LONG TIME BEHAVIOUR AND STATIONARY REGIME OF MEMORY GRADIENT DIFFUSIONS

Sébastien Gadat & Fabien Panloup
Institut de Mathématiques
Université de Toulouse (UMR 5219)
31062 Toulouse, Cedex 9, France
Sebastien.Gadat@math.univ-toulouse.fr,fabien.panloup@insa-toulouse.fr

Abstract

In this paper, we are interested in a diffusion process based on a gradient descent. The process is non Markov and has a memory term which is built as a weighted average of the drift term all along the past of the trajectory. For this type of diffusion, we study the long time behaviour of the process in terms of the memory. We exhibit some conditions for the long-time stability of the dynamical system and then provide, when stable, some convergence properties of the occupation measures and of the marginal distribution, to the associated steady regimes. When the memory is too long, we show that in general, the dynamical system has a tendency to explode, and in the particular gaussian case, we explicitly obtain the rate of divergence.

Keywords: Stochastic differential equation; Memory diffusions; Ergodic processes; Lyapunov function.

AMS classification (2000): 60J60, 60G10, 37A25, 93D30, 35H10.

1 Introduction

We are interested in this work in the evolution of a diffusion with a drift defined as an average over all past positions of a gradient of a functional. If k and h are two positive and increasing maps, the process can be written as

$$dX_t = -\left(\frac{1}{k(t)}\int_0^t h(s)\nabla U(X_s)ds\right)dt + \sigma(t,X_t)dW_t. \tag{1.1}$$

Such diffusions are naturally derived from the family of deterministic ordinary differential equations given by

$$x'_s = -\frac{1}{k(s)}\int_0^s h(u)\nabla U(x_u)du. \tag{1.2}$$

For optimization procedure, this deterministic equation may be useful since the solution $(x_s)_{s\geq 0}$ behaves like an inertial gradient descent: this point enables the solution to avoid some local traps although this property is of course false with a simple gradient descent. In some recent works of Cabot, Engler, and Gadat (2009a) and Cabot (2009), the authors propose a tricky change of variable to link the behaviour of (1.2) with the second order differential equation with a damping coefficient a:

$$\forall t \geq 0 \qquad z''(t) + a(t)z'(t) + \nabla U[z(t)] = 0. \tag{1.3}$$

More precisely, if τ is the solution initialized with $\tau(0) = 0$ of the ordinary differential equation

$$\tau'(t) = \sqrt{\frac{k(\tau(t))}{h(\tau(t))}},$$

Cabot (2009) shows that if x satisfies (1.2), then $z = x \circ \tau$ satisfies (1.3) with a damping effect a given by

$$a(t) = \left(\frac{k'h + kh'}{2h^{3/2}k} \right) \circ \tau(t) = \frac{1}{2}\sqrt{\frac{k}{h}} \left(\frac{h'}{h} + \frac{k'}{k} \right) (\tau(t)).$$

Equation (1.3) is indeed a generalization of the so-called dissipative Heavy Ball with Friction (HBF) system whose equations were first introduced by Polyak (1987) and Antipin (1994). If U is a real positive map from $\mathbb{R}^d$ to $\mathbb{R}_+$, the HBF models the evolution of a ball left on the graph of U and which is submitted to the action of the gravity with some friction resistance proportional to its speed, the friction here is described through the application a. There exists a large bibliography on equation (1.3) among the convex and optimization community, and most of these works are concerned with the convergence of the trajectory and its optimization properties. Most of the past works dealt with these properties depending on U and the repelling coefficient $a(t)$ that may (or not) depend on time t. When a is constant, some old result of Haraux (1991) establishes the convergence of solutions through some critical point of U in the one dimensional case when U is a bounded from below and coercive potential. More recently, Alvarez (2000) shows the asymptotic convergence of solutions of equation (1.3) to some minimum point of the potential U in Hilbert spaces with constant damping function a and any convex potential U. In much more general cases, Alvarez, Attouch, Bolte, and Redont (2002) yield some weak convergence of the trajectory to some minimizer (resp. critical point) of U (always for constant friction effect a) when the potential U is convex (resp. analytic) and the trajectory is weakly compact. At last, in a recent paper, Cabot et al. (2009a) establish some convergence results and optimization properties of the trajectories with general vanishing (or not) damping effect a and potential U.

In the sequel, we will be interested in the stochastic evolution of equations similar to (1.2) in the special case of $h = k'$ for k any positive increasing application. It can be shown that if k increases at least as $\sqrt{t}$, then a is likely to be positive. In this situation, it is easy to compute τ for special memory functions.

- If $k(t) = e^{\lambda t}$, one can show that $\tau'(t) = t/\sqrt{\lambda}$ and $a(t) = \sqrt{\lambda}$.
- If $k(t) = t^\alpha$ with $\alpha \geq 1/2$, it is also immediate to see that $\tau(t) = t^2/4\alpha$ and $a(t) = \frac{2\alpha - 1}{t}$.

Note that for each of these two situations, Cabot et al. (2009a) and Cabot, Engler, and Gadat (2009b) have shown that the deterministic trajectory $(z_s)_{s\geq 0}$ solution of (1.3) converges to a critical point of U which is generically a local minimum of U (the set of initialization points for whom $(z_s)_{s\geq 0}$ converges to a minimum is open and dense). Hence, we will be interested in this work in the behaviour of the system (1.1) for these two typical cases of memory.
One may rely the behaviour of (1.1) to a stochastic HBF as follows. Indeed, a stochastic version of (1.3) with any variance $\Sigma_{HBF}(s,.)$ can be expressed as a couple (z_1, z_2) satisfying

$$\left\{ \begin{array}{rcl} dz_1(s) & = & -z_2(s)ds + \Sigma_{HBF}(s, z_1(s))dW_s, \\ dz_2(s) & = & -a(s)z_2(s)ds - \nabla U(z_1(s))ds. \end{array} \right.$$

Following the reparametrization $X_t = z_1 \circ \tau^{-1}$, it is easy to show that $(X_t)_{t\geq 0}$ and $(z_1(t))_{t\geq 0}$ are equivalent up to the change of parametrisation τ if σ and Σ_{HBF} satisfy

$$\Sigma^2_{HBF}(t,.) = \sigma^2(t,.)\sqrt{\tau'(\tau^{-1}(t))}.$$

For instance, when $k(t) = e^{\lambda t}$, a time homogeneous σ in the average gradient descent (1.1) is equivalent to a time homogeneous Σ_{HBF} in the stochastic HBF. At last, in the case $k(t) = t^\alpha$, a time homogeneous Σ_{HBF} in the stochastic (HBF) system corresponds to an annealing situation where $\sigma^2(t) = \Sigma^2_{HBF}\{\alpha/t\}^{1/4}$ although conversely, a time homogeneous σ in the average gradient system (1.1) implies an increasing amount of noise in the stochastic HBF. In this work, we

will only consider the case of time-homogeneous σ since our main objective is to understand the effect of the memory on the dynamical system with a fixed level of noise.

Regarding now the probabilistic past works, in a sense, our work belongs to the large class of self-interacting diffusions introduced by Coppersmith and Diaconis (1987) that describe some non Markovian dynamical system whose evolution depends on the whole past of the trajectory. Such processes have been extensively studied in the *discrete* settings within the case of Random Walks with Reinforcement by Pemantle (1992) for the evolution of a growing polymer model.

In the *continuous* settings, Cranston and LeJan (1995) have studied a process (denoted $(X_t)_{t\geq 0}$) that looks similar to the one introduced in (1.1) and have established the almost sure behaviour of X_t/t for a special drift based on a functional of the differences $(X_t - X_s)_{0\leq s\leq t}$. The process is then reinforced by the occupation measure $\int_0^t \delta_{X_s} ds$, (see also the work of Raimond (1997) for an averaged drift based on $\frac{X_t - X_s}{\|X_t - X_s\|}$ as well as the initial work of Durrett and Rogers (1992)). Further works of Benaïm, Ledoux, and Raimond (2002), Benaïm and Raimond (2003) provided some complete study of self-attractive or interactive diffusions with values in a compact set when the process is reinforced by its *normalized* occupation measure (μ_t). They obtained some convergence of (μ_t) towards a measure defined as a fixed point of an equation derived through a Gibbs field. Their work is mainly based on the powerful tool of asymptotic pseudo-trajectory introduced by Benaïm and Hirsh (1996) and a compactness assumption. Further results can be obtained in the special case of symmetric self-interactions as pointed by Benaïm and Raimond (2005). In some very recent works (Kurtzman, 2009; Kurtzman & Chambeu, 2009), some study of the asymptotics of such types of non-homogeneous Markov processes has been extended to the non-compact setting. At last, one may also refer to the interesting works of Bakhtin (2002) and Bakhtin and Mattingly (2005) where the authors define in a very general case some diffusion for which the drift depends on the whole past trajectory: the drift coefficient of the equation is a nonlinear functional of the past history of the solution and they provide sufficient conditions for the existence and uniqueness of such solutions. Another common point with this work is the intensive use of Lyapunov function of the system. Note that such infinite memory diffusions may have some applications for stochastic Navier-Stokes equations (see *e.g.* Bakhtin (2006) for further details).

From a pure technical point of view, we will use a dimensional increment to treat (1.1) with Markovian tools. Hence, this space enlargement will naturally yield some coupled Langevin equations on the position and speed of a particle. Some recent works have dealt with the study of some processes $(X_t, V_t)_{t\geq 0}$ based on:

$$\begin{cases} dX_t &= V_t dt, \\ dV_t &= F(X_t, V_t)dt + \sigma(X_t, V_t)dW_t, \end{cases}$$

and such coupled equations cover a large number of situations such that the kinetic Fokker Planck equation for instance (one may find many details and references in Villani (2009), page 11 and the section 7 of chapter 1). To the best of our knowledge, the noise term dW always acts directly on the speed component and not on the position increment. Note that in our work, the noise will act on the position of the particle itself but not on its speed.

Let us now describe the main objective of the paper that is to study the long time behaviour of the dynamical system defined by (1.1), in terms of U, σ and especially to $t \mapsto k(t)$ that plays the role of the "memory" of the system. More precisely, we will be interested in the long-time stability of this process (*i.e.* existence and uniqueness of a steady regime, convergence properties to this steady regime including rate of convergence), and a description of this steady regime when it is possible. The paper is organized as follows. In Section 2, we first state our basic definitions and a description of (X_t) (solution to (1.1)) as a component of a generally non-homogeneous $\mathbb{R}^d \times \mathbb{R}^d$-Markov process that we denote $(X_t, Y_t)_{t\geq 0}$ (see (2.3)).

Then, we give some preliminary results about the existence of solutions and on the hypoelliptic nature of $(X_t, Y_t)_{t\geq 0}$. Under some non-degeneracy conditions, this second result leads to uniqueness of the invariant distribution in the homogeneous case. In Section 3, we state our main results about the long-time stability of the of our process in terms of the memory function $t \mapsto k(t)$ or more precisely of $r_\infty := \liminf(\frac{k'}{k})(t)$ as $t \to +\infty$. Throughout the paper, we assume that k is a positive increasing function. Thus, r_∞ belongs to $[0, +\infty]$. In Subsection 3.1, we focus on the (stable) case: $r_\infty \in (0, +\infty]$. Under some repelling condition on U, we build a Lyapunov function for the dynamical system and state a series of results about the long-time weak convergence of the occupation measures, some properties of the invariant distribution and convergence rates for the marginal distribution of $(X_t, Y_t)_{t\geq 0}$ to the steady regime. Then, in Subsection 3.2, we show that when $r_\infty = 0$ (*i.e.* when the dynamical system has too much memory), $(X_t)_{t\geq 0}$ has some long-time explosion properties. More precisely, we show that there exists a subsequence $(t_n)_{n\geq 0}$ such that $t_n \to +\infty$ and $\mathbb{E}[|X_{t_n}|^2] \to +\infty$. Furthermore, when $U(x) = x^2/2$, we obtain a CLT that gives the explicit rate of divergence of $(X_t)_{t\geq 0}$ in this particular case. Finally, Sections 4, 5, 6 and 7 are devoted to the proofs of the main results.

Aknowledgements We would like to gratefully acknowledge Patrick Cattiaux, François Delarue, Arnaud Guillin, Laurent Miclo and Stéphane Menozzi for their interests, their helpful discussions and their warm support. We are very much indebted to the anonymous referees and Alice Guionnet for their constructive comments that resulted in a significant improvement of the original manuscript.

2 Setting and General Statements

Before a precise definition of the dynamical system, let us list a short series of notations. The scalar product and the Euclidean norm on $\mathbb{R}^d$ are respectively denoted by $\langle\,,\,\rangle$ and $|\,.\,|$. The set of $d \times q$ matrices is denoted by $\mathbb{M}_{d,q}$ and we adopt the notation $\|\,.\,\|$ for every non-explicit norm on this finite-dimensional vector space. For a $\mathcal{C}^3$-function $f : \mathbb{R}^d \to \mathbb{R}$, ∇f, $D^2 f$ denote respectively the gradient of f and the Hessian matrix of f and $D^3 f$ is defined for every i, j, $k \in \{1, \ldots, d\}$ by $(D^3 f(x))_{i,j,k} = \partial^3_{x_i,x_j,x_k} f(x)$. For every $x \in \mathbb{R}^d$, we set

$$\|\|D^3 f(x)\|\| = \left(\sum_{i,j,k} \left|\partial^3_{x_i,x_j,x_k} f(x)\right|^2\right)^{\frac{1}{2}}.$$

Given any $\mathcal{C}^2$-function $f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, $\nabla_x f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ and $D^2_x f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{M}_{d,d}$ denote the functions respectively defined by $(\nabla_x f(x,y))_i = \partial_{x_i} f(x,y)$ and $(D^2_x f(x,y))_{i,j} = \partial_{x_i}\partial_{x_j} f(x,y)$. For a measure μ and a μ-integrable function f, we set $\mu(f) = \int f d\mu$. The Lebesgue measure on $\mathbb{R}^d$ is denoted by λ_d. Finally, we will denote by C every non explicit positive constant.

Throughout this paper, we denote by $U : \mathbb{R}^d \mapsto \mathbb{R}$ a smooth (at least $\mathcal{C}^2$) function on $\mathbb{R}^d$ satisfying the following coercivity conditions:

$$\lim_{|x|\to+\infty} U(x) = +\infty, \quad \inf_{x\in\mathbb{R}^d} U(x) > 0 \quad \text{and,} \quad \liminf_{|x|\to+\infty} \langle x, \nabla U(x)\rangle > 0. \tag{2.1}$$

We consider the following SDE:

$$d\mathcal{X}_t = \sigma(\mathcal{X}_t) dW_t - \frac{1}{k(t)}\left(\int_0^t k'(s)\nabla U(\mathcal{X}_s) ds\right) dt, \tag{2.2}$$

where $\sigma : \mathbb{R}^d \to \mathbb{M}_{d,d}$ is a continuous function, $(W_t)_{t\geq 0}$ is a d-dimensional standard Brownian motion and $(k(t))_{t\geq 0}$ is a deterministic positive increasing $\mathcal{C}^2$-function. Denoting by $(\mathcal{Y}_t)_{t\geq 0}$ the

process defined by

$$\mathcal{Y}_t = \frac{1}{k(t)} \int_0^t k'(s) \nabla U(\mathcal{X}_s) ds,$$

we observe that $d\mathcal{Y}_t = (k'/k)(t)(\nabla U(\mathcal{X}_t) - \mathcal{Y}_t)dt$. This means that SDE (2.2) can be viewed as a $2d$-dimensional non-homogeneous Markovian dynamical system given by the following SDE:

$$\begin{cases} dX_t = \sigma(X_t)dW_t - Y_t dt. \\ dY_t = r(t)(\nabla U(X_t) - Y_t)dt, \end{cases} \tag{2.3}$$

where $r(t) = \frac{k'}{k}(t)$ is a $\mathcal{C}^1$-function, $\sigma : \mathbb{R}^d \to \mathbb{M}_{d,d}$ is at least continuous and U satisfies (2.1). These assumptions will hold throughout the paper. Note that r is a non negative function on $\mathbb{R}_+$ owing to our assumption on k. We denote by $(Z_t)_{t\geq 0}$, the coupled general solution to (2.3): $Z_t = (X_t, Y_t)$, and by, $(Z_t^z)_{t\geq 0}$, the coupled solution starting from $z = (x, y)$ for $x, y \in \mathbb{R}^d$. Integrating by parts equations (2.3), one checks that:

$$Y_t^z = \frac{yk(0)}{k(t)} + \frac{1}{k(t)} \int_0^t k'(s) \nabla U(X_s^z) ds. \tag{2.4}$$

This means in particular that the previously defined process $(\mathcal{X}_t, \mathcal{Y}_t)_{t\geq 0}$ with $\mathcal{X}_0 = x$ corresponds to the solution of (2.3) starting from $z = (x, 0)$.

Under some classical conditions about existence and uniqueness of the solutions (see Subsection (2.1)), $(X_t, Y_t, t)_{t\geq 0}$ is an homogeneous Markov process whose infinitesimal generator $\mathcal{A}$ is defined for every $f \in \mathcal{C}_K^2(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_+)$, by:

$$\mathcal{A}f(x, y, t) = -\langle y, \nabla_x f\rangle + r(t)\langle \nabla U(x) - y, \nabla_y f\rangle + \frac{1}{2}\mathrm{Tr}\left(\sigma^*(x) D_x^2 f(x, y)\sigma(x)\right) + \partial_t f. \tag{2.5}$$

With a slight abuse of notation, in the particular case $r(t) = \lambda > 0$ for every $t \geq 0$, we will also denote by $\mathcal{A}$ the infinitesimal generator of the homogeneous Markov process $(X_t, Y_t)_{t\geq 0}$. Note also that in the proofs, we will write (X_t, Y_t) instead of (X_t^z, Y_t^z) in order to alleviate the notations.

2.1 Existence of solutions

First, let us state a result about existence of solutions for (2.3). In this way, we denote by $(\mathbf{H_0})$ the following growth assumption:

$(\mathbf{H_0})$: There exists $C > 0$ such that $\mathrm{Tr}\left[\sigma^*(x) D^2 U(x)\sigma(x)\right] \leq CU(x)$ for every $x \in \mathbb{R}^d$.

This assumption is satisfied for a very large class of potentials U (including potentials with non sublinear gradient). $(\mathbf{H_0})$ is true for potential U with asymptotic behaviour $U(x) \sim_\infty C_1|x|^p$ and $D^2U(x) \sim_\infty C_2|x|^{p-2}$ as soon as $\|\sigma(x)\| = O(|x|)$. It is even true for potential U with very weak growth: $U(x) \sim_\infty C_1 \ln|x|$ and $D^2U(x) \sim_\infty C_2|x|^{-2}$ as soon as $\|\sigma(x)\| = O(1 + |x|)$ also satisfies $(\mathbf{H_0})$.

Proposition 2.1 *Assume $(\mathbf{H_0})$, then strong existence holds for SDE (2.3). Moreover, if (X_0, Y_0) satisfies $\mathbb{E}[U(X_0) + |Y_0|^2] < +\infty$, then for every $T > 0$, $\sup_{t\in[0,T]} \mathbb{E}[U(X_t) + |Y_t|^2] < +\infty$.*

Remark 2.1 *Furthermore, if ∇U and σ are locally Lipschitz continuous functions, one checks classically that pathwise uniqueness also holds for (2.3).*

Proof : Consider $h : \mathbb{R}^{2d+1} \to \mathbb{R}_+$ defined by $h(x, y, t) = U(x) + |y|^2/(2r(t))$. Let $T > 0$. Then, for every $t \in [0, T]$, one checks that (2.5) applied to h implies

$$\mathcal{A}h(x, y, t) = \frac{1}{2}\mathrm{Tr}\left(\sigma^*(x) D^2 U(x)\sigma(x)\right) + |y|^2\left(-1 - \frac{r'(t)}{2r^2(t)}\right) \leq C_T h(x, y, t), \tag{2.6}$$

under $(\mathbf{H_0})$. Then, a classical Picard iteration leads to the strong existence of the solutions. Now, let $(X_t, Y_t)_{t\geq 0}$ be a solution of (2.3) starting from (X_0, Y_0) with $\mathbb{E}[U(X_0) + Y_0^2] < +\infty$. Then, Itô formula yields:

$$h(X_t, Y_t, t) = h(X_0, Y_0, 0) + \int_0^t \mathcal{A}h(X_s, Y_s, s)ds + M_t,$$

where $(M_t)_{t\geq 0}$ is the local martingale defined by

$$M_t := \int_0^t \langle \nabla U(X_s), \sigma(X_s)dW_s \rangle = \sum_{i,j} \int_0^t \partial_{x_i} U(X_s)\sigma_{i,j}(X_s)dW_s^j.$$

Let $(T_n)_{n\geq 1}$ denote an increasing sequence of stopping times such that $(M_{t\wedge T_n})_{t\geq 0}$ is a martingale for every $n \geq 1$. Using Fatou's lemma and the monotone convergence Theorem, we deduce from (2.6) that

$$\mathbb{E}[h(X_t, Y_t, t)] \leq \mathbb{E}[h(X_0, Y_0, 0)] + C_T \int_0^t \mathbb{E}[h(X_s, Y_s, s)]ds, \quad \forall t \leq T. \tag{2.7}$$

Now, $\mathbb{E}[h(X_0, Y_0, 0)] < +\infty$ and the second result follows from the Gronwall lemma. □

In the proof of the previous proposition, we observe that the function h leads to a finite-time control of the behaviour of $(Z_t)_{t\geq 0}$ but, owing to (2.6), it appears that this function will not be adapted for the study of the long-time stability of the dynamical system because there is only a mean-repelling effect for the second component Y. In other words, one can say that h is not a Lyapunov function for (2.3). In order to generate a mean-repelling effect for the first component X, we will have to consider a more complex function V that will be introduced in Proposition 3.1.

