THERMALISATION FOR STOCHASTIC SMALL RANDOM PERTURBATIONS OF HYPERBOLIC DYNAMICAL SYSTEMS

BARRERA, GERARDO AND JARA, MILTON

Abstract. In this article we study the so-called thermalisation or cut-off for a family of stochastic small perturbations of a given dynamical system. We focus in a semi-flow of a deterministic differential equation with a unique hyperbolic fixed point. We add to the deterministic dynamics a Brownian motion of small variance. Assuming that the vector field is strongly coercive, we prove that the family of perturbed dynamical system always presents a thermalisation (cut-off) in total variation distance. Moreover, we give a necessary and sufficient condition to have profile thermalisation (profile cut-off).

1. Introduction

Our main goal is the study of the convergence to equilibrium for a family of stochastic small perturbations of a given dynamical system. More precisely, we are concerned about the abrupt convergence to equilibrium of systems of the form:

$$d x^\varepsilon(t) = -F(x^\varepsilon(t)) + \sqrt{\varepsilon} dB(t), \ t \geq 0,$$

where $F$ is a given vector field with a unique hyperbolic fixed point and \{B(t) : t \geq 0\} is a Brownian motion.

When the intensity $\varepsilon$ of the noise goes to zero, the total variation distance between the law of the stochastic dynamics and the law of its equilibrium comes from one to zero abruptly.

Dynamical systems subjected to small Gaussian perturbations have been studied extensively, see the book of M. Freidling & A. Wentzell [17] which discusses this problem in great detail; see also M. Freidling & A. Wentzell [18], [19], M. Day [20], [21] and W. Siegert [24]. This treatment has inspired many works and considerable effort was concerned about purely local phenomena, i.e., on the computation of exit times and exit probabilities from neighborhoods of fixed points that are carefully stipulated not to contain any other fixed point of the dynamics.
The theory of large deviations allows to solve the exit problem from the domain of attraction of a stable point. It turns out that the mean exit time is exponentially large in the small noise parameter, and its logarithmic rate is proportional to the height of the potential barrier the trajectories have to overcome. Consequently, for a multi-well potential one can obtain a series of exponentially non-equivalent time scales given by the wells mean exit times. Moreover, one can prove that the normalised exit times are exponentially distributed and have a memoryless property, see A. Galves, E. Olivieri & M. Vares \[2\], E. Olivieri & M. Vares \[9\] and C. Kipnis & C. Newman \[4\].

The term “cut-off” was introduced by D. Aldous and P. Diaconis in \[6\] and \[7\] in the early eighties to describe the phenomenon of abrupt convergence of Markov chains modeling card shuffling. Since the appearance of \[6\] many families of stochastic processes have been shown to have similar properties. Various notions of cut-off have been proposed; see J. Barrera & B. Ycart \[14\] and P. Diaconis \[22\] for an account. In \[12\], L. Saloff-Coste gives an extensive list of random walks for which the cut-off phenomenon holds.

Roughly speaking, thermalisation or cut-off holds for a family of stochastic systems, when convergence to equilibrium happens in a time window which is small compared to the total running time of the system. Before a certain “cut-off time” those processes stay far from equilibrium with respect to some suitable distance; in a time window of smaller order the processes get close to equilibrium, and after that convergence to equilibrium happens exponentially fast.

Alternative names are threshold phenomenon and abrupt convergence. When the distance to equilibrium at the time window can be well approximated by some profile function, we speak about profile cut-off. Sequences of stochastic processes for which an explicit profile cut-off can be determined are scarce. Explicit profiles are usually out of reach, in particular for the total variation distance. In general, the existence of the phenomenon is proven through a precise estimation of the sequence of cut-off times and this precision comes at a high technical price, for more details see J. Barrera, O. Bertoncini & R. Fernández \[15\].

The main result of this article, Theorem 2.2, states that when the deterministic dynamics is strongly coercive, the family of perturbed dynamics presents a thermalisation (cut-off) as we describe in Section 2. Moreover, in Corollary 2.5 we give a necessary and sufficient condition to have profile thermalisation (profile cut-off). We point out that our sufficient condition is satisfied by reversible dynamics; i.e., when \( F(x) = \nabla V(x) \), but also for dynamics that are non-reversible. Non-reversible dynamics naturally appear for example in polymeric fluid flow or Wigner-Fokker-Planck equations, see A. Arnold, J. Carrillo & C. Manzini \[11\] and B. Jourdain, C. Le Bris, T. Lelièvre & F. Otto \[3\].

Notice that the set of symmetric matrices is not open. In particular, reversibility is not a generic property of dynamical systems. In the other hand, hyperbolicity is an open property, meaning that it is stable under
small perturbations of the vector field. Moreover, in the non reversible case, there is not explicit formula for the invariant measure as in the reversible case (Gibbs measure). Therefore it is desirable to have a treatment that does not rely on these properties, namely reversibility and/or explicit knowledge of invariant measures.

This material will be organized as follows. Section 2 describes the model and states the main result besides establishing the basic notation and definitions. Section 3 provides sharp estimates on the asymptotics of a related linear approximation which is the main ingredient in order to prove the main result in the end of this section. The Appendix is divided in three sections as follows: Section A gives useful properties for the total variation distances between Gaussian distributions. Section B and C provides the rigorous arguments about the deterministic dynamics and the stochastic dynamics respectively that we omit in Section 3 to make the presentation more fluid.

2. Notation and results

In this section we rigorously state the family of stochastically perturbed dynamical systems that we are considering and the results we prove.

2.1. The dynamical system. Let \( F : \mathbb{R}^d \to \mathbb{R}^d \) be a vector field of class \( C^2(\mathbb{R}^d, \mathbb{R}^d) \). For each \( x \in \mathbb{R}^d \), let \( \{ \varphi(t, x) : t \in [0, \tau_x) \} \) be the solution of the deterministic differential equation:

\[
\begin{align*}
\frac{d}{dt} \varphi(t) &= -F(\varphi(t)), \quad t \geq 0 \\
\varphi(0) &= x.
\end{align*}
\]

Due to \( F \) is smooth, this equation has a unique solution. Since we have not imposed any growth condition on \( F \), the explosion time \( \tau_x \) may be finite. Denoted by \( \lVert \cdot \rVert \) the Euclidean norm in \( \mathbb{R}^d \) and by \( \langle \cdot, \cdot \rangle \) the standard inner product of \( \mathbb{R}^d \). Under the condition

\[
\sup_{z \in \mathbb{R}^d} \frac{\langle z, -F(z) \rangle}{1 + \|z\|^2} < +\infty,
\]

the explosion time \( \tau_x \) is infinite for any \( x \in \mathbb{R}^d \). Later on, we will make stronger assumptions on \( F \), so we will assume that the explosion time is always infinite without further comments.

We call the family \( \{ \varphi(t, x) : t \geq 0, x \in \mathbb{R}^d \} \) the dynamical system associated to \( F \). We say that a point \( y \in \mathbb{R}^d \) is a fixed point of (2.1) if \( F(y) = 0 \). In that case \( \varphi(t, y) = y \) for any \( t \geq 0 \).

Let \( y \) be a fixed point of (2.1). We say that \( x \in \mathbb{R}^d \) belongs to the basin of attraction of \( y \) if

\[
\lim_{t \to +\infty} \varphi(t, x) = y.
\]

We say that \( y \) is an attractor of (2.1) if the set

\[
U_y = \{ x \in \mathbb{R}^d : x \text{ is in the basin of attraction of } y \}
\]
contains an open ball centered at \( y \). If \( U_y = \mathbb{R}^d \) we say that \( y \) is a **global attractor** of (2.1). We say that \( y \) is a hyperbolic fixed point of (2.1) if \( \text{Re}(\lambda) \neq 0 \) for any eigenvalue \( \lambda \) of the Jacobian matrix \( DF(y) \). By the Hartman-Grobman Theorem (see Theorem (Hartman) page 127 of [16] or the celebrated paper of P. Hartman [23]), a hyperbolic fixed point \( y \) of (2.1) is an attractor if and only if \( \text{Re}(\lambda) > 0 \) for any eigenvalue \( \lambda \) of the matrix \( DF(y) \). From now on, we will always assume that

\( 0 \) is a hyperbolic attractor of (2.1).

In that case, for any \( x \in U_0 \) the asymptotic behaviour of \( \varphi(t, x) \) at \( t \to +\infty \) can be described in a very precise way.

A sufficient condition for \( 0 \) to be a global attractor of (2.1) is the following **coercivity condition**: there exists a positive constant \( \delta \) such that

\[
\langle x, F(x) \rangle \geq \delta \| x \|^2 \text{ for any } x \in \mathbb{R}^d. \tag{C1}
\]

In this case, it is not very difficult to show that

\[
\|\varphi(t, x)\| \leq \| x \| e^{-\delta t} \text{ for any } x \in \mathbb{R}^d \text{ and any } t \geq 0.
\]

In other words, \( \varphi(t, x) \) converges to 0 exponentially fast as \( t \to +\infty \). In that case we have \( \text{Re}(\lambda) \geq \delta \) for any eigenvalue \( \lambda \) of the matrix \( DF(0) \).

**Lemma 2.1.** Let us suppose that the vector field \( F \) of (2.1) satisfies the coercivity condition (C1). For any \( x_0 \in \mathbb{R}^d \setminus \{0\} \) there exist \( \lambda := \lambda(x_0) > 0 \), \( \ell := \ell(x_0), m := m(x_0) \in \mathbb{N}, \theta_1 := \theta_1(x_0), \ldots, \theta_m := \theta_m(x_0) \in [0, 2\pi), v_1 := v_1(x_0), \ldots, v_m := v_m(x_0) \in \mathbb{C}^d \) linearly independent and \( \tau := \tau(x_0) > 0 \) such that

\[
\lim_{t \to +\infty} \left\| \frac{e^{\lambda t}}{\ell^{\ell-1}} \varphi(t + \tau, x_0) - \sum_{k=1}^m e^{i\theta_k t} v_k \right\| = 0.
\]

This lemma will be proved in Appendix B where we give a more detailed description of the constants and vectors appearing in this lemma. We can anticipate that the numbers \( \lambda \pm i\theta_k, \ k = 1, \ldots, m \) are eigenvalues of \( DF(0) \) and that the vectors \( v_k, \ k = 1, \ldots, m \) are elements of the Jordan decomposition of the matrix \( DF(0) \).

**2.2. The stochastic perturbations.** Let \( \{B(t) : t \geq 0\} \) be a standard Brownian motion in \( \mathbb{R}^d \) and let \( \epsilon \in (0, 1] \) be a scaling parameter. Let \( x_0 \in U_0 \setminus \{0\} \) and let \( \{x^\epsilon(t, x_0) : t \geq 0\} \) be the solution of the following stochastic differential equation:

\[
\begin{aligned}
\left\{ \begin{array}{ll}
dx^\epsilon(t) = -F(x^\epsilon(t))dt + \sqrt{\epsilon}dB(t), & t \geq 0, \\
x^\epsilon(0) = x_0.
\end{array} \right.
\tag{2.2}
\end{aligned}
\]

Stochastic differential equation (2.2) is used in molecular modeling. In that context \( \epsilon = 2\kappa\tau \), where \( \tau \) is the temperature, and \( \kappa \) is the Boltzmann constant. We will denote by \( (\Omega, \mathcal{F}, \mathbb{P}) \) the probability space where \( \{B(t) : t \geq 0\} \) is defined and we will denote by \( \mathbb{E} \) the expectation with respect to \( \mathbb{P} \). Notice that (2.2) has a unique strong solution (see Remark 2.1.2 page 57 of
In this case, we will prove that the law of \( x^\epsilon(t, x_0) : t \geq 0 \) can be taken as a stochastic process in the same probability space \((\Omega, \mathcal{F}, \mathbb{P})\). In order to avoid unnecessary notation, we will write \( \{x^\epsilon(t) : t \geq 0\} \) instead of \( \{x^\epsilon(t, x_0) : t \geq 0\} \) and \( \{\varphi(t) : t \geq 0\} \) instead of \( \{\varphi(t, x_0) : t \geq 0\} \).

Our aim is to describe in detail the asymptotic behaviour of the law of \( x^\epsilon(t) \) for large times \( t \), as \( \epsilon \to 0 \). In particular, we will be interested in the law of \( x^\epsilon(t) \) for times \( t \) of order \( \mathcal{O}(\log(1/\epsilon)) \), where thermalisation or cut-off phenomena will appear, depending on whether 0 is a local or global cut-off of \( (2.1) \).

Under \((C1)\), the process \( \{x^\epsilon(t) : t \geq 0\} \) has a unique invariant measure \( \mu^\epsilon \) which is absolutely continuous with respect to the Lebesgue measure in \( \mathbb{R}^d \).

In this case, we will prove that the law of \( x^\epsilon(t) \) converges in total variation distance to \( \mu^\epsilon \) in a time window

\[
w^\epsilon := \frac{1}{\lambda} + o(1)
\]

of order \( \mathcal{O}(1) \) around the mixing time

\[
t^\epsilon_{\text{mix}} := \frac{1}{2\lambda} \ln (1/\epsilon) + \frac{\ell - 1}{\lambda} \ln (\ln (1/\epsilon)) + \tau,
\]

where \( \lambda, \ell \) and \( \tau \) are the constants associated to \( x_0 \) in Lemma 2.1.

If we only assume that 0 is a hyperbolic attractor of \((2.1)\), we can not rule out the existence of other attractors. These attractors are accessible to the stochastic dynamics \((2.2)\) (a great part of the celebrated book of M. Freidlin & A. Wentzell [17] is devoted to the study of this problem).

