THERMALISATION FOR SMALL RANDOM PERTURBATIONS OF DYNAMICAL SYSTEMS

BY GERARDO BARRERA[1] AND MILTON JARA[2]

[1]*Department of Mathematical and Statistical Sciences, University of Alberta, barrerav@ualberta.ca*
[2]*Instituto Nacional de Matemática Pura e Aplicada, IMPA, monets@impa.br*

We consider an ordinary differential equation with a unique hyperbolic attractor at the origin, to which we add a small random perturbation. It is known that under general conditions, the solution of this stochastic differential equation converges exponentially fast to an equilibrium distribution. We show that the convergence occurs abruptly: in a time window of small size compared to the natural time scale of the process, the distance to equilibrium drops from its maximal possible value to near zero, and only after this time window the convergence is exponentially fast. This is what is known as the cut-off phenomenon in the context of Markov chains of increasing complexity. In addition, we are able to give general conditions to decide whether the distance to equilibrium converges in this time window to a universal function, a fact known as profile cut-off.

1. Introduction. This paper is a multidimensional generalisation of the previous work [10] by the same authors. Our main goal is the study of the convergence to equilibrium for a family of stochastic small random perturbations of a given dynamical system in $\mathbb{R}^d$. Consider an ordinary differential equation with a unique hyperbolic global attractor. Without loss of generality, we assume that the global attractor is located at the origin. Under general conditions, as time goes to infinity, any solution of this differential equation approaches the origin exponentially fast. We perturb the deterministic dynamics by a Brownian motion of small intensity. It is well known that, again under very general conditions, as time goes to infinity, any solution of this stochastic differential equation converges in distribution to an equilibrium law. The convergence can be improved to hold with respect to the total variation distance. The theory of Lyapunov functions allows us to show that this convergence, for each fixed perturbation, is again exponentially fast. We show that the convergence occurs abruptly: when the intensity of the noise goes to zero, the total variation distance between the law of the stochastic dynamics and the law of its equilibrium in a time window around the cut-off time decreases from one to near zero abruptly, and only after this time window the convergence is exponentially fast. This fact is known as cut-off phenomenon. Moreover, when a properly normalised ω-limit set of the initial datum of the deterministic differential equation is contained in a sphere, we are able to prove convergence of the distance to equilibrium to a universal function, a fact known as profile cut-off or profile thermalisation in the context of ergodic Markov processes.

To be more precise, we are concerned about the abrupt convergence to equilibrium in the total variation distance for systems of the form

$$(1.1) \qquad \begin{cases} \mathrm{d}x^\epsilon(t) = -F\big(x^\epsilon(t)\big)\,\mathrm{d}t + \sqrt{\epsilon}\,\mathrm{d}B(t) & \text{for } t \geq 0, \\ x^\epsilon(0) = x_0, \end{cases}$$

where F is a given vector field with a unique hyperbolic attractor at 0 and $\{B(t) : t \geq 0\}$ is a standard Brownian motion in $\mathbb{R}^d$. Notice that systems described by the stochastic differential equation (1.1) are not necessarily reversible. In statistical physics, equation (1.1) is known as an *overdamped Langevin dynamics*, and it is used to model fluctuations of stationary states.

In the small noise asymptotics, the stochastic dynamics (1.1) fluctuates around the attractor of the deterministic dynamics which is called *relaxation dynamics* or *zero-noise dynamics*. Assuming that the deterministic dynamics is *strongly coercive* together with some *growth condition* on F, when the intensity ϵ of the noise goes to zero, in a time windows of small size compared to the natural time scale of the process, the total variation distance to equilibrium drops from near one to near zero.

Dynamical systems subjected to small Gaussian perturbations have been studied extensively, see the book of M. Freidlin and A. Wentzell [24] which discusses this problem in great detail; see also M. Freidlin and A. Wentzell [22, 23], M. Day [18, 19] and W. Siegert [49]. This treatment has inspired many works and considerable effort was concerned about purely local phenomena, that is, on the computation of exit times and exit probabilities from neighbourhoods of fixed points that are carefully stipulated not to contain any other fixed point of the deterministic dynamics.

The theory of large deviations allows us 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 that the trajectories have to overcome. Consequently, for a multi-well potential one can obtain a series of exponentially nonequivalent time scales given by the wells-mean exit times. Moreover, the normalised exit times are asymptotically exponentially distributed and have a memoryless property, for further details see A. Galves, E. Olivieri and M. Vares [25], E. Olivieri and M. Vares [44] and C. Kipnis and C. Newman [32]. There are situations in which the analysis at the level of large deviations is not enough, and it is necessary to study the distributional scaling limits for the exit distributions; for more details see Y. Bakhtin [7] and [8].

The cut-off phenomenon was extensively studied in the eighties to describe the phenomenon of abrupt convergence that appears in models of card shuffling, Ehrenfest urns and random transpositions, see for instance D. Aldous and P. Diaconis [1] and [2]. In general, it is a challenging problem to prove that a specific family of stochastic models exhibit or does not exhibit a cut-off phenomenon. It requires a complete understanding of the dynamics of the specific random process.

Since the appearance of [1] many families of stochastic processes have been shown to have similar properties. Various notions of cut-off have been proposed; see J. Barrera and B. Ycart [12] and P. Diaconis [20] for an account. We refer to the book of D. Levin et al. ([37], Chapter 18) for an introduction of the subject in the Markov chain setting, L. Saloff-Coste [47] provides an extensive list of random walks for which the cut-off phenomenon holds, P. Diaconis [20] for a review on the finite Markov chain case, S. Martínez and B. Ycart [42] for the case of Markov chains with countably infinite state space, G. Chen and L. Saloff-Coste [17] for Brownian motions on a compact Riemann manifold, B. Lachaud [34] and G. Barrera [9] for Ornstein–Uhlenbeck processes on the line and G. Barrera and M. Jara [10] for stochastic small perturbations of one-dimensional dynamical systems.

Roughly speaking, thermalisation or window 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 time window, the 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 and R. Fernández [11].

The main result of this article, Theorem 2.2, states that when the deterministic dynamics is *strongly coercive* and satisfies some *growth condition*, the family of perturbed dynamics presents thermalisation (windows cut-off) as we describe in Section 2. Moreover, in Corollary 2.9 and Corollary 2.11 we give a necessary and sufficient condition for having profile thermalisation (profile cut-off). We point out that our condition is always satisfied by reversible dynamics; that is, when $F(x) = \nabla V(x)$, $x \in \mathbb{R}^d$, and also for a large class of dynamics that are nonreversible. Moreover, our condition is *open* in the sense that if it holds for a given field F, then it holds in an open neighbourhood of F with respect to the $\mathcal{C}^\infty$ topology.

Nonreversible dynamics naturally appear, for example, in polymeric fluid dynamics or Wigner–Fokker–Planck equations; see A. Arnold, J. Carrillo and C. Manzini [3] and B. Jourdain, C. Le Bris, T. Lelièvre and F. Otto [30]. Nonreversible systems arise in the theory of activated process in glasses and other disordered materials, chemical reactions far from equilibrium, stochastic modelled computer networks, evolutionary biology and theoretical ecology, see R. Maier and D. Stein [38] and [39].

Notice that the set of symmetric matrices is not open. In particular, reversibility is not a generic property of dynamical systems. On the other hand, hyperbolicity is an open property, meaning that it is stable under small perturbations of the vector field. Moreover, in general for the nonreversible case, there is not an explicit formula for the invariant measure of the random dynamics (1.1) as in the reversible case. For reversible dynamics, analytic methods from quantum mechanics have been used to compute asymptotic expansions in the diffusivity $\sqrt{\epsilon}$. The strong point is that full asymptotic expansions in $\sqrt{\epsilon}$ and sharp estimates can be done. However, so far only applicable for reversible diffusion process. For more details, see [15] and [16]. Therefore, it is desirable to have a treatment that does not rely on these properties, namely reversibility and/or explicit knowledge of invariant measures.

Our idea is to carry out this asymptotic expansion in $\sqrt{\epsilon}$ by probabilistic methods. It turns out that the hyperbolic contracting nature of the underlying dynamics can be used to show that a first-order expansion gives a description of the original dynamics which is good for times much larger than the time at which equilibration occurs. This expansion is natural in the context of perturbed dynamical systems and it is known in Physics as Van Kampen's approximation. It is also the same expansion introduced in [10]. For further details see Section 3 of [4].

Roughly speaking, our main results hold true since they also hold for the Ornstein–Uhlenbeck processes. Let A be a square matrix of dimension d. It is well known that the Ornstein–Uhlenbeck equation:

$$\begin{cases} d\gamma^\epsilon(t) = -A\gamma^\epsilon(t) \, dt + \sqrt{\epsilon} \, dB(t) & \text{for any } t \geq 0, \\ \gamma^\epsilon(0) = x_0 \end{cases}$$

can be solved explicitly. One gets the expression

$$\gamma^\epsilon(t) = e^{-At} x_0 + \sqrt{\epsilon} e^{-At} \int_0^t e^{As} \, dB(s)$$

for the solution. Therefore, for any $t > 0$, $\gamma^\epsilon(t)$ has a Gaussian distribution with mean vector $e^{-At} x_0$ and covariance matrix $\epsilon \Sigma(t)$, where $\Sigma(t)$ has an integral representation. When A is

a symmetric matrix, $\Sigma(t)$ is explicit. When A is not symmetric, for generic cases it is not possible to find an explicit formula for $\Sigma(t)$. However, if we assume that $\mathrm{Re}(\lambda) > 0$ for any eigenvalue λ of A, then it is not hard to prove that $\Sigma(t)$ converges as t goes to infinity to Σ which is symmetric and positive definite. We point out that Σ can be computed explicitly just in few cases. Bearing all this in mind, a refined analyses of the convergence as t goes by of the process $\{\gamma^\epsilon(t) : t \geq 0\}$ to its equilibrium in the total variation distance can be done.

On the other hand, in [10] we smoothed our vector field F. We assumed that F' and F'' are bounded. Therefore, our *a priori estimates* were straightforward. Later, we removed those assumptions by constructing a vector field $\tilde{F}$ such that $\tilde{F}$ agrees with F in a neighbourhood of zero, $\tilde{F}$ is linear around infinity and strongly coercive. In the multidimensional setting the latter is not straightforward as in dimension one. Since we need to carry out computations in a more appropriated way, we assume that our original vector field F satisfies some growth condition together with a strong coercivity condition (see Section 2.4). We used these conditions to carry out second-moment estimates from our original dynamics with its corresponding nonhomogeneous linear approximations which are meaningful for times much larger than the time at which equilibration happens.

Notice that in the one-dimensional case, the stochastic dynamics is always reversible. Therefore, profile cut-off always holds. In our multidimensional, nonreversible setting, a more refined analysis of the first-order expansion is needed in order to be able to discern whether profile cut-off holds or not. In particular, a more refined analysis of the nonhomogeneous Ornstein–Uhlenbeck and of the unperturbed dynamical system are required. This analysis reveals that window cut-off does not always imply profile cut-off. A consequence of our analysis is a L^1-version of the local central limit theorem (see Proposition 3.7) for the invariant measure of (1.1), which could be of independent interest.

This material is organised 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 related linear approximations which are the main ingredient in order to prove the main result in the end of this section. Finally, we provide an Appendix which is divided in three sections as follows: Section A gives useful properties for the total variation distances between Gaussian distributions. Sections B and C provide 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 $\mathcal{C}^2(\mathbb{R}^d, \mathbb{R}^d)$. For each $x \in \mathbb{R}^d$, let $\{\varphi(t, x) : t \in [0, \tau_x)\}$ be the solution of the deterministic differential equation

$$(2.1) \qquad \begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\varphi(t) = -F\big(\varphi(t)\big) & \text{for } 0 \leq t < \tau_x, \\ \varphi(0) = x \end{cases}$$

where τ_x denotes the *explosion time*. Since F is smooth, this equation has a unique solution. Since we have not imposed any growth condition on F, τ_x may be finite. We denote by $\|\cdot\|$ the Euclidean norm in $\mathbb{R}^d$ and by $\langle \cdot, \cdot \rangle$ the standard inner product of $\mathbb{R}^d$. Under the condition

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

a straightforward application of the Lemma C.7 (Gronwall's inequality) implies that the explosion time τ_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.

Let us recall the following terminology from dynamical systems. 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 in 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 $\mathrm{Re}(\lambda) \neq 0$ for any eigenvalue λ of the Jacobian matrix $DF(y)$ of F at y. By the Hartman–Grobman Theorem (see Theorem (Hartman) page 127 of [45] or the celebrated paper of P. Hartman [26]), a hyperbolic fixed point y of (2.1) is an attractor if and only if $\mathrm{Re}(\lambda) > 0$ for any eigenvalue λ of the matrix $DF(y)$.

From now on, we will always assume that

$$0 \text{ is a fixed point of (2.1).}$$

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

(C) $\langle x, F(x) \rangle \geq \delta \|x\|^2$ for any $x \in \mathbb{R}^d$.

Indeed, notice that

$$\frac{d}{dt} \|\varphi(t)\|^2 = 2 \left\langle \varphi(t), \frac{d}{dt} \varphi(t) \right\rangle = \langle \varphi(t), -F(\varphi(t)) \rangle \leq -2\delta \|\varphi(t)\|^2$$

for any $t \geq 0$. Then Lemma C.7 allows us to deduce that

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

In other words, $\varphi(t, x)$ converges to 0 exponentially fast as $t \to +\infty$. Notice that the eigenvalues of the Jacobian matrix of F at zero, $DF(0)$, might be complex numbers. From (C) we have $\mathrm{Re}(\lambda) \geq \delta$ for any eigenvalue λ of $DF(0)$. In other words, 0 is a hyperbolic attracting point for (2.1).

Recall that for any $\lambda \in \mathbb{C}$ and $v \in \mathbb{C}^d$,

$$\lambda v = (\mathrm{Re}(\lambda) + i \, \mathrm{Im}(\lambda))(\mathrm{Re}(v) + i \, \mathrm{Im}(v))$$
$$= \mathrm{Re}(\lambda) \, \mathrm{Re}(v) - \mathrm{Im}(\lambda) \, \mathrm{Im}(v) + i (\mathrm{Im}(\lambda) \, \mathrm{Re}(v) + \mathrm{Re}(\lambda) \, \mathrm{Im}(v)).$$

Let $v \in \mathbb{C}^d$ an eigenvector associated to the eigenvalue λ of $DF(0)$. Then

(2.3) $-(\mathrm{Re}(\lambda) - \delta) \|\mathrm{Im}(v)\|^2 \leq \mathrm{Im}(\lambda) \langle \mathrm{Re}(v), \mathrm{Im}(v) \rangle \leq (\mathrm{Re}(\lambda) - \delta) \|\mathrm{Re}(v)\|^2.$

Particularly, from (2.3) we have that (C) does not allow to control the imaginary part of the eigenvalues of $DF(0)$. Typically and roughly speaking, the dynamical system associated to (2.1) is a "uniformly contracting spiral".

The following lemma provides us the asymptotics of $\varphi(t)$ as t goes to $+\infty$. It will be important for determining the cut-off time and time window.

LEMMA 2.1.   *Assume that* (C) *holds. Then for any* $x_0 \in \mathbb{R}^d \setminus \{0\}$ *there exist* $\lambda := \lambda(x_0) > 0$, $\ell := \ell(x_0)$, $m := m(x_0) \in \{1, \ldots, d\}$, $\theta_1 := \theta_1(x_0), \ldots, \theta_m := \theta_m(x_0) \in [0, 2\pi)$, $v_1 := v_1(x_0), \ldots, v_m := v_m(x_0)$ *in* $\mathbb{C}^d$ *linearly independent and* $\tau := \tau(x_0) > 0$ *such that*

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

To made sense the multiplication by $e^{i\theta_k t}$ in the statement of Lemma 2.1, we point out that we are working in the complexified space. This lemma will be proved in Appendix B, where we give 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 $v_k \in \mathbb{C}^d$, $k = 1, \ldots, m$ are elements of the Jordan decomposition of the matrix $DF(0)$.

2.2. *The cut-off phenomenon.* Let μ, ν be two probability measures in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. We say that a probability measure π in $(\mathbb{R}^d \times \mathbb{R}^d, \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d))$ is a *coupling* between μ and ν if for any Borel set $B \in \mathcal{B}(\mathbb{R}^d)$,

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

In that case we say that $\pi \in \mathcal{C}(\mu, \nu)$. The *total variation* distance between μ and ν is defined as

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

Notice that the diameter with respect to $d_{\mathrm{TV}}(\cdot, \cdot)$ of the set $\mathcal{M}_1^+(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ of probability measures defined 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 $d_{\mathrm{TV}}(X, Y)$ instead of $d_{\mathrm{TV}}(\mathbb{P}(X \in \cdot), \mathbb{P}(Y \in \cdot))$.

For simplicity, we also write $d_{\mathrm{TV}}(X, \mu_Y)$ in place of $d_{\mathrm{TV}}(X, Y)$, where μ_Y is the distribution of the random variable Y. For an account of the equivalent formulations of the total variation distance (normalised or not normalised), we recommend the book of A. Kulik ([33], Chapter 2).

For any $\epsilon \in (0, 1]$, let x^ϵ be the continuous time stochastic process $\{x^\epsilon(t) : t \geq 0\}$. We say that a family of stochastic processes $\{x^\epsilon\}_{\epsilon \in (0,1]}$ has *thermalisation* at position $\{t^\epsilon\}_{\epsilon \in (0,1]}$, window $\{w^\epsilon\}_{\epsilon \in (0,1]}$ and state $\{\mu^\epsilon\}_{\epsilon \in (0,1]}$ if:

(i)
$$\lim_{\epsilon \to 0} t^\epsilon = +\infty \quad \text{and} \quad \lim_{\epsilon \to 0} \frac{w^\epsilon}{t^\epsilon} = 0,$$

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

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

If, for any $\epsilon \in (0, 1]$, x^ϵ is a Markov process with a unique invariant measure and μ^ϵ is the invariant measure of the process x^ϵ we say that the family $\{x^\epsilon\}_{\epsilon \in (0,1]}$ presents *thermalisation* or *window 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

(ii') $\quad \lim_{\epsilon \to 0} d_{\mathrm{TV}}(x^\epsilon(t^\epsilon + cw^\epsilon), \mu^\epsilon) =: G(c) \quad$ for any $c \in \mathbb{R}$,

we say that there is *profile thermalisation* or *profile cut-off*. Notice that (ii') implies (ii) and (iii), and therefore profile thermalisation (respectively profile cut-off) is a stronger notion than thermalisation (respectively window cut-off).

2.3. *The overdamped Langevin dynamics.* 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:

(2.4)
$$\begin{cases} dx^\epsilon(t) = -F(x^\epsilon(t)) \, dt + \sqrt{\epsilon} \, dB(t) & \text{for } t \geq 0, \\ x^\epsilon(0) = x_0. \end{cases}$$

Stochastic differential equation (2.4) is used in molecular modelling. In that context $\epsilon = 2\kappa\tau$, where τ is the temperature of the system and κ is the Boltzmann constant. In statistical physics, equation (2.4) has a computational interest to modelling a sample of a Gibbs measure in high-dimensional Euclidean spaces. Denote by $(\Omega, \mathcal{F}, \mathbb{P})$ the probability space where $\{B(t) : t \geq 0\}$ is defined and denote by $\mathbb{E}$ the expectation with respect to $\mathbb{P}$. Notice that (2.4) has a unique strong solution (see Remark 2.1.2 page 57 of [49] or Theorem 10.2.2 of [50]), and therefore $\{x^\epsilon(t, x_0) : t \geq 0\}$ can be taken as a stochastic process in the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

In order to avoid unnecessary notation, we 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 $\{\varphi(t, x_0) : t \geq 0\}$. Since $\epsilon \in (0, 1]$, for simplicity we write $\lim_{\epsilon \to 0}$ instead of $\lim_{\epsilon \to 0^+}$.

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 are interested in the law of $x^\epsilon(t)$ for times t of order $\mathcal{O}(\log(1/\epsilon))$, where *thermalisation* or *window cut-off* phenomenon appears.

Under (C), for any $\epsilon \in (0, 1]$, the process $\{x^\epsilon(t) : t \geq 0\}$ is uniquely ergodic with stationary measure μ^ϵ; see Lemma C.3 for details. Moreover, the process is strongly Feller. In particular, the process visits infinitely often every nonempty open set of the state space $\mathbb{R}^d$. The stationary measure μ^ϵ is absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^d$. The density ρ^ϵ of μ^ϵ is smooth and solves the stationary Fokker–Planck equation

$$\frac{\epsilon}{2} \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2} (\rho^\epsilon(x)) + \sum_{j=1}^{d} \frac{\partial}{\partial x_j} (F_j(x) \rho^\epsilon(x)) = 0 \quad \text{for any } x \in \mathbb{R}^d,$$

where $F = (F_1, \ldots, F_d)^T$, for details see [49] (pages 60–63). When the process is reversible, that is, $F(x) = \nabla V(x)$, $x \in \mathbb{R}^d$, for some scalar function V (also called potential), the stationary measure μ^ϵ is of the Gibbs type:

(2.5)
$$\mu^\epsilon(dx) = \frac{1}{\mathcal{Z}^\epsilon} e^{-\frac{2V(x)}{\epsilon}} \, dx \quad \text{where } \mathcal{Z}^\epsilon = \int_{\mathbb{R}^d} e^{-\frac{2V(x)}{\epsilon}} \, dx < +\infty.$$

The normalised constant $\mathcal{Z}^\epsilon$ is called the partition function. If the vector field F can be decomposed as

$$F(x) = \nabla V(x) + b(x) \quad \text{for any } x \in \mathbb{R}^d,$$

where $V : \mathbb{R}^d \to \mathbb{R}$ is a scalar function and $b : \mathbb{R}^d \to \mathbb{R}^d$ is a vector field which satisfies the divergence-free condition

$$\operatorname{div}(e^{-\frac{2}{\epsilon}V(x)} b(x)) = 0 \quad \text{for any } x \in \mathbb{R}^d,$$

then under some appropriate assumptions on V at infinity, that is,

$$\frac{1}{2} \|\nabla V(x)\|^2 - \Delta V(x) \to +\infty \quad \text{as } \|x\| \to +\infty,$$

the probability measure μ^ϵ given by (2.5) remains stationary for (2.4). For details see [28, 29, 51] and [36]. In this situation, using the Laplace Method, asymptotics as $\epsilon \to 0$ for μ^ϵ can be obtained; see [27] and [5] for further details. In general, the equilibrium measure can

be expressed as an integral of a Green function, but aside from a few simple cases, there are no closed expressions for it. In this case, the Freidlin–Wentzell theory implies that the non-Gibbs measure μ^ϵ is equivalent to a Gibbs measure with a "quasipotential" $\tilde{V}$ playing the role of the potential energy, see for instance [38, 40, 43] and [48]. However, the study of the regularity of the quasipotential is a nontrivial mathematical issue, for details see [6]. For our purposes, no transverse condition on the vector field F is assumed and also we do not need that the Gibbs measure remains stationary for (2.4); for further details see [14] and the references therein.