2.2 Density with respect to the Lebesgue measure

In this part, we focus on the smoothness of the semi-group associated with the homogeneous Markov process $(X_t, Y_t, t)_{t\geq 0}$ and deduce a uniqueness property of the stationary distribution of the Markov process $(X_t, Y_t)_{t\geq 0}$ in the homogeneous case $r(t) = \lambda$ for every $t \geq 0$. Let us first remind some tool of hypoelliptic theory for inhomogeneous Markov stochastic processes described in Cattiaux and Mesnager (2002) and Chaleyat-Maurel and Michel (1984). Note that in some special cases, r may not depend on time t and the coupled Markov process may be homogeneous. These very special cases occur only for exponential memory terms $k(t) = \lambda_1 e^{\lambda_2 t}$, $(\lambda_1, \lambda_2) \in \mathbb{R}_+^2$. In the sequel, we will avoid any distinction between the homogeneous and inhomogeneous setting and treat directly the general inhomogeneous case.

We first state some elementary notations for the vector fields which govern our equation (2.3). As the process may be inhomogeneous, these vector fields depend on the three variables (t, x, y). We denote by $\sigma_1, \dots \sigma_d$ the vector fields defined as

$$\forall j \in \{1 \dots d\} : \qquad \sigma_j(x) = \sum_{i=1}^{d} \sigma_j^i(x)\partial_{x_i}. \tag{2.8}$$

We also introduce the drift vector field L_D defined by

$$L_D(t, x, y) = -\langle y, \nabla_x \rangle + r(t)\langle \nabla U(x) - y, \nabla_y \rangle,$$

as well as the diffusion one:

$$L_\sigma(x)(f) = \frac{1}{2} \sum_{j=1}^{d} \langle \nabla_x(\sigma_j)(x), \sigma_j(x)(f) \rangle.$$

Following the convention of Cattiaux and Mesnager (2002), we define the vector field L_Z as

$$L_Z(t,x,y) = L_D(t,x,y) - L_\sigma(x).$$

If $A_1, \dots A_p$ are a set of p vector fields, we denote *Span Lie*$(A_1, \dots, A_p)$ the Lie algebra generated by the Lie bracket of vector fields $[A_i, A_j], [A_i, [A_j, A_k]], [A_i, [A_j, [A_k, A_l]]] \dots$

Let us define $\mathcal{E}_U$ as

$$\mathcal{E}_U = \Big\{ x \in \mathbb{R}^d,\ \det\left(D^2U(x)\right) \neq 0 \Big\}, \tag{2.9}$$

and $\mathcal{M}_U$ the complementary manifold $\mathcal{M}_U = \mathbb{R}^d \setminus \mathcal{E}_U$. We next state two hypothesis needed to obtain hypoellipticity of the process.

$(\mathbf{I_1})$: σ and U are $\mathcal{C}^\infty$ and there exists $\varepsilon_0 > 0$ such that $\sigma\sigma^* \geq \varepsilon_0 \mathrm{Id}$, (uniformly elliptic on $\mathbb{R}^d$).

$(\mathbf{I_2})$: $\dim(\mathcal{M}_U) \leq d-1$.

We are now able to state the following theorem whose proof is deferred to Subsection 4.1.

Theorem 2.1 *Assume* $(\mathbf{I_1})$ *and* $(\mathbf{I_2})$. *Then, for any* z, *the process* $(X_t^z, Y_t^z)_{t\geq 0}$ *is hypoelliptic and for any* $z \in \mathbb{R}^d \times \mathbb{R}^d$ *and* $t > 0$, *the density* $p_t(z,.)$ *of* $(X_t^z, Y_t^z)_{t\geq 0}$ *(w.r.t. the Lebesgue measure on* $\mathbb{R}^d \times \mathbb{R}^d$*) is* $\mathcal{C}^\infty$. *Furthermore, if* $\lim_{|x|\to+\infty} \frac{U(x)}{|x|} > 0$, *then for every* $z \in \mathbb{R}^d \times \mathbb{R}^d$, $\mathrm{Supp} P_t(z,.) = \mathbb{R}^d \times \mathbb{R}^d$ *and when* $r(t) = r_\infty > 0$ *for every* $t \geq 0$, *there is at most one invariant distribution for the homogeneous Markov process* $(X_t^z, Y_t^z)_{t\geq 0}$.

Remark 2.2 *1. Assumption* $(\mathbf{I_1})$ *is somewhat more classical and the uniform ellipticity of* σ *seems necessary. Let us briefly discuss on the technical hypothesis* $(\mathbf{I_2})$. *It is possible to state some less restrictive condition. We denote* $\partial_{i_1,\dots i_p} U(x) = \partial_{x_{i_1}} \partial_{x_{i_2}} \dots \partial_{x_{i_p}} U(x)$ *and we define* $\tilde{\mathcal{E}}_U$ *the set of* $x \in \mathbb{R}^d$ *such that there exists* d *finite sequence* $s_1 := (i_{1,1}, \dots i_{1,p_1}), \dots, s_d := (i_{d,1} \dots i_{d,p_d})$ *for which the matrix*

$$M_{s_1,\dots,s_d}(x) = \begin{pmatrix} \partial^{p_1+1}_{s_1,1} U(x) & \dots & \partial^{p_1+1}_{s_1,j} U(x) & \dots & \partial^{p_1+1}_{s_1,d} U(x) \\ \partial^{p_2+1}_{s_2,1} U(x) & \dots & \partial^{p_2+1}_{s_2,j} U(x) & \dots & \partial^{p_2+1}_{s_2,d} U(x) \\ \vdots & & \partial^{p_i+1}_{s_i,j} U(x) & & \partial^{p_i+1}_{s_i,d} U(x) \\ \partial^{p_d+1}_{s_d,1} U(x) & \dots & \partial^{p_d+1}_{s_d,j} U(x) & \dots & \partial^{p_d+1}_{s_d,d} U(x) \end{pmatrix}$$

is invertible. Indeed, $\tilde{\mathcal{E}}_U$ *corresponds to the special case* $s_1 = \{1\}, \dots s_d = \{d\}$ *in the above definition and thus* $\mathcal{E}_U \subset \tilde{\mathcal{E}}_U$. *Assumption* $(\mathbf{I_2})$ *may be replaced by the less restrictive one on* $\tilde{\mathcal{E}}_U$*:*

$(\tilde{\mathbf{I}}_\mathbf{2})$: $\dim(\mathbb{R}^d \setminus \tilde{\mathcal{E}}_U) \leq d-1$.

If we set $d = 1$ *and assume that* σ *is constant, the condition* $(\tilde{\mathbf{I}}_\mathbf{2})$ *states that the set of points* x *where all the derivatives of the potential* U *are vanishing is Lebesgue-negligible.*

2. Our theorem provides some smoothness properties of $z \mapsto p_t(z_0, z)$. *As concerns* $z_0 \mapsto p_t(z_0, z)$, *it seems that under some polynomial growth assumptions for the vector fields, such properties could be obtained using some Malliavin calculus arguments (see Hairer (2011) for instance) but we will not focus on this point in the sequel.*

3. The fact that $\mathrm{Supp}(P_t(z_0,.)) = \mathbb{R}^d \times \mathbb{R}^d$ *(for every* $z_0 \in \mathbb{R}^{2d}$*) and the uniqueness of the invariant distribution are proved in Lemma 4.2. The proof is strongly based on the surjectivity of* $x \mapsto \nabla U(x)$ *and* $\lim_{|x|\to+\infty} \frac{U(x)}{|x|} = +\infty$ *is a convenient assumption to ensure this property (see proof of Lemma 4.2 for details). Note that when* $x \mapsto \nabla U(x)$ *is bounded on* $\mathbb{R}^d$, $(Y_t)_{t\geq 0}$ *is a bounded process (see* (2.4)*) and thus,* $\mathrm{Supp}(P_t(z,.)) \neq \mathbb{R}^d \times \mathbb{R}^d$ *in this case.*

3 Asymptotic behaviour

We now focus on the main objective of this paper: the study of the ergodic properties of the process solution to (2.3), these properties strongly rely on the asymptotic behaviour of $t \mapsto r(t)$. In this way, we set $r_\infty = \liminf_{t\to+\infty} r(t)$ and divide this section into two parts corresponding to the cases $r_\infty > 0$ and $r_\infty = 0$ respectively.

3.1 The stable case: $r_\infty > 0$:

First, note that $r_\infty > 0$ occurs in the two following cases:

- $k(t) = \exp(\lambda t)$: in this case, $r(t) = r_\infty = \lambda$ and $(X_t^z, Y_t^z)_{t\geq 0}$ is an homogeneous Markov process.
- $k(t) = \exp(t^\alpha)$ with $\alpha > 1$: in this case, $r_\infty = \lim_{t\to+\infty} r(t) = +\infty$.

Even if a part of the main results about the asymptotic behaviour of the process is stated together, the reader has to keep in mind that there is an important difference for the two previous cases. Under some mean-repelling assumptions, we will show in particular, that in the first case, the stochastic process has its own stationary regime while in the second case, the non-homogeneous Markov process has some convergence properties to the stationary regime of the memoryless stochastic differential equation

$$dS_t = -\nabla U(S_t)dt + \sigma(S_t)dW_t, \tag{3.1}$$

whose infinitesimal generator $\mathcal{L}$ is defined for every $g \in \mathcal{C}_K^2(\mathbb{R}^d)$ by

$$\mathcal{L}g(x) = -\langle \nabla U(x), \nabla g(x)\rangle + \frac{1}{2}\mathrm{Tr}\left(\sigma(x)D^2 g(x)\sigma^*(x)\right). \tag{3.2}$$

Let us now introduce a Lyapunov-type stability assumption $(\mathbf{H_1})$ and an assumption on the asymptotic behaviour of the function $t \mapsto r(t)$:

$(\mathbf{H_1})$: There exist $m \in (0, r_\infty)$ and $\varepsilon \in (0, r_\infty - m)$ such that

$$\limsup_{|x|\to+\infty} \left(-m\langle x, \nabla U(x)\rangle + \frac{1}{2}\mathrm{Tr}\left(\sigma^*(x)(D^2U(x) + (m+\varepsilon)I_d)\sigma(x)\right)\right) = -\infty.$$

Note that in the homogeneous case ($r(t) = r_\infty$ for every $t \geq 0$), we could take $\varepsilon = 0$.

$(\mathbf{R_1})$: $\frac{r'(t)}{r^2(t)} \xrightarrow{t\to+\infty} 0$.

Remark 3.1 *Condition $(\mathbf{H_1})$ is not restrictive and is satisfied for a large class of potentials U. For instance, for constant covariance matrix σ, $(\mathbf{H_1})$ is satisfied for all potentials $U(x) \sim_{|x|\to+\infty} |x|^q$ as soon as $q > 0$. This is even true for all $U(x) \sim_{|x|\mapsto+\infty} \ln(|x|+1)^\beta$ for $\beta > 1$. For varying σ, $(\mathbf{H_1})$ is satisfied provided σ is not asymptotically too large:*

- *For asymptotic polynomial U: $U(x) \sim_{|x|\to+\infty} |x|^q$ with $q > 0$, $(\mathbf{H_1})$ is true if $\|\sigma(x)\sigma^*(x)\| = o(|x|^{q\wedge 2})$ as $|x| \to +\infty$.*
- *For asymptotic logarithmic U: $U(x) \sim_{|x|\to+\infty} \ln(|x|+1)^\beta$ with $\beta > 1$, $(\mathbf{H_1})$ is true if $\|\sigma(x)\sigma^*(x)\| = o(\ln(|x|+1)^{\beta-1})$ as $|x| \to +\infty$.*

As concerns $(\mathbf{R_1})$, note that this assumption is satisfied in the two cases mentioned before : $k(t) = \exp(\lambda t)$ and $k(t) = \exp(t^\alpha)$, $\alpha \geq 1$. Indeed, $(\mathbf{R_1})$ is true as soon as $\dot{k} \geq Ck$ for t large enough. This is generally true in our case $r_\infty > 0$.

The next proposition (whose proof is given in Subsection 4.2) establishes the existence of a Lyapunov function V for $(Z_t)_{t\geq 0} = (X_t, Y_t)_{t\geq 0}$. In this way, we need to introduce $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ defined by

$$\rho(t) = \left(\int_t^{+\infty} \frac{k(t)}{k(s)} ds \right)^{-1}. \tag{3.3}$$

Owing to Lemma 4.3, ρ is well-defined when $r_\infty \in (0, +\infty]$ and is a positive $\mathcal{C}^1$-solution to $\dot{u}(t) = u^2(t) - r(t)u(t)$ that satisfies $\rho(t) \sim r(t)$ as $t \to +\infty$.

Proposition 3.1 *Assume* $(\mathbf{H_1})$ *and* $(\mathbf{R_1})$ *and suppose that* $\liminf_{t\to+\infty} r(t) = r_\infty > 0$. *Set* $m_\varepsilon = m + \varepsilon$. *Then,* $V : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$ *defined by*

$$V(x, y, t) = U(x) + \frac{|y|^2}{2r(t)} + m_\varepsilon \left(\frac{|x|^2}{2} - \frac{\langle x, y \rangle}{\rho(t)} \right), \tag{3.4}$$

is a Lyapunov function for the SDE in the following sense: there exists $t_0 \geq 0$ *such that* V *is positive for every* $t \geq t_0$ *and,*

$$\limsup_{|(x,y)|\to+\infty} \left(\sup_{t\geq t_0} \mathcal{A}V(x, y, t) \right) = -\infty. \tag{3.5}$$

If moreover, $\limsup_{t\to+\infty} r(t) < +\infty$, *there exists* $t_1 > 0$ *such that*

$$\lim_{|(x,y)|\to+\infty} \left(\inf_{t\geq t_1} V(x, y, t) \right) = +\infty. \tag{3.6}$$

Remark 3.2 *The construction of this non trivial Lyapunov function is a key step for the sequel of the paper. The reader will remark a non classical point in the proof of this lemma: the mean-repelling effect of the first coordinate is generated by* $-\frac{\langle x,y \rangle}{\rho(t)}$.
We should note that this Lyapunov function is very similar to the one used in several works on the Vlasov-Fokker-Planck equation. For instance, a recent paper by Bolley, Guillin, and Malrieu (2010) (see also Bakry, Cattiaux, and Guillin (2008) and Wu (2001)) used a coupling argument with a Lyapunov function of the form $Q(x, y) = a|y|^2 + b\langle x, y \rangle + c|x|^2$ *to establish exponential rates of convergence to equilibrium of solutions of Vlasov-Fokker-Planck equations with respect to the Wasserstein distance. Conversely, Villani (2009) obtains lower bounds for the solutions of the kinetic Fokker-Planck equations using another function* $Q(t, x, y) = a(t)|y|^2 + b(t)\langle x, y \rangle + c(t)|x|^2$ *with a suitable choice of* a, b *and* c *(see theorem A.19 of Villani (2009)). Thus, in such coupled stochastic equation, the term implying* $\langle x, y \rangle$ *(or* $\langle x, v \rangle$ *with the standard notations of Fokker-Planck equations) seems to play a key role to obtain lower and upper bounds.*

3.1.1 Convergence properties of the occupation measures

For a fixed initial value $z = (x, y) \in \mathbb{R}^d \times \mathbb{R}^d$, we consider some families of occupation measures $(\nu_t^z(\omega, dx, dy))_{t\geq 1}$ and $(\mu_t^z(dx, dy))_{t\geq 1}$ respectively defined for every bounded continuous function $f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, for every $t \geq 0$ by:

$$\nu_t^z(\omega, f) = \frac{1}{t} \int_0^t f(X_s^z, Y_s^z) ds,$$

and by

$$\mu_t^z(f) = \frac{1}{t} \int_0^t \mathbb{E}[f(X_s^z, Y_s^z)] ds = \mathbb{E}[\nu_t^z(\omega, f)].$$

We first focus on the asymptotic behaviour of $(\mu_t^z)_{t\geq 0}$ and obtain the following result:

Theorem 3.1 *Assume that* $r_\infty \in \mathbb{R}_+^* \cup \{+\infty\}$. *Assume* $(\mathbf{H_0})$, $(\mathbf{H_1})$ *and* $(\mathbf{R_1})$. *Then, for every* $z = (x, y) \in \mathbb{R}^d \times \mathbb{R}^d$, *the family of probabilities* $(\mu_t^z)_{t\geq 1}$ *is tight on* $\mathbb{R}^d \times \mathbb{R}^d$. *Let* μ_∞ *denote an accumulation point of* $(\mu_t^z)_{t\geq 1}$ *when* $t \to +\infty$:

(i) If $r_\infty = +\infty$, the first marginal of μ_∞ is an invariant distribution for the stochastic differential equation (3.1).

(ii) If $r(t) \xrightarrow{t\to+\infty} r_\infty < +\infty$, μ_∞ is an invariant distribution of the homogeneous Markov process solution to (2.3) *with $r(t) = r_\infty, \forall t \geq 0$.*

Remark 3.3 *In particular, the second statement implies that under $(\mathbf{H_0})$ and $(\mathbf{H_1})$, existence holds for the invariant distribution in the homogeneous case.*

We now focus on the family of random occupation measures $(\nu_t^z(\omega, dx, dy))_{t\geq 1}$ for which we want to obtain a "quenched" version of Theorem 3.1. In this way, we need to introduce a little stronger assumption $(\mathbf{H'_a})$ (compared to $(\mathbf{H_1})$):

$(\mathbf{H'_a})$: There exists $a \in (0,1]$, $\beta \in \mathbb{R}$ and $\alpha > 0$ such that

$$(i) \quad -\langle x, \nabla U(x)\rangle \leq \beta - \alpha \left(U(x) \vee |x|^2\right)^a, \forall x \in \mathbb{R}^d$$

$$(ii) \quad (1 + \mathrm{Tr}(\sigma\sigma^*)(x)) \left(1 + \frac{|\nabla U(x)|^2}{U(x)} + \|D^2U(x)\| + |||D^3U(x)|||\right) \overset{|x|\to+\infty}{=} o((U(x) \vee |x|^2)^a).$$

Remark 3.4 *Assumption (i) is a mean-repelling assumption whose intensity depends on the parameter a. Assumption (ii) is a growth assumption that is essentially needed to control the part of the dynamical system that hampers the mean-repelling effect. Coming back to the examples of Remark 3.1, one checks that if U is a $\mathcal{C}^3$-function such that $U(x) = |x|^q$ $(q > 0)$ for $|x|$ large enough, Assumption $(\mathbf{H'_a})$ is fulfilled with $a = (q/2) \wedge 1$ if $\|\sigma(x)\sigma^*(x)\| = o(|x|^{q\wedge 2})$ as $|x| \to +\infty$. However, when $U(x) \sim \ln(1+|x|)^\beta$, one observes that $(\mathbf{H'_a})(i)$ is not satisfied, i.e. that the mean-repelling effect is too weak.*

Theorem 3.2 *Assume that $r_\infty \in \mathbb{R}_+^* \cup \{+\infty\}$. Assume $(\mathbf{H'_a})$ and $(\mathbf{R_1})$. Then, for every $z = (x,y) \in \mathbb{R}^d \times \mathbb{R}^d$, for every $p \geq 1$,*

$$\sup_{t\geq 1} \frac{1}{t} \int_0^t \left(\left(U(X_s^z) \vee |X_s^z|^2\right)^{p+a-1} + |Y_s^z|^{2p}\right) ds < +\infty \quad a.s. \tag{3.7}$$

In particular, the family of probabilities $(\nu_t^z)_{t\geq 1}$ is a.s. tight on $\mathbb{R}^d \times \mathbb{R}^d$. Let ν_∞ denote an accumulation point of $(\nu_t^z)_{t\geq 1}$ when $t \to +\infty$:

- *(i) If $r_\infty = +\infty$, $\nu_\infty(dx,dy) = \delta_{\nabla U(x)}(dy)\pi(dx)$ where π is a.s an invariant distribution for the stochastic differential equation (3.1).*
- *(ii) If $r(t) \xrightarrow{t\to+\infty} r_\infty \in \mathbb{R}_+^*$, ν_∞ is a.s an invariant distribution of the homogeneous Markov process solution to* (2.3) *with $r(t) = r_\infty, \forall t \geq 0$.*

Remark 3.5
- *Under the hypotheses of Theorem 2.1, uniqueness holds for the invariant distribution ν of* (2.3) *with $r(t) = r_\infty \in \mathbb{R}_+^*$, $\forall t \geq 0$. In this case, it follows from Theorems 3.1(ii) and 3.2(ii), that when $r(t) \to r_\infty \in \mathbb{R}_+^*$, for every bounded continuous function $f : \mathbb{R}^d \to \mathbb{R}$, $(\mu_t^z(f))_{t\geq 1}$ and $(\nu_t^z(\omega, f)_{t\geq 1}$ converge (a.s. in the second case) to $\nu(f)$ as $t \to +\infty$. Obviously, the same remark holds when $r_\infty = +\infty$ and when uniqueness holds for the invariant distribution of* (3.1).
  Furthermore, these convergence properties can be extended to non bounded continuous functions using (3.7) *and uniform integrability arguments.*
- *Note also that the condition on D^3U in $(\mathbf{H'_a})$ is only necessary for the identification of ν_∞ when $r_\infty = +\infty$ (for more details, see the proof of Proposition 4.2).*

3.1.2 Properties of the "invariant distribution" and convergence rates

We will only talk of invariant distributions within the classical homogeneous case when $k(t) = \exp(r_\infty t)$. In the non-homogeneous setting, we will be interested in the set of accumulation points of mean occupation measures $(\mu_t^z)_{t\geq 0, z\in\mathbb{R}^d\times\mathbb{R}^d}$.
In the next proposition, we focus on the homogeneous case and provide some properties of the invariant distribution unde Assumptions $(\mathbf{I_1})$ and $(\mathbf{I_2})$ introduced for Theorem 2.1.

