma 3.14.** *For any  $\rho, \varepsilon > 0$  small enough the following inequality holds*

$$\lim_{t \rightarrow \infty} \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}}(t < \tau_0 < \gamma) = -\infty. \quad (3.18)$$

### Stabilization of the occupation measure

Given that  $\gamma := \inf\{t \geq T_{\text{st}}^\rho : \mu_t^\sigma \notin \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)\}$ , the following lemma establishes control of the occupation measure  $\mu^\sigma$  until exit-time  $\tau_G^\sigma$ . We recall that  $H = \inf_{z \in \partial G} \{V(z) + F(z - a) - V(a)\}$  and present the following lemma.

**Lemma 3.15.** *For any  $\rho, \varepsilon > 0$  the following limit holds:*

$$\lim_{\sigma \rightarrow 0} \mathbb{P}_{x_0} \left( \gamma \leq \tau_G^\sigma \wedge \exp \left\{ \frac{2(H+1)}{\sigma^2} \right\} \right) = 0.$$

**Exit before nearing  $a$** 

We recall that  $H = \inf_{z \in \partial G} \{V(z) + F(z - a) - V(a)\}$  and state that the following asymptotic upper bound for the probability of exiting the domain  $G$  before approaching even smaller neighbourhood of  $a$  holds. Let  $\mathcal{C}_2 := \{\mathbf{x} \in \mathfrak{X} : T_{st}^\rho \leq T\mathbf{x} \leq \infty, M\mathbf{x} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a), \text{ and } P\mathbf{x} \in S_{(1+\varepsilon)\rho}(a)\}$ . Consider the following lemma:

**Lemma 3.16.** *For any  $\varepsilon > 0$  small enough*

$$\lim_{\rho \rightarrow 0} \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}}(\tau_0 = \tau_G^\sigma, \tau_0 < \gamma) \leq -H.$$

**Control of dynamics for small time intervals**

The following result claims that for any  $\varepsilon > 0$  and  $c > 0$  there exists small enough time  $T(\varepsilon, c)$  such that during this time the probability that the diffusion (3.5) drifts farther away than  $\varepsilon$  from any starting point  $T\mathbf{x} \geq T_{st}^\rho$ ,  $P\mathbf{x} \in G$  and  $M\mathbf{x} \in \mathbb{B}_\rho(\delta_a)$  decreases exponentially with the given rate  $c$ .

**Lemma 3.17.** *For any  $\varepsilon > 0$  and  $c > 0$  there exists time  $T(\varepsilon, c) > 0$  such that for any  $\rho, \varepsilon > 0$  small enough*

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}} \left( \sup_{t \in [0, T(\varepsilon, c)]} |X_t^\sigma - P\mathbf{x}| \geq \varepsilon \right) \leq -c,$$

where  $\mathcal{C}_1 := \{\mathbf{x} \in \mathfrak{X} : T_{st}^\rho \leq T\mathbf{x} \leq \infty, M\mathbf{x} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a), \text{ and } P\mathbf{x} \in G\}$ .

**Uniform lower bound for probability of exit from  $G$** 

The following lemma provides a uniform lower bound for the probability of an exit from domain  $G$  starting from a position that is close to  $a$ , given that the empirical measure of the process does not move far away from  $\delta_a$ .

**Lemma 3.18.** *For any  $\eta > 0$  there exists time  $T_0 > 0$  such that for any  $\rho > 0$  small enough there exists  $T_{st}^\rho > 0$  big enough such that*

$$\liminf_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \inf_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}}(\tau_G^\sigma \leq T_0, \tau_G^\sigma < \gamma) > -(H + \eta),$$

where  $\mathcal{C}_2 := \{\mathbf{x} \in \mathfrak{X} : T_{st}^\rho \leq T\mathbf{x} \leq \infty, M\mathbf{x} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a), \text{ and } P\mathbf{x} \in S_{(1+\varepsilon)\rho}(a)\}$ .

## Proof of the main theorem

### Kramers' law

Given the results of Lemmas 3.12 and 3.15, we expect our process to spend most of its time near  $a$  with  $\sigma$  small enough. In order to have more information about this behaviour, we introduce the following stopping times

$$\begin{aligned}\tau_1 &:= \inf\{t \geq T_{\text{st}}^\rho : X_t^\sigma \in B_\rho(a) \cup \partial G\}, \\ \theta_m &:= \inf\{t \geq \tau_m : X_t^\sigma \in S_{(1+\varepsilon)\rho}(a)\}, \\ \tau_{m+1} &:= \inf\{t \geq \theta_m : X_t^\sigma \in B_\rho(a) \cup \partial G\};\end{aligned}\tag{3.19}$$

for  $m \in \mathbb{N}$  with the convention that  $\theta_{m+1} = \infty$  if  $\tau_m = \tau_G^\sigma$ . We hereby can consider separately the intervals where  $X^\sigma$  is close to point  $a$  and those where it is not the case, i.e. the intervals of the type  $[\tau_k; \theta_k]$  and  $[\theta_k; \tau_{k+1}]$ . Besides, parameter  $\rho$  controls the desired closeness of  $X^\sigma$  to  $a$ . By  $\tau_0^\mathbf{x}$  and  $\theta_0^\mathbf{x}$  we denote the corresponding stopping times regardless of the number of exits that we had before, i.e.

$$\begin{aligned}\tau_0^\mathbf{x} &:= \inf\{t \geq 0 : X_t^{\mathbf{x},\sigma} \in B_\rho(a) \cup \partial G\}; \\ \theta_0^\mathbf{x} &:= \inf\{t : X_t^{\mathbf{x},\sigma} \in S_{(1+\varepsilon)\rho}(a)\}.\end{aligned}$$

*Lower bound.* Take the time  $T_0 := T(\varepsilon\rho, H)$ , where  $T(\varepsilon, c)$  is defined by Lemma 3.17, i.e., it is small enough time such that the probability that the process  $X^\sigma$  moves further away than  $\varepsilon\rho$  from its starting point within time  $T_0$  decays exponentially with rate at least  $H$ .

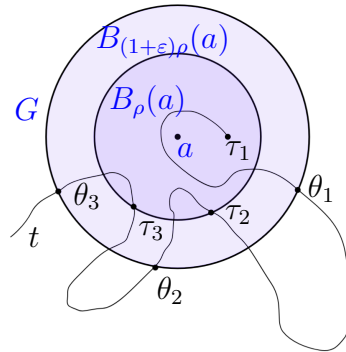


Figure 3.2: Illustration of the definitions of  $\tau_k$  and  $\theta_k$  for  $d = 2$ .

Consider the following event  $\{\tau_G^\sigma \leq kT_0\}$ . We can split this event into two parts accordingly to whether  $\{\gamma \leq \tau_G^\sigma\}$  holds or not. Define  $\mathcal{C}_2 = \{\mathbf{x} \in \mathfrak{X} : T_{\text{st}}^\rho \leq T_{\mathbf{x}} \leq \infty, M_{\mathbf{x}} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a), \text{ and } P_{\mathbf{x}} \in S_{(1+\varepsilon)\rho}(a)\}$ . Note that if the event  $\{\tau_G^\sigma \leq kT_0, \gamma > \tau_G^\sigma\}$  takes place, then either first  $k$  of disjoint events  $\{\tau_i = \tau_G^\sigma, \gamma > \tau_G^\sigma\}$  occur, or at least one period  $\theta_{i+1} - \tau_i$  was less than  $T_0$ , i.e.,

$$\begin{aligned} \mathbb{P}_{x_0}(\tau_G^\sigma \leq kT_0) &\leq \mathbb{P}_{x_0}(\tau_G^\sigma \leq T_{\text{st}}^\rho) + \sum_{i=1}^k \mathbb{P}_{x_0}(\tau_G^\sigma = \tau_i, \gamma > \tau_G^\sigma) \\ &\quad + \mathbb{P}_{x_0}\left(\min_{1 \leq i \leq k-1} \{\theta_i - \tau_{i-1}\} < T_0, \gamma > \tau_G^\sigma\right) \\ &\quad + \mathbb{P}_{x_0}(\gamma \leq \tau_G^\sigma \leq kT_0). \end{aligned}$$

That gives us:

$$\begin{aligned} \mathbb{P}_{x_0}(\tau_G^\sigma \leq kT_0) &\leq \mathbb{P}_{x_0}(\tau_G^\sigma \leq T_{\text{st}}^\rho) + k \sup_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}}(\tau_G^\sigma = \tau_0, \gamma > \tau_G^\sigma) \\ &\quad + (k-1) \sup_{\substack{T_{\text{st}}^\rho \leq T_{\mathbf{x}} \leq \infty; \\ M_{\mathbf{x}} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a); \\ P_{\mathbf{x}} \in B_\rho(a)}} \mathbb{P}_{\mathbf{x}}(\theta_0 < T_0, \gamma > \tau_G^\sigma) \\ &\quad + \mathbb{P}_{x_0}(\gamma \leq \tau_G^\sigma \leq kT_0). \end{aligned}$$

By Lemmas 3.16 and 3.17, we can choose  $\sigma$  small enough such that the respective  $\sup_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}}(\tau_G^\sigma = \tau_0, \gamma > \tau_G^\sigma)$  and  $\sup_{\mathbf{x}} \mathbb{P}_{\mathbf{x}}(\theta_0 < T_0, \gamma > \tau_G^\sigma)$  are less or equal than  $\exp\left\{-\frac{2(H-\delta/2)}{\sigma^2}\right\}$ . We also use Lemma 3.11 in order to deal with the probability of exit before stabilization time  $T_{\text{st}}^\rho$ . Therefore, we establish the following upper bound:

$$\mathbb{P}_{x_0}(\tau_G^\sigma \leq kT_0) \leq o_\sigma + ke^{-\frac{2(H-\delta/2)}{\sigma^2}} + (k-1)e^{-\frac{2(H-\delta/2)}{\sigma^2}} + \mathbb{P}_{x_0}(\gamma \leq \tau_G^\sigma \wedge kT_0),$$

where  $o_\sigma$  is an infinitesimal w.r.t.  $\sigma$ . Finally, choose  $k = \left\lfloor \frac{1}{T_0} \exp\left\{\frac{2(H-\delta)}{\sigma^2}\right\} \right\rfloor$ , where  $\delta > 0$  is small enough. By Lemma 3.15 and simple calculations we establish the following result:

$$\lim_{\sigma \rightarrow 0} \mathbb{P}_{x_0}(\tau_G^\sigma \leq e^{-\frac{2(H-\delta)}{\sigma^2}}) = 0. \quad (3.20)$$

*Upper bound.* Fix  $\delta > 0$ , let  $\eta = \delta/2$  and  $\rho > 0$ ,  $T_0 > 0$  be defined by Lemma 3.18. Let us define event  $A := \left\{ \gamma > \tau_G^\sigma \wedge \exp\left\{\frac{2(H+1)}{\sigma^2}\right\} \right\}$ . Since

$\{\gamma > \tau_G^\sigma\} \subset A$ , the result of the Lemma 3.18 takes the form:

$$\liminf_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \inf_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}}(\tau_G^\sigma \leq T_0, A) > -(H + \eta).$$

By Lemma 3.13, there exists  $T_1^\rho > 0$  such that

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}}(\tau_0 > T_1^\rho, A) \leq \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}}(\tau_0 > T_1^\rho) < 0,$$

where  $\mathcal{C}_1 = \{\mathbf{x} \in \mathfrak{X} : T_{\text{st}}^\rho \leq T_{\mathbf{x}} \leq \infty; M_{\mathbf{x}} \in B_{(1+\varepsilon)\rho}(\delta_a) \text{ and } P_{\mathbf{x}} \in G\}$  and  $\tau_0^\mathbf{x} := \inf\{t \geq 0 : X_t^{\mathbf{x}, \sigma} \in B_\rho(a) \cup \partial G\}$ . Let  $T := T_0 + T_1^\rho$ . Then, for  $\sigma$  small enough we can provide the following bound.

$$\begin{aligned} q &:= \inf_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}}(\tau_G^\sigma \leq T, A) \geq \inf_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}}(\tau_0 \leq T_1^\rho, A) \inf_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}}(\tau_G^\sigma \leq T_0, A) \\ &\geq \exp\left\{-\frac{2(H + \eta)}{\sigma^2}\right\}. \end{aligned} \quad (3.21)$$

This lower bound will help us to calculate the probability of event  $\{\tau_G^\sigma > kT, A\}$  for any  $k \in \mathbb{N}$  in the following way. Consider the following inequalities:

$$\begin{aligned} \mathbb{P}_{\mathbf{x}}(\tau_G^\sigma > (k+1)T, A) &= (1 - \mathbb{P}_{\mathbf{x}}(\tau_G^\sigma \leq (k+1)T, A | \tau_G^\sigma > kT, A)) \\ &\quad \times \mathbb{P}_{\mathbf{x}}(\tau_G^\sigma > kT, A) \\ &\leq (1 - q)\mathbb{P}_{\mathbf{x}}(\tau_G^\sigma > kT, A). \end{aligned}$$

Therefore, by induction in  $k$ , we get:

$$\sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}}(\tau_G^\sigma > kT, A) \leq (1 - q)^k. \quad (3.22)$$

Note that  $\tau_G^\sigma \mathbf{1}_A \leq \left(T + \sum_{k=1}^{\infty} T \mathbf{1}_{\{\tau_G^\sigma > kT\}}\right) \mathbf{1}_A$ . Therefore, by (3.22) and (3.21), we can express the following expectation as:

$$\sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{E}_{\mathbf{x}}(\tau_G^\sigma \mathbf{1}_A) \leq T \sum_{k=0}^{\infty} (1 - q)^k = \frac{T}{q} \leq T \exp\left\{\frac{2(H + \eta)}{\sigma^2}\right\}.$$

By Markov's inequality, we get:

$$\begin{aligned} \sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}}\left(\tau_G^\sigma > \exp\left\{\frac{2(H + \delta)}{\sigma^2}\right\}, A\right) &= \sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}}\left(\tau_G^\sigma \mathbf{1}_A > \exp\left\{\frac{2(H + \delta)}{\sigma^2}\right\}\right) \\ &\leq \exp\left\{-\frac{2(H + \delta)}{\sigma^2}\right\} \sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{E}_{\mathbf{x}}(\tau_G^\sigma \mathbf{1}_A) \\ &\leq T e^{-\delta/\sigma^2} \xrightarrow{\sigma \rightarrow 0} 0. \end{aligned}$$

We remark that by Lemmas 3.11, 3.15 and derivations above, we have

$$\begin{aligned} \mathbb{P}_{x_0} \left( \tau_G^\sigma \geq \exp \left\{ \frac{2(H+\delta)}{\sigma^2} \right\} \right) &\leq \mathbb{P}_{x_0}(\tau_G^\sigma \leq T_{\text{st}}^\rho) + \mathbb{P}_{x_0}(A^c) \\ &\quad + \sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}} \left( \tau_G^\sigma \geq \exp \left\{ \frac{2(H+\delta)}{\sigma^2} \right\}, A \right) \\ &\xrightarrow{\sigma \rightarrow 0} 0. \end{aligned}$$

This concludes the proof.

### Exit-location

Before proving the exit-location result, we provide some observations about the geometry of the domain  $G$  with respect to the effective potential. Let  $W_a(x) = V(x) + F(x-a) - V(a)$  be the effective potential. Note that the level sets  $L_\alpha := \{x \in \mathbb{R}^d : W_a(x) = \alpha\}$  are smooth hypersurfaces, since both  $V$  and  $F$  belong to  $C^2(\mathbb{R}^d; \mathbb{R})$ . Here, solely for simpler notation, we denote by  $L_\alpha$  only the parts of the level sets that intersect with  $\bar{G}$ . By Assumption A-2.3, for  $\alpha = H$  there is in fact only one “part” of the hypersurface intersecting the set  $\bar{G}$ . Namely, it can be shown that  $L_H$  is the boundary of the set  $L_H^- := \{x \in \mathbb{R}^d : W_a(x) \leq H\}$  that itself belongs to  $\bar{G}$  and is a bounded *connected* set (otherwise the Assumption A-2.3 is violated).

Note that there exists a constant  $C_H > 0$  such that for any point on the surface  $x \in L_H$ :  $\langle n(x); -\nabla W_a(x) \rangle \geq C_H$ , where  $n(x)$  is the inner unit normal vector to the surface  $L_H$  at the respective point. Indeed, if it is not the case, then either by the continuity argument there exists  $x^* \in L_H$  such that  $n(x^*) = 0$ , which violates Assumption A-2.3, or all the vectors  $-\nabla W_a(x)$  pin outside of the set  $L_H^-$ , which also contradicts the assumption above. It means that if we plug interior  $\text{Int}(L_H^-)$  and the set  $L_H$  in Assumptions A-2 instead of  $G$  and  $\partial G$ , all the conditions will be satisfied. Then, by the proof of Kramers’ law above, for  $\tau_{\text{Int}(L_H^-)}^{\mathbf{x}, \sigma} := \inf\{t \geq 0 : X_t^{\mathbf{x}, \sigma} \notin \text{Int}(L_H^-)\}$ , we have for any  $\delta > 0$ :

$$\sup_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}} \left( e^{\frac{2(H-\delta)}{\sigma^2}} < \tau_{\text{Int}(L_H^-)}^\sigma < e^{\frac{2(H+\delta)}{\sigma^2}} \right),$$

where  $\mathcal{C}_2 := \{\mathbf{x} \in \mathfrak{X} : T_{\text{st}}^\rho \leq T_{\mathbf{x}} \leq \infty, M_{\mathbf{x}} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a), \text{ and } P_{\mathbf{x}} \in S_{(1+\varepsilon)\rho}(a)\}$ .

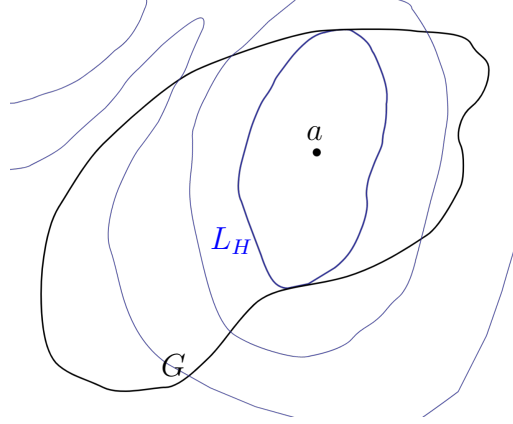


Figure 3.3: Domain  $G$  with level sets of  $W_a = V + F(\cdot - a) - V(a)$ .  $L_H = \{x \in \mathbb{R}^d : W_a(x) = H\}$  is the smallest level set that touches the boundary  $\partial G$

Since  $W_a$  is continuous, there exists a small enlargement of  $L_H$  such that the property  $\langle n(x); -\nabla W_a(x) \rangle \geq \text{Const} > 0$  is still satisfied. Namely, there exists  $\bar{\eta} > 0$  small enough such that for any  $\eta < \bar{\eta}$  there exists a constant  $C_{H+\eta} > 0$  such that  $\langle n(x); -\nabla W_a(x) \rangle \geq C_{H+\eta}$ . By analogy with the previous statement, it means that Assumptions A-2 are satisfied for  $\eta > 0$  small enough and thus, for any  $\delta > 0$ ,

$$\sup_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}} \left( e^{\frac{2(H+\eta-\delta)}{\sigma^2}} < \tau_{\text{Int}(L_{H+\eta}^-)}^\sigma < e^{\frac{2(H+\eta+\delta)}{\sigma^2}} \right).$$

Now we are ready to prove the exit-location result. Take the set  $N \subset \partial G$  described in the theorem. By definition,  $\inf_{z \in N} \{V(z) + F(z - a) - V(a)\} > H$ . That in particular means that we can choose  $\eta < \bar{\eta}$  such that  $H + \eta$  is smaller than  $\inf_{z \in N} \{V(z) + F(z - a) - V(a)\}$ . It means that  $N \cap \text{Cl}(L_{H+\eta}^-)$  is an empty set.

The fact that  $X_{\tau_G^\sigma}^\sigma \in N$  means that, after converging first towards  $a$ , we left the domain  $L_{H+\eta}^-$  before time  $\tau_G^\sigma$ , which is the exit-time from  $G$ . The following inequalities show that it is unlikely. Consider:

$$\mathbb{P}_{x_0}(X_{\tau_G^\sigma}^\sigma \in N) \leq \mathbb{P}_{x_0}(\tau_G^\sigma \leq T_{\text{st}}^\rho) + \sup_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}}(\tau_{L_{H+\eta}^-}^\sigma \leq \tau_G^\sigma).$$

Taking  $\delta = \eta/3$ , it gives us:

$$\begin{aligned}
\mathbb{P}_{x_0}(X_{\tau_G^\sigma}^\sigma \in N) &\leq \mathbb{P}_{x_0}(\tau_G^\sigma \leq T_{st}^\rho) + \sup_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}} \left( \tau_G^\sigma \geq e^{\frac{2(H+\delta)}{\sigma^2}} \right) \\
&\quad + \sup_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}} \left( \tau_{L_{H+\eta}^-}^\sigma \leq \tau_G^\sigma \leq e^{\frac{2(H+\delta)}{\sigma^2}} \right) \\
&\leq \mathbb{P}_{x_0}(\tau_G^\sigma \leq T_{st}^\rho) + \sup_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}} \left( \tau_G^\sigma \geq e^{\frac{2(H+\delta)}{\sigma^2}} \right) \\
&\quad + \sup_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}} \left( \tau_{L_{H+\eta}^-}^\sigma \leq e^{\frac{2(H+\eta-\delta)}{\sigma^2}} \right) \\
&\xrightarrow{\sigma \rightarrow 0} 0,
\end{aligned}$$

by Lemma 3.15 and Kramers' law for  $G$  and  $L_{H+\eta}^-$  proved above.

## Proofs of auxiliary lemmas

### Proof of Lemma 3.11: Initial descent to the point of attraction

Let us fix  $l := \inf_{t>0, z \in \partial G} |X_t^0 - z|$ , the distance between the deterministic path starting at  $x_0$  ( $X^0$ ) and the frontier of the domain  $G$ . Let us also define as  $T_{st}^\rho$  the first time when  $\mu_t^0 \in \mathbb{B}_{\rho/2}(\delta_a)$  as well as  $X_t^0 \in B_{\rho/2}(a)$ , where  $\mu_t^0 := \frac{1}{t} \int_0^t \delta_{X_s^0} ds$  is the empirical measure of the deterministic process. This time obviously exists and is finite, since, by Assumption A-2.2,  $X_t^0 \xrightarrow[t \rightarrow \infty]{} a$ . Let us decrease  $l = l_\rho$  if necessary (it will now depend on  $\rho$ ) such that  $l_\rho < \frac{\rho}{2}$ . Define  $\Phi := \{\phi \in C([0; T_{st}^\rho]) : \|X^0 - \phi\|_\infty \geq l_\rho\}$ . Then, by Theorem 3.3,

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}_{x_0}(X_{T_{st}^\rho}^\sigma \notin B_\rho(a)) \leq - \inf_{\phi \in \Phi} I_{T_{st}^\rho}^{x_0}(\phi)$$

and

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}_{x_0}(\mu_{T_{st}^\rho}^\sigma \notin \mathbb{B}_\rho(\delta_a)) \leq - \inf_{\phi \in \Phi} I_{T_{st}^\rho}^{x_0}(\phi).$$

Since  $I_{T_{st}^\rho}^{x_0}$  has only one minimum which is given by  $X^0$ ,  $I_{T_{st}^\rho}^{x_0}(X^0) = 0$ , and as the distance between  $X^0$  and the set  $\Phi$  is strictly positive, we conclude that  $\inf_{\phi \in \Phi} I_{T_{st}^\rho}^{x_0}(\phi) > 0$ . That completes the proof.



**Proof of Lemma 3.12: Convergence of deterministic process towards  $a$**

Let us recall the reader that  $\Delta_x > 0$  is the positive constant introduced in Assumption A-2.4. We separate the dynamics of the system into the following three parts (see Figure 3.4). First, we show that in time  $\hat{T}_1 > 0$  for any  $P_{\mathbf{x}}$  the process converges inside  $B_{\Delta_x}(a)$ , whereas its occupation measure  $\mu$  does not move far from  $\delta_a$  and stays inside  $\mathbb{B}_{(1+2\varepsilon)\rho}(\delta_a)$  for some small  $\varepsilon > 0$ , that we can decrease by increasing  $t_0$ . Second, we show that inside the set  $B_{\Delta_x}(a)$ , the attraction force of the potential  $V + F * \delta_a$  becomes so strong that, in some time  $\hat{T}_2 > 0$ , it drags  $X^0$  inside small ball  $B_{(1-\vartheta)\rho}(a)$  for some  $0 < \vartheta < 1$ , whereas its occupation measure, as before, does not move far from  $\delta_a$  and still stays inside  $\mathbb{B}_{(1+2\varepsilon)\rho}(\delta_a)$ . Third, we show that for small enough  $\rho > 0$  and after hitting  $B_{(1-\vartheta)\rho}(a)$ ,  $X^0$  stays inside this ball for a long time  $\hat{T}_3$  that is enough to attract the occupation measure  $\mu$  back inside  $\mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)$ . Thus, the required time is represented through the following sum  $T_1^{\rho} := \hat{T}_1 + \hat{T}_2 + \hat{T}_3$ .

*Convergence inside  $B_{\Delta_x}(a)$ .* By Assumption A-2.3, for any  $x \in \overline{G}$ , the flow  $\phi^x$ , generated by the effective potential  $V + F(\cdot - a)$ , converges towards  $a$ . Since, by Assumption A-1.2, both  $\nabla V$  and  $\nabla F$  are Lipschitz continuous and, by Assumption A-2.1,  $\overline{G}$  is a compact set, there exists a uniform upper bound  $\hat{T}_1$  for the time in which  $\phi^x$  converges inside  $B_{\Delta_x/2}(a)$ .

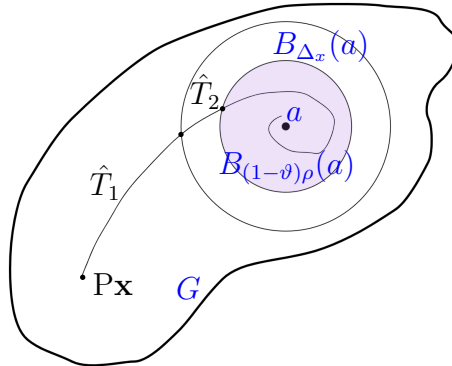


Figure 3.4: Dynamics of  $X^{\mathbf{x},0}$

Let us use this fact to establish the time of convergence of  $X^0$  inside. Define  $\eta_t := |X_t^{\mathbf{x},0} - \phi_t^{\mathbf{P}\mathbf{x}}|$ . Applying Assumption A-1.2, we get

$$\begin{aligned} \eta_t &\leq \text{Lip}_{\nabla V} \int_0^t \eta_s \, ds + \text{Lip}_{\nabla F} \int_0^t \int_{\mathbb{R}^d} |(X_s^{\mathbf{x},0} - z) - (\phi_s^{\mathbf{P}\mathbf{x}} - a)| \mu_s^{\mathbf{x},0}(dz) \, ds \\ &\leq \text{Lip}_{\nabla V} \int_0^t \eta_s \, ds + \text{Lip}_{\nabla F} \left( \int_0^t \eta_s \, ds + \int_0^t \mathbb{W}_2(\mu_s^{\mathbf{x},0}; \delta_a) \, ds \right). \end{aligned}$$

Then, by Grönwall's inequality:

$$\eta_t \leq \int_0^t \mathbb{W}_2(\mu_s^{\mathbf{x},0}; \delta_a) \, ds \cdot \exp\{(\text{Lip}_{\nabla V} + \text{Lip}_{\nabla F})t\}. \quad (3.23)$$

Since  $\mathbf{x} \in \mathcal{C}_1$ ,  $\mathbb{W}_2(\mu_0^{\mathbf{x},0}; \delta_a) \leq (1 + \varepsilon)\rho$ . We can show, by (3.23), that if we take  $\rho$  to be small enough, then  $\mathbb{W}_2(\mu_t^{\mathbf{x},0}; \delta_a) \leq (1 + 1.5\varepsilon)\rho$  at least for  $t \leq \hat{T}_1$ . Indeed, if we introduce a generic time  $T' := \inf\{t \geq 0 : \mathbb{W}_2(\mu_t^{\mathbf{x},0}; \delta_a) > (1 + 1.5\varepsilon)\rho\}$ , we can express, using (3.23), for  $t \leq T' \wedge \hat{T}_1$ :

$$\eta_t \leq (1 + 1.5\varepsilon)\rho \hat{T}_1 e^{(\text{Lip}_{\nabla V} + \text{Lip}_{\nabla F})\hat{T}_1}. \quad (3.24)$$

Since  $\phi_t$  belongs to  $\overline{G}$  for any  $t \geq 0$ , by inequality above, there exists a constant  $C > 0$  such that  $|X_t^{\mathbf{x},0} - a| \leq C$  for any  $t \leq T' \wedge \hat{T}_1$ . At the same time, by the definition of  $\mu^{\mathbf{x},0}$ ,

$$\mathbb{W}_2(\mu_t^{\mathbf{x},0}; \delta_a) \leq \frac{\mathbf{T}\mathbf{x}}{\mathbf{T}\mathbf{x} + t} \mathbb{W}_2(\mathbf{M}\mathbf{x}; \delta_a) + \frac{t}{\mathbf{T}\mathbf{x} + t} \mathbb{W}_2\left(\frac{1}{t} \int_0^t \delta_{X_s^{\mathbf{x},0}} \, ds; \delta_a\right).$$

Therefore, for any  $t \leq T' \wedge \hat{T}_1$ ,

$$\mathbb{W}_2(\mu_t^{\mathbf{x},0}; \delta_a) \leq (1 + \varepsilon)\rho + \frac{\hat{T}_1}{T_{\text{st}}^\rho + \hat{T}_1} C.$$

Note that, without loss of generality, we can increase  $T_{\text{st}}^\rho$  if necessary such that  $\frac{\hat{T}_1}{T_{\text{st}}^\rho + \hat{T}_1} C \leq 0.5\varepsilon\rho$  for any  $\rho > 0$ . Thus, for any  $t \leq T' \wedge \hat{T}_1$ :

$$\mathbb{W}_2(\mu_t^{\mathbf{x},0}; \delta_a) \leq (1 + 1.5\varepsilon)\rho.$$

That includes the time  $t = T' \wedge \hat{T}_1$ . At the same time, by the definition of  $T'$ ,  $\mathbb{W}_2(\mu_{T'}^{\mathbf{x},0}; \delta_a) > (1 + 1.5\varepsilon)\rho$  that means that  $T' \wedge \hat{T}_1 = \hat{T}_1$ .

Now we can conclude that equation (3.24) holds for any  $t \leq \hat{T}_1$ . Thus, we can choose  $\rho > 0$  small enough such that  $\eta_t < \Delta_x/2$  for any  $t \leq \hat{T}_1$ . That proves the uniform (in initial point  $\mathbf{P}\mathbf{x} \in G$ ) convergence inside  $B_{\Delta_x}(a)$ .

*Convergence inside  $B_{(1-\vartheta)\rho}(a)$ .* By Assumption A-2.4, in the set  $B_{\Delta_x}(a)$  attraction forces towards the point  $a$  prevail over the interaction forces. First of all, we decrease  $\rho$  to be small enough such that  $(1+2\varepsilon)\rho < \Delta_\mu$ , where  $\Delta_\mu$  is defined in A-2.4. As was shown before,  $\mathbb{W}_2(\mu_{\hat{T}_1}^{\mathbf{x},0}; \delta_a) \leq (1+1.5\varepsilon)\rho$  and thus  $\mu_{\hat{T}_1}^{\mathbf{x},0} \in \mathbb{B}_{\Delta_\mu}(\delta_a)$ . We can show that moreover  $\mu_{\hat{T}_1+t}^{\mathbf{x},0} \in \mathbb{B}_{(1+2\varepsilon)\rho}(\delta_a)$  long enough such that in  $X^{\mathbf{x},0}$  converges inside  $B_{(1-\vartheta)\rho}(a)$ .