In many theoretical or applied problems involving ergodic processes, it is important to estimate the time until the distribution of the process is close to its equilibrium distribution. Under some *strong coercivity condition* and *growth condition* that we will state precisely in Section 2.4, we will prove that the law of $x^\epsilon(t)$ converges in total variation distance to μ^ϵ in a *time window*

$$(2.6) \qquad w^\epsilon := \frac{1}{\lambda} + o(1)$$

of order $\mathcal{O}(1)$ around the *cut-off time*

$$(2.7) \qquad t^\epsilon_{\mathrm{mix}} := \frac{1}{2\lambda} \ln(1/\epsilon) + \frac{\ell - 1}{\lambda} \ln\big(\ln(1/\epsilon)\big) + \tau,$$

where λ, ℓ and τ are the positive constants associated to x_0 in Lemma 2.1.

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

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 Ξ. Let I_d be the identity $d \times d$-matrix. Given a matrix A, denote by A^* the transpose matrix of A. Recall that for any $y \in \mathbb{R}^d$, $DF(y)$ denotes the Jacobian matrix of F at y.

A sufficient condition that allows us to uniformly push back to the origin the dynamics of (2.4) is the following *strong coercivity condition*: there exists $\delta > 0$ such that

$$(\mathrm{H}) \qquad \langle x, DF(y)x \rangle \geq \delta \|x\|^2 \quad \text{for any } x, y \in \mathbb{R}^d.$$

At the beginning of Section 3 we will see that (H) implies (C). To control the growth of the vector field F around infinity, we assume the following *growth condition*: there exist positive constants c_0 and c_1 such that

$$(\mathrm{G}) \qquad \|F(x)\| \leq c_0 e^{c_1\|x\|^2} \quad \text{for any } x \in \mathbb{R}^d.$$

Since we use the Itô formula several times, in order to avoid technicalities we always assume that $F \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}^d)$. In the case of a stochastic perturbation of a dynamical system satisfying the strongly coercivity condition (H) and the growth condition (G) we prove thermalisation.

THEOREM 2.2. *Assume that* (H) *and* (G) *hold. Let* $\{x^\epsilon(t, x_0) : t \geq 0\}$ *be the solution of* (2.4) *and denote by* μ^ϵ *the unique invariant probability measure for the evolution given by* (2.4). *Denote by*

$$d^\epsilon(t) = \mathrm{d}_{\mathrm{TV}}\big(x^\epsilon(t, x_0), \mu^\epsilon\big) \quad \text{for any } t \geq 0$$

the total variation distance between the law of the random variable $x^\epsilon(t, x_0)$ *and its invariant probability* μ^ϵ. *Consider the cut-off time* $t^\epsilon_{\mathrm{mix}}$ *given by* (2.7) *and the time window given by* (2.6). *Let* $x_0 \neq 0$. *Then for any* $c \in \mathbb{R}$ *we have*

$$\lim_{\epsilon \to 0} \big|d^\epsilon\big(t^\epsilon_{\mathrm{mix}} + cw^\epsilon\big) - D^\epsilon\big(t^\epsilon_{\mathrm{mix}} + cw^\epsilon\big)\big| = 0,$$

where

$$(2.8) \qquad D^{\epsilon}(t) = d_{TV}\left(\mathcal{G}\left(\frac{(t-\tau)^{\ell-1}}{e^{\lambda(t-\tau)}\sqrt{\epsilon}} \Sigma^{-1/2} \sum_{k=1}^{m} e^{i\theta_k(t-\tau)} v_k, I_d \right), \mathcal{G}(0, I_d) \right)$$

for any $t \geq \tau$ with $m, \lambda, \ell, \tau, \theta_1, \ldots, \theta_m, v_1, \ldots, v_m$ are the constants and vectors associated to x_0 in Lemma 2.1, *and the matrix Σ is the unique solution of the matrix Lyapunov equation*

$$(2.9) \qquad DF(0)X + X\big(DF(0)\big)^* = I_d.$$

REMARK 2.3. The last theorem tells us that the total variation distance between the law of $x^{\epsilon}(t)$ and its equilibrium μ^{ϵ} can be well approximated in a time window (2.6) around the cut-off time (2.7) by the total variation distance between two Gaussian distributions (2.8).

REMARK 2.4. From Lemma A.2 we deduce an "explicit" formula for the distance (2.8), that is,

$$D^{\epsilon}(t) = \sqrt{\frac{2}{\pi}} \int_0^{\|m^{\epsilon}(t)\|/2} e^{-\frac{x^2}{2}} \, dx,$$

where $m^{\epsilon}(t) = \frac{(t-\tau)^{\ell-1}}{e^{\lambda(t-\tau)}\sqrt{\epsilon}} \Sigma^{-1/2} \sum_{k=1}^{m} e^{i\theta_k(t-\tau)} v_k$ for any $t \geq \tau$.

REMARK 2.5. Since the linear differential equation

$$dx(t) = -DF(0)x(t) \, dt \quad \text{for any } t \geq 0$$

is asymptotically stable, then the matrix Lyapunov equation (2.9) has a unique solution Σ which is symmetric and positive definite and it is given by the formula:

$$\Sigma = \int_0^{\infty} e^{-DF(0)s} e^{-(DF(0))^*s} \, ds.$$

For more details, see Theorem 1, page 443 of [35]. If in addition, $DF(0)$ is symmetric then Σ is easily computable and it is given by $\Sigma = \frac{1}{2}(DF(0))^{-1}$.

From Theorem 2.2 we have the following consequences that we write as corollaries. To made the presentation more fluent, in all the corollaries below, we will assume the same hypothesis of Theorem 2.2 and keep the same notation. For any $\epsilon \in (0, 1]$ and $x_0 \in \mathbb{R}^d$, denote by x^{ϵ, x_0} the Markov process $\{x^{\epsilon}(t, x_0) : t \geq 0\}$.

COROLLARY 2.6. *Suppose that $x_0 \neq 0$. Thermalisation for the distance D^{ϵ} at cut-off time $t^{\epsilon}_{\text{mix}}$ and time window w^{ϵ} is equivalent to thermalisation for the distance d^{ϵ} at cut-off time $t^{\epsilon}_{\text{mix}}$ and time window w^{ϵ}. The same holds true for profile thermalisation.*

PROOF. It follows easily from Theorem 2.2 and the following inequalities

$$D^{\epsilon}(t^{\epsilon}_{\text{mix}} + cw^{\epsilon}) \leq |D^{\epsilon}(t^{\epsilon}_{\text{mix}} + cw^{\epsilon}) - d^{\epsilon}(t^{\epsilon}_{\text{mix}} + cw^{\epsilon})| + d^{\epsilon}(t^{\epsilon}_{\text{mix}} + cw^{\epsilon})$$

and

$$d^{\epsilon}(t^{\epsilon}_{\text{mix}} + cw^{\epsilon}) \leq |D^{\epsilon}(t^{\epsilon}_{\text{mix}} + cw^{\epsilon}) - d^{\epsilon}(t^{\epsilon}_{\text{mix}} + cw^{\epsilon})| + D^{\epsilon}(t^{\epsilon}_{\text{mix}} + cw^{\epsilon}). \qquad \square$$

COROLLARY 2.7 (Thermalisation). *Suppose that $x_0 \neq 0$. Theorem 2.2 implies thermalisation for the family of Markov processes $\{x^{\epsilon, x_0}\}_{\epsilon \in (0,1]}$.*

PROOF. From Corollary 2.6 we only need to analyse the distance D^ϵ. Notice

$$0 < L$$

$$:= \liminf_{t \to +\infty} \left\| \sum_{k=1}^m e^{i\theta_k(t-\tau)} v_k \right\|$$

$$\leq \limsup_{t \to +\infty} \left\| \sum_{k=1}^m e^{i\theta_k(t-\tau)} v_k \right\| \leq \sum_{k=1}^m \|v_k\| =: U < +\infty,$$

where first inequality follows from the Cantor diagonal argument and the fact that $v_1, \ldots, v_m$ are linearly independent. From Remark 2.4 we have

$$D^\epsilon(t) = \sqrt{\frac{2}{\pi}} \int_0^{\|m^\epsilon(t)\|/2} e^{-\frac{x^2}{2}} \, dx,$$

where $m^\epsilon(t) = \frac{(t-\tau)^{\ell-1}}{e^{\lambda(t-\tau)}\sqrt{\epsilon}} \Sigma^{-1/2} \sum_{k=1}^m e^{i\theta_k(t-\tau)} v_k$ for any $t \geq \tau$. Straightforward computations led us to

$$\lim_{\epsilon \to 0} \frac{(t_{\text{mix}}^\epsilon - \tau + cw^\epsilon)^{\ell-1} e^{-\lambda(t_{\text{mix}}^\epsilon - \tau + cw^\epsilon)}}{\sqrt{\epsilon}} = (2\gamma)^{1-\ell} e^{-c}$$

for any $c \in \mathbb{R}$. Notice that there exist positive constants b_0 and b_1 such that $b_0\|v\| \leq \|\Sigma^{-1/2}v\| \leq b_1\|v\|$ for any $v \in \mathbb{R}^d$. Then for any $c \in \mathbb{R}$ we obtain

$$\tilde{L}e^{-c} \leq \liminf_{\epsilon \to 0} \|m^\epsilon(t_{\text{mix}}^\epsilon + cw^\epsilon)\| \leq \limsup_{\epsilon \to 0} \|m^\epsilon(t_{\text{mix}}^\epsilon + cw^\epsilon)\| \leq \tilde{U}e^{-c},$$

where $\tilde{L} = L b_0 (2\gamma)^{1-\ell}$ and $\tilde{U} = U b_0 (2\gamma)^{1-\ell}$. From Lemma A.6 and Lemma A.2 we deduce

$$\sqrt{\frac{2}{\pi}} \int_0^{\tilde{L}e^{-c}/2} e^{-\frac{x^2}{2}} \, dx \leq \liminf_{\epsilon \to 0} D^\epsilon(t_{\text{mix}}^\epsilon + cw^\epsilon)$$

$$\leq \limsup_{\epsilon \to 0} D^\epsilon(t_{\text{mix}}^\epsilon + cw^\epsilon) \leq \sqrt{\frac{2}{\pi}} \int_0^{\tilde{U}e^{-c}/2} e^{-\frac{x^2}{2}} \, dx$$

for any $c \in \mathbb{R}$. Therefore

$$\lim_{c \to -\infty} \liminf_{\epsilon \to 0} D^\epsilon(t_{\text{mix}}^\epsilon + cw^\epsilon) = 1$$

and

$$\lim_{c \to +\infty} \limsup_{\epsilon \to 0} D^\epsilon(t_{\text{mix}}^\epsilon + cw^\epsilon) = 0. \qquad \square$$

REMARK 2.8. Recall that $\{v_1, \ldots, v_m\}$ are linearly independent in $\mathbb{C}$. If in addition

$$\lim_{t \to +\infty} \left\| \Sigma^{-1/2} \sum_{k=1}^m e^{i\theta_k t} v_k \right\| \quad \text{is well defined,}$$

then

$$\lim_{t \to +\infty} \left\| \Sigma^{-1/2} \sum_{k=1}^m e^{i\theta_k t} v_k \right\| = \left\| \Sigma^{-1/2} \sum_{k=1}^m v_k \right\| > 0.$$

In this case, we define

$$r(x_0) := \left\| \Sigma^{-1/2} \sum_{k=1}^m v_k \right\| > 0.$$

COROLLARY 2.9 (Profile thermalisation).   *Suppose that $x_0 \neq 0$. There is profile thermalisation for the family of Markov processes $\{x^{\epsilon, x_0}\}_{\epsilon \in (0,1]}$ if and only if*

$$\lim_{t \to +\infty} \left\| \Sigma^{-1/2} \sum_{k=1}^{m} e^{i\theta_k t} v_k \right\| \quad \text{is well defined.}$$

PROOF.   Suppose that there is profile thermalisation for $\{x^{\epsilon, x_0}\}_{\epsilon \in (0,1]}$. Then $\lim_{\epsilon \to 0} d^{\epsilon}(t_{\text{mix}}^{\epsilon} + cw^{\epsilon})$ exists for any $c \in \mathbb{R}$. From Corollary 2.6 we have that $\lim_{\epsilon \to 0} D^{\epsilon}(t_{\text{mix}}^{\epsilon} + cw^{\epsilon})$ also exists for any $c \in \mathbb{R}$. From Remark 2.4 we deduce

$$\lim_{t \to +\infty} \left\| \Sigma^{-1/2} \sum_{k=1}^{m} e^{i\theta_k (t_{\text{mix}}^{\epsilon} + cw^{\epsilon})} v_k \right\| \quad \text{is well defined.}$$

Therefore,

$$\lim_{t \to +\infty} \left\| \Sigma^{-1/2} \sum_{k=1}^{m} e^{i\theta_k t} v_k \right\| \quad \text{is well defined.}$$

On the other hand, if $\lim_{t \to +\infty} \| \Sigma^{-1/2} \sum_{k=1}^{m} e^{i\theta_k t} v_k \| = r(x_0)$, from Remark 2.8 we have that $r(x_0) > 0$. From Remark 2.4, for any $c \in \mathbb{R}$ we deduce

$$\lim_{\epsilon \to 0} D^{\epsilon}(t_{\text{mix}}^{\epsilon} + cw^{\epsilon}) = \sqrt{\frac{2}{\pi}} \int_0^{\gamma(c)} e^{-\frac{x^2}{2}} \, dx,$$

where $\gamma(c) = (2\gamma)^{1-\ell} e^{-c} r(x_0)/2$. The latter together with Corollary 2.6 imply profile thermalisation for $\{x^{\epsilon, x_0}\}_{\epsilon \in (0,1]}$.   □

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

COROLLARY 2.10.   *Suppose that $x_0 \neq 0$. If all the eigenvalues of $DF(0)$ are real then the family of Markov processes $\{x^{\epsilon, x_0}\}_{\epsilon \in (0,1]}$ has profile thermalisation.*

PROOF.   The proof follows from Corollary 2.9 observing that $\theta_j = 0$ for any $j = 1, \ldots, m$ and the fact that $\{v_1, \ldots, v_m\}$ are linearly independent in $\mathbb{C}$.   □

Moreover, in [10], we study the case when $d = 1$ which follows immediately from Corollary 2.10.

We also have a *dynamical characterisation* of profile thermalisation. Define "a normalised" ω-limit set of x_0 as follows:

$$\omega(x_0) := \left\{ y \in \mathbb{R}^d : \text{there exists a sequence of positive numbers } \{t_n : n \in \mathbb{N}\} \right.$$

$$\left. \text{such that } \lim_{n \to +\infty} t_n = +\infty \text{ and } \lim_{n \to +\infty} \frac{e^{\lambda t_n}}{t_n^{\ell-1}} \Sigma^{-1/2} e^{-DF(0)t_n} x_0 = y \right\}.$$

From Lemma B.1, it is not hard to see that $\Sigma^{-1/2} \sum_{k=1}^{m} v_k \in \omega(x_0)$. When all the eigenvalues of $DF(0)$ are real, then again by Lemma B.1, we get that $\omega(x_0)$ consists of a nonzero element which is given by $\Sigma^{-1/2} \sum_{k=1}^{m} v_k$.

COROLLARY 2.11.   *Suppose that $x_0 \neq 0$. The family of Markov processes $\{x^{\epsilon, x_0}\}_{\epsilon \in (0,1]}$ has profile thermalisation if and only if $\omega(x_0)$ is contained in a d-sphere with radius $r(x_0) = \| \Sigma^{-1/2} \sum_{k=1}^{m} v_k \|$, that is, $\omega(x_0) \subset \mathbb{S}^{d-1}(r(x_0))$, where $\mathbb{S}^{d-1}(r(x_0)) := \{x \in \mathbb{R}^d : \|x\| = r(x_0)\}$.*

PROOF.   Suppose that $\{x^{\epsilon,x_0}\}_{\epsilon\in(0,1]}$ has profile thermalisation. By Corollary 2.9 we have

$$\lim_{t\to+\infty}\left\|\Sigma^{-1/2}\sum_{k=1}^{m}e^{i\theta_k t}v_k\right\| \quad \text{is well defined.}$$

From Remark 2.8 we know

$$\lim_{t\to+\infty}\left\|\Sigma^{-1/2}\sum_{k=1}^{m}e^{i\theta_k t}v_k\right\| = r(x_0) > 0.$$

The latter together with Lemma B.1 allows us to deduce that

$$\lim_{t\to+\infty}\left\|\frac{e^{\lambda t}}{t^{\ell-1}}\Sigma^{-1/2}e^{-DF(0)t}x_0\right\| = r(x_0).$$

Consequently, $\omega(x_0)\subset\mathbb{S}^{d-1}(r(x_0))$.

On the other hand, suppose that $\omega(x_0)\subset\mathbb{S}^{d-1}(r(x_0))$. Then

$$\lim_{t\to+\infty}\left\|\frac{e^{\lambda t}}{t^{\ell-1}}e^{-DF(0)t}x_0\right\| = r(x_0).$$

From Lemma B.1 we get

$$\lim_{t\to+\infty}\left\|\Sigma^{-1/2}\sum_{k=1}^{m}e^{i\theta_k t}v_k\right\| = r(x_0).$$

The latter together with Corollary 2.9 allow us to deduce the statement. $\square$

In dimension 2 and 3, we can state a *spectral characterisation* of profile thermalisation. Remind that if all the eigenvalues of $DF(0)$ are real, we have profile thermalisation as Corollary 2.10 stated, so we do not consider that case.

COROLLARY 2.12.   *Suppose that $x_0 \neq 0$ and $d = 2$. Let γ be a complex eigenvalue of $DF(0)$ with nonzero imaginary part and let $u_1 + iu_2$ be its eigenvector, where $u_1, u_2 \in \mathbb{R}^2$. Then the family of Markov processes $\{x^{\epsilon,x_0}\}_{\epsilon\in(0,1]}$ has profile thermalisation if and only if $\langle u_1, \Sigma^{-1}u_1\rangle = \langle u_2, \Sigma^{-1}u_2\rangle$ and $\langle u_1, \Sigma^{-1}u_2\rangle = 0$.*

PROOF.   Write $\gamma = \lambda + i\theta$, where $\lambda > 0$ with $\theta \neq 0$. To the eigenvalue γ we associated an eigenvector $u_1 + iu_2$, where $u_1, u_2 \in \mathbb{R}^2$. A straightforward computation shows

$$e^{\lambda t}e^{-DF(0)t}x_0 = \big(c_1\cos(\theta t) - c_2\sin(\theta t)\big)u_1 + \big(c_1\sin(\theta t) + c_2\cos(\theta t)\big)u_2$$

for any $t \geq 0$, where $c_1 := c_1(x_0)$ and $c_2 := c_2(x_0)$ are not both zero. Notice that $c := \sqrt{c_1^2 + c_2^2} > 0$ and let $\cos(\alpha) = c_1/c$ and $\sin(\alpha) = c_2/c$. Then

$$e^{\lambda t}e^{-DF(0)t}x_0 = c\cos(\theta t + \alpha)u_1 + c\sin(\theta t + \alpha)u_2 \quad \text{for any } t \geq 0.$$

Therefore,

(2.10)
$$\begin{aligned}
\big\|\Sigma^{-1/2}&e^{\lambda t}e^{-DF(0)t}x_0\big\|^2 \\
&= c^2\cos^2(\theta t + \alpha)\langle u_1, \Sigma^{-1}u_1\rangle \\
&\quad + c^2\sin^2(\theta t + \alpha)\langle u_2, \Sigma^{-1}u_2\rangle \\
&\quad + 2c^2\cos(\theta t + \alpha)\sin(\theta t + \alpha)\langle u_1, \Sigma^{-1}u_2\rangle
\end{aligned}$$

for any $t \geq 0$. If $\langle u_1, \Sigma^{-1} u_1 \rangle = \langle u_2, \Sigma^{-1} u_2 \rangle$ and $\langle u_1, \Sigma^{-1} u_2 \rangle = 0$ then

$$\| \Sigma^{-1/2} e^{\lambda t} e^{-DF(0)t} x_0 \|^2 = c^2 \langle u_1, \Sigma^{-1} u_1 \rangle \quad \text{for any } t \geq 0.$$

Notice that $\langle u_1, \Sigma^{-1} u_1 \rangle \neq 0$ since $u_1 \neq 0$ and Σ^{-1} is a positive definite symmetric matrix. The conclusion follows easily from Lemma B.1 and Corollary 2.9.

On the other hand, if $\{x^{\epsilon, x_0}\}_{\epsilon \in (0,1]}$ has profile thermalisation then Lemma B.1 and Corollary 2.9 imply

$$\lim_{t \to +\infty} \| \Sigma^{-1/2} e^{\lambda t} e^{-DF(0)t} x_0 \|^2 \quad \text{is well defined.}$$

Now, using (2.10) and taking different subsequences we deduce $\langle u_1, \Sigma^{-1} u_1 \rangle = \langle u_2, \Sigma^{-1} u_2 \rangle$ and $\langle u_1, \Sigma^{-1} u_2 \rangle$. $\square$

Notice that in dimension 3, at least one eigenvalue of $DF(0)$ is real. Therefore the interesting case is when the others eigenvalues are complex numbers with nonzero imaginary part. Let γ_1 be a real eigenvalue of $DF(0)$ with eigenvector $v \in \mathbb{R}^3$. Let γ be a complex eigenvalue of $DF(0)$ with nonzero imaginary part and let $u_1 + i u_2$ be its eigenvector, where $u_1, u_2 \in \mathbb{R}^3$. In this case,

$$e^{-DF(0)t} x_0 = c(x_0) e^{-\gamma_1 t} v$$
$$+ e^{-\lambda t} \big(c_1 \cos(\theta t) - c_2 \sin(\theta t) \big) u_1 + \big(c_1 \sin(\theta t) + c_2 \cos(\theta t) u_2 \big)$$

for any $t \geq 0$, where $c_0 := c_0(x_0)$, $c_1 := c_1(x_0)$ and $c_2 := c_2(x_0)$ are not all zero. Notice that $c := \sqrt{c_1^2 + c_2^2} > 0$ and let $\cos(\alpha) = c_1/c$ and $\sin(\alpha) = c_2/c$. Then

$$(2.11) \qquad e^{-DF(0)t} x_0 = c_0 e^{-\gamma_1 t} v + e^{-\lambda t} \big(c \cos(\theta t + \alpha) u_1 + c \sin(\theta t + \alpha) u_2 \big)$$

for any $t \geq 0$.