Proposition 3.2 *Assume* $(\mathbf{H_0})$, $(\mathbf{H_1})$ *and* $r(t) = r_\infty \in \mathbb{R}_+^*$. *Assume* $(\mathbf{I_1})$, $(\mathbf{I_2})$ *and* $\lim_{|x|\to+\infty} \frac{U(x)}{|x|} = +\infty$. *Then, there exists a unique invariant distribution* ν *satisfying the following properties:*

- *(i)* ν *is absolutely continuous w.r.t. the Lebesgue measure. Let* p_{r_∞} *denote the associated* $\mathcal{C}^\infty(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}_+)$ *density. Then,* p_{r_∞} *is the unique non-negative solution to*

$$\langle y, \nabla_x p_{r_\infty}\rangle + \frac{1}{2}\mathrm{Tr}\left(\sigma^* D_x^2 p_{r_\infty}\sigma)\right) + r_\infty\left[\langle y - \nabla U(x), \nabla_y p_{r_\infty}\rangle + p_{r_\infty}\right] = 0. \tag{3.8}$$

  which satisfies $\int_{\mathbb{R}^{2d}} p_{r_\infty}(x,y)dxdy = 1$.

- *(ii) Assume* $d = 1$, $U(x) = x^2/2$, $\sigma(x) = \sigma > 0\ \forall x \in \mathbb{R}$, *and* $r(t) = r_\infty \in]0;+\infty[$. *Then,* p_{r_∞} *is the centered Gaussian measure on* $\mathbb{R}\times\mathbb{R}$ *whose covariance matrix is given by*

$$\Sigma^2(r_\infty) = \frac{\sigma^2}{2}\begin{pmatrix} \frac{r_\infty+1}{r_\infty} & 1 \\ 1 & 1\end{pmatrix}.$$

Remark 3.6 *The proof of* (3.8) *is based on the fact that* $\int \mathcal{A}g(x,y)p_{r_\infty}(x,y)\lambda_{2d}(dx,dy) = 0$ *for every* $g \in \mathcal{C}_K^2(\mathbb{R}^d\times\mathbb{R}^d)$. *Extending and applying this identity to the non-bounded particular functions* $f(x,y) = x$ *and* $f(x,y) = y$, *one obtains the following simple properties:*

$$\int y p_{r_\infty}(x,y)\lambda_{2d}(dx,dy) = 0 \quad \text{and} \quad \int \nabla U(x) p_{r_\infty}(x,y)\lambda_{2d}(dx,dy) = 0.$$

As concerns (ii), *remark that when* $r_\infty \to +\infty$, *the limit variance of* $(X_t^z)_{t\geq 0}$ *is equal to* $\sigma^2/2$. *We recover here the standard variance of the Ornstein-Uhlenbeck process that corresponds to the non-memory case. Note also that when* r_∞ *decreases, the limit variance increases. This means that more the process remembers the past, less the dynamical system is long-time stable. This point will be emphasized in the next subsection in which we focus on the case* $r_\infty = 0$.

We now want to state some results about the convergence in distribution as $t \to +\infty$.
Let us first focus on the homogeneous case $r(t) = r_\infty$ where we derive some controls of the distance in total variation between the semi-group $P_t^\lambda((x,y),.)$ (associated $(X_t, Y_t)_{t\geq 0}$) and the invariant distribution from "Meyn-Tweedie"-type techniques.

Theorem 3.3 *Assume* $r(t) = r_\infty > 0$, $(\mathbf{H_a'})$ $(a \in (0,1])$, $(\mathbf{I_1})$ $(\mathbf{I_2})$, $\lim_{|x|\to+\infty} \frac{U(x)}{|x|} = +\infty$. *Assume that there exists a local minimum* x^* *of* U *such that* $D^2U(x^*) \neq 0$. *Then, for every* $p \geq 1$ *and for any* $t \geq 0$:

$$\sup_{\{f,|f|\leq 1\}} |P_t^{r_\infty}(z_0, f) - \nu(f)| \leq C_{a,p,r_\infty} V_\infty^p(z_0) \begin{cases} \exp(-\gamma_{p,r_\infty} t) & \text{if } a = 1 \\ t^{-\frac{p+a-1}{1-a}} & \text{if } a \in (0,1). \end{cases}$$

where $z = (x,y)$, V_∞ *is a positive function defined by* $V_\infty(z) = U(x) + \frac{r_\infty}{2}\left|x - \frac{y}{r_\infty}\right|^2$, γ_{p,r_∞} *and* C_{a,p,r_∞} *are some positive constants which do not depend on* z_0 *and* t.

Proof : The proofs in cases $a = 1$ and $a < 1$ rely respectively on Theorem 5.2 of Down, Meyn, and Tweedie (1995) and Theorem 3.10 of Douc, Fort, and Guillin (2009). The first assumption to check is that compact sets are *petite*, *i.e.* to show that for every compact set K of $\mathbb{R}^{2d}$, there

exist a probability a on $\mathbb{R}_+$ and a non-trivial σ-finite measure ν_a on $\mathcal{B}(\mathbb{R}^{2d})$ such that for every $z_0 \in K$, $\int_{t\geq 0} P_t(z_0,.)a(dt) \geq \nu_a(.)$. This point is a straightforward consequence of Lemma 4.2(ii). Following the assumptions of Down et al. (1995) (when $a = 1$) and Douc et al. (2009) (when $a < 1$), we want now to prove that there exists a compact set C and some suitable positive $\tilde{\alpha}$ and $\tilde{\beta}$ such that

$$\mathcal{A}V_\infty^p \leq \tilde{\beta} - \tilde{\alpha} V_\infty^{p+a-1}. \tag{3.9}$$

First, by Inequality (4.23) (applied with $r(t) = r_\infty$), there exist some positive α_1 and β_1 such that

$$\mathcal{A}V_\infty^p \leq \beta_1 - \alpha_1[(U(x) \vee |x|^2)^{p+a-1} + |y|^{2p}].$$

Second, one checks easily that $V_\infty(x) \leq C(1 + U(x) \vee |x|^2 + |y|^2)$. Thus, using that $|y|^{2p} \geq (|y|^{2(p+a-1)} - 1)$ and the elementary inequality $|u+v|^p \leq c_p(|u|^p + |v|^p)$, it follows that there exist some positive α_2 and β_2 such that,

$$\mathcal{A}V_\infty^p \leq \beta_2 - \alpha_2 V_\infty^{p+a-1}(x,y) \quad \forall x, y \in \mathbb{R}^d.$$

Finally, the fact that $\lim V_\infty(x,y) = +\infty$ as $|(x,y)| \to +\infty$ (see (3.6)) implies that we can build a compact set K such that $\beta_2 \leq \beta_2 1_K + \frac{\alpha_2}{2} V_\infty^{p+a-1}(x,y)$ for every $x, y \in \mathbb{R}^d$. Thus, we deduce (3.9) (with $\tilde{\beta} = \beta_2$ and $\tilde{\alpha} = \frac{\alpha_2}{2}$) and this concludes the proof when $a = 1$ owing to Theorem 5.2 of Down et al. (1995). When $a < 1$, we remark that (3.9) can be written

$$\mathcal{A}V_\infty^p \leq \tilde{\beta} 1\!\!1_C - \phi(V_\infty^p)$$

with $\phi(u) = \tilde{\alpha} u^{\frac{p+a-1}{p}}$. A simple computation shows that

$$H_\phi(u) := \int_1^u \frac{ds}{\phi(s)} = \frac{p}{\tilde{\alpha}(1-a)} \left[u^{\frac{1-a}{p}} - 1 \right].$$

Thus, a simple equivalent to $r_\star(s) := \phi \circ H_\phi^{-1}(s)$ is given by

$$r_\star(s) \sim \frac{\tilde{\alpha}^{p/(1-a)} (1-a)^{\frac{p+a-1}{1-a}}}{p^{\frac{p+a-1}{1-a}}} s^{\frac{p+a-1}{1-a}} \quad \text{as} \quad s \to +\infty,$$

The second result follows applying (3.6) of Douc et al. (2009) with $\Psi_1 = \mathrm{Id}$ and $\Psi_2 = 1$. □

Finally, we focus on the non-homogeneous case $r(t) \to +\infty$. In this case, by Theorems 3.1 and 3.2, it seems that $(X_t)_{t\geq 0}$ has some convergence properties to the invariant distribution π of the classical diffusion $dS_t = -\nabla U(S_t)dt + \sigma(S_t)dW_t$. In the following theorem, we derive a result about the Wasserstein distance between $\mathbb{P}_{X_t}$ and π from a coupling with the classical diffusion $(S_t)_{t\geq 0}$ under the additional following Assumption:

(**AC**) : There exists $\rho > 0$ such that for every $x_1, x_2 \in \mathbb{R}^d$,

$$\langle x_1 - x_2, \nabla U(x_2) - \nabla U(x_1) \rangle + \frac{1}{2} \mathrm{Tr}\left((\tilde{\sigma}\tilde{\sigma}^*)(x_1, x_2) \right) \leq -\rho |x_2 - x_1|^2. \tag{3.10}$$

where $\tilde{\sigma}(x,y) = \sigma(y) - \sigma(x)$.

Remark 3.7 *Note that under* (**AC**)*, existence and uniqueness hold for the invariant distribution π of the classical diffusion $dS_t = -\nabla U(S_t)dt + \sigma(S_t)dW_t$ (see e.g. Down et al. (1995)). Here, this assumption will have two roles: we will use it classically to control the (Wasserstein) distance between the distribution of $(S_t^x)_{t\geq 0}$ and π but also to control the L^2-distance between the trajectory of $(S_t^x)_{t\geq 0}$ and $(X_t^x)_{t\geq 0}$ (built with the same Brownian Motion).*

For any $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$, we recall that the 2^{th}-Wasserstein distance is defined by $W_2(\mu,\nu) = \inf \mathbb{E}[|X-Y|^2]^{\frac{1}{2}}$ where the infimum is taken over every couple (X,Y) with marginal distribution μ and ν respectively.

Theorem 3.4 *Assume $r_\infty = +\infty$, $(\mathbf{H}_1')$, $(\mathbf{AC})$ and that $t \to \int_0^t k(s)ds/k(t)$ is a non-increasing function that tends to 0 as $t \to +\infty$. Suppose that ∇U and σ are locally lispchitz functions and denote by π the unique invariant distribution of* (3.1)*. Then, there exist some positive C_1 and C_2 such that for every positive t,*

$$W_2(\mathbb{P}_{X_t}, \pi) \leq C_1 e^{-\rho(t-u)} + \frac{C_2}{k(u)} \int_0^u k(v)dv, \quad \forall u \in [0,t].$$

In particular, $W_2(\mathbb{P}_{X_t}, \pi) \Longrightarrow 0$ as $t \to +\infty$.

Owing to an optimization of the preceding inequality, the next corollary follows.

Corollary 3.1 *Under the assumptions of Theorem 3.4,*

$$W_2(\mathbb{P}_{X_t}, \pi) \leq C_1 \left(1 + \frac{\rho}{r \circ H^{-1}(t)}\right) e^{-\rho(t - H^{-1}(t))} + \frac{C_2}{r \circ H^{-1}(t)}$$

where H is defined by:

$$H(u) = u - \frac{1}{\rho} \log \left[\frac{C_2}{C_1 \rho} \left(\frac{r(u)}{k(u)} \int_0^u k(v)dv + \frac{k(0)}{k(u)} - 1 \right) \right].$$

Remark 3.8 *We can provide some explicit bound of W_2 for some specific memory functions k.*

- *If $k(t) = \alpha t^{\alpha-1} e^{t^\alpha}$ for $\alpha > 1$, then $r(t) \sim \alpha t^{2\alpha-2}$ and $H^{-1}(t) \sim t$ whereas*

  $$t - H^{-1}(t) \sim \frac{\alpha}{\rho} \log t,$$

  and the bound obtained is

  $$W_2(\mathbb{P}_{X_t}, \pi) \leq C t^{1-\alpha}$$

  for a suitable constant $C > 0$.

- *If $k(t) = e^t e^{e^t}$, then $H^{-1}(t) \sim \frac{\rho}{\rho+1} t$ and for a suitable constant $C > 0$*

  $$W_2(\mathbb{P}_{X_t}, \pi) \leq C e^{-\frac{\rho}{\rho+1} t}.$$

3.2 The non-stable case: $r_\infty = 0$

In this part, we focus on the long memory case: $r(t) \to 0$. For instance, one can think to $k(t) = (1+t)^\alpha$ with $\alpha > 0$ or to $k(t) = e^{(1+t)^\alpha}, \alpha \in (0;1)$. We prove first that in this case, the dynamical system is not stable for subquadratic case.

3.2.1 Subquadratic case

We show that when U has at most quadratic growth and the diffusion part is not degenerated (in a sense being precised below), the dynamical system has a tendency to explode.

Theorem 3.5 *Assume that $|\nabla U|^2 \leq C(1+U)$ and that there exists $\lambda_0 > 0$ such that $\mathrm{Tr}(\sigma^* D^2 U \sigma)(x) \geq \lambda_0 > 0$. Assume that $r(t) \to 0$ and that one can find $t_0 \geq 0$ such that $r'(t) + 2r^2(t) \geq 0$ for every $t \geq t_0$. Then, for every initial value $z = (x,y)$,*

$$\limsup_{t \to +\infty} r(t) \mathbb{E}[|X_t^z|^2] > 0.$$

In particular, there exists a subsequence $(t_n)_{n \geq 1}$ such that $\mathbb{E}[|X_{t_n}^z|^2] \to +\infty$.

Remark 3.9 *One observes that the condition on $r(t)$ is satisfied as soon as $k(t) = e^{(1+t)^\alpha}$ with $\alpha \in (0;1)$ or $k(t) = (1+t)^\alpha$ with $\alpha > \frac{1}{2}$. In particular, it contains the "non- weighted" averaged case where $Y_t = \frac{1}{1+t} \int_0^t \nabla U(X_s)ds$.*

3.2.2 The quadratic case

In this second part of the long memory case, we want to specify the previous result in the very particular quadratic case. More precisely, we assume that $U(x) = x^2/2$, that $d = 1$ and that the memory is polynomial: $k(t) = (1+t)^\alpha$ and $r(t) = \alpha/(1+t)$. In fact, in this case, the long-time behaviour of the process is given by that of its covariance matrix. Setting $f(t) = \mathbb{E}[X_t^2]$, $g(t) = \mathbb{E}[Y_t^2]$ and $h(t) = \mathbb{E}[X_t Y_t]$, we derive from Itô formula that

$$(\mathcal{S}) \begin{cases} f'(t) = & 1 - 2h(t) \\ g'(t) = & 2r(t)[h(t) - g(t)] \\ h'(t) = & -g(t) + r(t)[f(t) - h(t)]. \end{cases}$$

Then, some sharp computations on this differential system (see Section 7) yield the following result:

Theorem 3.6 *Let $d = 1$, $U(x) = x^2/2$ and $k(t) = (1+t)^\alpha$ with $\alpha > 1/2$, we have:*
i) The process $(X_t, Y_t)_{t\geq 0}$ is asymptotically centered.
ii) The process $(X_t, Y_t)_{t\geq 0}$ satisfies

$$\lim_{t\to\infty} \mathbb{E}Y_t^2 = \frac{\alpha}{2\alpha+1}, \qquad \text{and} \qquad \mathbb{E}X_t^2 \sim \frac{t}{2\alpha+1} \quad \text{as} \quad t \to +\infty.$$

iii)

$$\left(\sqrt{\frac{2\alpha+1}{t}} X_t, \sqrt{\frac{2\alpha+1}{\alpha}} Y_t\right) \overset{\mathcal{L}}{\Longrightarrow} \mathcal{N}(0, I_2) \quad \text{as} \quad t \to +\infty.$$

4 Proofs of Hypoellipticity and convergence to steady regime

4.1 Proof of Theorem 2.1 (Hypoellipticity)

The first part of Theorem 2.1 is given by the conclusions of Lemma 4.1 and 4.2 whereas the uniqueness of the invariant distribution (in the homogeneous case) comes from Lemma 4.2.

Lemma 4.1 *Assume $(\mathbf{I_1})$ and $(\mathbf{I_2})$. Then, for any $z \in \mathbb{R}^d \times \mathbb{R}^d$ and any $t > 0$, $P_t(z, .)$ is absolutely continuous with respect to the Lebesgue measure over $\mathbb{R}^d \times \mathbb{R}^d$. Moreover, for any time $t > 0$ and every $z_0 \in \mathbb{R}^d \times \mathbb{R}^d$, $z \mapsto p_t(z_0, z)$ is $\mathcal{C}^\infty$ on $\mathbb{R}^d \times \mathbb{R}^d$.*

Proof : Recall the notation $Z_t = (X_t, Y_t)$ for a solution of Equation (2.3) and the definition of $\mathcal{E}_U$ and $\overline{\mathcal{M}_U}$ given by (2.9). Our coupled process can be written in an homogeneous way considering (Z_t, ξ_t) with ξ_t defined through $d\xi_t = dt$. The system is thus equivalent to

$$\begin{cases} dX_t &= -Y_t dt + \sigma(X_t) dW_t \\ dY_t &= r(\xi_t)(\nabla U(X_t) - Y_t) dt \\ d\xi_t &= dt \end{cases}$$

Fix any $z = Z_0$, we first show that for all t, $P_t(z, .)$ is absolutely continuous with respect to the Lebesgue measure. As x and y belong to $\mathbb{R}^d$, in order to apply the Hörmander's sum of square theorem (see *e.g.* Hormander (1967)), we must establish that

$$\forall (x, y, \xi) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_+ \qquad \dim \text{ span Lie} \left(\frac{\partial}{\partial \xi} + L_Z, \sigma_1, \ldots, \sigma_d\right)(x, y, \xi) = 2d + 1,$$

where $\sigma_1, \ldots, \sigma_d$ are defined by (2.8). Note that following the Kohn's argument (detailed in Kohn (1978); Trèves (1980)), no assumption on the growth of the vector fields is needed for

the existence $p_t(z_0,.)$ of the density with respect to the Lebesgue measure and the regularity of $z \mapsto p_t(z_0,z)$.

Remark first that since σ is uniformly elliptic over $\mathbb{R}^d$, we have the simplification:

$$\text{span } (\sigma_1,\dots\sigma_d) = \text{ span } (\partial_{x_1},\dots\partial_{x_d}).$$

Thus, we obtain

$$\begin{aligned}\text{Lie}\left(\frac{\partial}{\partial\xi}+L_Z,\sigma_1,\dots,\sigma_d\right) &= \text{Lie}\left(\frac{\partial}{\partial\xi}+L_D-L_\sigma,\sigma_1,\dots,\sigma_d\right)\\ &= \text{Lie}\left(\frac{\partial}{\partial\xi}+L_D-L_\sigma,\partial_{x_1},\dots,\partial_{x_d}\right)\\ &= \text{Lie}\left(\frac{\partial}{\partial\xi}+L_D,\partial_{x_1},\dots,\partial_{x_d}\right)\end{aligned}$$

The last equality is true since from the definition of the vector field L_σ, this vector field belongs to span $(\partial_{x_1},\dots\partial_{x_d})$. Thus, $\sigma_1,\dots\sigma_d$ trivially provides the $\frac{\partial}{\partial x}$ component (of dimension d) as well as $\frac{\partial}{\partial\xi}+L_D$ ensures the presence of $\frac{\partial}{\partial\xi}$ (of dimension 1). Thus it remains to obtain the $\frac{\partial}{\partial y}$ component (of dimension d). A simple computation yields for any $f\in\mathcal{C}^\infty(\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}_+)$:

$$\left[\partial_{x_i},\frac{\partial}{\partial\xi}+L_D\right](f) = r(\xi)\sum_{j=1}^d \partial^2_{x_i,x_j}U\partial_{y_j}(f).$$

Now, suppose $x\in\mathcal{E}_U$, then $D^2U(x)$ is invertible and it implies that

$$\forall(x,y,\xi)\in\left(\mathbb{R}^d\setminus\mathcal{M}_U\right)\times\mathbb{R}^d\times\mathbb{R}_+ \qquad \dim \text{ span Lie}\left(\frac{\partial}{\partial\xi}+L_Z,\sigma_1,\dots,\sigma_d\right)(x,y,\xi)=2d+1.$$

As the manifold $\mathcal{M}_U$ has a vanishing Lebesgue measure, it shows that for any time $t>0$, $P_t(z_0,.)$ is absolutely continuous with respect to the Lebesgue measure and $\mathcal{C}^\infty$ on $\mathbb{R}^d\times\mathbb{R}^d$. □

If we replace $(\mathbf{I_2})$ by $(\tilde{\mathbf{I}}_2)$, the same result holds. One just have to use a convenient set of sequences of multi-indices which depends on the $x\in\tilde{\mathcal{E}}_U$. These sequences $(s_1,\dots,s_d)$ are defined in Remark 2.2, and it is enough to compute the several bracketing rules

$$\left(\left[\partial_{s_i\cup\{i\}},\frac{\partial}{\partial\xi}+L_D\right]\right)_{i=1\dots d},$$

to obtain the hypo-elliptic condition

$$\forall(x,y,\xi)\in\left(\mathbb{R}^d\setminus\mathcal{M}_U\right)\times\mathbb{R}^d\times\mathbb{R}_+ \qquad \dim \text{ span Lie}\left(\frac{\partial}{\partial\xi}+L_Z,\sigma_1,\dots,\sigma_d\right)(x,y,\xi)=2d+1.$$

Lemma 4.2 *Assume* $(\mathbf{I_1})$, $(\mathbf{I_2})$ *and* $\lim_{|x|\to+\infty}\frac{U(x)}{|x|}=+\infty$.
(i) For any $T>0$, for every $z_0\in\mathbb{R}^d\times\mathbb{R}^d$ and any open set $\mathcal{O}\subset\mathbb{R}^d\times\mathbb{R}^d$, the transition kernel associated to a solution of (2.3) satisfies $P_T(z_0,\mathcal{O})>0$. As a consequence, for every $z_0\in\mathbb{R}^{2d}$, $\text{Supp}(P_T(z_0,.))=\mathbb{R}^d\times\mathbb{R}^d$ and there exists at most one invariant probability measure for (X_t,Y_t) in the homogeneous setting and when $r(t)\longmapsto r_\infty\in(0;+\infty)$.
(ii) Furthermore, if $r(t)=\lambda>0$ and if there exists a local minimum x^* of U such that $D^2U(x^*)$ is an invertible matrix, then, there exists $T>0$ such that for every compact set K of $\mathbb{R}^{2d}$, there exists $r_K>0$ and $\alpha(T,K)>0$ such that

$$\forall z\in K,\quad P_T(z,.)\geq\alpha(T,K)\lambda_{2d}(.\cap B(z^*,r_K))$$

with $z^*=(x^*,0)$.