The exact way on which this convergence takes place is the content of the following section.

2.3. The cut-off phenomenon. Let \( \mu, \nu \) be two probability measures in \((\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))\). We say that a probability measure \( \pi \) in \((\mathbb{R}^d \times \mathbb{R}^d, \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d))\) is a coupling between \( \mu \) and \( \nu \) if for any Borel set \( B \in \mathcal{B}(\mathbb{R}^d) \),

\[
\pi(B \times \mathbb{R}^d) = \mu(B) \quad \text{and} \quad \pi(\mathbb{R}^d \times B) = \nu(B).
\]

We say in that case that \( \pi \in \mathcal{C}(\mu, \nu) \). The total variation distance between \( \mu \) and \( \nu \) is defined as

\[
d_{TV}(\mu, \nu) = \inf_{\pi \in \mathcal{C}(\mu, \nu)} \pi\{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d : x \neq y\}.
\]

Notice that the diameter with respect to \( d_{TV}\) of the set \( \mathcal{M}^+_\text{ad} (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \) of probability measures in \((\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))\) is equal to 1. If \( X \) and \( Y \) are two random variables in \( \mathbb{R}^d \) which are defined in the same measurable space \((\Omega, \mathcal{F})\), we write \( d_{TV}(X, Y) \) instead of \( d_{TV}(\mathbb{P}(X \in \cdot), \mathbb{P}(Y \in \cdot)) \).

We say that a family of stochastic processes \( \{x^\epsilon(t) : t \geq 0\}_{\epsilon \in [0,1]} \) has thermalisation at position \( \{t^\epsilon\}_{\epsilon \in (0,1]} \) if

\[
\{x^\epsilon(t) : t \geq 0\}_{\epsilon \in [0,1]} \quad \text{and} \quad \{\omega^\epsilon\}_{\epsilon \in (0,1]} \quad \text{for large times} \quad \mathcal{O}(\log(1/\epsilon)) \]

if
Thermalisation for Perturbations of Dynamical Systems

i) $$\lim_{\epsilon \to 0} \frac{\omega^\epsilon}{t^\epsilon} = 0,$$

ii) $$\lim_{\epsilon \to 0} \limsup_{t \to \infty} d_{TV}(x^\epsilon(t + c\omega^\epsilon), \mu^\epsilon) = 0,$$

iii) $$\lim_{\epsilon \to 0} \liminf_{t \to -\infty} d_{TV}(x^\epsilon(t + c\omega^\epsilon), \mu^\epsilon) = 1.$$

If \( \{x^\epsilon(t) : t \geq 0\} \) is a Markov process with a unique invariant measure and \( \mu^\epsilon \) is the invariant measure of the process \( \{x^\epsilon(t) : t \geq 0\} \) we say that \( \{x^\epsilon(t) : t \geq 0\}_{\epsilon \in (0,1]} \) presents thermalisation or cut-off.

If in addition to i) there is a continuous function \( G : \mathbb{R} \to [0,1] \) such that \( G(-\infty) = 1 \), \( G(+\infty) = 0 \) and

$$\text{ii') } \lim_{\epsilon \to 0} d_{TV}(x^\epsilon(t + c\omega^\epsilon), \mu^\epsilon) = G(c)$$

for every \( c \in \mathbb{R} \). We say that there is profile thermalisation or profile cut-off. Notice that ii') implies ii),iii) and therefore profile thermalisation (respectively profile cut-off) is a stronger notion than thermalisation (respectively cut-off).

2.4. Results. Denote by \( \mathcal{G}(v, \Xi) \) the Gaussian distribution in \( \mathbb{R}^d \) with vector mean \( v \) and positive definite covariance matrix \( \Xi \). Let \( I_d \) be the identity \( d \times d \)-matrix. Given a matrix \( A \), denote by \( A^\ast \) the transpose matrix of \( A \). In the case of a stochastic perturbation of a dynamical system with a strongly coercive hyperbolic attractor, we prove thermalisation:

**Theorem 2.2.** Assume that the vector field \( F \) of (2.2) satisfies (3.1), (3.2), and (3.3). Let \( \{x^\epsilon(t,x_0) : t \geq 0\} \) be the solution of (2.2) and denote by \( \mu^\epsilon \) the unique invariant probability measure for the evolution given by (2.2). Let \( d^\epsilon(t) = d_{TV}(x^\epsilon(t,x_0), \mu^\epsilon(t)) \), \( t \geq 0 \) and suppose that \( x_0 \neq 0 \). Let us consider the mixing time \( t_{mix}^\epsilon \) which is given by (2.4) and the time window which given by (2.3). Let \( \vartheta \in (0,1/16) \) and define \( \delta_\epsilon = \epsilon^{\vartheta} \). For any \( c \in \mathbb{R} \) we have

$$\lim_{\epsilon \to 0} \left| d^\epsilon(t_{mix}^\epsilon + \delta_\epsilon + c\omega^\epsilon) - \bar{D}^\epsilon(t_{mix}^\epsilon + \delta_\epsilon + c\omega^\epsilon) \right| = 0,$$

where

$$\bar{D}^\epsilon(t) = d_{TV}\left( \mathcal{G}\left( \frac{(t - \tau)\ell - 1}{e^{\lambda(t - \tau)} \sqrt{c}} \sum_{k=1}^{m} e^{i\theta_k(t - \tau)} v_k, I_d \right), \mathcal{G}(0, I_d) \right), \ t \geq \tau$$

with \( \lambda, \ell, \tau, \theta_1, \ldots, \theta_m \in [0,2\pi) \), \( v_1, \ldots, v_m \) are the constants and vectors associated to \( x_0 \) in Lemma 2.1 and the matrix \( \Sigma \) is the unique solution of the matrix Lyapunov equation

$$DF(0)X + X(DF(0))^* = I_d.$$  \tag{2.5}
Remark 2.3. Notice that we $DF(0)$ is symmetric, $\Sigma$ is easily computable $\Sigma = \frac{1}{2}(DF(0))^{-1}$. When $DF(0)$ is not symmetric, the solution of the matrix Lyapunov equation (2.5) is unique, symmetric and positive definite matrix.

From the last theorem we have the following consequences that we write as a corollaries.

Corollary 2.4 (Termalisation). Suppose that $x_0 \neq 0$. Theorem 2.2 implies termalisation for $\{x^f(t,x_0) : t \geq 0\}$.

Proof. Notice
\[
0 < \liminf_{t \to +\infty} \left\| \sum_{k=1}^m e^{i\theta_k(t-\tau)} v_k \right\| \leq \limsup_{t \to +\infty} \left\| \sum_{k=1}^m e^{i\theta_k(t-\tau)} v_k \right\| \leq \sum_{k=1}^m \|v_k\|,
\]
where first inequality follows from the fact that $v_1, \ldots, v_m$ are linearly independent and the others inequalities are straightforward. By using Theorem 2.2, Lemma C.4, Lemma C.2 and Lemma C.3 we get the statement.

Corollary 2.5 (Profile termalisation). Suppose that $x_0 \neq 0$. There is profile termalisation for $\{x^f(t,x_0) : t \geq 0\}$ if and only if $\left\| \Sigma^{-1/2} \sum_{k=1}^m e^{i\theta_k t} v_k \right\|$ is constant for any $t \geq 0$.

Proof. This follows immediately from Theorem 2.2 together with Lemma C.3 and the fact that
\[
d_{TV}(G(v,I_d),G(0,I_d)) = d_{TV}(G(\bar{v},I_d),G(0,I_d))
\]
for any $v, \bar{v} \in \mathbb{R}^d$ such that $\|v\| = \|\bar{v}\|$.

The following corollary includes the case when the dynamics is reversible, i.e., $F = \nabla V$ for some scalar function $V : \mathbb{R}^d \to \mathbb{R}$.

Corollary 2.6. Suppose that $x_0 \neq 0$. If all the eigenvalues of $DF(0)$ are real then $\{x^f(t,x_0) : t \geq 0\}$ has profile termalisation.

Proof. The proof follows from the Corollary 2.5 observing that $\theta_j = 0$ for any $j = 1, \ldots, m$.

In dimension 2 and 3, we have always profile termalisation when $\Sigma$ is the identity matrix as the following corollary states.

Corollary 2.7. Suppose that $x_0 \neq 0$. If $d \in \{2,3\}$ and $\Sigma = I_d$ then $\{x^f(t,x_0) : t \geq 0\}$ has profile termalisation.

Proof. If we do not have a complex eigenvalue of $DF(0)$ then by Corollary 2.6 the statement follows. We suppose that we have a complex eigenvalue $\lambda + i\theta$ of $DF(0)$. Then $\lambda - i\theta$ is an eigenvalue of $DF(0)$. Since we suppose that $\Sigma = I_d$ and $d \in \{2,3\}$, then
\[
\left\| \Sigma^{-1/2} \sum_{k=1}^m e^{i\theta_k t} v_k \right\| = \left\| \sum_{k=1}^2 e^{i\theta_k t} v_k \right\|
\]
which is constant for any $t \geq 0$. The corollary follows from Corollary 2.5. □

In [10], we study the case when $d = 1$ which follows from Corollary 2.6.

3. The multiscale analysis

In this section, we prove that the process $\{x^\epsilon(t) : t \geq 0\}$ can be well approximated by the solution of linear non-homogeneous process in a time window that will include the time scale on which we are interested. In order to avoid technicalities, we will assume an additional set of strong conditions on $F$:

i) **Strong coercivity:** There exists $\delta > 0$ such that

$$\langle x, DF(y)x \rangle \geq \delta \|x\|^2$$

for any $x, y \in \mathbb{R}^d$.

ii) **Boundedness:** There exists a finite constant $C_0$ such that

$$\|F(y) - F(x) - DF(x)(y - x)\| \leq C_0 \|y - x\|^2$$

for any $x, y \in \mathbb{R}^d$.

iii) **Lipschitz:** There exists a finite constant $C_1$ such that

$$\|F(y) - F(x)\| \leq C_1 \|y - x\|$$

for any $x, y \in \mathbb{R}^d$.

The first condition basically says that (C1) is satisfied around any point $y$. In fact, writing

$$F(y) - F(x) = \int_0^1 \frac{d}{dt} F(x + t(y - x))dt = \int_0^1 DF(x + t(y - x))(y - x)dt$$

we obtain the seemingly stronger condition

$$\langle y - x, F(y) - F(x) \rangle \geq \delta \|y - x\|^2$$

for any $x, y \in \mathbb{R}^d$. (C2)

We call this condition **strong coercivity**, since it is basically saying that (C1) is satisfied around any point $y \in \mathbb{R}^d$. The second condition is more technical and will be used to obtain an extremely useful **a priori** bound on the convergence of $x^\epsilon(t)$ to $\mu^\epsilon$. A good example of a vector field $F$ satisfying these assumptions is $F(x) = Ax + H(x)$, where $A$ is a matrix, $H$ is vector function such that $F$ satisfies (3.1) and it satisfies $H(0) = 0$, $DH(0) = 0$, $\|DH\|_\infty < +\infty$ and $\|D^2H\|_\infty < +\infty$. Since our original field $F$ satisfies these conditions in a neighborhood of the origin, it is reasonable to expect that a localisation argument will remove this stronger set of assumption.
3.1. **Zeroth order approximation.** It is fairly easy to see that for any \( t \geq 0 \), as \( \epsilon \to 0 \), \( x^\epsilon(t) \) converges to \( \phi(t) \). The convergence can be proved to be almost surely uniform in compacts. But for our purposes, we need a quantitative estimate on the distance between \( x^\epsilon(t) \) and \( \phi(t) \). The idea is fairly simple: condition (C2) says that the dynamical system (2.1) is uniformly contracting. Therefore, it is reasonable that fluctuations are pushed back to the solution of (2.1) and therefore the difference between \( x^\epsilon(t) \) and \( \phi(t) \) has a short time dependence on the noise \( \{ B(s) : 0 \leq s \leq t \} \). This heuristics can be made precise computing the Itô derivative of \( \| x^\epsilon(t) - \phi(t) \|^2 \):

\[
d\| x^\epsilon(t) - \phi(t) \|^2 = -2 \langle x^\epsilon(t) - \phi(t), F(x^\epsilon(t)) - F(\phi(t)) \rangle dt + \sqrt{\epsilon} (x^\epsilon(t) - \phi(t)) dB(t) + d\epsilon dt
\]

\[
\leq -2\delta \| x^\epsilon(t) - \phi(t) \|^2 dt + 2\sqrt{\epsilon} (x^\epsilon(t) - \phi(t)) dB(t) + d\epsilon dt.
\]

Taking expectations together with the Gronwall trick we obtain the uniform bound

\[
\mathbb{E} \left[ \| x^\epsilon(t) - \phi(t) \|^2 \right] \leq \frac{de}{2\delta}, \quad t \geq 0.
\]

We call this bound the *zeroth order* approximation of \( x^\epsilon(t) \).

We have just proved that the distance between \( x^\epsilon(t) \) and \( \phi(t) \) is of order \( O(\sqrt{\epsilon}) \), uniformly in \( t \geq 0 \). However, this estimate is meaningful only while \( \| \phi(t) \| \gg \sqrt{\epsilon} \). By Lemma 2.1, \( \| \phi(t) \| \) is of order \( O(t^{\lambda-1}e^{-\lambda t}) \), which means that (3.4) is meaningful for times \( t \) of order \( o(t_{\text{mix}}^\epsilon) \), which fall just short of what we need. This is very natural, because at times of order \( t_{\text{mix}}^\epsilon \) we expect that fluctuations play a predominant role.