COROLLARY 2.13.   *Suppose that $x_0 \neq 0$ and $d = 3$. Let γ_1 be a real eigenvalue of $DF(0)$ with eigenvector $v \in \mathbb{R}^3$. Let γ be a complex eigenvalue of $DF(0)$ with nonzero imaginary part and let $u_1 + i u_2$ be its eigenvector, where $u_1, u_2 \in \mathbb{R}^3$. Let c_0, c and α the constants that appears in (2.11).*

(i) *Assume $c_0 = 0$. $\{x^{\epsilon, x_0}\}_{\epsilon \in (0,1]}$ has profile thermalisation if and only if $\langle u_1, \Sigma^{-1} u_1 \rangle = \langle u_2, \Sigma^{-1} u_2 \rangle$ and $\langle u_1, \Sigma^{-1} u_2 \rangle = 0$.*

(ii) *Assume $c_0 \neq 0$ and $\gamma_1 < \lambda$. Then the family $\{x^{\epsilon, x_0}\}_{\epsilon \in (0,1]}$ has profile thermalisation.*

(iii) *Assume $c_0 \neq 0$, $\gamma_1 = \lambda$. The family $\{x^{\epsilon, x_0}\}_{\epsilon \in (0,1]}$ has profile thermalisation if and only if $\langle u_1, \Sigma^{-1} u_1 \rangle = \langle u_2, \Sigma^{-1} u_2 \rangle$ and $\langle u_1, \Sigma^{-1} u_2 \rangle = \langle v, \Sigma^{-1} u_1 \rangle = \langle v, \Sigma^{-1} u_2 \rangle = 0$.*

(iv) *Assume $c_0 \neq 0$, $\gamma_1 > \lambda$. The family $\{x^{\epsilon, x_0}\}_{\epsilon \in (0,1]}$ has profile thermalisation if and only if $\langle u_1, \Sigma^{-1} u_1 \rangle = \langle u_2, \Sigma^{-1} u_2 \rangle$ and $\langle u_1, \Sigma^{-1} u_2 \rangle = 0$.*

PROOF.   (i) This case can be deduced using the same arguments as Corollary 2.12.
(ii) Using relation (2.11) we obtain

$$\lim_{t \to +\infty} \Sigma^{-1/2} e^{\gamma_1 t} e^{-DF(0)t} x_0 = c_0 \Sigma^{-1/2} v \neq 0.$$

The latter together with Corollary 2.9 and Lemma B.1 allow us to deduce profile thermalisation.

(iii) Using relation (2.11) for any $t \geq 0$ we get

$$
\begin{aligned}
\big\| \Sigma^{-1/2} & e^{\gamma_1 t} e^{-DF(0)t} x_0 \big\|^2 \\
&= c_0^2 \langle v, \Sigma^{-1/2} v \rangle + c^2 \cos^2(\theta t + \alpha) \langle u_1, \Sigma^{-1/2} u_1 \rangle \\
&\quad + c^2 \sin^2(\theta t + \alpha) \langle u_2, \Sigma^{-1/2} u_2 \rangle + 2 c_0 c \cos(\theta t + \alpha) \langle v, \Sigma^{-1/2} u_1 \rangle \\
&\quad + 2 c_0 c \sin(\theta t + \alpha) \langle v, \Sigma^{-1/2} u_2 \rangle \\
&\quad + 2 c^2 \cos(\theta t + \alpha) \sin(\theta t + \alpha) \langle u_1, \Sigma^{-1/2} u_2 \rangle.
\end{aligned}
$$

(2.12)

If $\langle u_1, \Sigma^{-1} u_1 \rangle = \langle u_2, \Sigma^{-1} u_2 \rangle$ and

$$
\langle u_1, \Sigma^{-1} u_2 \rangle = \langle v, \Sigma^{-1} u_1 \rangle = \langle v, \Sigma^{-1} u_2 \rangle = 0,
$$

then from (2.12) we deduce

$$
\big\| \Sigma^{-1/2} e^{\gamma_1 t} e^{-DF(0)t} x_0 \big\| = c_0^2 \langle v, \Sigma^{-1/2} v \rangle + c^2 \langle u_1, \Sigma^{-1} u_1 \rangle > 0.
$$

The latter together with Lemma B.1 and Corollary 2.9 imply profile thermalisation.

On the other hand, if $\{x^{\epsilon, x_0}\}_{\epsilon \in (0,1]}$ has profile thermalisation then Lemma B.1 and Corollary 2.9 imply

$$
\lim_{t \to +\infty} \big\| \Sigma^{-1/2} e^{\lambda t} e^{-DF(0)t} x_0 \big\|^2 \quad \text{is well defined.}
$$

Now, using (2.12) and taking different subsequences we deduce

$$
\langle u_1, \Sigma^{-1} u_1 \rangle = \langle u_2, \Sigma^{-1} u_2 \rangle
$$

and

$$
\langle u_1, \Sigma^{-1} u_2 \rangle = \langle v, \Sigma^{-1} u_1 \rangle = \langle v, \Sigma^{-1} u_2 \rangle = 0.
$$

(iv) Using relation (2.11) for any $t \geq 0$ we have

$$
\begin{aligned}
\big\| \Sigma^{-1/2} & e^{\lambda t} e^{-DF(0)t} x_0 \big\|^2 \\
&= c_0^2 e^{-2(\gamma_1 - \lambda)t} \langle v, \Sigma^{-1/2} v \rangle + c^2 \cos^2(\theta t + \alpha) \langle u_1, \Sigma^{-1/2} u_1 \rangle \\
&\quad + c^2 \sin^2(\theta t + \alpha) \langle u_2, \Sigma^{-1/2} u_2 \rangle \\
&\quad + 2 c_0 c e^{-(\gamma_1 - \lambda)t} \cos(\theta t + \alpha) \langle v, \Sigma^{-1/2} u_1 \rangle \\
&\quad + 2 c_0 c e^{-(\gamma_1 - \lambda)t} \sin(\theta t + \alpha) \langle v, \Sigma^{-1/2} u_2 \rangle \\
&\quad + 2 c^2 \cos(\theta t + \alpha) \sin(\theta t + \alpha) \langle u_1, \Sigma^{-1/2} u_2 \rangle.
\end{aligned}
$$

(2.13)

If $\langle u_1, \Sigma^{-1} u_1 \rangle = \langle u_2, \Sigma^{-1} u_2 \rangle$ and $\langle u_1, \Sigma^{-1} u_2 \rangle = 0$ from (2.13) we obtain

$$
\lim_{t \to +\infty} \big\| \Sigma^{-1/2} e^{\lambda t} e^{-DF(0)t} x_0 \big\| = c^2 \langle u_1, \Sigma^{-1} u_1 \rangle \neq 0
$$

which together with Corollary 2.9 imply profile thermalisation.

On the other hand, if $\{x^{\epsilon, x_0}\}_{\epsilon \in (0,1]}$ has profile thermalisation then Lemma B.1 and Corollary 2.9 imply

$$
\lim_{t \to +\infty} \big\| \Sigma^{-1/2} e^{\lambda t} e^{-DF(0)t} x_0 \big\|^2 \quad \text{is well defined.}
$$

Now, using (2.13) and taking different subsequences one can deduce

$$
\langle u_1, \Sigma^{-1} u_1 \rangle = \langle u_2, \Sigma^{-1} u_2 \rangle \quad \text{and} \quad \langle u_1, \Sigma^{-1} u_2 \rangle = 0. \qquad \square
$$

When Σ is the identity matrix, roughly speaking Corollary 2.12 and Corollary 2.13 state that profile thermalisation is equivalent to "norm" preserving and orthogonality of the real and imaginary parts of the eigenvectors of $DF(0)$. When Σ is not the identity, the latter is still true under a change of basis.

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 a linear nonhomogeneous process in a time window that will include the time scale on which we are interested. It is not hard to see that (H) basically says that (C) 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

$$\langle y - x, F(y) - F(x) \rangle \geq \delta \|y - x\|^2 \quad \text{for any } x, y \in \mathbb{R}^d.$$

The latter is basically saying that (C) is satisfied around any point $y \in \mathbb{R}^d$. A good example of a vector field F satisfying (H) and (G) is $F(x) = Ax + H(x)$, $x \in \mathbb{R}^d$, where A is a matrix, H is a vector valued function such that F satisfies (H) and it satisfies $H(0) = 0$, $DH(0) = 0$, $\|DH\|_\infty < +\infty$ and $\|D^2 H\|_\infty < +\infty$. In dimension one, a good example to keep in mind is

$$(3.1) \qquad\qquad F(x) = \sum_{j=1}^n a_j x^{2j-1} \quad \text{for any } x \in \mathbb{R},$$

where $n \in \mathbb{N}$, $a_1 > 0$ and $a_j \geq 0$ for any $j \in \{2, \ldots, n\}$. It is fairly easy to see that (3.1) satisfies (H) and (G). If $a_j > 0$ for some $j \in \{2, \ldots, n\}$ then (3.1) is not globally Lipschitz continuous.

Recall that $F \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}^d)$. Notice that for any $x, y \in \mathbb{R}^d$ we have

$$F(x) - F(y) = \int_0^1 DF(x + t(y - x))(y - x) \, dt.$$

Therefore, for any $x, y \in \mathbb{R}^d$ we get

$$F(x) - F(y) - DF(y)(x - y) = \int_0^1 (DF(y + t(x - y)) - DF(y))(x - y) \, dt.$$

Note

$$DF(y + t(x - y)) - DF(y) = \int_0^1 D^2 F(y + st(x - y))t(x - y) \, ds$$

for any $x, y \in \mathbb{R}^d$. For any $r_0 > 0$ and $r_1 > 0$, define $C := \sup_{|z| \leq 2r_1 + r_0} \|D^2 F(z)\|$. Then

$$(3.2) \qquad\qquad \|F(x) - F(y) - DF(y)(x - y)\| \leq C \|x - y\|^2$$

for any $\|x\| \leq r_0$ and $\|y\| \leq r_1$. Inequality (3.2) will allow us to control the random dynamics $\{x^\epsilon(t) : t \geq 0\}$ on compacts sets and it will be very useful in our a priori estimates.

3.1. *Zeroth-order approximations.* It is fairly easy to see that for any $t \geq 0$, as $\epsilon \to 0$, $x^\epsilon(t)$ converges to $\varphi(t)$. 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 $\varphi(t)$. The idea is fairly simple: (H) 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 $\varphi(t)$ 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 $\|x^\epsilon(t) - \varphi(t)\|^2$ as follows:

$$
\begin{aligned}
\mathrm{d}\|x^\epsilon(t) - \varphi(t)\|^2 &= -2\langle x^\epsilon(t) - \varphi(t), F(x^\epsilon(t)) - F(\varphi(t))\rangle\,\mathrm{d}t \\
&\quad + 2\sqrt{\epsilon}\langle x^\epsilon(t) - \varphi(t), \mathrm{d}B(t)\rangle + d\epsilon\,\mathrm{d}t \\
&\leq -2\delta\|x^\epsilon(t) - \varphi(t)\|^2\,\mathrm{d}t \\
&\quad + 2\sqrt{\epsilon}\langle x^\epsilon(t) - \varphi(t), \mathrm{d}B(t)\rangle + d\epsilon\,\mathrm{d}t,
\end{aligned}
$$

(3.3)

where the last inequality follows from (H). After a localisation argument we get

$$
\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big[\|x^\epsilon(t) - \varphi(t)\|^2\big] \leq -2\delta\mathbb{E}\big[\|x^\epsilon(t) - \varphi(t)\|^2\big] + \epsilon d \quad \text{for any } t \geq 0.
$$

From Lemma C.7, we obtain the following uniform bound

(3.4) $$ \mathbb{E}\big[\|x^\epsilon(t) - \varphi(t)\|^2\big] \leq \frac{d\epsilon}{2\delta}(1 - e^{-2\delta t}) \leq \frac{d\epsilon}{2\delta} \quad \text{for any } t \geq 0. $$

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

3.2. *First-order approximations.*

Notice that (3.4) can be seen as a law of large numbers for $x^\epsilon(t)$. Therefore, it is natural to look at the corresponding central limit theorem. Define $\{y^\epsilon(t) : t \geq 0\}$ as

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

As above, it is not very difficult to prove that for every $T > 0$, the process $\{y^\epsilon(t) : t \in [0, T]\}$ converges in distribution to the solution $\{y(t) : t \in [0, T]\}$ of the linear nonhomogeneous stochastic differential equation (also known as nonhomogeneous Ornstein–Uhlenbeck process in stochastic analysis):

(3.5) $$ \begin{cases} \mathrm{d}y(t) = -DF(\varphi(t))y(t)\,\mathrm{d}t + \mathrm{d}B(t) & \text{for } t \geq 0, \\ y(0) = 0. \end{cases} $$

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

(3.6) $$ \mathbb{E}\big[\|y^\epsilon(t)\|^2\big] \leq \frac{d}{2\delta} \quad \text{for any } t \geq 0. $$

We will also need an upper bound for $\mathbb{E}[\|y^\epsilon(t)\|^4]$. From the Itô formula and (H) we have

$$
\begin{aligned}
\mathrm{d}\|y^\epsilon(t)\|^4 &= -4\|y^\epsilon(t)\|^2\langle y^\epsilon(t), DF(\varphi(t))y^\epsilon(t)\rangle\,\mathrm{d}t + 4\|y^\epsilon(t)\|^2\langle y^\epsilon(t), \mathrm{d}B(t)\rangle \\
&\quad + (2d + 4)\|y^\epsilon(t)\|^2\,\mathrm{d}t \\
&\leq -4\delta\|y^\epsilon(t)\|^4\,\mathrm{d}t + 4\|y^\epsilon(t)\|^2\langle y^\epsilon(t), \mathrm{d}B(t)\rangle + (2d + 4)\|y^\epsilon(t)\|^2\,\mathrm{d}t.
\end{aligned}
$$

After a localisation argument we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big[\|y^\epsilon(t)\|^4\big] \le -4\delta\mathbb{E}\big[\|y^\epsilon(t)\|^4\big] + (2d+4)\mathbb{E}\big[\|y^\epsilon(t)\|^2\big].$$

From (3.6) and Lemma C.7 we get the uniformly bound

$$(3.7) \qquad \mathbb{E}\big[\|y^\epsilon(t)\|^4\big] \le \frac{d(d+2)}{4\delta^2}(1 - e^{-4\delta t}) \le \frac{d(d+2)}{4\delta^2} \quad \text{for any } t \ge 0.$$

Notice that $x^\epsilon(t) = \varphi(t) + \sqrt{\epsilon}y^\epsilon(t)$ for any $t \ge 0$ and the difference $y^\epsilon(t) - y(t)$ has bounded variation. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}(y^\epsilon(t) - y(t)) = -\frac{1}{\sqrt{\epsilon}}\big(F(x^\epsilon(t)) - F(\varphi(t)) - \sqrt{\epsilon}DF(\varphi(t))y(t)\big)$$

$$= -\frac{1}{\sqrt{\epsilon}}\big(F(\varphi(t) + \sqrt{\epsilon}y^\epsilon(t)) - F(\varphi(t) + \sqrt{\epsilon}y(t))\big)$$

$$\quad - \frac{1}{\sqrt{\epsilon}}\big(F(\varphi(t) + \sqrt{\epsilon}y(t)) - F(\varphi(t)) - \sqrt{\epsilon}DF(\varphi(t))y(t)\big).$$

Define $h^\epsilon(t) := F(\varphi(t) + \sqrt{\epsilon}y(t)) - F(\varphi(t)) - \sqrt{\epsilon}DF(\varphi(t))y(t)$ for any $t \ge 0$. Therefore, using the chain rule for $\|y^\epsilon(t) - y(t)\|^2$ we obtain the differential equation

$$(3.8) \qquad \begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\|y^\epsilon(t) - y(t)\|^2 \\ &= 2\Big\langle y^\epsilon(t) - y(t), \frac{\mathrm{d}}{\mathrm{d}t}(y^\epsilon(t) - y(t))\Big\rangle \\ &= -\frac{2}{\sqrt{\epsilon}}\langle y^\epsilon(t) - y(t), F(\varphi(t) + \sqrt{\epsilon}y^\epsilon(t)) - F(\varphi(t) + \sqrt{\epsilon}y(t))\rangle \\ &\quad - \frac{2}{\sqrt{\epsilon}}\langle y^\epsilon(t) - y(t), h^\epsilon(t)\rangle \\ &\le -2\delta\|y^\epsilon(t) - y(t)\|^2 - \frac{2}{\sqrt{\epsilon}}\langle y^\epsilon(t) - y(t), h^\epsilon(t)\rangle, \end{aligned}$$

where the last inequality follows from (H). From the Cauchy–Schwarz inequality we observe

$$(3.9) \qquad \big|(2/\sqrt{\epsilon})\langle y^\epsilon(t) - y(t), h^\epsilon(t)\rangle\big| \le (2/\sqrt{\epsilon})\|y^\epsilon(t) - y(t)\|\|h^\epsilon(t)\|.$$

Recall the well-known Young-type inequality $2|ab| \le \varrho a^2 + (1/\varrho)b^2$ for any $a, b \in \mathbb{R}$ and $\rho > 0$. From inequality (3.9) we have

$$(3.10) \qquad \big|(2/\sqrt{\epsilon})\langle y^\epsilon(t) - y(t), h^\epsilon(t)\rangle\big| \le \delta\|y^\epsilon(t) - y(t)\|^2 + \frac{1}{\epsilon\delta}\|h^\epsilon(t)\|^2.$$

From inequality (3.8) and inequality (3.10) we deduce

$$\frac{\mathrm{d}}{\mathrm{d}t}\|y^\epsilon(t) - y(t)\|^2 \le -\delta\|y^\epsilon(t) - y(t)\|^2 + \frac{1}{\epsilon\delta}\|h^\epsilon(t)\|^2.$$

By taking expectation in both sides of the last inequality, we obtain

$$(3.11) \qquad \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big[\|y^\epsilon(t) - y(t)\|^2\big] \le -\delta\mathbb{E}\big[\|y^\epsilon(t) - y(t)\|^2\big] + \frac{1}{\epsilon\delta}\mathbb{E}\big[\|h^\epsilon(t)\|^2\big].$$

Define $H^\epsilon(t) := \mathbb{E}[\|h^\epsilon(t)\|^2]$ for any $t \ge 0$. From inequality (3.11) and Lemma C.7 we deduce

$$\mathbb{E}\big[\|y^\epsilon(t) - y(t)\|^2\big] \le \frac{(1 - e^{-\delta t})}{\epsilon\delta^2}\int_0^t H^\epsilon(s)\,\mathrm{d}s \quad \text{for any } t \ge 0.$$

Therefore, we need to get an upper bound for $\int_0^t H^\epsilon(s)\,\mathrm{d}s$ for any $t \geq 0$. From Lemma 3.1 we have

$$\int_0^t H^\epsilon(s)\,\mathrm{d}s \leq C(\eta, \|x_0\|, d, \delta)\epsilon^2 t + \tilde{C}(\eta, \|x_0\|, d, \delta)\epsilon^{3/2} t^{7/4}$$

for any $\eta > 0$ and $t \in [0, \frac{\eta^2}{2\epsilon d})$, where $C(\eta, \|x_0\|, d, \delta)$ and $\tilde{C}(\eta, \|x_0\|, d, \delta)$ are positive constants that only depend on η, $\|x_0\|$, d and δ. The latter implies

(3.12)
$$\mathbb{E}\big[\|x^\epsilon(t) - (\varphi(t) + \sqrt{\epsilon}y(t))\|^2\big]$$
$$\leq \frac{1}{\delta^2}\big(C(\eta, \|x_0\|, d, \delta)\epsilon^2 t + \tilde{C}(\eta, \|x_0\|, d, \delta)\epsilon^{3/2} t^{7/4}\big)$$

for any $t \in [0, \frac{\eta^2}{2\epsilon d})$. We call this bound the *first-order* approximation of $x^\epsilon(t)$. Roughly speaking, for $t = \mathcal{O}(\ln(1/\epsilon))$ we have just proved that the distance between $x^\epsilon(t)$ and $\varphi(t) + \sqrt{\epsilon}y(t)$ is of order $\mathcal{O}(\epsilon^{3/4-\wp})$ for any $\wp \in (0, 3/4)$ which will be enough for our purposes.