Let  $T'' := \inf\{t \geq 0 : \mathbb{W}_2(\mu_{\hat{T}_1+t}^{\mathbf{x},0}; \delta_a) > (1+2\varepsilon)\rho, \text{ or } |X_{\hat{T}_1+t}^{\mathbf{x},0} - a| > \Delta_x\}$  and consider  $\xi_t := |X_{\hat{T}_1+t}^{\mathbf{x},0} - a|^2$ . By Assumption A-2.4, for any  $t \leq T''$ , its derivative is bounded by:

$$\dot{\xi}_t \leq -2K\xi_t.$$

That guarantees exponentially fast convergence towards 0:

$$\xi_t \leq \Delta_x e^{-2Kt}.$$

Thus, all we need to prove is that  $\hat{T}_2 := \inf\{t \geq 0 : \Delta_x \exp\{-2Kt\} \leq (1-\vartheta)\rho\}$  is less or equal then  $T''$ . We can prove that fact by contradiction. For any  $t \leq T'' \wedge \hat{T}_2$ :

$$\begin{aligned} \mathbb{W}_2(\mu_{\hat{T}_1+t}^{\mathbf{x},0}; \delta_a) &\leq \mathbb{W}_2(\mu_{\hat{T}_1}^{\mathbf{x},0}; \delta_a) + \frac{t}{T_{\mathbf{x}} + \hat{T}_1 + t} \mathbb{W}_2\left(\frac{1}{t} \int_0^t \delta_{X_{\hat{T}_1+s}^{\mathbf{x},0}} ds; \delta_a\right) \\ &\leq (1+1.5\varepsilon)\rho + \frac{\hat{T}_2}{T_{\text{st}}^\rho + \hat{T}_1 + \hat{T}_2} \Delta_x. \end{aligned}$$

For any  $\rho > 0$  we can choose, without loss of generality,  $T_{\text{st}}^\rho$  to be big enough such that  $\frac{\hat{T}_2}{T_{\text{st}}^\rho + \hat{T}_1 + \hat{T}_2} \Delta_x \leq 0.5\varepsilon\rho$ . That gives, for any  $t \leq T'' \wedge \hat{T}_2$ :

$$\mathbb{W}_2(\mu_{\hat{T}_1+t}^{\mathbf{x},0}; \delta_a) \leq (1+2\varepsilon)\rho.$$

Thus, as before,  $\hat{T}_2 < T''$  or else we get a contradiction between the definition of  $T''$  and the inequality above along with the fact that  $\xi$  decreases for any  $t \leq T'' \wedge \hat{T}_2$ .

*Return of the occupation measure back inside*  $\mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)$ . The last time period can be found easily. Note that since  $X^{\mathbf{x},0}$  belongs to  $B_{(1-\vartheta)\rho}(a)$  at time  $\hat{T}_1 + \hat{T}_2$ , as well as  $\mu^{\mathbf{x},0}$  belongs to  $\mathbb{B}_{(1+2\varepsilon)\rho}(\delta_a)$ , then by Assumption A-2.4,  $X_t^{\mathbf{x},0}$  will not leave  $B_{(1-\vartheta)\rho}(a)$  for any  $t \geq \hat{T}_1 + \hat{T}_2$ . Using this fact and estimations that we had on  $\mu^{\mathbf{x},0}$  for time intervals  $[0; \hat{T}_1]$  and  $[\hat{T}_1; \hat{T}_2]$ , we can provide the following bound for  $\mathbb{W}_2(\mu_{\hat{T}_1 + \hat{T}_2 + t}^{\mathbf{x},0}; \delta_a)$ :

$$\begin{aligned} & \frac{\mathbf{T}\mathbf{x}}{\mathbf{T}\mathbf{x} + \hat{T}_1 + \hat{T}_2 + t}(1 + \varepsilon)\rho + \frac{\hat{T}_1}{\mathbf{T}\mathbf{x} + \hat{T}_1 + \hat{T}_2 + t}(1 + 1.5\varepsilon)\rho \\ & + \frac{\hat{T}_2}{\mathbf{T}\mathbf{x} + \hat{T}_1 + \hat{T}_2 + t}(1 + 2\varepsilon)\rho + \frac{t}{\mathbf{T}\mathbf{x} + \hat{T}_1 + \hat{T}_2 + t}(1 - \vartheta)\rho \\ & = \rho + \frac{\mathbf{T}\mathbf{x}\varepsilon + 1.5\hat{T}_1\varepsilon + 2\hat{T}_2\varepsilon - t\vartheta}{\mathbf{T}\mathbf{x} + \hat{T}_1 + \hat{T}_2 + t}\rho \\ & \leq \rho + \frac{(\mathbf{T}\mathbf{x} + \hat{T}_1 + \hat{T}_2 + t) + \hat{T}_2 - t(1 + \frac{\vartheta}{\varepsilon})}{\mathbf{T}\mathbf{x} + \hat{T}_1 + \hat{T}_2 + t}\varepsilon\rho \\ & \leq (1 + \varepsilon)\rho + \frac{\hat{T}_2 + 0.5\hat{T}_1 - t(1 + \frac{\vartheta}{\varepsilon})}{T_{\text{st}}^\rho + \hat{T}_1 + \hat{T}_2 + t}\varepsilon\rho. \end{aligned}$$

We just need to choose  $\hat{T}_3$  big enough such that  $\left(\hat{T}_2 + 0.5\hat{T}_1 - \hat{T}_3(1 + \frac{\vartheta}{\varepsilon})\right) < 0$ .

We complete the prove by choosing  $T_1^\rho = \hat{T}_1 + \hat{T}_2 + \hat{T}_3$ .

### **Proof of Lemma 3.13: Attraction of stochastic process towards $a$**

According to Lemma 3.12, there exists an upper bound for time of convergence of deterministic process  $X^{\mathbf{x},0}$  inside  $B_{(1-\vartheta)\rho}$  for some  $0 < \vartheta < 1$  and for any  $\mathbf{T}\mathbf{x} \geq T_{\text{st}}^\rho$ ,  $\mathbf{M}\mathbf{x} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)$ , and  $\mathbf{P}\mathbf{x} \in G$ . Denote this time as  $T_1^\rho$ .

Let us define

$$\Psi^{x_0} := \{\psi \in C([0, T_1^\rho]; \mathbb{R}^d), \psi(0) = x_0, \psi(s) \in \text{Cl}(G \setminus B_\rho(a)) \forall s \leq T_1^\rho\}.$$

The following inclusion of the events holds:  $\{\tau_0^{\mathbf{x}} \geq T_1^\rho\} \subset \{X^{\mathbf{x},\sigma} \in \Psi^{\mathbf{P}\mathbf{x}}\}$  (the definition of  $\tau_0^{\mathbf{x}}$  was presented on page 63). Let  $\Psi = \bigcup_{x_0 \in G} \Psi^{x_0}$ . Note that  $\Psi$  is a closed set and the following enlargement of  $\mathcal{C}_1$ :  $\mathcal{C}'_1 := \{\mathbf{x} \in$

$\mathfrak{X} : T_{\text{st}}^\rho \leq T_{\mathbf{x}} \leq \infty, P_{\mathbf{x}} \in \overline{G}$ , and,  $M_{\mathbf{x}} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)$  is a compact set by Lemma 3.9. By Lemma 3.6 and the inclusion of the events, we get

$$\begin{aligned} \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}}(\tau_0 > T_1^\rho) \\ \leq \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup_{\mathbf{x} \in \mathcal{C}'_1} \mathbb{P}_{\mathbf{x}}(X^\sigma \in \Psi) \leq - \inf_{\mathbf{x} \in \mathcal{C}'_1} \inf_{\phi \in \Psi} I_{T_1^\rho}^{\mathbf{x}}(\phi). \end{aligned}$$

What is left to prove is that

$$\inf_{\mathbf{x} \in \mathcal{C}'_1} \inf_{\phi \in \Psi} I_{T_1^\rho}^{\mathbf{x}}(\phi) > 0. \quad (3.25)$$

We can not apply Corollary 3.8 directly since  $\Psi$  is not necessarily compact. Let  $\Psi_1^{\mathbf{x}} := \{\psi \in \Psi : I_{T_1^\rho}^{\mathbf{x}}(\psi) \leq 1\}$ . Of course, if all  $\Psi_1^{\mathbf{x}}$  are empty, then expression in (3.25) is indeed strictly greater than 0. If it is not the case, then the sets  $\{\Psi_1^{\mathbf{x}}\}_{\mathbf{x} \in \mathcal{C}'_1}$  satisfy conditions of Lemma 3.10, and, therefore, union  $\Psi_1 := \bigcup_{\mathbf{x} \in \mathcal{C}'_1} \Psi_1^{\mathbf{x}}$  is precompact, and  $\Psi_2 = \overline{\Psi_1}$  is a compact set. Thus, by Corollary 3.8, there exists some  $\mathbf{x}^* \in \mathcal{C}'_1$ ,  $\phi^* \in \Psi_2$  such that

$$\inf_{\mathbf{x} \in \mathcal{C}'_1} \inf_{\phi \in \Psi_2} I_{T_1^\rho}^{\mathbf{x}}(\phi) = I_{T_1^\rho}^{\mathbf{x}^*}(\phi^*).$$

By Lemma 3.12, for any  $\mathbf{x} \in \mathcal{C}'_1$ , the corresponding deterministic trajectories converge inside  $B_{(1-\vartheta)\rho}(a)$  in time  $T_1^\rho$  for some  $\vartheta$ . Thus, for any  $\psi \in \Psi$ :  $\|\psi - X^{\mathbf{x},0}\|_\infty > \vartheta$  for any  $\mathbf{x} \in \mathcal{C}'_1$ . The same bound obviously holds for  $\Psi_2$  since it is a closure of some subset of  $\Psi$ . Since the deterministic trajectories are the unique functions in  $C([0, T_1^\rho]; \mathbb{R}^d)$  for which  $I_{T_1^\rho}^{\mathbf{x}}(X^{\mathbf{x},0}) = 0$ , we conclude that

$$\inf_{\mathbf{x} \in \mathcal{C}'_1} \inf_{\phi \in \Psi} I_{T_1^\rho}^{\mathbf{x}}(\phi) > I_{T_1^\rho}^{\mathbf{x}^*}(\phi^*) > 0.$$

And this concludes the proof of the lemma.

### **Proof of Lemma 3.14: Behaviour in the annulus between $B_\rho(a)$ and $\partial G$**

The idea of the proof is to show that, since  $T_1^\rho$  represents the time in which the noiseless process converges inside  $B_\rho(a)$ , after each time interval of length  $T_1^\rho$  it should be more and more unlikely that diffusion  $X^\sigma$  did not follow the deterministic path even once.

We introduce the following set of functions whose path stay inside the annulus  $\text{Cl}(G \setminus B_\rho(a))$ :

$$\begin{aligned} \Psi_t &:= \{\psi \in C([0, t]; \mathbb{R}^d) : \psi_s \in \text{Cl}(G \setminus B_\rho(a)) \\ &\quad \text{and } \frac{1}{s} \int_0^s \delta_{\psi_u} du \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a) \forall s \in [0, t]\}. \end{aligned}$$

Obviously, the following inclusion of events takes place  $\{\tau_0^{\mathbf{x}} > t, \gamma > \tau_0^{\mathbf{x}}\} \subset \{X^{\mathbf{x}, \sigma} \in \Psi_t\}$ . By Lemma 3.6 for any  $t < \infty$ :

$$\lim_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}}(X^\sigma \in \Psi_t) \leq - \inf_{\psi \in \Psi_t} I_t(\psi),$$

where  $I_t(\psi) := \inf_{\mathbf{x} \in \mathcal{C}_1} I_t^{\mathbf{x}}(\psi)$ . That means that it is enough to show that  $\lim_{t \rightarrow \infty} \inf_{\psi \in \Psi_t} I_t(\psi) = \infty$ . Let us assume that it is not true and there exists  $M < \infty$  such that for any  $n$  there exists some function  $\psi^n \in \Psi_{nT_1^\rho}$  such that  $I_{nT_1^\rho}(\psi^n) \leq M$ , where  $T_1^\rho$  is defined by Lemma 3.13. Then, we can separate the path of  $\psi^n$  into  $n$  parts and establish the following lower bound:

$$M \geq I_{nT_1^\rho}(\psi^n) \geq \sum_{k=0}^{n-1} I_{T_1^\rho}^{\mathbf{x}_k^n}(\psi^{n,k}) \geq n \min_{k \leq n} I_{T_1^\rho}^{\mathbf{x}_k^n}(\psi^{n,k}), \quad (3.26)$$

where  $\mathbf{x}_k^n \in \mathfrak{X}$  are such that:

$$\begin{aligned} \text{T}\mathbf{x}_k^n &= \text{T}\mathbf{x}_0 + kT_1^\rho, \\ \text{P}\mathbf{x}_k^n &= \psi^{n,k-1}(kT_1^\rho), \\ \text{M}\mathbf{x}_k^n &= \frac{\text{T}\mathbf{x}_0}{\text{T}\mathbf{x}_0 + kT_1^\rho} \text{M}\mathbf{x}_0 + \frac{1}{\text{T}\mathbf{x}_0 + kT_1^\rho} \int_0^{kT_1^\rho} \delta_{\psi_s^n} ds; \end{aligned}$$

and  $\psi^{n,k} \in C([0, T_1^\rho]; \mathbb{R}^d)$  are the corresponding peaces of  $\psi^n$  of length  $T_1^\rho$ , i.e.,  $\psi^{n,k}(s) = \psi^n(kT_1^\rho + s)$  for  $s \in [0, T_1^\rho]$ .

For equation (3.26) to hold for any  $n$  it is required that there is a sequence of functions  $\{\psi^{n,k_n}\}_{n=1}^\infty$  such that  $I_{T_1^\rho}^{\mathbf{x}_k^n}(\psi^{n,k_n}) \rightarrow 0$ . Thus, after some  $n_0$  all of these functions belong to the set

$$\Phi = \bigcup_{\mathbf{x} \in \mathcal{C}_1} \Phi^{\mathbf{x}},$$

where  $\Phi^{\mathbf{x}} := \{f \in C([0, T_1^\rho]) : I^{\mathbf{x}}(f) \leq 1\}$ . It also means that we have  $\inf_{\phi \in \Phi} I_{T_1^\rho}(\phi) = 0$ . We use the same logic as in the proof of Lemma 3.13 and Lemma 3.19. If all the sets  $\{\Phi_{\mathbf{x}}\}_{\mathbf{x} \in \mathcal{C}}$  are empty, then we get the contradiction with the fact that infimum of rate functions over this set should be equal to 0. In the other case, we can apply Lemma 3.10 to conclude that  $\Phi$  is precompact set, which makes  $\Phi_1 := \overline{\Phi}$  to be a compact set. By Corollary 3.8, there exist  $x^* \in \mathcal{C}_1$  and  $\phi^* \in \Phi_1$  such that

$$\inf_{\phi \in \Phi_1} \inf_{\mathbf{x} \in \mathcal{C}_1} I_{T_1^\rho}^{\mathbf{x}}(\phi) = I_{T_1^\rho}^{\mathbf{x}^*}(\phi^*).$$

By Lemma 3.11, all deterministic processes  $X^{\mathbf{x},0}$  converge inside  $B_{(1-\vartheta)\rho}(a)$  for some  $\vartheta > 0$  in time  $T_1^\rho$ . Thus, for any  $\phi \in \Psi_{T_1^\rho}$  (and as a consequence in  $\Phi_1$ ):  $\|\phi - X^{\mathbf{x},0}\|_\infty \geq \vartheta$ . That means that  $I_{T_1^\rho}^{\mathbf{x}^*}(\phi^*)$  is strictly positive, which contradicts existence of  $M < \infty$  in (3.26) and, thus, proves the lemma.

### Proof of Lemma 3.15: Stabilization of the occupation measure

The proof of this result requires an additional Lemma 3.19. This lemma claims that for any choice of constant  $T_2 > 0$  we can make  $\sigma$  small enough such that with high probability the stochastic process  $X^{\mathbf{x},\sigma}$  spends inside small neighbourhood of the point of attraction  $a$  at least time  $T_2$ .

We remind that  $\theta_0^{\mathbf{x}} := \inf\{t : X_t^{\mathbf{x},\sigma} \in S_{(1+\varepsilon)\rho}(a)\}$  is the time that the process starting in some point inside  $B_\rho(a)$  spends inside the ball  $B_{(1+\varepsilon)\rho}(a)$ . Let  $\mathcal{C}_3 := \{\mathbf{x} \in \mathfrak{X} : T_{st}^\rho \leq T_{\mathbf{x}} \leq \infty; M_{\mathbf{x}} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a); P_{\mathbf{x}} \in B_\rho(a)\}$ . The following lemma holds.

**Lemma 3.19.** *For any  $\rho, \varepsilon > 0$  small enough and for any constant  $T_2 > 0$  the following limit holds*

$$\lim_{\sigma \rightarrow 0} \sup_{\mathbf{x} \in \mathcal{C}_3} \mathbb{P}_{\mathbf{x}}(\theta_0 < T_2) = 0. \quad (3.27)$$

*Proof.* The proof of the following lemma follow the same logic as the one of Lemma 3.13.

By Lemma 3.12, the deterministic process  $X^{\mathbf{x},0}$  starting with any initial conditions  $\mathbf{x} \in \mathfrak{X}$ :  $T_{st}^\rho \leq T_{\mathbf{x}} \leq \infty$ ,  $P_{\mathbf{x}} \in B_\rho(a)$ , and  $M_{\mathbf{x}} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)$  stays inside  $B_\rho(a)$  for any  $t \geq 0$ .

Fix  $T_2$  and define  $\Psi^y := \{\psi \in C([0, T_2]; \mathbb{R}^d) : \psi(0) = y, |\psi(s) - a| \geq (1 + \varepsilon)\rho, \text{ for some } s \leq T_2\}$ . Obviously, the following inclusion holds:  $\{\theta_0^x \geq T_2\} \subset \{X^{x, \sigma} \in \Psi^{P^x}\}$ . Let  $\Psi := \text{Cl}(\bigcup_{y \in B_\rho(a)} \Psi^y)$ . Note that, by Lemma 3.9, the set  $\mathcal{C}_3$  is compact. Then, by Lemma 3.6, we get

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup_{\mathbf{x} \in \mathcal{C}_3} \mathbb{P}_{\mathbf{x}}(\theta_0 > T_2) \leq - \inf_{\mathbf{x} \in \mathcal{C}_3} \inf_{\phi \in \Psi} I_{T_2}^{\mathbf{x}}(\phi).$$

We have to show now that

$$\inf_{\mathbf{x} \in \mathcal{C}_3} \inf_{\phi \in \Psi} I_{T_2}^{\mathbf{x}}(\phi) > 0.$$

In order to apply Corollary 3.8, consider  $\Psi_1^x := \{\psi \in \Psi : I_{T_2}^x(\psi) \leq 1\}$ . If all the sets are empty then the inequality above is trivially satisfied. If not, the family of sets  $\{\Psi_1^x\}_{x \in \mathcal{C}_3}$  satisfies conditions of Lemma 3.10. Thus, the union  $\bigcup_{x \in \mathcal{C}_3} \Psi_1^x$  is precompact and  $\Psi_2 := \overline{\Psi_1}$  is a compact set. Thus, by Corollary 3.8, there exist some  $\mathbf{x}^* \in \mathcal{C}_3$  and  $\phi^* \in \Psi_2$  such that

$$\inf_{\mathbf{x} \in \mathcal{C}_3} \inf_{\phi \in \Psi} I_{T_2}^{\mathbf{x}}(\phi) = I_{T_2}^{\mathbf{x}^*}(\phi^*).$$

As was pointed out before,  $X^{x,0}$  starting with initial conditions  $\mathbf{x} \in \mathcal{C}_3$  stays inside  $B_\rho(a)$  for any  $0 \leq t \leq T_2$ . By definition of  $\Psi$ , any function  $X^{x,0}$  has positive distance of at least  $\varepsilon\rho$  to  $\Psi$ . Same holds for  $\Psi_2$ . Since  $X^{x,0}$  are unique minimizers of  $I_{T_2}^x$  for corresponding initial conditions  $\mathbf{x}$  and  $I_{T_2}^x(X^{x,0}) = 0$ ,

$$I_{T_2}^{\mathbf{x}^*}(\phi^*) > 0.$$

This completes the proof. □

Now we are ready to prove Lemma 3.15 itself.

*Proof of Lemma 3.15. Dynamics of  $\mu_t^\sigma$ .* We separate the path of  $\mu_t^\sigma$  into two parts: when it belongs to the ball of radius  $\rho$  and when big excursions occur. Let  $\mathbf{x} \in \mathcal{C}_4 := \{\mathbf{x} \in \mathfrak{X} : T_{st}^\rho \leq T\mathbf{x} \leq \infty, M\mathbf{x} \in \mathbb{B}_\rho(\delta_a), \text{ and } P\mathbf{x} \in$



$B_\rho(a)$ . Let  $\mu_0 = M\mathbf{x}$  and  $t_0 = T\mathbf{x}$ . Consider the following equations for  $t \in [\theta_m; \tau_{m+1}]$  (the definitions of  $\theta_k$  and  $\tau_k$  were presented on page 63).

$$\begin{aligned} \mathbb{W}_2(\mu_t^\sigma; \delta_a) &\leq \frac{t_0}{t_0 + t} \mathbb{W}_2(\mu_0; \delta_a) + \frac{t}{t_0 + t} \mathbb{W}_2\left(\frac{1}{t} \int_0^t \delta_{X_s} ds; \delta_a\right) \\ &= \frac{t_0}{t_0 + t} \mathbb{W}_2(\mu_0; \delta_a) \\ &\quad + \sum_{k=1}^{m-1} \frac{\tau_{k+1} - \theta_k}{t_0 + t} \mathbb{W}_2\left(\frac{1}{\tau_{k+1} - \theta_k} \int_{\theta_k}^{\tau_{k+1}} \delta_{X_s} ds; \delta_a\right) \\ &\quad + \sum_{k=1}^m \frac{\theta_k - \tau_k}{t_0 + t} \mathbb{W}_2\left(\frac{1}{\theta_k - \tau_k} \int_{\tau_k}^{\theta_k} \delta_{X_s} ds; \delta_a\right) \\ &\quad + \frac{t - \theta_m}{t_0 + t} \mathbb{W}_2\left(\frac{1}{t - \theta_m} \int_{\theta_m}^t \delta_{X_s} ds; \delta_a\right). \end{aligned}$$

That gives us:

$$\begin{aligned} \mathbb{W}_2(\mu_t^\sigma; \delta_a) &\leq \frac{t_0}{t_0 + t} \rho + \left( \sum_{k=1}^{m-1} \frac{\tau_{k+1} - \theta_k}{t_0 + t} + \frac{t - \theta_m}{t_0 + t} \right) R \\ &\quad + \left( \sum_{k=1}^m \frac{\theta_k - \tau_k}{t_0 + t} \right) (1 + \varepsilon) \rho, \end{aligned}$$

where  $R := \sup_{z \in \partial G} |z - a|$  is the maximal distance between the point  $a$  and the frontier of  $G$ . Let us define  $\mathcal{T}_{\text{out}}(m) := \sum_{k=1}^m (\tau_{k+1} - \theta_k)$  that is the total amount of time spent significantly outside of the ball  $B_\rho(a)$  after  $m$  full exits, as well as  $\mathcal{T}_{\text{in}}(m) := \sum_{k=1}^m (\theta_k - \tau_k)$ . Taking into account this notation, for  $t \in [\theta_m; \tau_{m+1}]$  we express

$$\begin{aligned} \mathbb{W}_2(\mu_t^\sigma; \delta_a) &\leq \frac{t_0}{t_0 + t} \rho + \frac{\mathcal{T}_{\text{out}}(m)}{t_0 + t} R + \frac{t - \mathcal{T}_{\text{out}}(m-1)}{t_0 + t} (1 + \varepsilon) \rho \\ &\leq \rho + \frac{\mathcal{T}_{\text{out}}(m)}{t_0 + t} R \leq \rho + \frac{\mathcal{T}_{\text{out}}(m)}{\mathcal{T}_{\text{in}}(m)} R. \end{aligned} \tag{3.28}$$

Here we emphasize that  $\mu_t^\sigma$ ,  $\tau_k$ ,  $\theta_k$  and, as a consequence,  $\mathcal{T}_{\text{out}}(m)$  depend also on elementary event  $\omega$ . Now, in order to prove that  $\gamma > \tau_G^\sigma \wedge \exp\left\{\frac{2(H+1)}{\sigma^2}\right\}$  with high probability, it suffices to show that  $\mathcal{T}_{\text{out}}(m)$  constitutes such a small part of  $t$  that it will not be able to move the occupation measure  $\mu_t^\sigma$  significantly away from  $\delta_a$ .

*Control of  $\mathcal{T}_{\text{in}}$ .* By Lemma 3.14 if we choose  $\sigma$  small enough, there exists  $T_1^\rho > 0$  big enough such that

$$\sup_{\mathbf{x} \in \mathcal{C}_4} \mathbb{P}(\tau_0 > T_1^\rho, \gamma > \tau_0) < \exp\left\{-\frac{8(H+1)}{\sigma^2}\right\}. \quad (3.29)$$

Define  $T_2^\rho > 0$  as a number that is big enough such that:

$$\frac{T_1^\rho}{T_2^\rho} R < \varepsilon \rho. \quad (3.30)$$

We recall that Lemma 3.19 establishes the following asymptotic behaviour for small  $\sigma$ :

$$\sup_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}}(\theta_0 < T_2^\rho) = o_\sigma, \quad (3.31)$$

where  $o_\sigma$  is an infinitesimal w.r.t.  $\sigma$ . Thus, we can get the following lower bound for time spent inside the ball  $B_{(1+\varepsilon)\rho}(a)$ . For any  $\mathbf{x} \in \mathcal{C}_4$ ,

$$\begin{aligned} \mathbb{P}_{\mathbf{x}}(\mathcal{T}_{\text{in}}(m) < \frac{m}{2} T_2^\rho, \gamma \geq \tau_{m+1}) \\ &\leq \mathbb{P}_{\mathbf{x}}(\#\{i \leq m : \theta_i - \tau_i < T_2^\rho(\sigma)\} > \frac{m}{2}, \gamma \geq \tau_{m+1}) \\ &\leq \sum_{k=\lceil \frac{m}{2} \rceil}^m \sum_{(i_1, \dots, i_k)} \mathbb{P}_{\mathbf{x}}\left(\bigcap_j \{\theta_{i_j} - \tau_{i_j} < T_2^\rho(\sigma)\}, \gamma \geq \tau_{m+1}\right), \end{aligned} \quad (3.32)$$

where the summation is taken with respect to all tuples of the form  $(i_1, \dots, i_k)$  for  $i_1 < \dots < i_k$ . Note that we can roughly estimate the number of such tuples to be less than  $2^m$ . We also emphasize that each respective probability can be expressed as

$$\begin{aligned} &\mathbb{P}_{\mathbf{x}}\left(\bigcap_{j=1}^k \{\theta_{i_j} - \tau_{i_j} < T_2^\rho\}, \gamma \geq \tau_{m+1}\right) \\ &= \prod_{j=1}^k \mathbb{P}_{\mathbf{x}}\left(\theta_{i_j} - \tau_{i_j} < T_2^\rho, \gamma \geq \tau_{m+1} \mid \bigcap_{z \leq j} \{\theta_{i_z} - \tau_{i_z} < T_2^\rho\}\right) \\ &\leq \left(\sup_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}}(\theta_0 < T_2^\rho)\right)^k. \end{aligned}$$

By equation (3.31) we conclude the following bound for probability (3.32). For any  $m \in \mathbb{N}$  and  $\mathbf{x} \in \mathcal{C}_4$ ,

$$\begin{aligned} \mathbb{P}_{\mathbf{x}}(\mathcal{T}_{\text{in}}(m) < \frac{m}{2}T_2^\rho, \gamma \geq \tau_{m+1}) &\leq \sum_{k=\lceil \frac{m}{2} \rceil}^m 2^m \left( \sup_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}}(\theta_0 < T_2^\rho) \right)^k \\ &\leq 2^m \frac{o_\sigma^{\lceil \frac{m}{2} \rceil} (o_\sigma^{\lceil \frac{m}{2} \rceil} - 1)}{o_\sigma - 1} \leq \frac{o_\sigma^{\lceil \frac{m}{2} \rceil}}{1 - o_\sigma}. \end{aligned} \quad (3.33)$$

*Control of the number of excursion.* Inequality (3.33) also provides us with the following upper bound for  $\tau_m$ . Let  $m^* := \lceil \frac{2}{T_2^\rho} \rceil \exp\left\{\frac{2(H+1)}{\sigma^2}\right\}$  and consider

$$\begin{aligned} \mathbb{P}_{\mathbf{x}}\left(\tau_{m^*} < \exp\left\{\frac{2(H+1)}{\sigma^2}\right\}\right) &\leq \mathbb{P}_{\mathbf{x}}\left(\mathcal{T}_{\text{in}}(m^*) < \exp\left\{\frac{2(H+1)}{\sigma^2}\right\}\right) \\ &\leq \frac{o_\sigma^{\lceil \frac{m^*}{2} \rceil}}{1 - o_\sigma}. \end{aligned} \quad (3.34)$$

Note that  $m^*$  tends to infinity with  $\sigma \rightarrow 0$ . It means that the probability in (3.34) tends to zero, which provides us with an asymptotic upper bound for  $\tau_m$  knowing that  $m$  is large enough. Inequality (3.34) in particular means that the probability that there were more than  $m^*$  (that grows exponentially fast with  $\sigma$ ) excursions before time  $\exp\left\{\frac{2(H+1)}{\sigma^2}\right\}$  is very small.

*Control of  $\mathcal{T}_{\text{out}}$ .* For time spent significantly outside of  $B_\rho(a) - \mathcal{T}_{\text{out}}(m)$  – we provide the following simple bound. For any  $\mathbf{x} \in \mathcal{C}_1$ ,

$$\begin{aligned} \mathbb{P}_{\mathbf{x}}(\mathcal{T}_{\text{out}}(m) > mT_1^\rho, \gamma \geq \tau_m) &\leq \mathbb{P}_{\mathbf{x}}(\{\exists i \leq m : \tau_i - \theta_i > T_1^\rho\}, \gamma \geq \tau_m) \\ &\leq m \sup_{\mathbf{x} \in \mathcal{C}_1} \mathbb{P}_{\mathbf{x}}(\tau_0 > T_1^\rho) \leq m \exp\left\{-\frac{8(H+1)}{\sigma^2}\right\}, \end{aligned} \quad (3.35)$$

where we obtain the last inequality by (3.29).

*Control of  $\gamma$ .* Note that, by definition of  $\gamma$ ,  $\mathbb{W}_2(\mu_\gamma; \delta_a) \geq (1 + \varepsilon)\rho$  a.s.