3.2. **First order approximation.** Notice that (3.4) can be seen as a law of large numbers for \( x^\epsilon(t) \). In fact, \( \mathbb{E}[x^\epsilon(t)] = \phi(t) \) for every \( t \geq 0 \) and for \( t \ll t_{\text{mix}}^\epsilon, \epsilon/\| \phi(t) \|^2 \to 0 \) and by the second-moment method, \( x^\epsilon(t) \) satisfies a law of large numbers when properly renormalised. Therefore, it is natural to look at the corresponding central limit theorem. Let us define \( \{ y^\epsilon(t) : t \geq 0 \} \) as

\[
y^\epsilon(t) = \frac{x^\epsilon(t) - \phi(t)}{\sqrt{\epsilon}} \quad \text{for any } t \geq 0.
\]

As above, it is not very difficult to prove that for every \( T > 0 \), \( \{ y^\epsilon(t) : t \in [0, T] \} \) converges in distribution to the solution \( \{ y(t) : t \in [0, T] \} \) of the linear non-homogeneous stochastic differential equation:

\[
\begin{aligned}
dy(t) &= -DF(\phi(t))y(t) dt + dB(t), \quad t \geq 0 \\
y(0) &= 0.
\end{aligned}
\]

Notice that this equation is linear and in particular \( y(t) \) has a Gaussian law for any \( t > 0 \). As in the previous section, our aim is to obtain good quantitative bounds for the distance between \( y^\epsilon(t) \) and \( y(t) \). First, we notice
that the estimate (3.4) can be rewritten as
\[ E \left[ \| y'(t) \|^2 \right] \leq \frac{d}{2\delta} \quad \text{for any } t \geq 0. \] (3.5)

We will also need an upper bound for \( E \left[ \| y'(t) \|^4 \right] \). We have that
\[
d\| y'(t) \|^4 = -4\| y'(t) \|^2 (y(t), DF(\varphi(t)) y'(t)) dt + 4\| y'(t) \|^2 (y(t), dB(t))
+ (2d + 4)\| y'(t) \|^2 dt
\leq -4\delta\| y'(t) \|^4 dt + 4\| y'(t) \|^2 (y(t), dB(t)) + (2d + 4)\| y'(t) \|^2 dt.\]

Therefore, taking expectations we get the bound
\[
\frac{d}{dt} E \left[ \| y'(t) \|^4 \right] \leq -4\delta E \left[ \| y'(t) \|^4 \right] + (2d + 4) E \left[ \| y'(t) \|^2 \right].
\]

Multiplying this inequality by \( e^{4\delta t} \) and using the Gronwall trick, we get the bound
\[
E \left[ \| y'(t) \|^4 \right] \leq \frac{d(d + 2)}{4\delta^2} (1 - e^{-4\delta t}) \leq \frac{d(d + 2)}{4\delta^2} \quad \text{for any } t \geq 0. \] (3.6)

Notice that the difference \( y'(t) - y(t) \) has finite variation, and

\[
\frac{d}{dt} (y'(t) - y(t)) = -\frac{1}{\sqrt{\epsilon}} (F(x'(t)) - F(\varphi(t)) - \sqrt{\epsilon} DF(\varphi(t)) y(t))
= -\frac{1}{\sqrt{\epsilon}} (F(x'(t)) - F(\varphi(t)) - DF(\varphi(t))(x'(t) - \varphi(t)))
- DF(\varphi(t))(y'(t) - y(t)).
\]

Therefore, using (3.2) and the chain rule for \( \| y'(t) - y(t) \|^2 \) we get the bound
\[
\frac{d}{dt} \| y'(t) - y(t) \|^2 \leq 2C_0 \sqrt{\epsilon} \| y'(t) \|^2 \| y'(t) - y(t) \|
- 2(y'(t) - y(t), DF(\varphi(t))(y'(t) - y(t))) \quad \text{(3.7)}
\leq 2C_0 \sqrt{\epsilon} \| y'(t) \|^2 \| y'(t) - y(t) \| - 2\delta \| y'(t) - y(t) \|^2.
\]

Using the Gronwall trick, the Cauchy-Schwarz inequality and the \textit{a priori} estimates (3.5) and (3.6) we get the bound
\[
E \left[ \| y'(t) - y(t) \|^2 \right] \leq C_1 \sqrt{\epsilon} \quad \text{for any } t \geq 0, \] (3.8)
with a constant \( C_1 := C_1(d, \delta) > 0 \). Using the inequality (3.8) and differential inequality (3.7) together with the Gronwall trick we get the bound
\[
E \left[ \| y'(t) - y(t) \|^2 \right] \leq C_2 e^{\frac{3}{2}t} \quad \text{for any } t \geq 0, \] (3.9)
with a constant \( C_2 := C_2(d, \delta) > 0 \). From the last trick we can notice that given any \( \theta \in (0, 1/2) \) there exists a constant \( C := C(d, \delta, \theta) > 0 \) such that
\[
E \left[ \| y'(t) - y(t) \|^2 \right] \leq C e^{1-\theta} \quad \text{for any } t \geq 0. \quad (3.10)
\]
We call this bound the first order approximation of \( x'(t) \). We have just proved that the distance between \( y'(t) \) and \( y(t) \) is of order \( O(\epsilon^{1/2-\theta/2}) \), uniformly in \( t \geq 0 \).

In Lemma 3.5. we will prove that the linear non-homogeneous process \( \{y(t) : t \geq 0\} \) has a limiting, non-degenerate law which is Gaussian with mean vector zero and covariance matrix \( \Sigma \) which is the unique solution of the Lyapunov matrix equation (2.5).

### 3.3. An \( \epsilon \)-3 proof.

We will approximate the process \( \{x'(t) : t \geq 0\} \) by a linear non-homogeneous process \( \{z'(t) := \varphi(t) + \sqrt{\epsilon}y(t) : t \geq 0\} \) in which we can make “explicit” computations. Since we will need to compare solutions of various Stochastic Differential Equations with different initial conditions, we will introduce some notation. Let \( \xi \) be a random variable in \( \mathbb{R}^d \) and let \( T > 0 \). Let \( \{\varphi(t, \xi) : t \geq 0\} \) denote the solution of

\[
\begin{align*}
\frac{d\varphi(t, \xi)}{dt} &= -F(\varphi(t, \xi))dt, \quad t \geq 0, \\
\varphi(0, \xi) &= \xi.
\end{align*}
\]

Let \( \{y(t, \xi, T) : t \geq 0\} \) be the solution of the stochastic differential equation

\[
\begin{align*}
\frac{dy(t, \xi, T)}{dt} &= -DF(\varphi(t, \xi))y(t, \xi, T)dt + dB(t + T), \quad t \geq 0, \\
y(0, \xi, T) &= 0
\end{align*}
\]

and define \( \{z'(t, \xi, T) : t \geq 0\} \) as \( z'(t, \xi, T) := \varphi(t, \xi) + \sqrt{\epsilon}y(t, \xi, T) \) for any \( t \geq 0 \).

Let \( c \in \mathbb{R} \) and \( \delta_c > 0 \) such that \( \delta_c = o(1) \). In what follows, we will always take \( T = t^{mix}_{\epsilon} + cw_c > 0 \) for every \( \epsilon > 0 \) small enough, so we will omit it from the notation. Notice that

\[
\begin{align*}
d_{TV}(x'(t^{mix}_{\epsilon} + cw_c + \delta_c, x_0), \mu^\epsilon) &= d_{TV}(x'(\delta_c, x'(t^{mix}_{\epsilon} + cw_c, x_0)), \mu^\epsilon) \\
&\leq d_{TV}(x'(\delta_c, x'(t^{mix}_{\epsilon} + cw_c, x_0))), z'(\delta_c, x'(t^{mix}_{\epsilon} + cw_c, x_0))) + \\
d_{TV}(z'(\delta_c, x'(t^{mix}_{\epsilon} + cw_c, x_0))), z'(\delta_c, x'(t^{mix}_{\epsilon} + cw_c, x_0))) + \\
d_{TV}(z'(\delta_c, z'(t^{mix}_{\epsilon} + cw_c, x_0), G(0, \epsilon \Sigma)) + d_{TV}(G(0, \epsilon \Sigma), \mu^\epsilon).
\end{align*}
\]

By reversing argument we obtain

\[
\begin{align*}
|d_{TV}(x'(t^{mix}_{\epsilon} + cw_c + \delta_c, x_0), \mu^\epsilon) - d_{TV}(z'(\delta_c, x'(t^{mix}_{\epsilon} + cw_c, x_0)), G(0, \epsilon \Sigma))| &\leq \\
ed_{TV}(x'(\delta_c, x'(t^{mix}_{\epsilon} + cw_c, x_0)), z'(\delta_c, x'(t^{mix}_{\epsilon} + cw_c, x_0))) + \\
ed_{TV}(z'(\delta_c, x'(t^{mix}_{\epsilon} + cw_c, x_0)), z'(\delta_c, z'(t^{mix}_{\epsilon} + cw_c, x_0))) + d_{TV}(G(0, \epsilon \Sigma), \mu^\epsilon).
\end{align*}
\]

In what follows, we will prove that the upper bound is negligible as \( \epsilon \to 0 \).

### 3.3.1. Short time change of measure.

**Proposition 3.1.** Let \( \delta_c > 0 \) such that \( \delta_c = o(1) \). Then for any \( c \in \mathbb{R} \)

\[
\lim_{\epsilon \to 0} d_{TV}(x'(\delta_c, x'(t^{mix}_{\epsilon} + cw_c, x_0)), z'(\delta_c, x'(t^{mix}_{\epsilon} + cw_c, x_0))) = 0.
\]
Proof. Let $\epsilon > 0$, $t \geq 0$ and $c \in \mathbb{R}$ be fixed. Let us define $\gamma^\epsilon(t) := \frac{F(x^\epsilon(t))}{\sqrt{\epsilon}}$ and $\Gamma^\epsilon(t) := \frac{F(\varphi(t)) - DF(\varphi(t))\varphi(t) + DF(\varphi(t))z^\epsilon(t)}{\sqrt{\epsilon}}$. Let $t^\epsilon(c) := t^\epsilon_{\text{mix}} + cw^\epsilon$. For any $\rho > 0$, we have

$$
\mathbb{E} \left[ \exp \left( \rho \int_{t^\epsilon(c)}^{t^\epsilon(c) + \delta^\epsilon} \| \gamma^\epsilon(s) \|^2 \, ds \right) \right] < +\infty
$$

and

$$
\mathbb{E} \left[ \exp \left( \rho \int_{t^\epsilon(c) + \delta^\epsilon}^{t^\epsilon(c) + \delta^\epsilon + \delta^\epsilon} \| \Gamma^\epsilon(s) \|^2 \, ds \right) \right] < +\infty
$$

for $\epsilon > 0$ small enough. From Cameron-Martin-Girsanov Theorem and the Novikov Theorem, it follows that

$$
\frac{d\mathbb{P}^1_{t^\epsilon(c) + \delta^\epsilon}}{d\mathbb{P}_{t^\epsilon(c) + \delta^\epsilon}} := \exp \left( \int_{t^\epsilon(c)}^{t^\epsilon(c) + \delta^\epsilon} \gamma^\epsilon(s) dW(s) - \frac{1}{2} \int_{t^\epsilon(c)}^{t^\epsilon(c) + \delta^\epsilon} \| \gamma^\epsilon(s) \|^2 \, ds \right),
$$

$$
\frac{d\mathbb{P}^2_{t^\epsilon(c) + \delta^\epsilon}}{d\mathbb{P}_{t^\epsilon(c) + \delta^\epsilon}} := \exp \left( \int_{t^\epsilon(c)}^{t^\epsilon(c) + \delta^\epsilon} \Gamma^\epsilon(s) dW(s) - \frac{1}{2} \int_{t^\epsilon(c)}^{t^\epsilon(c) + \delta^\epsilon} \| \Gamma^\epsilon(s) \|^2 \, ds \right)
$$

are well-defined Radon-Nikodym derivatives and they define true probability measures $\mathbb{P}^i_{t^\epsilon(c) + \delta^\epsilon}$, $i \in \{1, 2\}$. From now to the end of this proof we will use the notations $\mathbb{P}^i := \mathbb{P}^i_{t^\epsilon(c) + \delta^\epsilon}$, $i \in \{1, 2\}$ and $\mathbb{P} := \mathbb{P}_{t^\epsilon(c) + \delta^\epsilon}$. Under the probability measure $\mathbb{P}^1$, $B^1(t) := B(t) - \int_{t^\epsilon(c)}^{t} \gamma^\epsilon(s) \, ds$, where $t^\epsilon(c) \leq t \leq t^\epsilon(c) + \delta^\epsilon$ is a Brownian motion. Also, under the probability measure $\mathbb{P}^2$, $B^2(t) := B(t) - \int_{t^\epsilon(c)}^{t} \Gamma^\epsilon(s) \, ds$, where $t^\epsilon(c) \leq t \leq t^\epsilon(c) + \delta^\epsilon$ is a Brownian
motion. Consequently,

\[
\frac{d\mathbb{P}^1}{d\mathbb{P}^2} = \frac{\exp \left( \int_{t^*(c)}^{t^*(c)+\delta_t} \gamma^\epsilon(s) dB(s) - \frac{1}{2} \int_{t^*(c)}^{t^*(c)+\delta_t} \|\gamma^\epsilon(s)\|^2 ds \right)}{\exp \left( \int_{t^*(c)}^{t^*(c)+\delta_t} \Gamma^\epsilon(s) dB(s) - \frac{1}{2} \int_{t^*(c)}^{t^*(c)+\delta_t} \|\Gamma^\epsilon(s)\|^2 ds \right)} = \exp \left( \int_{t^*(c)}^{t^*(c)+\delta_t} (\gamma^\epsilon(s) - \Gamma^\epsilon(s)) dB(s) - \frac{1}{2} \int_{t^*(c)}^{t^*(c)+\delta_t} \left( \|\gamma^\epsilon(s)\|^2 - \|\Gamma^\epsilon(s)\|^2 \right) ds \right) = \exp \left( \int_{t^*(c)}^{t^*(c)+\delta_t} (\gamma^\epsilon(s) - \Gamma^\epsilon(s)) dB^1(s) + \frac{1}{2} \int_{t^*(c)}^{t^*(c)+\delta_t} \|\gamma^\epsilon(s) - \Gamma^\epsilon(s)\|^2 ds \right) .
\]