LEMMA 3.1. *Assume that* (H) *and* (G) *hold. Let* $\epsilon \in (0, \frac{\delta}{32c_1})$. *For any* $\eta > 0$ *and* $t \in [0, \frac{\eta^2}{2\epsilon d})$ *we have*

$$\int_0^t H^\epsilon(s)\,\mathrm{d}s \leq C(\eta, \|x_0\|, d, \delta)\epsilon^2 t + \tilde{C}(\eta, \|x_0\|, d, \delta)\epsilon^{3/2} t^{7/4},$$

where $C(\eta, \|x_0\|, d, \delta)$ *and* $\tilde{C}(\eta, \|x_0\|, d, \delta)$ *only depend on* η, $\|x_0\|$, d *and* δ. *Moreover, for any* $\wp \in (0, 6/7)$ *we have*

$$\lim_{\epsilon \to 0} \sup_{0 \leq t \leq \mathcal{O}(1/\epsilon^\wp)} \mathbb{E}\big[\|x^\epsilon(t) - (\varphi(t) + \sqrt{\epsilon}y(t))\|^2\big] = 0.$$

PROOF. Recall that $H^\epsilon(t) = \mathbb{E}[\|h^\epsilon(t)\|^2]$, where

$$h^\epsilon(t) = F\big(\varphi(t) + \sqrt{\epsilon}y(t)\big) - F\big(\varphi(t)\big) - \sqrt{\epsilon}DF\big(\varphi(t)\big)y(t) \quad \text{for any } t \geq 0.$$

Take any $\eta > 0$ and $t > 0$ and define the event

$$A_{\eta,\epsilon,t} = \left[\sup_{0 \leq s \leq t} \|y(s)\| \leq \frac{\eta}{\sqrt{\epsilon}}\right].$$

Denote by $A_{\eta,\epsilon,t}^c$ its complement. By inequality (3.2) we have

$$\mathbb{E}\big[\|h^\epsilon(t)\|^2 \mathbb{1}_{A_{\eta,\epsilon,t}}\big] \leq C_0^2(\eta, \|x_0\|)\epsilon^2 \mathbb{E}\big[\|y(t)\|^4\big]$$

for any $t \geq 0$, where the positive constant $C_0(\eta, \|x_0\|)$ depends on η and $\|x_0\|$. By a similar argument using in inequality (3.7) we deduce $\mathbb{E}[\|y(t)\|^4] \leq \frac{d(d+2)}{4\delta^2}$ for any $t \geq 0$. Then

$$\mathbb{E}\big[\|h^\epsilon(t)\|^2 \mathbb{1}_{A_{\eta,\epsilon,t}}\big] \leq C_0^2(\eta, \|x_0\|)\frac{d(d+2)}{4\delta^2}\epsilon^2 \quad \text{for any } t \geq 0.$$

On the other hand, recall the well-known inequality

$$(x + y + z)^2 \leq 4(x^2 + y^2 + z^2) \quad \text{for any } x, y, z \in \mathbb{R}.$$

Then

$$\mathbb{E}\big[\|h^\epsilon(t)\|^2 \mathbb{1}_{A_{\eta,\epsilon,t}^c}\big] \leq 4\mathbb{E}\big[\|F(\varphi(t) + \sqrt{\epsilon}y(t))\|^2 \mathbb{1}_{A_{\eta,\epsilon,t}^c}\big]$$
$$+ 4\mathbb{E}\big[\|F(\varphi(t))\|^2 \mathbb{1}_{A_{\eta,\epsilon,t}^c}\big] + 4\epsilon\mathbb{E}\big[\|DF(\varphi(t))y(t)\|^2 \mathbb{1}_{A_{\eta,\epsilon,t}^c}\big]$$

for any $t \geq 0$. We will analyse the upper bound of the last inequality. Since $\|\varphi(s)\| \leq \|x_0\|$ for any $s \geq 0$, then

$$\mathbb{E}\big[\|F(\varphi(t))\|^2 \mathbb{1}_{A^c_{\eta,\epsilon,t}}\big] \leq C_1^2(\|x_0\|)\mathbb{P}(A^c_{\eta,\epsilon,t}),$$

for any $t \geq 0$, where $C_1(\|x_0\|)$ is a positive constant that only depends on $\|x_0\|$. We also observe that

$$\mathbb{E}\big[\|DF(\varphi(t))y(t)\|^2 \mathbb{1}_{A^c_{\eta,\epsilon,t}}\big] \leq C_2^2(\|x_0\|)\mathbb{E}\big[\|y(t)\|^2 \mathbb{1}_{A^c_{\eta,\epsilon,t}}\big]$$

for any $t \geq 0$, where $C_2(\|x_0\|)$ is a positive constant that only depends on $\|x_0\|$. From the Cauchy–Schwarz inequality we get

$$\begin{aligned}
\mathbb{E}\big[\|y(t)\|^2 \mathbb{1}_{A^c_{\eta,\epsilon,t}}\big] &= \mathbb{E}\big[\big(\|y(t)\|^2 \mathbb{1}_{A^c_{\eta,\epsilon,t}}\big)\mathbb{1}_{A^c_{\eta,\epsilon,t}}\big] \\
&\leq \big(\mathbb{E}\big[\big(\|y(t)\|^4 \mathbb{1}_{A^c_{\eta,\epsilon,t}}\big)\big]\big)^{1/2}\big(\mathbb{P}(A^c_{\eta,\epsilon,t})\big)^{1/2} \\
&\leq \big(\mathbb{E}\big[\big(\|y(t)\|^8\big)\big]\big)^{1/4}\big(\mathbb{P}(A^c_{\eta,\epsilon,t})\big)^{3/4},
\end{aligned}$$

for any $t \geq 0$. Following similar computations as we did in (3.7) or by item (ii) of Proposition C.2 we deduce

$$\mathbb{E}\big[\big(\|y(t)\|^8\big)\big] \leq \frac{d(d+2)(d+4)(d+6)}{16\delta^4} \quad \text{for any } t \geq 0.$$

Therefore

$$\mathbb{E}\big[\|DF(\varphi(t))y(t)\|^2 \mathbb{1}_{A^c_{\eta,\epsilon,t}}\big] \leq C_3^2(\|x_0\|, \delta, d)\big(\mathbb{P}(A^c_{\eta,\epsilon,t})\big)^{3/4}$$

for any $t \geq 0$, where $C_3(\|x_0\|, \delta, d)$ is a positive constant that only depends on $\|x_0\|$, δ and d. Finally, we analise $\mathbb{E}[\|F(\varphi(t) + \sqrt{\epsilon}y(t))\|^2 \mathbb{1}_{A^c_{\eta,\epsilon,t}}]$. From (G) we have

$$\mathbb{E}\big[\|F(\varphi(t) + \sqrt{\epsilon}y(t))\|^2 \mathbb{1}_{A^c_{\eta,\epsilon,t}}\big] \leq c_0^2 e^{4c_1\|x_0\|^2}\mathbb{E}\big[e^{4c_1\epsilon\|y(t)\|^2}\mathbb{1}_{A^c_{\eta,\epsilon,t}}\big]$$

for any $t \geq 0$. Again, from the Cauchy-Schwarz inequality we deduce

$$\begin{aligned}
\mathbb{E}\big[\big(e^{4c_1\epsilon\|y(t)\|^2}\mathbb{1}_{A^c_{\eta,\epsilon,t}}\big)\mathbb{1}_{A^c_{\eta,\epsilon,t}}\big] &\leq \big(\mathbb{E}\big[e^{8c_1\epsilon\|y(t)\|^2}\mathbb{1}_{A^c_{\eta,\epsilon,t}}\big]\big)^{1/2}\big(\mathbb{P}(A^c_{\eta,\epsilon,t})\big)^{1/2} \\
&\leq \big(\mathbb{E}\big[e^{16c_1\epsilon\|y(t)\|^2}\big]\big)^{1/4}\big(\mathbb{P}(A^c_{\eta,\epsilon,t})\big)^{3/4}
\end{aligned}$$

for any $t \geq 0$. From item (iv) of Proposition C.2, for any $\epsilon \in (0, \frac{\delta}{32c_1})$ we have

$$\mathbb{E}\big[e^{16c_1\epsilon\|y(t)\|^2}\big] \leq e^{16c_1 d\epsilon t} \quad \text{for any } t \geq 0.$$

Therefore,

$$\begin{aligned}
\mathbb{E}\big[\|h^\epsilon(t)\|^2\big] &\leq C_4^2(\eta, \|x_0\|, d, \delta)\epsilon^2 + 4C_1^2(\|x_0\|)\mathbb{P}(A^c_{\eta,\epsilon,t}) \\
&\quad + 4\big(C_3^2(\|x_0\|, \delta, d) + C_5(\|x_0\|)e^{4c_1 d\epsilon t}\big)\big(\mathbb{P}(A^c_{\eta,\epsilon,t})\big)^{3/4} \\
&\leq C_4^2(\eta, \|x_0\|, d, \delta)\epsilon^2 + 4\big(C_1^2(\|x_0\|) \\
&\quad + C_3^2(\|x_0\|, \delta, d) + C_5(\|x_0\|)e^{4c_1 d\epsilon t}\big)\big(\mathbb{P}(A^c_{\eta,\epsilon,t})\big)^{3/4},
\end{aligned}$$

where $C_4^2(\eta, \|x_0\|, d, \delta) = C_0^2(\eta, \|x_0\|)\frac{d(d+2)}{4\delta^2}$ and $C_5(\|x_0\|) = c_0^2 e^{4c_1\|x_0\|^2}$. From item (ii) of Lemma C.1 we have

$$\mathbb{P}(A^c_{\eta,\epsilon,t}) \leq \frac{2d\epsilon^2 t}{\delta(\eta^2 - \epsilon dt)^2} \quad \text{for any } 0 \leq t < \frac{\eta^2}{\epsilon d}.$$

Notice that

$$\mathbb{P}(A_{\eta,\epsilon,t}^c) \leq \frac{4d\epsilon^2 t}{\delta\eta^2} \quad \text{for any } 0 \leq t < \frac{\eta^2}{2\epsilon d}.$$

Consequently,

$$\mathbb{E}[\|h^\epsilon(t)\|^2] \leq C_4^2(\eta, \|x_0\|, d, \delta)\epsilon^2$$

$$+ 4(C_1^2(\|x_0\|) + C_3^2(\|x_0\|, \delta, d) + C_5(\|x_0\|)e^{2c_1\eta^2})\left(\frac{4d\epsilon^2 t}{\delta\eta^2}\right)^{3/4}$$

for any $0 \leq t < \frac{\eta^2}{2\epsilon d}$. By integration we deduce the first part of the statement. The second part follows immediately from inequality (3.12). □

In Lemma C.6, we will prove that the linear nonhomogeneous process $\{y(t) : t \geq 0\}$ has a limiting, nondegenerate law which is Gaussian with mean vector zero and covariance matrix Σ which is the unique solution of the Lyapunov matrix equation (2.9).

3.3. *An $\epsilon/3$ proof.* We approximate the process $\{x^\epsilon(t) : t \geq 0\}$ by a linear nonhomogeneous process $\{z^\epsilon(t) := \varphi(t) + \sqrt{\epsilon}y(t) : t \geq 0\}$ in which we can carry out "explicit" computations. This process is also known in Physics as Van Kampen's approximation, see [4] for more details. Since we need to compare solutions of various stochastic differential equations with different initial conditions, we will introduce some notation. Let ξ be a random variable in $\mathbb{R}^d$ and let $T > 0$. Let $\{\varphi(t, \xi) : t \geq 0\}$ denotes the solution of

$$\begin{cases} d\varphi(t, \xi) = -F(\varphi(t, \xi)) \, dt & \text{for any } t \geq 0, \\ \varphi(0, \xi) = \xi. \end{cases}$$

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

$$\begin{cases} dy(t, \xi, T) = -DF(\varphi(t, \xi))y(t, \xi, T) \, dt + dB(t + T) & \text{for any } t \geq 0, \\ y(0, \xi, T) = 0 \end{cases}$$

and define $\{z^\epsilon(t, \xi, T) : t \geq 0\}$ as $z^\epsilon(t, \xi, T) := \varphi(t, \xi) + \sqrt{\epsilon}y(t, \xi, T)$ for any $t \geq 0$. Take $c \in \mathbb{R}$ and let $\delta_\epsilon > 0$ such that $\delta_\epsilon = o(1)$. For $\epsilon \ll 1$ define

$$(3.13) \qquad t_{\text{shift}}^\epsilon := t_{\text{mix}}^\epsilon - \delta_\epsilon > 0.$$

In what follows, we will always take $T = t_{\text{shift}}^\epsilon + cw^\epsilon > 0$ for every $\epsilon > 0$ small enough, so for simplicity, we will omit it from the notation.

The following lemma is the key of the proof. Roughly speaking, from Lemma 3.1 we see that the processes $\{x^\epsilon(t) : t \geq 0\}$ and $\{z^\epsilon(t) : t \geq 0\}$ are close enough for times of order $\mathcal{O}(\ln(1/\epsilon))$. Therefore, we can shift the processes for a small time δ_ϵ and then we coupled the remainder differences in a small time interval $[0, \delta_\epsilon]$. Since $\{z^\epsilon(t) : t \geq 0\}$ is linear, then thermalisation and/or profile thermalisation will be concluded from it.

LEMMA 3.2. *For any $c \in \mathbb{R}$ and $\epsilon \ll 1$ we have*

$$\begin{aligned} \big| &d_{\text{TV}}\big(x^\epsilon(t_{\text{mix}}^\epsilon + cw^\epsilon, x_0), \mu^\epsilon\big) \\ &- d_{\text{TV}}\big(z^\epsilon(\delta_\epsilon, z^\epsilon(t_{\text{shift}}^\epsilon + cw^\epsilon, x_0)), \mathcal{G}(0, \epsilon\Sigma))\big| \\ (3.14) \qquad &\leq d_{\text{TV}}\big(x^\epsilon(\delta_\epsilon, x^\epsilon(t_{\text{shift}}^\epsilon + cw^\epsilon, x_0)), z^\epsilon(\delta_\epsilon, x^\epsilon(t_{\text{shift}}^\epsilon + cw^\epsilon, x_0))\big) \\ &+ d_{\text{TV}}\big(z^\epsilon(\delta_\epsilon, x^\epsilon(t_{\text{shift}}^\epsilon + cw^\epsilon, x_0)), z^\epsilon(\delta_\epsilon, z^\epsilon(t_{\text{shift}}^\epsilon + cw^\epsilon, x_0))\big) \\ &+ d_{\text{TV}}\big(\mathcal{G}(0, \epsilon\Sigma), \mu^\epsilon\big). \end{aligned}$$

PROOF.   Notice that

$$\begin{aligned}
&d_{TV}\big(x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon + \delta_\epsilon, x_0\big), \mu^\epsilon\big) \\
&= d_{TV}\big(x^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big), \mu^\epsilon\big) \\
&\leq d_{TV}\big(x^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big), z^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big)\big) \\
&\quad + d_{TV}\big(z^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big), z^\epsilon\big(\delta_\epsilon, z^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big)\big) \\
&\quad + d_{TV}\big(z^\epsilon\big(\delta_\epsilon, z^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big), \mathcal{G}(0, \epsilon\Sigma)\big) + d_{TV}\big(\mathcal{G}(0, \epsilon\Sigma), \mu^\epsilon\big).
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&d_{TV}\big(z^\epsilon\big(\delta_\epsilon, z^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big), \mathcal{G}(0, \epsilon\Sigma)\big) \\
&\leq d_{TV}\big(z^\epsilon\big(\delta_\epsilon, z^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big), z^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big)\big) \\
&\quad + d_{TV}\big(z^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big), x^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big)\big) \\
&\quad + d_{TV}\big(x^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big), \mu^\epsilon\big) + d_{TV}\big(\mu^\epsilon, \mathcal{G}(0, \epsilon\Sigma)\big).
\end{aligned}$$

Gluing both inequalities we deduce

$$\begin{aligned}
&\big|d_{TV}\big(x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon + \delta_\epsilon, x_0\big), \mu^\epsilon\big) - d_{TV}\big(z^\epsilon\big(\delta_\epsilon, z^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big), \mathcal{G}(0, \epsilon\Sigma)\big)\big| \\
&\leq d_{TV}\big(x^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big), z^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big)\big) \\
&\quad + d_{TV}\big(z^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big), z^\epsilon\big(\delta_\epsilon, z^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big)\big) \\
&\quad + d_{TV}\big(\mathcal{G}(0, \epsilon\Sigma), \mu^\epsilon\big). \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \square
\end{aligned}$$

In what follows, we prove that the upper bound of inequality (3.14) is negligible as $\epsilon \to 0$. To be precise, we prove

CLAIM A (Short-time coupling).

$$\lim_{\epsilon \to 0} d_{TV}\big(x^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big), z^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big)\big) = 0.$$

This is the content of Section 3.3.1.

CLAIM B (Linear nonhomogeneous coupling).

$$\lim_{\epsilon \to 0} d_{TV}\big(z^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big), z^\epsilon\big(\delta_\epsilon, z^\epsilon\big(t^\epsilon_{shift} + cw^\epsilon, x_0\big)\big)\big) = 0.$$

This is the content of Section 3.3.2.

CLAIM C (Local central limit theorem).

$$\lim_{\epsilon \to 0} d_{TV}\big(\mathcal{G}(0, \epsilon\Sigma), \mu^\epsilon\big) = 0.$$

This is the content of Section 3.3.4.

CLAIM D (Windows cut-off).   *The family of processes*

$$\{z^\epsilon := \{z^\epsilon(t) : t \geq 0\} : \epsilon \in (0, 1]\} \quad \text{presents windows cut-off.}$$

This is the content of Section 3.3.3 together with Corollary 2.7.

3.3.1. *Short-time coupling.*   A natural question arising is how to obtain explicit "good" bounds for the total variation distance between $x^\epsilon(t)$ and $z^\epsilon(t)$. Using the celebrated Cameron–Martin–Girsanov Theorem, a coupling on the path space can be done and it is possible to establish bounds on the total variation distance using the Pinsker inequality of such diffusions. This method only provides a coupling over short time intervals. For more details see [10, 21, 31] and the references therein. On the other hand, "explicit" bounds for the total variation distance between transition probabilities of diffusions with different drifts are derived using analytic arguments. This approach also works for the stationary measures of the diffusions. For further details see [13] and the references therein.

In order to avoid homogenisation arguments for F, we use the Hellinger approach developed in [31] to obtain an upper bound for the total variation distance between the nonlinear model $x^\epsilon(t)$ with the linear nonhomogeneous model $z^\epsilon(t)$ in a short time interval. That upper bound is enough for our purposes. As we can notice in Theorem 5.1 in [31], we need to carry out second-moment estimates of the distance between the vector fields associated to the diffusions $\{x^\epsilon(t) : t \geq 0\}$ and $\{z^\epsilon(t) : t \geq 0\}$, respectively. It is exactly the estimate that we did in Lemma 3.1.

PROPOSITION 3.3.   *Assume that* (H) *and* (G) *hold. Let* $\delta_\epsilon > 0$ *such that* $\delta_\epsilon = o(1)$. *Then for any* $c \in \mathbb{R}$

$$\lim_{\epsilon \to 0} \mathrm{d}_{\mathrm{TV}}\big(x^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{\mathrm{shift}} + cw^\epsilon, x_0\big)\big), z^\epsilon\big(\delta_\epsilon, x^\epsilon\big(t^\epsilon_{\mathrm{shift}} + cw^\epsilon, x_0\big)\big)\big) = 0,$$

where $t^\epsilon_{\mathrm{shift}}$ *is given by* (3.13).

PROOF.   Let $T^\epsilon = t^\epsilon_{\mathrm{shift}} + cw^\epsilon > 0$ for $\epsilon \ll 1$. Notice that

$$\mathrm{d}_{\mathrm{TV}}\big(x^\epsilon\big(\delta_\epsilon, x^\epsilon\big(T^\epsilon, x_0\big)\big), z^\epsilon\big(\delta_\epsilon, x^\epsilon\big(T^\epsilon, x_0\big)\big)\big)$$
$$\leq \int_{\mathbb{R}^d} \mathrm{d}_{\mathrm{TV}}\big(x^\epsilon(\delta_\epsilon, u), z^\epsilon(\delta_\epsilon, u)\big)\mathbb{P}\big(x^\epsilon\big(T^\epsilon, x_0\big) \in \mathrm{d}u\big).$$

For short, denote by $\mathbb{P}^\epsilon(\mathrm{d}u)$ the probability measure $\mathbb{P}(x^\epsilon(T^\epsilon, x_0) \in \mathrm{d}u)$. Let K be a positive constant. Then

$$\mathrm{d}_{\mathrm{TV}}\big(x^\epsilon\big(\delta_\epsilon, x^\epsilon\big(T^\epsilon, x_0\big)\big), z^\epsilon\big(\delta_\epsilon, x^\epsilon\big(T^\epsilon, x_0\big)\big)\big)$$

(3.15)

$$\leq \int_{\|u\| \leq K} \mathrm{d}_{\mathrm{TV}}\big(x^\epsilon(\delta_\epsilon, u), z^\epsilon(\delta_\epsilon, u)\big)\mathbb{P}^\epsilon(\mathrm{d}u) + \mathbb{P}\big(\big\|x^\epsilon\big(T^\epsilon, x_0\big)\big\| > K\big).$$

Now, we prove that the upper bound of (3.15) is negligible as $\epsilon \to 0$. From the Markov inequality we get

$$\mathbb{P}\big(\big\|x^\epsilon\big(T^\epsilon, x_0\big)\big\| > K\big) \leq \frac{\mathbb{E}[\|x^\epsilon(T^\epsilon, x_0)\|^2]}{K^2}.$$

Recall the well-known inequality $(x + y)^2 \leq 2(x^2 + y^2)$ for any $x, y \in \mathbb{R}$. Then

$$\mathbb{E}\big[\big\|x^\epsilon\big(T^\epsilon, x_0\big)\big\|^2\big] \leq 2\mathbb{E}\big[\big\|x^\epsilon\big(T^\epsilon, x_0\big) - \varphi\big(T^\epsilon, x_0\big)\big\|^2\big]$$
$$+ 2\big\|\varphi\big(T^\epsilon, x_0\big)\big\|^2.$$

From inequality (3.4) and inequality (2.2) we have

$$\mathbb{E}\big[\big\|x^\epsilon\big(T^\epsilon, x_0\big)\big\|^2\big] \leq \frac{\epsilon d}{\delta} + 2e^{-2\delta T^\epsilon}\|x_0\|^2,$$

which allows us to deduce

(3.16) $$\lim_{\epsilon \to 0} \mathbb{P}\big(\big\|x^\epsilon\big(T^\epsilon, x_0\big)\big\| > K\big) = 0.$$

Now, we analyse $\int_{\|u\|\leq K} d_{TV}(x^\epsilon(\delta_\epsilon, u), z^\epsilon(\delta_\epsilon, u))\mathbb{P}^\epsilon(du)$. From Theorem 5.1 in [31] we obtain

$$\int_{\|u\|\leq K} d_{TV}\big(x^\epsilon(\delta_\epsilon, u), z^\epsilon(\delta_\epsilon, u)\big)\mathbb{P}^\epsilon(du)$$
$$\leq \frac{1}{\epsilon}\int_{\|u\|\leq K}\int_0^{\delta_\epsilon} \mathbb{E}\big[\|I^\epsilon(s, u)\|^2\big]\,ds\mathbb{P}^\epsilon(du),$$

where

$$I^\epsilon(s, u) := F(x^\epsilon(s, u) - DF\big(\varphi(s, u)\big)z^\epsilon(s, u) + DF\big(\varphi(s, u)\big)\varphi(s, u) - F\big(\varphi(s, u)\big)$$

for any $s \geq 0$ and $u \in \mathbb{R}^d$. Following the same argument using Lemma 3.1, for any $0 \leq s \leq \delta_\epsilon$ and $\epsilon \ll 1$ we deduce

$$\mathbb{E}\big[\|I^\epsilon(s, u)\|^2\big] \leq C(K, d, \delta)\epsilon^2\delta_\epsilon + \tilde{C}(K, d, \delta)\epsilon^{3/2}\delta_\epsilon^{7/4}$$

for any $u \in \mathbb{R}^d$ such that $\|u\| \leq K$, where $C(K, d, \delta)$ and $\tilde{C}(K, d, \delta)$ only depend on K, d and δ. Therefore,

$$(3.17) \qquad \lim_{\epsilon\to 0}\int_{\|u\|\leq K} d_{TV}\big(x^\epsilon(\delta_\epsilon, u), z^\epsilon(\delta_\epsilon, u)\big)\mathbb{P}^\epsilon(du) = 0.$$

Inequality (3.15) together with relations (3.16) and (3.17) allow us to deduce the desired result. □

3.3.2. *Linear nonhomogeneous coupling.* In this part, we couple two nonhomogeneous solutions $z^\epsilon(t, x)$ and $z^\epsilon(t, y)$ for short time $t \ll 1$ and initials conditions x and y such that $\|x - y\|$ small enough.

PROPOSITION 3.4. *Assume that* (H) *and* (G) *hold. Let* $\delta_\epsilon = \epsilon^\theta$ *for some* $\theta \in (0, 1/2)$. *Then for any* $c \in \mathbb{R}$

$$\lim_{\epsilon\to 0} d_{TV}\big(z^\epsilon\big(\delta_\epsilon, x^\epsilon(t_{shift}^\epsilon + cw^\epsilon, x_0)\big), z^\epsilon\big(\delta_\epsilon, z^\epsilon(t_{shift}^\epsilon + cw^\epsilon, x_0)\big)\big) = 0,$$

where t_{shift}^ϵ *is given by* (3.13).