Remark 4.1 *In fact, these properties are strongly linked to the controllability of the dynamical system. In the first part where we end the proof of Theorem 2.1, the result relies on an approximate controllability property of the dynamical system.*
The second one which is based on a local lower bound of the density of the semi-group involves some (local) exact controllability. This property is ensured by the nondegeneracy of $D^2U(x^*)$. Note that (ii) implies that compact sets are petite (see Down et al. (1995)), which is a fundamental point of the proof of Theorem 3.3.

Proof : Proof of (i) : Let $T > 0$. Our first objective is to show that the dynamical is *approximatively controllable*, *i.e.* that, for every $z_0 = (x_0, y_0)$ and $z_1 = (x_1, y_1) \in \mathbb{R}^d \times \mathbb{R}^d$, for every $\eta > 0$, there exists a continuous function $(u_\eta(t))_{t\in[0,T]}$ such that $(z^{u_\eta}(t))$ defined as the unique solution starting from z_0 to

$$\left\{ \begin{array}{rcl} \dot{x}(t) &=& -y(t) + \sigma(x(t))u_\eta(t) \\ \dot{y}(t) &=& r(t)(\nabla U(x(t)) - y(t)), \end{array} \right. \tag{4.1}$$

satisfies $|z^{u_\eta}(T) - z_1| \leq \eta$. In the sequel, when z_0, z_1 and η are fixed, we will say that $(z^{u_\eta}(t))$ is a solution to the (z_0, z_1, η)-controllability problem. Note that it is enough to prove this approximate controllability problem when $\sigma = I_d$. Indeed, if $(z^{u_\eta}(t))_{t\in[0,T]}$ is a solution to the problem with $\sigma = I_d$, then $(\tilde{z}^{\sigma^{-1}(x(t))u_\eta(t)})_{t\in[0,T]}$ is a solution to the corresponding problem of approximate controllability with $\sigma(.)$.
Then, let us assume that $\sigma = I_d$ and consider a trajectory $(x_\eta(t))_{t\geq 0}$ that belongs to $\mathcal{C}^\infty(\mathbb{R}_+, \mathbb{R}^d)$ such that

$$x_\eta(0) = x_0, \quad \dot{x}_\eta(0) = -y_0 \quad x_\eta(T) = x_1 \quad \text{and} \quad \dot{x}_\eta(T) = -y_1. \tag{4.2}$$

We deduce from (4.1) that for a function $(u_\eta(t))$, the process $(x_\eta(t), y_\eta(t))$ with $y_\eta(t) := u_\eta(t) - \dot{x}(t)$, is a solution to the (z_0, z_1, η) controllability problem if, $u_\eta(0) = 0$, $|u_\eta(T)| \leq \eta$ and

$$\forall t \in [0, T], \quad \dot{u}_\eta(t) + \lambda u_\eta(t) = \ddot{x}_\eta(t) + r(t)\dot{x}_\eta(t) + r(t)\nabla U(x_\eta(t)).$$

Solving the above differential equation with initial condition $u_\eta(0) = 0$, we obtain

$$\begin{aligned} u_\eta(t) &= \frac{1}{k(t)} \int_0^t k(s) \left(\ddot{x}_\eta(s) + r(s)\dot{x}_\eta(s) + r(s)\nabla U(x_\eta(s)) \right) ds \\ &= \frac{1}{k(t)} \int_0^t \overbrace{(k(s)\dot{x}_\eta(s))}^{\cdot} + \dot{k}(s)\nabla U(x_\eta(s))ds. \end{aligned}$$

In particular, we have

$$u_\eta(T) = \frac{y_0}{k(T)} - y_1 + \int_0^T \frac{k(s)}{k(T)} \nabla U(x_\eta(s))ds.$$

Thus, it is enough to show that there exists $(x_\eta(t))_{t\geq 0}$ satisfying (4.2) such that

$$\left| \frac{y_0}{k(T)} - y_1 + \int_0^T \frac{k(s)}{k(T)} \nabla U(x_\eta(s))ds \right| \leq \eta. \tag{4.3}$$

In this way, we first check that $x \mapsto \nabla U(x)$ is surjective on $\mathbb{R}^d$ using the Fenchel-Legendre transform. Actually, for every $v \in \mathbb{R}^d$, let F_v be defined for every $x \in \mathbb{R}^d$ by $F_v(x) = \langle x, v\rangle - U(x)$. Owing to the assumption $\lim_{|x|\to+\infty} U(x)/|x| = +\infty$, we have $\lim_{|x|\to+\infty} F_v(x) = -\infty$. As a consequence, the function F_v has a global maximum x_v. In particular, $\nabla F_v(x_v) = 0$ and thus, $v = \nabla U(x_v)$.
Then, set $v_0 = \frac{k(T)y_1 - y_0}{k(T) - k(0)}$ and let $x_{v_0} \in \mathbb{R}^d$ be such that $\nabla U(x_{v_0}) = v_0$. Note that for this choice, we have

$$\frac{y_0}{k(T)} - y_1 + \int_0^T \frac{k(s)}{k(T)} \nabla U(x_{v_0})ds = 0. \tag{4.4}$$

The idea is then to consider a trajectory $(x_\eta(t))_{t\in[0,T]}$ which spends almost all the time in x_{v_0}. In this way, let $\delta_\eta > 0$ and consider $(x_\eta(t))_{t\in[0,T]}$ be a $\mathcal{C}^2$ function which satisfies (4.2),

$$x_\eta(t) = x_{v_0} \text{ on } [\delta_\eta, T-\delta_\eta] \text{ and } \forall t \in [0,T] \quad |x(t)| \leq M := 1 + \max\{|x_0|, |x_1|, |x_{v_0}|\}. \tag{4.5}$$

Such a function clearly exists and owing to (4.4), the associated function u_η satisfies:

$$|u_\eta(T)| \leq C_T \sup_{|x|\leq M} |\nabla U(x)| \delta_\eta.$$

Choosing δ_η small enough yields the (z_0, z_1, η)-controllability.

Now, let $\mathcal{O}$ denote an open set of $\mathbb{R}^d \times \mathbb{R}^d$, choose $z_1 = (x_1, y_1)$ and $\eta > 0$ such that $\mathcal{B}(z_1, 2\eta) \subset \mathcal{O}$ and let $T > 0$. We want to show that for every $z_0 \in \mathbb{R}^d \times \mathbb{R}^d$, $P_T(z_0, \mathcal{O}) > 0$. Let $(z^{u_\eta}(t))_{t\in[0,T]}$ be a solution to the (z_0, z_1, η)-controllability problem, *i.e.* such that $|z^{u_\eta}(T) - z_1| \leq \eta$. Then, we can deduce that $P_T(z_0, \mathcal{O}) > 0$ if $\mathbb{P}(|Z_T^{z_0} - z^{u_\eta}(T)| < \eta) > 0$. This point is implied by the Support theorem of Stroock and Varadhan (1972) (see *e.g.* Ikeda and Watanabe (1981)). Now, under $(\mathbf{I_1})$ and $(\mathbf{I_2})$, we know that $P_t(z_0, .)$ has a density $p_t(z_0, .)$ with respect to the Lebesgue measure and that $(P_t)_{t\geq 0}$ is almost strong Feller in the following sense: for any $\Gamma \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d), z \longmapsto P_t(z, \Gamma)$ is continuous on $\mathcal{E}_U$, which is an open set whose complementary $\mathcal{M}_U$ has a vanishing Lebesgue measure. Then, owing to a straightforward adaptation of Proposition 4.1.1. of Da Prato and Zabczyk (1996), we obtain that for any $(z, z') \in (\mathbb{R}^d \times \mathbb{R}^d)^2$, for any postive t and T, $P_{t+T}(z, .)$ and $P_{t+T}(z', .)$ are equivalent. The uniqueness of the invariant distribution then follows from Doob's Theorem (see *e.g.* Da Prato and Zabczyk (1996), Theorem 4.2.1). This concludes the proof of (i).

Proof of (ii) : We first need to show that the system is *locally exactly controllable* near $z^* = (x^*, 0)$ in a sense made precise below.

Let us define

$$F(x,y) = \begin{pmatrix} -y \\ \lambda(\nabla U(x) - y) \end{pmatrix}, \qquad B = \begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad A = \begin{pmatrix} 0 & -I_d \\ \lambda D^2 U(x^*) & -\lambda I_d \end{pmatrix}. \tag{4.6}$$

Recall that we denote $z := (x, y)$, the linear system $\dot{z} = Az + B\sigma(x^*)u$ is called the linearized system of $\dot{z} = F(z) + \Sigma(x)u$ at z^*. In fact, since $\sigma(x^*)$ and $D^2U(x^*)$ are invertible matrices, one checks easily that $\text{Span}(B\sigma(x^*)u, A\Sigma(x^*)u, u \in \mathbb{R}^{2d}) = \mathbb{R}^{2d}$. As a consequence, the linearized system solves the so-called Kalman condition. Thus, it follows from Theorems 1.16 and 3.8 of (Coron, 2007) that the system $\dot{z} = F(z) + \sigma(x)u$ is locally exactly controllable at z^*. More formally, it means that for every positive T and ε, there exists $\eta > 0$ such that for every $(z_1, z_2) \in B(z^*, \eta)$,

$$\exists\, (z(t), u(t))_{t\in[0,T]} \quad \text{with} \quad \|u\|_{\infty,T} := \sup_{t\in[0,T]} |u(t)| \leq \varepsilon \quad \text{and} \quad \begin{cases} z(0) = z_1, z(T) = z_2 \\ \dot{z} = F(z) + B\sigma(x)u. \end{cases}$$

Owing to the exact controllability, the idea is now to obtain some lower bounds for the transition density around z^* by localizing the work of Delarue and Menozzi (2010) (see also Bally and Kohatsu-Higa (2010) for other results on the subject). In this view, we follow the notations and the numbering of Delarue and Menozzi (2010) and observe that our equation (2.3) is a particular case of (Delarue & Menozzi, 2010)(1.1) with $F_1(t,x,y) = -y$ and $F_2(t,x,y) = \nabla U(x) - y$.

In particular, since $D^2U(x^*)$ is invertible and $x \to D^2U(x)$ is continuous, since σ is uniformly elliptic and locally Lipschitz continuous, we can check that Assumption $(\mathbf{A})$ of (Delarue & Menozzi, 2010) is satisfied on a sufficiently small ball $B(z^*, \rho)$ but not obviously on the whole space $\mathbb{R}^d$. However, following carefully the proofs of Delarue and Menozzi (2010), we can check that, owing to the local exact controllability around z^*, the lower bound obtained in Theorem 1.1 of Delarue and Menozzi (2010) remains true if T, ε and η are small enough (see the appendix for details).

As a consequence, there exists $T > 0$, $\eta_T > 0$ and $C_T > 0$ such that for every $t \in (0, T]$, for every $z_1, z_2 \in B(z^*, \eta_T)$,

$$p_t(z_1, z_2) \geq C_T^{-1} t^{-2d} \exp(-C_T t |\mathbb{T}_t^{-1} (\theta_t(z_1) - z_2)|^2) \quad \text{where } \mathbb{T}_t = \begin{pmatrix} tI_d & 0 \\ 0 & t^2 I_d \end{pmatrix},$$

and $(\theta_t(z_1))_{t\geq 0}$ denotes the solution to $\dot{z} = F(z)$. It follows in particular that there exists positive C_T^1 and C_T^2 such that for every $z_1, z_2 \in B(z^*, \eta_T)$,

$$p_T(z_1, z_2) \geq C_T^1 \exp(-C_T^2 |\theta_T(z_1) - z_2|^2). \tag{4.7}$$

We denote in the sequel $\theta_t = (\theta_t^1, \theta_t^2)$ and remark that for every $t \geq 0$ and any trajectory initialised at z_1:

$$U(\theta_t^1(z_1)) + \frac{|\theta_t^2(z_1)|^2}{2\lambda} = U(x_1) + \frac{|y_1|^2}{2\lambda} - \int_0^t |\theta_s^2(z_1)|^2 ds \leq U(x_1) + \frac{|y_1|^2}{2\lambda}. \tag{4.8}$$

Hence, we deduce easily that

$$\sup_{z_1, z_2 \in B(z^*, \eta_T)} |\theta_T(z_1) - z_2|^2 < +\infty.$$

By (4.7), it follows that there exists $\alpha_T > 0$ such that for every $z_1, z_2 \in B(z^*, \eta_T)$,

$$p_T(z_1, z_2) \geq \alpha_T > 0.$$

We are now able to end the proof of (ii). Let B be a Borel subset of $\mathcal{O} := B(z^*, \eta_T)$. We have for every $z \in K$:

$$P_{2T}(z, B) \geq \int_B \int_{\mathcal{O}} p_T(z, z_1) p_T(z_1, z_2) \lambda_{2d}(dz_1) \lambda_{2d}(dz_2) \geq \alpha_T \lambda_{2d}(B) \inf_{z \in K} P_T(z, \mathcal{O}).$$

Then, it remains to show that $\inf_{z\in K} P_T(z, \mathcal{O}) > 0$ and this point follows again from controllability argument: denote by $(u_{\frac{\eta_T}{2}}^{(z,z^*)})_{t\in[0,T]}$ the control built in (i) that yields the $(z, z^*, \frac{\eta_T}{2})$-controlability. Since $z \in K$, we deduce from the construction of x_η defined in (4.5) that $u_{\frac{\eta_T}{2}}^{(z,z^*)}$ can be built such that $\sup_{z\in K} \int_0^T |u_{\frac{\eta_T}{2}}^{(z,z^*)}|^2 ds < +\infty$. As a consequence, it follows from the support Theorem that $\inf_{z\in K} \mathbb{P}(|Z_T^z - z_2| \leq \frac{\eta_T}{2}) > 0$. The result follows. □

4.2 Building the Lyapunov function

We first show a key lemma for the construction of the Lyapunov function in Proposition 3.1.

Lemma 4.3 *Assume* $(\mathbf{R_1})$ *and* $r_\infty > 0$. *Then,* $(\rho(t))_{t\geq 0}$ *defined by* (3.3) *is a positive* $\mathcal{C}^1$*-solution to the differential equation* $\dot{u}(t) = -r(t)u(t) + u^2(t)$ *which satisfies* $\frac{\rho(t)}{r(t)} \mapsto 1$ *as* $t \to +\infty$.

Proof : First, $(\rho(t))_{t\geq 0}$ is a positive solution to $\dot{u}(t) = -r(t)u(t) + u^2(t)$ on $\mathbb{R}_+$ if and only if $z(t) = \frac{1}{\rho(t)}$ is solution to $v'(t) - r(t)v(t) = -1$. The general solution is given by:

$$v(t) = e^{\int_0^t r(s)ds} \left(C - \int_0^t e^{-\int_0^s r(x)dx} ds \right),$$

with $C \in \mathbb{R}$. Since $r_\infty > 0$, we can set $C = \int_0^{+\infty} e^{-\int_0^s r(x)dx} ds$ and we obtain the following positive particular solution on $\mathbb{R}_+$:

$$z(t) = e^{\int_0^t r(s)ds} \int_t^{+\infty} e^{-\int_0^s r(x)dx} ds = \int_t^{+\infty} f_t(s) ds,$$

where $f_t(s) = e^{-\int_t^s r(x)dx} = k(t)/k(s)$. Since $f_t'(s) = -r(s)f_t(s)$, an integration by parts yields:

$$z(t) = \int_t^{+\infty} \frac{f_t'(s)}{r(s)} ds = \left[\frac{f_t(s)}{r(s)}\right]_t^{+\infty} + \int_t^{+\infty} \frac{f_t(s)r'(s)}{r^2(s)} ds = \frac{1}{r(t)} + \int_t^{+\infty} \frac{f_t(s)r'(s)}{r^2(s)} ds.$$

By Assumption $(\mathbf{R_1})$,

$$\int_t^{+\infty} \frac{f_t(s)r'(s)}{r^2(s)} = o(z(t)) \quad \text{as } t \to +\infty.$$

It follows that $r(t)z(t) \to 1$ as $t \mapsto +\infty$. Finally, $\rho = 1/z$ satisfies $\frac{\rho(t)}{r(t)} \to 1$ as $t \to +\infty$. This completes the proof. □

With the result of Lemma 4.3, one can choose a suitable ρ to build a Lyapunov functional V as pointed in Proposition 3.1 whose proof can be found below.

Proof of Proposition 3.1:

(i) First, one checks that V can be written:

$$V(x,y,t) = U(x) + \frac{|y|^2}{2}\left(\frac{1}{r(t)} - \frac{m_\varepsilon}{\rho^2(t)}\right) + \frac{m_\varepsilon}{2}\left|x - \frac{y}{\rho(t)}\right|^2. \tag{4.9}$$

Then, since $m_\varepsilon < r_\infty$ and $\frac{\rho(t)}{r(t)} \mapsto 1$ as $t \to +\infty$ we deduce that

$$\liminf_{t\to+\infty}\left(\frac{1}{r(t)} - \frac{m_\varepsilon}{\rho^2(t)}\right) \geq 0.$$

It follows that V is positive for t large enough. If moreover, $\limsup_{t\to+\infty} r(t) < +\infty$, we have

$$\liminf_{t\to+\infty}\left(\frac{1}{r(t)} - \frac{m_\varepsilon}{\rho^2(t)}\right) > 0,$$

and (3.6) follows.

Second,

$$\begin{aligned}\mathcal{A}V(x,y,t) =& -m_\varepsilon\frac{r(t)}{\rho(t)}\langle x, \nabla U(x)\rangle - |y|^2\left(1 - \frac{m_\varepsilon}{\rho(t)} + \frac{r'(t)}{2r^2(t)}\right)\\ &+ m_\varepsilon\langle x,y\rangle\left(-1 + \frac{r(t)}{\rho(t)} + \frac{\rho'(t)}{\rho^2(t)}\right) + \frac{1}{2}\mathrm{Tr}\left(\sigma^*(x)(D^2U(x) + m_\varepsilon I_d)\sigma(x)\right).\end{aligned}$$

On the one hand, ρ satisfies

$$-1 + \frac{r(t)}{\rho(t)} + \frac{\rho'(t)}{\rho^2(t)} = 0.$$

On the other hand, $m_\varepsilon \in (0, r_\infty)$. Thus, since $r'(t)/r^2(t) \to 0$ and $\rho(t) \sim r(t)$ as $t \to +\infty$,

$$\liminf_{t\to+\infty}\left(1 - \frac{m_\varepsilon}{\rho(t)} + \frac{r'(t)}{2r^2(t)}\right) > 0. \tag{4.10}$$

We deduce that there exist $t_1 \geq 0$ and $\alpha_1 > 0$ such that

$$\mathcal{A}V(x,y,t) \leq -m_\varepsilon\frac{r(t)}{\rho(t)}\langle x, \nabla U(x)\rangle + \frac{1}{2}\mathrm{Tr}\left(\sigma^*(x)(D^2U(x) + m_\varepsilon I_d)\sigma(x)\right) - \alpha_1|y|^2 \quad \forall t \geq t_1. \tag{4.11}$$

Now, since $m_\varepsilon > m$ and $\frac{r(t)}{\rho(t)} \to 1$, there exists $t_0 \geq t_1$ such that $m_\varepsilon\frac{r(t)}{\rho(t)} \geq m$ for every $t \geq t_0$. Using that $\liminf_{|x|\to+\infty}\langle x, \nabla U(x)\rangle > 0$, we deduce that there exists a compact subset K of $\mathbb{R}^d$ such that

$$-m_\varepsilon\frac{r(t)}{\rho(t)}\langle x, \nabla U(x)\rangle \leq -m\langle x, \nabla U(x)\rangle, \quad \forall x \in K^c.$$

Finally, using the Lyapunov stability assumption $(\mathbf{H_1})$ and (4.11), we obtain that for every $t \geq t_0$,

$$\limsup_{|(x,y)|\to+\infty} \mathcal{A}V(x,y,t) = -\infty.$$

This ends the proof. □

4.3 Proofs of Theorems 3.1 and 3.2

Both theorems rely on a first step concerning the tightness of $(\mu_t^z)_{t\geq 1}$ and $(\nu_t^z)_{t\geq 1}$. This step is detailed in Proposition 4.1. The proof of Proposition 4.1 requires some technical results detailed in Lemma 4.4 (which concerns exclusively the tightness of the stochastic occupation measures $(\nu_t^z)_{t\geq 1}$ and not $(\mu_t^z)_{t\geq 1}$). Next, the identification steps of Theorems 3.1 and 3.2 are provided by Proposition 4.2.

Lemma 4.4 *Assume* $(\mathbf{R_1})$ *and* $\liminf_{t\to+\infty} r(t) = r_\infty \in (0, +\infty]$.