Consider the following inequalities. For any  $\mathbf{x} \in \mathcal{C}_1$ ,

$$\begin{aligned}
& \mathbb{P}_{\mathbf{x}} \left( \gamma \leq \tau_G^\sigma \wedge \exp \left\{ \frac{2(H+1)}{\sigma^2} \right\} \right) \\
&= \mathbb{P}_{\mathbf{x}} \left( \mathbb{W}_2(\mu_\gamma; \delta_a) \geq (1+\varepsilon)\rho, \gamma \leq \tau_G^\sigma \wedge \exp \left\{ \frac{2(H+1)}{\sigma^2} \right\} \right) \\
&\leq \sum_{m=1}^{\infty} \mathbb{P}_{\mathbf{x}} \left( \mathbb{W}_2(\mu_\gamma; \delta_a) \geq (1+\varepsilon)\rho, \tau_m \leq \gamma \leq \exp \left\{ \frac{2(H+1)}{\sigma^2} \right\} \right) \\
&\leq \sum_{m=1}^{m^*} \mathbb{P}_{\mathbf{x}} \left( \mathbb{W}_2(\mu_\gamma; \delta_a) \geq (1+\varepsilon)\rho, \tau_m \leq \gamma \leq \exp \left\{ \frac{2(H+1)}{\sigma^2} \right\} \right) \\
&\quad + \sum_{m=m^*}^{\infty} \mathbb{P}_{\mathbf{x}} \left( \tau_m \leq \exp \left\{ \frac{2(H+1)}{\sigma^2} \right\} \right) =: A + B.
\end{aligned}$$

Let us first deal with the term  $B$ . By (3.34), we have

$$B \leq \sum_{m=m^*}^{\infty} \frac{o_\sigma^{\lceil \frac{m}{2} \rceil}}{o_\sigma - 1} \leq \frac{o_\sigma^{\lceil \frac{m^*}{2} \rceil}}{(o_\sigma - 1)^2} \xrightarrow{\sigma \rightarrow 0} 0.$$

Consider now  $A$ . Separate each probability inside the sum the following way

$$\begin{aligned}
& \mathbb{P}_{\mathbf{x}} \left( \mathbb{W}_2(\mu_\gamma; \delta_a) \geq (1+\varepsilon)\rho, \tau_m \leq \gamma \leq \exp \left\{ \frac{2(H+1)}{\sigma^2} \right\} \right) \\
&\leq \sum_{k=m}^{m^*} \mathbb{P}_{\mathbf{x}} \left( \mathbb{W}_2(\mu_\gamma; \delta_a) \geq (1+\varepsilon)\rho, \mathcal{T}_{\text{out}}(k) < kT_1^\rho, \right. \\
&\quad \left. \mathcal{T}_{\text{in}}(k) > \frac{k}{2}T_2^\rho, \gamma \in [\theta_k; \tau_{k+1}] \right) \\
&\quad + \sum_{k=m}^{m^*} \mathbb{P}_{\mathbf{x}}(\mathcal{T}_{\text{out}}(k) \geq kT_1^\rho, \gamma \geq \tau_k) + \sum_{k=m}^{m^*} \mathbb{P}_{\mathbf{x}}(\mathcal{T}_{\text{in}}(k) \leq \frac{k}{2}T_2^\rho, \gamma \geq \tau_k).
\end{aligned}$$

Note that all the probabilities of the following form  $\mathbb{P}_{\mathbf{x}}(\mathbb{W}_2(\mu_\gamma; \delta_a) \geq (1+\varepsilon)\rho, \mathcal{T}_{\text{out}}(k) < kT_1^\rho, \mathcal{T}_{\text{in}}(k) > \frac{k}{2}T_2^\rho, \gamma \in [\theta_k; \tau_{k+1}])$  are equal to 0 by (3.28) and . For the latter two sums we can use (3.33) and (3.35). Finally, we can conclude that the twice summation of infinitesimals above gives us

$$A \leq \frac{o_\sigma}{(1-o_\sigma)^3} + (1+m^*)^3 \exp \left\{ -\frac{8(H+1)}{\sigma^2} \right\}.$$

We remind that, by definition,  $m^* = \lceil \frac{2}{T_2^\rho} \rceil \exp\left\{\frac{2(H+1)}{\sigma^2}\right\}$ . It means that  $A$  is also bounded by some function that tends to 0. Combining the bounds above for  $A$  and  $B$  as well as Lemma 3.11 (decrease  $\sigma$  if necessary such that  $\exp\left\{\frac{2(H+1)}{\sigma^2}\right\} > T_{\text{st}}^\rho$ ), we finally get

$$\begin{aligned} \mathbb{P}_{x_0}(\gamma \leq \tau_G^\sigma) &\leq \mathbb{P}_{x_0}(\tau_G^\sigma \leq T_{\text{st}}^\rho) + \sup_{\mathbf{x} \in \mathcal{C}_4} \mathbb{P}_{\mathbf{x}}\left(\gamma \leq \tau_G^\sigma \wedge \exp\left\{\frac{2(H+1)}{\sigma^2}\right\}\right) \\ &\xrightarrow{\sigma \rightarrow 0} 0. \end{aligned}$$

□

### Proof of Lemma 3.16: Exit before nearing $a$

First, we show that it is possible to establish a lower bound for

$$\inf_{z \in \partial G} Q^\rho(x, z) := \frac{1}{2} \inf_{z \in \partial G} \inf_{t > 0} \inf_{\phi} I_t^{\mathbf{x}}(\phi),$$

where  $\mathbf{x} \in \mathcal{C}_1$  and  $\mathcal{C}_1 = \{\mathbf{x} \in \mathfrak{X} : T_{\text{st}}^\rho \leq T_{\mathbf{x}} \leq \infty; M_{\mathbf{x}} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a); P_{\mathbf{x}} \in G\}$ , and infimum is taken over functions  $\phi \in C([0, t]; \mathbb{R}^d)$  such that  $\phi_t = z$  and  $\mu_s^\phi := \frac{1}{s} \int_0^s \delta_{\phi_z} dz \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)$  for any  $s \leq t$ . Moreover, we show that this lower bound approaches  $H$  with  $\rho \rightarrow 0$ . Indeed, for any  $f \in C([0, T]; \mathbb{R}^d)$ , such that  $\mu_t^f \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)$  for any  $t \leq T$ , and  $f_0 \in G$ ,  $f_T \in \partial G$ :

$$\begin{aligned} &\frac{1}{4} \int_0^T |\dot{f}_t + \nabla V(f_t) + \nabla F * \delta_a(f_t)|^2 dt \\ &= \frac{1}{4} \int_0^T |\dot{f}_t + \nabla V(f_t) + \nabla F * \mu_t^f(f_t)|^2 dt \\ &\quad + \frac{1}{4} \int_0^T |\nabla F * \mu_t^f(f_t) - \nabla F * \delta_a(f_t)|^2 dt \\ &\quad + \frac{1}{2} \int_0^T \langle \dot{f}_t + \nabla V(f_t) + \nabla F * \delta_a(f_t); \nabla F * \mu_t^f(f_t) - \nabla F * \delta_a(f_t) \rangle dt \\ &\leq \frac{1}{4} \int_0^T |\dot{f}_t + \nabla V(f_t) + \nabla F * \mu_t^f(f_t)|^2 dt \\ &\quad + \frac{\text{Lip}_{\nabla F}}{4} T(1 + 2\varepsilon)^2 \rho^2 \\ &\quad + \frac{1}{2} \sqrt{\int_0^T |\dot{f}_t + \nabla V(f_t) + \nabla F * \mu_t^f(f_t)|^2 dt} \cdot \text{Lip}_{\nabla F}(1 + \varepsilon)\rho. \end{aligned}$$

We can first take the infimum of both sides of the inequality above over all  $f$  such that  $\mu_t^f \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)$  and  $f_0 \in G$ ,  $f_T \in \partial G$ , then over all  $T > 0$ . Note that  $H$  will be less than the infimum in the left-hand side. Thus, we conclude that

$$\begin{aligned} H &\leq \inf_{z \in \partial G} Q^\rho(a, z) + \text{Lip}_{\nabla F}(1 + \varepsilon)\rho \sqrt{\inf_{z \in \partial G} Q^\rho(a, z)} \\ &\leq Q^\rho(a, y) + \inf_{z \in \partial G} Q^\rho(y, z) + \text{Lip}_{\nabla F}(1 + \varepsilon)\rho \sqrt{\inf_{z \in \partial G} Q^\rho(a, z)}, \end{aligned}$$

or, equivalently,

$$\inf_{z \in \partial G} Q^\rho(y, z) \geq H - C(\rho), \quad (3.36)$$

where  $C(\rho) > 0$  is some function of  $\rho$ , that tends to 0, whenever  $\rho \rightarrow 0$ .

By Lemma 3.14, there exists time  $T > 0$  big enough such that,

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup \mathbb{P}_{\mathbf{x}}(\tau_0 > T, \gamma > \tau_0) < -H,$$

where supremum is taken over  $\mathbf{P}_{\mathbf{x}} \in S_{(1+\varepsilon)\rho}(a)$ . Consider the following set:

$$\begin{aligned} \Phi := \{ \phi \in C([0, T]; \mathbb{R}^d) : \exists t \in [0, T] : \phi_t \in \partial G, \\ \text{and } \mu_s^\phi \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a) \forall s \in [0, T] \}. \end{aligned}$$

By the definition of  $Q^\rho(x, z)$  and proved above facts,

$$\inf_{\substack{T_{\text{st}}^\rho \leq T_{\mathbf{x}} \leq \infty \\ \mathbf{P}_{\mathbf{x}} \in S_{(1+\varepsilon)\rho}(a) \\ \mathbf{M}_{\mathbf{x}} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)}} \inf_{\phi \in \Phi} I_T^{\mathbf{x}}(\phi) \geq \inf_{\substack{y \in S_{(1+\varepsilon)\rho}(a) \\ z \in \partial G}} Q^\rho(y, z) \geq H - C(\rho).$$

By Lemma 3.6,

$$\begin{aligned} \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup_{\substack{T_{\text{st}}^\rho \leq T_{\mathbf{x}} \leq \infty \\ \mathbf{P}_{\mathbf{x}} \in S_{(1+\varepsilon)\rho}(a) \\ \mathbf{M}_{\mathbf{x}} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)}} \sup_{\phi \in \Phi} \mathbb{P}_{\mathbf{x}}(X^\sigma \in \Phi) \\ \leq - \inf_{\substack{T_{\text{st}}^\rho \leq T_{\mathbf{x}} \leq \infty \\ \mathbf{P}_{\mathbf{x}} \in S_{(1+\varepsilon)\rho}(a) \\ \mathbf{M}_{\mathbf{x}} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)}} \inf_{\phi \in \Phi} I_T^{\mathbf{x}}(\phi) \leq -H + C(\rho). \end{aligned}$$

And, since  $\mathbb{P}_{\mathbf{x}}(\tau_0 = \tau_G^\sigma, \tau_0 < \gamma) \leq \mathbb{P}_{\mathbf{x}}(\tau_0 < T, \tau_0 < \gamma) + \mathbb{P}_{\mathbf{x}}(X^\sigma \in \Phi)$ , we get:

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup_{\substack{T_{\text{st}}^\rho \leq T_{\mathbf{x}} \leq \infty \\ \mathbf{P}\mathbf{x} \in \mathcal{S}_{(1+\varepsilon)\rho}(a) \\ \mathbf{M}\mathbf{x} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)}} \sup_{\phi \in \Phi} \mathbb{P}_{\mathbf{x}}(\tau_0 = \tau_G^\sigma, \tau_0 < \gamma) \leq -H + C(\rho),$$

which proves the lemma by taking  $\rho \rightarrow 0$ .

### Proof of Lemma 3.17: Control of dynamics for small time intervals

Let us investigate the dynamics of  $|X^\sigma - \mathbf{P}\mathbf{x}|$ . We plug in the expression describing  $X_t^\sigma$  from (3.5), use Lipschitzness of  $\nabla V$  and  $\nabla F$  (Assumption A-1.2) as well as  $\nabla V(a) = 0$ ,  $\nabla F(0) = 0$  (Assumption A-2.3), and express for any  $t$ :

$$\begin{aligned} |X_t^\sigma - \mathbf{P}\mathbf{x}| &\leq \sigma |W_t| + \text{Lip}_{\nabla V} \int_0^t |X_s^\sigma - \mathbf{P}\mathbf{x}| ds + \text{Lip}_{\nabla V} |\mathbf{P}\mathbf{x} - a|t \\ &\quad + \int_0^t \frac{T_{\mathbf{x}}}{T_{\mathbf{x}} + s} \text{Lip}_{\nabla F} \left( |X_s^\sigma - \mathbf{P}\mathbf{x}| + \int |z - \mathbf{P}\mathbf{x}| \mathbf{M}\mathbf{x}(dz) \right) ds \\ &\quad + \int_0^t \frac{1}{T_{\mathbf{x}} + s} \text{Lip}_{\nabla F} \int_0^s \left( |X_s^\sigma - \mathbf{P}\mathbf{x}| + |X_u^\sigma - \mathbf{P}\mathbf{x}| \right) du ds \end{aligned}$$

Use the Jensen's inequality, integrate the second part of the last integral over  $s$ , and get:

$$\begin{aligned} |X_t^\sigma - \mathbf{P}\mathbf{x}| &\leq \sigma |W_t| + \text{Lip}_{\nabla V} \int_0^t |X_s - \mathbf{P}\mathbf{x}| ds + \text{Lip}_{\nabla V} |\mathbf{P}\mathbf{x} - a|t \\ &\quad + \text{Lip}_{\nabla F} \int_0^t |X_s^\sigma - \mathbf{P}\mathbf{x}| ds + \text{Lip}_{\nabla F} t \left( \mathbb{W}_2(\mathbf{M}\mathbf{x}; \delta_a) + |\mathbf{P}\mathbf{x} - a| \right) \\ &\quad + \text{Lip}_{\nabla F} \left( \int_0^t |X_s^\sigma - \mathbf{P}\mathbf{x}| ds + \int_0^t \log \left( \frac{T_{\mathbf{x}} + t}{T_{\mathbf{x}} + s} \right) |X_s^\sigma - \mathbf{P}\mathbf{x}| ds \right) \end{aligned}$$

Finally, introducing  $R := \sup_{z \in \partial G} |z - a|$  and using the fact that  $\mathbf{x}$  is assumed, in the lemma, to belong to  $\mathcal{C}_4$ , we get the bound:

$$\begin{aligned} |X_t^\sigma - \mathbf{P}\mathbf{x}| &\leq \sigma |W_t| + \left( (\text{Lip}_{\nabla V} + \text{Lip}_{\nabla F})R + \text{Lip}_{\nabla F}(1 + \varepsilon)\rho \right) t \\ &\quad + \left( \text{Lip}_{\nabla V} + 2\text{Lip}_{\nabla F} + \log \left( \frac{T_{\mathbf{x}} + t}{T_{\mathbf{x}}} \right) \right) \int_0^t |X_s^\sigma - \mathbf{P}\mathbf{x}| ds. \end{aligned}$$

We apply Grönwall's inequality and get:

$$\begin{aligned} |X_t^\sigma - \mathbf{P}\mathbf{x}| &\leq \left( \sigma |W_t| + \left( (\text{Lip}_{\nabla V} + \text{Lip}_{\nabla F})R + \text{Lip}_{\nabla F}\rho \right) t \right) \\ &\times \exp \left\{ (\text{Lip}_{\nabla V} + 2\text{Lip}_{\nabla F})t + \mathbf{T}\mathbf{x} \frac{\mathbf{T}\mathbf{x} + t}{\mathbf{T}\mathbf{x}} \log \left( \frac{\mathbf{T}\mathbf{x} + t}{\mathbf{T}\mathbf{x}} \right) - t \right\}. \end{aligned} \quad (3.37)$$

Thus, it follows that, for the absolute value of the Brownian motion itself,

$$\sigma |W_t| \geq -C_1 t + C_2(t) |X_t^\sigma - \mathbf{P}\mathbf{x}|,$$

where, for simplicity of further derivations, we introduced a positive constant  $C_1 := (\text{Lip}_{\nabla V} + \text{Lip}_{\nabla F})R + \text{Lip}_{\nabla F}\rho$  and a function  $C_2(t)$  that is one over the exponent that appears in the equation (3.37). For our purposes, it is not the form of  $C_2$  itself that is important, but rather its following properties:  $C_2(t) \geq 0$  for any  $t \geq 0$  and  $C_2(t) \xrightarrow[t \rightarrow 0]{} 1$ .

The infimum of  $C_2$  over any time interval  $[0, T]$ , for  $T < 1$ , is either equal to  $C_2(T)$ , if  $\text{Lip}_{\nabla V} + 2\text{Lip}_{\nabla F} \geq 1$ , or can be bounded by  $C_2(T) \exp\{\text{Lip}_{\nabla V} + 2\text{Lip}_{\nabla F} - 1\}$  otherwise. This observation can be expressed in the following form:

$$\inf_{t \in [0, T]} C_2(t) \geq \min\{C_2(T); C_2(T) \exp\{\text{Lip}_{\nabla V} + 2\text{Lip}_{\nabla F} - 1\}\}.$$

Moreover, it is easy to check that  $\lim_{T \rightarrow 0} C_2^2(T)/T = \infty$ .

Taking into account these remarks, we can now use the Schilder theorem [DZ10, Lemma 5.2.1] that provides the LDP for the path of the Brownian motion. Hence, for some fixed  $\epsilon$  and  $0 < T < 1$ :

$$\begin{aligned} \mathbb{P}_{\mathbf{x}} \left( \sup_{t \in [0, T]} |X_t^\sigma - \mathbf{P}\mathbf{x}| \geq \epsilon \right) &\leq \mathbb{P}_{\mathbf{x}} \left( \sigma \sup_{t \in [0, T]} |W_t| \geq -C_1 t + C_2(t)\epsilon \right) \\ &\leq 4d \exp \left\{ - \frac{(-C_1 t + \inf_{t \in [0, T]} C_2(t)\epsilon)^2}{4dT} \cdot \frac{2}{\sigma^2} \right\}, \end{aligned}$$

where  $d$  is the dimension of the space.

Note that for any  $\epsilon > 0$  the following limit holds:

$$\frac{(-C_1 T + \inf_{t \in [0, T]} C_2(t)\epsilon)^2}{T} \xrightarrow[T \rightarrow 0]{} \infty.$$



Thus, for any given in advance  $c$  and  $\epsilon$ , we can choose  $T = T(\epsilon, c)$  to be small enough, such that the rate  $-( -C_1 T(\epsilon, c) + \inf_{t \in [0, T(\epsilon, c)]} C_2(t) \epsilon )^2 / (4dT(\epsilon, c))$  will be less or equal to  $-c$ , which completes the proof.

### Proof of Lemma 3.18: Uniform lower bound for probability of exit from $G$

In order to prove the lemma, we need to find for some  $\delta > 0$  small enough, for some  $T_0 > 0$  and for any  $\mathbf{x} \in \mathcal{C}_2$  a function  $\psi^{\mathbf{x}} \in C([0, T_0]; \mathbb{R}^d)$  such that  $\psi^{\mathbf{x}}(0) = \mathbf{P}_{\mathbf{x}}$ ,  $\inf_{z \in G} |\psi^{\mathbf{x}}(s) - z| > \delta$  for some  $s \leq T_0$  and  $I_{T_0}^{\mathbf{x}}(\psi^{\mathbf{x}}) < H + \eta$ . Moreover, this function should posses an empirical measure that is close to  $\delta_a$  at any point of time, i.e.  $\mathbb{W}_2\left(\frac{1}{t} \int_0^t \delta_{\psi_s^{\mathbf{x}}} ds; \delta_a\right) \leq \rho$ . Given such a function, by a simple inclusion of events, we can get the following bound for probability of leaving domain  $G$  before time  $T_0$ :

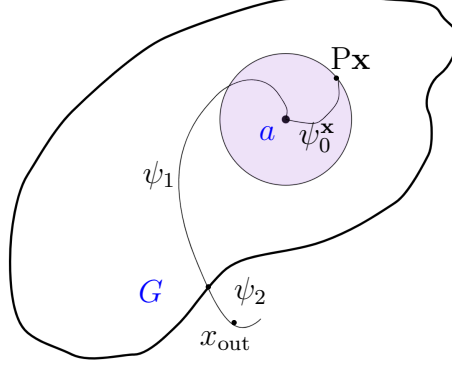
$$\mathbb{P}_{\mathbf{x}}(\tau_G^\sigma \leq T_0, \gamma > \tau_G^\sigma) \geq \mathbb{P}_{\mathbf{x}}(X^{\mathbf{x}, \sigma} \in \Psi),$$

where  $\Psi := \bigcup_{\mathbf{x} \in \mathcal{C}_2} \Psi^{\mathbf{x}}$  and  $\Psi^{\mathbf{x}} := \{\phi \in C([0, T_0]; \mathbb{R}^d) : \|\phi - \psi^{\mathbf{x}}\|_\infty < \delta\}$ . Of course, we should take  $\delta$  small enough such that any  $\phi \in \Psi$  has an occupation measure that satisfies  $\mathbb{W}_2\left(\frac{1}{t} \int_0^t \delta_{\phi_s} ds; \delta_a\right)$ . Note that  $\Psi$  is an open subset of  $C([0, T_0]; \mathbb{R}^d)$ , as a union of open sets. Therefore, we can use Theorem 3.3 and get

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}_{\mathbf{x}}(\tau_G^\sigma \leq T_0, \gamma > \tau_G^\sigma) \geq - \inf_{\phi \in \Psi} I_{T_0}^{\mathbf{x}}(\phi) \geq -I_{T_0}^{\mathbf{x}}(\psi^{\mathbf{x}}) > -(H + \eta).$$

The same lower bound obviously holds for infimum  $\inf_{\mathbf{x} \in \mathcal{C}_2} \mathbb{P}_{\mathbf{x}}(\tau_G^\sigma \leq T_0, \gamma > \tau_G^\sigma)$ , which is what the lemma claims. Thus, we only need to find a  $\psi^{\mathbf{x}}$  with the properties given above.

*Construction of  $\psi^{\mathbf{x}}$ .* The function  $\psi^{\mathbf{x}}$  will be represented as a consecutive gluing of four functions (see Figure 3.5):  $\psi_0^{\mathbf{x}} \in C([0, 1]; \mathbb{R}^d)$ ,  $\psi_a \in C([0, T_a]; \mathbb{R}^d)$ ,  $\psi_1 \in C([0, T_1]; \mathbb{R}^d)$ , and  $\psi_2 \in C([0, 1]; \mathbb{R}^d)$ , for some  $T_a, T_1$  that will be defined below and will determine  $T_0$  as  $T_0 := T_a + T_1 + 2$ .

Figure 3.5: Construction of  $\psi^{\mathbf{x}}$ 

1. The first function  $\psi_0^{\mathbf{x}}$  will bring us from point  $\mathbf{P}\mathbf{x}$  to point  $a$  in at most time 1 with small rate function. Define

$$\psi_0^{\mathbf{x}}(s) = \begin{cases} \mathbf{P}\mathbf{x} + \frac{a - \mathbf{P}\mathbf{x}}{|a - \mathbf{P}\mathbf{x}|}s, & \text{for } 0 \leq s \leq |a - \mathbf{P}\mathbf{x}|, \\ a, & \text{for } |a - \mathbf{P}\mathbf{x}| \leq s \leq 1. \end{cases}$$

We can establish the following bound for its rate function. Let  $\mathbf{x}_0 = (\infty, \delta_a, \mathbf{P}\mathbf{x})$  and  $\mu_s^{\psi_0^{\mathbf{x}}} := \frac{\mathbf{T}\mathbf{x}}{\mathbf{T}\mathbf{x} + s} \mathbf{M}\mathbf{x} + \frac{1}{\mathbf{T}\mathbf{x} + s} \int_0^s \delta_{\psi_0^{\mathbf{x}}(s)} ds$ . We want to separate the interaction. Then:

$$I_1^{\mathbf{x}}(\psi_0^{\mathbf{x}}) \leq 2I_1^{\mathbf{x}_0}(\psi_0^{\mathbf{x}}) + \frac{1}{2} \int_0^1 \left| \nabla F * \mu_s^{\psi_0^{\mathbf{x}}}(\psi_0^{\mathbf{x}}(s)) - \nabla F(\psi_0^{\mathbf{x}}(s) - a) \right|^2 ds.$$

Note that, since  $\psi_0^{\mathbf{x}}$  never leaves  $B_\rho(a)$  and since  $\mathbf{M}\mathbf{x} \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)$ , then  $\mathbb{W}_2(\mu_s^{\psi_0^{\mathbf{x}}}; \delta_a) \leq (1 + \varepsilon)\rho$  for any  $0 \leq s \leq 1$ . We use Lipschitz continuity of  $\nabla F$  (Assumption A-1.2) and get

$$\begin{aligned} I_1^{\mathbf{x}}(\psi_0^{\mathbf{x}}) &\leq 2I_1^{\mathbf{x}_0}(\psi_0^{\mathbf{x}}) + \frac{\text{Lip}_{\nabla F}^2}{2} \max_{0 \leq s \leq 1} \mathbb{W}_2^2(\mu_s^{\psi_0^{\mathbf{x}}}; \mu_1^{\psi_0^{\mathbf{x}}}) \\ &\leq 2I_1^{\mathbf{x}_0}(\psi_0^{\mathbf{x}}) + \frac{\text{Lip}_{\nabla F}^2}{2} (1 + \varepsilon)^2 \rho^2. \end{aligned}$$

As for  $I_1^{\mathbf{x}_0}(\psi_0^{\mathbf{x}})$ ,

$$\begin{aligned} 4I_1^{\mathbf{x}_0} &= \int_0^{|a-\mathbf{P}\mathbf{x}|} \left| \frac{a-\mathbf{P}\mathbf{x}}{|a-\mathbf{P}\mathbf{x}|} + \nabla V(\psi_0^{\mathbf{x}}(s)) \right|^2 ds \\ &\leq 2|a-\mathbf{P}\mathbf{x}| + 2\text{Lip}_{\nabla V} \int_0^{|a-\mathbf{P}\mathbf{x}|} |\psi_0^{\mathbf{x}}(s)|^2 ds \\ &\leq 2\rho + 4\text{Lip}_{\nabla V}(|a|^2 + \rho^2)\rho. \end{aligned} \tag{3.38}$$

Therefore, we can choose  $\rho > 0$  small enough such that

$$I_1^{\mathbf{x}}(\psi_0^{\mathbf{x}}) \leq \frac{\eta}{5}.$$

2. The second function  $\psi_a$  is defined to be constant and equal to  $a$  for some time  $T_a$ , which will be increased later if necessary:

$$\psi_a \equiv a.$$

We need this segment in order to put enough ‘‘mass’’ in  $\delta_a$  and balance the later path. Of course, its occupation measure never exceeds  $(1 + \varepsilon)\rho$ . Moreover, if we denote  $\mathbf{a} = (\infty, \delta_a, a)$ ,  $\mathbf{x}_a := (T_a + 1, \mu_1^{\psi_0^{\mathbf{x}}}, a)$ , and  $\mu_s^{\psi_a} := \frac{T_a+1}{T_a+1+s}\mu_1^{\psi_0^{\mathbf{x}}} + \frac{1}{T_a+1+s} \int_0^s \delta_{\psi_a(s)} ds$ , we can, as before, establish the following bound for the rate function:

$$\begin{aligned} I_{T_a}^{\mathbf{x}_a}(\psi_a) &\leq 2I_{T_a}^{\mathbf{a}}(\psi_a) + \frac{\text{Lip}_{\nabla F}^2}{2} \max_{0 \leq s \leq T_a} \mathbb{W}_2^2(\mu_s^{\psi_a}; \delta_a) \\ &\leq \frac{\text{Lip}_{\nabla F}^2}{2} (1 + \varepsilon)^2 \rho^2 \leq \frac{\eta}{5}, \end{aligned}$$

for  $\rho$  small enough. Note that this bound is independent of the choice of  $T_a$ .

3. In order to construct the third function  $\psi_1$ , we first remind that  $H = \inf_{z \in \partial G} \{V(z) + F(z - a) - V(a)\}$  is the height of the effective potential  $V + F(\cdot - a)$  within domain  $G$ . By the classical result,  $H$  can be also expressed as an infimum

$$\inf_{t>0} \inf_{\phi \in \Phi_t^a} I_t^{\mathbf{a}}(\phi) = H,$$

where  $\mathbf{a} := (\infty, \delta_a, a)$  and  $\Phi_t^a := \{\phi \in C([0, t]; \mathbb{R}^d) : \phi(0) = a, \phi(t) \in \partial G\}$ . Hence, there exists  $T_1 > 0$  and  $\psi_1 \in \Phi_{T_1}^a$  (which among other things implies  $\psi_1(T) \in \partial G$ ) such that  $I_{T_1}^{\mathbf{a}}(\psi_1) \leq H + \eta/10$ . Define

$$\begin{aligned} \mu_s^{\psi_1} &:= \frac{\mathbf{T}\mathbf{x} + 1}{\mathbf{T}\mathbf{x} + 1 + T_a + s} \mu_1^{\psi_0^{\mathbf{x}}} + \frac{T_a}{\mathbf{T}\mathbf{x} + 1 + T_a + s} \delta_a \\ &\quad + \frac{s}{\mathbf{T}\mathbf{x} + 1 + T_a + s} \cdot \frac{1}{s} \int_0^s \delta_{\psi_1(s)} \, ds. \end{aligned}$$

So, if we define  $R = \sup_{z \in G} |z - a|$ , we can express:

$$\mathbb{W}_2(\mu_s^{\psi_1}; \delta_a) \leq (1 + \varepsilon)\rho + \frac{-(1 + \varepsilon)\rho T_a + T_1 R}{\mathbf{T}\mathbf{x} + 1 + T_a + s}.$$

And we can increase without loss of generality  $T_a$  to be big enough such that  $-(1 + \varepsilon)\rho T_a + T_1 R < 0$ , which guarantees that

$$\mathbb{W}_2(\mu_s^{\psi_1}; \delta_a) \leq (1 + \varepsilon)\rho$$

for any  $s \leq T_1$ .

We can also achieve the following upper bound for the rate function of  $\psi_1$  with initial condition  $\mathbf{x}_1 = (\mathbf{T}\mathbf{x} + 1 + T_a, \mu_1^{\psi_a}, a)$ . For any  $c \in (0, 1)$ ,

$$\begin{aligned} I_{T_1}^{\mathbf{x}_1}(\psi_1) &\leq (1 + c)I_{T_1}^{\mathbf{a}}(\psi_1) \\ &\quad + \frac{1 + c}{4c} \int_0^{T_1} |\nabla F * \mu_s^{\psi_1}(\psi_1(s)) - \nabla F(\psi_1(s) - a)|^2 \, ds \\ &\leq (1 + c)I_{T_1}^{\mathbf{a}}(\psi_1) + \frac{\text{Lip}_{\nabla F} T_1 (1 + c)}{4c} \mathbb{W}_2^2(\mu_s^{\psi_1}; \delta_a) \\ &\leq (1 + c)I_{T_1}^{\mathbf{a}}(\psi_1) + \frac{\text{Lip}_{\nabla F} T_1 (1 + c)}{2c} (1 + \varepsilon)^2 \rho^2 \xrightarrow{\rho \rightarrow 0} (1 + c)I_{T_1}^{\mathbf{a}}(\psi_1). \end{aligned}$$

Since the limit above holds for any  $c \in (0, 1)$ , we can, without loss of generality, make  $\rho$  small enough such that  $I_{T_1}^{\mathbf{x}_1}(\psi_1) \leq I_{T_1}^{\mathbf{a}}(\psi_1) + \eta/10$ . That leads to

$$I_{T_1}^{\mathbf{x}_1}(\psi_1) \leq H + \frac{\eta}{5},$$

for any  $\mathbf{x} \in \mathcal{C}_2$  (we remind that  $\mathbf{x}_1$  depends on  $\mathbf{x}$ ).