PROOF. Let $T^\epsilon = t_{shift}^\epsilon + cw^\epsilon > 0$ for $\epsilon \ll 1$. Notice that

$$d_{TV}\big(z^\epsilon\big(\delta_\epsilon, x^\epsilon(T^\epsilon, x_0)\big), z^\epsilon\big(\delta_\epsilon, z^\epsilon(T^\epsilon, x_0)\big)\big)$$
$$\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} d_{TV}\big(z^\epsilon(\delta_\epsilon, u), z^\epsilon(\delta_\epsilon, \tilde{u})\big)\mathbb{P}\big(x^\epsilon(T^\epsilon, x_0) \in du, z^\epsilon(T^\epsilon, x_0) \in d\tilde{u}\big).$$

For short, we denote by $\mathbb{P}^\epsilon(du, d\tilde{u})$ for the coupling

$$\mathbb{P}\big(x^\epsilon(T^\epsilon, x_0) \in du, z^\epsilon(T^\epsilon, x_0) \in d\tilde{u}\big).$$

Let K and $\tilde{K}$ any positive constants. Then

$$d_{TV}\big(z^\epsilon\big(\delta_\epsilon, x^\epsilon(T^\epsilon, x_0)\big), z^\epsilon\big(\delta_\epsilon, z^\epsilon(T^\epsilon, x_0)\big)\big)$$
$$(3.18) \qquad \leq \int_{\|u\|\leq K, \|\tilde{u}\|\leq\tilde{K}} d_{TV}\big(z^\epsilon(\delta_\epsilon, u), z^\epsilon(\delta_\epsilon, \tilde{u})\big)\mathbb{P}^\epsilon(du, d\tilde{u})$$
$$+ \mathbb{P}\big(\|x^\epsilon(T^\epsilon, x_0)\| > K\big) + \mathbb{P}\big(\|z^\epsilon(T^\epsilon, x_0)\| > \tilde{K}\big).$$

Now, we prove that the upper bound of (3.18) is negligible as $\epsilon \to 0$. From relation (3.16) we have

$$\lim_{\epsilon\to 0}\mathbb{P}\big(\|x^\epsilon(T^\epsilon, x_0)\| > K\big) = 0.$$

Similar ideas as in the proof of relation (3.16) and item (ii) of Proposition C.2 yield

$$\lim_{\epsilon \to 0} \mathbb{P}\big(\|z^\epsilon(T^\epsilon, x_0)\| > \tilde{K}\big) = 0.$$

Since the stochastic differential equation associated to $\{y(t, u, T^\epsilon) : t \geq 0\}$ is linear then the variation of parameters formula allows us to deduce that

$$z^\epsilon(\delta_\epsilon, u) = \varphi(\delta_\epsilon, u) + \sqrt{\epsilon}\Phi(\delta_\epsilon)u + \sqrt{\epsilon}\Phi(\delta_\epsilon)\int_0^{\delta_\epsilon} \big(\Phi(s)\big)^{-1}\,\mathrm{d}B\big(T^\epsilon + s\big)$$

for any $u \in \mathbb{R}^d$, where $\{\Phi(t) : t \geq 0\}$ is the solution of the matrix differential equation

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\Phi(t) = -DF\big(\varphi(t + T^\epsilon)\big)\Phi(t) & \text{for } t \geq 0, \\ \Phi(0) = I_d. \end{cases}$$

Observe that for any $\upsilon \in \mathbb{R}^d$, $z^\epsilon(\delta_\epsilon, \upsilon)$ has Gaussian distribution with mean vector $\varphi(\delta_\epsilon, \upsilon)$ and covariance matrix $\epsilon\Sigma(\delta_\epsilon)$, where $\Sigma(\delta_\epsilon)$ is the covariance matrix of the random vector

$$\Phi(\delta_\epsilon)\int_0^{\delta_\epsilon} \big(\Phi(s)\big)^{-1}\,\mathrm{d}B\big(T^\epsilon + s\big)$$

which does not depend on υ. Moreover, using the Itô formula we deduce

$$\Sigma(\delta_\epsilon) = \Phi(\delta_\epsilon)\int_0^{\delta_\epsilon} \big(\Phi(s)\big)^{-1}\big(\big(\Phi(s)\big)^{-1}\big)^*\,\mathrm{d}s\big(\Phi(\delta_\epsilon)\big)^*.$$

Since $\lim_{\epsilon \to 0}\Phi(\delta_\epsilon) = I_d$, then one can deduce

$$\lim_{\epsilon \to 0}\frac{\Sigma(\delta_\epsilon)}{\delta_\epsilon} = I_d.$$

The latter allows us to deduce that $\|(\Sigma(\delta_\epsilon))^{-1/2}\| \leq (\delta_\epsilon)^{-1/2}C(d)$, where $C(d)$ is a positive constant that only depends on d.

Remind that $\mathcal{G}(v, \Xi)$ denotes the Gaussian distribution in $\mathbb{R}^d$ with vector mean v and positive definite covariance matrix Ξ. From item (ii), item (iii) of Lemma A.1 and Lemma A.2 we have

$$d_{\mathrm{TV}}\big(z^\epsilon(\delta_\epsilon, u), z^\epsilon(\delta_\epsilon, \tilde{u})\big)$$

$$= d_{\mathrm{TV}}\big(\mathcal{G}\big(\varphi(\delta_\epsilon, u) + \sqrt{\epsilon}\Phi(\delta_\epsilon)u, \epsilon\Sigma(\delta_\epsilon)\big), \mathcal{G}\big(\varphi(\delta_\epsilon, \tilde{u}) + \sqrt{\epsilon}\Phi(\delta_\epsilon)\tilde{u}, \epsilon\Sigma(\delta_\epsilon)\big)\big)$$

$$\leq \frac{1}{\sqrt{2\pi\epsilon}}\big\|(\Sigma(\delta_\epsilon))^{-1/2}\big(\varphi(\delta_\epsilon, u) - \varphi(\delta_\epsilon, \tilde{u}) + \sqrt{\epsilon}\Phi(\delta_\epsilon)(u - \tilde{u})\big)\big\|$$

for any $u, \tilde{u} \in \mathbb{R}^d$. From Condition (H) we obtain

$$\big\|\varphi(\delta_\epsilon, u) - \varphi(\delta_\epsilon, \tilde{u})\big\| \leq e^{-\delta\delta_\epsilon}\|u - \tilde{u}\| \leq \|u - \tilde{u}\|$$

for any $u, \tilde{u} \in \mathbb{R}^d$. Then

$$d_{\mathrm{TV}}\big(z^\epsilon(\delta_\epsilon, u), z^\epsilon(\delta_\epsilon, \tilde{u})\big) \leq \frac{C_1(d)}{\sqrt{\epsilon\delta_\epsilon}}\|u - \tilde{u}\| \quad \text{for any } u, \tilde{u} \in \mathbb{R}^d,$$

where $C_1(d)$ is a positive constant that only depends on d.

Therefore,

$$\int_{\|u\| \leq K, \|\tilde{u}\| \leq \tilde{K}} d_{\mathrm{TV}}\big(z^\epsilon(\delta_\epsilon, u), z^\epsilon(\delta_\epsilon, \tilde{u})\big)\mathbb{P}^\epsilon(\mathrm{d}u, \mathrm{d}\tilde{u})$$

$$\leq \frac{C_1(d)}{\sqrt{\epsilon\delta_\epsilon}}\mathbb{E}\big[\|x^\epsilon(T^\epsilon, x_0) - z^\epsilon(T^\epsilon, x_0)\|\big]$$

$$\leq \frac{C_1(d)}{\sqrt{\epsilon\delta_\epsilon}}\big(\mathbb{E}\big[\|x^\epsilon(T^\epsilon, x_0) - z^\epsilon(T^\epsilon, x_0)\|^2\big]\big)^{1/2}.$$

From Lemma 3.1 we deduce

$$\lim_{\epsilon \to 0} \int_{\|u\| \leq K, \|\tilde{u}\| \leq \tilde{K}} d_{TV}\big(z^\epsilon(\delta_\epsilon, u), z^\epsilon(\delta_\epsilon, \tilde{u})\big) \mathbb{P}^\epsilon(du, d\tilde{u}) = 0.$$

Putting all pieces together, we get the statement. $\quad\square$

REMARK 3.5. Under the same assumptions we can notice that Proposition 3.3 and Proposition 3.4 also hold when $T^\epsilon = t^\epsilon_{\text{shift}} + cw^\epsilon > 0$ is replaced by $\frac{1}{\epsilon^\gamma}$ with $\gamma \in (0, \frac{2-4\theta}{7})$.

3.3.3. *Window cut-off.* Remind that $z^\epsilon(t) = \varphi(t) + \sqrt{\epsilon} y(t)$, $t \geq 0$, where $\{y(t) : t \geq 0\}$ satisfies the linear nonhomogeneous stochastic differential equation

$$\begin{cases} dy(t) = -DF(\varphi(t))y(t)\,dt + dB(t) & \text{for } t \geq 0, \\ y(0) = 0. \end{cases}$$

Therefore, for any $t > 0$, $z^\epsilon(t)$ has Gaussian distribution with mean vector $\varphi(t)$ and covariance matrix $\epsilon \Sigma(t)$, where $\{\Sigma(t) : t \geq 0\}$ is the solution to the deterministic matrix differential equation

$$\begin{cases} \dfrac{d}{dt}\Sigma(t) = -DF(\varphi(t))\Sigma(t) - \Sigma(t)\big(DF(\varphi(t))\big)^* + I_d & \text{for } t \geq 0, \\ \Sigma(0) = 0. \end{cases}$$

Under (H), we can prove that $\varphi(t) \to 0$ and $\Sigma(t) \to \Sigma$ as $t \to +\infty$, where Σ is a symmetric and positive definite matrix (See Lemma C.6). Therefore, $z^\epsilon(t)$ converges in distribution to a random vector $z^\epsilon(\infty)$ as $t \to +\infty$, where $z^\epsilon(\infty)$ has Gaussian law with zero mean vector and covariance matrix $\epsilon \Sigma$. Using item (iii) of Lemma A.1, Lemma A.3, Lemma A.5 the convergence can be easily improved to be in total variation distance. Bearing all this in mind, we can analyse the convergence of $z^\epsilon(t)$ to its equilibrium in the total variation distance. Define

$$\bar{D}^\epsilon(t) := d_{TV}\big(z^\epsilon(t), z^\epsilon(\infty)\big) = d_{TV}\big(\mathcal{G}(\varphi(t), \epsilon\Sigma(t)), \mathcal{G}(0, \epsilon\Sigma)\big)$$

for any $t > 0$.

PROPOSITION 3.6. *Assume that* (H) *holds. Let* $\delta_\epsilon \geq 0$ *such that* $\delta_\epsilon = o(1)$. *For any* $c \in \mathbb{R}$ *we have*

$$\lim_{\epsilon \to 0}\big|\bar{D}^\epsilon(t^\epsilon_{\text{shift}} + \delta_\epsilon + cw^\epsilon) - D^\epsilon(t^\epsilon_{\text{shift}} + \delta_\epsilon + cw^\epsilon)\big| = 0,$$

where $t^\epsilon_{\text{shift}}$ *is given by* (3.13),

$$D^\epsilon(t) := d_{TV}\left(\mathcal{G}\left(\frac{(t-\tau)^{\ell-1}}{e^{\lambda(t-\tau)}\sqrt{\epsilon}}\Sigma^{-1/2}\sum_{k=1}^m e^{i\theta_k(t-\tau)}v_k, I_d\right), \mathcal{G}(0, I_d)\right)$$

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, *and the matrix* Σ *is the unique solution of the matrix Lyapunov equation*:

$$DF(0)X + X\big(DF(0)\big)^* = I_d.$$

PROOF. Let $t > 0$. From the triangle inequality and item (ii), item (iii) of Lemma A.1 we obtain

$$\bar{D}^\epsilon(t) \leq d_{TV}\big(\mathcal{G}(\varphi(t), \epsilon\Sigma(t)), \mathcal{G}(\varphi(t), \epsilon\Sigma)\big) + d_{TV}\big(\mathcal{G}(\varphi(t), \epsilon\Sigma), \mathcal{G}(0, \epsilon\Sigma)\big)$$

$$\leq d_{TV}\big(\mathcal{G}(0, \Sigma(t)), \mathcal{G}(0, \Sigma)\big) + d_{TV}\left(\mathcal{G}\left(\frac{1}{\sqrt{\epsilon}}\varphi(t), \Sigma\right), \mathcal{G}(0, \Sigma)\right).$$

From a similar argument we obtain

$$
\begin{aligned}
d_{TV} &\left(\mathcal{G}\left(\frac{1}{\sqrt{\epsilon}} \varphi(t), \Sigma \right), \mathcal{G}(0, \Sigma) \right) \\
&= d_{TV}(\mathcal{G}(\varphi(t), \epsilon \Sigma), \mathcal{G}(0, \epsilon \Sigma)) \\
&\leq d_{TV}(\mathcal{G}(\varphi(t), \epsilon \Sigma), \mathcal{G}(\varphi(t), \epsilon \Sigma(t))) + d_{TV}(\mathcal{G}(\varphi(t), \epsilon \Sigma(t)), \mathcal{G}(0, \Sigma)) \\
&= d_{TV}(\mathcal{G}(0, \Sigma), \mathcal{G}(0, \Sigma(t))) + \bar{D}^\epsilon(t).
\end{aligned}
$$

Putting all pieces together we deduce

$$
(3.19) \qquad \left| \bar{D}^\epsilon(t) - d_{TV}\left(\mathcal{G}\left(\frac{1}{\sqrt{\epsilon}} \varphi(t), \Sigma \right), \mathcal{G}(0, \Sigma) \right) \right| \leq d_{TV}(\mathcal{G}(0, \Sigma(t)), \mathcal{G}(0, \Sigma)).
$$

Using Lemma A.5 and Lemma C.6 we get

$$
(3.20) \qquad \lim_{t \to +\infty} d_{TV}(\mathcal{G}(0, \Sigma(t)), \mathcal{G}(0, \Sigma)) = 0.
$$

Therefore, the cut-off phenomenon can be deduced from the distance

$$
\begin{aligned}
\hat{D}^\epsilon(t) &:= d_{TV}\left(\mathcal{G}\left(\frac{1}{\sqrt{\epsilon}} \varphi(t), \Sigma \right), \mathcal{G}(0, \Sigma) \right) \\
&= d_{TV}\left(\mathcal{G}\left(\Sigma^{-1/2} \frac{1}{\sqrt{\epsilon}} \varphi(t), I_d \right), \mathcal{G}(0, I_d) \right)
\end{aligned}
$$

for any $t > 0$, where the last equality follows from item (iii) of Lemma A.1. Using the constants and vectors associated to x_0 in Lemma 2.1, for any $t > \tau$ define

$$
D^\epsilon(t) := d_{TV}\left(\mathcal{G}\left(\Sigma^{-1/2} \frac{(t - \tau)^{\ell-1}}{e^{\lambda(t-\tau)} \sqrt{\epsilon}} \sum_{k=1}^m e^{i\theta_k(t-\tau)} v_k, I_d \right), \mathcal{G}(0, I_d) \right)
$$

and

$$
\begin{aligned}
R^\epsilon(t) & \\
&:= d_{TV}\left(\mathcal{G}\left(\Sigma^{-1/2} \frac{1}{\sqrt{\epsilon}} \varphi(t), I_d \right), \mathcal{G}\left(\Sigma^{-1/2} \frac{(t - \tau)^{\ell-1}}{e^{\lambda(t-\tau)} \sqrt{\epsilon}} \sum_{k=1}^m e^{i\theta_k(t-\tau)} v_k, I_d \right) \right).
\end{aligned}
$$

From item (ii) of Lemma A.1 we deduce that

$$
\begin{aligned}
R^\epsilon(t) & \\
&= d_{TV}\left(\mathcal{G}\left(\frac{1}{\sqrt{\epsilon}} \Sigma^{-1/2} \left(\varphi(t) - \frac{(t - \tau)^{\ell-1}}{e^{\lambda(t-\tau)}} \sum_{k=1}^m e^{i\theta_k(t-\tau)} v_k \right), I_d \right), \mathcal{G}(0, I_d) \right).
\end{aligned}
$$

A similar argument used to deduce inequality (3.19) allows us to show that

$$
(3.21) \qquad \left| \hat{D}^\epsilon(t) - D^\epsilon(t) \right| \leq R^\epsilon(t) \quad \text{for any } t > \tau.
$$

From inequalities (3.19) and (3.21) we obtain

$$
(3.22) \qquad \left| D^\epsilon(t) - \bar{D}^\epsilon(t) \right| \leq R^\epsilon(t) + d_{TV}(\mathcal{G}(0, \Sigma(t)), \mathcal{G}(0, \Sigma))
$$

for any $t > \tau$. Straightforward computations led us to

$$
(3.23) \qquad \lim_{\epsilon \to 0} \frac{e^{-\lambda(t_{\text{shift}}^\epsilon + \delta_\epsilon + cw^\epsilon - \tau)} (t_{\text{shift}}^\epsilon + \delta_\epsilon + cw^\epsilon - \tau)^{\ell-1}}{\sqrt{\epsilon}} = (2\lambda)^{1-\ell} e^{-c}
$$

for any $c \in \mathbb{R}$. Therefore, Lemma 2.1 together relation (3.23) and Lemma A.3 allow to deduce that

$$\lim_{\epsilon \to 0} R^{\epsilon} \big(t_{\text{shift}}^{\epsilon} + \delta_{\epsilon} + c w^{\epsilon} \big) = 0 \tag{3.24}$$

for any $c \in \mathbb{R}$. Consequently, from inequality (3.22) together with relation (3.20) and relation (3.24) we obtain the statement. $\square$

3.3.4. *The invariant measure.* In this section, we prove a L^1-local central limit theorem. We prove that the invariant measure μ^{ϵ} of the evolution (2.4) is well approximated in total variation distance by a Gaussian distribution with zero mean vector and covariance matrix $\epsilon \Sigma$, where Σ is the unique solution of the matrix Lyapunov equation

$$DF(0)X + X(DF(0))^* = I_d.$$

PROPOSITION 3.7. *Assume that* (H) *and* (G) *hold. Then*

$$\lim_{\epsilon \to 0} d_{\text{TV}} \big(\mathcal{G}(0, \epsilon \Sigma), \mu^{\epsilon} \big) = 0.$$

PROOF. Recall that $z^{\epsilon}(t) = \varphi(t) + \sqrt{\epsilon} y(t)$ for any $t \geq 0$. Note that for any $s, t \geq 0$ and $x \in \mathbb{R}^d$ we have

$$\begin{aligned} &d_{\text{TV}} \big(\mathcal{G}(0, \epsilon \Sigma), \mu^{\epsilon} \big) \\ &\quad \leq d_{\text{TV}} \big(\mathcal{G}(0, \epsilon \Sigma), z^{\epsilon}(s + t, x) \big) \\ &\quad\quad + d_{\text{TV}} \big(z^{\epsilon}(s + t, x), x^{\epsilon}(s + t, x) \big) + d_{\text{TV}} \big(x^{\epsilon}(s + t, x), \mu^{\epsilon} \big). \end{aligned} \tag{3.25}$$

Observe that

$$d_{\text{TV}} \big(\mathcal{G}(0, \epsilon \Sigma), z^{\epsilon}(s + t, x) \big) = d_{\text{TV}} \big(\mathcal{G}(0, \epsilon \Sigma), \mathcal{G}(\varphi(s + t, x), \epsilon \Sigma(s + t)) \big),$$

where $\Sigma(t)$ is the covariance matrix of $y(t)$. Therefore, using the triangle inequality together with item (ii) and item (iii) of Lemma A.1 we obtain

$$\begin{aligned} d_{\text{TV}} \big(\mathcal{G}(0, \epsilon \Sigma), z^{\epsilon}(s + t, x) \big) &\leq d_{\text{TV}} \big(\mathcal{G}(0, \Sigma), \mathcal{G}(0, \Sigma(s + t)) \big) \\ &\quad + d_{\text{TV}} \bigg(\mathcal{G}\bigg(\frac{\varphi(s + t, x)}{\sqrt{\epsilon}}, \Sigma \bigg), \mathcal{G}(0, \Sigma) \bigg). \end{aligned} \tag{3.26}$$

Let $s^{\epsilon} \ll \epsilon^{1/2019}$ and $t_{\text{mix}}^{\epsilon} \ll t^{\epsilon} := \frac{1}{\epsilon^{1/8}}$. By Lemma A.5 and Lemma C.6 we obtain

$$\lim_{\epsilon \to 0} d_{\text{TV}} \big(\mathcal{G}(0, \Sigma), \mathcal{G}(0, \Sigma(s^{\epsilon} + t^{\epsilon})) \big) = 0.$$

From (H) we obtain $\|\varphi(s^{\epsilon} + t^{\epsilon}, x)\| \leq \|x\| e^{-\delta(s^{\epsilon} + t^{\epsilon})}$. Straightforward computations led us to deduce that

$$\lim_{\epsilon \to 0} \frac{\|\varphi(s^{\epsilon} + t^{\epsilon}, x)\|}{\sqrt{\epsilon}} = 0.$$

The latter together with item (iii) of Lemma A.1 and Lemma A.3 imply

$$\lim_{\epsilon \to 0} d_{\text{TV}} \bigg(\mathcal{G}\bigg(\frac{\varphi(s^{\epsilon} + t^{\epsilon}, x)}{\sqrt{\epsilon}}, \Sigma \bigg), \mathcal{G}(0, \Sigma) \bigg) = 0.$$

Therefore, from inequality (3.26) we obtain

$$\lim_{\epsilon \to 0} d_{\text{TV}} \big(\mathcal{G}(0, \epsilon \Sigma), z^{\epsilon}(s^{\epsilon} + t^{\epsilon}, x) \big) = 0.$$

Since the stochastic differential equation associated to $\{y^\epsilon(t) : t \geq 0\}$ *is not homogeneous* we should improve the notation as we did in the beginning of Section 3.3. Following such notation, we always use $T = t^\epsilon$. Therefore, for simplicity, we can omit as we did in Proposition 3.3 but we should always keep it in mind.