By the Pinsker inequality and the mean-zero martingale property of the stochastic integral, we have

\[
dTV^2(x^\epsilon(\delta_t, x^\epsilon(t^*(c) + cw_\epsilon, x_0)), z^\epsilon(\delta_t, x^\epsilon(t^*(c) + cw_\epsilon, x_0))) \leq \mathbb{E}_{\mathbb{P}^1} \left[ \int_{t^*(c)}^{t^*(c)+\delta_t} \|\gamma^\epsilon(s) - \Gamma^\epsilon(s)\|^2 ds \right] = \mathbb{E} \left[ \frac{d\mathbb{P}^1}{d\mathbb{P}} \int_{t^*(c)}^{t^*(c)+\delta_t} \|\gamma^\epsilon(s) - \Gamma^\epsilon(s)\|^2 ds \right].
\]

Define \( I^\epsilon(c) := \|t^*(c), t^*(c) + \delta_t \|. \) Using the Cauchy-Schwarz inequality and the mean-one Doléans exponential martingale property, we have

\[
\mathbb{E} \left[ \frac{d\mathbb{P}^1}{d\mathbb{P}} \int_{t^*(c)}^{\cdot} \left( \frac{d\mathbb{P}^1}{d\mathbb{P}} \int_{t^*(c)}^{\cdot} \|\gamma^\epsilon(s) - \Gamma^\epsilon(s)\|^2 ds \right)^{\frac{1}{2}} \right] \leq \left\{ \mathbb{E} \left[ \exp \left( \int_{t^*(c)}^{\cdot} \|\gamma^\epsilon(s)\|^2 ds \right) \left( \int_{t^*(c)}^{\cdot} \|\gamma^\epsilon(s) - \Gamma^\epsilon(s)\|^2 ds \right)^{\frac{1}{2}} \right] \right\} \frac{1}{2} \leq \left\{ \mathbb{E} \left[ \exp \left( 2 \int_{t^*(c)}^{\cdot} \|\gamma^\epsilon(s)\|^2 ds \right) \right] \mathbb{E} \left[ \left( \int_{t^*(c)}^{\cdot} \|\gamma^\epsilon(s) - \Gamma^\epsilon(s)\|^2 ds \right)^{\frac{1}{4}} \right] \right\} \leq \left\{ \mathbb{E} \left[ \left( \int_{t^*(c)}^{\cdot} \|\gamma^\epsilon(s) - \Gamma^\epsilon(s)\|^2 ds \right)^{\frac{1}{4}} \right] \right\} \frac{1}{4}.
\]
Then, using the Jensen inequality and the Lipschitz inequality \((3.3)\), we have

\[
\exp \left( 2 \int_{I^\epsilon(c)} \| \gamma^\epsilon(s) \|^2 \, ds \right) \leq \frac{1}{\delta \epsilon} \int_{I^\epsilon(c)} \exp \left( 2 \delta \epsilon \| \gamma^\epsilon(s) \|^2 \right) \, ds \leq \\
\frac{1}{\delta \epsilon} \int_{I^\epsilon(c)} \exp \left( \frac{2 \delta \epsilon^2 C_1^2 \| x^\epsilon(s) \|^2}{\epsilon} \right) \, ds \leq \\
\frac{1}{\delta \epsilon} \int_{I^\epsilon(c)} \exp \left( \frac{4 \delta \epsilon C_1^2 (\| x^\epsilon(s) - \varphi(s) \|^2 + \| \varphi(s) \|^2)}{\epsilon} \right) \, ds.
\]

Therefore,

\[
\mathbb{E} \left[ \exp \left( 2 \int_{I^\epsilon(c)} \| \gamma^\epsilon(s) \|^2 \, ds \right) \right] \leq \\
\frac{1}{\delta \epsilon} \int_{I^\epsilon(c)} \mathbb{E} \left[ \exp \left( \frac{4 \delta \epsilon C_1^2 (\| x^\epsilon(s) - \varphi(s) \|^2 + \| \varphi(s) \|^2)}{\epsilon} \right) \right] \, ds \leq \\
\exp \left( 4 \delta \epsilon C_1^2 d(t^\epsilon(c) + \delta \epsilon) \right) \exp \left( \frac{4 \delta \epsilon C_1^2 \| \varphi(t^\epsilon(c)) \|^2}{\epsilon} \right).
\]

From Lemma \((B.3)\), we have \( \lim_{\epsilon \to 0} \frac{\delta \epsilon \| \varphi(t^\epsilon(c)) \|^2}{\epsilon} = 0 \) for any function \( \delta \epsilon = o(1) \). Consequently,

\[
\lim_{\epsilon \to 0} \mathbb{E} \left[ \exp \left( \int_{I^\epsilon(c)} \| \gamma^\epsilon(s) \|^2 \, ds \right) \right] = 1.
\]

Now, we will prove that

\[
\lim_{\epsilon \to 0} \mathbb{E} \left[ \left( \int_{I^\epsilon(c)} \| \Gamma^\epsilon_s - \gamma^\epsilon_s \|^2 \, ds \right)^4 \right] = 0.
\]

By the Jensen inequality, we have

\[
\mathbb{E} \left[ \left( \int_{I^\epsilon(c)} \| \Gamma^\epsilon_s - \gamma^\epsilon_s \|^2 \, ds \right)^4 \right] \leq \delta \epsilon^3 \int_{I^\epsilon(c)} \mathbb{E} \left[ \| \Gamma^\epsilon_s - \gamma^\epsilon_s \|^8 \right] \, ds.
\]
Using the Boundedness inequality [3.2] and Lipschitz inequality [3.3], we get
\[ \|\Gamma^\varepsilon(s) - \gamma^\varepsilon(s)\| \leq C_0 \left\| \frac{x^\varepsilon(s) - \varphi(s)}{\sqrt{\varepsilon}} \right\|^2 + C_1 \left\| \frac{x^\varepsilon(s) - \varphi(s) - \sqrt{\varepsilon}y(s)}{\sqrt{\varepsilon}} \right\| \]
for every \( s \geq 0 \). From the last inequality, the Jensen inequality, the zeroth order approximation [3.3.3] and the first order approximation [3.10] we obtain
\[
\mathbb{E} \left[ \left( \int_{I^\varepsilon(c)} \|\Gamma_s^\varepsilon - \gamma_s^\varepsilon\|^2 \, ds \right)^4 \right] \leq \delta^3 \int_{I^\varepsilon(c)} \mathbb{E} \left[ \|\Gamma_s^\varepsilon - \gamma_s^\varepsilon\|^8 \right] \, ds \leq \tilde{C} \delta^3 \varepsilon^{4(1-\theta)},
\]
where \( \tilde{C} > 0 \) is a constant which implies that we need. □

3.3.2. Linear non-homogeneous coupling.

**Proposition 3.2.** Let \( \theta \in (0, 1/2] \) and take \( \delta_\varepsilon > 0 \) such that \( \varepsilon^{1-\theta} = o(\delta_\varepsilon) \). Then for any \( c \in \mathbb{R} \)
\[
\lim_{\varepsilon \to 0} d_{TV}(z^\varepsilon(\delta_\varepsilon, x^\varepsilon(t^\varepsilon_{\text{mix}} + cw_\varepsilon, x_0)), z^\varepsilon(\delta_\varepsilon, z^\varepsilon(t^\varepsilon_{\text{mix}} + cw_\varepsilon, x_0))) = 0.
\]

**Proof.** Due to the linear non-homogeneous approximation has Gaussian distribution, then using items ii) and item iii) of Lemma A.1, Lemma A.2 and the Cauchy-Schwarz inequality, we have
\[
d_{TV}(z^\varepsilon(\delta_\varepsilon, x^\varepsilon(t^\varepsilon_{\text{mix}} + cw_\varepsilon, x_0)), z^\varepsilon(\delta_\varepsilon, z^\varepsilon(t^\varepsilon_{\text{mix}} + cw_\varepsilon, x_0)))
\leq C_3(d, \delta) \sqrt{\frac{1}{\varepsilon \delta_\varepsilon}} \left( \mathbb{E} \left[ \|x^\varepsilon(t^\varepsilon_{\text{mix}} + cw_\varepsilon, x_0) - z^\varepsilon(t^\varepsilon_{\text{mix}} + cw_\varepsilon, x_0)\|^2 \right] \right)^{1/2}
\leq \tilde{C} \delta_\varepsilon C(d, \delta, \theta) \sqrt{\varepsilon^{1-\theta} \delta_\varepsilon},
\]
where the last inequality comes from the inequality [3.10] and the constants \( C_3(d, \delta), \tilde{C}(d, \delta, \theta) \) are positive. Therefore we get the statement. □

3.3.3. Profile function. Let us remind that \( z^\varepsilon(t) = \varphi(t) + \sqrt{\varepsilon}y(t), t \geq 0 \) where \( \{y(t) : t \geq 0\} \) satisfied the linear non-homogeneous stochastic differential equation
\[
\begin{cases}
    dy(t) = -DF(\varphi(t))y(t) \, dt + dB(t), & t \geq 0, \\
y(0) = 0.
\end{cases}
\] (3.12)
Therefore, for any \( t > 0, \ z^\varepsilon(t) \) is a Gaussian process with mean zero \( \varphi(t) \) and covariance matrix \( \varepsilon \Sigma(t) \), where \( \Sigma(t) \) can be write as the solution to the deterministic matrix differential equation:
\[
\begin{cases}
    \frac{d}{dt} \Sigma(t) = -DF(\varphi(t)) \Sigma(t) - \Sigma(t)(DF(\varphi(t)))^* + I_d, & t \geq 0, \\
\Sigma(0) = 0.
\end{cases}
\] (3.13)

Under condition [C2], we can prove that \( \varphi(t) \to 0 \) and \( \Sigma(t) \to \Sigma \) as \( t \to +\infty \), where \( \Sigma \) is a symmetric and positive definite matrix (See Lemma
Therefore, \( z^\epsilon(t) \) converges in distribution to \( z^\epsilon(\infty) \) as \( t \to +\infty \) where \( z^\epsilon(\infty) \) has Gaussian law with mean zero and covariance matrix \( \epsilon \Sigma \). The latter together with the item \( iii \) of Lemma \[A.1\] \ref{LemmaA1} Lemma \[A.3\] \ref{LemmaA3} \ref{LemmaA3} Lemma \[A.5\] \ref{LemmaA5} together with the triangle inequality improved easily the convergence to be in total variation distance. Let us measure how drastic is the convergence to the equilibrium.

\[
D^\epsilon(t) := d_{TV}(z^\epsilon(t), z^\epsilon(\infty)) = d_{TV}(G(\varphi(t), \epsilon \Sigma(t)), G(0, \epsilon \Sigma)).
\]

**Proposition 3.3.** Let \( \delta_\epsilon > 0 \) such that \( \delta_\epsilon = o(1) \) For any \( c \in \mathbb{R} \) we have

\[
\lim_{c \to 0} \left| D^\epsilon(t_{\text{mix}}^\epsilon + \delta_\epsilon + cw^\epsilon) - D^\epsilon(t_{\text{mix}}^\epsilon + \delta_\epsilon + cw^\epsilon) \right| = 0,
\]

where

\[
D(t) := d_{TV} \left( G \left( \frac{(t - \tau)^{l-1}}{e^{\lambda(t-\tau)}\sqrt{\epsilon}} \sum_{k=1}^{m} e^{i\theta_k(t-\tau)} v_k, I_d \right), G(0, I_d) \right)
\]

for any \( t \geq \tau \), with \( \lambda, \ell, \tau, \theta_1, \ldots, \theta_m \in [0, 2\pi) \), \( v_1, \ldots, v_m \) are the constants and vectors associated to \( x_0 \) in Lemma \[2.1\] \ref{Lemma2.1} and the matrix \( \Sigma \) is the unique solution of the matrix Lyapunov equation

\[
DF(0) \Sigma + \Sigma (DF(0))^* = I_d.
\]

**Proof.** For any \( t > 0 \), by the triangle inequality and item \( ii \), item \( iii \) of Lemma \[A.1\] \ref{LemmaA1} we get

\[
D^\epsilon(t) \leq d_{TV}(G(\varphi(t), \epsilon \Sigma(t)), G(\varphi(t), \epsilon \Sigma)) + d_{TV}(G(\varphi(t), \epsilon \Sigma), G(0, \epsilon \Sigma)) \leq d_{TV}(G(0, \Sigma(t)), G(0, \Sigma)) + d_{TV} \left( G \left( \frac{1}{\sqrt{\epsilon}} \varphi(t), \Sigma \right), G(0, \Sigma) \right).
\]

By the same facts we have

\[
\left| D^\epsilon(t) - d_{TV} \left( G \left( \frac{1}{\sqrt{\epsilon}} \varphi(t), \Sigma \right), G(0, \Sigma) \right) \right| \leq d_{TV}(G(0, \Sigma(t)), G(0, \Sigma))
\]

for any \( t > 0 \). Using Lemma \[A.5\] \ref{LemmaA5} we see that

\[
\lim_{t \to +\infty} d_{TV}(G(0, \Sigma(t)), G(0, \Sigma)) = 0.
\]

Therefore, the cut-off phenomena can be study from the distance

\[
\tilde{D}^\epsilon(t) := d_{TV} \left( G \left( \frac{1}{\sqrt{\epsilon}} \varphi(t), \Sigma \right), G(0, \Sigma) \right) = d_{TV} \left( G \left( \Sigma^{-1/2} \frac{1}{\sqrt{\epsilon}} \varphi(t), I_d \right), G(0, I_d) \right)
\]

for any \( t > 0 \), where the last equality follow from item \( iii \) of Lemma \[A.1\] \ref{LemmaA1}.