4. For the last part, we choose a point  $x_{\text{out}} \notin G$  such that the line segment starting from point  $\psi_1(T_1)$  and finishing at  $x_{\text{out}}$  is included in  $\mathbb{R}^d \setminus G$  (this point exists since  $\partial G$  is smooth by Assumption A-2.1). Define

$$\psi_2(s) := \psi_1(T_1) + \frac{\psi_1(T_1) - x_{\text{out}}}{|\psi_1(T_1) - x_{\text{out}}|} s,$$

for  $0 \leq s \leq \delta := |\psi_1(T_1) - x_{\text{out}}|$ . After that, let this function follow a path that does not contribute to the rate function for the rest of the time  $\delta < s \leq 1$ . Precisely, if we define the cumulative empirical measure of  $\psi_2$  as

$$\mu_s^{\psi_2} := \frac{\mathbf{T}\mathbf{x} + 1 + T_a + T_1}{\mathbf{T}\mathbf{x} + 1 + T_a + T_1 + s} \mu_{T_1}^{\psi_1} + \frac{1}{\mathbf{T}\mathbf{x} + 1 + T_a + T_1 + s} \int_0^s \delta_{\psi_2(s)} ds,$$

we can also construct a deterministic process  $X_s^{\mathbf{x}_3, 0}$  given by (3.6) with initial conditions  $\mathbf{x}_3 := (T_3, \mu_3, x_3)$ , where  $T_3 = \mathbf{P}\mathbf{x} + 1 + T_a + T_1 + \delta$ ,  $\mu_3 = \mu_\delta^{\psi_2}$  and  $x_3 = x_{\text{out}}$ . Lemma 3.12 provides the upper bound for the occupation measure of  $X^{\mathbf{x}_3, 0}$ :  $\mathbb{W}_2(\mu_s^{\mathbf{x}_3, 0}; \delta_a) \leq (1 + \varepsilon)\rho$ , for some  $\varepsilon > 0$  that is defined in the lemma. Hence, similarly to previous computations, we can control the cumulative empirical measure  $\mu_s^{\psi_2}$ . We can increase  $T_a$  if necessary such that

$$\mathbb{W}_2(\mu_s^{\psi_2}; \delta_a) \leq (1 + \varepsilon)\rho$$

for any  $s \leq T_1$

Define  $\mathbf{x}_2 := (\mathbf{P}\mathbf{x} + 1 + T_a + T_1, \mu_{T_1}^{\psi_1}, \psi_1(T_1))$ . We remind that  $\psi_2$  was constructed in a way that the second part of it does not contribute to the rate function  $I_1^{\mathbf{x}_2}(\psi_2)$ . For the first part the computations are similar to those of (3.38). As a result, without loss of generality, we can choose  $\delta > 0$  to be small enough ( $x_{\text{out}}$  closer to  $\partial G$ ) and get:

$$I_1^{\mathbf{x}_2}(\psi_2) \leq \frac{\eta}{5}.$$

5. Finally, by letting

$$\psi^{\mathbf{x}} = \begin{cases} \psi_0^{\mathbf{x}}(s), & \text{for } 0 \leq s \leq 1, \\ \psi_a(s - 1), & \text{for } 1 < s \leq 1 + T_a, \\ \psi_1(s - 1 - T_a), & \text{for } 1 + T_a < s \leq 1 + T_a + T_1, \\ \psi_2(s - 1 - T_a - T_1), & \text{for } 1 + T_a + T_1 < s \leq T_a + T_1 + 2; \end{cases} \quad (3.39)$$

we can observe that  $I_{T_0}^{\mathbf{x}}(\psi^{\mathbf{x}}) = I_1^{\mathbf{x}}(\psi_0^{\mathbf{x}}) + I_{T_a}^{\mathbf{x}_a}(\psi_a) + I_{T_1}^{\mathbf{x}_1}(\psi_1) + I_1^{\mathbf{x}_2}(\psi_2) \leq H + \frac{4\eta}{5} < H + \eta$  and  $\sup_{s \in [0, T_0]} \inf_{z \in G} |\psi^{\mathbf{x}}(s) - z| > \delta$ . Moreover, by construction of all the function, we have ensured that  $\frac{1}{t} \int_0^t \delta_{\psi^{\mathbf{x}}(s)} ds \in \mathbb{B}_{(1+\varepsilon)\rho}(\delta_a)$  for any  $t \leq T_0$ , which is what is needed.

### 3.3 Generalization

In this section we present the possible generalisation of Assumptions A-1 and A-2 as well as the exit-time result for this more general process. In this chapter we did not talk in details about the problem of existence and uniqueness of the self-interacting diffusion. While in the Lipschitz case (Assumption A-1) these questions are standard, locally Lipschitzness of the potentials add some complication. In [AdMKT23], the authors provide a proof for existence and uniqueness of self-interacting diffusion under the following assumptions that we take as a baseline:

#### Assumptions A-1'.

1. (regularity) Potentials  $V$  and  $F$  belong to the space  $C^2(\mathbb{R}^d; \mathbb{R})$ .
2. (growth) There exist a constant  $C_{\text{gr}}$  and an order  $n > 0$  such that for any  $x, y \in \mathbb{R}^d$ ,  $|\nabla V(x) - \nabla V(y)| \leq C_{\text{gr}}|x - y|(1 + |x|^{2n} + |y|^{2n})$  as well as  $|\nabla F(x) - \nabla F(y)| \leq C_{\text{gr}}|x - y|(1 + |x|^{2n} + |y|^{2n})$ .
3. (confinement at infinity)  $\lim_{|x| \rightarrow \infty} V(x) = +\infty$ ,  $\lim_{|x| \rightarrow \infty} \frac{|\nabla V(x)|^2}{V(x)} = +\infty$  and there exists  $\alpha > 0$  such that  $\Delta V(x) \leq \alpha V(x)$  and  $\Delta F(x) \leq \alpha F(x)$ .

It is worth noting that Assumption A-1'.2 is a combination of locally Lipschitz and polynomial growth conditions. Assumption A-1'.3 has been introduced in Chapter 2 (this chapter is based on [AdMKT23]) to control the Lyapunov functional of the diffusion process and thus to ensure that there is no explosion within a finite time. It is important to mention that, if required, this assumption can be replaced by another one that ensures the process's global existence and uniqueness without any impact on the exit-time result.

We can also relax Assumptions A-2. Indeed, the domain  $G$  does not necessarily have to be bounded, but it is necessary to add another assumption on the level sets instead. Consider:

**Assumptions A-2'.**

1. (domain  $G$ )  $G \subset \mathbb{R}^d$  is an open connected set such that  $\partial G = \partial \overline{G}$ . The boundary  $\partial G$  is a smooth  $(d - 1)$ -dimensional hypersurface.
2. (bounded sublevel set) Let  $L_H^- := \{x \in V(x) + F(x - a) \leq H\}$  be the sublevel set of height  $H := \inf_{x \in \partial G} \{V(x) + F(x - a) - V(a)\}$ . Then  $L_H^- \cap G$  is a bounded connected set.
3. (point of attraction  $a$ ) There exists  $a \in G$  such that  $\{X_t^{x_0, 0}\}_{t \geq 0} \subset G$  and  $X_t^{x_0, 0} \xrightarrow[t \rightarrow \infty]{} a$ .
4. (stability of  $\overline{G}$  under the effective potential) Let  $\phi$  be defined as  $\phi_t^x = x - \int_0^t \nabla V(\phi_s) ds - \int_0^t \nabla F(\phi_s - a) ds$ . For any  $x \in \overline{G}$ ,  $\{\phi_t^x\}_{t > 0} \subset G$  and  $\phi_t^x \xrightarrow[t \rightarrow \infty]{} a$ . Moreover,  $\nabla V(a) = \nabla F(0) = 0$ .
5. (strong attraction around  $a$ ) There exist  $\Delta_\mu, \Delta_x > 0$  small enough and a constant  $K > 0$ , such that for any  $\mu \in \mathbb{B}_{\Delta_\mu}(\delta_a)$  and for any  $x \in B_{\Delta_x}(a)$ ,  $\langle \nabla V(x) + \nabla F * \mu(x); x - a \rangle \geq K|x - a|^2$ .

Under these assumptions we can obtain Kramers' type law and the exit-location result. Consider

**Corollary 3.20.** *Let Assumptions A-1' and A-2' be fulfilled. Let the process  $X^\sigma$  be the unique strong solution of the system (3.1) with initial condition  $x_0 \in G$ . Let  $\tau_G^\sigma := \inf\{t : X_t^\sigma \notin G\}$  denote the first time when  $X^\sigma$  exits the domain  $G$ . Let  $H := \inf_{x \in \partial G} \{V(x) + F(x - a) - V(a)\}$  be the height of the effective potential. Then, the following two results hold:*

1. *Kramers' law: for any  $\delta > 0$*

$$\lim_{\sigma \rightarrow 0} \mathbb{P}_{x_0} \left( e^{\frac{2(H-\delta)}{\sigma^2}} < \tau_G^\sigma < e^{\frac{2(H+\delta)}{\sigma^2}} \right) = 1; \quad (3.40)$$

2. *Exit-location: for any closed set  $N \subset \partial G$  such that  $\inf_{z \in N} \{V(z) + F(z - a) - V(a)\} > H$  the following limit holds:*

$$\lim_{\sigma \rightarrow 0} \mathbb{P}_{x_0}(X_{\tau_G^\sigma}^\sigma \in N) = 0. \quad (3.41)$$

We do not prove this corollary rigorously but rather give an idea of the proof. The main insight comes from the proof of the exit-location result—Section 3.2. Namely, after the stabilization around  $\delta_a$ , the exit from the domain  $G$  happens around the level set that touches its boundary  $L_H$  (see page 66). Despite  $G$  is not bounded, we assumed  $L_H^- \cap G$  to be bounded. That means that we can introduce a diffusion

$$\begin{cases} dY_t^\sigma = -\nabla \bar{V}(Y_t^\sigma) dt - \frac{1}{t} \int_0^t \nabla \bar{F}(Y_t^\sigma - Y_s^\sigma) ds dt + \sigma dW_t, \\ Y_0^\sigma = x_0 \in \mathbb{R}^d \text{ a.s.} \end{cases}$$

driven by the same Brownian motion as (3.1), where  $\bar{V}$  and  $\bar{F}$  are modifications of  $V$  and  $F$  respectively defined in the following way. Let  $B_{\text{mod}} \subset \mathbb{R}^d$  be a big enough closed ball that contains the sets  $\{X_t^0\}_{t \geq 0}$  and  $L_H^-$ . First, we let  $\bar{V}(x) = V(x)$  and  $\bar{F}(x) = F(x)$  for any  $x \in B_{\text{mod}}$ . Then, we extend  $\bar{V}$  and  $\bar{F}$  on other points of  $\mathbb{R}^d$  such that in the end the potentials  $\bar{V}$  and  $\bar{F}$  are Lipschitz continuous and belong to  $C^2(\mathbb{R}^d; \mathbb{R})$ .

It means that  $\bar{V}$ ,  $\bar{F}$  and  $B_{\text{mod}} \cap G$  satisfy Assumptions A-1 and A-2 and we can establish the Kramers' law for  $\bar{\tau}^\sigma := \inf\{t \geq 0 : Y_t^\sigma \notin B_{\text{mod}} \cap G\}$  and exit-location result for  $Y^\sigma$ . Note also that for any  $t \geq 0$  we have:

$$X_{t \wedge \bar{\tau}^\sigma}^\sigma = Y_{t \wedge \bar{\tau}^\sigma}^\sigma \text{ a.s.}$$

The last observation that one needs to do in order to prove the corollary is the following. By the exit-location result, the probability that  $Y^\sigma$  exits domain  $B_{\text{mod}} \cap G$  around  $\partial G \cap L_H$  tends to one as  $\sigma$  tends to zero. It means that the probability that  $\tau_G^\sigma = \bar{\tau}^\sigma$  should also tend to one as  $\sigma$  tends to zero, since leaving domain  $B_{\text{mod}} \cap G$  on the boundary  $\partial G$  also means leaving  $G$  itself.



# Chapter 4

## Exit-problem for self-stabilizing diffusion: General case

Let us consider a measure-dependent stochastic process  $(X_t^\sigma, t \geq 0)$  (also called McKean-Vlasov diffusion [McK66, McK67]), solution to the following stochastic differential equation (SDE):

$$dX_t^\sigma = \sigma dW_t - \nabla V(X_t^\sigma) dt - \nabla F * \mu_t^\sigma(X_t^\sigma) dt, \quad X_0^\sigma = x_{\text{init}} \in \mathbb{R}^d. \quad (4.1)$$

Here  $\mu_t^\sigma = \mathcal{L}(X_t^\sigma)$ ,  $(W_t, t \geq 0)$  stands for a  $d$ -dimensional Brownian motion,  $V$  represents the environment which is assumed to be a multi-well function (also called confinement potential in this work) and  $F$  is the interaction potential corresponding to the form and strength of interaction of the process with its law. This specific form of the McKean-Vlasov diffusion is also known in the literature under the name of self-stabilizing diffusion or SSD (see [HIP08]).

The aim of this study is to describe how long the stochastic process stays in a domain  $\mathcal{D}$ , which is a neighborhood of a local minimum of  $V$ , before its first exit from this neighborhood. Therefore, the main object of interest in this chapter is the following stopping time:

$$\tau_{\mathcal{D}}^\sigma := \inf\{t \geq 0 : X_t^\sigma \notin \mathcal{D}\}. \quad (4.2)$$

The precise assumptions under consideration are given later on.

## Organization of the chapter

The current section is followed by presenting and discussing assumptions on potentials  $V$  and  $F$  and domain  $\mathcal{D}$ , exit-time from which is considered. An existing result on existence and uniqueness of the process under almost identical assumptions are provided on page 100 with a discussion on how its proof can be adapted to our case.

Main results of this chapter are formulated in Section 4.1. Namely, the large deviation principle for self-stabilizing diffusion with general initial condition and Kramers' type law for exit-time as well as exit-location result for both cases of bounded and unbounded domain  $\mathcal{D}$ . These theorems are followed by a section comparing them to previously known results for exit-time problem in the case of Self-stabilizing diffusion (page 104) and a section discussing open questions and possible extensions of our findings (page 105).

Section 4.2 contains intermediate lemmas that are necessary for the proof of the main theorems of the chapter. These lemmas are proved in Section 4.4.

Section 4.3 contains the proof of the main theorems provided in Section 4.1.

## Assumptions

Here, we give the assumptions on the potentials and on the domain.

**Assumptions A-1.** Let us consider the following hypotheses concerning the confinement potential:

(V – 1) The confinement potential is a regular function  $V \in C^2(\mathbb{R}^d; \mathbb{R})$ .

(V – 2)  $V$  is uniformly convex at infinity. Namely, there exists  $\theta_1 > 0$  and  $R > 0$  such that for all  $x \in \mathbb{R}^d$  satisfying  $|x| > R$  we have

$$\nabla^2 V(x) \succeq \theta_1 \text{Id},$$

where  $\text{Id}$  is the identity matrix.

(V – 3) There exist  $r \in \mathbb{Z}_+$  and a constant  $C > 0$  such that

$$|\nabla V(x)| \leq C(1 + |x|^{2r-1}), \quad \text{for all } x \in \mathbb{R}^d.$$

(V – 4) There exists  $a \in \mathbb{R}^d$  such that  $\nabla V(a) = 0$  and  $\nabla^2 V(a) \succeq \rho_1 \text{Id}$  for some  $\rho_1 > 0$ , where  $\text{Id}$  is the identity matrix.

(V – 5) The function  $\nabla V$  is locally Lipschitz. More precisely, for any  $x \in \mathbb{R}^d$  and  $y \in \mathbb{R}^d$ , we have:

$$|\nabla V(x) - \nabla V(y)| \leq C|x - y|(1 + |x|^{2r-1} + |y|^{2r-1}), \quad (4.3)$$

where  $r$  has been introduced in (V – 3).

Assumption (V – 1) is natural since we will use Itô calculus to obtain some of our results. Thus, we require that  $V$  is of class  $C^2$ . Assumption (V – 2) is taken to ensure that the confinement potential forces the diffusion to stay in a compact set and thus that the process does not explode. Assumptions (V – 3) and (V – 5) are required to comply with the theory developed in [BRTV98] for ensuring the existence of the self-stabilizing diffusion when the drift is superlinear. Assumption (V – 4) means that there is a local minimizer with a non-degenerate Hessian. We point out that  $\nabla V$  is not assumed to be globally Lipschitz.

Assumption A-1 covers a wide range of possible multi-well potentials. An analytical example of such a potential  $V$  that satisfies Assumption A-1 in dimension  $d = 1$  could be the classical double-well potential (see Fig. 4.1)

$$V(x) := \frac{x^4}{4} - \frac{x^2}{2}.$$

In dimension two, the following function

$$\begin{aligned} V(x_1, x_2) := & \frac{3}{2} (1 - x_1^2 - x_2^2)^2 + \frac{1}{3} (x_1^2 - 2)^2 + \frac{1}{6} ((x_1 + x_2)^2 - 1)^2 \\ & + \frac{1}{6} ((x_1 - x_2)^2 - 1)^2 \end{aligned}$$

could be an example of a double-well potential satisfying these assumptions. Fig. 4.2 shows its level sets.

We now give the assumptions on the interaction potential.

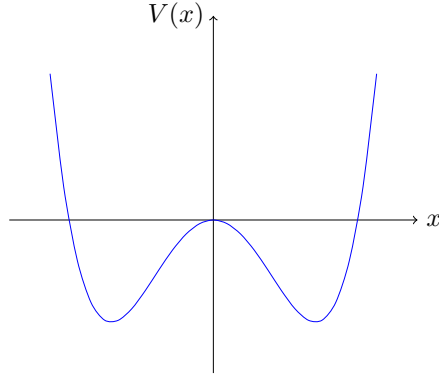


Figure 4.1: Example of  $V$  in dimension  $d = 1$ .

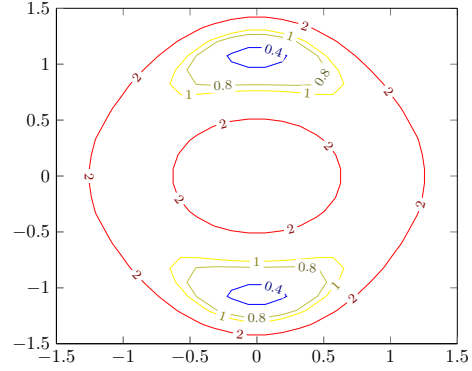


Figure 4.2: Example of  $V$  in dimension  $d = 2$ .

**Assumptions A-2.** Let  $\theta_1$  and  $r$  be the positive constants introduced in (V-2) and (V-3). Consider the following hypotheses concerning the interaction:

- (F-1) The interaction potential is a regular function  $F \in C^2(\mathbb{R}^d; \mathbb{R})$ .
- (F-2)  $F(0) = 0$  and  $\nabla F$  is rotationally invariant, that is there exists a continuous function  $\phi : [0; \infty) \rightarrow \mathbb{R}$  with  $\phi(0) = 0$  such that

$$\nabla F(x) = \frac{x}{|x|} \phi(|x|).$$

- (F-3) There exists a constant  $C' > 0$  such that

$$|\nabla F(x)| \leq C'(1 + |x|^{2r-1}), \quad \text{for all } x \in \mathbb{R}^d.$$

- (F-4) The function  $\nabla F$  is locally Lipschitz. More precisely, for any  $x \in \mathbb{R}^d$  and  $y \in \mathbb{R}^d$ , we have:

$$|\nabla F(x) - \nabla F(y)| \leq C'|x - y|(1 + |x|^{2r-1} + |y|^{2r-1}). \quad (4.4)$$

- (F-5) There exists a constant  $\theta_2 > 0$  such that for any  $x \in \mathbb{R}^d$  we have

$$\nabla^2 F(x) \succeq -\theta_2 \text{Id},$$

where Id is the identity matrix. Moreover,  $\theta_1 > \theta_2$ .

Again, Assumption  $(F - 1)$  is natural since we will use Itô calculus. Assumption  $(F - 2)$  is taken to ensure existence and uniqueness of the process following the work [HIP08], where a similar assumption was introduced. We point out that the exact value of  $F(0)$  does not have any effect on our methods, however, taking it equal to 0 simplifies the writing. Note that we do not use assumption  $(F - 2)$  for proving the exit-time result. Assumption  $(F - 3)$  is required for using the method developed in [BRTV98, HIP08] about the existence of the self-stabilizing diffusion when the drift is super-linear. We point out that  $\nabla F$  is not assumed to be globally Lipschitz. Assumption  $(F - 5)$  is taken in order to guarantee that the attractive behaviour at infinity of  $V$  will not be overcome by  $F$ , which is essential for existence and uniqueness results (we provide this result on page 100).

Assumption A-2 covers a wide range of possible interaction potentials defining various behaviour with respect to the law of the process (attractive, repulsive or the combination of two). A classical analytical example of the interaction potential in general dimension  $d$  is

$$F(x) := \pm \frac{\alpha}{2} |x|^2,$$

with  $\alpha > 0$ . In the case of  $F(x) = \frac{\alpha}{2} |x|^2$  (see Fig. 4.3 for its depiction in  $d = 1$ ), the interacting potential is globally convex and induces attracting behaviour, whereas it is globally concave and thus repulsive with the negative sign. Another possible example of a potential is

$$F(x) := C e^{-\frac{\theta}{2} |x|^2},$$

with  $\theta > 0$  (for its graph in  $d = 1$  see Fig. 4.4). In this case, the function is neither convex nor concave, but, after a careful examination, we can see that it still exhibits repulsive behaviour, though dissipating at infinity. Note that here, despite assumption  $(F - 2)$ ,  $F(0) \neq 0$ . As was pointed out above, the translations of  $F$  do not influence the dynamic of (4.1).

In the following, we introduce the assumptions on the domain. First, we define the effective (in the small-noise limit) potential.

**Definition 4.1.** Let  $a$  be the local minimizer of  $V$  introduced in A-1. Then,  $W_a \in C^2(\mathbb{R}^d; \mathbb{R})$  such that  $W_a := V + F * \delta_a = V + F(\cdot - a)$  is called the effective potential.

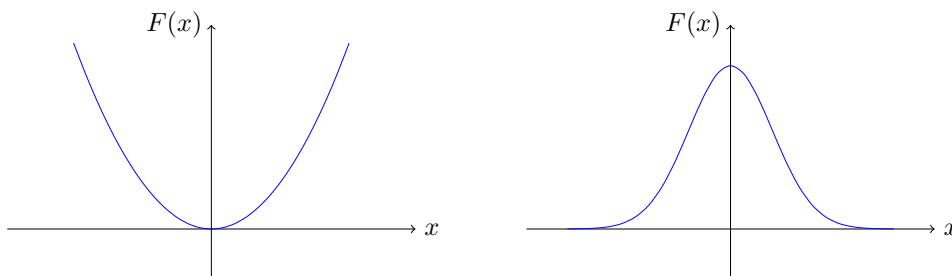


Figure 4.3: Example of a convex  $F$ . Figure 4.4: Example of a non-convex  $F$ .

The name “effective” comes from the fact that, as will be proved below, before the exit-time from the stable domain  $\mathcal{D}$ , for small  $\sigma$ , the potential  $V + F * \mu_t^\sigma$ , inducing the drift term of our process, is well approximated by  $W_a$ .

In order to ensure that, in the small-noise limit, our process behaves well around the attractor  $a$ , we need to assume that  $a$  is also a stable local minimizer of the effective potential. Consider the following assumption

**Assumptions A-4.** The matrix  $\nabla^2 W_a(a) = \nabla^2 V(a) + \nabla^2 F(0)$  is positive definite.

Note that Assumption A-4, along with the continuity assumptions on  $\nabla^2 V$  and  $\nabla^2 F$  (Assumptions A-1 and A-2), leads to the fact that we can find an open neighborhood of the point  $a$  such that  $W_a$  is convex inside it. Consider:

**Definition 4.2.** Let  $\rho > 0$  be a small enough positive number such that  $W_a$  is convex inside  $B_\rho(a)$ . Let  $C_W > 0$  be a constant such that for any  $x \in B_\rho(a)$ :

$$\nabla^2 W_a(x) = \nabla^2 V(x) + \nabla^2 F(x - a) \succeq C_W \text{Id},$$

where  $\text{Id}$  is the identity matrix.

Let us now introduce assumptions regarding the domain of interest  $\mathcal{D} \subset \mathbb{R}^d$ , exit-time from which will be considered in the future. First assumption on domain  $\mathcal{D}$  is the following:

**Assumptions A-5.**  $\mathcal{D}$  is a bounded connected open subset of  $\mathbb{R}^d$  containing the point  $a$ .

**Remark 4.3.** Without loss of generality, we choose  $\rho > 0$  from Definition 4.2 to be small enough such that we have the following strict inclusion  $B_\rho(a) \subset \mathcal{D}$ .

The boundedness of the domain  $\mathcal{D}$  will be relaxed later. However, the fact that  $\mathcal{D}$  is connected and open is mandatory and classical from [DZ10, FW98].

The following assumptions on  $\mathcal{D}$  are mandatory:

**Assumptions A-6.** The domain  $\mathcal{D}$  contains the deterministic path  $(\gamma_t, t \geq 0)$  solution of the following dynamical system

$$\frac{d}{dt}\gamma_t = -\nabla V(\gamma_t), \quad \gamma_0 = x_{\text{init}}. \quad (4.5)$$

We assume furthermore that  $\lim_{t \rightarrow \infty} \gamma_t = a$ .

This assumption is important for the type of exit-problem that we consider here, which is exit created by the small noise from a domain of attraction. We will see further, using the large deviations principle (LDP), that for any  $T > 0$ , the processes  $(X_t^\sigma, 0 \leq t \leq T)$  and  $(\gamma_t, 0 \leq t \leq T)$  are close in supremum norm with high probability when  $\sigma$  is small enough. In the case where  $T_1 := \inf\{t \geq 0 : \gamma_t \notin \mathcal{D}\} < \infty$ , it is easy to show, using LDP, that  $\tau_{\mathcal{D}}^\sigma \approx T_1$  for small  $\sigma$ . In other words, it is impossible to obtain the Kramers' type law without Assumption A-6.

Now, we present the definition of a stable domain.

**Definition 4.4.** We say that an open connected subset  $\mathcal{G}$  of  $\mathbb{R}^d$  is stable by the vector field  $-\nabla W_a$  if for any  $t \geq 0$ , for any  $x \in \mathcal{G}$ ,  $\psi_t(x) \in \mathcal{G}$  where the process  $\psi(x)$  is the solution to the following dynamical system:

$$\psi_t(x) = x - \int_0^t \nabla W_a(\psi_s(x)) \, ds.$$

This leads to a classical assumptions on the domain  $\mathcal{D}$  that is standard for the Freidlin-Wentzell theory, see [DZ10, FW98].

**Assumptions A-7.** The open domain  $\mathcal{D}$  is stable by the vector field  $-\nabla W_a$ . Moreover, for any  $z \in \partial\mathcal{D}$ ,  $\lim_{t \rightarrow +\infty} \psi_t(z) = a$ .

**Remark 4.5.** Note that by continuity argument we can expand domain  $\mathcal{D}$  such that Assumptions A-6 and A-7 still hold in the enlargement. Namely, for any  $\kappa > 0$  small enough there exists an open connected bounded set  $\mathcal{D}_\kappa^e \subseteq \{x \in \mathbb{R}^d : \inf_{z \in \mathcal{D}} |z - x| < \kappa\}$  such that Assumptions A-6 and A-7 are satisfied for  $\mathcal{D}_\kappa^e$ . Obviously, the same holds for constrictions: for any  $\kappa > 0$  small enough there exists an open set  $\mathcal{D}_\kappa^c \subseteq \{x \in \mathcal{D} : \inf_{z \in \partial\mathcal{D}} |z - x| > \kappa\}$  satisfying Assumptions A-6 and A-7.

We can also define their exit-costs as  $H_\kappa^e := \inf_{z \in \partial\mathcal{D}_\kappa^e} \{W_a(z) - W_a(a)\}$  and  $H_\kappa^c := \inf_{z \in \partial\mathcal{D}_\kappa^c} \{W_a(z) - W_a(a)\}$  respectively.

## Existence of the process

First, we consider the initial SDE with any initial condition.

$$X_t^\sigma = X_0 + \sigma W_t - \int_0^t \nabla V(X_s^\sigma) ds - \int_0^t \nabla F * \mu_s^\sigma(X_s^\sigma) ds. \quad (4.6)$$

Mutatis mutandis from [HIP08, Theorem 2.13], we get the following proposition:

**Proposition 4.6.** Let  $r$  be the positive constant introduced in (V-3) of Assumptions A-1. For any random variable  $X_0$  such that  $\mathbb{E}[|X_0|^{8r^2}] < +\infty$ , for any  $\sigma \geq 0$ , under Assumptions A-1 and A-2, the SDE (4.6) has a unique strong solution that we denote by  $(X_t^\sigma, t \geq 0)$ . Moreover, there exists a constant  $M > 0$  which only depends on  $\mathbb{E}[|X_0|^{8r^2}]$  such that

$$\sup_{0 \leq \sigma \leq 1} \sup_{t \geq 0} \mathbb{E}[|X_t^\sigma|^{8r^2}] \leq M. \quad (4.7)$$

Note that the assumptions used in [HIP08, Theorem 2.13] are slightly different from ours, particularly for the interaction term. Assumption (F-



2) of A-2 allows  $\phi$  to be negative and thus to exhibit repulsive behaviour, while in [HIP08]  $\phi$  is set to be a positive increasing function. To neutralise possible problems that this relaxation could pose, we introduce assumption (F-5). The fact that  $\theta_1 > \theta_2$  guarantees that, regardless of  $\mu^\sigma$ , the drift term of our process is always attractive outside of a compact set. Namely, for any  $\mu \in \mathcal{P}(\mathbb{R}^d)$  and for any  $x \in \mathbb{R}^d$  such that  $|x| > R$ , we have  $\nabla^2 V(x) + \nabla^2 F * \mu(x) \succeq (\theta_1 - \theta_2)\text{Id}$  and thus

$$\langle x; -\nabla V(x) - \nabla F * \mu(x) \rangle \leq -(\theta_1 - \theta_2)|x|^2.$$

This guarantees non-explosiveness of the process in finite time. After this observation, the proof in [HIP08] can be easily adapted for the case of Assumptions A-1 and A-2.

Note that the process (4.1) that we study in this chapter corresponds to (4.6) with for  $X_0 = x_{\text{init}} \in \mathbb{R}^d$ .

Basically, by Proposition 4.6, unique solutions to (4.6), obtained for different  $0 \leq \sigma \leq 1$  and fixed random variables  $X_0$ , form a family of stochastic processes to which the processes of our interest of the form (4.1) also belong. In the following, it is this process that we denote as  $X^\sigma$  if not stated otherwise.

Note, that the case  $\sigma = 0$  corresponds to the following ODE:

$$dX_t^0 = -\nabla V(X_t^0) dt - \nabla F * \mu_t^0(X_t^0) dt, \quad (4.8)$$

where, as it was before,  $\mu_t^0 = \mathcal{L}(X_t^0)$ .

## 4.1 Main results

In this paragraph, we list the main results of the chapter.

### Large deviations principle

For the case of random initial condition, define  $X_t^{x,\sigma} := \mathbb{E}[X_t^\sigma | X_0 = x]$ , where processes  $(X_t^\sigma, t \geq 0)$ , defined for different  $0 < \sigma < 1$ , are the unique strong solutions to (4.6) with some fixed initial condition  $X_0$  and respective noise level  $\sigma$ . We recall that  $\mu_t^0 = \mathcal{L}(X_t^0)$  and state the following

**Theorem 4.1.** *For any  $T > 0$  and for any  $x \in \mathbb{R}^d$ , the measures induced by  $(X_t^{x,\sigma}, 0 \leq t \leq T)$  on  $C([0, T]; \mathbb{R}^d)$  satisfy the LDP with convergence rate  $\frac{\sigma^2}{2}$  and with the following good rate function:*

$$I_T(\varphi) := \frac{1}{4} \int_0^T |\dot{\varphi}_t + \nabla V(\varphi_t) + \nabla F * \mu_t^0(\varphi_t)|^2 dt, \quad (4.9)$$

for any  $\varphi \in \mathcal{H}_1$ , the set of absolutely continuous functions  $\varphi$  from  $[0; T]$  to  $\mathbb{R}^d$  such that  $\varphi(0) = x$ . Otherwise,  $I_T(\varphi) := +\infty$ .

Proof is provided in Section 4.3.

**Remark 4.7.** Most authors use the convergence rate  $\sigma^2$  and consequently, the term in front of the integral in (4.9) is  $\frac{1}{2}$  instead of  $\frac{1}{4}$ . However, we choose to take as convergence rate the coefficient in front of the Laplacian in the associated partial differential equation.

Note that, in the case of (4.1), the theorem above takes the form:

**Corollary 4.8.** Let  $\gamma$  be the unique solution of the following ODE

$$\frac{d}{dt}\gamma_t = -\nabla V(\gamma_t), \quad \gamma_0 = x_{\text{init}}.$$

Then for any  $T > 0$ , the process  $(X_t^\sigma, 0 \leq t \leq T)$  defined by the SDE (4.1) satisfies a LDP with convergence rate  $\frac{\sigma^2}{2}$  with the following good rate function:

$$I_T(\varphi) := \frac{1}{4} \int_0^T |\dot{\varphi}_t + \nabla V(\varphi_t) + \nabla F(\varphi_t - \gamma_t)|^2 dt, \quad (4.10)$$

for any  $\varphi \in \mathcal{H}_1$ , the set of absolutely continuous functions  $\varphi$  from  $[0; T]$  to  $\mathbb{R}^d$  such that  $\varphi(0) = x_{\text{init}}$ . Otherwise,  $I_T(\varphi) := +\infty$ .