(i) Let $a \in (0,1]$ *such that* $(\mathbf{H}'_\mathbf{a})$ *holds. Let* $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ *be a decreasing* $\mathcal{C}^1$ *function such that* $\int_0^{+\infty} \eta_s ds < +\infty$. *Then, for every* $z \in \mathbb{R}^d \times \mathbb{R}^d$, *for every* $p \geq 1$,

$$\int_0^{+\infty} \eta_s \left(\mathbb{E}[(U(X_s^z) \vee (X_s^z)^2)^{p+a-1}] + \frac{\mathbb{E}[|Y_s^z|^{2p}]}{r(s)^{p-1}} \right) ds < +\infty, \tag{4.12}$$

$$\sup_{t\geq 0} \eta_t \left[\left(U(X_t^z) \vee |X_t^z|^2\right)^p + |Y_t^z|^{2p} \right] < +\infty \quad a.s. \tag{4.13}$$

and,

$$\sup_{t\geq 0} \eta_t \mathbb{E}\left[\left(U(X_t^z) \vee |X_t^z|^2\right)^p + |Y_t^z|^{2p} \right] < +\infty. \tag{4.14}$$

(ii) Assume $(\mathbf{H}'_\mathbf{1})$. *Then, for every* $z \in \mathbb{R}^d \times \mathbb{R}^d$,

$$\sup_{t\geq 0} \mathbb{E}\left[\left(U(X_t^z) \vee |X_t^z|^2\right)^p + |Y_t^z|^{2p} \right] < +\infty. \tag{4.15}$$

Proof : We recall that throughout the paper, C denotes a positive constant whose value may change from line to line and that in the proofs, we write (X_t, Y_t) instead of (X_t^z, Y_t^z). Let V be defined by (3.4) and let $t_0 \geq 0$ such that (3.5) holds. Fix any $t \geq t_0$. By Itô formula, we have:

$$\eta_t V^p(X_t, Y_t, t) = \eta_{t_0} V^p(X_{t_0}, Y_{t_0}, t_0) + \int_{t_0}^t \left(\eta_s \mathcal{A}V^p(X_s, Y_s, s) + \eta_s' V^p(X_s, Y_s, s)\right) ds + (M_t - M_{t_0}), \tag{4.16}$$

where $(M_t)_{t\geq 0}$ is the local martingale defined by

$$M_t = \int_0^t p\eta_s V^{p-1}(X_s, Y_s, s) \langle \nabla U(X_s) + m_\varepsilon \left(X_s - \frac{Y_s}{\rho(s)} \right), \sigma(X_s) dW_s \rangle.$$

Setting $\sigma_i(V) = \sum_{k=1}^d (\nabla_x V)_k \sigma_i^k$, one checks that

$$\mathcal{A}V^p = pV^{p-1} \left(\mathcal{A}V + \frac{p-1}{2} \sum_{i=1}^d \frac{\sigma_i(V)}{V} \right).$$

Owing to (4.11), it follows that there exists $t_1 > 0$ and $\alpha_1 > 0$ such that

$$\begin{aligned} \frac{\mathcal{A}V^p}{pV^{p-1}}(x,y,t) \leq & -m_\varepsilon \frac{r(t)}{\rho(t)} \langle x, \nabla U(x) \rangle - \alpha_1 |y|^2 \\ & + \frac{1}{2} \left[\mathrm{Tr}\left(\sigma^*(x)(D^2U(x) + m_\varepsilon I_d)\sigma(x)\right) + (p-1) \sum_{i=1}^d \frac{\sigma_i(V)}{V}(x,y,t) \right]. \end{aligned} \tag{4.17}$$

First, using Assumption $(\mathbf{H}'_\mathbf{a})$ and the fact that $r(t)/\rho(t) \to 1$ as $t \to +\infty$, we obtain the existence of $\tilde{\alpha}_1 > 0$ and $\tilde{\beta}_1 \in \mathbb{R}$ such that,

$$-m_\varepsilon \frac{r(t)}{\rho(t)} \langle x, \nabla U(x) \rangle - \alpha_1 |y|^2 \leq \tilde{\beta}_1 - \tilde{\alpha}_1 \left((U(x) \vee |x|^2)^a + |y|^2 \right). \tag{4.18}$$

Second, we focus on (4.17).

$$\mathrm{Tr}\left(\sigma^*(x)(D^2U(x)+m_\varepsilon I_d)\sigma(x)\right)+(p-1)\sum_{i=1}^{d}\frac{\sigma_i(V)}{V}(x,y,t)$$
$$\leq C\mathrm{Tr}(\sigma\sigma^*)(x)\left(\frac{|\nabla_x V(x,y,t)|^2}{V(x,y,t)}+\|D^2U(x)\|+1\right).$$

Let us control the above right term. On the one hand, it follows from (4.9) that for t large enough,

$$V(x,y,t)\geq C\frac{|y|^2}{r(t)},$$

where C is a positive constant. As well, remark that V can be written as follows:

$$V(x,y,t)=U(x)+\frac{1}{2r(t)}\left|y-m_\varepsilon\frac{xr(t)}{\rho(t)}\right|^2+\frac{m_\varepsilon}{2}|x|^2\left(1-m_\varepsilon\frac{r(t)}{\rho^2(t)}\right).$$

Then, by Lemma 4.3 and the definition of m_ε, we deduce that $1-m_\varepsilon r(t)/\rho^2(t)\geq C>0$ for t large enough. It follows that $V(x,y,t)\geq C(U(x)+|x|^2)$ for t large enough. As a consequence,

$$V(x,y,t)\geq C\max\left\{U(x)+|x|^2;\frac{|y|^2}{r(t)}\right\}\quad\text{for } t \text{ large enough.}\tag{4.19}$$

On the other hand, one easily derives from (4.9) that

$$|\nabla_x V(x,y,t)|^2\leq C\left(|\nabla U(x)|^2+|x|^2+\frac{|y|^2}{\rho^2(t)}\right).$$

Thus, using that $\liminf_{t\to+\infty} r(t)/\rho^2(t)=1/r_\infty<+\infty$, it follows from the two previous inequalities that, for t large enough,

$$\mathrm{Tr}\left(\sigma^*(x)(D^2U(x)+m_\varepsilon I_d)\sigma(x)\right)+(p-1)\sum_{i=1}^{d}\frac{\sigma_i(V)}{V}(x,y,t)$$
$$\leq C\mathrm{Tr}(\sigma\sigma^*)(x)\left(1+\frac{|\nabla U(x)|^2}{U(x)}+\|D^2U(x)\|\right)\overset{|x|\to+\infty}{=}o((U(x)\vee|x|^2)^a),$$

by Assumption $(\mathbf{H}'_\mathbf{a})$. Then, we derive from (4.18) that for t large enough, one can find $\alpha_2>0$ and $\beta_2\in\mathbb{R}$ such that

$$\mathcal{A}V^p(x,y,t)\leq V^{p-1}(x,y,t)\left(\beta_2-\alpha_2\left((U(x)\vee|x|^2)^a+|y|^2\right)\right).\tag{4.20}$$

Now, by (4.9) and (4.19), there exist $0<C_1\leq C_2$ such that for t large enough,

$$C_1\left(U(x)\vee|x|^2\vee\frac{|y|^2}{r(t)}\right)\leq V(x,y,t)\leq C_2\left(U(x)\vee|x|^2\vee\frac{|y|^2}{r(t)}\right).\tag{4.21}$$

It is thus immediate to check that for any $\delta>0$, one can find a suitable β_3 such that

$$\beta_1V^{p-1}(x,y,t)\leq\delta V^{p-1}(x,y,t)\left((U(x)\vee|x|^2)^a+|y|^2\right)+\beta_3.$$

Hence, setting $\delta=\alpha_2/2$, we derive from equation (4.20) that:

$$\mathcal{A}V^p(x,y,t)\leq\beta_3-\frac{\alpha_2}{2}\left((U(x)\vee|x|^2)^a+|y|^2\right)V^{p-1}(x,y,t).\tag{4.22}$$

Now, we deduce from (4.21) that

$$V^{p-1}(x,y,t) \geq C\left([U(x)\vee|x|^2]^{p-1}\vee\frac{|y|^{2(p-1)}}{r(t)^{p-1}}\right)\geq\frac{C}{2}\left([U(x)\vee|x|^2]^{p-1}+\frac{|y|^{2(p-1)}}{r(t)^{p-1}}\right).$$

Consequently, for a suitable choice of $\tilde{\beta}\in\mathbb{R}_+$, $\tilde{\alpha}>0$ and $t_0\geq 0$, one can check that for every $t\geq t_0$, for every $x,y\in\mathbb{R}^d$,

$$\mathcal{A}V^p(x,y,t)\leq\tilde{\beta}-\tilde{\alpha}\left((U(x)\vee|x|^2)^{p+a-1}+\frac{|y|^{2p}}{r(t)^{p-1}}\right), \tag{4.23}$$

and that $V^p(x,y,t)$ is positive.
Owing to (4.23), we are going to prove (i) by exhibiting a non-negative supermartingale. This argument can be viewed as a continuous adaptation of Lemma 4 of Lamberton and Pagès (2003). Set $\psi_{a,p}(x,y,t)=(U(x)\vee|x|^2)^{p+a-1}+\frac{|y|^{2p}}{r(t)^{p-1}}$ and let $(G_t)_{t\geq t_0}$ be the non-negative process defined by

$$G_t=\eta_t V^p(X_t,Y_t,t)+\tilde{\alpha}\int_{t_0}^t\eta_s\psi_{a,p}(X_s,Y_s,s)ds+\tilde{\beta}\int_t^{+\infty}\eta_s ds\quad\forall t\geq t_0.$$

With the notations of (4.16), $\forall t\geq t_0$,

$$G_t=G_{t_0}+\int_{t_0}^t\eta_s\left[\mathcal{A}V^p(X_s,Y_s,s)+\tilde{\alpha}\psi_{a,p}(X_s,Y_s,s)-\tilde{\beta}\right]ds+\int_{t_0}^t\eta_s' V^p(X_s,Y_s,s)ds+(M_t-M_{t_0}). \tag{4.24}$$

Since $\eta'\leq 0$, it follows from (4.23) that

$$\eta_s\left[\mathcal{A}V^p(X_s,Y_s,s)+\tilde{\alpha}\psi_{a,p}(X_s,Y_s)-\tilde{\beta}\right]+\eta_s' V(X_s,Y_s,s)\leq 0,\quad\forall s\geq t_0.$$

Finally, checking that $\mathcal{A}V^p(x,y,t)\leq CV^p(x,y,t)$ for every $t\leq t_0$, for every $x,y\in\mathbb{R}^d$, we get by similar arguments as those developed in Proposition 2.1:

$$\sup_{t\leq t_0}\mathbb{E}[V^p(X_t,Y_t,t)]<+\infty. \tag{4.25}$$

In particular, $\mathbb{E}[V^p((X_{t_0},Y_{t_0},t_0)]<+\infty$ is finite and it follows that $(G_t)_{t\geq t_0}$ is a non-negative supermartingale. Thus, (G_t) is *a.s* convergent and $\sup_{t\geq t_0}\mathbb{E}[G_t]<+\infty$. As a consequence,

$$\sup_{t\geq t_0}\eta_t V^p(X_t,Y_t,t)<+\infty\ a.s.,\ \sup_{t\geq t_0}\eta_t\mathbb{E}[V^p(X_t,Y_t,t)]<+\infty,\ \int_{t_0}^{+\infty}\eta_s\mathbb{E}[\psi_{a,p}(X_s,Y_s)]ds<+\infty. \tag{4.26}$$

Now, by the *a.s.* local boundedness of $(Z_t)_{t\geq 0}$, (4.25) and the fact that $\psi_{a,p}\leq CV^p$, we deduce that (4.26) holds with $t_0=0$. Thus, (4.12) follows. (4.13) and (4.14) follow when $\limsup_{t\to+\infty}r(t)<+\infty$. When $\limsup_{t\to+\infty}r(t)=+\infty$, the controls for (X_t) in (4.13) and (4.14) are also true but we still have to prove that for every $p\geq 1$:

$$\sup_{t\geq 0}\eta_t|Y_t|^{2p}<+\infty\quad a.s.\quad\text{and}\quad\sup_{t\geq 0}\eta_t\mathbb{E}[|Y_t|^{2p}]<+\infty. \tag{4.27}$$

By $(\mathbf{H}_\mathbf{a}')(ii)$ and the fact that $\min_{\mathbb{R}^d}U>0$, we get $|\nabla U(x)|\leq U(x)\vee|x|^2$. Thus, we deduce from (2.4) and Jensen's inequality that

$$|Y_t|^{2p}\leq C\left(1+\frac{1}{k(t)}\int_0^t k'(s)(U(X_s)\vee|X_s|^2)^p ds\right)\quad a.s. \tag{4.28}$$

The statements of (4.27) follow easily.
(ii) When $a = 1$, it follows from (4.21), (4.22) and $\liminf_{t\to+\infty} r(t) = r_\infty > 0$ that there exists $\beta' \in \mathbb{R}$, $\alpha' > 0$ and $t_0 \geq 0$ such that for every $(x, y, t) \in \mathbb{R}^d \times \mathbb{R}^d \times [t_0, +\infty[$,

$$\mathcal{A}V^p(x, y, t) \leq \beta' - \alpha' V^p(x, y, t).$$

Then, Itô formula yields,

$$\begin{aligned}\mathbb{E}[e^{\alpha' t} V^p(X_t, Y_t, t)] &= e^{\alpha' t_0}\mathbb{E}[V^p(X_{t_0}, Y_{t_0}, t_0)] + \int_{t_0}^t e^{\alpha' s}\mathbb{E}\left[\alpha' V^p(X_s, Y_s, s) + \mathcal{A}V^p(X_s, Y_s, s)\right] ds \\ &\leq e^{\alpha' t_0}\mathbb{E}[V^p(X_{t_0}, Y_{t_0}, t_0)] + \beta' \int_{t_0}^t e^{\alpha' s} ds.\end{aligned}$$

Using (4.25), we deduce that

$$\sup_{t\geq 0} \mathbb{E}[V^p(X_t, Y_t, t)] \leq C\left(1 + \frac{\beta'}{\alpha'}\right) < +\infty.$$

Thus, by (4.21), it follows that $\sup_{t\geq 0} \mathbb{E}[(U(X_t) \vee |X_t|^2)^p] < +\infty$ and then from (4.28) that, $\sup_{t\geq 0} \mathbb{E}[|Y_t|^{2p}] < +\infty$. □
We are now able to establish tightness of mean and stochastic occupation measures as announced in the next proposition.

Proposition 4.1 *Assume* $(\mathbf{R_1})$ *and fix any* $z = (x, y) \in \mathbb{R}^d \times \mathbb{R}^d$. *Then,*

(i) If $(\mathbf{H_0})$ *and* $(\mathbf{H_1})$ *hold,* $(\mu_t^z)_{t\geq 1}$ *is tight and* $\sup_{t\geq 1} \frac{1}{t}\int_0^t \mathbb{E}[|Y_s^z|^2] ds < +\infty$.

(ii) If there exists $a \in (0, 1]$ *such that* $(\mathbf{H'_a})$ *holds, for every* $q \geq 1$

$$\sup_{t\geq 1} \frac{1}{t}\int_0^t \left(U(X_s^z) \vee |X_s^z|^2\right)^q + \frac{\{|Y_s^z|^2\}^q}{r(s)^{q-1}} ds < +\infty \quad a.s.$$

In particular, $(\nu_t^z(\omega, .)_{t\geq 1}$ *is a.s. tight.*

Proof : The proof of (i) is almost standard and we provide it for sake of completeness, (ii) requires more technicalities.

(i) On the one hand, the tightness of $(\mu_t^z)_{t\geq 1}$ follows from (3.5) and from a straightforward adaptation of Lemma 9.7 of Ethier and Kurtz (1986) (chapter 9, p. 242) to this non-homogeneous Markovian framework. On the other hand, by (4.11) $(\mathbf{H_1})$ and the fact that ∇U, D^2U and σ are locally bounded, we deduce that there exist $t_1 > 0$, $\beta_1 \in \mathbb{R}$ and $\alpha_1 > 0$ such that for every $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$, for every $t \geq t_1$, $\mathcal{A}V(x, y, t) \leq \beta_1 - \alpha_1|y|^2$. Then, applying Itô formula, we have

$$\alpha_1 \limsup_{t\to+\infty} \frac{1}{t}\int_{t_1}^t \mathbb{E}[|Y_s|^2] ds \leq \limsup_{t\to+\infty} \left(\frac{\mathbb{E}[V(X_{t_1}, Y_{t_1}, t_1)] - \mathbb{E}[V(X_t, Y_t, t)]}{t} + \beta_1\right) \leq \beta_1$$

and the second statement of (i) follows.
(ii) Set $p \geq 1$. Using Itô formula and (4.23), we have for every $t \geq t_0 > 0$:

$$\frac{1}{t}\int_{t_0}^t \left(\left(U(X_s) \vee |X_s|^2\right)^{p+a-1} + \frac{\{|Y_s|^2\}^p}{r(s)^{p-1}}\right) ds \leq \frac{1}{\tilde{\alpha}} \frac{V^p(X_{t_0}, Y_{t_0}, t_0) - V^p(X_t, Y_t, t)}{t} + \beta + \frac{N_t - N_{t_0}}{t},$$

where (N_t) is defined for every $t \geq 0$ by

$$N_t = p\int_0^t V^{p-1}(X_s, Y_s, s)\langle \nabla U(X_s) + m_\varepsilon\left(X_s - \frac{Y_s}{\rho(s)}\right), \sigma(X_s) dW_s\rangle.$$

Note that by Lemma 4.4, (N_t) is a martingale. We have now to show that $N_t/t \to 0$ as $t \to +\infty$. Set $M_t = \int_{t_0}^t (1/s) dN_s$. We want to show that $\langle M \rangle_\infty < +\infty$. For every $t \geq t_0$, we have:

$$\langle M \rangle_t \leq C \int_{t_0}^t s^{-2} V^{2p-2}(X_s, Y_s, s) \left(|\nabla U(X_s)|^2 + |X_s|^2 + \frac{|Y_s|^2}{\rho^2(s)} \right) \mathrm{Tr}(\sigma\sigma^*)(X_s) ds.$$

Owing to the elementary inequality $|u+v|^p \leq C_p(|u|^p + |v|^p)$ for $u, v \in \mathbb{R}$, we have:

$$V^{2p-2}(x,y,t) \leq C \left(1 + \left(U(x) \vee |x|^2 \right)^{2p-2} + \frac{\{|y|^2\}^{2p-2}}{r(t)^{2p-2}} \right).$$

Using $(\mathbf{H}'_\mathbf{a})$, it follows that, since $\rho(t) \sim r(t)$,

$$\begin{aligned} \langle M \rangle_t &\leq C \int_{t_0}^t s^{-2} \left(1 + \frac{|Y_s|^2}{r(s)^2} \right) \left(U(X_s) \vee |X_s|^2 \right)^{2p+a-1} ds \\ &\quad + C \int_{t_0}^t s^{-2} \left(1 + \frac{\{|Y_s|^2\}^{2p-1}}{r(s)^{2p}} + \frac{\{|Y_s|^2\}^{2p-2}}{r(s)^{2p-2}} \right) \left(U(X_s) \vee |X_s|^2 \right)^a ds. \end{aligned} \tag{4.29}$$

By the elementary inequality $|uv| \leq (u^2+v^2)/2$ for $u, v \in \mathbb{R}$, and after some standard inequalities, one can find suitable constants $(C, \tilde{C})$ satisfying *a.s.*,

$$\begin{aligned} \langle M \rangle_t &\leq C \int_{t_0}^t s^{-2} \left[1 + \left(U(X_s) \vee |X_s|^2 \right)^{4p+2a-2} + \frac{\{|Y_s|^2\}^2}{r(s)^4} + \frac{\{|Y_s|^2\}^{4p-2}}{r(s)^{4p}} + \frac{\{|Y_s|^2\}^{4p-4}}{r(s)^{4p-4}} \right] ds \\ &\leq \tilde{C} \int_{t_0}^t s^{-2} \left[1 + \left(U(X_s) \vee |X_s|^2 \right)^{4p+2a-2} + \frac{\{|Y_s|^2\}^2}{r(s)} + \frac{\{|Y_s|^2\}^{4p-2}}{r(s)^{4p-3}} + \frac{\{|Y_s|^2\}^{4p-4}}{r(s)^{4p-5}} \right] ds. \end{aligned}$$

The second inequality holds since $\liminf r(t) > 0$. By Lemma 4.4, it follows that $\langle M \rangle_\infty < +\infty$ *a.s.* Then, (M_t) is a convergent martingale. Setting $M_\infty = \lim_{t \to +\infty} M_t$, we derive from an integration by parts and from Cesaro's Lemma that

$$\begin{aligned} \frac{N_t - N_{t_0}}{t} = \frac{1}{t} \int_{t_0}^t s dM_s &= (M_t - M_{t_0}) - \frac{1}{t} \int_{t_0}^t (M_s - M_{t_0}) ds \\ &\xrightarrow{t \to +\infty} (M_\infty - M_{t_0}) - (M_\infty - M_{t_0}) = 0. \end{aligned}$$

As a consequence, $\frac{N_t}{t} \to 0$ *a.s.* and it follows that $(\nu_t^z)_{t \geq 1}(\omega, .)$ is *a.s.*-tight. □

The next proposition identifies the adherence of the tight occupation measures in the several cases addressed by Theorem 3.2.

Proposition 4.2 *Assume* $(\mathbf{R_1})$.

(i) Assume $(\mathbf{H_0})$ *and* $(\mathbf{H_1})$ *and denote by* μ_∞ *an accumulation point of the (tight) family* $(\mu_t^z)_{t \geq 1}$. *Then,*

- (i_1) *If* $r_\infty = +\infty$, *the first marginal of* μ_∞ *is an invariant distribution for SDE (3.1).*
- (i_2) *If* $r(t) \to r_\infty \in \mathbb{R}_+^*$, $\mu_\infty(dx, dy)$ *is an invariant distribution of the homogeneous Markov process solution to* (2.3) *with* $r(t) = r_\infty$ *for every* $t \geq 0$.

(ii) Let $a \in (0, 1]$ *such that* $(\mathbf{H}'_\mathbf{a})$ *holds, and denote by* $\nu_\infty(\omega)$ *an accumulation point of the (a.s. tight) family* $(\nu_t^z(\omega))_{t \geq 1}$.

- (ii_1) *If* $r_\infty = +\infty$, $\nu_\infty(\omega, dx, dy) = \delta_{\nabla U(x)}(dy) \pi(dx)$ *where* π *is a.s. an invariant distribution for SDE (3.1).*
- (ii_2) *If* $r(t) \to r_\infty \in \mathbb{R}_+^*$, $\nu_\infty(\omega, dx, dy)$ *is an invariant distribution of the homogeneous Markov process solution to* (2.3) *with* $r(t) = r_\infty$ *for every* $t \geq 0$.