Using the constants associated to \( x_0 \) given by Lemma \[2.1\] \ref{Lemma2.1} for any \( t > \tau \) define

\[
\bar{D}(t) := d_{TV} \left( G \left( \Sigma^{-1/2} \frac{(t - \tau)^{l-1}}{e^{\lambda(t-\tau)}\sqrt{\epsilon}} \sum_{k=1}^{m} e^{i\theta_k(t-\tau)} v_k, I_d \right), G(0, I_d) \right)
\]
and
\[
R^\epsilon(t) := d_{TV} \left( \mathcal{G} \left( \Sigma^{-1/2} \frac{1}{\sqrt{\epsilon}} \varphi(t), I_d \right), \mathcal{G} \left( \Sigma^{-1/2} \frac{(t-\tau)^{1-1}}{\epsilon} \sqrt{\epsilon} \sum_{k=1}^{m} e^{i\theta_k(t-\tau)} v_k, I_d \right) \right).
\]

An straightforward calculation shows
\[
\left| \tilde{D}^\epsilon(t) - D^\epsilon(t) \right| \leq R^\epsilon(t)
\]
for any \( t > \tau \). Therefore, by Lemma 2.1
\[
\lim_{\epsilon \to 0} R^\epsilon(t^\epsilon_{\text{mix}} + \delta \epsilon + cw^\epsilon) = 0
\]
for any \( c \in \mathbb{R} \). □

### 3.3.4. The invariant measure

In this section, we will prove that the invariant measure of the evolution (2.2) is well approximated in total variation distance by a Gaussian distribution as the following proposition states.

**Proposition 3.4.**
\[
\lim_{\epsilon \to 0} d_{TV}(\mathcal{G}(0, \epsilon \Sigma), \mu^\epsilon) = 0.
\]

**Proof.** Let us recall that \( y^\epsilon(t) = \varphi(t) + \sqrt{\epsilon}y(t) \). Note that for any \( s, t \geq 0 \) we have
\[
d_{TV}(\mathcal{G}(0, \epsilon \Sigma), \mu^\epsilon) \leq d_{TV}(\mathcal{G}(0, \epsilon \Sigma), y^\epsilon(s + t)) + d_{TV}(y^\epsilon(s + t, x_0), x^\epsilon(s + t, x_0)) + d_{TV}(x^\epsilon(s + t, x_0), \mu^\epsilon)
\]
Then
\[
d_{TV}(\mathcal{G}(0, \epsilon \Sigma), y^\epsilon(s + t)) = d_{TV}(\mathcal{G}(0, \epsilon \Sigma), \mathcal{G}(\varphi(s + t), \epsilon \Sigma(s + t))),
\]
where \( \Sigma(t) \) is the covariance matrix of \( y(t) \). Therefore using the invariant translation and the scaling invariance of the total variation distance, we get
\[
d_{TV}(\mathcal{G}(0, \epsilon \Sigma), y^\epsilon(s + t)) \leq d_{TV}(\mathcal{G}(0, \Sigma), \mathcal{G}(0, \Sigma(s + t)))+ d_{TV} \left( \mathcal{G} \left( \frac{\varphi(s + t)}{\sqrt{\epsilon}}, \Sigma \right), \mathcal{G}(0, \Sigma) \right).
\]
Let \( s^\epsilon \ll \epsilon^{1/2017} \) and \( t^\epsilon_{\text{mix}} \ll t^\epsilon \). By Lemma [A.5] and Lemma [B.2] we obtain
\[
\lim_{\epsilon \to 0} d_{TV}(\mathcal{G}(0, \Sigma), \mathcal{G}(0, \Sigma(s^\epsilon + t^\epsilon))) = 0.
\]
Now, using the same ideas as in Proposition 3.1 (much easier due to \( t^\epsilon \gg t^\epsilon_{\text{mix}} \)) we get
\[
\lim_{\epsilon \to 0} d_{TV}(y^\epsilon(s^\epsilon + t^\epsilon, x_0), x^\epsilon(s^\epsilon + t^\epsilon, x_0)) = 0.
\]
It remains to prove that
\[
\lim_{\epsilon \to 0} d_{TV}(x^\epsilon(s^\epsilon + t^\epsilon, x_0), \mu^\epsilon) = 0.
\]
Notice
\[ d_{TV}(x'(s + t, x_0), \mu^\epsilon) \leq \int_{\mathbb{R}^d} d_{TV}(x'(s + t, x_0), x'(s + t, x)) \mu^\epsilon(dx). \]

Due to the stochastic differential equation associated to \( \{y^\epsilon(t) : t \geq 0\} \) is not homogeneous we need to improve the notation as we did just before Proposition 3.1. Nevertheless, we can omit as we did in Proposition 3.1.

Then
\[ \int_{\mathbb{R}^d} d_{TV}(x'(s + t, x_0), x'(s + t, x)) \mu^\epsilon(dx) \leq d_{TV}(x'(s, x'(t, x_0)), y'(s, x'(t, x_0))) + d_{TV}(y'(s, x'(t, x_0)), y'(s, y'(t, x_0))) + \int_{\mathbb{R}^d} d_{TV}(y'(s, y'(t, x_0)), y'(s, y'(t, x))) \mu^\epsilon(dx) + \int_{\mathbb{R}^d} d_{TV}(y'(s, y'(t, x)), x'(s, x'(t, x))) \mu^\epsilon(dx). \]

Again, using the same ideas as in Proposition 3.1 and Proposition 3.2 (much easier due to \( t^\epsilon \gg t_{\text{mix}}^\epsilon \)) we get
\[ \lim_{\epsilon \to 0} d_{TV}(x'(s, x'(t, x_0)), y'(s, x'(t, x_0))) = 0. \]

and
\[ \lim_{\epsilon \to 0} d_{TV}(y'(s, x'(t, x_0)), y'(s, y'(t, x_0))) = 0. \]

Fix \( R > 0 \). We split the remainders integrals as follows
\[ \int_{\mathbb{R}^d} d_{TV}(y'(s, y'(t, x_0)), y'(s, y'(t, x))) \mu^\epsilon(dx) \leq \int_{\|x\| \leq R} d_{TV}(y'(s, y'(t, x_0)), y'(s, y'(t, x))) \mu^\epsilon(dx) + \mu^\epsilon(\|x\| > R) \]

and
\[ \int_{\mathbb{R}^d} d_{TV}(y'(s, y'(t, x)), x'(s, x'(t, x))) \mu^\epsilon(dx) \leq \int_{\|x\| \leq R} d_{TV}(y'(s, y'(t, x)), x'(s, x'(t, x))) \mu^\epsilon(dx) + \mu^\epsilon(\|x\| > R). \]

Notice
\[ \int_{\|x\| \leq R} d_{TV}(y'(s, y'(t, x_0)), y'(s, y'(t, x))) \mu^\epsilon(dx) \leq \kappa(R) \frac{1}{\sqrt{\epsilon}} e^{-\delta(t+s)}, \]

where \( \kappa(R) \) is a constant depending on \( R \).
where $\kappa(R)$ is a non-negative constant and $\delta > 0$ coming from the strong coercivity (3.2). Taking $t^\epsilon \gg t^\epsilon_{\text{mix}}$ such that $e^{-\delta t^\epsilon} = o(\sqrt{\epsilon})$. Therefore

$$\lim_{\epsilon \to 0} \int_{\|x\| \leq R} d_{TV}(y^\epsilon(s', y^\epsilon(t^\epsilon, x_0)), y^\epsilon(s^\epsilon, y^\epsilon(t^\epsilon, x)))\mu^\epsilon(dx) = 0.$$ 

Following the proof of Proposition 3.1, we have

$$\int_{\|x\| \leq R} d_{TV}(y^\epsilon(s^\epsilon, y^\epsilon(t^\epsilon, x)), x^\epsilon(s^\epsilon, x^\epsilon(t^\epsilon, x)))\mu^\epsilon(dx) \leq \tilde{\kappa}(R)o(1)$$

where $\tilde{\kappa}(R)$ is a non-negative constant. Now, we only need to prove that $\mu^\epsilon(\|x\| > R)$ is negligible when $\epsilon$ vanish. Following the ideas from [8] (page 122, Section 5, Step 1), the invariant measure $\mu^\epsilon$ has finite p-moments for any $p \geq 0$. Moreover, we have $\int \|x\|^2\mu^\epsilon(dx) \leq \frac{\epsilon^2}{\delta}$ that together with the Chebyshev inequality is enough to conclude that the $\mu^\epsilon(\|x\| > R) = o(1)$.

3.3.5. Proof of Theorem 2.2. Now, we are ready to prove Theorem 2.2. To stress the fact that Theorem 2.2 is just a consequence of what we have proved up to here, let us state this as a Lemma.

**Lemma 3.5.** Assume that the vector field $F$ of (2.2) satisfies (3.1), (3.2) and (3.3). Let $\{x^\epsilon(t, x_0) : t \geq 0\}$ be the solution of (2.2) and denote by $\mu^\epsilon$ the unique invariant probability measure for the evolution given by (2.2). Let $d^\epsilon(t) = d_{TV}(x^\epsilon(t, x_0), \mu^\epsilon)$, $t \geq 0$ and suppose that $x_0 \neq 0$. Let us consider the mixing time $t^\epsilon_{\text{mix}}$ which is given by (2.4) and the time window which given by (2.3). Let $\vartheta \in (0, 1/5)$ and define $\delta_\epsilon = \epsilon^2$. For any $c \in \mathbb{R}$ we have

$$\lim_{\epsilon \to 0} \left| d^\epsilon(t^\epsilon_{\text{mix}} + \delta_\epsilon + cw^\epsilon) - \tilde{d}^\epsilon(t^\epsilon_{\text{mix}} + \delta_\epsilon + cw^\epsilon) \right| = 0,$$

where

$$\tilde{d}^\epsilon(t) = d_{TV}\left(\mathcal{G}\left(\frac{(t - \tau)^{\ell-1}}{\epsilon^{\ell}} \sum_{k=1}^{m} e^{i\theta_k(t-\tau)}v_k, I_d\right), \mathcal{G}(0, I_d)\right), \quad t \geq \tau$$

with $\lambda, \ell, \tau, \theta_1, \ldots, \theta_m \in [0, 2\pi)$, $v_1, \ldots, v_m$ are the constants and vectors associated to $x_0$ in Lemma 2.1 and the matrix $\Sigma$ is the unique solution of the matrix Lyapunov equation $DF(0)X + X(DF(0))^* = I_d$.

**Proof.** Firstly, from Lemma C.2 we have that exist a unique invariant probability measure for the evolution (2.2). Let call the invariant measure by $\mu^\epsilon$. We follow the notation as in Proposition 3.3. From Proposition 3.1, Proposition 3.2 and Proposition 3.4 we get that the upper bound of inequality (3.11) is negligible, i.e.,

$$\left| d^\epsilon(t^\epsilon_{\text{mix}} + \delta_\epsilon + cw^\epsilon) - \tilde{d}^\epsilon(t^\epsilon_{\text{mix}} + \delta_\epsilon + cw^\epsilon) \right| = o(1).$$
By the triangle inequality, we have
\[ \left| d^\epsilon(t^\epsilon_{\text{mix}} + \delta^\epsilon + cw^\epsilon) - D^\epsilon(t^\epsilon_{\text{mix}} + \delta^\epsilon + cw^\epsilon) \right| \leq \left| D^\epsilon(t^\epsilon_{\text{mix}} + \delta^\epsilon + cw^\epsilon) - D^\epsilon(t^\epsilon_{\text{mix}} + \delta^\epsilon + cw^\epsilon) \right| + o(1). \]

From the last inequality and Proposition 3.3 we get the result. \( \square \)

**Appendix A. Properties of the Total Variation Distance for Gaussian Distributions**

Recall that \( \mathcal{G}(v, \Xi) \) denote the Gaussian distribution in \( \mathbb{R}^d \) with vector mean \( v \) and positive definite covariance matrix \( \Xi \).