## Exit-time and exit-location

We now give the main results concerning the exit-time, for the case when  $\mathcal{D}$  is a bounded domain.

**Theorem 4.2.** *Let  $H$  be the exit-cost introduced in Assumption A-7. Under Assumptions A-1–A-7, the following two results hold*

1. *Kramers' law: for any  $\delta > 0$ , the following limit holds:*

$$\lim_{\sigma \rightarrow 0} \mathbb{P} \left[ \exp \left\{ \frac{2}{\sigma^2} (H - \delta) \right\} \leq \tau_{\mathcal{D}}^{\sigma} \leq \exp \left\{ \frac{2}{\sigma^2} (H + \delta) \right\} \right] = 1. \quad (4.11)$$

2. *Exit-location: for any closed set  $N \in \partial \mathcal{D}$  such that  $\inf_{z \in N} W_a(z) > H$  the following limit holds:*

$$\lim_{\sigma \rightarrow 0} \mathbb{P}(X_{\tau_{\mathcal{D}}^{\sigma}}^{\sigma} \in N) = 0. \quad (4.12)$$

Proof of Theorem 4.2 is provided in Section 4.3.

## Control of the law

We now present a result on the control of the law in the case where  $\mathcal{D}$  is bounded. The following theorem rigorously states that, starting from some uniformly bounded in  $\sigma$  time, the law of the process  $\mu^{\sigma}$  stays close to  $\delta_a$  long enough to obtain the result of Theorem 4.2.

**Theorem 4.3.** *Under Assumptions A-1–A-7, for any  $\kappa > 0$  small enough there exist  $\bar{T}_{\text{st}}(\kappa) > 0$  and  $\sigma_{\kappa} > 0$  such that*

$$\sup_{0 < \sigma < \sigma_{\kappa}} \sup_{t \in \left[ \bar{T}_{\text{st}}(\kappa); e \frac{2H}{\sigma^2} \right]} \mathbb{W}_2(\mu_t^{\sigma}; \delta_a) \leq \kappa.$$

This theorem can be easily proven using Lemmas 4.11 and 4.16 provided in Section 4.2. It is left for the reader

## Unbounded case

We now present the generalisation of the results above to the case where  $\mathcal{D}$  is not bounded.

**Corollary 4.9.** *If  $\mathcal{D}$  is an open and connected subset of  $\mathbb{R}^d$ , under Assumptions A-1–A-4 and Assumptions A-6, A-7, the statements of Theorem 4.2 hold.*

The control of the law also holds immediately even if  $\mathcal{D}$  is unbounded.

**Corollary 4.10.** If  $\mathcal{D}$  is an open and connected subset of  $\mathbb{R}^d$ , under Assumptions A-1–A-4 and Assumptions A-6, A-7, the statement of Theorem 4.3 holds.

Proofs of Corollary 4.9 and Corollary 4.10 are postponed to Section 4.3.

## Comparison to previous results

In the seminal work [HIP08], S. Herrmann, P. Imkeller, and D. Peithmann proved the existence of the self-stabilizing diffusion in the irreversible case. The assumptions they used correspond to A-1 and A-2 if confinement and interaction were gradients of some regular potentials, except for a slight difference in the interaction term (this difference was discussed on page 100). In the same work, the authors show the exit-time result for SSD, but, in order to do that, they had to assume convexity of confinement and interaction. Removal of this assumption, that we present in this thesis is a big improvement of previous results. Note, that, unlike in [HIP08], we solve the exit-time problem for the reversible case (confinement and interaction are gradients of some regular functions). Nevertheless, we could treat the general situation, see Section 4.1 on the possible extensions of our results.

Another difference between our approach and the one presented in the paper [HIP08] is that, after controlling the law of the process  $X^\sigma$ , we use coupling techniques to prove the exit-time, while the approach used by S. Herrmann, P. Imkeller, and D. Peithmann consists in reconstructing the Freidlin-Wentzell techniques and taking advantage of the contractivity of the drift.

In [Tug12], J. Tugaut focused on reversible case of SSD with potentials  $V$  and  $F$  being convex. He proved a similar to ours result by using another method than in [HIP08]. The approach of [Tug12] was to apply the Freidlin-Wentzell theory without adapting it to the McKean-Vlasov diffusions. In this work, the classical large deviations principle theory for processes is used to the associated system of particles

$$dX_t^{i,N} = \sigma dW_t - \left( \nabla V(X_t^{i,N}) + \frac{1}{N} \sum_{j=1}^N \nabla F(X_t^{i,N} - X_t^{j,N}) \right) dt, \quad (4.13)$$

after which a trajectorial uniform propagation of chaos is established. Using the propagation of chaos, the author obtained the Kramers' type law.

In [Tug16] J. Tugaut employed a different method, applicable to the case where the drifts are not necessarily assumed to be gradients of a regular function, although they remain uniformly convex. This method primarily revolves around controlling the law at time  $t$  of  $X^\sigma$ , denoted as  $\mu_t^\sigma$ . Notably, J. Tugaut demonstrated that this law converges to  $\delta_a$  in Wasserstein distance for  $t \rightarrow +\infty$ . Subsequently, a synchronous coupling with a diffusion, where the drift is represented as  $x \mapsto -\nabla V(x) - \nabla F * \delta_a(x)$  instead of  $x \mapsto -\nabla V(x) - \nabla F * \mu_t^\sigma(x)$ , is employed. Exploiting the contractivity, it is straightforward to prove that the two diffusions remain close. Consequently, the exit-time of  $X^\sigma$  behaves similarly to that of the coupled diffusion.

This approach has been extended to non-convex scenarios with the reversible case, as described in [Tug18a]. In this context,  $V$  is not necessarily convex, although  $F$  exhibits sufficient convexity to ensure convexity of the effective potential  $W_a = V + F(\cdot - a)$ . As a result, coupling between the two diffusions is straightforward, allowing us to infer the exit-time of  $X^\sigma$  from that of the coupled diffusion.

The convexity assumption on  $W_a$  has been removed in [Tug19], though this result is limited to the one-dimensional case. Unfortunately, the method used there cannot be directly extended to the general-dimensional case. Thus, it becomes essential to find an alternative way to control the law.

In [Tug18b], J. Tugaut demonstrated that  $\mu^\sigma$  does not always converge to  $\delta_a$ . This limitation arises when  $W_a$  fails to reach its global minimum at  $a$ , therefore, in order to control the law of the process (at least until exit-time) other methods should be used.

Despite all these developments, the exit-time problem for SSD with general (non-convex) coefficients was an open problem throughout all these years. We solve it in this chapter by significantly improving the coupling method introduced in [Tug16].

## Discussions on extension

In this section, we provide some possible extensions to our results.

### Non-identity matrix as the diffusion coefficient

In this work, we have simplified the study by assuming that the diffusion coefficient takes the form  $\sigma \text{Id}$ . However, for certain algorithmic applications such as molecular dynamics, it could be beneficial to consider scenarios where the diffusion coefficient is not directly proportional to the identity matrix, as discussed for example in [dRDM<sup>+</sup>].

To make further progress, it would be a significant improvement to include the scenario where the diffusion coefficient is given by  $\sigma M$ , with  $M$  being a non-degenerate matrix. This particular situation has been studied in, for instance, [DT16, DT18, Mon17].

The techniques developed in the present work can be readily adapted for this non-identity diffusion coefficient case.

However, a more challenging extension would involve considering cases where  $M$  is degenerate. This would allow us to address the Langevin kinetic diffusion, where both position and velocity play crucial roles. Combining techniques we have developed with those from [dRDM<sup>+</sup>], we firmly believe that we can obtain valuable insights into the asymptotic behaviour of the first exit-time.

### Initial random variable

Another possible extension is related to the initial random variable. In the current work, we establish the asymptotic behaviour of the exit-time for  $X_0 := x_{\text{init}} \in \mathbb{R}^d$ . However, for studying the basins of attraction, as was done in [Tug21], it is crucial to consider scenarios where  $\mu_0^\sigma := \mathcal{L}(X_0^\sigma)$  is not necessarily a Dirac measure. Specifically, we may be interested in cases where  $\mu_0^\sigma := \mu_0$ , with the measure  $\mu_0$  being compactly supported in  $\mathcal{D}$ .

In this situation, we need to make a slight modification to Assumptions A-6. Instead of considering  $\gamma'(t) = -\nabla V(\gamma_t)$ , we would need to consider the partial differential equation:

$$\frac{\partial}{\partial t} \mu_t^0 = \text{div} (\mu_t^0 (\nabla V + \nabla F * \mu_t^0)),$$

with  $\mu_0^0 = \mu_0$ . This corresponds to the granular media equation with zero noise. The associated dynamical system that approximates the diffusion

$X^0$  on  $[0; T]$  (with  $T > 0$ ) due to the large deviations principle is thus given by:

$$\rho_t(x_{\text{init}}) = x_{\text{init}} - \int_0^t \nabla V(\rho_s(x_{\text{init}})) ds - \int_0^t \nabla F * \mu_s^0(\rho_s(x_{\text{init}})) ds,$$

for any  $x_{\text{init}} \in \text{supp}(\mu_0)$ . In this case, Assumptions A-6 would be: for any  $x_{\text{init}} \in \mathcal{D} \cap \text{supp}(\mu_0)$  and for any  $t \geq 0$ , we have  $\rho_t(x_{\text{init}}) \in \mathcal{D}$ .

The techniques developed in the present work can be seamlessly adapted to handle this situation.

### Reflexion on the boundary

In this work, the diffusion process takes place in the entire phase space  $\mathbb{R}^d$ . However, we can consider a subspace of  $\mathbb{R}^d$  instead. This could be achieved by introducing a reflection on certain boundaries, as it was done, for example, in [Tan79]. Such an extension would be a significant improvement compared to [AdRR<sup>+</sup>22], where the uniform convexity of both confinement and interaction potentials was assumed.

In the mentioned article, the domain  $\mathcal{G}$  in which the diffusion takes place satisfies  $d(\overline{\mathcal{D}}; \partial\mathcal{G}) > 0$ , which simplifies the study. We believe that techniques we have developed could treat this case. However, considering scenarios where  $\mathcal{D} \cap \mathcal{G}^c \neq \emptyset$  is more challenging. This could require extending the large deviation techniques for processes with reflection, something that is not done yet even for linear case.

### More accurate estimates

In this chapter, our focus has been on establishing the Kramers' law, that is a limit in probability of  $\frac{\sigma^2}{2} \log(\tau_{\mathcal{D}}^\sigma)$  as  $\sigma$  approaches 0, as well as the exit-location result. However, in [HIP08], the authors have obtained a more precise estimate, which could be of interest in our context. For example, the so-called Arrhenius law was established, i.e. the convergence of

$$\frac{\sigma^2}{2} \log \mathbb{E}(\tau_{\mathcal{D}}^\sigma) \xrightarrow{\sigma \rightarrow 0} H > 0.$$

Unfortunately, since we do not provide the control of the law of the process after the exit-time, we could not use the standard method to show the Arrhenius law in the current work.

Additionally, it is well-known, as discussed in [NRS85], that the first exit-time  $\tau_{\mathcal{D}}^{\sigma}$  for a linear (Itô) diffusion satisfies the following limit:

$$\frac{\tau_{\mathcal{D}}^{\sigma}}{\mathbb{E}[\tau_{\mathcal{D}}^{\sigma}]} \xrightarrow[\sigma \rightarrow 0]{\mathcal{L}} \mathcal{E}(1),$$

where the convergence is meant in law, and  $\mathcal{E}(1)$  is the exponential law with a parameter equal to 1. The same behaviour for self-stabilizing diffusions is not established yet even in the case where both  $V$  and  $F$  are convex.

In [BEGK04, BGK05], A. Bovier, M. Eckhoff, V. Gaynard, and M. Klein studied the exit-time problem for linear reversible diffusion process using potential theory approach. Using these techniques, the authors could not only establish the Arrhenius law for multi-well potential in  $\mathbb{R}^d$ , but also prefactor of the convergence. Namely, the following equality was established:

$$\mathbb{E}[\tau_{\mathcal{D}}^{\sigma}] = C^* e^{\frac{2H}{\sigma^2}} (1 + O(\sigma |\log(\sigma)|)),$$

where the constant  $C^* > 0$  depends on the derivatives of the potential  $V$  at the point of attraction  $a$  as well as the saddle points surrounding the well under consideration. For the explicit form of the prefactor see [BEGK04].

Similar methods could be also used for the self-stabilizing diffusion. However, that would imply studying the associated PDE for the law of the process:

$$\frac{\partial}{\partial t} \mu_t^{\sigma} = \frac{\sigma^2}{2} \Delta \mu_t^{\sigma} + \operatorname{div} (\mu_t^{\sigma} (\nabla V + \nabla F * \mu_t^{\sigma})),$$

which is considered to be a hard problem due to its non-linearity. These questions could be the focus of future studies.

### Non-reversible case

In this work, we have focused on the case where both the confinement and the interaction terms are gradients of some potentials. However, it would be valuable to consider non-reversible situations of the form:



$$X_t = X_0 + \sigma MB_t + \int_0^t a(X_s) ds + \int_0^t b * \mu_s^\sigma(X_s) ds,$$

where  $a$  and  $b$  are general vector fields on  $\mathbb{R}^d$ . It is worth noting that in previous works such as [dRDM<sup>+</sup>, HIP08, Tug16], the authors have successfully addressed this problem, but in the contractive (convex confinement and interaction) case.

The techniques developed in this chapter can readily be adapted to handle the non-reversible case. However, the exit-cost is not explicit in this situation, which is why we have described the reversible case here.

### More general McKean-Vlasov diffusions

A broader class of nonlinear diffusion processes can be considered. For example:

$$dX_t = \sigma dB_t - \nabla V(X_t) dt - b(X_t, \mu_t^\sigma) dt,$$

where the nonlinear drift  $b$  takes the form

$$b(x, \mu) := \int_{\mathbb{R}^d} B(x, y) \mu(dy).$$

Here, the function  $B$  is required to be regular and maps from  $\mathbb{R}^d \times \mathbb{R}^d$  to  $\mathbb{R}^d$ . Such a generalization would have significant implications for theoretical purposes (as shown in [GS20]) as well as applications (see e.g. [GGM<sup>+</sup>18]). We firmly believe that the techniques developed in this work can be adapted to handle a wide range of situations within this framework.

For algorithmic applications, it would be also interesting to include jumps in the process, as discussed in [Gra92, GM89]. This could be a subject of future research endeavors.

### Extension on the domain $\mathcal{D}$ and metastability

An important yet challenging extension concerns the domain  $\mathcal{D}$  itself. In this work, we have confined our study to cases where  $\overline{\mathcal{D}}$  is stable under the effective potential  $W_a$ . However, the most interesting scenario arises when the saddle point lies on the boundary of  $\mathcal{D}$ .

Moreover, it would be interesting to establish some metastable properties of  $X^\sigma$ , that is considering  $t(\sigma)$  as a function of  $\sigma$  and investigating  $X_{t(\sigma)}^\sigma$  in metastable confinement as it was done in [FW98]. Complexity of this problem in the case of SSD is that the drift itself (the effective potential) may change after the transition of the process from one metastable state to another. These questions could be the focus of future studies.

### System of Particles

For algorithmic applications, it is essential to consider the associated system of particles described by equation (4.13). In this system, the measure  $\mu_t^X$  is replaced by  $L_t^\sigma = \frac{1}{N} \sum_{j=1}^N \delta_{X_t^j}$ .

In [Tug12], J. Tugaut has obtained the exit-time of the McKean-Vlasov diffusion from the system of particles in the convex case. Consequently, it appears feasible to do the opposite and establish the exit-time of the system of particles based on the exit-time of the McKean-Vlasov diffusion. Similar techniques like a trajectorial uniform propagation of chaos (see for example [BRTV98, CGM08, Szn91]) can be used. However, in [Tug12], convexity was essential for controlling the law, which is now also available in the general situation due to the current work.

## 4.2 Intermediate results

In this preliminary section, we will give the key results which allow us to prove the main results related to exit-time in Section 4.3. Their proofs are given in Section 4.4.

## Stabilisation in finite time

Let us define the following two deterministic times for any  $\kappa > 0$  small enough:

$$\begin{aligned} T_{\text{st}}^\sigma(\kappa) &:= \inf \{t \geq 0 : \mathbb{W}_2(\mu_t^\sigma; \delta_a) \leq \kappa\}, \\ S_{\text{st}}^\sigma(\kappa) &:= \inf \{t \geq T_{\text{st}}^\sigma(\kappa) : \mathbb{W}_2(\mu_t^\sigma; \delta_a) > \kappa\}; \end{aligned}$$

and we let the infima to be equal to  $+\infty$  if respective sets are empty.

First key result consists in obtaining the existence of a time  $T$  such that  $\mathbb{W}_2(\mu_t^\sigma; \delta_a)$  is small and such that  $X_T^\sigma$  is concentrated around  $a$ .

**Lemma 4.11.** *Under Assumptions A-1–A-7, for any  $\kappa > 0$  there exist  $\bar{T}_{\text{st}}(\kappa) > 0$  and  $\sigma_\kappa > 0$  such that:*

$$T_{\text{st}}^\sigma(\kappa) \leq \bar{T}_{\text{st}}(\kappa) \quad \text{for any } 0 < \sigma < \sigma_\kappa.$$

Moreover,

$$\lim_{\sigma \rightarrow 0} \mathbb{P} \left( \left| X_{\bar{T}_{\text{st}}(\kappa)}^\sigma - a \right| > \kappa \right) = 0.$$

An important implication of this lemma is that, with high probability, the exit from the domain  $\mathcal{D}$  does not occur before time  $\bar{T}_{\text{st}}(\kappa)$  (see Section 4.4 for the proof). Consider the following corollary.

**Corollary 4.12.** Under Assumptions A-1–A-7, for any  $\kappa > 0$  the following limit holds:

$$\lim_{\sigma \rightarrow 0} \mathbb{P}(\tau_{\mathcal{D}}^\sigma \leq \bar{T}_{\text{st}}(\kappa)) = 0.$$

## The coupling method

We now introduce the diffusion  $Y^\sigma := (Y_t^\sigma, t \geq T_{\text{st}}^\sigma(\kappa))$  solution to the following linear SDE:

$$\begin{aligned} Y_t^\sigma &= X_{T_{\text{st}}^\sigma(\kappa)}^\sigma + \sigma(W_t - W_{T_{\text{st}}^\sigma(\kappa)}) - \int_{T_{\text{st}}^\sigma(\kappa)}^t \nabla V(Y_s^\sigma) ds \\ &\quad - \int_{T_{\text{st}}^\sigma(\kappa)}^t \nabla F(Y_s^\sigma - a) ds, \end{aligned} \tag{4.14}$$

where  $(W_t, t \geq 0)$  is the same Brownian motion that drives the main equation (4.1). Note, that this SDE has a unique solution (see for example [SV79, Theorem 10.2.2, p. 255]).

Note also that  $Y^\sigma$  is a linear diffusion. As a consequence, we can apply the classical Freidlin-Wentzell theory, see [DZ10, FW98], for estimating the first exit-time as the diffusion coefficient tends to 0.

Apart from the processes  $(Y^\sigma)_{0 < \sigma < 1}$  that is defined by SDE (4.14), we also define the following family of processes that constitute Itô diffusions and will help us to study stochastic properties of  $Y^\sigma$ . For any  $y \in \mathbb{R}^d$  and for any  $0 < \sigma < 1$  define  $(Y_t^{y,\sigma}, t \geq 0)$  as the unique solution to the following SDE:

$$Y_t^{y,\sigma} = y + \sigma W_t - \int_0^t \nabla V(Y_s^{y,\sigma}) ds - \int_0^t \nabla F(Y_s^{y,\sigma} - a) ds. \quad (4.15)$$

Following the standard notation for diffusions, we will drop the initial point  $y$  for  $Y^{y,\sigma}$ , as well as for all random variables that are functions of  $Y^{y,\sigma}$ , and put it as a subscript under the probability measure. Namely, for any  $y \in \mathbb{R}^d$  we introduce a probability measure  $\mathbb{P}_y$  that is simply a restriction of  $\mathbb{P}$  to the measurable space  $(\Omega, \sigma(Y_t^{y,\sigma} : t \geq 0))$ .

The following proposition is a classical result of Freidlin–Wentzell theory for the exit-time of linear diffusions of the type (4.15). Consider:

**Proposition 4.13** ([DZ10], Theorem 5.7.11). Let Assumption A-1 be satisfied and let  $G \subset \mathbb{R}^d$  be a domain such that Assumptions A-5–A-7 are satisfied for it and its exit-cost  $H_G := \inf_{z \in \partial G} \{W_a(z) - W_a(a)\}$ . Let  $K \subset G$  be a compact set. Define  $\tau_G^{Y,\sigma} := \inf\{t \geq 0 : Y_t^{y,\sigma} \notin G\}$ . Then, for any  $\delta > 0$  we have

$$\limsup_{\sigma \rightarrow 0} \sup_{y \in K} \mathbb{P}_y \left( \exp \left\{ \frac{2(H_G - \delta)}{\sigma^2} \right\} \leq \tau_G^{Y,\sigma} \leq \exp \left\{ \frac{2(H_G + \delta)}{\sigma^2} \right\} \right) = 1.$$

Obviously, this theorem also holds when  $G$  is the domain  $\mathcal{D}_\kappa^e$  defined as in Remark 4.5 and  $H_G = H_\kappa^e := \inf_{x \in \partial \mathcal{D}_\kappa^e} \{W_a(x) - W_a(a)\}$  respectively.

Let us now describe how both diffusion processes  $X$  (the targeted diffusion) and  $Y$  (the auxiliary one) are coupled. We are especially interested in describing the distance between them.

**Proposition 4.14.** Under Assumptions A-1–A-7 there exists  $\eta > 0$  such that for any  $\kappa > 0$  small enough, we have

$$\lim_{\sigma \rightarrow 0} \mathbb{P}(\sup |X_t^\sigma - Y_t^\sigma| > \kappa) = 0,$$

where supremum is taken over  $t \in \left[ T_{\text{st}}^\sigma(\kappa); S_{\text{st}}^\sigma(\kappa) \wedge \exp\left\{ \frac{2(H+\eta)}{\sigma^2} \right\} \right]$ .

As it is shown below (Corollary 4.17), this result can be improved by removing the time  $S_{\text{st}}^\sigma(\kappa)$ , since, as it turns out, the destabilization of the law of the process can not happen before its exit-time from the domain  $\mathcal{D}$ .

The following lemma is an important result stating that, at each point of time, the diffusion  $Y^\sigma$  is close to  $a$  with high probability.

**Lemma 4.15.** *Let  $\rho$  be a positive constant introduced in Definition 4.2. Under Assumptions A-1–A-7 there exists  $\eta > 0$  small enough such that for any  $\kappa > 0$  small enough:*

$$\sup \mathbb{P}(Y_t^\sigma \notin B_{\rho/2}(a)) = o_\sigma(1),$$

where supremum is taken over  $t \in \left[ T_{\text{st}}^\sigma(\kappa); \exp\left\{ \frac{2(H+\eta)}{\sigma^2} \right\} \right]$ .

Note that the position of supremum in Lemma 4.15 is important. Indeed, according to Freidlin-Wentzell theory for Itô diffusions, the exit-time of  $Y^\sigma$  from  $B_{\rho/2}(a)$  is, with high probability, of order  $\exp\{2H_{\rho/2}/\sigma^2\}$ , where  $H_{\rho/2} := \inf_{z \in \partial B_{\rho/2}(a)} \{V(z) + F(z - a) - V(a)\}$ , which means, among other things, that we can not expect  $\mathbb{P}(\sup |Y_t^\sigma - a| > \frac{\rho}{2})$  to be equal to  $o_\sigma(1)$ . Instead, what Lemma 4.15 states is that, with high probability, we find  $Y_t^\sigma$  around  $a$  for all time  $t$  until the exit-time from a stable domain of attraction  $\mathcal{D}_\kappa^e$ . We come back to this description on page 137.

## Control of the law

In this section we present a result regarding the control of the law of the process after the stabilisation time. Consider the following lemma.

**Lemma 4.16.** *Under Assumptions A-1–A-7 there exists  $\eta > 0$  such that for any  $\kappa > 0$  small enough there exists  $\sigma_\kappa$  such that for any  $0 < \sigma < \sigma_\kappa$  we have*

$$S_{\text{st}}^\sigma(\kappa) > \exp\left\{\frac{2(H + \eta)}{\sigma^2}\right\}.$$

This lemma together with Proposition 4.14 immediately gives us the following corollary:

**Corollary 4.17.** Under Assumptions A-1–A-7 there exists  $\eta > 0$  such that for any  $\kappa > 0$  small enough, we have

$$\lim_{\sigma \rightarrow 0} \mathbb{P}(\sup |X_t^\sigma - Y_t^\sigma| > \kappa) = 0,$$

where supremum is taken over  $t \in \left[T_{\text{st}}^\sigma(\kappa); \exp\left\{\frac{2(H+\eta)}{\sigma^2}\right\}\right]$ .

## 4.3 Proofs of the main results

Here, we give the proofs of the main results.

### Large Deviations Principle

Note that the system (4.6) has several non-classical parts in it. First of all, initial condition is random. Second, the law of the process itself is a part of the solution. In order to deal with this separately, we first introduce the process

$$Z_t^\sigma = X_0 + \sigma W_t - \int_0^t \nabla V(Z_s^\sigma) ds - \int_0^t \nabla F * \mu_s^0(Z_s^\sigma) ds. \quad (4.16)$$

Since  $\mu^0$  does not depend on  $Z^\sigma$ , but is given by the unique solution to (4.8), equation (4.16), being a classic SDE with at most polynomial drift, admits a unique solution. In order to prove LDP for (4.6), we will first show

it for equation (4.16) and then we will prove the exponential equivalence between the two.

Consider the following lemma

**Lemma 4.18.** *Under Assumptions A-1 and A-2, for any  $T > 0$  and for any  $x \in \mathbb{R}^d$  the family of laws induced by the process  $Z^{x,\sigma} = \mathbb{E}[Z^\sigma | Z_0^\sigma = x]$  on  $C([0, T]; \mathbb{R}^d)$  satisfies the LDP with the following good rate function:*

$$I_T(\varphi) := \frac{1}{4} \int_0^T |\dot{\varphi}_t + \nabla V(\varphi_t) + \nabla F * \mu_t^0(\varphi_t)|^2 dt,$$

for any  $\varphi \in \mathcal{H}_1^x$ , the set of absolutely continuous functions  $\varphi$  from  $[0; T]$  to  $\mathbb{R}^d$  such that  $\varphi(0) = x$ . Otherwise,  $I_T(\varphi) := +\infty$ . This corresponds to the good rate function in Equation (4.9).

*Proof.* Let us denote by  $G$  the function that maps the trajectory of Brownian motion  $g \in C([0, T]; \mathbb{R}^d)$  to the unique solution  $f \in C([0, T]; \mathbb{R}^d)$  to the following ODE:

$$f(t) = x - \int_0^t \nabla V(f(s)) ds - \int_0^t \nabla F * \mu_s^0(f(s)) ds + g(t).$$

By existence and uniqueness result, non-explosiveness of  $f$  in finite time is guaranteed. Moreover, it can be shown that the bound on  $f$  can be controlled in terms of function  $g$ . Namely, for any  $g$  there exists  $R > 0$  such that for any  $g' \in C([0, T]; \mathbb{R}^d)$  satisfying  $\|g - g'\|_\infty \leq 1$  we have  $G(g')(s) \in B_R(0)$  for any  $s \leq T$ . Therefore, given  $g$  and thus  $f$ , we can restrict ourselves to the ball  $B_R(0)$ . Note that both  $\nabla V$  and  $\nabla F * \mu_s^0$  are locally Lipschitz by A-1 and A-2 and the properties of convolution. Moreover, the Lipschitz constant of the latter does not depend on time  $s$ , since the integrals  $\int |x|^{8r^2} \mu_s^0(dx)$  are uniformly in time bounded. Therefore, there exists a constant  $C > 0$  such that  $\nabla V + \nabla F * \mu_s^0$  is  $C$ -Lipschitz continuous inside the set  $B_R(0)$ .