Notice that

$$d_{TV}\big(z^\epsilon\big(s^\epsilon + t^\epsilon, x\big), x^\epsilon\big(s^\epsilon + t^\epsilon, x\big)\big) \leq d_{TV}\big(z^\epsilon\big(s^\epsilon, z^\epsilon(t, x)\big), z^\epsilon\big(s^\epsilon, x^\epsilon(t, x)\big)\big)$$
$$+ d_{TV}\big(z^\epsilon\big(s^\epsilon, x^\epsilon(t, x)\big), x^\epsilon\big(s^\epsilon, x^\epsilon(t, x)\big)\big).$$

Now, using the same ideas as in Proposition 3.3 and Proposition 3.4 together with Remark 3.5 (much easier since $t^\epsilon \gg t^\epsilon_{\text{mix}}$) we deduce

$$\lim_{\epsilon \to 0} d_{TV}\big(z^\epsilon\big(s^\epsilon + t^\epsilon, x\big), x^\epsilon\big(s^\epsilon + t^\epsilon, x\big)\big) = 0.$$

From inequality (3.25), it remains to prove that

$$\lim_{\epsilon \to 0} d_{TV}\big(x^\epsilon\big(s^\epsilon + t^\epsilon, x\big), \mu^\epsilon\big) = 0.$$

Notice that

$$d_{TV}\big(x^\epsilon(s + t, x), \mu^\epsilon\big) \leq \int_{\mathbb{R}^d} d_{TV}\big(x^\epsilon(s + t, x), x^\epsilon(s + t, \bar{x})\big)\mu^\epsilon(d\bar{x}).$$

Then

$$\int_{\mathbb{R}^d} d_{TV}\big(x^\epsilon(s + t, x), x^\epsilon(s + t, \bar{x})\big)\mu^\epsilon(d\bar{x})$$

$$\leq d_{TV}\big(x^\epsilon\big(s, x^\epsilon(t, x)\big), z^\epsilon\big(s, x^\epsilon(t, x)\big)\big) + d_{TV}\big(z^\epsilon\big(s, x^\epsilon(t, x)\big), z^\epsilon\big(s, z^\epsilon(t, x)\big)\big)$$

$$+ \int_{\mathbb{R}^d} d_{TV}\big(z^\epsilon\big(s, z^\epsilon(t, x)\big), z^\epsilon\big(s, z^\epsilon(t, \bar{x})\big)\big)\mu^\epsilon(d\bar{x})$$

$$+ \int_{\mathbb{R}^d} d_{TV}\big(z^\epsilon\big(s, z^\epsilon(t, \bar{x})\big), x^\epsilon\big(s, x^\epsilon(t, \bar{x})\big)\big)\mu^\epsilon(d\bar{x}).$$

Again, using the same ideas as in Proposition 3.3 and Proposition 3.4 together with Remark 3.5 we deduce

$$\lim_{\epsilon \to 0} d_{TV}\big(x^\epsilon\big(s^\epsilon, x^\epsilon(t^\epsilon, x)\big), z^\epsilon\big(s^\epsilon, x^\epsilon(t^\epsilon, x)\big)\big) = 0$$

and

$$\lim_{\epsilon \to 0} d_{TV}\big(z^\epsilon\big(s^\epsilon, x^\epsilon(t^\epsilon, x)\big), z^\epsilon\big(s^\epsilon, z^\epsilon(t^\epsilon, x)\big)\big) = 0.$$

Fix $R > 0$. We split the remainders integrals as follows:

$$\int_{\mathbb{R}^d} d_{TV}\big(z^\epsilon\big(s, z^\epsilon(t, x)\big), z^\epsilon\big(s, z^\epsilon(t, \bar{x})\big)\big)\mu^\epsilon(d\bar{x})$$

$$\leq \int_{\|\bar{x}\| \leq R} d_{TV}\big(z^\epsilon\big(s, z^\epsilon(t, x)\big), z^\epsilon\big(s, z^\epsilon(t, \bar{x})\big)\big)\mu^\epsilon(d\bar{x}) + \mu^\epsilon\big(\|x\| > R\big)$$

and

$$\int_{\mathbb{R}^d} d_{TV}\big(z^\epsilon\big(s, z^\epsilon(t, \bar{x})\big), x^\epsilon\big(s, x^\epsilon(t, \bar{x})\big)\big)\mu^\epsilon(d\bar{x})$$

$$\leq \int_{\|\bar{x}\| \leq R} d_{TV}\big(z^\epsilon\big(s, z^\epsilon(t, \bar{x})\big), x^\epsilon\big(s, x^\epsilon(t, \bar{x})\big)\big)\mu^\epsilon(d\bar{x}) + \mu^\epsilon\big(\|x\| > R\big).$$

Again, following the same ideas as in the proof of Proposition 3.3 together with Remark 3.5 we deduce

$$\lim_{\epsilon \to 0} \int_{\|\bar{x}\| \leq R} \mathrm{d}_{\mathrm{TV}}\big(z^{\epsilon}\big(s^{\epsilon}, z^{\epsilon}\big(t^{\epsilon}, \bar{x}\big)\big), x^{\epsilon}\big(s^{\epsilon}, x^{\epsilon}\big(t^{\epsilon}, \bar{x}\big)\big)\big)\mu^{\epsilon}(\mathrm{d}\bar{x}) = 0.$$

Now, we only need to prove that $\mu^{\epsilon}(\|x\| > R)$ is negligible when $\epsilon \to 0$. Following the same ideas in [46] (page 122, Section 5, Step 1), the invariant measure μ^{ϵ} has finite p-moments for any $p \geq 0$. Moreover, we have $\int_{\mathbb{R}^d} \|x\|^2 \mu^{\epsilon}(\mathrm{d}x) \leq \frac{\epsilon d}{\delta}$. Indeed, from inequality (3.4) we have

$$\mathbb{E}\big[\|x^{\epsilon}(t, x)\|^2\big] \leq \|x\|^2 e^{-2\delta t} + \frac{d\epsilon}{2\delta} \quad \text{for any } t \geq 0 \text{ and } x \in \mathbb{R}^d.$$

For any two numbers $a, b \in \mathbb{R}$, denote by $a \wedge b$ the minimum between a and b. Recall that:

  (i) If $a \leq b$ then $a \wedge c \leq b \wedge c$ for any $c \in \mathbb{R}$.
 (ii) $(a + b) \wedge c \leq a \wedge c + b \wedge c$ for any $a, b, c \geq 0$.

Notice that for any $t \geq 0$, $n \in \mathbb{N}$ and $x \in \mathbb{R}^d$ we have

$$\mathbb{E}\big[\|x^{\epsilon}(t, x)\|^2 \wedge n\big] \leq \mathbb{E}\big[\|x^{\epsilon}(t, x)\|^2\big] \wedge n.$$

Then

$$\mathbb{E}\big[\|x^{\epsilon}(t, x)\|^2 \wedge n\big] \leq \left(\|x\|^2 e^{-2\delta t} + \frac{d\epsilon}{2\delta}\right) \wedge n \leq \big(\|x\|^2 e^{-2\delta t}\big) \wedge n + \frac{d\epsilon}{2\delta} \wedge n$$

for any $t \geq 0$, $n \in \mathbb{N}$ and $x \in \mathbb{R}^d$. Integrating this inequality against $\mu^{\epsilon}(\mathrm{d}x)$ we obtain

$$\int_{\mathbb{R}^d}\big(\|x\|^2 \wedge n\big)\mu^{\epsilon}(\mathrm{d}x) \leq \int_{\mathbb{R}^d}\big(\big(\|x\|^2 e^{-2\delta t}\big) \wedge n\big)\mu^{\epsilon}(\mathrm{d}x) + \frac{d\epsilon}{2\delta} \wedge n$$

for any $t \geq 0$ and $n \in \mathbb{N}$. Passing to the limit first as $t \to \infty$ and using the dominated convergence theorem we have

$$\int_{\mathbb{R}^d}\big(\|x\|^2 \wedge n\big)\mu^{\epsilon}(\mathrm{d}x) \leq \frac{d\epsilon}{2\delta} \wedge n$$

for any $n \in \mathbb{N}$. Now, taking $n \to \infty$ and using the monotone convergence theorem we have

$$\int_{\mathbb{R}^d} \|x\|^2 \mu^{\epsilon}(\mathrm{d}x) \leq \frac{d\epsilon}{2\delta}.$$

The latter together with the Chebyshev inequality imply

$$\mu^{\epsilon}\big(\|x\| \geq R\big) \leq \frac{d\epsilon}{2R^2\delta} \quad \text{for any } R > 0. \qquad \square$$

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, we state this as a lemma.

LEMMA 3.8. *Assume that* (H) *and* (G) *hold. Let* $\{x^{\epsilon}(t, x_0) : t \geq 0\}$ *be the solution of* (2.4) *and denote by* μ^{ϵ} *the unique invariant probability measure for the evolution given by* (2.4). *Denote by*

$$d^{\epsilon}(t) = \mathrm{d}_{\mathrm{TV}}\big(x^{\epsilon}(t, x_0), \mu^{\epsilon}\big) \quad \text{for any } t \geq 0$$

the total variation distance between the law of the random variable $x^\epsilon(t, x_0)$ and its invariant probability μ^ϵ. Consider the cut-off time t^ϵ_{mix} given by (2.7) and the time window given by (2.6). Let $x_0 \neq 0$. Then for any $c \in \mathbb{R}$ we have

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

where

$$D^\epsilon(t) = d_{\text{TV}}\left(\mathcal{G}\left(\frac{(t-\tau)^{\ell-1}}{e^{\lambda(t-\tau)}\sqrt{\epsilon}} \Sigma^{-1/2} \sum_{k=1}^{m} e^{i\theta_k(t-\tau)} v_k, \mathrm{I}_d \right), \mathcal{G}(0, \mathrm{I}_d) \right)$$

for any $t \geq \tau$ with $m, \lambda, \ell, \tau, \theta_1, \ldots, \theta_m, v_1, \ldots, v_m$ are the constants and vectors associated to x_0 in Lemma 2.1, *and the matrix Σ is the unique solution of the matrix Lyapunov equation*:

$$DF(0)X + X\left(DF(0)\right)^* = \mathrm{I}_d.$$

PROOF. First, from Lemma C.3 we have that there exists a unique invariant probability measure for the evolution (2.4). Let call the invariant measure by μ^ϵ. From Lemma 3.2 together with Proposition 3.3, Proposition 3.4 and Proposition 3.7 we deduce

$$\left| d^\epsilon\left(t^\epsilon_{\text{mix}} + cw^\epsilon\right) - \bar{D}^\epsilon\left(t^\epsilon_{\text{mix}} + cw^\epsilon\right) \right| = o(1) \quad \text{as } \epsilon \to 0.$$

From the triangle inequality we obtain

$$\left| d^\epsilon\left(t^\epsilon_{\text{mix}} + cw^\epsilon\right) - D^\epsilon\left(t^\epsilon_{\text{mix}} + cw^\epsilon\right) \right|$$
$$\leq \left| D^\epsilon\left(t^\epsilon_{\text{mix}} + cw^\epsilon\right) - \bar{D}^\epsilon\left(t^\epsilon_{\text{mix}} + cw^\epsilon\right) \right| + o(1) \quad \text{as } \epsilon \to 0.$$

The latter together with Proposition 3.6 allows us to deduce the statement. $\square$

APPENDIX A: PROPERTIES OF THE TOTAL VARIATION DISTANCE FOR GAUSSIAN DISTRIBUTIONS

Recall that $\mathcal{G}(v, \Xi)$ denotes the Gaussian distribution in $\mathbb{R}^d$ with vector mean v and positive definite covariance matrix Ξ. Since the proofs are straightforward, we left most of details to the interested reader.

LEMMA A.1. *Let $v, \tilde{v} \in \mathbb{R}^d$ be two fixed vectors and $\Xi, \tilde{\Xi}$ be two fixed symmetric positive definite $d \times d$ matrices. Then*:

(i) *For any scalar $c \neq 0$ we have*

$$d_{\text{TV}}\left(\mathcal{G}(cv, c^2\Xi), \mathcal{G}(c\tilde{v}, c^2\tilde{\Xi}) \right) = d_{\text{TV}}\left(\mathcal{G}(v, \Xi), \mathcal{G}(\tilde{v}, \tilde{\Xi}) \right).$$

(ii) $d_{\text{TV}}\left(\mathcal{G}(v, \Xi), \mathcal{G}(\tilde{v}, \tilde{\Xi}) \right) = d_{\text{TV}}\left(\mathcal{G}(v - \tilde{v}, \Xi), \mathcal{G}(0, \tilde{\Xi}) \right).$
(iii) $d_{\text{TV}}\left(\mathcal{G}(v, \Xi), \mathcal{G}(\tilde{v}, \Xi) \right) = d_{\text{TV}}\left(\mathcal{G}(\Xi^{-1/2}v, \mathrm{I}_d), \mathcal{G}(\Xi^{-1/2}\tilde{v}, \mathrm{I}_d) \right).$
(iv) $d_{\text{TV}}\left(\mathcal{G}(0, \Xi), \mathcal{G}(0, \tilde{\Xi}) \right) = d_{\text{TV}}\left(\mathcal{G}(0, \tilde{\Xi}^{-1/2}\Xi\tilde{\Xi}^{-1/2}), \mathcal{G}(0, \mathrm{I}_d) \right).$

PROOF. The proofs follow from the characterisation of the total variation distance between two probability measures with densities, that is,

$$d_{\text{TV}}(\mathbb{P}_1, \mathbb{P}_2) = \frac{1}{2} \int_{\mathbb{R}^d} |f_1(x) - f_2(x)| \, dx,$$

where f_1 and f_2 are the densities of $\mathbb{P}_1$ and $\mathbb{P}_2$, respectively, and using the change of variable theorem. $\square$

LEMMA A.2. *For any $v \in \mathbb{R}^d$ we have*

$$d_{\mathrm{TV}}(\mathcal{G}(v, \mathrm{I}_d), \mathcal{G}(0, \mathrm{I}_d)) = \sqrt{\frac{2}{\pi}} \int_0^{\|v\|/2} e^{-\frac{x^2}{2}} \, dx \le \frac{1}{\sqrt{2\pi}} \|v\|.$$

PROOF. The proof in dimension one is a straightforward computation. We left the details to the interested reader. For dimension bigger than one, the idea is to reduce the proof to dimension one. To do that, we use the following fact: for any $v, \tilde{v} \in \mathbb{R}^d$ such that $\|v\| = \|\tilde{v}\|$ there exists an orthogonal matrix A such that $\tilde{v} = A(v)$. Recall that the law of $\mathcal{G}(0, \mathrm{I}_d)$ is invariant under orthogonal transformations, that is, $O\mathcal{G}(0, \mathrm{I}_d) = \mathcal{G}(0, \mathrm{I}_d)$ for any orthogonal matrix O. Then for any $v, \tilde{v} \in \mathbb{R}^d$ with $\|v\| = \|\tilde{v}\|$ we have

$$d_{\mathrm{TV}}(\mathcal{G}(\tilde{v}, \mathrm{I}_d), \mathcal{G}(0, \mathrm{I}_d)) = d_{\mathrm{TV}}(\mathcal{G}(Av, \mathrm{I}_d), \mathcal{G}(0, \mathrm{I}_d)) = d_{\mathrm{TV}}(\mathcal{G}(Av, \mathrm{I}_d), A\mathcal{G}(0, \mathrm{I}_d))$$
$$= d_{\mathrm{TV}}(A\mathcal{G}(v, \mathrm{I}_d), A\mathcal{G}(0, \mathrm{I}_d)) = d_{\mathrm{TV}}(\mathcal{G}(v, \mathrm{I}_d), \mathcal{G}(0, \mathrm{I}_d)),$$

where the last equality follows from the characterisation of the total variation distance between two probability measures with densities and the Change of Variable Theorem. The latter allows us to reduce the proof to dimension one by observing that the vectors v and $(\|v\|, 0, \ldots, 0)^* \in \mathbb{R}^d$ have the same norm and the statement follows from a straightforward computation. $\square$

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*

$$\lim_{\epsilon \to 0} d_{\mathrm{TV}}(\mathcal{G}(v_\epsilon, \mathrm{I}_d), \mathcal{G}(0, \mathrm{I}_d)) = d_{\mathrm{TV}}(\mathcal{G}(v, \mathrm{I}_d), \mathcal{G}(0, \mathrm{I}_d)).$$

PROOF. The idea of the proof follows from Lemma A.2 together with the dominated convergence theorem. $\square$

LEMMA A.4. *Let $\{v_\epsilon : \epsilon > 0\} \subset \mathbb{R}^d$ such that $\lim_{\epsilon \to 0} \|v_\epsilon\| = +\infty$. Then*

$$\lim_{\epsilon \to 0} d_{\mathrm{TV}}(\mathcal{G}(v_\epsilon, \mathrm{I}_d), \mathcal{G}(0, \mathrm{I}_d)) = 1.$$

PROOF. The idea of the proof follows from Lemma A.2 together with the dominated convergence theorem. $\square$

LEMMA A.5. *Let $\mathcal{S}_d$ denotes the set of $d \times d$ symmetric and positive definite matrices. Let $\{\Xi_\epsilon : \epsilon > 0\} \subset \mathcal{S}_d$ such that $\lim_{\epsilon \to 0} \Xi_\epsilon = \Xi \in \mathcal{S}_d$. Then*

$$\lim_{\epsilon \to 0} d_{\mathrm{TV}}(\mathcal{G}(0, \Xi_\epsilon), \mathcal{G}(0, \Xi)) = 0.$$

PROOF. The proof follows from the characterisation of the total variation distance between two probability measures with densities together with the Scheffé Lemma. An alternative proof can be done using the dominated convergence theorem. $\square$

For $m \in \mathbb{R}$, $\mathcal{N}(m, 1)$ denotes the Gaussian distribution on $\mathbb{R}$ with mean m and unit variance.

LEMMA A.6. *Let $\{v_t : t \ge 0\} \subset \mathbb{R}^d$.*

(i) *If $\limsup_{t \to +\infty} \|v_t\| \le C_0 \in [0, +\infty)$ then*

$$\limsup_{t \to +\infty} d_{\mathrm{TV}}(\mathcal{G}(v_t, \mathrm{I}_d), \mathcal{G}(0, \mathrm{I}_d)) \le d_{\mathrm{TV}}(\mathcal{N}(C_0, 1), \mathcal{N}(0, 1)).$$

(ii) *If* $\liminf_{t\to+\infty} \|v_t\| \geq C_1 \in [0, +\infty)$ *then*

$$\liminf_{t\to+\infty} d_{TV}\big(\mathcal{G}(v_t, I_d), \mathcal{G}(0, I_d)\big) \geq d_{TV}\big(\mathcal{N}(C_1, 1), \mathcal{N}(0, 1)\big).$$

PROOF. From Lemma A.2 we deduce

$$d_{TV}\big(\mathcal{G}(v_t, I_d), \mathcal{G}(0, I_d)\big) = d_{TV}\big(\mathcal{N}(\|v_t\|, 1), \mathcal{N}(0, 1)\big) = \sqrt{\frac{2}{\pi}} \int_0^{\|v_t\|/2} e^{-\frac{x^2}{2}} \, dx$$

which allows us to reduce the proof for $d = 1$. The proof proceeds from the following straight-forward argument: after passing a subsequence, we use the continuity of the total variation distance (Lemma A.3 and Lemma A.5) and the monotonicity property

$$d_{TV}\big(\mathcal{N}(m_1, 1), \mathcal{N}(0, 1)\big) \leq d_{TV}\big(\mathcal{N}(m_2, 1), \mathcal{N}(0, 1)\big)$$

for any $0 \leq |m_1| \leq |m_2| < +\infty$ in order to deduce item (i) and item (ii) of the statement. □

APPENDIX B: THE DETERMINISTIC DYNAMICAL SYSTEM

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

LEMMA B.1. *Assume that* (C) *holds. Then for any* $x_0 \in \mathbb{R}^d \setminus \{0\}$ *there exist* $\lambda := \lambda(x_0) > 0$, $\ell := \ell(x_0)$, $m := m(x_0) \in \{1, \ldots, d\}$, $\theta_1 := \theta_1(x_0), \ldots, \theta_m := \theta_m(x_0) \in [0, 2\pi)$ *and* $v_1 := v_1(x_0), \ldots, v_m := v_m(x_0)$ *in* $\mathbb{C}^d$ *linearly independent such that*

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

PROOF. Write $\Lambda = DF(0)$ and let $t \geq 0$. By (C), all eigenvalues of Λ have positive real parts. Denote by $\{\phi(t, x) : t \geq 0\}$ the solution of the linear system:

$$\begin{cases} \dfrac{d}{dt}\phi(t) = -\Lambda\phi(t) & \text{for } t \geq 0, \\ \phi(0) = x. \end{cases}$$

Let $(w_{j,k} : j = 1, \ldots, N; k = 1, \ldots, N_j)$ be a Jordan basis of $-\Lambda$, that is,

$$-\Lambda w_{j,k} = -\lambda_j w_{j,k} + w_{j,k+1}$$

for any $j = 1, \ldots, N; k = 1, \ldots, N_j$. In this formula we use the convention $w_{j,N_j+1} = 0$. Since $(w_{j,k} : j = 1, \ldots, N; k = 1, \ldots, N_j)$ is a basis the decomposition

$$\phi(t, x) = \sum_{j=1}^N \sum_{k=1}^{N_j} \phi_{j,k}(t, x) w_{j,k}$$

defines the functions $\phi_{j,k}(t, x)$ in a unique way. Then

$$\sum_{j=1}^N \sum_{k=1}^{N_j} \frac{d}{dt}\phi_{j,k}(t, x) w_{j,k} = \sum_{j=1}^N \sum_{k=1}^{N_j} \phi_{j,k}(t, x)(-\lambda_j w_{j,k} + w_{j,k+1}),$$

and the aforementioned uniqueness implies

$$\frac{d}{dt}\phi_{j,k}(t, x) = -\lambda_j \phi_{j,k}(x, t) + \phi_{j,k-1}(t, x)$$

for any $j = 1, \ldots, N; k = 1, \ldots, N_j$, 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

$$x = \sum_{j=1}^{N} \sum_{k=1}^{N_j} x_{j,k} w_{j,k}.$$

For each $j \in \{1, \ldots, N\}$, the system of equations for $\{\phi_{j,k}(t, x) : k = 1, \ldots, N_j\}$ is autonomous, as well as the equation for $\phi_{j,1}(t, x)$. Notice 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

$$\phi_{j,k}(t, x) = x_{j,k} e^{-\lambda_j t} + \int_0^t e^{-\lambda_j(t-s)} \phi_{j,k-1}(s, x) \, ds.$$

Applying this formula for $k = 2$ we see

$$\phi_{j,2}(t, x) = x_{j,2} e^{-\lambda_j t} + x_{j,1} t e^{-\lambda_j t}$$

and from this expression we can guess and check the formula

$$\phi_{j,k}(t, x) = \sum_{i=1}^{k} x_{j,i} \frac{t^{k-i} e^{-\lambda_j t}}{(k-i)!}.$$

Here, we use the convention $0^0 = 1$. We conclude that

(B.1) $$\phi(t, x) = \sum_{j=1}^{N} \sum_{k=1}^{N_j} \sum_{i=1}^{k} \frac{t^{k-i} e^{-\lambda_j t}}{(k-i)!} x_{j,i} w_{j,k}.$$

With this expression in hand, we are ready to prove Lemma B.1. Let $x_0 \in \mathbb{R}^d$ be fixed. Assume that $x_0 \neq 0$ and write

$$x_0 = \sum_{j=1}^{N} \sum_{k=1}^{N_j} x_{j,k}^0 w_{j,k}.$$

Take

$$\lambda = \min\{\mathrm{Re}(\lambda_j) : x_{j,k}^0 \neq 0 \text{ for some } j \in \{1, \ldots, N\} \text{ and } k \in \{1, \ldots, N_j\}\}$$

and define

$$J_0 = \{j \in \{1, \ldots, N\} : \mathrm{Re}(\lambda_j) = \lambda \text{ and } x_{j,k}^0 \neq 0 \text{ for some } k \in \{1, \ldots, N_j\}\}.$$

In other words, we identify in (B.1) the smallest exponential rate of decay and we collect in J_0 all the indices with that exponential decay. Now, define

$$\ell_0 = \max\{N_j - k : j \in J_0 \text{ and } x_{j,k}^0 \neq 0\}$$

and

$$J = \{j \in J_0 : x_{j,N_j-\ell_0}^0 \neq 0\}.$$

We see that for $j \in J$,

$$\lim_{t \to \infty} |\phi_{j,N_j}(t, x_0)| \frac{e^{\lambda t}}{t^{\ell_0}} = \frac{|x_{j,N_j-\ell_0}|}{\ell_0!},$$

while for $j \notin J$ and k arbitrary or $j \in J$ and $k \neq N_j$,

$$\lim_{t\to\infty} |\phi_{j,k}(t, x_0)| \frac{e^{\lambda t}}{t^{\ell_0}} = 0.$$

Therefore,

$$\lim_{t\to\infty} \left\| \frac{e^{\lambda t}}{t^{\ell_0}} \phi(t, x_0) - \sum_{j\in J} \frac{e^{-(\lambda_j-\lambda)t}}{\ell_0!} x_{j,N_j-\ell_0} w_{j,N_j} \right\| = 0.$$

Let $m = \#J$ and let $\sigma : \{1, \ldots, m\} \to J$ be a numbering of J. By definition of λ and J, for any $j \in J$ the numbers $\lambda_j - \lambda$ are imaginary. Therefore, Lemma B.1 is proved choosing $\theta_k = i(\lambda_{\sigma_k} - \lambda)$, $v_k = \frac{x_{\sigma_k,N_{\sigma_k}-\ell_0} w_{\sigma_k,N_{\sigma_k}}}{\ell_0!}$ for any $k \in \{1, \ldots, m\}$ and $\ell = \ell_0 + 1$. $\quad\square$

Now, we are ready to prove Lemma 2.1. The proof is based in the Hartman–Grobman Theorem (see Theorem (Hartman) page 127 of [45] or the celebrated paper of P. Hartman [26]) that guarantees that the conjugation around the hyperbolic fixed point 0 of (2.1) is $\mathcal{C}^1$-local diffeomorphism under some resonance conditions which are fulfilled when all the eigenvalues of the matrix $DF(0)$ have negative (or positive) real parts. Recall that $\{\varphi(t, x_0) : t \geq 0\}$ is the solution of the differential equation (2.1).