**Lemma A.1.** Let \( v, \tilde{v} \in \mathbb{R}^d \) be two fixed vectors and \( \Xi, \tilde{\Xi} \) be two fixed matrices \( d \times d \) symmetric positive definite matrices. It follows
\[ i) \quad d_{TV} \left( \mathcal{G}(cv, c^2 \Xi), \mathcal{G} \left( c\tilde{v}, c^2 \tilde{\Xi} \right) \right) = d_{TV} \left( \mathcal{G}(v, \Xi), \mathcal{G} \left( \tilde{v}, \tilde{\Xi} \right) \right). \]
\[ ii) \quad d_{TV} \left( \mathcal{G}(v, \Xi), \mathcal{G} \left( \tilde{v}, \tilde{\Xi} \right) \right) = d_{TV} \left( \mathcal{G}(v - \tilde{v}, \Xi), \mathcal{G}(0, \tilde{\Xi}) \right). \]
\[ iii) \quad d_{TV} \left( \mathcal{G}(v, \Xi), \mathcal{G} \left( \tilde{v}, \Xi \right) \right) = d_{TV} \left( \mathcal{G} \left( \Xi^{-\frac{1}{2}} v, I_d \right), \mathcal{G} \left( \Xi^{-\frac{1}{2}} \tilde{v}, I_d \right) \right). \]
\[ iv) \quad d_{TV} \left( \mathcal{G}(0, \Xi), \mathcal{G}(0, \tilde{\Xi}) \right) = d_{TV} \left( \mathcal{G} \left( 0, \Xi^{-\frac{1}{2}} \Xi^{-\frac{1}{2}} \right), \mathcal{G}(0, I_d) \right). \]

**Proof.** The proof is just an straightforward calculations. \( \square \)

**Lemma A.2.** Let \( v = (v_1, \ldots, v_d)^* \in \mathbb{R}^d \) then
\[ d_{TV} \left( \mathcal{G}(v, I_d), \mathcal{G}(0, I_d) \right) \leq \sqrt{\frac{\sum_{i=1}^{d} |v_i|}{2\pi}}. \]

**Proof.** This is done using the classical coupling argument. \( \square \)

**Lemma A.3.** Let \( \{v_\epsilon\}_{\epsilon>0} \subset \mathbb{R}^d \) such that \( \lim_{\epsilon \to 0} v_\epsilon = v \in \mathbb{R}^d \). Then,
\[ \lim_{\epsilon \to 0} d_{TV} \left( \mathcal{G}(v_\epsilon, I_d), \mathcal{G}(0, I_d) \right) = d_{TV} \left( \mathcal{G}(v, I_d), \mathcal{G}(0, I_d) \right) \]

**Proof.** This is done using the triangle inequality, the item ii) of Lemma A.1 and Lemma A.2. \( \square \)

**Lemma A.4.** Let \( \{v_\epsilon\}_{\epsilon>0} \subset \mathbb{R}^d \) such that \( \lim_{\epsilon \to 0} \|v_\epsilon\| = +\infty \). Then,
\[ \lim_{\epsilon \to 0} d_{TV} \left( \mathcal{G}(v_\epsilon, I_d), \mathcal{G}(0, I_d) \right) = 1. \]
Proof. By definition
\[
d_{TV}(\mathcal{G}(v, I_d), \mathcal{G}(0, I_d)) = \frac{1}{2(2\pi)^{d/2}} \int_{\mathbb{R}^d} |f(x - v) - f(x)| \, dx,
\]
where \( f(x) = \exp(-\|x\|^2/2) \). Using a classical analysis trick we know
\[
\lim_{\epsilon \to 0} \int_{\mathbb{R}^d} |f(x - v_\epsilon) - f(x)| \, dx = \int_{\mathbb{R}^d} |f(x)| \, dx,
\]
when \( \int_{\mathbb{R}^d} |f(x)| \, dx < +\infty \). The last statement implies the result. Now, we will prove the relation (A.1). Let us define \( M := \int_{\mathbb{R}^d} |f(x)| \, dx < +\infty \). Let \( \eta > 0 \) be fixed. Let \( B(x, r) \) denotes the Euclidean open ball centered in \( x \in \mathbb{R}^d \) with radius \( r > 0 \). Then, there exists \( r = r(\eta) > 0 \) large enough such that
\[
M - \int_{B(0, r)} |f(x)| \, dx < \frac{\eta}{4}.
\]
Therefore,
\[
M - \int_{B(v_\epsilon, r)} |f(x - v_\epsilon)| \, dx < \frac{\eta}{4}.
\]
Due to \( \lim_{\epsilon \to 0} \|v_\epsilon\| = +\infty \), then there exists \( \epsilon_0 := \epsilon_0(r) > 0 \) such that for every \( 0 < \epsilon < \epsilon_0 \), we have \( B(0, r) \cap B(v_\epsilon, r) = \emptyset \). Moreover, we can take \( \epsilon_0 \) such that \( \int_{B(v_\epsilon, r)} |f(x)| \, dx \leq \frac{\eta}{4} \) for every \( 0 < \epsilon < \epsilon_0 \). Consequently,
\[
\int_{\mathbb{R}^d} |f(x - v_\epsilon) - f(x)| \, dx \geq \int_{B(0, r)} |f(x - v_\epsilon) - f(x)| \, dx + \int_{B(v_\epsilon, r)} |f(x - v_\epsilon) - f(x)| \, dx \geq \int_{B(0, r)} (|f(x)| - |f(x - v_\epsilon)|) \, dx + \int_{B(v_\epsilon, r)} |f(x - v_\epsilon)| - |f(x)| \, dx \geq 2M - \eta.
\]
Consequently, for every \( \eta > 0 \), we have
\[
2M - \eta \leq \int_{\mathbb{R}^d} |f(x - v_\epsilon) - f(x)| \, dx \leq 2M.
\]
Now, taking \( \epsilon \to 0 \) and then \( \eta \to 0 \) we obtain the statement. □
Lemma A.5. Let $S_d$ denote the set of $d \times d$ symmetric and positive definite matrices. Let $\{\Xi_\epsilon\}_{\epsilon > 0} \subset S_d$ such that $\lim_{\epsilon \to 0} \Xi_\epsilon = \Xi \in S_d$. Then
\[
\lim_{\epsilon \to 0} d_{TV}(G(0, \Xi_\epsilon), G(0, \Xi)) = 0.
\]

Proof. By item iv) of Lemma A.1, for every $\epsilon > 0$, we have
\[
d_{TV}(G(0, \Xi_\epsilon), G(0, \Xi)) = d_{TV}(G(0, \Xi_\epsilon^{-\frac{1}{2}}\Xi_\epsilon^{-\frac{1}{2}}) - G(0, I_d)).
\]
Consequently, it suffices to prove, when $\lim_{\epsilon \to 0} \Xi_\epsilon = I_d \in S_d$. By definition, we have
\[
d_{TV}(G(0, \Xi_\epsilon), G(0, I_d)) = \frac{1}{2(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \exp \left( -\frac{x^* \Xi_\epsilon^{-\frac{1}{2}} x}{2} \right) \left( \frac{\det(\Xi_\epsilon)}{\det(\Xi_\epsilon^{-\frac{1}{2}})} \right)^{\frac{1}{2}} - \exp \left( -\frac{x^* x}{2} \right) \right| dx.
\]
Define the function
\[
f_\epsilon(x) = \exp \left( -\frac{x^* \Xi_\epsilon^{-\frac{1}{2}} x}{2} \right) \left( \frac{\det(\Xi_\epsilon)}{\det(\Xi_\epsilon^{-\frac{1}{2}})} \right)^{\frac{1}{2}} - \exp \left( -\frac{x^* x}{2} \right), \quad x \in \mathbb{R}^d.
\]
For every $x \in \mathbb{R}^d$, we have $\lim_{\epsilon \to 0} f_\epsilon(x) = 0$. Also, for $\epsilon > 0$ small enough, it follows that
\[
f_\epsilon(x) \leq K_1 \exp \left( -K_2 \|x\|^2 \right)
\]
for any $x \in \mathbb{R}^d$, where $K_1 > 0$ and $K_2 > 0$ are constants that does not depend on $\epsilon$. Consequently, the result follows from the Dominated Convergence Theorem.

Appendix B. The Deterministic Dynamical System

In this section we present a proof of Lemma 2.1. We start analyzing the linear differential equation associated to the linearisation of the non-linear deterministic differential equation (2.1) around the hyperbolic fixed point 0.

Lemma B.1. Let us suppose that the vector field $F$ of (2.1) satisfies the coercivity condition (C1). For any $x_0 \in \mathbb{R}^d \setminus \{0\}$ there exist $\lambda := \lambda(x_0) > 0$, $\ell := \ell(x_0), m := m(x_0) \in \mathbb{N}, \theta_1 := \theta_1(x_0), \ldots, \theta_m := \theta_m(x_0) \in [0, 2\pi)$ and $v_1 := v_1(x_0), \ldots, v_m := v_m(x_0)$ in $\mathbb{C}^d$ linearly independent such that
\[
\lim_{t \to +\infty} \left\| \frac{e^{\lambda t}}{t^\ell} e^{-DF(0)t} x_0 - \sum_{k=1}^m e^{i\theta_k t} v_k \right\| = 0.
\]

Proof. Let us write $\Lambda = DF(0)$. Notice that by coercivity condition (C1), all eigenvalues of $\Lambda$ have positive real part. Let us denote by $\{\phi(t, x) : t \geq 0\}$ the solution of the linear system
\[
\left\{ \begin{array}{l}
\frac{d}{dt} \phi(t) = -\Lambda \phi(t) \\
\phi(0) = x.
\end{array} \right.
\]
Let \((w_{j,k} : j = 1, \ldots, N; k = 1, \ldots, N_j)\) be a Jordan basis of \(-\Lambda\), that is,
\[-\Lambda w_{j,k} = -\lambda_j w_{j,k} + w_{j,k+1}.
\]
In this formula we use the convention \(w_{j,N_j+1} = 0\). Since \((w_{j,k})_{j,k}\) is a basis of \(\mathbb{R}^d\), the decomposition
\[\phi(t, x) = \sum_{j,k} \phi_{j,k}(t, x)w_{j,k}\]
defines the functions \(\phi_{j,k}(t, x)\) in a unique way. We have that
\[\sum_{j,k} \frac{d}{dt} \phi_{j,k}(t, x)w_{j,k} = \sum_{j,k} \phi_{j,k}(t, x)(-\lambda_j w_{j,k} + w_{j,k+1}),\]
and the aforementioned uniqueness implies that
\[\frac{d}{dt} \phi_{j,k}(t, x) = -\lambda_j \phi_{j,k}(x, t) + \phi_{j,k-1}(t, x),\] (B.1)
where we use the convention \(\phi_{j,0}(t, x) = 0\). In addition, we have that \(\phi_{j,k}(0, x) = x_{j,k}\), where
\[x = \sum_{j,k} x_{j,k}w_{j,k}.
\]
For each \(j\), the system of equations for \(\{\phi_{j,k}(t, x) : k = 1, \ldots, N_j\}\) is autonomous, as well as the equation for \(\phi_{j,1}(t, x)\). We have that
\[\phi_{j,1}(t, x) = x_{j,1}e^{-\lambda_j t}\]
and by the method of variation of parameters, for \(k = 2, \ldots, N_j\) we have that
\[\phi_{j,k}(t, x) = x_{j,k}e^{-\lambda_j t} + \int_0^t e^{-\lambda_j(t-s)} \phi_{j,k-1}(s, x)ds.
\]
Applying this formula for \(k = 2\) we see that
\[\phi_{j,2}(t, x) = x_{j,2}e^{-\lambda_j t} + x_{j,1}te^{-\lambda_j t}\]
and from this expression we can guess and check the formula
\[\phi_{j,k}(t, x) = \sum_{i=1}^k x_{j,i} \frac{t^{k-i}e^{-\lambda_j t}}{(k-i)!}.\]
We conclude that
\[\phi(t, x) = \sum_{j=1}^N \sum_{k=1}^{N_j} \sum_{i=1}^k x_{j,k}^{0} \frac{t^{k-i}e^{-\lambda_j t}}{(k-i)!} x_{j,i}w_{j,k}.\] (B.2)

With this expression in hand, we are ready to prove Lemma [B.1]. Let \(x_0 \in \mathbb{R}^d\) be fixed. Let us assume that \(x_0 \neq 0\) and write
\[x_0 = \sum_{j,k} x_{j,k}^{0}w_{j,k}.
\]
We take
\[\lambda = \min\{\Re(\lambda_j) : x_{j,k}^{0} \neq 0 \text{ for some } k\}\]
and we define
\[ J_0 = \{ j : \text{Re}(\lambda_j) = \lambda \text{ and } x_{j,k}^0 \neq 0 \text{ for some } k \}. \]
In other words, we identify in (B.2) the smallest exponential rate of decay and we collect in \( J_0 \) all the indices with that exponential decay. Now we define
\[ \ell = \max\{N_j - k : j \in J_0 \text{ and } x_{j,k}^0 \neq 0 \} \]
and
\[ J = \{ j \in J_0 : x_{j,N_j-\ell}^0 \neq 0 \}. \]
We see that for \( j \in J \),
\[ \lim_{t \to \infty} \left| \phi_{j,N_j}(t,x_0) \right| \frac{e^{\lambda t}}{t!} = \frac{|x_{j,N_j-\ell}|}{\ell!}, \]
while for \( j \notin J \) and \( k \) arbitrary or \( j \in J \) and \( k \neq N_j \),
\[ \lim_{t \to \infty} \left| \phi_{j,k}(t,x_0) \right| \frac{e^{\lambda t}}{t!} = 0. \]
Therefore,
\[ \lim_{t \to \infty} \left\| \frac{e^{\lambda t}}{t!} \phi(t,x_0) - \sum_{j \in J} e^{-(\lambda_j - \lambda)t} \frac{1}{\ell!} x_{j,N_j-\ell} w_{j,N_j} \right\| = 0. \]
Let \( m = \# J \) and let \( \sigma : \{1, \ldots, m\} \to J \) be a numbering of \( J \). By definition of \( \lambda \) and \( J \), the numbers \( \lambda_j - \lambda \) are imaginary. Therefore, Lemma B.1 is proved choosing \( \theta_k = i(\lambda_{\sigma_k} - \lambda) \) and \( v_k = \frac{x_{\sigma_k,N_{\sigma_k}}}{\ell!} \).