Remark 4.2 *Oppositely to (ii_1), we only identify the first marginal of the accumulation point in (i_1). This point may appear a little surprising since $\mu_t(f) = \mathbb{E}[\nu_t(f)]$ but is due to the weaker stability assumption $(\mathbf{H_1})$ in the first part of the proposition. Note that we could obtain the whole identification in (i_1) under $(\mathbf{H'_a})$.*

Proof : The proof of (i_1) and (i_2) being an adaptation of that of (ii_1) and (ii_2) in a simpler case, we choose to mainly detail the second ones and to give some elements of the first ones at the end of the proof.
(ii_1) We assume that $r_\infty = +\infty$. Let $\nu_\infty(\omega)$ be a weak limit of $(\nu_t^z(\omega))_{t\geq 0}$. By the factorization theorem for probability measures, $\nu_\infty(\omega) = \nu_\omega^1(x,dy)\nu_\omega^2(dx)$ where *a.s.*, ν_ω^1 is a transition probability and ν_ω^2 is a probability distribution. We first prove that ν_ω^2 is *a.s.* an invariant distribution for SDE (3.1). Recall the notation $\mathcal{L}$ for the generator of $(S_t)_{t\geq 0}$ solution to the classical SDE (3.1). Owing to the Echeverria-Weiss Theorem (see *e.g.* Theorem 9.17, chapter 9, p. 248 of Ethier and Kurtz (1986)), ν_ω^2 is an invariant distribution for $(S_t)_{t\geq 0}$ if *a.s.*, $\int \mathcal{L}f(x)\nu_\infty(\omega,dx,dy) = 0$ for every $f \in \mathcal{C}_K^3(\mathbb{R}^d)$. A countability argument shows that it is enough to show that for every $f \in \mathcal{C}_K^3(\mathbb{R}^d)$, $\int \mathcal{L}f(x)\nu_\infty(\omega,dx,dy) = 0$ *a.s.*, *i.e.* that for every $f \in \mathcal{C}_K^3(\mathbb{R}^d)$,

$$\frac{1}{t}\int_0^t \mathcal{L}f(X_s)ds \xrightarrow{t\to+\infty} 0 \quad a.s.$$

The idea of the proof is thus to compare the generator $\mathcal{A}$ of (2.3) to $\mathcal{L}$, the one corresponding to SDE (3.1). For $f \in \mathcal{C}_K^3(\mathbb{R}^d)$, set $g_1(x,y,t) = f(x) - \frac{\langle y,\nabla f(x)\rangle}{r(t)}$. By Itô formula,

$$-\frac{1}{t}\int_0^t Ag_1(X_s,Y_s,s)ds = \frac{g_1(x,y,0) - g_1(X_t,Y_t,t)}{t} + \frac{\tilde{N}_t}{t}$$

with $\tilde{N}_t = \int_0^t \langle \nabla_x g_1(X_s,Y_s,s), \sigma(X_s)dW_s\rangle$. On the one hand, by (4.13) applied with $\eta_t = 1/t^2$ and $p = 1$, $\sup_{t\geq 1}|Y_t|^2/t^2 < +\infty$ *a.s.* Then, since f is compactly supported,

$$\frac{|g_1(X_t,Y_t,t)|}{t} \leq C\left(\frac{1}{t} + \frac{|Y_t|}{r(t)t}\right) \xrightarrow{t\to+\infty} 0 \quad a.s., \tag{4.30}$$

and it follows that

$$\frac{g_1(X_t,Y_t,t) - g_1(x,y,0)}{t} \xrightarrow{t\to+\infty} 0 \quad a.s.$$

On the other hand, using that $\inf_{t\geq 0} r(t) > 0$, we have

$$|\nabla_x g_1(x,y,t)|^2 = |\nabla f(x) - \frac{D^2 f(x)y}{r(t)}|^2 \leq C(1+|y|^2)).$$

Then, it follows from (4.12) that the martingale $(\tilde{M}_t)_{t\geq 1}$ defined by $\tilde{M}_t = \int_1^t (1/s)d\tilde{N}_s$ and a similar method to that of (4.29) that

$$\frac{1}{t}\int_0^t \langle \nabla_x g_1(X_s,Y_s,s), \sigma(X_s)dW_s\rangle \xrightarrow{n\to+\infty} 0 \quad a.s.$$

Finally, $\nu_t^z(\mathcal{A}g_1) \xrightarrow{t\to+\infty} 0$ *a.s.*. Now, $\mathcal{A}$ can be decomposed as $\mathcal{A}g_1(x,y,t) = \mathcal{L}f(x) + \mathcal{H}(x,y,t)$ where

$$\mathcal{H}(x,y,t) = \frac{1}{r(t)}\left(-y^*D^2f(x)y + \frac{1}{2}\mathrm{Tr}(\sigma^*(x)C_y(x)\sigma(x))\right) - g_1(x,y,t)\frac{r'(t)}{r^2(t)},$$

and $(C_y(x))_{i,j} = \sum_{l=1}^d y_l \partial^3_{x_i,x_j,x_l} f(x)$. Then, ν_ω^2 is *a.s.* an invariant distribution for SDE (3.1) if

$$\frac{1}{t}\int_0^t \mathcal{H}(X_s,Y_s,s)ds \xrightarrow{t\to+\infty} 0 \quad a.s.$$

and this point follows from the control below

$$|\mathcal{H}(x,y,t)| \leq \left(\frac{1}{r(t)}(|y|^2+|y|)+\frac{|r'(t)|}{r^2(t)}(1+|y|)\right),$$

the fact that $r'(t)/r^2(t) \to 0$, $1/r(t) \to 0$ (under the assumptions of (ii_1)), and Proposition 4.1(ii) that yields : $\sup_{t\geq 1}\frac{1}{t}\int_0^t |Y_s|^2 ds < +\infty \quad a.s.$

Finally, let us check that $\nu_\omega^2(dx)$-$a.s$ one has $\nu_\omega^1(x,dy) = \delta_{\nabla U(x)}(dy)$ $a.s$. It is enough to prove that for every Lipschitz bounded continuous function $f : \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}$,

$$\frac{1}{t}\int_0^t f(X_s,Y_s)ds - f(X_s,\nabla U(X_s))ds \xrightarrow{t\to+\infty} 0 \quad a.s.,$$

and this property will be true if

$$\frac{1}{t}\int_0^t |\nabla U(X_s)-Y_s|ds \xrightarrow{t\to+\infty} 0 \quad a.s. \tag{4.31}$$

Applying Itô formula to $\varphi(t,\omega) = |\nabla U(X_t)-Y_t|^2$, we get

$$\begin{aligned}\frac{1}{t}\int_0^t r(s)|\nabla U(X_s)-Y_s|^2 ds = \frac{|\nabla U(x)-y|^2-\varphi(t,\omega)}{2t} + \frac{1}{2t}\int_0^t F(X_s,Y_s)ds \\ + \frac{1}{t}\int_0^t \langle D^2U(X_s)(\nabla U(X_s)-Y_s), \sigma(X_s)dW_s\rangle\end{aligned}$$

where F is a function that depends on ∇U, D^2U, D^3U and σ that satisfies

$$|F(x,y)| \leq C(1+(U(x)\vee|x|^2)^{p_1}+|y|^2) \quad \text{with } p_1 \geq 1. \tag{4.32}$$

Now, it follows from Proposition 4.1 and (4.32) that $\sup_{t\geq 1}\frac{1}{t}\int_0^t |F(X_s,Y_s)|ds < +\infty \quad a.s.$ and similar arguments as those developed in (4.29) combined with (4.12) yield

$$\frac{1}{t}\int_0^t \langle D^2U(X_s)(\nabla U(X_s)-Y_s), \sigma(X_s)dW_s\rangle \xrightarrow{t\to+\infty} 0 \quad a.s.$$

As a consequence,

$$\sup_{t\geq 1}\frac{1}{t}\int_0^t r(s)|\nabla U(X_s)-Y_s|^2 ds < +\infty \quad a.s.$$

and (4.31) follows from Jensen's inequality and the fact that $r(t) \to +\infty$.
(ii_2): $r(t) \to r_\infty < +\infty$. Let $\nu_\infty(\omega)$ be a weak limit of $(\nu_t^z(\omega))_{t\geq 1}$. Let $f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. Following a similar strategy, it is enough to show that

$$\frac{1}{t}\int_0^t \mathcal{A}_{r_\infty} f(X_s,Y_s)ds \xrightarrow{t\to+\infty} 0 \quad a.s.$$

where $\mathcal{A}_{r_\infty}$ denotes the infinitesimal generator of the homogeneous Markov process $(Z_t)_{t\geq 0}$ when r is constant ($r(t) = r_\infty$ for every $t \geq 0$). In this case, we set $g_2(x,y,t) = f(x,y)$ and derive from similar arguments to those developed in (ii_1) that,

$$\frac{1}{t}\int_0^t \mathcal{A}g_2(X_s,Y_s,s)ds \xrightarrow{t\to+\infty} 0 \quad a.s.$$

Then, we can show that

$$\frac{1}{t}\int_0^t \mathcal{A}_{r_\infty} f(X_s,Y_s)ds - \mathcal{A}g_2(X_s,Y_s,s)ds \xrightarrow{t\to+\infty} 0 \quad a.s.$$

using that

$$\mathcal{A}g_2(x,y,t) - \mathcal{A}_{r_\infty} f(x,y) = (r(t) - r_\infty)\langle \nabla U(x) - y, \nabla_y f(x,y)\rangle$$

combined with the fact that supp f is compact and $r(t) \xrightarrow{t\to+\infty} r_\infty$. This ends the proof of (ii_2).

(i_1) and (i_2): in this case, it is enough to prove that for every $\mathcal{C}^2$-function f with compact support,

$$\frac{1}{t}\int_0^t \mathbb{E}[\mathcal{L}f(X_s)]ds \xrightarrow{t\to+\infty} 0 \quad \text{if } r(t) \to +\infty \quad \text{and,} \tag{4.33}$$

$$\frac{1}{t}\int_0^t \mathbb{E}[\mathcal{A}_{r_\infty} f(X_s, Y_s)]ds \xrightarrow{t\to+\infty} 0 \quad \text{if } r(t) \to r_\infty \in \mathbb{R}_+^*. \tag{4.34}$$

Following carefully the proof of (ii_2), one checks that (4.34) is true using in particular that g_2 is compactly supported. In the same way, one observes that under $(\mathbf{H_0})$ and $(\mathbf{H_1})$ (only), a direct adaptation of the proof of the first part of (i_1) is available to obtain (4.33) if

$$\sup_{t\geq 1}\frac{1}{t}\int_0^t \mathbb{E}[|Y_s|^2]ds < +\infty \quad \text{and,} \quad \frac{\mathbb{E}[|Y_t|]}{r(t)t} \xrightarrow{t\to+\infty} 0 \quad (\text{when } r(t)\to+\infty).$$

The first point has already been proven in Proposition 4.1. Let us focus on the second one. Since for t large enough $\limsup_{|x,y)|\to+\infty} \mathcal{A}V(x,y,t) = -\infty$, it follows that there exist C and $t_0 > 0$ such that for every $t \geq t_0$ for every $x, y \in \mathbb{R}^d$, $\mathcal{A}V(x,y,t) \leq C$. Then, an adaptation of the proof of Lemma 4.4 (see (4.24)) shows that $(\tilde{G}_t)_{t\geq t_0}$ defined by $\tilde{G}_t = \eta_t V(X_t, Y_t, t) + C\int_t^{+\infty}\eta_s ds$ is a non-negative supermartingale when $t \mapsto \eta_t$ is a non-increasing positive function such that $\int_0^{+\infty}\eta_s ds < +\infty$. Since $V(x,y,t) \geq |y|^2/(r(t))$, this implies $\sup_{t\geq t_0}\{\eta_t \mathbb{E}[|Y_t|^2]r(t)^{-1}\} < +\infty$. Applying this property with $\eta_t = t^{-2}$ yields

$$\left(\frac{\mathbb{E}[|Y_t|]}{r(t)t}\right)^2 \leq \frac{1}{r(t)}\left(\frac{\mathbb{E}[|Y_t|^2]}{t^2 r(t)}\right) \leq \frac{C}{r(t)}\left(1 + \sup_{t\geq t_0}\frac{1}{t^2}\left(\frac{\mathbb{E}[|Y_t|^2]}{r(t)}\right)\right) \xrightarrow{t\to+\infty} 0.$$

□

4.4 Some properties of the invariant distribution

In this short paragraph, we provide some identification clues for the adherence points of $(\nu_t^z(\omega,.))_{t\geq 0}$ and $(\mu_t^z)_{t\geq 0}$ when $r(t) = r_\infty \in \mathbb{R}_+^*$. This results are summarized in Proposition 3.2 in the first part of the paper.

Proof of Proposition 3.2:

(ii) We study first the identification problem when $d = 1$, σ a positive constant, $r(t) = r_\infty > 0$ and $U(x) = x^2/2$. Note first that the assumptions of Theorem 3.3 are fulfilled with $a = 1$. Then, in particular, $(X_t, Y_t)_{t\geq 0}$ converges weakly to the unique invariant distribution ν. Since $(X_t, Y_t)_{t\geq 0}$ is a Gaussian process, ν is a Gaussian random variable whose parameters are the limits of the expectation and variance/covariance of the process. Let us compute these limits. We set for any $t \geq 0$: $\phi_1(t) = \mathbb{E}[X_t]$ and $\phi_2(t) = \mathbb{E}[Y_t]$. One easily checks that ϕ_1 and ϕ_2 satisfy a simple coupled differential equation

$$\begin{cases} \phi_1'(t) = & -\phi_2(t) \\ \phi_2'(t) = & r_\infty[\phi_1(t) - \phi_2(t)]. \end{cases}$$

When $r_\infty \in]0,4[$, the eigenvalues of the matrix of the system are complex with negative real part and when $r_\infty \geq 4$, the eigenvalues are real and negative. In the two cases, it follows that

$$\mathbb{E}[X_t] \xrightarrow{t\to+\infty} 0 \quad \text{and,} \quad \mathbb{E}[Y_t] \xrightarrow{t\to+\infty} 0.$$

Set now $f(t) = \mathbb{E}[|X_t|^2], g(t) = \mathbb{E}[|Y_t|^2]$ and $h(t) = \mathbb{E}[X_t Y_t]$. Simple computations yield the following first order differential system:

$$\begin{cases} f'(t) = & \sigma^2 - 2h(t) \\ g'(t) = & 2r_\infty[h(t) - g(t)] \\ h'(t) = & -g(t) + r_\infty[f(t) - h(t)]. \end{cases}$$

The homogeneous system associated to the preceding one can be written $Z'(t) = M_{r_\infty} Z(t)$ where $Z(t) = (f(t), g(t), h(t))^T$ and

$$M_{r_\infty} = \begin{pmatrix} 0 & 0 & -2 \\ 0 & -2r_\infty & 2r_\infty \\ r_\infty & -1 & -r_\infty \end{pmatrix}.$$

Computations on the characteristic polynom of this matrix show that for every $r_\infty > 0$, M_{r_∞} has a negative real eigenvalue that we denote by α and two complex eigenvalues β_1 and β_2 whose real part is negative when $r_\infty < 4$. When $r_\infty \geq 4$, M_{r_∞} has a 3 real negative eigenvalues. Denoting by $\Delta_{r_\infty} = \text{diag}(\alpha, \beta_1, \beta_2)$, by $(v_\alpha, v_{\beta_1}, v_{\beta_2})$ a basis of eigen vectors and by P_{r_∞} the matrix of the coordinates of this set of vectors in the canonical basis, we have:

$$Z'(t) = P_{r_\infty} \Delta_{r_\infty} P_{r_\infty}^{-1} Z(t) + \begin{pmatrix} \sigma^2 \\ 0 \\ 0 \end{pmatrix}.$$

Consider now $\tilde{Z}(t) = Z(t) + M_{r_\infty}^{-1} \begin{pmatrix} \sigma^2 \\ 0 \\ 0 \end{pmatrix}$, we check immediately that

$$\tilde{Z}'(t) = P_{r_\infty} \Delta_{r_\infty} P_{r_\infty}^{-1} \tilde{Z}(t),$$

and $\tilde{Z}$ is given by

$$P_{r_\infty}^{-1} \tilde{Z}(t) = e^{\Delta_{r_\infty} t} P_{r_\infty}^{-1} \tilde{Z}(0).$$

Hence, using that the real part of the eigenvalues is negative, it follows that $\tilde{Z}(t) \to \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ as $t \to +\infty$. Thus,

$$Z(t) \xrightarrow{t \to +\infty} -M_{r_\infty}^{-1} \begin{pmatrix} \sigma^2 \\ 0 \\ 0 \end{pmatrix} = - \begin{pmatrix} -\frac{r_\infty + 1}{2r_\infty} & -\frac{1}{2{r_\infty}^2} & \frac{1}{r_\infty} \\ -\frac{1}{2} & -\frac{1}{2r_\infty} & 0 \\ -\frac{1}{2} & 0 & 0 \end{pmatrix} \begin{pmatrix} \sigma^2 \\ 0 \\ 0 \end{pmatrix}.$$

Thus, we obtain that $\mathbb{E}[(X_t)^2] \xrightarrow{t \to +\infty} \frac{r_\infty + 1}{r_\infty} \frac{\sigma^2}{2}$, $\mathbb{E}[(Y_t)^2] \xrightarrow{t \to +\infty} \frac{\sigma^2}{2}$, $\mathbb{E}[X_t Y_t] \xrightarrow{t \to +\infty} \frac{\sigma^2}{2}$. This ends the proof of (ii).

(i) We consider now a more general case when $r(t) = r_\infty$ and U satisfies the hypo-ellipticity condition $\dim(\mathcal{M}_U) \leq d - 1$. Then, from Theorem 2.1 and 3.1 there exists a unique invariant distribution of the coupled markovian process $(X_t, Y_t)_{t \geq 0}$ whose density is denoted p_{r_∞} with respect to the Lebesgue measure. We know that p_{r_∞} is characterized by the balance property:

$$\forall f \in \mathcal{C}_K^2(\mathbb{R} \times \mathbb{R}) : \qquad \int_{\mathbb{R} \times \mathbb{R}} \mathcal{A} f(x,y) p_{r_\infty}(x,y) dx dy = 0.$$

Since we can choose f and all its derivatives of order 1 and 2 vanishing on ∂K, very simple integration by parts ensures that p_{r_∞} satisfies the following partial differential equation:

$$\langle y, \nabla_x p_{r_\infty}(x,y) \rangle + \frac{1}{2} \text{Tr} \left(D_x^2 (p_{r_\infty(x,y)} \sigma(x,y) \sigma^*(x,y)) \right) + r_\infty \left[\langle y - \nabla U(x), \nabla_y p_{r_\infty}(x,y) \rangle + p_{r_\infty} \right] = 0$$

This ends the proof of the proposition. □

5 Convergence rate when $r_\infty = +\infty$

Proof of Theorem 3.4: We establish in this proof that under the conditions of Theorem 3.4, there exist some positive constants C_1 and C_2 such that for every positive s, t,

$$W_2(\mathbb{P}_{X_{t+s}}, \pi) \leq C_1 \exp(-\rho s) + \frac{C_2}{k(t)} \int_0^t k(v)dv,$$

which is obviously equivalent to the conclusion of Theorem 3.4. First, note that under the assumptions of the theorem, π has a moment of any order owing to (3.5) combined with (3.7). Then, let (W_t) denote a d-dimensional Brownian Motion and let (X_t, Y_t) denote the unique (strong) solution to (2.3) adapted to the filtration generated by W. For every $t > 0$, we also denote by $(S_s^{t,x})$ the unique (strong) solution to the following SDE

$$dS_s = -\nabla U(S_s)ds + \sigma(S_s)dW_{t+s}, \quad S_0 = x, \tag{5.1}$$

where (W_t) is the same Brownian Motion as previously. Let $(S_s^{t,\pi})$ denote a stationary solution to (5.1) whose initial value is independent of X_t (this is possible owing to a potential filtration enlargement). We have:

$$W_2(\mathbb{P}_{X_{t+s}}, \pi) \leq \mathbb{E}[|X_{t+s} - S_s^{t,\pi}|^2]^{\frac{1}{2}} \leq \mathbb{E}[|X_{t+s} - S_s^{t,X_t}|^2]^{\frac{1}{2}} + \mathbb{E}[|S_s^{t,X_t} - S_s^{t,\pi}|^2]^{\frac{1}{2}}, \tag{5.2}$$

thanks to the Minkowski inequality. First, we deduce from Itô formula that

$$\exp(2\rho s)|S_s^{t,x} - S_s^{t,y}|^2 = |x-y|^2 + \int_0^s \exp(2\rho u)\left(2\rho|S_u^{t,x} - S_u^{t,y}|^2 + 2\varphi(S_u^{t,x}, S_u^{t,y})\right)du + M_s^t,$$

where $(M_s^t)_{s\geq 0}$ is a martingale and φ denotes the left-hand side of Assumption (**AC**) (see (3.10)) defined by

$$\varphi(x_1, x_2) = \langle x_1 - x_2, \nabla U(x_2) - \nabla U(x_1)\rangle + \frac{1}{2}\mathrm{Tr}\left((\tilde{\sigma}\tilde{\sigma}^*)(x_1, x_2)\right)$$

with $\tilde{\sigma}(x_1, x_2) = \sigma(x_2) - \sigma(x_1)$. Then, owing to (**AC**), it follows that

$$\mathbb{E}[|S_s^{t,x} - S_s^{t,y}|^2] \leq \exp(-2\rho s)|x-y|^2.$$

Thus, using that X_t and $S_0^{t,\pi}$ are independent, we deduce that

$$\mathbb{E}[|S_s^{t,X_t} - S_s^{t,\pi}|^2]^{\frac{1}{2}} \leq \exp(-\rho s)\left(\int\int |x-y|^2\pi(dy)\mathbb{P}_{X_t}(dx)\right)^{\frac{1}{2}} \leq C_1\exp(-\rho s), \tag{5.3}$$

where $C_1 = \left(\sup_{t\geq 0}\int\int |x-y|^2\pi(dy)\mathbb{P}_{X_t}(dx)\right)^{\frac{1}{2}}$ is finite since π has a moment of order 2 and $\sup_{t\geq 0}\mathbb{E}[|X_t|^2] < +\infty$ by (4.15).
Second, we focus on the first term of the right-hand side of (5.2). Set $\delta = 2\rho - \varepsilon$ where $\varepsilon \in (0, 2\rho)$. As previously, setting $H_s = X_{t+s} - S_s^{t,X_t}$, we deduce from Itô formula that

$$\begin{aligned}\exp(\delta s)\mathbb{E}[|H_s|^2] &= \int_0^s \delta\exp(\delta u)\mathbb{E}[|H_u|^2]du \\ &\quad + 2\int_0^s \exp(\delta u)\left(\mathbb{E}[\langle H_u, \nabla U(S_u^{t,X_t}) - Y_{t+u}\rangle] + \frac{1}{2}\mathbb{E}[\mathrm{Tr}(\tilde{\sigma}\tilde{\sigma}^*(X_{t+u}, S_u^{t,X_t}))]\right)du. \end{aligned}\tag{5.4}$$