LEMMA B.2.   *Assume that* (C) *holds. Then for any* $x_0 \in \mathbb{R}^d \setminus \{0\}$ *there exist* $\lambda := \lambda(x_0) > 0$, $\ell := \ell(x_0)$, $m := m(x_0) \in \{1, \ldots, d\}$, $\theta_1 := \theta_1(x_0), \ldots, \theta_m := \theta_m(x_0) \in [0, 2\pi)$, $v_1 := v_1(x_0), \ldots, v_m := v_m(x_0)$ *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.   Since all the eigenvalues of $DF(0)$ have real positive real parts, there exist open sets U, V of $\mathbb{R}^d$ 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 $x \in U$. From (C) we obtain

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

Observe that there exists $\tau := \tau(x_0) > 0$ such that $\varphi(t, x_0) \in U$ for any $t \geq \tau$. Let $x_\tau = \varphi(\tau, x_0)$, then

$$\varphi(t + \tau, x_0) = \varphi(t, x_\tau) = h^{-1}(e^{-DF(0)t}h(x_\tau)) \quad \text{for any } t \geq 0.$$

Let $\tilde{x} := h(x_\tau)$. By Lemma B.1 there exist $\lambda(\tilde{x}) := \lambda > 0$, $\ell(\tilde{x}) := \ell$, $m(\tilde{x}) := m \in \{1, \ldots, d\}$, $\theta_1(\tilde{x}) := \theta_1, \ldots, \theta_m(\tilde{x}) := \theta_m \in [0, 2\pi)$ and $v_1(\tilde{x}) := v_1, \ldots, v_m(\tilde{x}) := v_m$ in $\mathbb{C}^d$ linearly independent such that

(B.2)
$$\lim_{t\to+\infty} \left\| \frac{e^{\lambda t}}{t^{\ell-1}} e^{-DF(0)t} \tilde{x} - \sum_{k=1}^m e^{i\theta_k t} v_k \right\| = 0.$$

From the triangle inequality we obtain

(B.3)
$$\left\| \frac{e^{\lambda t}}{t^{\ell-1}} \varphi(t + \tau, x_0) - \sum_{k=1}^m e^{i\theta_k t} v_k \right\| \leq \left\| \frac{e^{\lambda t}}{t^{\ell-1}} \varphi(t + \tau, x_0) - \frac{e^{\lambda t}}{t^{\ell-1}} e^{-DF(0)t} \tilde{x} \right\|$$
$$+ \left\| \frac{e^{\lambda t}}{t^{\ell-1}} e^{-DF(0)t} \tilde{x} - \sum_{k=1}^m e^{i\theta_k t} v_k \right\|.$$

Observe that

$$
\frac{e^{\lambda t}}{t^{\ell-1}} \|\varphi(t + \tau, x_0) - e^{-DF(0)t}\tilde{x}\|
$$

$$
= \frac{e^{\lambda t}}{t^{\ell-1}} \|h^{-1}(e^{-DF(0)t}h(x_\tau)) - e^{-DF(0)t}\tilde{x}\|
$$

$$
= \frac{e^{\lambda t}}{t^{\ell-1}} o(\|e^{-DF(0)t}\tilde{x}\|) = \frac{e^{\lambda t}\|e^{-DF(0)t}\tilde{x}\|}{t^{\ell-1}} o(1)
$$

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

where $o(1)$ goes to zero as t goes by. The latter together with inequality (B.3) and relation (B.2) allow us to deduce

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

$\square$

LEMMA B.3.   *Assume that* (C) *holds. Let* $\delta_\epsilon = o(1)$. *Then*

$$
\lim_{\epsilon \to 0} \frac{\delta_\epsilon \|\varphi(t_{\mathrm{mix}}^\epsilon + \delta_\epsilon + cw^\epsilon, x_0)\|^2}{\epsilon} = 0 \quad \text{for any } c \in \mathbb{R}.
$$

PROOF.   Remember that

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

and

$$
w^\epsilon = \frac{1}{\lambda} + o(1),
$$

where λ, ℓ and τ are the constants associated to x_0 in Lemma 2.1 and $o(1)$ goes to zero as $\epsilon \to 0$. Define $t^\epsilon := t_{\mathrm{mix}}^\epsilon - \tau + \delta_\epsilon + cw^\epsilon$. Note

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

$$
+ \frac{(t^\epsilon)^{\ell-1}}{e^{\lambda t^\epsilon}\sqrt{\epsilon}} \sum_{k=1}^{m} \|v_k\|.
$$

From the last inequality, using the fact that $\lim_{\epsilon \to 0} \frac{(t^\epsilon)^{\ell-1}}{e^{\lambda t^\epsilon}\sqrt{\epsilon}} = \frac{e^{-c}}{(2\lambda)^{\ell-1}}$ and Lemma B.2 we deduce the desired result. $\square$

APPENDIX C: THE STOCHASTIC DYNAMICAL SYSTEM

In this Appendix we analyse the zeroth and first order approximations for the Itô diffusion $\{x^\epsilon(t) : t \geq 0\}$ given by (2.4). Recall that $\{\varphi(t) : t \geq 0\}$ is the solution of the differential equation (2.1), $\{y(t) : t \geq 0\}$ is the solution of the stochastic differential equation (3.5), and $\delta > 0$ is the constant that appears in (H).

LEMMA C.1. *Assume that* (H) *holds. For any* $\eta > 0$ *and* $t \in [0, \frac{\eta^2}{\epsilon d})$ *we have*

$$\mathbb{P}\Big(\sup_{0 \leq s \leq t} \|x^\epsilon(t) - \varphi(t)\| \geq \eta\Big) \leq \frac{2d\epsilon^2 t}{\delta(\eta^2 - \epsilon dt)^2}$$

and

$$\mathbb{P}\Big(\sup_{0 \leq s \leq t} \|\sqrt{\epsilon} y(t)\| \geq \eta\Big) \leq \frac{2d\epsilon^2 t}{\delta(\eta^2 - \epsilon dt)^2}.$$

PROOF. Let $\epsilon > 0$ and $t \geq 0$ be fixed. From (3.3) we have

(C.1)
$$\begin{aligned} \mathrm{d}\|x^\epsilon(t) - \varphi(t)\|^2 &\leq -2\delta \|x^\epsilon(t) - \varphi(t)\|^2 \, \mathrm{d}t + 2\sqrt{\epsilon}\langle (x^\epsilon(t) - \varphi(t)), \mathrm{d}B(t)\rangle \\ &\quad + d\epsilon \, \mathrm{d}t. \end{aligned}$$

Let $M^\epsilon(t) := 2\sqrt{\epsilon}(x^\epsilon(t) - \varphi(t))^*$ for every $t \geq 0$. Notice that

$$\Big\{N^\epsilon(t) := \int_0^t M^\epsilon(s) \, \mathrm{d}B(s) : t \geq 0\Big\} \quad \text{is a local martingale.}$$

Then, there exists a sequence of increasing stopping times $\{\tau_n^\epsilon\}_{n \in \mathbb{N}}$ such that almost surely $\tau_n^\epsilon \uparrow \infty$ as n goes to infinity and for each $n \in \mathbb{N}$,

$$\{N^{\epsilon,n}(t) = N^\epsilon(\min\{\tau_n^\epsilon, t\}) : t \geq 0\} \quad \text{is a true martingale.}$$

Taking expectation on (C.1) and using the fact that $\{N^{\epsilon,n}(t) : t \geq 0\}$ is a zero-mean martingale, we deduce

$$\mathbb{E}\big[\|x^\epsilon(\min\{\tau_n^\epsilon, t\}) - \varphi(\min\{\tau_n^\epsilon, t\})\|^2\big] \leq \epsilon d \min\{\tau_n^\epsilon, t\} \leq \epsilon dt$$

for every $t \geq 0$. Consequently, by the well-known Fatou lemma we obtain

$$\mathbb{E}\big[\|x^\epsilon(t) - \varphi(t)\|^2\big] \leq \epsilon dt \quad \text{for any } t \geq 0.$$

The latter implies

$$\Big\{N^\epsilon(t) = \int_0^t M^\epsilon(s) \, \mathrm{d}B(s) : t \geq 0\Big\} \quad \text{is a true martingale.}$$

From inequality (C.1) we have

$$\|x^\epsilon(t) - \varphi(t)\|^2 \leq \epsilon dt + N^\epsilon(t) \quad \text{for any } t \geq 0.$$

For any $\eta > 0$ and $0 \leq t < \eta^2/(\epsilon d)$ we have

$$\mathbb{P}\Big(\sup_{0 \leq s \leq t} \|x^\epsilon(s) - \varphi(s)\|^2 \geq \eta^2\Big) \leq \mathbb{P}\Big(\sup_{0 \leq s \leq t} \|N^\epsilon(s)\| \geq \eta^2 - \epsilon dt\Big).$$

From the Doob inequality for submartingales we obtain

$$\mathbb{P}\Big(\sup_{0 \leq s \leq t} \|N^\epsilon(s)\| \geq \eta^2 - \epsilon dt\Big) \leq \frac{\mathbb{E}[\|N^\epsilon(t)\|^2]}{(\eta^2 - \epsilon dt)^2}.$$

The Itô isometry allows us to deduce that

$$\mathbb{E}\big[\|N^\epsilon(t)\|^2\big] = 4\epsilon \int_0^t \mathbb{E}\big[\|x^\epsilon(s) - \varphi(s)\|^2\big] \, \mathrm{d}s.$$

From inequality (3.4) we obtain $\mathbb{E}[\|N^\epsilon(t)\|^2] \leq 2d\epsilon^2 t/\delta$. Therefore

$$\mathbb{P}\Big(\sup_{0 \leq s \leq t} \|x^\epsilon(s) - \varphi(s)\| \geq \eta\Big) \leq \frac{2d\epsilon^2 t}{\delta(\eta^2 - \epsilon dt)^2}$$

for $0 \leq t < \eta^2/(\epsilon d)$. The proof for the second part proceeds from the same ideas as the first part. We left the details to the interested reader. □

PROPOSITION C.2. *Assume that* (H) *holds. For* $t \geq 0$, *write* $W(t) := \sup_{0 \leq s \leq t} \|B(s)\|$. *For any* $t \geq 0$, *the following hold true*:

(i) $\mathbb{E}[\|x^{\epsilon}(t) - \varphi(t)\|^{2}] \leq \frac{d\epsilon}{2\delta}$ *and* $\mathbb{E}[\|y(t)\|^{2}] \leq \frac{d}{2\delta}$.

(ii) *For each* $n \in \mathbb{N}$, *define* $c_{n} := \prod_{j=0}^{n-1}(d + 2j)$. *Then*

$$\mathbb{E}[\|x^{\epsilon}(t) - \varphi(t)\|^{2n}] \leq \frac{c_{n}\epsilon^{n}}{2^{n}\delta^{n}} \quad \text{and} \quad \mathbb{E}[\|y(t)\|^{2n}] \leq \frac{c_{n}}{2^{n}\delta^{n}}.$$

(iii) *For any* $0 \leq r < \delta$ *we have*

$$\mathbb{E}\left[\exp\left(\frac{r\|x^{\epsilon}(t) - \varphi(t)\|^{2}}{\epsilon}\right)\right] < +\infty$$

and

$$\mathbb{E}[\exp(r\|y(t)\|^{2})] < +\infty.$$

(iv) *Let* $r \in (0, \delta/2]$. *Then*

$$\mathbb{E}\left[\exp\left(\frac{r\|x^{\epsilon}(t) - \varphi(t)\|^{2}}{\epsilon}\right)\right] \leq \exp(drt)$$

and

$$\mathbb{E}[\exp(r\|y(t)\|^{2})] \leq \exp(drt).$$

PROOF. (i) The first part follows from inequality (3.4). The second part follows exactly as inequality (3.4). We left the details to the interested reader.

(ii) We provide the proof for the first part. The second part proceeds exactly as the first part and we left the details to the interested reader.

Let $\epsilon > 0$ and $t \geq 0$ be fixed. Notice that

$$x^{\epsilon}(t) - \varphi(t)$$

$$= -\int_{0}^{t}\left[F(x^{\epsilon}(s)) - F(\varphi(s))\right]ds + \sqrt{\epsilon}B(t)$$

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

$$= -\int_{0}^{t}A^{\epsilon}(s)(x^{\epsilon}(s) - \varphi(s))ds + \sqrt{\epsilon}B(t),$$

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 basis had already proved in item (i) of this proposition. Consider $f_{n+1}(x) = \|x\|^{2(n+1)}$, $x \in \mathbb{R}^{d}$. By the Itô formula, it follows that

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

$$= -2(n+1)\|x^{\epsilon}(t) - \varphi(t)\|^{2n}\langle x^{\epsilon}(t) - \varphi(t), A^{\epsilon}(t)(x^{\epsilon}(t) - \varphi(t))\rangle dt$$

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

$$+ 2(n+1)\sqrt{\epsilon}\|x^{\epsilon}(t) - \varphi(t)\|^{2n}\langle x^{\epsilon}(t) - \varphi(t), dB(t)\rangle.$$

From (H) we obtain

$$d\|x^{\epsilon}(t) - \varphi(t)\|^{2(n+1)} \leq -2\delta(n+1)\|x^{\epsilon}(t) - \varphi(t)\|^{2(n+1)}dt$$

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

$$+ 2(n+1)\sqrt{\epsilon}\|x^{\epsilon}(t) - \varphi(t)\|^{2n}\langle x^{\epsilon}(t) - \varphi(t), dB(t)\rangle.$$

After a localisation argument, we can take expectation in both sides of the last differential inequality and deduce that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big[\big\|x^\epsilon(t) - \varphi(t)\big\|^{2(n+1)}\big] \leq -2\delta(n+1)\mathbb{E}\big[\big\|x^\epsilon(t) - \varphi(t)\big\|^{2(n+1)}\big]$$

$$+ \epsilon(d+2n)(n+1)\mathbb{E}\big[\big\|x^\epsilon(t) - \varphi(t)\big\|^{2n}\big].$$

By the induction hypothesis we have

$$\mathbb{E}\big[\big\|x^\epsilon(t) - \varphi(t)\big\|^{2n}\big] \leq \frac{c_n\epsilon^n}{2^n\delta^n} \quad \text{for any } t \geq 0.$$

Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big[\big\|x^\epsilon(t) - \varphi(t)\big\|^{2(n+1)}\big] \leq -2\delta(n+1)\mathbb{E}\big[\big\|x^\epsilon(t) - \varphi(t)\big\|^{2(n+1)}\big]$$

$$+ (n+1)\frac{c_{n+1}\epsilon^{n+1}}{2^n\delta^n}.$$

From Lemma C.7 we obtain

$$\mathbb{E}\big[\big\|x^\epsilon(t) - \varphi(t)\big\|^{2(n+1)}\big] \leq \frac{c_{n+1}\epsilon^{n+1}}{2^{n+1}\delta^{n+1}} \quad \text{for any } t \geq 0.$$

(iii) We provide the proof for the first part. The second part follows exactly as the first part and again we left the details to the interested reader.

Let $\epsilon > 0$ and $t \geq 0$ be fixed. By the monotone convergence theorem it follows that

$$\mathbb{E}\big[e^{\frac{r\|x^\epsilon(t)-\varphi(t)\|^2}{\epsilon}}\big] = \sum_{n=0}^\infty \mathbb{E}\Big[\frac{r^n\|x^\epsilon(t) - \varphi(t)\|^{2n}}{\epsilon^n n!}\Big].$$

By item (i) of this Proposition, we have

$$\sum_{n=0}^\infty \mathbb{E}\Big[\frac{r^n\|x^\epsilon(t) - \varphi(t)\|^{2n}}{\epsilon^n n!}\Big] \leq 1 + \sum_{n=1}^\infty \frac{r^n c_n}{2^n \delta^n n!}.$$

Since $\sum_{n=1}^\infty \frac{r^n c_n}{2^n\delta^n n!} < +\infty$ when $0 \leq r < \delta$, then we deduce the statement.

(iv) We give the proof for the first part. The second part proceeds exactly as the first part and again we left the details to the interested reader.

Let $\epsilon > 0$ and $t \geq 0$ be fixed. We will use the Itô formula for the function $g_\epsilon(x) = e^{\kappa_\epsilon\|x\|^2}$, $x \in \mathbb{R}^d$, where $\kappa_\epsilon := \frac{r}{\epsilon}$. Then

$$\mathrm{d}e^{\kappa_\epsilon\|x^\epsilon(t)-\varphi(t)\|^2}$$

$$= -2\kappa_\epsilon e^{\kappa_\epsilon\|x^\epsilon(t)-\varphi(t)\|^2}\big\langle A^\epsilon(t)\big(x^\epsilon(t) - \varphi(t)\big), x^\epsilon(t) - \varphi(t)\big\rangle\mathrm{d}t$$

$$+ \epsilon\big(2\kappa_\epsilon^2 e^{\kappa_\epsilon\|x^\epsilon(t)-\varphi(t)\|^2}\big\|x^\epsilon(t) - \varphi(t)\big\|^2 + \kappa_\epsilon d e^{\kappa_\epsilon\|x^\epsilon(t)-\varphi(t)\|^2}\big)\mathrm{d}t$$

$$+ 2d\sqrt{\epsilon}\kappa_\epsilon e^{\kappa_\epsilon\|x^\epsilon(t)-\varphi(t)\|^2}\big\langle x^\epsilon(t) - \varphi(t), \mathrm{d}B(t)\big\rangle.$$

Using (H) we obtain

$$\mathrm{d}e^{\kappa_\epsilon\|x^\epsilon(t)-\varphi(t)\|^2}$$

$$\leq -2\kappa_\epsilon\delta e^{\kappa_\epsilon\|x^\epsilon(t)-\varphi(t)\|^2}\big\|x^\epsilon(t) - \varphi(t)\big\|^2\mathrm{d}t$$

$$+ \epsilon\big(2\kappa_\epsilon^2 e^{\kappa_\epsilon\|x^\epsilon(t)-\varphi(t)\|^2}\big\|x^\epsilon(t) - \varphi(t)\big\|^2 + \kappa_\epsilon d e^{\kappa_\epsilon\|x^\epsilon(t)-\varphi(t)\|^2}\big)\mathrm{d}t$$

$$+ 2d\sqrt{\epsilon}\kappa_\epsilon e^{\kappa_\epsilon\|x^\epsilon(t)-\varphi(t)\|^2}\big\langle x^\epsilon(t) - \varphi(t), \mathrm{d}B(t)\big\rangle.$$

Since $0 < r \le \frac{\delta}{2}$ then

$$\mathrm{d}e^{\kappa_\epsilon \|x^\epsilon(t) - \varphi(t)\|^2}$$

$$\le -\kappa_\epsilon \delta e^{\kappa_\epsilon \|x^\epsilon(t) - \varphi(t)\|^2} \|x^\epsilon(t) - \varphi(t)\|^2 \, \mathrm{d}t$$

$$+ \epsilon\kappa_\epsilon d e^{\kappa_\epsilon \|x^\epsilon(t) - \varphi(t)\|^2} \, \mathrm{d}t + 2d\sqrt{\epsilon}\kappa_\epsilon e^{\kappa_\epsilon \|x^\epsilon(t) - \varphi(t)\|^2} \langle x^\epsilon(t) - \varphi(t), \mathrm{d}B(t)\rangle.$$

By item (i) and item (ii) of this proposition and using a localisation argument we deduce

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big[e^{\kappa_\epsilon \|x^\epsilon(t) - \varphi(t)\|^2}\big] \le \epsilon\kappa_\epsilon d\mathbb{E}\big[e^{\kappa_\epsilon \|x^\epsilon(t) - \varphi(t)\|^2}\big] \quad \text{for any } t \ge 0.$$

Now, using Lemma C.7 we obtain

$$\mathbb{E}\big[e^{\frac{r\|x^\epsilon(t) - \varphi(t)\|^2}{\epsilon}}\big] \le e^{drt} \quad \text{for any } t \ge 0. \qquad \square$$

LEMMA C.3 (Uniquely ergodic). *Assume that* (C) *holds. For any* $\epsilon \in (0, 1]$ *there exists a unique invariant measure* μ^ϵ *for the dynamics* (2.4). *The unique probability invariant measure* μ^ϵ *has exponential moments*

$$\int_{\mathbb{R}^d} e^{\beta\|y\|} \mu^\epsilon(\mathrm{d}y) < +\infty \quad \text{for any } \beta \ge 0.$$

In addition, for any $\beta > 0$ *there exist positive constants* $C_1^{\epsilon,\beta}$ *and* $C_2^{\epsilon,\beta}$ *such that for any initial condition* $x_0 \in \mathbb{R}^d$ *we have*

$$d_{\mathrm{TV}}\big(x^\epsilon(t, x_0), \mu^\epsilon\big) \le C_1^{\epsilon,\beta} e^{-tC_2^{\epsilon,\beta}} \left(e^{\beta\|x_0\|} + \int_{\mathbb{R}^d} e^{\beta\|y\|} \mu^\epsilon(\mathrm{d}y)\right)$$

for any $t \ge 0$. *In particular,*

$$\lim_{t \to +\infty} d_{\mathrm{TV}}\big(x^\epsilon(t, x_0), \mu^\epsilon\big) = 0.$$