Consider the following estimate for fixed  $\delta > 0$ ,  $g, f \in C([0; T]; \mathbb{R}^d)$  and

$g', f' := G(g')$  chosen such that  $\|g - g'\|_\infty \leq \delta$ .

$$\begin{aligned} |f(t) - f'(t)| &\leq \int_0^t \left| \nabla V(f(s)) + \nabla F * \mu_s^0(f(s)) \right. \\ &\quad \left. - \nabla V(f'(s)) - \nabla F * \mu_s^0(f'(s)) \right| ds + |g(t) - g'(t)| \\ &\leq C \int_0^t |f(s) - f'(s)| ds + \|g - g'\|_\infty. \end{aligned}$$

Thus, by the Grönwall inequality:

$$|f(t) - f'(t)| \leq \delta e^{Ct},$$

which establishes the continuity of the mapping  $G$  and the Large deviations principle for diffusion (4.16) by the Contraction principle and the Schilder theorem (see [DZ10, Theorem 4.2.1, Theorem 5.2.3]).  $\square$

*Proof of Theorem 4.1.* Let us show the exponential equivalence of processes  $Z^{x,\sigma}$  defined by equation (4.16) and  $X^{x,\sigma}$ , for which it is sufficient to prove that for any  $\delta > 0$ :

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}(\|Z^{x,\sigma} - X^{x,\sigma}\|_\infty \geq \delta) = -\infty.$$

**Step 1.1.** Let us first show that until both  $Z^{x,\sigma}$  and  $X^{x,\sigma}$  leave the ball of radius  $R > 0$  their trajectories stay close to each other. In order to do so, consider first only those  $\omega \in \Omega$  such that  $t \leq \tilde{\tau}_R := \inf\{s \geq 0 : X^{x,\sigma} \notin B_R(0) \text{ or } Z^{x,\sigma} \notin B_R(0)\}$ . As was pointed out before, both  $\nabla V$  and  $\nabla F * \mu_t^\sigma$  are Lipschitz continuous inside the set  $B_R(0)$  with Lipschitz constant that does not depend on either  $\sigma$  or  $t$  as long as  $0 \leq \sigma \leq 1$  and  $0 \leq t \leq T$ . Let us call this constant  $C_R$ . Consider the following equation for  $t \leq \tilde{\tau}_R$ :

$$\begin{aligned} |Z_t^{x,\sigma} - X_t^{x,\sigma}| &\leq \int_0^t |\nabla V(Z_s^{x,\sigma}) - \nabla V(X_s^{x,\sigma})| ds \\ &\quad + \int_0^t |\nabla F * \mu_s^0(Z_s^{x,\sigma}) - \nabla F * \mu_s^\sigma(X_s^{x,\sigma})| ds \\ &\leq C_R \int_0^t |Z_s^{x,\sigma} - X_s^{x,\sigma}| ds + I_1, \end{aligned}$$



where  $I_1$  is the second term of the equation above that we consider separately

**Step 1.2.** Let us consider the term  $I_1$ . Let us introduce random variables  $\tilde{X}_s^0$  and  $\tilde{X}_s^\sigma$  that are defined on some auxiliary probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ , have distributions  $\mu_s^0$  and  $\mu_s^\sigma$  respectively, and are coupled in the optimal way in the sense that  $\tilde{\mathbb{E}}|\tilde{X}_s^0 - \tilde{X}_s^\sigma|^2 = \mathbb{W}_2^2(\mu_s^0; \mu_s^\sigma)$ . Then,  $I_1$  can be expressed as:

$$\begin{aligned} I_1 &= \int_0^t \left| \tilde{\mathbb{E}} \left[ \nabla F(Z_s^{x,\sigma} - \tilde{X}_s^0) - \nabla F(X_s^{x,\sigma} - \tilde{X}_s^\sigma) \right] \right| ds \\ &\leq C_R \int_0^t |Z_s^{x,\sigma} - X_s^{x,\sigma}| ds + C_R \tilde{\mathbb{E}} \left[ \int_0^t |\tilde{X}_s^0 - \tilde{X}_s^\sigma| ds \right] \\ &\leq C_R \int_0^t |Z_s^{x,\sigma} - X_s^{x,\sigma}| ds + TC_R \sup_{0 \leq u \leq T} \mathbb{W}_2(\mu_u^0; \mu_u^\sigma). \end{aligned}$$

**Step 2.** In order to finalise the first step, we have to show that the distance  $\sup_{0 \leq u \leq T} \mathbb{W}_2(\mu_u^0; \mu_u^\sigma)$  tends to 0 with  $\sigma \rightarrow 0$ . To do so, we use the similar approach as above and consider for coupled random variables  $\tilde{X}^{x,0}$ ,  $\tilde{X}^{x,\sigma}$  the following representation of the processes  $X^0$  and  $X^\sigma$ :

$$\begin{aligned} X_t^0 &= X_0 - \int_0^t \nabla V(X_s^0) ds - \tilde{\mathbb{E}} \left[ \int_0^t \nabla F(X_s^0 - \tilde{X}_s^0) ds \right], \\ X_t^\sigma &= X_0 - \int_0^t \nabla V(X_s^\sigma) ds - \tilde{\mathbb{E}} \left[ \int_0^t \nabla F(X_s^\sigma - \tilde{X}_s^\sigma) ds \right] + \sigma W_t. \end{aligned}$$

Let us define  $\xi(t) = \mathbb{E}|X_t^0 - X_t^\sigma|^2$ . By the Itô formula, we get:

$$\begin{aligned} |X_t^0 - X_t^\sigma|^2 &= -2 \int_0^t \langle X_s^0 - X_s^\sigma; \nabla V(X_s^0) - \nabla V(X_s^\sigma) \rangle ds \\ &\quad - 2\tilde{\mathbb{E}} \left[ \int_0^t \langle X_s^0 - X_s^\sigma; \nabla F(X_s^0 - \tilde{X}_s^0) - \nabla F(X_s^\sigma - \tilde{X}_s^\sigma) \rangle ds \right] \\ &\quad + 2\sigma \int_0^t \langle X_s^0 - X_s^\sigma; dW_s \rangle + d\sigma^2 t. \end{aligned}$$

After taking expectation, we can add and subtract expression of the form  $\nabla F(X_s^\sigma - \tilde{X}_s^0)$  in the second integral above and get the following

representation for  $\xi(t)$ :

$$\begin{aligned} \xi(t) &= d\sigma^2 t - 2\mathbb{E} \int_0^t \langle X_s^0 - X_s^\sigma; \nabla V(X_s^0) - \nabla V(X_s^\sigma) \rangle ds \\ &\quad - 2\mathbb{E} \left[ \tilde{\mathbb{E}} \int_0^t \langle X_s^0 - X_s^\sigma; \nabla F(X_s^0 - \tilde{X}_s^0) - \nabla F(X_s^\sigma - \tilde{X}_s^0) \rangle ds \right] \\ &\quad - 2\mathbb{E} \left[ \tilde{\mathbb{E}} \int_0^t \langle X_s^0 - X_s^\sigma; \nabla F(X_s^\sigma - \tilde{X}_s^\sigma) - \nabla F(X_s^\sigma - \tilde{X}_s^\sigma) \rangle ds \right]. \end{aligned}$$

Use (V-2) of Assumption A-1 for the first expression and assumption (F-5) of A-2 for the second one as well as Cauchy-Schwarz inequality and assumption (F-4) for the third expression. Then, we get the following bound:

$$\begin{aligned} \xi(t) &\leq C\sigma^2 + 2(\theta_1 - \theta_2) \int_0^t \xi(s) ds + 2\mathbb{E} \int_0^t |X_s^0 - X_s^\sigma| \tilde{\mathbb{E}} \left[ |\tilde{X}_s^0 - \tilde{X}_s^\sigma| \right. \\ &\quad \left. \times \left( 1 + |X_s^\sigma - \tilde{X}_s^0|^{2r-1} + |X_s^\sigma - \tilde{X}_s^\sigma|^{2r-1} \right) \right] ds. \end{aligned}$$

From now on, denote by C some generic constant that does not depend on  $\sigma$ . In the following step, we use once again Cauchy-Schwarz inequality, but this time for  $\tilde{\mathbb{E}}$ , bounds of the form  $|x+y|^q \leq C(|x|^q + |y|^q)$ , and, finally, Proposition 4.6 to bound uniformly expectations of the form  $\tilde{\mathbb{E}} \left| \tilde{X}_s^\sigma \right|^q$  by a constant. That gives us:

$$\begin{aligned} \xi(t) &\leq C\sigma^2 + C \int_0^t \xi(s) ds \\ &\quad + C \mathbb{E} \int_0^t |X_s^0 - X_s^\sigma| \sqrt{\tilde{\mathbb{E}} |\tilde{X}_s^0 - \tilde{X}_s^\sigma|^2} \sqrt{C + |X_s^0|^{4r-2}} ds. \end{aligned}$$

We note that  $\tilde{\mathbb{E}} |\tilde{X}_s^0 - \tilde{X}_s^\sigma|^2$  is also equal to  $\xi(s)$ . We use once again Cauchy-Schwarz inequality, this time in terms of  $\mathbb{E}$ , and Proposition 4.6 to bound  $\mathbb{E} |X_s^0|^{4r-2}$  by a constant. Final bound for  $\xi$  is of the form:

$$\xi(t) \leq C\sigma^2 + C \int_0^t \xi(s) ds.$$

We finish the following step by using Grönwall's inequality to bound  $\xi$  by:

$$\xi(t) \leq C\sigma^2 e^{CT},$$

for each  $t \in [0; T]$ , which also proves that:

$$\sup_{0 \leq t \leq T} \mathbb{W}_2(\mu_t^0; \mu_t^\sigma) \leq \sup_{0 \leq t \leq T} \sqrt{\xi(t)} \leq C\sigma \xrightarrow{\sigma \rightarrow 0} 0. \quad (4.17)$$

**Step 3.** Let us come back to Step 1. We can join the results of Steps 1.1 and 1.2, along with (4.17) and show for any  $\omega \in \Omega$  such that  $t \leq \tilde{\tau}_R$ :

$$|Z_t^{x,\sigma} - X_t^{x,\sigma}| \leq C \int_0^t |Z_s^{x,\sigma} - X_s^{x,\sigma}| ds + C\sigma.$$

Therefore, by the Grönwall inequality, for any  $t \leq T \wedge \tilde{\tau}_R$ :

$$|Z_t^{x,\sigma} - X_t^{x,\sigma}| \leq C\sigma e^{CT}$$

and thus for any  $\delta > 0$ :

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}\left(\|Z^{x,\sigma} - X^{x,\sigma}\|_\infty \geq \delta, \tilde{\tau}_R > T\right) = -\infty. \quad (4.18)$$

**Step 4.** Note that for any  $\delta > 0$  and for any  $R > 0$  big enough, the following inequality holds

$$\begin{aligned} & \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}(\|Z^{x,\sigma} - X^{x,\sigma}\|_\infty \geq \delta) \\ & \leq \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}(\|Z^{x,\sigma} - X^{x,\sigma}\|_\infty \geq \delta, \tilde{\tau}_R > T) \\ & \quad + \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}(\tilde{\tau}_R \leq T) \\ & \leq \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}(\tilde{\tau}_R \leq T), \end{aligned} \quad (4.19)$$

by equation (4.18). In order to deal with the last probability, note that, by (4.3), since both  $X^{x,\sigma}$  and  $Z^{x,\sigma}$  stay close enough to each other at least

until  $\tilde{\tau}$ ,  $\{\tilde{\tau}_R \leq T\} \subseteq \{\tau_{R/2}^Z \leq T\}$ , where  $\tau_{R/2}^Z := \inf\{t \geq 0 : Z_t^{x,\sigma} \notin B_{R/2}(0)\}$ . Since  $Z^{x,\sigma}$ , being a family of Itô diffusion processes, satisfies LDP with a good rate function  $I_T^x$ , then for any  $R > 0$  big enough we can say that

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}(\tilde{\tau}_R \leq T) \leq \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}(\tau_{R/2}^Z \leq T) \leq - \inf_{\phi} I_T^x(\phi),$$

where infimum is taken over functions  $\phi \in C([0, T]; \mathbb{R}^d)$  such that  $\phi(0) = x$  and  $\phi(T) \notin B_{R/2}(0)$ . It can be easily shown that, since, by assumptions A-1,  $\liminf_{|x| \rightarrow \infty} V(x) = \infty$ ,  $\inf_{\phi} I_T^x(\phi)$  also tends to  $\infty$  when  $R \rightarrow \infty$ . That finishes the proof of the theorem, since upper bound in (4.19) holds for any  $R > 0$  big enough. □

## Exit-time and exit-location

**Step 1.** To prove the lower bound of Kramers' law, consider the following inequality. For any  $\delta > 0$  and for fixed  $\kappa > 0$  small enough we have

$$\begin{aligned} \mathbb{P} \left( \tau_{\mathcal{D}}^{\sigma} < \exp \left\{ \frac{2(H - \delta)}{\sigma^2} \right\} \right) &\leq \mathbb{P}(\tau_{\mathcal{D}}^{\sigma} < T_{\text{st}}^{\sigma}(\kappa)) \\ &+ \mathbb{P} \left( \tau_{\mathcal{D}}^{\sigma} < \exp \left\{ \frac{2(H - \delta)}{\sigma^2} \right\}, \sup_{t \in [T_{\text{st}}^{\sigma}(\kappa); e^{\frac{2H}{\sigma^2}}]} |X_t^{\sigma} - Y_t^{\sigma}| \leq \kappa \right) \\ &+ \mathbb{P} \left( \sup_{t \in [T_{\text{st}}^{\sigma}(\kappa); e^{\frac{2H}{\sigma^2}}]} |X_t^{\sigma} - Y_t^{\sigma}| > \kappa \right). \end{aligned} \quad (4.20)$$

By the construction of the domain  $\mathcal{D}_{\kappa}^c$  (see Remark 4.5),  $d(\mathcal{D}_{\kappa}^c, \partial \mathcal{D}) \geq \kappa$ . Let us define  $\delta_{\kappa} := H - H_{\kappa}^c$ . Note that  $H_{\kappa}^c \xrightarrow{\kappa \rightarrow 0} H$  due to the continuity of the effective potential  $W_a$ . Therefore, we can choose  $\kappa$  to be small enough such that  $\delta_{\kappa} < \delta$ . Then the following inequality holds:

$$\begin{aligned} &\mathbb{P} \left( \tau_{\mathcal{D}}^{\sigma} < \exp \left\{ \frac{2(H - \delta)}{\sigma^2} \right\}, \sup |X_t^{\sigma} - Y_t^{\sigma}| \leq \kappa \right) \\ &\leq \mathbb{P} \left( \tau_{\mathcal{D}_{\kappa}^c}^{Y,\sigma} > \exp \left\{ \frac{2(H - \delta)}{\sigma^2} \right\} = \exp \left\{ \frac{2(H_{\kappa}^c + \delta_{\kappa} - \delta)}{\sigma^2} \right\} \right) \\ &\leq \mathbb{P}(|X_{T_{\text{st}}^{\sigma}(\kappa)}^{\sigma} - a| > \kappa) + \sup_{y \in B_{\kappa}(a)} \mathbb{P}_y \left( \tau_{\mathcal{D}_{\kappa}^c}^{Y,\sigma} > \exp \left\{ \frac{2(H_{\kappa}^c + \delta_{\kappa} - \delta)}{\sigma^2} \right\} \right) \xrightarrow{\sigma \rightarrow 0} 0, \end{aligned}$$

where the convergence to 0 is due to Lemma 4.11 and Proposition 4.13, since  $\delta_\kappa - \delta < 0$ .

The other probabilities in (4.20) converge to 0 by Corollaries 4.12 and 4.17.

**Step 2.** To prove the upper bound of Kramers' law, consider the set  $\mathcal{D}_\kappa^e$  (see Remark 4.5): enlargement of  $\mathcal{D}$  for small enough  $\kappa > 0$ . Let  $\eta > 0$  be the positive constant defined in Corollary 4.17. Without loss of generality, let us fix positive  $\delta < \eta$ . Consider the following inequalities.

$$\begin{aligned} \mathbb{P} \left( \tau_{\mathcal{D}}^\sigma > \exp \left\{ \frac{2(H + \delta)}{\sigma^2} \right\} \right) &\leq \mathbb{P}(\tau_{\mathcal{D}}^\sigma < T_{\text{st}}^\sigma(\kappa)) \\ &+ \mathbb{P} \left( \tau_{\mathcal{D}}^\sigma > \exp \left\{ \frac{2(H + \delta)}{\sigma^2} \right\}, \sup_{t \in [T_{\text{st}}^\sigma(\kappa); e^{\frac{2(H+\delta)}{\sigma^2}}]} |X_t^\sigma - Y_t^\sigma| \leq \kappa \right) \\ &+ \mathbb{P} \left( \sup_{t \in [T_{\text{st}}^\sigma(\kappa); e^{\frac{2(H+\delta)}{\sigma^2}}]} |X_t^\sigma - Y_t^\sigma| > \kappa \right). \end{aligned} \quad (4.21)$$

If  $\tau_{\mathcal{D}}^\sigma > \exp \left\{ \frac{2(H+\delta)}{\sigma^2} \right\}$  and  $\sup \left\{ |X_t^\sigma - Y_t^\sigma| : t \in [T_{\text{st}}^\sigma(\kappa); e^{\frac{2(H+\delta)}{\sigma^2}}] \right\} \leq \kappa$ , then at the point of time  $e^{\frac{2(H+\delta)}{\sigma^2}}$  the process  $Y^\sigma$  is still inside  $\mathcal{D}_\kappa^e$ . Define  $\delta_\kappa := H_\kappa^e - H$ , decrease  $\kappa$  if necessary such that  $\delta_\kappa < \delta$ , and consider

$$\begin{aligned} &\mathbb{P} \left( \tau_{\mathcal{D}}^\sigma > \exp \left\{ \frac{2(H + \delta)}{\sigma^2} \right\}, \sup_{t \in [T_{\text{st}}^\sigma(\kappa); e^{\frac{2(H+\delta)}{\sigma^2}}]} |X_t^\sigma - Y_t^\sigma| \leq \kappa \right) \\ &\leq \mathbb{P} \left( \tau_{\mathcal{D}_\kappa^e}^{Y, \sigma} > \exp \left\{ \frac{2(H + \delta)}{\sigma^2} \right\} = \exp \left\{ \frac{2(H_\kappa - \delta_\kappa + \delta)}{\sigma^2} \right\} \right) \\ &\leq \mathbb{P}(|X_{T_{\text{st}}^\sigma(\kappa)}^\sigma - a| > \kappa) + \sup_{y \in B_\kappa(a)} \mathbb{P}_y \left( \tau_{\mathcal{D}_\kappa^e}^{Y, \sigma} > \exp \left\{ \frac{2(H_\kappa - \delta_\kappa + \delta)}{\sigma^2} \right\} \right) \xrightarrow{\sigma \rightarrow 0} 0, \end{aligned}$$

where the convergence to 0 holds due to Lemma 4.11 and Proposition 4.13. We finalise the proof of Kramers' type law by observing that, as in Step 1, all the others probabilities in (4.21) also tend to 0 by Corollaries 4.12 and 4.17. That proves Kramers' type law.

**Step 3.** Let us now show the exit-location result. Fix a set  $N \subset \partial\mathcal{D}$  such that  $\inf_{z \in N} \{W_a(z) - W_a(a)\} > H$ . Let us choose  $\xi > 0$  to be small

enough such that  $\xi < (\inf_{z \in N} \{W_a(z) - W_a(a)\} - H)/2$ . Let us define the sublevel set  $L_{H+\xi}^- := \{x \in \mathbb{R}^d : W_a(x) - W_a(a) \leq H + \xi\}$  (without loss of generality by  $L_{H+\xi}^-$  we will denote the unique connected component of the sublevel set that contains  $a$ ). By geometric properties of the effective potential (regularity and convergence at infinity for big  $|x|$ ),  $L_{H+\xi}^-$  satisfies the Assumptions A-5–A-7. Thus, after the initial convergence of  $X^\sigma$  to  $a$  and its law  $\mu_t^\sigma$  to  $\delta_a$ , the Kramers' type law holds for the exit-time  $\tau_{L_{H+\xi}^-}^\sigma$ , that is, for any  $\delta > 0$ ,

$$\lim_{\sigma \rightarrow 0} \mathbb{P} \left( e^{\frac{2(H+\xi-\delta)}{\sigma^2}} \leq \tau_{L_{H+\xi}^-}^\sigma \leq e^{\frac{2(H+\xi+\delta)}{\sigma^2}} \right) = 0, \quad (4.22)$$

including for  $\delta = \xi/2$ . We could easily show geometrically that exiting  $\mathcal{D}$  in the set  $N$  means crossing the boundary  $L_{H+\xi}^- := \partial L_{H+\xi}^-$  before leaving the domain  $\mathcal{D}$ . Therefore, we get the following inequality:

$$\mathbb{P}(X_{\tau_{\mathcal{D}}^\sigma}^\sigma \in N) \leq \mathbb{P}(\tau_{\mathcal{D}}^\sigma \leq T_{\text{st}}^\sigma(\kappa)) + \mathbb{P}(\tau_{L_{H+\xi}^-}^\sigma \leq \tau_{\mathcal{D}}^\sigma).$$

The first probability converges to 0 by Corollary 4.12. Let us look at the second probability:

$$\mathbb{P}(\tau_{L_{H+\xi}^-}^\sigma \leq \tau_{\mathcal{D}}^\sigma) \leq \mathbb{P} \left( \tau_{\mathcal{D}}^\sigma \geq e^{\frac{2(H+\xi/2)}{\sigma^2}} \right) + \mathbb{P} \left( \tau_{L_{H+\xi}^-}^\sigma \leq \tau_{\mathcal{D}}^\sigma < e^{\frac{2(H+\xi/2)}{\sigma^2}} \right) \xrightarrow{\sigma \rightarrow 0} 0,$$

where the first probability tends to 0 by the Kramers' type law (Step 2) and the second probability tends to 0 by (4.22) if we take  $\delta = \xi/2$ .

## Proof of Corollaries 4.9 and 4.10

We consider an unbounded domain  $\mathcal{D}$  with finite exit-cost  $H > 0$ . Then, set  $L_{H+\xi}^- := \{x \in \mathbb{R}^d : W_a(x) - W_a(a) \leq H + \xi\}$ . Let us assume without loss of generality that  $x_{\text{init}} \in L_{H+\xi}^-$  (otherwise, the uniform in  $\sigma$  convergence in finite time inside  $L_{H+\xi}^-$  can be easily proven using LDP, similarly to Lemma 4.11).

Let us define  $\mathcal{D}' := \mathcal{D} \cap L_{H+\xi}^-$ . Immediately,  $\mathcal{D}'$  is bounded. Indeed, since  $W_a(x)$  tends to infinity as  $|x|$  goes to infinity, the level set  $L_{H+\xi}^-$  is compact. The domain  $\mathcal{D}'$  is also stable by  $-\nabla W_a$ , since both the domains  $\mathcal{D}$  and  $L_{H+\xi}^-$  are stable by definition. Thus, the domain  $\mathcal{D}'$  satisfies all the

assumptions of Theorem 4.2 with the height of  $W_a$  inside  $\mathcal{D}'$  being equal to  $H$ . Therefore, for any  $\xi > 0$  we have:

$$\lim_{\sigma \rightarrow 0} \mathbb{P} \left( e^{\frac{2}{\sigma^2}(H-\xi)} \leq \tau'(\sigma) \leq e^{\frac{2}{\sigma^2}(H+\xi)} \right) = 1,$$

where,  $\tau'(\sigma)$  is the first exit-time of  $X^\sigma$  from  $\mathcal{D}'$ . Indeed, the exit-cost is  $H$ .

Note that, by construction of the domain  $\mathcal{D}'$ , and by continuity of  $W_a$ , for any  $\xi > 0$  we have

$$\inf \{ W_a(z) - W_a(a) : z \in \text{Cl}(\partial\mathcal{D}' \setminus \partial\mathcal{D}) \} > H,$$

where Cl stands for closure. It means that the exit-location result of the main Theorem 4.2 holds for  $N = \text{Cl}(\partial\mathcal{D}' \setminus \partial\mathcal{D})$ , namely

$$\lim_{\sigma \rightarrow 0} \mathbb{P} (X_{\tau'(\sigma)}^\sigma \in \text{Cl}(\partial\mathcal{D}' \setminus \partial\mathcal{D})) = 0.$$

That essentially means that

$$\lim_{\sigma \rightarrow 0} \mathbb{P}(\tau'(\sigma) = \tau_{\mathcal{D}}^\sigma) = 1,$$

which proves Corollary 4.9.

The second corollary can be proved the same way by choosing  $\xi > 0$  to be small enough such that the set under consideration  $N \subset \mathcal{D}$  lies entirely beyond the level set  $L_{H+\xi}^-$ .

## 4.4 Proofs of the intermediate results

### Stabilisation in finite time: Proof of Lemma 4.11 and Corollary 4.12

The proof is based on LDP ideas and the fact that, for small  $\sigma$ , the process  $X^\sigma$  is attracted towards  $a$ . Fix some  $\kappa > 0$ . By Assumption A-6, the path of the deterministic solution to the following equation

$$\frac{d}{dt} \gamma_t = -\nabla V(\gamma_t), \quad \text{with } \gamma_0 = x_{\text{init}}, \quad (4.23)$$

is contained in  $\mathcal{D}$ , i.e.  $\{\gamma_t, t \geq 0\} \subset \mathcal{D}$ , and tends to  $a$ . Let us decrease  $\kappa > 0$  to be small enough such that the distance between the set  $(\gamma_t, t \geq 0)$  and  $\partial\mathcal{D}$  is strictly greater than  $\kappa/3$ . Let us define  $\bar{T}_{\text{st}}(\kappa)$  as the first time when  $\gamma_t \in B_{\kappa/3}(a)$ . The following inclusion of events takes place:

$$\mathbb{P}\left(\left|X_{\bar{T}_{\text{st}}(\kappa)}^\sigma - a\right| > \frac{2\kappa}{3}\right) \leq \mathbb{P}\left(\left|X_{\bar{T}_{\text{st}}(\kappa)}^\sigma - \gamma_{\bar{T}_{\text{st}}(\kappa)}\right| > \frac{\kappa}{3}\right) \leq \mathbb{P}(X^\sigma \in \Phi),$$

where  $\Phi := \{\varphi \in C([0; \bar{T}_{\text{st}}(\kappa)]; \mathbb{R}^d) : \|\varphi - \gamma\|_\infty \geq \kappa/3\}$ . By Corollary 4.8,

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}(X^\sigma \in \Phi) \leq - \inf_{\varphi \in \Phi} I_{\bar{T}_{\text{st}}(\kappa)}(\varphi). \quad (4.24)$$

Note that, by definition of the rate function  $I_T$ , and by uniqueness of solution to equation (4.23), function  $\gamma$  is its only minimizer such that  $I_{\bar{T}_{\text{st}}(\kappa)}(\gamma) = 0$ . Since  $I_T$  is a good rate function, its infima are achieved over closed sets. Note that  $\gamma \notin \Phi$ , thus  $A := I_{\bar{T}_{\text{st}}(\kappa)}(\varphi) > 0$ . That proves the second result of the Lemma 4.11, since it guarantees that there exists  $\sigma_\kappa > 0$  small enough such that for any  $0 < \sigma < \sigma_\kappa$ :

$$\mathbb{P}\left(\left|X_{\bar{T}_{\text{st}}(\kappa)}^\sigma - a\right| > \frac{2\kappa}{3}\right) \leq e^{-\frac{2A}{\sigma^2}}. \quad (4.25)$$

For the first statement, consider the following equality:

$$\begin{aligned} \mathbb{W}_2^2\left(\mu_{\bar{T}_{\text{st}}(\kappa)}^\sigma; \delta_a\right) &= \mathbb{E}\left|X_{\bar{T}_{\text{st}}(\kappa)}^\sigma - a\right|^2 = \mathbb{E}\left[\left|X_{\bar{T}_{\text{st}}(\kappa)}^\sigma - a\right|^2 \mathbf{1}_{\{X_{\bar{T}_{\text{st}}(\kappa)}^\sigma \in B_{\frac{2\kappa}{3}}(a)\}}\right] \\ &\quad + \mathbb{E}\left[\left|X_{\bar{T}_{\text{st}}(\kappa)}^\sigma - a\right|^2 \mathbf{1}_{\{X_{\bar{T}_{\text{st}}(\kappa)}^\sigma \notin B_{\frac{2\kappa}{3}}(a)\}}\right]. \end{aligned}$$

Therefore, by Cauchy–Schwarz inequality, we can bound the difference between the two measures by:

$$\mathbb{W}_2^2\left(\mu_{\bar{T}_{\text{st}}(\kappa)}^\sigma; \delta_a\right) \leq \frac{4\kappa^2}{9} + \sqrt{\mathbb{E}\left|X_{\bar{T}_{\text{st}}(\kappa)}^\sigma - a\right|^4} \sqrt{\mathbb{P}\left(\left|X_{\bar{T}_{\text{st}}(\kappa)}^\sigma - a\right| > \frac{2\kappa}{3}\right)}.$$

By Proposition 4.6, there exists  $M > 0$  such that  $\sup_{0 < \sigma < 1} \sup_{t \geq 0} \mathbb{E}|X_t^\sigma - a|^2 < M^2$ . This estimate along with equation (4.25) gives us:

$$\mathbb{W}_2^2\left(\mu_{\bar{T}_{\text{st}}(\kappa)}^\sigma; \delta_a\right) \leq \frac{4\kappa^2}{9} + Me^{-A/\sigma^2}.$$



That expression can be bounded by  $\kappa^2$  if we choose  $\sigma_\kappa > 0$  to be small enough, which proves Lemma 4.11.

Corollary 4.12 can be also easily proven by choosing  $\kappa$  such that

$$\inf_{t \geq 0} \inf_{z \in \partial \mathcal{D}} |\gamma_t - z| > \frac{\kappa}{3}.$$

In this case, the following estimate holds:

$$\mathbb{P}(\tau_{\mathcal{D}}^\sigma \leq \bar{T}_{\text{st}}(\kappa)) \leq \mathbb{P}(X \notin \Phi) \leq e^{-\frac{2A}{\sigma^2}} \xrightarrow{\sigma \rightarrow 0} 0.$$

### The coupling estimate: Proof of Proposition 4.14

In this section we prove Proposition 4.14. The idea of the proof is based on the fact that, since the processes  $X^\sigma$  and  $Y^\sigma$  are coupled by the same Brownian motion and by the properties of convex sets, whenever both  $X^\sigma$  and  $Y^\sigma$  belong to the set  $B_\rho(a)$  (Definition 4.2), the distance between them decreases a.s. (we show this in Lemma 4.19). At the same time, whenever the two processes belong to the region  $\mathcal{D} \setminus B_\rho(a)$ , their maximum scatter can be controlled in terms of the time spent inside  $\mathcal{D} \setminus B_\rho(a)$  (Lemma 4.20 below). The proof is finished by observing that, before exiting  $\mathcal{D}$ , the processes  $X^\sigma$  and  $Y^\sigma$  spend inside  $B_\rho(a)$  long enough time comparing to the total time spent inside  $\mathcal{D} \setminus B_\rho(a)$ , that the attracting effect surpasses the scattering one.

Before proving the proposition rigorously, let us present the following notions. Let us decrease without loss of generality  $\kappa > 0$  to be smaller than  $\rho/4$ . Let us also fix some enlargement of the domain  $\mathcal{D}$  of some radius  $R > 0$ :  $\mathcal{D}_R^e$  (see Remark 4.5 for the definition). Decrease  $\kappa$ , if necessary, so that  $\kappa < R/2$ . Consider the following sequence of stopping times:

$$\begin{aligned} \theta_1 &:= \inf\{t \geq T_{\text{st}}^\sigma(\kappa) : Y_t^\sigma \notin B_{\rho/2}(a)\}, \\ \tau_m &:= \inf\{t \geq \theta_m : Y_t^\sigma \in B_{\rho/4}(a) \cup \partial \mathcal{D}_R^e\}, \\ \theta_{m+1} &:= \inf\{t \geq \tau_m : Y_t^\sigma \notin B_{\rho/2}(a)\}. \end{aligned} \tag{4.26}$$

We also define the following stopping times that will allow us to study the behaviour of  $\theta_i, \tau_i$  for different  $i$  using the strong Markov property of

diffusion  $Y^\sigma$ . For any  $y \in \mathbb{R}^d$  consider:

$$\begin{aligned}\theta_0 &:= \inf\{t \geq 0 : Y_t^{y,\sigma} \notin B_{\rho/2}(a)\}, \\ \tau_0 &:= \inf\{t \geq 0 : Y_t^{y,\sigma} \in B_{\rho/4}(a) \cup \partial\mathcal{D}_R^e\}.\end{aligned}\tag{4.27}$$

Consider the following

**Lemma 4.19.** *Define for some  $K > 0$  the following family of mappings  $\varphi_T : x \mapsto xe^{-KT} + o_\kappa(1)$  for any  $T > 0$ , where  $o_\kappa(1) \xrightarrow{\kappa \rightarrow 0} 0$ . Then, under Assumptions A-1-A-7, there exists a constant  $K > 0$  such that for any  $\alpha < \rho/4$ , for any  $m \geq 1$ , and for any  $\kappa > 0$  small enough:*

$$\mathbb{P}\left(\sup_{t \in [\tau_m; \theta_{m+1}]} |X_t^\sigma - Y_t^\sigma| > \varphi_{\theta_{m+1} - \tau_m}(\alpha), A\right) = 0,$$

where  $A := \{\theta_{m+1} \leq S_{\text{st}}^\sigma(\kappa), \sup_{t \leq \tau_m} |X_t^\sigma - Y_t^\sigma| \leq \alpha\}$

*Proof.* Let us define random time  $\mathcal{T} := \inf\{t \geq \tau_m : X_t^\sigma \notin B_\rho(a)\}$  – first time when  $X^\sigma$  leaves the convexity area  $B_\rho(a)$ . Obviously, for almost every  $\omega \in A$ , we have  $\mathcal{T} > 0$ .

**Step 1.** Let us define  $\xi(t) := |X_t^\sigma - Y_t^\sigma|^2$ . The way functions  $X^\sigma$  and  $Y^\sigma$  are coupled provides us with the fact that  $\xi$  is differentiable in the usual sense. Its derivative is equal to:

$$\begin{aligned}\xi'(t) &= -2\langle X_t^\sigma - Y_t^\sigma; \nabla W_a(X_t^\sigma) - \nabla W_a(Y_t^\sigma) \rangle \\ &\quad - 2\langle X_t^\sigma - Y_t^\sigma; \nabla F * \mu_t^\sigma(X_t^\sigma) - \nabla F * \delta_a(X_t^\sigma) \rangle.\end{aligned}$$

Since in this lemma we consider only outcomes such that  $\mathbb{W}_2(\mu_t^\sigma; \delta_a) \leq \kappa$  and  $|X_{\tau_m}^\sigma - Y_{\tau_m}^\sigma| \leq \alpha$ , i.e.  $\omega \in A$ , after integrating over the time interval  $[\tau_m; \theta_{m+1} \wedge \mathcal{T}]$  and applying Assumption A-4 (see also Definition 4.2), we get the following estimate. For any  $t > 0$  and for  $\mathbb{P}$ -a.e.  $\omega \in A \cap \{t \in$

$[\tau_m; \theta_{m+1} \wedge \mathcal{T}]$ :

$$\begin{aligned}
\xi(t) &\leq |X_{\tau_m}^\sigma - Y_{\tau_m}^\sigma|^2 - 2 \int_{\tau_m}^t \langle X_s^\sigma - Y_s^\sigma; \nabla W_a(X_s^\sigma) - \nabla W_a(Y_s^\sigma) \rangle ds \\
&\quad + 2 \int_{\tau_m}^t |X_s^\sigma - Y_s^\sigma| |\nabla F * \mu_s^\sigma(X_s^\sigma) - \nabla F * \delta_a(X_s^\sigma)| ds \\
&\leq \alpha^2 - 2C_W \int_{\tau_m}^t \xi(s) ds \\
&\quad + 2 \int_{\tau_m}^t \sqrt{\xi(s)} \left| \nabla F * \mu_s^\sigma(X_s^\sigma) - \nabla F * \delta_a(X_s^\sigma) \right| ds.
\end{aligned} \tag{4.28}$$

Since the term  $\left| \nabla F * \mu_s^\sigma(X_s^\sigma) - \nabla F * \delta_a(X_s^\sigma) \right|$  is hard to analyse, we study it separately.