Now, we are ready to prove Lemma 2.1. The proof is in the Hartman-Grobman Theorem (see Theorem(Hartman) page 127 of [16] or the celebrate paper of P. Hartman [23]) that guarantees that the conjugation around the hyperbolic fixed point zero and \( \theta \) of \( (2.1) \) is \( C^1 \)-local diffeomorphism under some resonance conditions which are fulfilled when all the eigenvalues of the matrix \( DF(0) \) have negative (or positive) real part.

Lemma B.2. Let us suppose that the vector field \( F \) of (2.1) satisfies the coercivity condition (C.1). For any \( x_0 \in \mathbb{R}^d \setminus \{0\} \) there exist \( \lambda := \lambda(x_0) > 0 \), \( \ell := \ell(x_0), m := m(x_0) \in \mathbb{N}, \theta_1 := \theta_1(x_0), \ldots, \theta_m := \theta_m(x_0) \in [0, 2\pi), v_1 := v_1(x_0), \ldots, v_m := v_m(x_0) \) in \( \mathbb{C}^d \) linearly independent and \( \tau := \tau(x_0) > 0 \) such that
\[ \lim_{t \to +\infty} \left\| \frac{e^{\lambda t}}{t^{\ell-1}} \varphi(t+\tau,x_0) - \sum_{k=1}^m e^{i\theta_k t} v_k \right\| = 0. \]

Proof. Due to all the eigenvalues of \( DF(0) \) has real positive real part, there exist open sets \( U, V \) around the hyperbolic fixed point zero and \( h : U \to V \) a \( C^1(U, V) \) homeomorphism such that \( h(0) = 0 \) and \( h(x) = x + o(||x||) \) as \( ||x|| \to 0 \) such that \( \varphi(t,x) = h^{-1}(e^{-DF(0)t}h(x)) \) for any \( t \geq 0 \) and \( x \in U \). Because of
\[ ||\varphi(t,x)|| \leq ||x||e^{-\delta t} \text{ for any } x \in \mathbb{R}^d \text{ and any } t \geq 0. \]
There exists $\tau := \tau(x_0) > 0$ such that $\varphi(t, x_0) \in U$ for any $t \geq \tau$. Then $\varphi(t + \tau, x_0) = \varphi(t, x_\tau) = h^{-1}(e^{-DF(t)h(x_\tau)})$ for any $t \geq 0$. Let $\tilde{x} := h(x_\tau)$. By Lemma B.1 there exist $\lambda(\tilde{x}) := \lambda > 0$, $\ell(\tilde{x}) := \ell, m(\tilde{x}) := m \in \mathbb{N}$, $\theta_1(\tilde{x}) := \theta_1, \ldots, \theta_m(\tilde{x}) := \theta_m \in [0, 2\pi)$ and $v_1(\tilde{x}) := v_1, \ldots, v_m(\tilde{x}) := v_m$ in $\mathbb{C}^d$ linearly independent such that

$$
\lim_{t \to +\infty} \left\| \frac{e^{\lambda t}}{t^{\ell-1}} e^{-DF(t)\tilde{x}} - \sum_{k=1}^m e^{i\theta_k t} v_k \right\| = 0.
$$

Then

$$
\left\| \frac{e^{\lambda t}}{t^{\ell-1}} \varphi(t + \tau, x_0) - \sum_{k=1}^m e^{i\theta_k t} v_k \right\| \leq \left\| \frac{e^{\lambda t}}{t^{\ell-1}} \varphi(t + \tau, x_0) - \frac{e^{\lambda t}}{t^{\ell-1}} e^{-DF(t)\tilde{x}} \right\| + \left\| \frac{e^{\lambda t}}{t^{\ell-1}} e^{-DF(t)\tilde{x}} - \sum_{k=1}^m e^{i\theta_k t} v_k \right\|
$$

Let us observe that

$$
\left\| \frac{e^{\lambda t}}{t^{\ell-1}} e^{-DF(t)\tilde{x}} \right\| o(1) \leq \left\| \frac{e^{\lambda t}}{t^{\ell-1}} e^{-DF(t)\tilde{x}} - \sum_{k=1}^m e^{i\theta_k t} v_k \right\| o(1) + \left( \sum_{k=1}^m \|v_k\| \right) o(1).
$$

Therefore

$$
\lim_{t \to +\infty} \left\| \frac{e^{\lambda t}}{t^{\ell-1}} \varphi(t + \tau, x_0) - \sum_{k=1}^m e^{i\theta_k t} v_k \right\| = 0.
$$

\[ \square \]

**Lemma B.3.** Let $\delta_\epsilon = o(1)$. Then

$$
\lim_{\epsilon \to 0} \frac{\delta_\epsilon \| \varphi(t_{\text{mix}}^\epsilon + \delta_\epsilon, x_0) \|^2}{\epsilon} = 0.
$$

**Proof.** Let us remember that

$$
t_{\text{mix}}^\epsilon = \frac{1}{2\lambda} \ln (1/\epsilon) + \frac{\ell - 1}{\lambda} \ln (\ln (1/\epsilon)) + \tau,
$$

where $\lambda$, $\ell$ and $\tau$ are the constants associated to $x_0$ in Lemma B.1. Define $t^\epsilon := t_{\text{mix}}^\epsilon - \tau + \delta_\epsilon$.

Note that

$$
\frac{1}{\sqrt{\epsilon}} \| \varphi(t^\epsilon + \tau, x_0) \| \leq \frac{(t^\epsilon)^{\ell-1}}{e^{\lambda t^\epsilon} \sqrt{\epsilon}} \left\| \frac{e^{\lambda t^\epsilon}}{(t^\epsilon)^{\ell-1}} \varphi(t^\epsilon + \tau, x_0) - \sum_{k=1}^m e^{i\theta_k t^\epsilon} v_k \right\| + \frac{(t^\epsilon)^{\ell-1}}{e^{\lambda t^\epsilon} \sqrt{\epsilon}} \sum_{k=1}^m \|v_k\|.
$$
From the last inequality, using the fact that \( \lim_{\epsilon \to 0} \frac{\epsilon^{(n)} - 1}{\epsilon^{n}} \sqrt{\epsilon} = e^{\lambda \epsilon} \) and Lemma B.2 we get the result.

\[ \square \]

**APPENDIX C. THE STOCHASTIC DYNAMICAL SYSTEM**

The following proposition will give us the zeroth order and first order approximation for the Itô diffusion \( \{x^\epsilon(t) : t \geq 0\} \).

**Proposition C.1** (Zeroth order and first order approximation). Let us write

\[ W(t) := \sup_{0 \leq s \leq t} \|B(s)\| \text{ for } t \geq 0. \]

i) For every \( t \geq 0 \) we have

\[ E \left[ \left( \frac{\|x^\epsilon(t) - \varphi(t)\|}{\epsilon} \right)^{2n} \right] \leq \frac{c_n}{\epsilon^{2n}} \leq \frac{c_n}{\epsilon^{2n}}, \]

where \( c_n := \prod_{j=0}^{n-1} (d + 2j) \) for every \( n \in \mathbb{N} \).

ii) Let \( \delta_\epsilon = o(1) \). Then for \( \epsilon \ll 1 \) we have

\[ E \left( \frac{\|x^\epsilon(t) - \varphi(t)\|}{\epsilon} \right)^{2n} \leq \exp(\delta_\epsilon t) \text{ for any } t \geq 0. \]

iii) Let \( \delta_\epsilon = o(1) \). Then for \( \epsilon \ll 1 \) we have

\[ E \left[ \left( \frac{\|x^\epsilon(t) - \varphi(t)\|}{\epsilon} \right)^{2n} \right] \leq \exp(\delta_\epsilon t) \text{ for any } t \geq 0. \]

**Proof.** i) Let \( \epsilon > 0 \) and \( t \geq 0 \) be fixed. Note that

\[ x^\epsilon(t) - \varphi(t) = -\int_0^t [F(x^\epsilon(s)) - F(\varphi(s))] ds + \sqrt{\epsilon}B(t) = \]

\[ \int_0^t \left[ \frac{1}{0} DF(\varphi(s) + \theta (x^\epsilon(s) - \varphi(s))) d\theta \right] (x^\epsilon(s) - \varphi(s)) ds + \sqrt{\epsilon}B(t) = \]

\[ \int_0^t A^\epsilon(s) (x^\epsilon(s) - \varphi(s)) ds + \sqrt{\epsilon}B(t), \]

where \( A^\epsilon(s) := \int_0^1 DF(\varphi(s) + \theta (x^\epsilon(s) - \varphi(s))) d\theta \). We will use the induction method. The induction hypothesis had already proved in [3.41]. Let us consider \( f_{n+1}(x) = \|x\|^{2(n+1)} \), \( x \in \mathbb{R}^d \). By the Itô formula, it follows that

\[ d\|x^\epsilon(t) - \varphi(t)\|^{2(n+1)} = \]

\[ 2(n+1)\|x^\epsilon(t) - \varphi(t)\|^{2n} (x^\epsilon(t) - \varphi(t))^* A^\epsilon(t) (x^\epsilon(t) - \varphi(t)) dt + \]

\[ \epsilon(d + 2n)(n + 1)\|x^\epsilon(t) - \varphi(t)\|^{2n} dt + \]

\[ 2(n+1)\sqrt{\epsilon}\|x^\epsilon(t) - \varphi(t)\|^{2n} (x^\epsilon(t) - \varphi(t))^* dB(t). \]
Again by the **strong coercivity** (C2) together the Gronwall trick, we get
\[
E \left[ \left\| x^\varepsilon(t) - \varphi(t) \right\|^{2(n+1)} \right] \leq \frac{c_{n+1} \varepsilon^{n+1}}{2^n \delta^n} \quad \text{for any } t \geq 0.
\]

**ii)** Let \( t \geq 0 \) be fixed. By the Monotone Convergence Theorem, it follows that
\[
E \left[ e^{\delta_t \frac{2 R_{n+1}^2}{\varepsilon}} \right] = \sum_{n=0}^\infty E \left[ \delta_n \frac{2 R_{n+1}^2}{e^n n!} \right].
\]

By item i) of this Proposition, we have
\[
\sum_{n=0}^\infty E \left[ \delta_n \frac{2 R_{n+1}^2}{e^n n!} \right] \leq 1 + \sum_{n=1}^\infty \frac{\delta_n c_n}{2^n \delta^n n!} < +\infty,
\]
when \( \delta_t < \frac{\delta}{2} \) which is always satisfied due to \( \delta_t = o(1) \).

**iii)** We will use the Itô formula for the function \( g_\varepsilon(x) = e^{\delta_\varepsilon \frac{\|x\|^2}{\varepsilon}}, \ x \in \mathbb{R}^d \).

Let \( \kappa_\varepsilon := \frac{\delta_\varepsilon}{2} \). Then
\[
de^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} = -2\kappa_\varepsilon e^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \left( x^\varepsilon(t) - \varphi(t) \right)^* \dot{A} \left( x^\varepsilon(t) - \varphi(t) \right) dt + \epsilon \left( 2\kappa_\varepsilon e^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \left\| x^\varepsilon(t) - \varphi(t) \right\|^2 + \kappa_\varepsilon e^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \right) dt + 2d\sqrt{\epsilon} e^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \left( x^\varepsilon(t) - \varphi(t) \right)^* dB(t).
\]

Using the **strong coercivity** (C2), we obtain
\[
de^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \leq -2\kappa_\varepsilon \delta_\varepsilon e^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \|x^\varepsilon(t) - \varphi(t)\|^2 dt + \epsilon \left( 2\kappa_\varepsilon e^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \left\| x^\varepsilon(t) - \varphi(t) \right\|^2 + \kappa_\varepsilon e^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \right) dt + 2d\sqrt{\epsilon} e^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \left( x^\varepsilon(t) - \varphi(t) \right)^* dB(t).
\]

Therefore
\[
de^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \leq -\kappa_\varepsilon \delta_\varepsilon e^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \|x^\varepsilon(t) - \varphi(t)\|^2 dt + \epsilon \kappa_\varepsilon e^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} dt + 2d\sqrt{\epsilon} e^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \left( x^\varepsilon(t) - \varphi(t) \right)^* dB(t),
\]
when \( \delta_t \leq \frac{\delta}{2} \). By the item i) and the item ii) of this Proposition, we have
\[
\frac{d}{dt} E \left[ e^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \right] \leq \epsilon \kappa_\varepsilon d E \left[ e^{\kappa_\varepsilon \|x^\varepsilon(t) - \varphi(t)\|^2} \right],
\]
when \( \delta_t \leq \frac{\delta}{2} \). Now, using the Gronwall trick, when \( \delta_t \leq \frac{\delta}{2} \)
\[
E \left[ e^{\delta_t \frac{2 R_{n+1}^2}{\varepsilon}} \right] \leq e^{\delta_t \varepsilon} \quad \text{for any } t \geq 0.
\]

**Lemma C.2.** For any \( 0 < \epsilon < 1 \) there exists an unique invariant measure \( \mu_\varepsilon \) for the evolution given by (2.2).
Proof. By the strong coercivity (C2) it follows that
\[ \mathbb{E}[\|x^\varepsilon(t, x_1) - x^\varepsilon(t, x_2)\|^2] \leq e^{-2\delta t}\|x_1 - x_2\|^2 \]
for any \(x_1, x_2 \in \mathbb{R}^d\) and \(t \geq 0\). Using the Chebyshev inequality, we get easily that the process \(\{x^\varepsilon(t) : t \geq 0\}\) is bounded in probability at infinity, i.e., a relation (6) of [11] is fulfilled. Therefore, the existence of an invariant measure for the evolution given by (2.2) is given by Theorem 388 of [11] (or Theorem 11.4.2 page 306 of [13]). The uniqueness follows from Theorem 11.4.3, page 308 of [13]. □