PROOF. This follows immediately from Theorem 3.3.4 page 91 of [33]. $\square$

LEMMA C.4. *Assume that* (C) *holds. Consider the matrix differential equation*:

(C.2)
$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\Sigma(t) = -DF(0)\Sigma(t) - \Sigma(t)\big(DF(0)\big)^* + \mathrm{I}_d & \text{for } t \ge 0, \\ \Sigma(0) = \Sigma_0, \end{cases}$$

where Σ_0 *is any* $d \times d$ *matrix. Then*

$$\big\|\Sigma(t) - \Sigma\big\| \le e^{-2\delta t}\|\Sigma_0 - \Sigma\| \quad \text{for any } t \ge 0,$$

where $\delta > 0$ *is the constant that appears in* (C) *and* Σ *is the unique solution of the Lyapunov matrix equation*

(C.3)
$$DF(0)X + X\big(DF(0)\big)^* = \mathrm{I}_d.$$

PROOF. Write $\Lambda = DF(0)$ and let $t \ge 0$. Notice that all eigenvalues of Λ have positive real parts. Denote by $\{\phi(t, x) : t \ge 0\}$ the solution of the linear system:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\phi(t) = -\Lambda\phi(t) & \text{for } t \ge 0, \\ \phi(0) = x. \end{cases}$$

Then $\{\phi(t, x) : t \geq 0\}$ is globally asymptotic stable and consequently the Lyapunov matrix equation (C.3) has a unique solution Σ which is symmetric and positive definite. For more details, see Theorem 1, page 443 of [35]. From (C.3) it follows that Σ is a symmetric matrix. Let

$$r(t) := \left\| \Sigma(t) - \Sigma \right\|^2 = \sum_{i,j=1}^{d} \left(\Sigma_{i,j}(t) - \Sigma_{i,j} \right)^2 \quad \text{for any } t \geq 0.$$

Let $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ if $i \neq j$. Notice that

$$\sum_{k=1}^{d} \Lambda_{i,k} \Sigma_{k,j} + \sum_{k=1}^{d} \Sigma_{i,k} \Lambda_{j,k} = \delta_{i,j} \quad \text{for any } i, j \in \{1, \ldots, d\}.$$

Then

$$\frac{d}{dt} r(t) = 2 \sum_{i,j=1}^{d} \left(\Sigma_{i,j}(t) - \Sigma_{i,j} \right) \frac{d}{dt} \Sigma_{i,j}(t)$$

$$= 2 \sum_{i,j=1}^{d} \left(\Sigma_{i,j}(t) - \Sigma_{i,j} \right) \left(- \sum_{k=1}^{d} \Lambda_{i,k} \left(\Sigma_{k,j}(t) - \Sigma_{k,j} \right) \right.$$

$$\left. - \sum_{k=1}^{d} \left(\Sigma_{i,k}(t) - \Sigma_{i,k} \right) \Lambda_{j,k} \right).$$

After rearrangement the sums and using (C) we have

$$\frac{d}{dt} r(t) \leq -4\delta r(t) \quad \text{for } t \geq 0,$$

$$r(0) = \left\| \Sigma_0 - \Sigma \right\|^2.$$

By Lemma C.7 we deduce

$$\left\| \Sigma(t) - \Sigma \right\|^2 \leq e^{-4\delta t} \left\| \Sigma_0 - \Sigma \right\|^2 \quad \text{for any } t \geq 0$$

which implies the statement. $\square$

REMARK C.5.   If we take $F(x) = Ax$, $x \in \mathbb{R}^d$, in the stochastic differential equation (2.4), the covariance matrix associated to the solution of (2.4) satisfies the matrix differential equation (C.2) with initial datum Σ_0 the zero matrix of dimension $d \times d$.

LEMMA C.6.   *Assume that* (H) *holds. The covariance matrix of $y(t)$ converge as $t \to +\infty$ to a nondegenerate covariance matrix Σ, where Σ 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

$$\begin{cases} \dfrac{d}{dt} \Lambda(t) = -DF(\varphi(t))\Lambda(t) - \Lambda(t)(DF(\varphi(t)))^* + I_d & \text{for } t \geq 0, \\ \Lambda(0) = 0. \end{cases}$$

For further details, see item b) of Theorem 3.2, page 97 of [41].

Let $\mathcal{K}_{x_0} := \{x \in \mathbb{R}^d : \|x\| \leq \|x_0\|\}$. By (H) we have $\varphi(x,t) \in \mathcal{K}_{x_0}$ for any $x \in \mathcal{K}_{x_0}$ and $t \geq 0$. Since $F \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}^d)$ there exists a constant $L := L(\|x_0\|) > 0$ such that

$$\|DF(x) - DF(0)\| \leq L\|x\| \quad \text{for any } x \in \mathcal{K}_{x_0}.$$

Take $\eta \in (0, \|x_0\|)$ and $\tau_\eta := \frac{1}{\delta} \ln(\frac{\|x_0\|}{\eta})$ such that

$$(C.4) \qquad \|DF(\varphi(t)) - DF(0)\| \leq L\|\varphi(t)\| \leq L\|x_0\|e^{-\delta t} \leq L\eta$$

for every $t \geq \tau_\eta$. Call $\tau := \tau_\eta$. Then,

$$(C.5) \qquad \begin{cases} \dfrac{d}{dt}\Delta(t) = -DF(0)\Delta(t) - \Delta(t)(DF(0))^* + I_d & \text{for } t \geq 0, \\ \Delta(0) = \Lambda(\tau). \end{cases}$$

Let $\Pi(t) = \Lambda(t+\tau) - \Delta(t)$, $t \geq 0$. Then

$$\begin{cases} \dfrac{d}{dt}\Pi(t) = -DF(\varphi(t+\tau))\Pi(t) - \Pi(t)(DF(\varphi(t+\tau)))^* + g(t,\tau) & \text{for } t \geq 0, \\ \Pi(0) = 0, \end{cases}$$

where $g(t,\tau) := (DF(0) - DF(\varphi(t+\tau)))\Delta(t) + \Delta(t)(DF(0) - DF(\varphi(t+\tau)))^*$ for $t \geq 0$. Therefore

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

where

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

After rearrangement, from (H) we deduce

$$(C.6) \qquad \frac{d}{dt}\|\Pi(t)\|^2 \leq -4\delta\|\Pi(t)\|^2 + 2\sum_{i,j=1}^d |\Pi_{i,j}(t)R_{i,j}(t)| \quad \text{for any } t \geq 0.$$

Moreover, using the inequality $2|xy| \leq \rho x^2 + \frac{y^2}{\rho}$ for any $\rho > 0$ and $x, y \in \mathbb{R}$ we deduce

$$(C.7) \qquad 2\sum_{i,j=1}^d |\Pi_{i,j}(t)R_{i,j}(t)| \leq \delta\|\Pi(t)\|^2 + \frac{1}{\delta}\sum_{i,j=1}^d |R_{i,j}(t)|^2 \quad \text{for any } t \geq 0.$$

Using Lipschitz condition (C.4), for any $i, j \in \{1, \ldots, d\}$ we get

$$(C.8) \qquad \left|R_{i,j}(t)\right| \leq L\eta \sum_{k=1}^{d} \left(\left|\Delta_{k,j}(t)\right| + \left|\Delta_{i,k}(t)\right|\right) \leq 2dL\eta \|\Delta(t)\| \quad \text{for any } t \geq 0.$$

Recall that $\{\Delta(t) : t \geq 0\}$ satisfies (C.5). From Lemma C.4 we have

$$(C.9) \qquad \|\Delta(t) - \Sigma\| \leq e^{-2\delta t} \|\Lambda(\tau) - \Sigma\| \quad \text{for any } t \geq 0,$$

where the matrix Σ satisfies $DF(0)\Sigma + \Sigma(DF(0))^* = I_d$. Similar computations using in inequalities (C.6), (C.7) and (C.8) allows us to deduce that

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\Lambda(t) - \Sigma\|^2 \leq -3\delta \|\Lambda(t) - \Sigma\|^2 + \frac{C(\|x_0\|)\tilde{C}(d)}{\delta} \quad \text{for any } t \geq 0,$$

where

$$C(\|x_0\|) = \sup_{x \in \mathcal{K}_{x_0}} \|DF(x) - DF(0)\|^2$$

and

$$\tilde{C}(d) = \sum_{i,j=1}^{d} \left(\sum_{k=1}^{d} |\Sigma_{i,k}| + |\Sigma_{k,j}|\right)^2.$$

From Lemma C.7 we obtain that there exists a positive constant $C_0 := C_0(\|x_0\|, \delta, d)$ such that

$$(C.10) \qquad \|\Lambda(t)\| \leq C_0 \quad \text{for any } t \geq 0.$$

The latter together with inequality (C.9) imply that there exists a positive constant $C_1 := C_1(\|x_0\|, \delta, d)$ such that $\|\Delta(t)\| \leq C_1$ for any $t \geq 0$. From inequalities (C.6), (C.7) and (C.8) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\Pi(t)\|^2 \leq -3\delta \|\Pi(t)\|^2 + C_2 \eta^2 \quad \text{for any } t \geq 0,$$

where $C_2 = 4L^2 C_1^2 d^4 / \delta$. Then Lemma C.7 implies

$$(C.11) \qquad \|\Pi(t)\|^2 \leq C_2 \eta^2 \quad \text{for any } t \geq 0.$$

Now, we are ready to get the statement. Let $t \geq \tau$. From the triangle inequality and inequality (C.11)

$$\begin{aligned}
\|\Lambda(t) - \Sigma\| &= \|\Lambda((t-\tau) + \tau) - \Sigma\| \\
&\leq \|\Lambda((t-\tau) + \tau) - \Delta(t-\tau)\| + \|\Delta(t-\tau) - \Sigma\| \\
&= \|\Pi(t-\tau)\| + \|\Delta(t-\tau) - \Sigma\| \\
&\leq \sqrt{C_2}\eta + \|\Delta(t-\tau) - \Sigma\|.
\end{aligned}$$

Letting $t \to \infty$ and using Lemma C.4 we obtain

$$\limsup_{t \to \infty} \|\Lambda(t) - \Sigma\| \leq \sqrt{C_2}\eta.$$

Now, letting $\eta \to 0$ we deduce $\lim_{t \to \infty} \|\Lambda(t) - \Sigma\| = 0$. $\quad \square$

LEMMA C.7 (Gronwall inequality).   *Let $T > 0$ be fixed. Let $g : [0, T] \to \mathbb{R}$ be a $\mathcal{C}^1$-function and $h : [0, T] \to \mathbb{R}$ be a $\mathcal{C}^0$ function. If*

$$\frac{\mathrm{d}}{\mathrm{d}t} g(t) \leq -ag(t) + h(t) \quad \textit{for any } t \in [0, T],$$

where $a \in \mathbb{R}$, and the derivative at 0 and T are understanding as the right and left derivatives, respectively. Then

$$g(t) \leq e^{-at} g(0) + e^{-at} \int_0^t e^{as} h(s) \, \mathrm{d}s \quad \textit{for any } t \in [0, T].$$

Moreover, if $a \neq 0$ we have

$$\left| g(t) \right| \leq e^{-at} \left| g(0) \right| + \frac{(1 - e^{-at})}{a} \max_{s \in [0,t]} \left| h(s) \right| \quad \textit{for any } t \in [0, T].$$

REFERENCES

[1] ALDOUS, D. and DIACONIS, P. (1986). Shuffling cards and stopping times. *Amer. Math. Monthly* **93** 333–348. MR0841111 https://doi.org/10.2307/2323590

[2] ALDOUS, D. and DIACONIS, P. (1987). Strong uniform times and finite random walks. *Adv. in Appl. Math.* **8** 69–97. MR0876954 https://doi.org/10.1016/0196-8858(87)90006-6

[3] ARNOLD, A., CARRILLO, J. A. and MANZINI, C. (2010). Refined long-time asymptotics for some polymeric fluid flow models. *Commun. Math. Sci.* **8** 763–782. MR2730330

[4] ARNOLD, L. (2003). Linear and nonlinear diffusion approximation of the slow motion in systems with two time scales. In *IUTAM Symposium on Nonlinear Stochastic Dynamics*. *Solid Mech. Appl.* **110** 5–18. Kluwer Academic, Dordrecht. MR2049394 https://doi.org/10.1007/978-94-010-0179-3_1

[5] ATHREYA, K. B. and HWANG, C.-R. (2010). Gibbs measures asymptotics. *Sankhya A* **72** 191–207. MR2658170 https://doi.org/10.1007/s13171-010-0006-5

[6] BAEK, Y. and KAFRI, Y. (2015). Singularities in large deviation functions. *J. Stat. Mech. Theory Exp.* **8** P08026, 31. MR3400252 https://doi.org/10.1088/1742-5468/2015/08/p08026

[7] BAKHTIN, Y. (2010). Small noise limit for diffusions near heteroclinic networks. *Dyn. Syst.* **25** 413–431. MR2731621 https://doi.org/10.1080/14689367.2010.482520

[8] BAKHTIN, Y. (2011). Noisy heteroclinic networks. *Probab. Theory Related Fields* **150** 1–42. MR2800902 https://doi.org/10.1007/s00440-010-0264-0

[9] BARRERA, G. (2018). Abrupt convergence for a family of Ornstein–Uhlenbeck processes. *Braz. J. Probab. Stat.* **32** 188–199. MR3770869 https://doi.org/10.1214/16-BJPS337

[10] BARRERA, G. and JARA, M. (2016). Abrupt convergence for stochastic small perturbations of one dimensional dynamical systems. *J. Stat. Phys.* **163** 113–138. MR3472096 https://doi.org/10.1007/s10955-016-1468-1

[11] BARRERA, J., BERTONCINI, O. and FERNÁNDEZ, R. (2009). Abrupt convergence and escape behavior for birth and death chains. *J. Stat. Phys.* **137** 595–623. MR2565098 https://doi.org/10.1007/s10955-009-9861-7

[12] BARRERA, J. and YCART, B. (2014). Bounds for left and right window cutoffs. *ALEA Lat. Am. J. Probab. Math. Stat.* **11** 445–458. MR3265085

[13] BOGACHEV, V. I., RÖCKNER, M. and SHAPOSHNIKOV, S. V. (2016). Distances between transition probabilities of diffusions and applications to nonlinear Fokker–Planck–Kolmogorov equations. *J. Funct. Anal.* **271** 1262–1300. MR3522009 https://doi.org/10.1016/j.jfa.2016.05.016

[14] BOUCHET, F. and REYGNER, J. (2016). Generalisation of the Eyring–Kramers transition rate formula to irreversible diffusion processes. *Ann. Henri Poincaré* **17** 3499–3532. MR3568024 https://doi.org/10.1007/s00023-016-0507-4

[15] BOVIER, A., ECKHOFF, M., GAYRARD, V. and KLEIN, M. (2004). Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc.* (*JEMS*) **6** 399–424. MR2094397 https://doi.org/10.4171/JEMS/14

[16] BOVIER, A., GAYRARD, V. and KLEIN, M. (2005). Metastability in reversible diffusion processes. II. Precise asymptotics for small eigenvalues. *J. Eur. Math. Soc.* (*JEMS*) **7** 69–99. MR2120991 https://doi.org/10.4171/JEMS/22

[17] CHEN, G.-Y. and SALOFF-COSTE, L. (2008). The cutoff phenomenon for ergodic Markov processes. *Electron. J. Probab.* **13** 26–78. MR2375599 https://doi.org/10.1214/EJP.v13-474

[18] DAY, M. (1982). Exponential leveling for stochastically perturbed dynamical systems. *SIAM J. Math. Anal.* **13** 532–540. MR0661588 https://doi.org/10.1137/0513035

[19] DAY, M. V. (1983). On the exponential exit law in the small parameter exit problem. *Stochastics* **8** 297–323. MR0693886 https://doi.org/10.1080/17442508308833244

[20] DIACONIS, P. (1996). The cutoff phenomenon in finite Markov chains. *Proc. Natl. Acad. Sci. USA* **93** 1659–1664. MR1374011 https://doi.org/10.1073/pnas.93.4.1659

[21] EBERLE, A. and ZIMMER, R. Sticky couplings of multidimensional diffusions with different drifts. *Ann. Inst. Henri Poincaré B, Calc. Probab. Stat.*. To appear. (N.S.) Available at arXiv:1612.06125.

[22] FREIDLIN, M. and WENTZELL, A. (1970). On small random perturbations of dynamical systems. *Russian Math. Surveys* **25** 1–55. MR0267221

[23] FREIDLIN, M. and WENTZELL, A. (1972). Some problems concerning stability under small random perturbations. *Theory Probab. Appl.* **17** 269–283.

[24] FREIDLIN, M. I. and WENTZELL, A. D. (2012). *Random Perturbations of Dynamical Systems*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **260**. Springer, Heidelberg. Translated from the 1979 Russian original by Joseph Szücs. MR2953753 https://doi.org/10.1007/978-3-642-25847-3

[25] GALVES, A., OLIVIERI, E. and VARES, M. E. (1987). Metastability for a class of dynamical systems subject to small random perturbations. *Ann. Probab.* **15** 1288–1305. MR0905332

[26] HARTMAN, P. (1960). On local homeomorphisms of Euclidean spaces. *Bol. Soc. Mat. Mexicana* (2) **5** 220–241. MR0141856

[27] HWANG, C.-R. (1980). Laplace's method revisited: Weak convergence of probability measures. *Ann. Probab.* **8** 1177–1182. MR0602391

[28] HWANG, C.-R., HWANG-MA, S.-Y. and SHEU, S. J. (1993). Accelerating Gaussian diffusions. *Ann. Appl. Probab.* **3** 897–913. MR1233633

[29] JACQUOT, S. (1992). Comportement asymptotique de la seconde valeur propre des processus de Kolmogorov. *J. Multivariate Anal.* **40** 335–347. MR1150617 https://doi.org/10.1016/0047-259X(92)90030-J

[30] JOURDAIN, B., LE BRIS, C., LELIÈVRE, T. and OTTO, F. (2006). Long-time asymptotics of a multiscale model for polymeric fluid flows. *Arch. Ration. Mech. Anal.* **181** 97–148. MR2221204 https://doi.org/10.1007/s00205-005-0411-4

[31] KABANOV, YU. M., LIPTSER, R. SH. and SHIRYAEV, A. N. (1986). On the variation distance for probability measures defined on a filtered space. *Probab. Theory Related Fields* **71** 19–35. MR0814659 https://doi.org/10.1007/BF00366270

[32] KIPNIS, C. and NEWMAN, C. M. (1985). The metastable behavior of infrequently observed, weakly random, one-dimensional diffusion processes. *SIAM J. Appl. Math.* **45** 972–982. MR0813459 https://doi.org/10.1137/0145059

[33] KULIK, A. (2018). *Ergodic Behavior of Markov Processes: With Applications to Limit Theorems. De Gruyter Studies in Mathematics* **67**. de Gruyter, Berlin. MR3791835

[34] LACHAUD, B. (2005). Cut-off and hitting times of a sample of Ornstein–Uhlenbeck processes and its average. *J. Appl. Probab.* **42** 1069–1080. MR2203823 https://doi.org/10.1239/jap/1134587817

[35] LANCASTER, P. and TISMENETSKY, M. (1985). *The Theory of Matrices*, 2nd ed. *Computer Science and Applied Mathematics*. Academic Press, Orlando, FL. MR0792300

[36] LELIÈVRE, T., NIER, F. and PAVLIOTIS, G. A. (2013). Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion. *J. Stat. Phys.* **152** 237–274. MR3082649 https://doi.org/10.1007/s10955-013-0769-x

[37] LEVIN, D. A., PERES, Y. and WILMER, E. L. (2009). *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI. With a chapter by James G. Propp and David B. Wilson. MR2466937

[38] MAIER, R. and STEIN, D. (1993). Escape problem of irreversible systems. *Phys. Rev. E* **48** 931–938.

[39] MAIER, R. and STEIN, D. (1993). Transition-rate theory for nongradient drift fields. *Phys. Rev. Lett.* **69** 3691–3695.

[40] MAIER, R. S. and STEIN, D. L. (1997). Limiting exit location distributions in the stochastic exit problem. *SIAM J. Appl. Math.* **57** 752–790. MR1450848 https://doi.org/10.1137/S0036139994271753

[41] MAO, X. (2008). *Stochastic Differential Equations and Applications*, 2nd ed. Horwood Publishing Limited, Chichester. MR2380366 https://doi.org/10.1533/9780857099402

[42] MARTÍNEZ, S. and YCART, B. (2001). Decay rates and cutoff for convergence and hitting times of Markov chains with countably infinite state space. *Adv. in Appl. Probab.* **33** 188–205. MR1825322 https://doi.org/10.1239/aap/999187903

[43] MATKOWSKY, B. J. and SCHUSS, Z. (1977). The exit problem for randomly perturbed dynamical systems. *SIAM J. Appl. Math.* **33** 365–382. MR0451427 https://doi.org/10.1137/0133024

[44] OLIVIERI, E. and VARES, M. E. (2005). *Large Deviations and Metastability*. *Encyclopedia of Mathematics and Its Applications* **100**. Cambridge Univ. Press, Cambridge. MR2123364 https://doi.org/10.1017/CBO9780511543272

[45] PERKO, L. (2001). *Differential Equations and Dynamical Systems*, 3rd ed. *Texts in Applied Mathematics* **7**. Springer, New York. MR1801796 https://doi.org/10.1007/978-1-4613-0003-8

[46] PRIOLA, E., SHIRIKYAN, A., XU, L. and ZABCZYK, J. (2012). Exponential ergodicity and regularity for equations with Lévy noise. *Stochastic Process. Appl.* **122** 106–133. MR2860444 https://doi.org/10.1016/j.spa.2011.10.003

[47] SALOFF-COSTE, L. (2004). Random walks on finite groups. In *Probability on Discrete Structures*. *Encyclopaedia Math. Sci.* **110** 263–346. Springer, Berlin. MR2023654 https://doi.org/10.1007/978-3-662-09444-0_5

[48] SCHUSS, Z. and MATKOWSKY, B. J. (1979). The exit problem: A new approach to diffusion across potential barriers. *SIAM J. Appl. Math.* **36** 604–623. MR0531616 https://doi.org/10.1137/0136043

[49] SIEGERT, W. (2009). *Local Lyapunov Exponents*: *Sublimiting Growth Rates of Linear Random Differential Equations*. *Lecture Notes in Math.* **1963**. Springer, Berlin. MR2465137

[50] STROOCK, D. W. and VARADHAN, S. R. S. (2006). *Multidimensional Diffusion Processes*. *Classics in Mathematics*. Springer, Berlin. Reprint of the 1997 edition. MR2190038

[51] VILLANI, C. (2009). Hypocoercivity. *Mem. Amer. Math. Soc.* **202** iv+141. MR2562709 https://doi.org/10.1090/S0065-9266-09-00567-5