**Step 2.** Consider the following inequality. By Assumption (F – 4) of A-2, we can express:

$$\begin{aligned}
&\int_{\mathbb{R}^d} \left| \nabla F(X_s^\sigma - z) - \nabla F(X_s^\sigma - a) \right| \mu_s^\sigma(dz) \\
&\leq C' \int_{\mathbb{R}^d} |z - a| (1 + |X_s^\sigma - z|^{2r-1} + |X_s^\sigma - a|^{2r-1}) \mu_s^\sigma(dz) \\
&\leq C' \int_{\mathbb{R}^d} |z - a| (1 + 2^{2r-1} |X_s^\sigma|^{2r-1} + 2^{2r-2} |z|^{2r-1} + 2^{2r-2} |a|^{2r-1}) \mu_s^\sigma(dz).
\end{aligned}$$

In the following, we will denote by C the generic constant that may depend on  $r$ ,  $\rho$  and other parameters defined in assumptions. The bound thus takes the form:

$$\begin{aligned}
&C \int_{\mathbb{R}^d} |z - a| (C + C |X_s^\sigma|^{2r-1} + |z|^{2r-1} + |a|^{2r-1}) \mu_s^\sigma(dz) \\
&\leq C \sqrt{\int_{\mathbb{R}^d} |z - a|^2 \mu_s^\sigma(dz)} \sqrt{C + C |X_s^\sigma|^{4r-2} + |a|^{4r-2} + \int_{\mathbb{R}^d} |z|^{4r-2} \mu_s^\sigma(ds)}
\end{aligned}$$

Since we only consider  $\omega \in A \cap \{t \in [\tau_m; \theta_{m+1} \wedge \mathcal{T}]\}$ ,  $X^\sigma$  belongs to  $B_\rho(a)$  and is thus bounded by a constant. Moreover,  $\mathbb{W}_2(\mu_s^\sigma; \delta_a) \leq \kappa$  and

$|Y_t^\sigma| \leq \sup_{z \in \partial \mathcal{D}_R^e} |z - a|$  by the definition of the set  $A$ . At the same time, by Proposition 4.6, we know that  $\int |z|^{4r-2} d\mu_s^\sigma \leq M$  for any time  $t \geq 0$  and for any  $0 \leq \sigma \leq 1$ . Therefore, for any  $t > 0$  and for any  $\omega \in A \cap \{t \in [\tau_m; \theta_{m+1} \wedge \mathcal{T}]\}$  we have

$$\int_{\mathbb{R}^d} \left| \nabla F(X_s^\sigma - z) - \nabla F(X_s^\sigma - a) \right| \mu_s^\sigma(dz) \leq C\kappa.$$

**Step 3.** Let us come back to equation (4.28). Given the calculations in Step 2, the final bound takes the following form:

$$\xi(t) \leq \alpha^2 - 2C_W \int_{\tau_m}^t \xi(s) ds + 2\kappa C \int_{\tau_m}^t \sqrt{\xi(s)} ds.$$

It means that, if we introduce the deterministic function  $\psi$  that is the unique solution of equation

$$\psi(u) = \alpha^2 - 2C_W \int_0^u \psi(s) ds + 2\kappa C \int_0^u \sqrt{\psi(s)} ds,$$

then  $\xi_{\tau_m+u} \leq \psi_u$  for any positive  $u \leq t$  and for  $\mathbb{P}$ -a.e. point  $\omega \in A \cap \{t \in [\tau_m; \theta_{m+1} \wedge \mathcal{T}]\}$ .

If  $\alpha > \frac{C\kappa}{2C_W}$ , we can solve this equation explicitly and get:

$$\sqrt{\psi(u)} = \left( \alpha - \frac{C}{2C_W} \kappa \right) e^{-2C_W u} + \frac{C}{2C_W} \kappa. \quad (4.29)$$

Otherwise, we can simply bound  $\psi(u)$  by

$$\psi(u) \leq \frac{C}{4C_W^2} \kappa^2, \quad (4.30)$$

since  $\psi'(u) < 0$  whenever  $\psi(u) > C\kappa^2/(4C_W^2)$ . Thus,  $\psi$  can be expressed in the form:

$$\sqrt{\psi(u)} \leq \alpha e^{-2C_W u} + o_\kappa(1).$$

In particular, it means that if there is some random time  $\mathcal{S}$  defined for  $\omega \in A$  and such that for  $\mathbb{P}$ -a.e.  $\omega \in A$  we have  $\tau_m \leq \mathcal{S} \leq \theta_{m+1} \wedge \mathcal{T}$ , then:

$$\sqrt{\xi(\mathcal{S})} \leq \alpha e^{-2C_W(\mathcal{S}-\tau_m)} + o_\kappa(1)$$

for  $\mathbb{P}$ -a.e.  $\omega \in A$ .

**Step 4.** To finalise the proof, let us show that for  $\mathbb{P}$ -a.e.  $\omega \in A$ , we have  $\mathcal{T} > \theta_{m+1}$ . Indeed, if it is not true, then there exists a set  $B \subseteq A$  with  $\mathbb{P}(B) > 0$ , such that for any  $\omega \in B$ ,  $X_{\mathcal{T}}^\sigma \notin B_\rho(a)$ , but  $Y_{\mathcal{T}}^\sigma \in B_{\rho/2}(a)$ . Yet, by derivations of Step 3, for  $\mathbb{P}$ -a.e.  $\omega \in A$ :

$$|X_{\mathcal{T}}^\sigma - Y_{\mathcal{T}}^\sigma| \leq \max\left(\alpha; \frac{C}{2C_W}\kappa\right).$$

Therefore, without loss of generality, we can choose  $\kappa > 0$  to be small enough to get the contradiction. That proves the lemma.  $\square$

For control outside of the set  $B_\rho(a)$  consider the following lemma.

**Lemma 4.20.** *Define for some constant  $L > 0$  and for any  $T > 0$  the following mapping:  $\psi_T : x \mapsto xe^{LT}$ . Then, under Assumptions A-1–A-7, there exists a constant  $L > 0$  such that for  $\alpha < \rho/4$ , for any  $m \geq 1$ , and for any  $\kappa > 0$  small enough:*

$$\mathbb{P}\left(\sup_{t \in [\theta_m; \tau_m]} |X_t^\sigma - Y_t^\sigma| > \psi_{\tau_m - \theta_m}(\alpha), A\right) = 0,$$

where  $A := \{\tau_m \leq S_{\text{st}}^\sigma(\kappa), \sup_{t \leq \theta_m} |X_t^\sigma - Y_t^\sigma| \leq \alpha\}$

*Proof.* As in the proof of Lemma 4.19, we first introduce  $\xi(t) = |X_t^\sigma - Y_t^\sigma|^2$  and then differentiate this function with respect to time. The difference is that now we can not use convexity properties of the set  $B_\rho(a)$ . Moreover, we will not be able to provide a good upper bound for  $|X_t^\sigma|$ , since  $Y^\sigma$  and  $X^\sigma$  drift apart from each other.

**Step 1.** The following inequality holds for  $\mathbb{P}$ -a.e.  $\omega \in A \wedge \{t \in [\theta_m; \tau_m]\}$ :

$$\begin{aligned} \xi(t) &\leq |X_{\theta_m}^\sigma - Y_{\theta_m}^\sigma|^2 - 2 \int_{\theta_m}^t \langle X_s^\sigma - Y_s^\sigma; \nabla V(X_s^\sigma) - \nabla V(Y_s^\sigma) \rangle ds \\ &\quad - 2 \int_{\theta_m}^t \langle X_s^\sigma - Y_s^\sigma; \nabla F * \mu_s^\sigma(X_s^\sigma) - \nabla F(Y_s^\sigma - a) \rangle ds. \end{aligned}$$

Using Cauchy–Schwarz inequality, we can obtain the following bound:

$$\begin{aligned} \xi(t) &\leq \alpha^2 + 2 \int_{\theta_m}^t \sqrt{\xi(s)} |\nabla V(X_s^\sigma) - \nabla V(Y_s^\sigma)| ds \\ &\quad + 2 \int_{\theta_m}^t \sqrt{\xi(s)} |\nabla F * \mu_s^\sigma(X_s^\sigma) - \nabla F(Y_s^\sigma - a)| ds =: \alpha^2 + I_1 + I_2. \end{aligned}$$

Let us consider  $I_1$  and  $I_2$  separately. In the following,  $C$  will denote a generic constant that may depend on parameters defined in the assumptions.

**Step 2.** For the first expression  $I_1$ , we use Assumption (V – 5) of A-1 and get:

$$2 \int_{\theta_m}^t \sqrt{\xi(s)} |\nabla V(X_s^\sigma) - \nabla V(Y_s^\sigma)| ds \leq C \int_{\theta_m}^t \xi(s) (1 + |X_s^\sigma|^{2r-1} + |Y_s^\sigma|^{2r-1}) ds.$$

By adding and subtracting  $Y_s^\sigma$  in the expression above, we can upper bound it by

$$C \int_{\theta_m}^t \xi(s) \left( C + C\xi(s)^{\frac{2r-1}{2}} + |Y_s^\sigma|^{2r-1} \right) ds.$$

Moreover, since we consider only those  $\omega$  for which  $t \leq \tau_m$ ,  $Y_s^\sigma$  belongs to  $\mathcal{D}_R^e$ , which is a bounded set. Therefore, the upper bound takes the final form:

$$I_1 \leq C \int_{\theta_m}^t \xi(s) \left( C + \xi(s)^{\frac{2r-1}{2}} \right) ds. \quad (4.31)$$

**Step 3.** For the second expression  $I_2$ , let us use assumption (F – 4) of A-2 and get:

$$\begin{aligned} I_2 &\leq C \int_{\theta_m}^t \sqrt{\xi(s)} \int_{\mathbb{R}^d} |X_s^\sigma - z - Y_s^\sigma + a| \\ &\quad \times (1 + |X_s^\sigma - z|^{2r-1} + |Y_s^\sigma - a|^{2r-1}) \mu_s^\sigma(dz) ds \\ &\leq C \int_{\theta_m}^t \int_{\mathbb{R}^d} \xi(s) (1 + |X_s^\sigma - z|^{2r-1} + |Y_s^\sigma - a|^{2r-1}) \mu_s^\sigma(dz) ds \\ &\quad + C \int_{\theta_m}^t \int_{\mathbb{R}^d} \sqrt{\xi(s)} |z - a| (1 + |X_s^\sigma - z|^{2r-1} + |Y_s^\sigma - a|^{2r-1}) \mu_s^\sigma(dz) ds. \end{aligned}$$

Let us denote the two expressions above as  $A_1$  and  $A_2$ . For  $A_1$ , we add and subtract  $Y_s^\sigma$  inside  $|X_t^\sigma - z|^{2r-1}$  and get:

$$A_1 \leq C \int_{\theta_m}^t \int_{\mathbb{R}^d} \xi(s) \left( C + C\xi(s)^{\frac{2r-1}{2}} + C|Y_s^\sigma|^{2r-1} + |z|^{2r-1} + |a|^{2r-1} \right) \mu_s^\sigma(dz) ds.$$

As was pointed out above, since  $t \in [\theta_m; \tau_m]$ ,  $|Y_s^\sigma - a|$  is bounded for  $\mathbb{P}$ -a.e.  $\omega \in A \wedge \{t \in [\theta_m; \tau_m]\}$ . Moreover, by Proposition 4.6, there exists  $M > 0$  such that  $\int |z|^{2r-1} d\mu_s^\sigma < M$ . Thus:

$$A_1 \leq C \int_{\theta_m}^t \xi(s) \left( C + \xi(s)^{\frac{2r-1}{2}} \right) ds.$$

Similarly, for  $A_2$ :

$$A_2 \leq C \int_{\theta_m}^t \int_{\mathbb{R}^d} \sqrt{\xi(s)} |z - a| \times \left( C + C\xi(s)^{\frac{2r-1}{2}} + C|Y_s^\sigma|^{2r-1} + |z|^{2r-1} + |a|^{2r-1} \right) \mu_s^\sigma(dz) ds.$$

By Cauchy–Schwarz inequality and since both  $\int_{\mathbb{R}^d} |z| d\mu_s^\sigma$  and  $\int_{\mathbb{R}^d} |z|^{4r-2} d\mu_s^\sigma$  are bounded by a constant, we get:

$$A_2 \leq C \int_{\theta_m}^t \sqrt{\xi(s)} \sqrt{C + \xi(s)^{2r-1}} ds,$$

which gives the following bound for  $I_2$ :

$$I_2 \leq C \int_{\theta_m}^t \left( C\xi(s) \left( C + \xi(s)^{\frac{2r-1}{2}} \right) + \sqrt{\xi(s)} \sqrt{C + \xi(s)^{2r-1}} \right) ds.$$

Since for any  $\alpha > 1$  we have  $x^\alpha \leq \sqrt{x} + x^{\alpha+1}$  and since  $\sqrt{x} \leq 1 + x$ , we can roughly bound  $I_2$  by the following expression:

$$I_2 \leq C \int_{\theta_m}^t \left( \xi(s)^{r+1} + \sqrt{\xi(s)} \right) ds \quad (4.32)$$

**Step 4.** From (4.31) and (4.32) we get that for  $\mathbb{P}$ -a.e.  $\omega \in A \wedge \{t \in [\theta_m; \tau_m]\}$ :

$$\xi(t) \leq \alpha^2 + C \int_{\theta_m}^t \left( C\xi(s)^{r+1} + \sqrt{\xi(s)} \right) ds.$$

Obviously,  $\xi(t)$  is bounded for respective  $\omega$  and  $t$  by a function of the form:

$$\psi(u) = \alpha^2 + C \int_0^u \left( C\psi(s)^{r+1} + \sqrt{\psi(s)} \right) ds,$$

which in its term is bounded by the following expression. Note that for each period of time when  $\psi(u) \leq 1$ , it is simply bounded by a linear function:

$$\psi(u) \leq \alpha^2 + Cu.$$

Otherwise, its upper bound take the form:

$$\psi(u) \leq \alpha^2 + C \int_0^u \psi(s)^{r+1} ds,$$

which is a polynomial. By choosing the right constant  $L > 0$ , we can easily bound  $\psi$  by

$$\psi(u) \leq \alpha^2 e^{2Lu},$$

which proves the Lemma by using the same approach as in Steps 3 and 4 of the proof of Lemma 4.19.  $\square$

The following lemma establishes the maximum number of excursions of the process  $Y^\sigma$  from  $B_\rho(a)$ . Let us define the height of the effective potential inside the sets of the form  $B_{\rho/2}(a)$  as  $Q^c := \inf_{z \in S_{\rho/2}(a)} \{W_a(z) - W_a(a)\}$ . We remind that  $H_R^e := \inf_{z \in \partial \mathcal{D}_R^e} \{W_a(z) - W_a(a)\}$  is the height of the effective potential inside the set  $\mathcal{D}_R^e$ . Consider the following lemma:

**Lemma 4.21.** *Let  $N^* := 2 \lceil \exp\{\frac{2}{\sigma^2}(H_R^e - Q^c + \kappa)\} \rceil$ . Let  $\tau_{N^*}$  be defined as in (4.26). Then, for any  $\kappa > 0$  small enough:*

1.  $\mathbb{P}(\tau_{\mathcal{D}_R^e}^{Y, \sigma} > \tau_{N^*}) \xrightarrow{\sigma \rightarrow 0} 0$ .
2. *There exists  $T_1 > 0$  such that  $\mathbb{P}(\exists i \leq N^* : \theta_i - \tau_{i-1} > T_1) \xrightarrow{\sigma \rightarrow 0} 0$ .*

*Proof.* We separate the proof into 2 steps.

**Step 1.** Let us prove the first part of the lemma. Note, that if  $\tau_{N^*}$  is less or equal then  $\exp\{\frac{2}{\sigma^2}(H_R^e + \frac{\kappa}{2})\}$ , then necessarily the number of intervals



of the form  $[\tau_{i-1}; \theta_i]$  such that  $\theta_i - \tau_{i-1} \geq \exp\{\frac{2}{\sigma^2}(Q^c - \frac{\kappa}{2})\}$  can not exceed  $N^*/2$  by definition of the latter. Based on this observation and using Proposition 4.13, we have

$$\begin{aligned} \mathbb{P}(\tau_{\mathcal{D}_R^e}^{Y,\sigma} > \tau_{N^*}) &\leq \mathbb{P}\left(\tau_{N^*} < \tau_{\mathcal{D}_R^e}^{Y,\sigma} < e^{\frac{2}{\sigma^2}(H_R^e + \frac{\kappa}{2})}\right) \\ &\quad + \mathbb{P}\left(\tau_{\mathcal{D}_R^e}^{Y,\sigma} \geq e^{\frac{2}{\sigma^2}(H_R^e + \frac{\kappa}{2})}\right) \\ &\leq \mathbb{P}\left(\#\left\{i \leq N^* : \theta_i - \tau_{i-1} < e^{\frac{2}{\sigma^2}(Q^c - \frac{\kappa}{2})}\right\} > \frac{N^*}{2}\right) + o_\sigma(1), \end{aligned} \quad (4.33)$$

where  $o_\sigma(1)$  is an infinitesimal with respect to  $\sigma$ . Consider

$$\begin{aligned} &\mathbb{P}\left(\#\left\{i \leq N^* : \theta_i - \tau_{i-1} < e^{\frac{2}{\sigma^2}(Q^c - \frac{\kappa}{2})}\right\} > \frac{N^*}{2}\right) \\ &\leq \sum_{k=\lceil \frac{N^*}{2} \rceil}^{N^*} \sum_{(i_1, \dots, i_k)} \mathbb{P}\left(\bigcap_{j=1}^k \left\{\theta_{i_j} - \tau_{i_j-1} < e^{\frac{2}{\sigma^2}(Q^c - \frac{\kappa}{2})}\right\}\right) \\ &\leq \sum_{k=\lceil \frac{N^*}{2} \rceil}^{N^*} 2^{N^*} \left(\sup_{y \in B_{\rho/4}(a)} \mathbb{P}_y\left(\theta_0 < e^{\frac{2}{\sigma^2}(Q^c - \frac{\kappa}{2})}\right)\right)^k, \end{aligned} \quad (4.34)$$

where  $(i_1, \dots, i_k)$  stands for all possible choices of  $k$  numbers  $i_1 < \dots < i_k$  from the set  $\{1, \dots, N^*\}$ . Note that the number of such combinations can be roughly bounded by  $2^{N^*}$ . The last inequality in (4.34) we get due to the fact that  $Y^\sigma$  is a strong Markov process and  $\theta_0$  is defined in (4.27). By the exit-time result for diffusions of type  $Y^\sigma$  (see Proposition 4.13), for any  $\kappa < \rho/4$ :

$$\sup_{y \in B_{\rho/4}(a)} \mathbb{P}_y\left(\theta_0 < e^{\frac{2}{\sigma^2}(Q^c - \frac{\kappa}{2})}\right) = o_\sigma(1).$$

After adding this bound to equations (4.34) and (4.33), we get:

$$\mathbb{P}(\tau_{\mathcal{D}_R^e}^{Y,\sigma} > \tau_{N^*}) \leq 2^{N^*} o_\sigma(1)^{\lceil \frac{N^*}{2} \rceil} \frac{1 - o_\sigma(1)^{\lceil \frac{N^*}{2} \rceil}}{1 - o_\sigma} + o_\sigma(1) = o_\sigma(1).$$

**Step 2.** For the second part of the lemma, we use [DZ10, Lemma 5.7.19], that is the fact that there exists  $T_1 > 0$  big enough such that

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \sup_{y \in B_{\rho/2}(a)} \mathbb{P}_y(\tau_0 > T_1) < -(H_R^e - Q^c + 1). \quad (4.35)$$

Consider the following equations:

$$\begin{aligned} \mathbb{P}(\exists i \leq N^* : \tau_i - \theta_i > T_1) &\leq \sum_{i=1}^{N^*} \mathbb{P}(\tau_i - \theta_i > T_1) \\ &\leq N^* \sup_{y \in B_{\rho/2}(a)} \mathbb{P}_y(\tau_0 > T_1), \end{aligned}$$

where the last inequality is due to the Markov property of the diffusion  $Y^\sigma$ . Finally, by (4.35), we get:

$$\begin{aligned} \mathbb{P}(\exists i \leq N^* : \tau_i - \theta_i > T_1) &\leq 2 \left( e^{2(H_R^e - Q^c + \kappa)/\sigma^2} + 1 \right) e^{-2(H_R^e - Q^c + 1)/\sigma^2} \\ &\xrightarrow{\sigma \rightarrow 0} 0, \end{aligned}$$

which proves the lemma if  $\kappa$  is chosen to be small enough.  $\square$

Now we are ready to prove Proposition 4.14.

*Proof of Proposition 4.14.* Since, by Lemma 4.21, each time spent outside of  $B_{\rho/2}(a)$  is bounded by a constant  $T_1 > 0$  with high probability, we are interested in the composition

$$\psi_{T_1} \circ \phi_t(x) = x e^{(LT_1 - Kt)} + o_\kappa(1) e^{LT_1} \leq x e^{(LT_1 - Kt)} + o_\kappa(1). \quad (4.36)$$

Let us introduce the following mapping:

$$\Psi_t(x) := x e^{LT_1 - Kt} + o_\kappa(1).$$

Then the results of Lemmas 4.19 and 4.20 can be rewritten in the following form: for any  $\kappa > 0$  small enough, for any  $m \geq 1$  and for any  $\alpha < \kappa$ :

$$\begin{aligned} \mathbb{P} \left( \sup_{t \in [\tau_m; \tau_{m+1}]} |X_t^\sigma - Y_t^\sigma| > \Psi_{\theta_{m+1} - \tau_m}(\alpha); \right. \\ \left. \tau_{m+1} \leq S_{\text{st}}^\sigma(\kappa), \sup_{t \leq \tau_m} |X_t^\sigma - Y_t^\sigma| \leq \alpha \right) = 0. \end{aligned} \quad (4.37)$$

Let us now come back to the statement of the proposition. Fix some  $0 < \eta < H_R^e - H$ . Note that, if  $\sup_t |X_t^\sigma - Y_t^\sigma| > \alpha$ , for  $t \in [T_{\text{st}}^\sigma(\kappa); S_{\text{st}}^\sigma(\kappa) \wedge e^{\frac{2}{\sigma^2}(H+\eta)}]$ , then it should happen for  $t$  belonging to one of the periods of time of the form  $[\tau_{k-1}; \tau_k]$  that are before  $S_{\text{st}}^\sigma(\kappa) \wedge e^{\frac{2}{\sigma^2}(H+\eta)}$ . Moreover, since we know, by Lemma 4.21, that  $\tau_{N^*}$  happens after  $S_{\text{st}}^\sigma(\kappa) \wedge e^{\frac{2}{\sigma^2}(H+\eta)}$  with high probability, the number of periods of the form  $[\tau_{k-1}; \tau_k]$ , during which  $|X_t^\sigma - Y_t^\sigma|$  can surpass the level  $\alpha$ , is bounded by  $N^*$ . Given these observations, consider the following line of equations:

$$\begin{aligned}
& \mathbb{P} \left( \sup \left\{ |X_t^\sigma - Y_t^\sigma| : t \in \left[ T_{\text{st}}^\sigma(\kappa); S_{\text{st}}^\sigma(\kappa) \wedge \exp \left\{ \frac{2}{\sigma^2}(H + \eta) \right\} \right] \right\} > \alpha \right) \\
& \leq \mathbb{P} \left( \tau_{N^*} \leq S_{\text{st}}^\sigma(\kappa) \wedge \exp \left\{ \frac{2}{\sigma^2}(H + \eta) \right\} \right) + \mathbb{P}(\exists m \leq N^* : \tau_m - \theta_m > T_1) \\
& \quad + \mathbb{P} \left( \exists k^* \leq N^* : \sup_{t \in [\tau_{k^*-1}; \tau_{k^*}]} |X_t^\sigma - Y_t^\sigma| > \alpha \text{ and} \right. \\
& \quad \left. \tau_{k^*} \leq S_{\text{st}}^\sigma(\kappa) \wedge \exp \left\{ \frac{2}{\sigma^2}(H + \eta) \right\}, \forall m \leq N^* : \tau_m - \theta_m \leq T_1, \right) \\
& =: I_1 + I_2 + I_3.
\end{aligned} \tag{4.38}$$

For the first probability:

$$\begin{aligned}
I_1 & \leq \mathbb{P} \left( \tau_{N^*} \leq \tau_{\mathcal{D}_R^e}^{Y, \sigma} \right) + \mathbb{P} \left( \tau_{\mathcal{D}_R^e}^{Y, \sigma} < \tau_{N^*} \leq \exp \left\{ \frac{2}{\sigma^2}(H + \eta) \right\} \right) \\
& \leq \mathbb{P} \left( \tau_{N^*} \leq \tau_{\mathcal{D}_R^e}^{Y, \sigma} \right) + \mathbb{P}(|X_{T_{\text{st}}^\sigma(\kappa)} - a| > \kappa) \\
& \quad + \sup_{y \in B_\kappa(a)} \mathbb{P}_y \left( \tau_{\mathcal{D}_R^e}^{Y, \sigma} < \exp \left\{ \frac{2}{\sigma^2}(H + \eta) \right\} \right) \xrightarrow{\sigma \rightarrow 0} 0,
\end{aligned}$$

by Lemmas 4.21, 4.11 and Proposition 4.13, and since  $H + \eta < H_R^e$ . At the same time, by Lemma 4.21, the second expression:

$$I_2 \xrightarrow{\sigma \rightarrow 0} 0.$$

What is left is the third expression. Note that, by (4.37),  $I_3$  is bounded by:

$$\sum_{k^*=1}^{N^*} \mathbb{P}(\Psi_{\theta_{k^*} - \tau_{k^*-1}} \circ \cdots \circ \Psi_{\theta_2 - \tau_1} \circ \Psi_{\theta_1 - T_{\text{st}}^\sigma(\kappa)} \circ 0 > \kappa).$$

We get that expression by observing that, if there exists  $k^*$  such that inequality

$$\sup_{t \in [\tau_{k^*}; \tau_{k^*+1}]} |X_t^\sigma - Y_t^\sigma| > \kappa$$

holds, then, given that this difference is smaller than  $\kappa$  for times smaller than  $\tau_{k^*}$ , we can control this difference in terms of  $\Psi_{\theta_k - \tau_{k-1}}$  by (4.37).

Let us study the sum above. By definition of  $\Psi_T$ , we have:

$$\begin{aligned} & \sum_{k^*=1}^{N^*} \mathbb{P}(\Psi_{\theta_{k^*} - \tau_{k^*-1}} \circ \cdots \circ \Psi_{\theta_2 - \tau_1} \circ \Psi_{\theta_1 - T_{\text{st}}(\kappa)} \circ 0 > \kappa) \\ & \leq \sum_{k^*=1}^{N^*} \mathbb{P} \left( \sum_{i=1}^{k^*-1} o_\kappa^i(1) \exp \left\{ \sum_{j=i}^{k^*} (LT_1 - K(\theta_j - \tau_{j-1})) \right\} + o_\kappa^{k^*}(1) > \kappa \right). \end{aligned}$$

We can continue the calculations and get the following upper bound:

$$\begin{aligned} & \sum_{k^*=1}^{N^*} \mathbb{P} \left( \sup_{1 \leq i \leq k^*-1} \exp \left\{ \sum_{j=i}^{k^*} (LT_1 - K(\theta_j - \tau_{j-1})) \right\} > \frac{\kappa - o_\kappa^{k^*}(1)}{k^* - 1} \right) \\ & \leq \sum_{k^*=1}^{N^*} \mathbb{P} \left( \sup_{1 \leq i \leq k^*-1} \left\{ \sum_{j=i}^{k^*} (LT_1 - K(\theta_j - \tau_{j-1})) \right\} > -\log(k^* - 1) \right) \end{aligned}$$

Note that, if there are more than  $(k^* - i + 1)/2$  intervals of the size  $(\theta_j - \tau_{j-1}) > \exp\left\{\frac{2(Q^c - \kappa)}{\sigma^2}\right\}$ , then necessarily

$$\begin{aligned} \sum_{j=i}^{k^*} (LT_1 - K(\theta_j - \tau_{j-1})) & \leq (k^* - i + 1)LT_1 - \left\lceil \frac{(k^* - i + 1)}{2} \right\rceil K e^{\frac{2(Q^c - \kappa)}{\sigma^2}} \\ & \leq (k^* - i + 1) \left( LT_1 - \frac{K}{2} e^{\frac{2(Q^c - \kappa)}{\sigma^2}} \right) \\ & \leq k^* \left( LT_1 - \frac{K}{2} e^{\frac{2(Q^c - \kappa)}{\sigma^2}} \right). \end{aligned}$$

Since  $k^* > \log(k^* - 1)$  for any  $k^* \geq 1$  and since  $\left( LT_1 - \frac{K}{2} e^{\frac{2(Q^c - \kappa)}{\sigma^2}} \right)$  is negative for small enough  $\sigma$ , we get

$$\sup_{1 \leq i \leq k^*-1} \left\{ \sum_{j=i}^{k^*} (LT_1 - K(\theta_j - \tau_{j-1})) \right\} \leq -\log(k^* - 1),$$

which means that it is impossible to have more than  $(k^* - i + 1)/2$  intervals of the size  $\theta_j - \tau_{j-1} \geq \exp\left\{\frac{2(Q^c - \kappa)}{\sigma^2}\right\}$ . Therefore, we have:

$$\begin{aligned} & \mathbb{P}\left(\sup_{1 \leq i \leq k^* - 1} \left\{ \sum_{j=i}^{k^*} (LT_1 - K(\theta_j - \tau_{j-1})) \right\} > -\log(k^* - 1)\right) \\ & \leq \mathbb{P}\left(\forall i \leq k^* - 1 : \# \left\{ j : i \leq j \leq k^* : \theta_j - \tau_{j-1} \leq \exp\left\{\frac{2(Q^c - \kappa)}{\sigma^2}\right\} \right\} \right. \\ & \qquad \qquad \qquad \left. \geq \frac{k^* - i + 1}{2} \right) \\ & \leq \min_{1 \leq i \leq k^* - 1} \sum_{n=\lfloor \frac{k^* - i + 1}{2} \rfloor}^{k^*} \sum_{(j_1, \dots, j_n)} \mathbb{P}\left(\bigcap_{l=1}^n \left\{ \theta_{j_l} - \tau_{j_l - 1} \leq \exp\left\{\frac{2(Q^c - \kappa)}{\sigma^2}\right\} \right\}\right). \end{aligned}$$

Since the number of combinations of the form  $(j_1, \dots, j_n)$  can be roughly bounded by  $2^n$ , we can deduce

$$\begin{aligned} & \min_{1 \leq i \leq k^* - 1} \sum_{n=\lfloor \frac{k^* - i + 1}{2} \rfloor}^{k^*} \sum_{(j_1, \dots, j_n)} \mathbb{P}\left(\bigcap_{l=1}^n \left\{ \theta_{j_l} - \tau_{j_l - 1} \leq \exp\left\{\frac{2(Q^c - \kappa)}{\sigma^2}\right\} \right\}\right) \\ & \leq \sum_{n=\lfloor \frac{k^*}{2} \rfloor}^{k^*} 2^n \left( \sup_{y \in B_{\rho/4}(a)} \mathbb{P}_y\left(\theta_0 \leq \exp\left\{\frac{2(Q^c - \kappa)}{\sigma^2}\right\}\right) \right)^n \\ & \leq o_\sigma(1)^{\lfloor \frac{k^* + 1}{2} \rfloor} \frac{1 + o_\sigma(1)^{\lfloor \frac{k^* + 1}{2} \rfloor}}{1 - o_\sigma(1)}, \end{aligned}$$

by Proposition 4.13.