Lemma C.3. Let \(0 < \varepsilon < 1\). For any initial condition \(x_0 \in \mathbb{R}^d\), we have
\[ \lim_{t \to +\infty} d_{TV}(x^\varepsilon(t, x_0), \mu^\varepsilon) = 0. \]

Proof. Under strong coercivity (C2) and Lipschitz condition (3.3), the proof of Theorem 2.7 of [8] can be easily adapted when the noise is Gaussian (as our case) and we get exponential convergence to the unique probability invariant measure in total variation distance. □

Lemma C.4. Let us consider the following matrix differential equation
\[
\begin{aligned}
&\frac{d}{dt}\Sigma(t) = -DF(0)\Sigma(t) - \Sigma(t)(DF(0))^* + I_d, \quad t \geq 0, \\
&\Sigma(0) = \Sigma_0,
\end{aligned}
\]
where \(\Sigma_0\) is a \(d \times d\) matrix. Then \(\lim_{t \to +\infty} \|\Sigma(t) - \Sigma\| = 0\), where \(\Sigma\) is the unique solution of the Lyapunov matrix equation:
\[ DF(0)X + X(DF(0))^* = I_d. \]

Proof. Let us write \(\Lambda = DF(0)\). Notice that by (C1), all eigenvalues of \(\Lambda\) have positive real part. Let us denote by \(\{\phi(t, x) : t \geq 0\}\) the solution of the linear system
\[
\begin{aligned}
&\frac{d}{dt}\phi(t) = -\Lambda\phi(t), \quad t \geq 0, \\
&\phi(0) = x.
\end{aligned}
\]
Then \(\{\phi(t, x) : t \geq 0\}\) is globally asymptotic stable. Then the Lyapunov matrix equation (C2) has a unique positive definite solution \(\Sigma\). From (C2) it follows that \(\Sigma\) is a symmetric matrix. Let
\[ r(t) := \|\Sigma(t) - \Sigma\|^2 = \sum_{i,j=1}^d (\Sigma_{i,j}(t) - \Sigma_{i,j})^2 \text{ for any } t \geq 0. \]

Let us denote by \(\delta_{i,j} = 1\) if \(i = j\) and \(\delta_{i,j} = 0\) if \(i \neq j\) and note that
\[ \sum_{k=1}^d \Lambda_{i,k}\Sigma_{k,j} + \sum_{k=1}^d \Sigma_{i,k}\Lambda_{j,k} = \delta_{i,j} \text{ for any } i, j \in \{1, \ldots, d\}. \]
Then
\[
\frac{d}{dt} r(t) = 2 \sum_{i,j=1}^{d} (\Sigma_{i,j}(t) - \Sigma_{i,j}) \frac{d}{dt} \Sigma_{i,j}(t) = \\
2 \sum_{i,j=1}^{d} (\Sigma_{i,j}(t) - \Sigma_{i,j}) \left( -\sum_{k=1}^{d} \Lambda_{i,k}(\Sigma_{k,j}(t) - \Sigma_{k,j}) - \sum_{k=1}^{d} (\Sigma_{i,k}(t) - \Sigma_{i,k}) \Lambda_{j,k} \right).
\]

After rearrangement the sums and using the coercivity (C1) we have
\[
\frac{d}{dt} r(t) \leq -4 \delta r(t), \ t \geq 0, \\
r(0) = \|\Sigma_0 - \Sigma\|_2.
\]
Then by Gronwall trick we obtain
\[
\|\Sigma(t) - \Sigma\|_2 \leq e^{-4\delta t} \|\Sigma_0 - \Sigma\|_2 \text{ for any } t \geq 0
\]
which implies the statement.

**Remark C.5.** Let \( A \) be a \( d \)-squared matrix such that \( F(x) = Ax \) satisfies coercivity (C1). If we take \( F(x) = Ax \) in the stochastic differential equation (2.2), the covariance matrix associated to the solution of (2.2) satisfies the matrix differential equation (C.1) with initial datum \( \Sigma_0 \) the zero-matrix.

**Lemma C.6.** The covariance matrix of \( y(t) \) converge as \( t \to +\infty \) to a non-degenerate covariance matrix \( \Sigma \), where \( \Sigma \) is the unique solution of the Lyapunov matrix equation:

\[
DF(0)X + X(DF(0))^* = I_d.
\]

**Proof.** For any \( t \geq 0 \), let \( \Lambda(t) \) be the covariance matrix of the \( y(t) \). This matrix satisfies the matrix differential equation:

\[
\begin{cases}
\frac{d}{dt} \Lambda(t) = -DF(\varphi(t))\Lambda(t) - \Lambda(t)(DF(\varphi(t)))^* + I_d, \ t \geq 0, \\
\Lambda(0) = 0.
\end{cases}
\]

Let \( \mathcal{K}_{x_0} := \{x \in \mathbb{R}^d : \|x\| \leq \|x_0\|\} \). By the coercivity (C1) we have \( \varphi(x,t) \in \mathcal{K}_{x_0} \) for any \( x \in \mathcal{K}_{x_0} \) and \( t \geq 0 \). Due to \( F \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}^d) \) there exists a constant \( L := L_{x_0} > 0 \) such that \( \|DF(x) - DF(0)\| \leq L\|x\| \) for any \( x \in \mathcal{K}_{x_0} \).

Let us take \( \eta > 0 \), we can take \( \tau_\eta := \frac{1}{2} \ln \left( \frac{\|x_0\|}{\eta} \right) \) such that

\[
\|DF(\psi(t)) - DF(0)\| \leq L\|\psi(t)\| \leq L\|x_0\|e^{-\delta t} \leq L\eta
\]
for every \( t \geq \tau_\eta \). Let us call \( \tau := \tau_\eta \). Then,

\[
\begin{cases}
\frac{d}{dt} \Delta(t) = -DF(0)\Delta(t) - \Delta(t)(DF(0))^* + I_d, \ t \geq 0 \\
\Delta(0) = \Lambda(\tau).
\end{cases}
\]
Let $\Pi(t) = \Lambda(t + \tau) - \Delta(t)$, $t \geq 0$. Therefore

$$
\begin{align*}
\frac{d}{dt}\Lambda(t) &= -DF(\varphi(t + \tau))\Pi(t) - \Pi(t)(DF(\varphi(t + \tau)))^* + (DF(0) - DF(\varphi(t + \tau)))\Delta(t) + \Delta(t)(DF(0) - DF(\varphi(t + \tau)))^*, \\
\Pi(0) &= 0.
\end{align*}
$$

Therefore

$$
\frac{d}{dt}\|\Pi(t)\|^2 = 2 \sum_{i,j=1}^{d} \Pi_{i,j}(t) \frac{d}{dt}\Pi_{i,j}(t) = 2 \sum_{i,j=1}^{d} \Pi_{i,j}(t) \left( - \sum_{k=1}^{d} DF(\varphi(t + \tau))_{i,k} \Pi_{k,j}(t) - \sum_{k=1}^{d} \Pi_{i,k}(t) DF(\varphi(t + \tau))_{j,k} + R_{i,j}(t) \right),
$$

where

$$
R_{i,j}(t) = \sum_{k=1}^{d} (DF(0)_{i,k} - DF(\varphi(t + \tau))_{i,k})\Delta_{k,j}(t) + \sum_{k=1}^{d} \Delta_{i,k}(t)(DF(0) - DF(\varphi(t + \tau))_{j,k})^*.
$$

Due to the strong coercivity \( \text{(C2)} \) we have

$$
\frac{d}{dt}\|\Pi(t)\|^2 \leq -4\delta\|\Pi(t)\|^2 + \sum_{i,j=1}^{d} \left| \Pi_{i,j}(t) R_{i,j}(t) \right|.
$$

Moreover, using Lipschitz condition \( \text{(C.4)} \) and Lemma \( \text{C.4} \) we obtain

$$
\sum_{i,j=1}^{d} \left| \Pi_{i,j}(t) R_{i,j}(t) \right| \leq C\eta + C\eta\|\Pi(t)\|^2 \text{ for any } t \geq 0,
$$

where $C$ is a positive constant. \textit{A priori we can take} $0 < \eta < \frac{3\delta}{C}$ \text{ and using Gronwall trick we obtain}

$$
\|\Pi(t)\|^2 \leq \frac{C\eta}{\delta} (1 - e^{-\delta t}) < \frac{C\eta}{\delta} \text{ for any } t \geq 0.
$$

Letting $t \to +\infty$ and after $\eta \to 0$ we get $\lim_{t \to +\infty} \|\Pi(t)\|^2 = 0$ which together with Lemma \( \text{C.4} \) implies $\lim_{t \to +\infty} \Lambda(t) = \Sigma$. \( \square \)

\textbf{Acknowledgments}

G. Barrera would like to thank to CONACyT-México for the post-doctorate grant 2016-2017 at the Department of Probability and Statistics, CIMAT. He acknowledges financial support from CNPq-Brazil and the Instituto de Matemática Pura e Aplicada, IMPA for the grant for supporting his Ph.D. studies and the Summer postdoctoral stay 2017, where part of this work...
was done. Also, he would like to express his gratitude to FORDECyT-
CONACYT-México for the travel support to the Summer postdoctoral stay 2017 at IMPA.

REFERENCES

1. A. Arnold, J. Carrillo & C. Manzini. Refined long-time asymptotics for some polymeric fluid flow models, Communications in Mathematical Sciences, Volume 8, Number 3, 2010, 763-782.
2. A. Galves, E. Olivieri & M. Vares. Metastability for a class of dynamical systems subject to small random perturbations, Annals of Probability, Volume 15, 1987, 1288-1305.
3. B. Jourdain, C. Le Bris, T. Lelièvre & F. Otto, Long-time asymptotics of a multiscale model for polymeric fluid flows, Archive for Rational Mechanics and Analysis, Volume 181, Number 1, 2006, 97-148.
4. C. Kipnis & C. Newman. The metastable behavior of infrequently observed, weakly random, one-dimensional diffusion processes, SIAM Journal on Applied Mathematics, Volume 45, Number 6, 1985, 972-982.
5. D. Stroock & S. Varadhan. Multidimensional diffusion processes, Springer-Verlag, Berlin, 1997.
6. D. Aldous & P. Diaconis. Shuffling cards and stopping times, American Mathematical Monthly, Volume 93, Number 5, 1986, 333-348.
7. D. Aldous & P. Diaconis. Strong uniform times and finite random walks, Advances in Applied Mathematics, Volume 8, Number 1, 1987, 69-97.
8. E. Priola, A. Shirikyan, L. Xu & J. Zabczyk. Exponential ergodicity and regularity for equations with Lévy noise, Stochastic Processes and their Applications, Volume 122, 2012, 106-133.
9. E. Olivieri & M. Vares. Large deviations and metastability, Cambridge University Press, 2004.
10. G. Barrera & M. Jara. Abrupt convergence for stochastic small perturbations of one dimensional dynamical systems, Journal of Statistical Physics, Volume 163, Number 1, 2016, 113-138.
11. G. Da Prato, D. Gatarek & J. Zabczyk. Invariant measures for semilinear stochastic equations, Stochastic Analysis and Applications, Volume 10, Number 4, 1992, 387-408.
12. L. Saloff-Coste. Random walks on finite groups, Probability & Discrete Structures, Springer, 2004, 263-346.
13. G. Kallianpur & P. Sundar. Stochastic Analysis and Diffusion Processes, Oxford University Press, 2014.
14. J. Barrera & B. Yeart, Bounds for left and right window cutoffs, ALEA-Latin American Journal of Probability and Mathematical Statistics, Volume 11, Number 2, 2014, 445-458.
15. J. Barrera, O. Bertoncini & R. Fernández. Abrupt convergence and escape behavior for birth and death chains, Journal of Statistical Physics, Volume 137, Number 4, 2009, 595-623.
16. L. Perko. Differential equations and dynamical systems, Third edition, Springer, 2001.
17. M. Freidlin & A. Wentzell. Random perturbations of dynamical systems, Springer, 2012.
18. M. Freidlin & A. Wentzell. On small random perturbations of dynamical systems, Russian Mathematical Surveys, Volume 25, 1970, 1-55.
19. M. Freidlin & A. Wentzell. Some problems concerning stability under small random perturbations, Theory Probability Applied, Volume 17, 1972, 269-283.
20. M. Day. *Exponential leveling of stochastically perturbed dynamical systems*, SIAM Journal on Mathematical Analysis, Volume 13, 1982, 532-540.

21. M. Day. *On the exponential exit law in the small parameter exit problem*, Stochastics, Volume 8, 1983, 297-323.

22. P. Diaconis. *The cut-off phenomenon in finite Markov chains*, Proceedings of the National Academy of Sciences, USA, Volume 93, 1996, 1659-1664.

23. P. Hartman. *On local homeomorphisms of Euclidean spaces*, Bulletin of Mexican Mathematical Society, Volume 5, 1960, 220-241.

24. W. Siegert. *Local Lyapunov exponents: sublimiting growth rates of linear random differential equations*, Springer, 2009.

Mathematics Research Center, CIMAT. Jalisco S/N, Valenciana, ZIP: 36240, Guanajuato, Guanajuato, México.

E-mail address: bvargas@cimat.mx

Instituto de Matemática Pura e Aplicada, IMPA. Estrada Dona Castorina 110, Jardim Botânico, ZIP: 22460-320. Rio de Janeiro, Rio de Janeiro, Brazil.

E-mail address: monets@impa.br