Using that

$$\langle H_u, \nabla U(S_u^{t,X_t}) - Y_{t+u}\rangle = \langle X_{t+u} - S_u^{t,X_t}, \nabla U(S_u^{t,X_t}) - \nabla U(X_{t+u})\rangle + \langle H_u, \nabla U(X_{t+u}) - Y_{t+u}\rangle,$$

we deduce that

$$(5.4) = 2\int_0^s \exp(\delta u)\left(\frac{\delta}{2}\mathbb{E}[|H_u|^2] + \mathbb{E}[\varphi(X_{t+u}, S_u^{t,X_t})] + \mathbb{E}[\langle H_u, \nabla U(X_{t+u}) - Y_{t+u}\rangle]\right) du.$$

Using the fact that for every $\varepsilon > 0$, $|\langle H_u, \nabla U(X_{t+u}) - Y_u\rangle| \leq \frac{\varepsilon}{2}|H_u|^2 + |\nabla U(X_{t+u}) - Y_{t+u}|^2/(2\varepsilon)$, one can find sufficiently small δ and ε to ensure that $\delta - 2\rho + \varepsilon < 0$ and then it follows from Assumption (**AC**) that

$$\mathbb{E}[|H_s|^2] \leq \frac{\exp(-\delta s)}{\varepsilon}\int_0^s \exp(\delta u)\mathbb{E}[|Y_{t+u} - \nabla U(X_{t+u})|^2]du. \tag{5.5}$$

Now, by (2.4)

$$|Y_{t+u} - \nabla U(X_{t+u})|^2 \leq 2\left(\frac{|y|^2 k(0)^2}{(k(t+u))^2} + \left(\frac{1}{k(t+u)}\int_0^{t+u} k'(v)|\nabla U(X_v) - \nabla U(X_{t+u})|dv\right)^2\right).$$

Then, we deduce from Jensen's inequality that

$$\mathbb{E}[|Y_{t+u} - \nabla U(X_{t+u})|^2] \leq \frac{2|y|^2 k(0)^2}{(k(t+u))^2}|y| + \frac{1}{k(t+u)}\int_0^{t+u} k'(v)\mathbb{E}[|\nabla U(X_{t+u}) - \nabla U(X_v)|^2]dv.$$

Applying Itô formula, we have

$$\begin{aligned}|\nabla U(X_{t+u}) - \nabla U(X_v)|^2 = &-2\int_v^{t+u}\langle \nabla U(X_r) - \nabla U(X_v), D^2U(X_r)Y_r\rangle dr \\ &+ \frac{1}{2}\sum_{i=1}^d \int_v^{t+u} \mathrm{Tr}(\sigma^* D^3 U_{i,\cdot}\sigma^*)(X_r)dr + M_{t+u} - M_v,\end{aligned}$$

where $M_t = 2\int_0^t \langle \nabla U(X_r) - \nabla U(X_v), D^2U(X_r)\sigma(X_r)dW_r\rangle$ and $D^3U_{i,\cdot}$ is the $d \times d$ matrix defined by $(D^3U_{i,\cdot})_{j,k} = \partial_{i,j,k}U$. By Assumption $(\mathbf{H_1'})(ii)$ and the elementary inequality $|\langle a, b\rangle| \leq (|a|^2 + |b|^2)/2$, we obtain the existence of a positive number $\bar{p}$ such that:

$$\begin{aligned}\left|\langle \nabla U(X_r) - \nabla U(X_v), D^2U(X_r)Y_r\rangle\right. &+ \left.\frac{1}{2}\sum_{i=1}^d \mathrm{Tr}(\sigma^* D^3 U_{i,\cdot}\sigma^*)(X_r)\right| \\ &\leq C\left(1 + (U(X_v) \vee |X_v|^2)^{\bar{p}} + (U(X_r) \vee |X_r|^2)^{\bar{p}} + |Y_r|^2\right)\end{aligned}$$

Furthermore, by (2.4), $(\mathbf{H_1'})(ii)$ and Jensen's inequality,

$$\mathbb{E}[|Y_r|^2] \leq \sup_{0\leq l\leq r}\mathbb{E}[|\nabla U(X_l)|^2] \leq C(1 + \sup_{0\leq l\leq r}\mathbb{E}[(U(X_l) \vee |X_l|)^2]).$$

Thus, by Lemma 4.4 (see (4.15)), we deduce that there exists a positive C such that for every $t, u, v \in \mathbb{R}_+$ with $t + u \geq v$,

$$\mathbb{E}[|\nabla U(X_{t+u}) - \nabla U(X_v)|^2] \leq C(t + u - v).$$

As a consequence,

$$\mathbb{E}[|Y_{t+u} - \nabla U(X_{t+u})|^2] \leq \frac{C}{k(t+u)}\left(\int_0^{t+u} k'(v)(t+u-v)dv\right) \leq \frac{C}{k(t+u)}\left(\int_0^{t+u} k(v)dv\right),$$

by an integration by parts. Using that $t \mapsto \int_0^t k(v)dv)/k(t)$ is nonincreasing and plugging the previous inequality into (5.5), we obtain that for every positive s and t:

$$\mathbb{E}[|H_s|^2] \leq C\frac{\int_0^t k(v)dv}{k(t)}\exp(-\delta s)\int_0^s \exp(\delta u)du \leq C_2\frac{\int_0^t k(v)dv}{k(t)},$$

where C_2 does not depend on t and s. Theorem 3.4 follows. □

Proof of Corollary 3.1:

For any $t \geq t_0$, set

$$\varphi_t(u) = C_1 e^{-\rho(t-u)} + \frac{C_2}{k(u)} \int_0^u k(v)dv.$$

φ_t achieves its unique minimum on $u^*(t)$ which satisfies

$$\rho C_1 e^{-\rho(t-u)} + C_2 = C_2 \frac{k'(u)}{k^2(u)} \int_0^u k(v)dv.$$

Taking the logarithm, we obtain

$$t = u - \frac{1}{\rho} \log \left[\frac{C_2}{\rho C_1} \left(\frac{r(u)}{k(u)} \int_0^u k(v)dv + \frac{k(0)}{k(u)} - 1 \right) \right].$$

Using the notation H defined in Corollary 3.1, we thus obtain $u^*(t) = H^{-1}(t)$, and

$$W_2(\mathbb{P}_{X_t}, \pi) \leq \varphi_t(H^{-1}(t)) = C_1 e^{-\rho(t-H^{-1}(t))} \left(1 + \frac{\rho}{r(H^{-1}(t))} \right) + \frac{C_2}{r(H^{-1}(t))}.$$

□

6 Proof of general unstability ($r_\infty = 0$ and subquadratic potential)

Proof of Theorem 3.5: Let $J : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$ be defined by

$$J(x, y, t) = v(t) \left(r(t)U(x) + \frac{|y|^2}{2} \right),$$

where v is a $\mathcal{C}^1$-function on $\mathbb{R}_+$ which will be fixed later. We have:

$$\mathcal{A}J(x, y, t) = \frac{1}{2} r(t)v(t) \mathrm{Tr} \left[\sigma^*(x) D^2 U(x) \sigma(x) \right] + |y|^2 \left(-v(t)r(t) + \frac{v'(t)}{2} \right) + (r(t)v(t))' U(x).$$

From now, we take v as a solution of the o.d.e. given by $-v(t)r(t) + \frac{v'(t)}{2} = 0$. Thus, v can be chosen as $v(t) = \exp(2R(t))$ where $R(t) = \int_0^t r(s)ds$. Using that $\mathrm{Tr}(\sigma^* D^2 U \sigma)(x) \geq \lambda_0 > 0$, we deduce that

$$\mathcal{A}J(x, y, t) \geq \frac{\lambda_0}{4} v'(t) + (r(t)v(t))' U(x).$$

As a consequence, it follows from Itô formula applied between t_0 and t that

$$\mathbb{E}\left[J(X_t, Y_t, t) \right] \geq \mathbb{E}\left[J(X_{t_0}, Y_{t_0}, t_0) \right] + \frac{\lambda_0}{4} \left[v(t) - v(t_0) \right] + \int_{t_0}^t (r(s)v(s))' \mathbb{E}\left[U(X_s) \right] ds.$$

Dividing by $v(t)$, we deduce

$$r(t)\mathbb{E}[U(X_t)] + \frac{\mathbb{E}[|Y_t|^2]}{2} \geq \frac{C}{v(t)} + \frac{\lambda_0}{4} \left(1 - \frac{v(t_0)}{v(t)} \right) + \frac{1}{v(t)} \int_{t_0}^t (r(s)v(s))' \mathbb{E}[U(X_s)] ds.$$

Let $t_0 \in \mathbb{R}_+$ such that $r'(t) + 2r^2(t) \geq 0, \forall t \geq t_0$, then $(r(t)v(t))' \geq 0$ for $t \geq t_0$ and it follows that

$$\forall t \geq t_0 \qquad \frac{1}{v(t)} \int_{t_0}^t (r(s)v(s))' \mathbb{E}[U(X_s)] ds \geq 0.$$

Using that $v(t) \to +\infty$ (since $v(t) = (k(t)/k(0))^2$), we deduce that

$$\liminf_{t\to+\infty} \left(r(t)\mathbb{E}[U(X_t)] + \frac{\mathbb{E}[|Y_t|^2]}{2} \right) \geq \frac{\lambda_0}{4} > 0. \tag{6.1}$$

From now, let us argue by contradiction and assume that

$$\limsup_{t\to+\infty} r(t)\mathbb{E}[|X_t|^2] = 0. \tag{6.2}$$

From our hypothesis, there exists a suitable C such that $|\nabla U|^2 \leq C(1+U)$ and it follows that U is an under-quadratic potential: $U(x) \leq C(1+|x|^2)$. The hypothesis given by equation (6.2) trivially implies that $r(t)\mathbb{E}[U(X_t)] \to 0$. Thus,

$$\liminf_{t\to+\infty} \mathbb{E}[|Y_t|^2] \geq \frac{\lambda_0}{4} > 0. \tag{6.3}$$

We now focus on $K : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_+$ defined by $K(x,y,t) = \exp(R(t))\langle x, y\rangle$. First,

$$\mathcal{A}K(x,y,t) = e^{R(t)}(-|y|^2 + r(t)\langle x, \nabla U(x)\rangle).$$

Owing to Itô formula, we obtain

$$e^{R(t)}\mathbb{E}[\langle X_t, Y_t\rangle] = e^{R(0)}\langle x_0, y_0\rangle + \int_0^t e^{R(s)} \left(-\mathbb{E}[|Y_s|^2] + r(s)\mathbb{E}[\langle X_s, \nabla U(X_s)\rangle] \right) ds.$$

Since $|\langle x, \nabla U(x)\rangle| \leq C(1+|x|^2)$, it follows from (6.2) and (6.3) that there exists $\tilde{\alpha} > 0$ and $t_1 > 0$ such that for $s \geq t_1$,

$$\mathbb{E}[|Y_s|^2] - r(s)\mathbb{E}[\langle X_s, \nabla U(X_s)\rangle] \geq \tilde{\alpha}.$$

Then, since $r(t) \to 0$ and $\exp(R(t)) \to +\infty$, we deduce that:

$$\limsup_{t\to+\infty} \mathbb{E}[\langle X_t, Y_t\rangle] \leq -\tilde{\alpha} \liminf_{t\to+\infty} \frac{1}{e^{R(t)}} \int_0^t e^{R(s)} ds.$$

Recall that t_0 is such that $r'(t) + 2r^2(t) \geq 0$ for every $t \geq 0$. Then, an integration by parts yields

$$\frac{1}{e^{R(t)}} \int_{t_0}^t e^{R(s)} ds = \frac{1}{r(t)} - \frac{e^{R(t_0)-R(t)}}{r(t_0)} + \frac{1}{e^{R(t)}} \int_{t_0}^t \frac{r'(s)e^{R(s)}}{r(s)^2} ds \geq \frac{1}{r(t)} - \frac{e^{R(t_0)-R(t)}}{r(t_0)} - \frac{2}{e^{R(t)}} \int_{t_0}^t e^{R(s)} ds.$$

Thus,

$$\frac{1}{e^{R(t)}} \int_{t_0}^t e^{R(s)} ds \geq \frac{1}{3} \left(\frac{1}{r(t)} - \frac{e^{R(t_0)-R(t)}}{r(t_0)} \right) \xrightarrow{t\to+\infty} +\infty.$$

As a consequence, $\limsup_{t\to+\infty} \mathbb{E}[\langle X_t, Y_t\rangle] = -\infty$. Then, by Itô formula applied to $|X_t|^2$, we obtain that:

$$\mathbb{E}[|X_t|^2] = |x|^2 - 2\int_{t_0}^t \mathbb{E}[\langle X_s, Y_s\rangle] ds + \int_{t_0}^t \mathbb{E}[\mathrm{Tr}(\sigma\sigma^*)(X_s)] ds.$$

The fact that $r'(t) + 2r^2(t) \geq 0$ for t large enough implies that $r(t) \geq 1/(2t)$. Thus,

$$\limsup_{t\to+\infty} r(t)\mathbb{E}[|X_t|^2] \geq C \limsup_{t\to+\infty} \frac{1}{t}\mathbb{E}[|X_t|^2] \geq C \liminf_{t\to+\infty} (-\mathbb{E}[\langle X_t, Y_t\rangle]) = +\infty.$$

This is a contradiction with (6.2). □

7 Proof of unstability ($r_\infty = 0$ and quadratic potential)

Proof of Theorem 3.6: We set for any $t \geq 0$: $\phi_1(t) = \mathbb{E}[X_t]$ and $\phi_2(t) = \mathbb{E}[Y_t]$. One easily checks that ϕ_1 and ϕ_2 satisfy a simple coupled differential equation

$$\left\{ \begin{array}{l} \phi_1'(t) = \quad -\phi_2(t) \\ \phi_2'(t) = \frac{\alpha}{1+t}[\phi_1(t) - \phi_2(t)]. \end{array} \right.$$

Thus, ϕ_1 satisfies the second order differential equation $\phi_1''(t) + a(t)\phi_1'(t) + \phi_1(t) = 0$. This last equation is a particular case of second order differential equation with asymptotically small dissipation studied in Cabot et al. (2009b) where we set $a(t) = \frac{\alpha}{1+t}$. Since $\int_0^\infty a(s)ds = +\infty$, we can apply Corollary 3.2 of Cabot et al. (2009a) to obtain that

$$\mathbb{E}[X_t] \xrightarrow{t\to+\infty} 0 \quad \text{and,} \quad \mathbb{E}[Y_t] \xrightarrow{t\to+\infty} 0.$$

This shows i). Now, we study the second point ii) and we will show a sequence of technical lemmas. We use the notation f, g and h defined in Subsection 3.2.2.

Lemma 7.1 *Assume that $k(t) = (1+t)^\alpha$ with $\alpha \geq 1/2$. Then,*

$$\forall t > 0 \qquad r(t)f(t) + g(t) \geq \frac{1}{2}, \quad \text{and,} \quad \limsup_{t\to+\infty} r(t)f(t) + g(t) \leq \alpha.$$

Proof : We consider the application $F(t) = r(t)f(t) + g(t)$. F satisfies

$$\begin{aligned} F'(t) &= r'(t)f(t) + r(t) - 2r(t)h(t) + 2r(t)[h(t) - g(t)] \\ &= r'(t)f(t) + r(t) - 2r(t)g(t) \\ F'(t) &= \frac{r'(t)}{r(t)}F(t) + r(t) + g(t)\left[-2r(t) - \frac{r'(t)}{r(t)}\right]. \end{aligned} \tag{7.1}$$

Since $\alpha \geq 1/2$, that $r' + 2r^2 \geq 0$ for every $t \geq 0$ and F satisfies the inequality

$$F'(t) \leq \frac{r'(t)}{r(t)}F(t) + r(t) \quad \forall t \geq 0.$$

Thus, for any $t \geq 0$:

$$(F/r)'(t) = (F'/r)(t) - r'(t)F(t)/r^2(t) \leq 1,$$

and it follows that

$$F(t) = r(t)f(t) + g(t) \leq C(1+t)r(t) \xrightarrow{t\to+\infty} \alpha.$$

We focus now on the lower bound. We observe that

$$\forall t \geq 0 \qquad F'(t) = -2r(t)F(t) + f(t)(r'(t) + 2r^2(t)) + r(t).$$

Thus, for every $t \geq 0$, $F'(t) + 2r(t)F(t) \geq r(t)$ since $r' + 2r^2 \geq 0$. Hence,

$$\forall t \geq 0, \quad \left(F(t)e^{2\int_0^t r(s)ds} \right)' \geq r(t)e^{2\int_0^t r(s)ds}.$$

Using a simple integration and the fact that F is positive, we obtain

$$F(t) \geq F(0)e^{-2\int_0^t r(s)ds} + \int_0^t r(s)e^{-2\int_s^t r(u)du}ds \geq \frac{1}{2} \quad \forall t \geq 0.$$

This ends the proof of Lemma 7.1. □

The preceding lemma shows in particular that $(r(t)f(t))_{t\geq 0}$ and $(g(t))_{t\geq 0}$ are bounded functions. We now want to obtain the same property for h. This is the purpose of Lemma 7.2.

Lemma 7.2 *Assume that $k(t) = (1+t)^\alpha$ with $\alpha \geq 1/2$. Then, h is a bounded function on $\mathbb{R}_+$.*

Proof : First, one observes that

$$\begin{aligned}((1+t)^\alpha h(t))' &= (1+t)^{\alpha-1}\left(\alpha h(t) + (1+t)(-g(t) + \frac{\alpha}{1+t}(f(t)-h(t))\right)\\ &= (1+t)^\alpha[r(t)f(t) - g(t)] = (1+t)^{\alpha-1}(\alpha f(t) - (1+t)g(t)),\end{aligned}$$

and thus, that $h(t) = (1+t)^{-\alpha}\left(h(0) + \int_0^t s^{\alpha-1}\psi(s)ds\right)$ with $\psi(t) = \alpha f(t) - (1+t)g(t)$. This representation of h and the controls of f and g obtained previously suggest to study ψ. One checks that

$$\psi'(t) = \alpha + (2\alpha - 1)g(t) - 4\alpha h(t) \quad \forall t \geq 0,$$

and that ψ satisfies the second order differential equation

$$(1+t)\psi'' + (2\alpha - 1/2)\psi' + 4\alpha\psi = \alpha(2\alpha - 1/2) - (\alpha - \frac{1}{2})g.$$

We build now a Lyapunov function for the second order differential equation written above and consider $\mathcal{C}$ given as

$$\mathcal{C}(t) = \psi^2(t) + \frac{1+t}{4\alpha}\psi'^2(t).$$

A simple derivation shows that

$$\mathcal{C}'(t) = -(1 - \frac{1}{2\alpha})\psi'^2(t) + \psi'(t)B(t),$$

where B is the function defined as

$$B(t) = (\alpha - \frac{1}{4}) + (\frac{1}{4\alpha} - \frac{1}{2})g(t).$$

From Lemma 7.1, we know that B is bounded and the elementary inequality $|uv| \leq \varepsilon u^2/2 + v^2/(2\varepsilon)$ (with $u, v \in \mathbb{R}$ and $\varepsilon > 0$), applied with $u = \psi'(t)$, $v = B(t)$ and $\varepsilon = 1 - 1/(2\alpha) > 0$ since $\alpha > 1/2$, yields

$$\mathcal{C}'(t) \leq \frac{1}{2 - \frac{1}{\alpha}}B^2(t) \leq C < +\infty.$$

It follows that for every $t \geq 0$, $\mathcal{C}(t) \leq C(1+t)$. Then, the construction of $\mathcal{C}$ implies that ψ' is a bounded function. Since h satisfies $4\alpha h = \alpha + (2\alpha - 1)g - \psi'$, we easily conclude that h is also bounded. □

We have shown that rf, g and h are bounded functions (Lemmas 7.1 and 7.2). It is then natural to use $\tilde{f} = rf$ and to study the asymptotic behaviour of $\tilde{z}$ defined by $\tilde{z}(t) = (\tilde{f}, g, h)^T$ for every $t \geq 0$. Using $(\mathcal{S})$, we observe that $\tilde{z}$ is a solution to

$$\tilde{z}'(t) = \tilde{M}_t\tilde{z}(t) + r(t)\delta,$$

with $\delta = (1,0,0)^T$ and,

$$\tilde{M}_t = \begin{pmatrix} \frac{r'}{r} & 0 & -2r \\ 0 & -2r(t) & 2r(t) \\ 1 & -1 & -r(t) \end{pmatrix}.$$

Let us denote by $\tilde{z}^{(t)}$ the t-shifted trajectory of $\tilde{z}$, *i.e.* defined for every $s \geq 0$ by $\tilde{z}^{(t)}(s) = \tilde{z}(t+s)$. We have the next lemma.