Combining inequalities above, we can come back to (4.38) and conclude that  $I_3$  also tends to 0 with  $\sigma \rightarrow 0$ , for each  $\kappa > 0$  small enough, which finalizes the proof. □

### Control of $Y^\sigma$ : Proof of Lemma 4.15

We can show, using large deviations techniques, that there exists a uniform upper bound on the time of convergence of  $Y^\sigma$  inside  $B_{\rho/4}(a)$ . Namely, for

any  $r > 0$  small enough, there exists  $\bar{T} > 0$  such that

$$\sup_{y \in \mathcal{D}_r^e} \mathbb{P}_y(Y_{\bar{T}}^\sigma \notin B_{\rho/4}(a)) \xrightarrow{\sigma \rightarrow 0} 0.$$

The construction of such a  $\bar{T}$  can be found in [DZ10, Proof of Lemma 5.7.19].

Therefore, for small enough  $\sigma > 0$ , given only  $r$  and  $\rho$ , we can choose a continuous function  $\bar{o}(\sigma)$  such that  $\bar{o}(\sigma) \xrightarrow{\sigma \rightarrow 0} 0$  and we have

$$\sup_{y \in \mathcal{D}_r^e} \mathbb{P}_y(Y_{\bar{T}}^\sigma \notin B_{\rho/4}(a)) \leq \bar{o}(\sigma), \quad (4.39)$$

for all  $\sigma > 0$  small enough.

Moreover, by Proposition 4.13, we know that for any  $\delta > 0$  we have  $\mathbb{P}(\tau_{\mathcal{D}_r^e}^{Y, \sigma} \leq \exp\{\frac{2(H_r^e - \delta)}{\sigma^2}\}) \xrightarrow{\sigma \rightarrow 0} 0$ . After fixing some positive  $r > 0$  and choosing  $\delta$  to be small enough such that  $H < H_r^e - \delta$ , we can define  $\eta > 0$  as a small enough number such that  $H + \eta < H_r^e - \delta$ . In the following, we can restrict ourselves only to those trajectories that do not leave domain  $\mathcal{D}_r^e$  before time  $\exp\{\frac{2(H+\eta)}{\sigma^2}\} < \exp\{\frac{2(H_r^e - \delta)}{\sigma^2}\}$ . Define the event  $A := \{\tau_{\mathcal{D}_r^e}^{Y, \sigma} > \exp\{\frac{2(H+\eta)}{\sigma^2}\}\}$ .

Consider the following inequalities. By Lemma 4.11 and the definition of  $Y^\sigma$ , for any  $\kappa > 0$ , we can introduce  $\bar{o}_\kappa(\sigma)$ , the modification of function  $\bar{o}(\sigma)$  such that 4.39 still holds and also we have

$$\mathbb{P}(Y_{T_{\text{st}}^\sigma(\kappa)}^\sigma \notin B_{\rho/4}(a), A) \leq \bar{o}_\kappa(\sigma). \quad (4.40)$$

At the same time, using the Markov property of diffusion  $Y^\sigma$ , for small enough  $\sigma > 0$ , we have

$$\begin{aligned} & \mathbb{P}(Y_{T_{\text{st}}^\sigma(\kappa) + \bar{T}}^\sigma \notin B_{\rho/4}(a), A) \\ & \leq \sup_{y \in B_{\rho/4}(a)} \mathbb{P}_y(Y_{\bar{T}}^\sigma \notin B_{\rho/4}(a), A) \mathbb{P}(Y_{T_{\text{st}}^\sigma(\kappa)}^\sigma \in B_{\rho/4}(a), A) \\ & \quad + \sup_{y \in \mathcal{D}_r^e \setminus B_{\rho/4}(a)} \mathbb{P}_y(Y_{\bar{T}}^\sigma \notin B_{\rho/4}(a), A) \mathbb{P}(Y_{T_{\text{st}}^\sigma(\kappa)}^\sigma \notin B_{\rho/4}(a), A) \\ & \leq \bar{o}_\kappa(\sigma) + \bar{o}_\kappa^2(\sigma), \end{aligned}$$

by Equations (4.39) and (4.40), while  $\mathbb{P}(Y_{T_{\text{st}}^\sigma(\kappa)}^\sigma \in B_{\rho/4}(a), A)$  is bounded by 1. For the next step consider:

$$\begin{aligned} & \mathbb{P}(Y_{T_{\text{st}}^\sigma(\kappa)+2\bar{T}}^\sigma \notin B_{\rho/4}(a), A) \\ & \leq \sup_{y \in B_{\rho/4}(a)} \mathbb{P}_y(Y_{\bar{T}}^\sigma \notin B_{\rho/4}(a), A) \mathbb{P}(Y_{T_{\text{st}}^\sigma(\kappa)+\bar{T}}^\sigma \in B_{\rho/4}(a), A) \\ & \quad + \sup_{y \in \mathcal{D}_r^e \setminus B_{\rho/4}(a)} \mathbb{P}_y(Y_{\bar{T}}^\sigma \notin B_{\rho/4}(a), A) \mathbb{P}(Y_{T_{\text{st}}^\sigma(\kappa)+\bar{T}}^\sigma \notin B_{\rho/4}(a), A) \\ & \leq \bar{o}_\kappa(\sigma) [1 + \bar{o}_\kappa(\sigma) + \bar{o}_\kappa^2(\sigma)], \end{aligned}$$

similarly to the previous computations. For any fixed  $\kappa > 0$  and  $\sigma > 0$  small enough, we can repeat this procedure  $N(\sigma) := \left\lfloor \frac{1}{\bar{T}} \exp\left\{\frac{2(H+\eta)}{\sigma^2}\right\} \right\rfloor$  times, thus while  $A$  still holds. We finally get the following upper bound:

$$\sup_{n \leq N(\sigma)} \mathbb{P}(Y_{T_{\text{st}}^\sigma(\kappa)+n\bar{T}}^\sigma \notin B_{\rho/4}(a), A) \leq \bar{o}_\kappa(\sigma) \sum_{i=0}^{N(\sigma)} \bar{o}_\kappa^i(\sigma) \leq \frac{\bar{o}_\kappa(\sigma)}{1 - \bar{o}_\kappa(\sigma)}. \quad (4.41)$$

This allows us to confine with high probability  $Y^\sigma$  for points of time of the form  $T_{\text{st}}^\sigma(\kappa) + n\bar{T}$  inside the ball  $B_{\rho/4}(a)$ .

The last steps that one has to make in order to prove the lemma is, first, to control the probability  $\mathbb{P}(Y_t^\sigma \notin B_{\rho/2}(a))$  in between points of time of the form  $T_{\text{st}}^\sigma(\kappa) + k\bar{T}$  and  $T_{\text{st}}^\sigma(\kappa) + (k+1)\bar{T}$  and, second, remove event  $A$ . Note that

$$\sup_t \mathbb{P}(Y_t^\sigma \notin B_{\rho/2}(a)) \leq \mathbb{P}(\bar{A}) + \sup_t \mathbb{P}(Y_t^\sigma \notin B_{\rho/2}(a), A),$$

where the suprema are taken with respect to  $t \in \left[ T_{\text{st}}^\sigma(\kappa); \exp\left\{\frac{2(H+\eta)}{\sigma^2}\right\} \right]$ . The first probability tends to zero by Lemma 4.11 and Proposition 4.13, since

$$\mathbb{P}(\bar{A}) \leq \mathbb{P}(|X_{T_{\text{st}}^\sigma(\kappa)}^\sigma - a| > \kappa) + \sup_{y \in B_\kappa(a)} \mathbb{P}_y \left( \tau_{\mathcal{D}_r^e}^{Y, \sigma} < \exp\left\{\frac{2(H_r^e - \delta)}{\sigma^2}\right\} \right) \xrightarrow{\sigma \rightarrow 0} 0.$$

For the second probability, consider the following inequalities for any  $\kappa > 0$  and for any  $\sigma > 0$  small enough. Using the Markov property of  $Y^\sigma$ ,

we have

$$\begin{aligned} \sup_t \mathbb{P}(Y_t^\sigma \notin B_{\rho/2}(a), A) &\leq \sup_{n \leq N(\sigma)} \mathbb{P}(Y_{T_{\text{st}}^\sigma(\kappa) + n\bar{T}}^\sigma \in B_{\rho/4}(a), A) \\ &\quad \times \sup_{y \in B_{\rho/4}(a)} \sup_{t \leq \bar{T}} \mathbb{P}_y(Y_t^\sigma \notin B_{\rho/2}(a), A) \\ &+ \sup_{n \leq N(\sigma)} \mathbb{P}(Y_{T_{\text{st}}^\sigma(\kappa) + n\bar{T}}^\sigma \notin B_{\rho/4}(a), A) \\ &\quad \times \sup_{y \in \mathcal{D}_r^e} \sup_{t \leq \bar{T}} \mathbb{P}_y(Y_t^\sigma \notin B_{\rho/2}(a), A). \end{aligned}$$

Let us use (4.41) and bound by 1 the probabilities that are not needed for our derivations. Finally, we get for any  $\kappa > 0$  and  $\sigma > 0$  small enough:

$$\sup_t \mathbb{P}(Y_t^\sigma \notin B_{\rho/2}(a), A) \leq \sup_{y \in B_{\rho/4}(a)} \sup_{t \leq \bar{T}} \mathbb{P}_y(Y_t^\sigma \notin B_{\rho/2}(a)) + \frac{\bar{o}_\kappa(\sigma)}{1 - \bar{o}_\kappa(\sigma)}.$$

Note that  $\{Y_t^\sigma \notin B_{\rho/2}(a)\} \subseteq \{\tau_{B_{\rho/2}(a)}^{Y,\sigma} < t\}$ . Therefore, we have

$$\sup_{y \in B_{\rho/4}(a)} \sup_{t \leq \bar{T}} \mathbb{P}_y(Y_t^\sigma \notin B_{\rho/2}(a)) \leq \sup_{y \in B_{\rho/4}(a)} \mathbb{P}_y(\tau_{B_{\rho/2}(a)}^{Y,\sigma} < \bar{T}) \xrightarrow{\sigma \rightarrow 0} 0,$$

by Proposition 4.13. This finally shows that we can find  $\eta > 0$  such that for any  $\kappa > 0$  small enough, we have

$$\sup_{t \in [T_{\text{st}}^\sigma(\kappa); \exp\{\frac{2(H+\eta)}{\sigma^2}\}]} \mathbb{P}(Y_t^\sigma \notin B_{\rho/2}(a)) \xrightarrow{\sigma \rightarrow 0} 0,$$

which proves the lemma.

## Control of the law: Proof of Lemma 4.16

In this section, we prove Lemma 4.16. In order to do that, we first provide and prove Lemma 4.22 below, that is a modification of a technique introduced by J. Tugaut in [Tug18a]. Let  $\xi(t) := \mathbb{W}_2^2(\mu_t^\sigma; \delta_a)$ . Consider the following lemma:

**Lemma 4.22.** *Under Assumptions A-1–A-7, there exist  $K_1, K_2 > 0$  such that for any  $t > 0$ :*

$$\xi'(t) \leq -K_1 \xi(t) + d\sigma^2 + K_2 \sqrt{\mathbb{P}(X_t^\sigma \notin B_\rho(a))}.$$



*Proof.* The proof is similar to the one of [Tug18a, Lemma 4.1] although it is strongly different.

**Step 1.** First of all, by Itô's formula, we have

$$\begin{aligned} |X_t^\sigma - a|^2 &= |X_0 - a|^2 + 2\sigma \int_0^t \langle X_s^\sigma - a; dW_s \rangle - 2 \int_0^t \langle X_s^\sigma - a; \nabla V(X_s^\sigma) \rangle ds \\ &\quad - 2 \int_0^t \langle X_s^\sigma - a; \nabla F * \mu_s^\sigma(X_s^\sigma) \rangle ds + d\sigma^2 t. \end{aligned}$$

For the next step, we take the expectation and derivative with respect to  $t$ . We get:

$$\xi'(t) = d\sigma^2 - 2\mathbb{E}[\langle X_t^\sigma - a; \nabla V(X_t^\sigma) + \nabla F * \mu_t^\sigma(X_t^\sigma) \rangle].$$

**Step 2.** Let us introduce  $\tilde{F} \in C^2(\mathbb{R}^d; \mathbb{R})$  – a modification of the function  $F$  such that  $\tilde{F}$  is "convex enough" around 0. Namely, if  $\nabla^2 F(0) \succeq \frac{C_W}{2} \text{Id}$ , where  $C_W$  is the positive constant from Definition 4.2, then we simply let  $\tilde{F} = F$ . If not, we introduce a matrix  $\mathcal{M} := -\nabla^2 F(0) + \frac{C_W}{2} \text{Id}$  and define  $\tilde{F}(x) := F(x) + \frac{1}{2} \langle x; \mathcal{M}x \rangle$ .

In the following, without loss of generality, we stick to the case  $\nabla^2 F(0) \prec \frac{C_W}{2} \text{Id}$ . Moreover, without loss of generality, we assume that  $\tilde{F}$  is locally convex inside the ball  $B_\rho(0)$ , where  $\rho$  is the radius of convexity of the effective potential introduced in Definition 4.2. Indeed, since  $\nabla^2 F$  is continuous, we can always choose  $\rho$  in Definition 4.2 to be small enough such that  $\nabla^2 F(x) - \nabla^2 F(0) \succ -\frac{C_W}{2} \text{Id}$  for any  $x \in B_\rho(0)$ . Note, that, under these assumptions,  $\mathcal{M}$  is a positive definite matrix.

**Step 3.** By definition of  $\tilde{F}$ , we have:

$$\begin{aligned} &\mathbb{E}[\langle X_t^\sigma - a; \nabla F * \mu_t^\sigma(X_t^\sigma) \rangle] \\ &= \mathbb{E}[\langle X_t^\sigma - a; \nabla \tilde{F} * \mu_t^\sigma(X_t^\sigma) \rangle] - \mathbb{E}[\langle X_t^\sigma - a; \mathcal{M}(X_t^\sigma - \mathbb{E}[X_t^\sigma]) \rangle] \\ &= \mathbb{E}[\langle X_t^\sigma - a; \nabla \tilde{F} * \mu_t^\sigma(X_t^\sigma) \rangle] - \mathbb{E}[\langle X_t^\sigma - a; \mathcal{M}(X_t^\sigma - a) \rangle] \\ &\quad - \mathbb{E}[\langle X_t^\sigma - a; \mathcal{M}(a - \mathbb{E}[X_t^\sigma]) \rangle]. \end{aligned}$$

Let  $Y_t^\sigma$  be an independent copy of  $X_t^\sigma$ . Since  $\mathcal{M}$  is positive definite, this gives us the following lower bound:

$$\begin{aligned} & \mathbb{E}[\langle X_t^\sigma - a; \nabla F * \mu_t^\sigma(X_t^\sigma) \rangle] \\ & \geq \mathbb{E}[\langle X_t^\sigma - a; \nabla \tilde{F}(X_t^\sigma - Y_t^\sigma) \rangle] - \mathbb{E}[\langle X_t^\sigma - a; \mathcal{M}(X_t^\sigma - a) \rangle]. \end{aligned} \quad (4.42)$$

**Step 4.** We now focus on the first term of the inequality above. Let us consider separately the parts of the process  $X^\sigma$  lying outside and inside the ball  $B_{\rho/2}(a)$ . Using the polynomial growth (Assumption A-2), for some generic constant  $C$ , we get:

$$\begin{aligned} & \mathbb{E}[\langle X_t^\sigma - a; \nabla \tilde{F}(X_t^\sigma - Y_t^\sigma) \rangle] \\ & \geq \mathbb{E}[\langle X_t^\sigma - a; \nabla \tilde{F}(X_t^\sigma - Y_t^\sigma) \rangle \mathbf{1}_{X_t^\sigma \in B_{\rho/2}(a)} \mathbf{1}_{Y_t^\sigma \in B_{\rho/2}(a)}] \\ & \quad - C \mathbb{E}[(1 + |X_t^\sigma|^{2r}) \mathbf{1}_{X_t^\sigma \notin B_{\rho/2}(a)}]. \end{aligned}$$

Since  $\tilde{F}$  is convex inside  $B_\rho(0)$  and the moments are uniformly bounded (Proposition 4.6), by using the Cauchy–Schwarz inequality, for any  $0 < \sigma < 1$ , we immediately obtain the existence of a positive constant  $K > 0$  such that:

$$\mathbb{E}[\langle X_t^\sigma - a; \nabla \tilde{F}(X_t^\sigma - Y_t^\sigma) \rangle] \geq -K \sqrt{\mathbb{P}(X_t^\sigma \notin B_{\rho/2}(a))}.$$

We plug this inequality in Equation 4.42 and get:

$$\begin{aligned} \mathbb{E}[\langle X_t^\sigma - a; \nabla F * \mu_t^\sigma(X_t^\sigma) \rangle] & \geq -\mathbb{E}[\langle X_t^\sigma - a; \mathcal{M}(X_t^\sigma - a) \rangle] \\ & \quad - K \sqrt{\mathbb{P}(X_t^\sigma \notin B_{\rho/2}(a))}. \end{aligned} \quad (4.43)$$

**Step 5.** We now focus on the term involving  $\nabla V$ . According to Definition 4.2, for any  $x \in B_\rho(a)$ , we have:

$$\nabla^2 V(x) \succeq C_W \text{Id} - \nabla^2 F(x - a).$$

At the same time, by the definition of  $\mathcal{M}$ :

$$\nabla^2 F(0) = -\mathcal{M} + \frac{C_W}{2} \text{Id}.$$

Since  $\nabla^2 F$  is continuous, we can, without loss of generality, decrease  $\rho$  if necessary so that  $-\nabla^2 F(x) \succeq -\nabla^2 F(0) - \frac{C_W}{4}\text{Id}$  for any  $x \in B_\rho(0)$ . Therefore, for any  $x \in B_\rho(a)$ , we have:

$$\nabla^2 V(x) \succeq \frac{C_W}{4}\text{Id} + \mathcal{M}.$$

Using the same logic as in **Step 4**, we get:

$$\begin{aligned} & \mathbb{E}[\langle X_t^\sigma - a; \nabla V(X_t^\sigma) \rangle] \\ &= \mathbb{E}[\langle X_t^\sigma - a; \nabla V(X_t^\sigma) \rangle \mathbb{1}_{X_t^\sigma \in B_{\rho/2}(a)}] + \mathbb{E}[\langle X_t^\sigma - a; \nabla V(X_t^\sigma) \rangle \mathbb{1}_{X_t^\sigma \notin B_{\rho/2}(a)}] \\ &\geq \mathbb{E} \left[ \langle X_t^\sigma - a; \left( \frac{C_W}{4}\text{Id} + \mathcal{M} \right) (X_t^\sigma - a) \rangle \right] \\ &\quad - K \sqrt{\mathbb{P}(X_t^\sigma \notin B_{\rho/2}(a))}. \end{aligned} \tag{4.44}$$

**Final step.** As a consequence, putting 4.43 and 4.44 in the **Step 1**, we get:

$$\xi'(t) \leq d\sigma^2 - \frac{C_W}{2}\xi(t) + 2K \sqrt{\mathbb{P}(X_t^\sigma \notin B_{\rho/2}(a))},$$

which concludes the proof. □

Now we are ready to prove Lemma 4.16 itself.

*proof of Lemma 4.16.* In order to prove the lemma, we use Lemma 4.22 above, that is the inequality

$$\xi'(t) \leq -2\rho' \xi(t) + d\sigma^2 + K \sqrt{\mathbb{P}(X_t^\sigma \notin B_\rho(a))}.$$

After that, we use Lemma 4.15 along with Proposition 4.14 in order to show that the term  $\mathbb{P}(X_t^\sigma \notin B_\rho(a))$  tends to 0 with  $\sigma \rightarrow 0$  for any  $T_{\text{st}}^\sigma(\kappa) \leq t \leq S_{\text{st}}^\sigma(\kappa) \wedge \exp\left\{\frac{2(H+\eta)}{\sigma^2}\right\}$ , which, in its term, means that we can choose  $\sigma$  to be small enough such that  $\xi(t) \leq \kappa^2$  for all such  $t$ . Final step is to

show that  $S_{\text{st}}^\sigma(\kappa)$  can not be less or equal than  $\exp\left\{\frac{2(H+\eta)}{\sigma^2}\right\}$  or else we get contradiction between the fact that  $\xi(S_{\text{st}}^\sigma(\kappa)) \leq \kappa^2$  and definition of  $S_{\text{st}}^\sigma(\kappa)$ .

Consider the following inequalities. For any  $T_{\text{st}}^\sigma(\kappa) \leq t \leq S_{\text{st}}^\sigma(\kappa) \wedge \exp\left\{\frac{2(H+\eta)}{\sigma^2}\right\}$ :

$$\mathbb{P}(X_t^\sigma \notin B_\rho(a)) \leq \mathbb{P}(Y_t^\sigma \notin B_{\rho/2}(a)) + \mathbb{P}(|Y_t^\sigma - X_t^\sigma| > \rho/2) = o_\sigma(1),$$

by Lemma 4.15 and Proposition 4.14. Thus, by Lemma 4.22,  $\xi(t) = \mathbb{W}_2^2(\mu_t; \delta_a)$  is bounded for any  $t$  considered above in the following way:

$$\xi'(t) \leq -2\rho'\xi(t) + d\sigma^2 + K o_\sigma(1).$$

Therefore, we can decrease  $\kappa$  and then  $\sigma$  to be small enough such that  $\xi(t) \leq \kappa^2$  for any  $T_{\text{st}}^\sigma(\kappa) \leq t \leq S_{\text{st}}^\sigma(\kappa) \wedge \exp\left\{\frac{2(H+\eta)}{\sigma^2}\right\}$ .

The last step is to note that if  $S_{\text{st}}^\sigma(\kappa) < \exp\left\{\frac{2(H+\eta)}{\sigma^2}\right\}$ , then we get a contradiction between definition of  $S_{\text{st}}^\sigma(\kappa)$  and the fact that  $\xi(S_{\text{st}}^\sigma(\kappa)) \leq \kappa^2$ , which proves the lemma.

□

# Bibliography

- [AdMKT23] Ashot Aleksian, Pierre del Moral, Aline Kurtzmann, and Julian Tugaut. Self-interacting diffusions: long-time behaviour and exit-problem in the convex case, 2023. preprint.
- [AdRR<sup>+</sup>22] Daniel Adams, Gonalo dos Reis, Romain Ravaille, William Salkeld, and Julian Tugaut. Large deviations and exit-times for reflected McKean-Vlasov equations with self-stabilising terms and superlinear drifts. *Stochastic Process. Appl.*, 146:264–310, 2022.
- [AKT23] Ashot Aleksian, Aline Kurtzmann, and Julian Tugaut. Exit-problem for a class of non-Markov processes with path dependency, 2023. preprint.
- [Arr89a] Svante Arrhenius. ber die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Suren. *Zeitschrift fr physikalische Chemie*, 4(1):226–248, 1889.
- [Arr89b] Svante Arrhenius. ber die Dissociationswarme und den Einfluss der Temperatur auf den Dissociationsgrad der Elektrolyte. *Zeitschrift fr physikalische Chemie*, 4(1):96–116, 1889.
- [AT23] Ashot Aleksian and Julian Tugaut. Measure-dependent nonlinear diffusions with superlinear drifts: asymptotic behaviour of the first exit-times, 2023. preprint.
- [BEGK04] Anton Bovier, Michael Eckhoff, Veronique Gaynard, and Markus Klein. Metastability in reversible diffusion processes.

- I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc. (JEMS)*, 6(4):399–424, 2004.
- [BGK05] Anton Bovier, Véronique Gayraud, and Markus Klein. Metastability in reversible diffusion processes. II. Precise asymptotics for small eigenvalues. *J. Eur. Math. Soc. (JEMS)*, 7(1):69–99, 2005.
- [BLR02] Michel Benaïm, Michel Ledoux, and Olivier Raimond. Self-interacting diffusions. *Probab. Theory Related Fields*, 122(1):1–41, 2002.
- [Bol03] Ludwig Boltzmann. Further Studies on the Thermal Equilibrium of Gas Molecules. *History of Modern Physical Sciences*, 1:262–349, July 2003.
- [BR05] Michel Benaïm and Olivier Raimond. Self-interacting diffusions. III. Symmetric interactions. *Ann. Probab.*, 33(5):1717–1759, 2005.
- [BRTV98] S. Benachour, B. Roynette, D. Talay, and P. Vallois. Nonlinear self-stabilizing processes. I. Existence, invariant probability, propagation of chaos. *Stochastic Process. Appl.*, 75(2):173–201, 1998.
- [CD22a] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: A review of models, methods and applications. I. Models and methods. *Kinetic and Related Models*, 15(6):895, 2022.
- [CD22b] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: A review of models, methods and applications. II. Applications. *Kinetic and Related Models*, 15(6):1017, 2022.
- [CGM08] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields*, 140(1-2):19–40, 2008.
- [CMV03] José A. Carrillo, Robert J. McCann, and Cédric Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3):971–1018, 2003.

- [Col06a] Andrea Collecchio. Limit theorems for reinforced random walks on certain trees. *Probability theory and related fields*, 136(1):81–101, 2006.
- [Col06b] Andrea Collecchio. On the transience of processes defined on galton-watson trees. *The Annals of Probability*, pages 870–878, 2006.
- [Dia88] P. Diaconis. Recent progress on de Finetti’s notions of exchangeability. In *Bayesian statistics, 3 (Valencia, 1987)*, Oxford Sci. Publ., pages 111–125. Oxford Univ. Press, New York, 1988.
- [DR92] R. T. Durrett and L. C. G. Rogers. Asymptotic behavior of Brownian polymers. *Probab. Theory Related Fields*, 92(3):337–349, 1992.
- [dRDM<sup>+</sup>] Chaudru de Raynal, Hong Duong, Pierre Monmarché, Milica Tomasevic, and Julian Tugaut. Reducing exit-times of diffusions with repulsive interactions.
- [DT16] M. H. Duong and J. Tugaut. Stationary solutions of the Vlasov-Fokker-Planck equation: existence, characterization and phase-transition. *Appl. Math. Lett.*, 52:38–45, 2016.
- [DT18] Manh Hong Duong and Julian Tugaut. The Vlasov-Fokker-Planck equation in non-convex landscapes: convergence to equilibrium. *Electron. Commun. Probab.*, 23:Paper No. 19, 10, 2018.
- [DZ10] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.
- [FW98] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, second edition, 1998. Translated from the 1979 Russian original by Joseph Szücs.

- [GGM<sup>+</sup>18] Clemens Gohlke, Paul Gajewski, Mario Maurelli, Peter K. Friz, and Wolfgang Dreyer. Stochastic many-particle model for LFP electrodes. *Contin. Mech. Thermodyn.*, 30(3):593–628, 2018.
- [GM89] Carl Graham and Michel Métivier. System of interacting particles and nonlinear diffusion reflecting in a domain with sticky boundary. *Probab. Theory Related Fields*, 82(2):225–240, 1989.
- [Gra92] Carl Graham. Nonlinear diffusion with jumps. *Ann. Inst. H. Poincaré Probab. Statist.*, 28(3):393–402, 1992.
- [GS20] Rishabh S. Gvalani and André Schlichting. Barriers of the McKean-Vlasov energy via a mountain pass theorem in the space of probability measures. *J. Funct. Anal.*, 279(11):108720, 34, 2020.
- [HIP08] Samuel Herrmann, Peter Imkeller, and Dierk Peithmann. Large deviations and a Kramers’ type law for self-stabilizing diffusions. *Ann. Appl. Probab.*, 18(4):1379–1423, 2008.
- [HT10] S. Herrmann and J. Tugaut. Stationary measures for self-stabilizing processes: asymptotic analysis in the small noise limit. *Electron. J. Probab.*, 15:no. 69, 2087–2116, 2010.
- [JS17] Samy Jazaerli and Yuri F. Saporito. Functional Itô calculus, path-dependence and the computation of Greeks. *Stochastic Process. Appl.*, 127(12):3997–4028, 2017.
- [KK12] Victor Kleptsyn and Aline Kurtzmann. Ergodicity of self-attracting motion. *Electron. J. Probab.*, 17:no. 50, 37, 2012.
- [Kur10] Aline Kurtzmann. The ODE method for some self-interacting diffusions on  $\mathbb{R}^d$ . *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(3):618–643, 2010.
- [McK66] H. P. McKean, Jr. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. U.S.A.*, 56:1907–1911, 1966.



- [McK67] H. P. McKean, Jr. Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57. Air Force Office Sci. Res., Arlington, Va., 1967.
- [Mon17] Pierre Monmarché. Long-time behaviour and propagation of chaos for mean field kinetic particles. *Stochastic Process. Appl.*, 127(6):1721–1737, 2017.
- [NRS85] A. G. Nobile, L. M. Ricciardi, and L. Sacerdote. Exponential trends of first-passage-time densities for a class of diffusion processes with steady-state distribution. *J. Appl. Probab.*, 22(3):611–618, 1985.
- [NRW87] J. R. Norris, L. C. G. Rogers, and David Williams. Self-avoiding random walk: a Brownian motion model with local time drift. *Probab. Theory Related Fields*, 74(2):271–287, 1987.
- [Pem88a] Robin Pemantle. Phase transition in reinforced random walk and rwre on trees. *The Annals of Probability*, pages 1229–1241, 1988.
- [Pem88b] Robin Pemantle. *Random processes with reinforcement*. PhD thesis, M.I.T., 1988.
- [Pem90] Robin Pemantle. Nonconvergence to unstable points in urn models and stochastic approximations. *The Annals of Probability*, 18(2):698–712, 1990.
- [Rai09] Olivier Raimond. Self-interacting diffusions: a simulated annealing version. *Probab. Theory Related Fields*, 144(1-2):247–279, 2009.
- [RW00] L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.

- [RW19] Panpan Ren and Feng-Yu Wang. Bismut formula for Lions derivative of distribution dependent SDEs and applications. *J. Differential Equations*, 267(8):4745–4777, 2019.
- [SV79] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1979.
- [Szn91] A.-S. Sznitman. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.
- [Tan79] Hiroshi Tanaka. Stochastic differential equations with reflecting boundary condition in convex regions. *Hiroshima Math. J.*, 9(1):163–177, 1979.
- [TTV12] Pierre Tarrès, Bálint Tóth, and Benedek Valkó. Diffusivity bounds for 1D Brownian polymers. *Ann. Probab.*, 40(2):695–713, 2012.
- [Tug12] Julian Tugaut. Exit problem of McKean-Vlasov diffusions in convex landscapes. *Electron. J. Probab.*, 17:no. 76, 26, 2012.
- [Tug16] Julian Tugaut. A simple proof of a Kramers’ type law for self-stabilizing diffusions. *Electron. Commun. Probab.*, 21:Paper No. 11, 7, 2016.
- [Tug18a] Julian Tugaut. Exit-problem of McKean-Vlasov diffusions in double-well landscape. *J. Theoret. Probab.*, 31(2):1013–1023, 2018.
- [Tug18b] Julian Tugaut. Exit-time of granular media equation starting in a local minimum. *Commun. Stoch. Anal.*, 12(1):Art. 3, 31–36, 2018.
- [Tug19] Julian Tugaut. A simple proof of a Kramers’ type law for self-stabilizing diffusions in double-wells landscape. *ALEA Lat. Am. J. Probab. Math. Stat.*, 16(1):389–398, 2019.

- [Tug21] Julian Tugaut. Captivity of the solution to the granular media equation. *Kinet. Relat. Models*, 14(2):199–209, 2021.
- [VF70] A. D. Ventcel' and M. I. Freĭdlin. Small random perturbations of dynamical systems. *Uspehi Mat. Nauk*, 25(1 (151)):3–55, 1970.
- [Vil09] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.