Lemma 7.3 *The family $(\tilde{z}^{(t)}(.))_{t\geq 0}$ is relatively compact for the topology of uniform convergence on compact sets and every limit function $\tilde{z}^{(\infty)}$ is a stationary solution to $y' = \tilde{M}_\infty y$ with $\tilde{M}_\infty = \lim_{t\to+\infty} \tilde{M}_t$. As a consequence,*

$$\lim_{t\to\infty} r(t)f(t) - g(t) = 0.$$

Proof : We are going to apply the Ascoli Theorem to $(\tilde{z}^{(t)})_{t\geq 0}$. By Lemmas 7.1 and 7.2, $\tilde{z}$ is a bounded vector of $\mathbb{R}^3$ on $\mathbb{R}_+$. Thus,

$$\sup_{t\geq 0} \|\tilde{z}^{(t)}(0)\| < +\infty.$$

Furthermore, $(\tilde{z}^{(t)})_{t\geq 0}$ is equicontinuous. Actually,

$$\forall (t,T) \in \mathbb{R}_+^2, \quad \forall (u,v) \in [0;T]^2 \qquad \tilde{z}^{(t)}(u) - \tilde{z}^{(t)}(v) = \int_{t+v}^{t+u} [\tilde{M}_s \tilde{z}(s) + r(s)\delta] ds.$$

and it follows from the boundedness of $(\tilde{M}_t)_{t\geq 0}$, $(\tilde{z}_t)_{t\geq 0}$ and $(r(t))_{t\geq 0}$ that, there exists $C > 0$ such that

$$\forall (u,v) \in [0;T]^2 \qquad \|\tilde{z}^{(t)}(u) - \tilde{z}^{(t)}(v)\| \leq C|u-v|.$$

Then, by the Ascoli theorem, $(\tilde{z}^{(t)})_{t\geq 0}$ is relatively compact. Since $((\tilde{z}^{(t)})'(.))_{t\geq 0} = \tilde{M}_. \tilde{z}^{(t)}(.) + r^{(t)}(.)\delta$, it is immediate to check that $((\tilde{z}^{(t)})'(.))_{t\geq 0}$ is also a relatively compact. Denoting by $\tilde{z}^{(\infty)} = (\tilde{f}^{(\infty)}, g^{(\infty)}, h^{(\infty)})^T$, a limit point of $(\tilde{z}^{(t)})_{t\geq 0}$, we deduce that $(\tilde{z}^{(\infty)})'(.) = \tilde{M}_\infty \tilde{z}^{(\infty)}(.)$ (where we also used that $\lim_{t\to+\infty} r(t) = 0$). It follows that

$$(\tilde{f}^{(\infty)})' = 0 \qquad (g^{(\infty)})' = 0 \qquad (h^{(\infty)})' = f^{(\infty)} - g^{(\infty)}.$$

Moreover, there exists $(C_1, C_2, C_3) \in \mathbb{R}_+^2 \times \mathbb{R}$ such that $\tilde{f}^{(\infty)} = C_1$ and $g^{(\infty)} = C_2$ and $h^{(\infty)}(u) = (C_1 - C_2)u + C_3$. Now, by Lemma 7.2, $h^{(\infty)}$ is clearly bounded, and we conclude that $C_1 = C_2$. It follows that $\tilde{z}^{(\infty)}$ is stationary and that $0 = \tilde{f}^{(\infty)}(0) - g^{(\infty)}(0) = \lim_{t\to+\infty}(r(t)f(t) - g(t))$.

We now end the proof of the theorem. Recall that $F(t) = r(t)f(t) + g(t)$. Since $r(t)f(t) - g(t) \xrightarrow{t\to+\infty} 0$, we choose to express F' as follows:

$$\begin{aligned} F'(t) &= \frac{1}{2}\left(\frac{r'}{r}(t) - 2r(t)\right)(r(t)f(t) + g(t)) + \frac{1}{2}\left(\frac{r'}{r}(t) + 2r(t)\right)(r(t)f(t) - g(t)) + r(t) \\ &= -\frac{2\alpha+1}{2(1+t)}F(t) + \frac{2\alpha-1}{2(1+t)}(r(t)f(t) - g(t)) + \frac{a}{1+t} \end{aligned}$$

It follows that G defined by $G(t) = (1+t)^{\frac{2\alpha+1}{2}}F(t)$ satisfies

$$G'(t) = \frac{2\alpha-1}{2}(1+t)^{\frac{2\alpha-1}{2}}(r(t)f(t) - g(t)) + \alpha(1+t)^{\frac{2\alpha-1}{2}}.$$

Then, by an integration, we obtain

$$F(t) = (1+t)^{-\frac{2\alpha+1}{2}}F(0) + \alpha(1+t)^{-\frac{2\alpha+1}{2}} \times \frac{2}{2\alpha+1}\left(t^{\frac{2\alpha+1}{2}} - 1\right) + \frac{2\alpha-1}{2}t^{-\frac{2\alpha+1}{2}}\int_0^t s^{\frac{2\alpha-1}{2}}(r(s)f(s) - g(s))\,ds.$$

We observe that up to a constant, the last term can be written $(b(t))^{-1}\int_0^t b'(s)(r(s)f(s) - g(s))ds$ with $b(t) = t^{\frac{2\alpha+1}{2}} \xrightarrow{t\to+\infty} +\infty$. Then, thanks to a Cesaro-type argument, it follows that

$$\frac{2\alpha-1}{2}t^{-\frac{2\alpha+1}{2}}\int_0^t s^{\frac{2\alpha-1}{2}}(r(s)f(s) - g(s))\,ds \xrightarrow{t\to+\infty} 0.$$

As a consequence, $F(t) \longrightarrow_{t\to+\infty} \frac{2\alpha}{2\alpha+1}$, and Lemma 7.3 leads to

$$\mathbb{E}[X_t^2] \sim \frac{t}{2\alpha+1}, \qquad \text{and} \qquad \lim_{t\to+\infty} \mathbb{E}Y_t^2 = \frac{\alpha}{2\alpha+1}.$$

Finally, since h is bounded, $\frac{1}{\sqrt{t}}\mathbb{E}[X_tY_t] \xrightarrow{t\to+\infty} 0$. Thus, assertion iii) of the theorem follows using that in the Gaussian case, the convergence in distribution follows from that of the covariance matrix. This ends the proof of the theorem. □

8 Appendix

We detail the proof of Lemma 4.2(ii) or more precisely, we explain how the proof of Delarue and Menozzi (2010) can be used in our context. With the notations introduced in (4.6), (2.3) can be written as $dZ_t = F(Z_t)dt + B\sigma(X_t)dW_t$. Since Assumption $(\mathbf{A})$ is satisfied on a ball $B(z^*,\rho)$ ($\rho > 0$), we can easily build $(\tilde{Z}_t)$ solution to $d\tilde{Z}_t = \tilde{F}(\tilde{Z}_t)dt + B\tilde{\sigma}(\tilde{X}_t)dW_t$ (with the same Brownian Motion) such that $\tilde{F}(x) = F(x)$ and $\tilde{\sigma}(x) = \sigma(x)$ on $B(z^*,\rho)$.

For the process $(\tilde{Z}_t)_{t\geq 0}$, Theorem 1.1 of Delarue and Menozzi (2010) holds and one may observe that the lower bound obtained in this theorem is a consequence of equation numbered $(4.21-\mathbf{DM})$ and of the control of its remainder denoted by $R_{T-\varepsilon}$. Here, the remainder is denoted by $R^{\tilde{Z}}_{T-\varepsilon}$ in order to specify the process involved. Then, we first emphasize two points:

(a) $R^{\tilde{Z}}_{T-\varepsilon}$ is a functional of $(W_t)_{t\in[0,T-\varepsilon]}$, $(\tilde{\phi}_t)_{t\in[0,T-\varepsilon]}$, $(v_t)_{t\in[0,T-\varepsilon]}$ and $(\tilde{\chi}_t)_{t\in[0,T-\varepsilon]}$ where:

- $(\tilde{\phi}_t)_{t\in[0,T]}$ is a particular solution to the control problem $\dot{\tilde{\phi}} = \tilde{F}(\tilde{\phi}) + B\tilde{\sigma}(\tilde{\phi}^1)\varphi$ with $\tilde{\phi}(0) = z_1$, $\tilde{\phi}(T) = z_2$ (with $(\varphi_t)_{t\in[0,T]}$ being such that $\int_0^T |\varphi_s|^2 ds < +\infty$).
- $(v_t)_{t\in[0,T-\varepsilon]}$ is a progressively measurable stochastic process such that $\mathbb{E}[\int_0^{T-\varepsilon} |v_t|^2 dt]$.
- $(\tilde{\chi}_t)_{t\in[0,T]}$ is a solution to $d\tilde{\chi}_t = (\tilde{F}(\tilde{\chi}_t) + Bv_t)dt + B\tilde{\sigma}(\tilde{\chi}_t^1)dW_t$ with $\tilde{\chi}_0 = z_1$.

(b) If $(\tilde{\phi}_t)$ and (v_t) are such that Proposition 4.2 and 4.3 of (Delarue & Menozzi, 2010) hold, then there exists a measurable set $\bar{\mathcal{C}}$ such that $\mathbb{P}(\bar{\mathcal{C}}) \geq 1/2$ and such that for every measurable $\mathcal{C} \in \bar{\mathcal{C}}$,

$$\mathbb{E}[R^{\tilde{Z}}_{T-\varepsilon}\mathbb{1}_{\mathcal{C}}] \leq C_T(1 + |\theta_T(z_1) - z_2|^2),$$

where $(\theta_t(z_1))_{t\geq 0}$ always denotes the solution to $\dot{\theta} = F(\theta)$ starting from z_1.

Second, we show that we can find a sufficiently small ball $B(z^*,\delta)$ such that, for every $z_1, z_2 \in B(z^*,\delta)$, $(\tilde{\phi}_t)_{t\in[0,T]}$, built with a control $(\phi_t)_{t\in[0,T]}$ that satisfies the conclusions of Proposition 4.2 of Delarue and Menozzi (2010), is always included in $B(z^*,\rho)$ for any time $t \in [0;T]$. In this view, set $\psi(x,y) = U(x) + \frac{|y|^2}{2\lambda}$. We have

$$\psi(\tilde{\phi}(t)) - \psi(z_1) \leq C\int_0^t \langle \nabla\psi(\tilde{\phi}(s)), B\tilde{\sigma}(\tilde{\phi}^1(s))\varphi_s\rangle ds.$$

Let $\tau_\rho := \inf\{t \geq 0, |\tilde{\phi}(t) - z^*| = \rho\} \wedge T$. From Proposition 4.2 of (Delarue & Menozzi, 2010) and denoting by $M_\rho = \sup_{x\in B(x^*,\rho)} |\nabla\psi(x)|.\|\sigma(x)\|$, we deduce that,

$$\psi(\tilde{\phi}(t\wedge\tau_\rho)) - \psi(z_1) \leq C_T M_\rho |\theta_{t\wedge\tau_\rho}(z_1) - z_2|^2. \tag{8.1}$$

Now, by our non explosive result given in (4.8), we have that for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every $z_1, z_2 \in B(z^*,\delta)$, $\sup_{t\geq 0} |\theta_t(z_1) - z_2|^2 \leq \varepsilon$. By (8.1), and the strict convexity of U on $B(z^*,\rho)$, it follows that for a suitable choice of ϵ and δ

$$\sup_{|z_1-z^*|\leq\delta} \psi(z_1) + C_T M_\rho \epsilon < \inf_{|x-x^*|=\rho} U(x).$$

Hence, for $\varepsilon > 0$ and δ small enough,

$$\sup_{z_1,z_2\in B(z^*,\delta)} \sup_{t\in[0,T]} \psi(\tilde{\phi}(t\wedge\tau_\rho)) < \inf_{|x-x^*|=\rho} U(x).$$

It implies that $\tau_\rho = T$ and we deduce that for every $z_1, z_2 \in B(z^*,\delta)$, $(\tilde{\phi}_t)_{t\in[0,T]}$ is always included in $B(z^*,\rho)$ for any $t \in [0;T]$.

We can now focus on $(Z_t)_{t\geq 0}$ itself. From the latter argument, one deduces that $(\tilde{\phi}(t))_{t\in[0,T]}$ is a solution to $\dot{z}_t = F(z_t) + B\sigma(x_t)\varphi_t$ with $\tilde{\phi}(0) = z_1$ and $\tilde{\phi}(T) = z_2$. We then denote it by $(\phi_t)_{t\in[0,T]}$. Following carefully the construction of $(4.21 - \mathbf{DM})$, one then checks that this equation also holds for (Z_t).

If (v_t) is built as in Proposition 4.3 of (Delarue & Menozzi, 2010), then $R^Z_{T-\varepsilon} = R^{\tilde{Z}}_{T-\varepsilon}$ on the set $\mathcal{D} = \{\omega \in \Omega, \forall t \in [0, T-\varepsilon], \chi_t(\omega) = \tilde{\chi}_t(\omega)\}$ (where $(\chi_t)_{t\in[0,T]}$ denotes the solution to $d\chi_t = (F(\chi_t) + Bv_t)dt + B\sigma(\chi^1_t)dW_t$ with $\chi_0 = z_1$). From (b), one deduces that the lower bound of Theorem 1.1 remains true for (Z_t) itself if $\mathbb{P}(\mathcal{D}) > 1/2$ (so that $\mathbb{P}(\mathcal{D}\cap\mathcal{C}) > 0$).

In fact,

$$\mathcal{D} = \{\omega \in \Omega, \forall t \in [0, T-\varepsilon], \tilde{\chi}_t \in B(z^*,\rho)\}$$

and it remains to prove that

$$\mathbb{P}(\sup_{t\in[0,T-\varepsilon]} |\chi_t - z^*| \geq \rho) < 1/2.$$

We can find δ small enough such that for every $z_1, z_2 \in B(z^*,\delta)$, $(\tilde{\phi}_t)_{t\in[0,T-\varepsilon]} \in B(z^*,\rho/2)$, we only have to check that $\mathbb{P}(\sup_{t\in[0,T-\varepsilon]} |\tilde{\chi}_t - \tilde{\phi}_t\| \geq \frac{\rho}{2}) < 1/2$.
Writing $\tilde{\chi}_t - \tilde{\phi}_t = \tilde{\Theta}_t - \Gamma_t + \Gamma_t$ with $\tilde{\Theta}_t = \tilde{\chi}_t - \tilde{\phi}_t$ and $(\Gamma_t)_{t\geq 0}$ being a Gaussian process defined by (3.15-**DM**), we have

$$\mathbb{P}(\sup_{t\in[0,T-\varepsilon]} |\tilde{\chi}_t - \tilde{\phi}_t\| \geq \frac{\rho}{2}) < \mathbb{P}(\sup_{t\in[0,T-\varepsilon]} |\tilde{\Theta}_t - \Gamma_t| \geq \frac{\rho}{4}) + C_\rho \sup_{t\in[0,T]} \mathbb{E}[|\Gamma_t|^2].$$

Then, we deduce from Proposition 4.3 and Lemma 3.9 of Delarue and Menozzi (2010) that for T small enough, $\mathbb{P}(\mathcal{D}) > 1/2$. This concludes the proof. □

References

Alvarez, F. (2000). On the minimizing property of a second order dissipative system in hilbert spaces. *SIAM J. Control Optim.*, *38*, 110–1119.

Alvarez, F., Attouch, H., Bolte, J., & Redont, P. (2002). A second-order gradient-like dissipative dynamical system with hessian-driven damping. application to optimization and mechanics. *Journal des Mathématiques Pures et Appliquées*, *81*, 747–779.

Antipin, A. (1994). Minimization of convex functions on convex sets by means of differential equations (in russian). *Differential Equations*, *30*, 1365–1375.

Bakhtin, Y. (2002). Existence and uniqueness of stationary solution of nonlinear stochastic differential equation with memory. *Theory Probab. Appl.*, *47*, 764–769.

Bakhtin, Y. (2006). Lyapunov exponents for stochastic differential equations with infinite memory and applucation to stochastic navier-stokes equations. *Discrete and Continuous dynamical systems, series B*, *6*, 697–709.

Bakhtin, Y., & Mattingly, J. (2005). Stationary solutions of stochastic differential equations with memory and stochastic partial differential equations. *Communications in Contemporary Mathematics*, *7*, 553–582.

Bakry, D., Cattiaux, P., & Guillin, A. (2008). Rate of convergence for ergodic continuous markov processes : Lyapunov versus poincare. *Journal of Functional Analysis*, *254*, 727–759.

Bally, V., & Kohatsu-Higa, A. (2010). Lower bounds for densities of Asian type stochastic differential equations. *J. Funct. Anal.*, *258*(9), 3134–3164. Available from http://dx.doi.org/10.1016/j.jfa.2009.10.027

Benaïm, M., & Hirsh, M. (1996). Asymptotic pseudotrajectories and chain recurrent flows, with applications. *J. Dynam. Differential Equations*, *8*, 141–176.

Benaïm, M., Ledoux, M., & Raimond, O. (2002). Self-interacting diffusions. *Probab. Theory Related Fields*, *122*, 1–41.

Benaïm, M., & Raimond, O. (2003). Self-interacting diffusions ii: Convergence in law. *Ann. Inst. H. Poincaré Probab. Statist.*, *39*, 1043–1055.

Benaïm, M., & Raimond, O. (2005). Self-interacting diffusions iii: Symmetric interactions. *Ann. Proba.*, *33*, 1716–1759.

Bolley, F., Guillin, A., & Malrieu, F. (2010). Trend to equilibrium and particle approximation for a weakly selfconsistent vlasov-fokker-planck equation. *Mathematical Modelling and Numerical Analysis, to appear*.

Cabot, A. (2009). Asymptotics for a gradient system with memory term. *Proc. of the American Mathematical Society*, *9*, 3013–3024.

Cabot, A., Engler, H., & Gadat, S. (2009a). On the long time behavior of second order differential equations with asymptotically small dissipation. *Trans. of the American Mathematical Society*, *361*, 5983–6017.

Cabot, A., Engler, H., & Gadat, S. (2009b). Second-order differential equations with asymptotically small dissipation and piecewise flat potentials. *Electronic Journal of Differential Equation*, *17*, 33–38.

Cattiaux, P., & Mesnager, L. (2002). Hypoelliptic non-homogeneous diffusions. *Probability Theory and Related Fields*, *123*, 453–483.

Chaleyat-Maurel, M., & Michel, D. (1984). Hypoellipticity theorems and conditionnal laws. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, *65*, 573–597.

Coppersmith, D., & Diaconis, P. (1987). Random walk with reinforcement. *Unpublished*.

Coron, J.-M. (2007). *Control and nonlinearity* (Vol. 136). Providence, RI: American Mathematical Society.

Cranston, M., & LeJan, Y. (1995). Self-attracting diffusions: two case studies. *Math. Ann.*, *303*, 87–93.

Da Prato, G., & Zabczyk, J. (1996). *Ergodicity for infinite-dimensional systems* (Vol. 229). Cambridge: Cambridge University Press. Available from http://dx.doi.org/10.1017/CBO9780511662829

Delarue, F., & Menozzi, S. (2010). Density estimates for a random noise propagating through a chain of differential equations. *J. Funct. Anal.*, *259*(6), 1577–1630. Available from http://dx.doi.org/10.1016/j.jfa.2010.05.002

Douc, R., Fort, G., & Guillin, A. (2009). Subgeometric rates of convergence of f-ergodic strong markov processes. *Stochastic Processes and their Applications*, *119*, 897–923.

Down, D., Meyn, S., & Tweedie, R. (1995). Exponential and uniform ergodicity of markov processes. *The Annals of Probability*, *23*, 1671–1691.

Durrett, R., & Rogers, L. (1992). Asymptotic behavior of brownian polymers. *Probab. Theory Related Fields*, *3*, 337–349.

Ethier, S., & Kurtz, T. (1986). *Markov processes, first edition*. John Wiley & Sons Inc., New York.

Hairer, M. (2011). On malliavin's proof of hörmander's theorem. *Bull. Sci. Math.*, *165*.

Haraux, A. (1991). *Systèmes dynamiques dissipatifs et applications*. R.M.A., Masson, Paris.

Hormander, L. (1967). Hypoelliptic second order differential equations. *Acta Mathematica*, *117*, 147–171.

Ikeda, N., & Watanabe, S. (1981). *Stochastic differential equations and diffusion processes*. North Holland, Amsterdam/Kodansha, Tokyo.

Kohn, J. J. (1978). Lectures on degenerate elliptic problems. In *Pseudodifferential operator with applications (Bressanone, 1977)* (pp. 89–151). Naples: Liguori.

Kurtzman, A. (2009). The ode method for some self-interacting diffusions. *Ann. Inst. H. Poincaré Probab. Statist., to appear*.

Kurtzman, A., & Chambeu, S. (2009). Some particular self-interacting diffusions: ergodic behavior and almost sure convergence. *Preprint*.

Lamberton, D., & Pagès, G. (2003). Recursive computation of the invariant distribution of a diffusion: the case of a weakly mean reverting drift. *Stochastics and Dynamics*, *3*(4), 435–451.

Pemantle, R. (1992). Vertex-reinforced random walk. *Probab. Theory Related Fields*, *1*, 117–136.

Polyak, B. (1987). *Introduction to optimization*. Optimization Software, New York.

Raimond, O. (1997). Self-attracting diffusions: Case of the constant interaction. *Probab. Theory Related Fields*, *107*.

Stroock, D. W., & Varadhan, S. R. S. (1972). On the support of diffusion processes with applications to the strong maximum principle. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory* (pp. 333–359). Berkeley, Calif.: Univ. California Press.

Trèves, F. (1980). *Introduction to pseudodifferential and Fourier integral operators. Vol. 1*. New York: Plenum Press. (Pseudodifferential operators, The University Series in Mathematics)

Villani, C. (2009). Hypocoercivity. *Mem. Amer. Math. Soc.*, *202*(950), iv+141.

Wu, L.-M. (2001). Large and moderate deviations and exponential convergence for stochastic damping hamilton systems. *Stochastic Processes and their Applications*, *91*, 205–238.