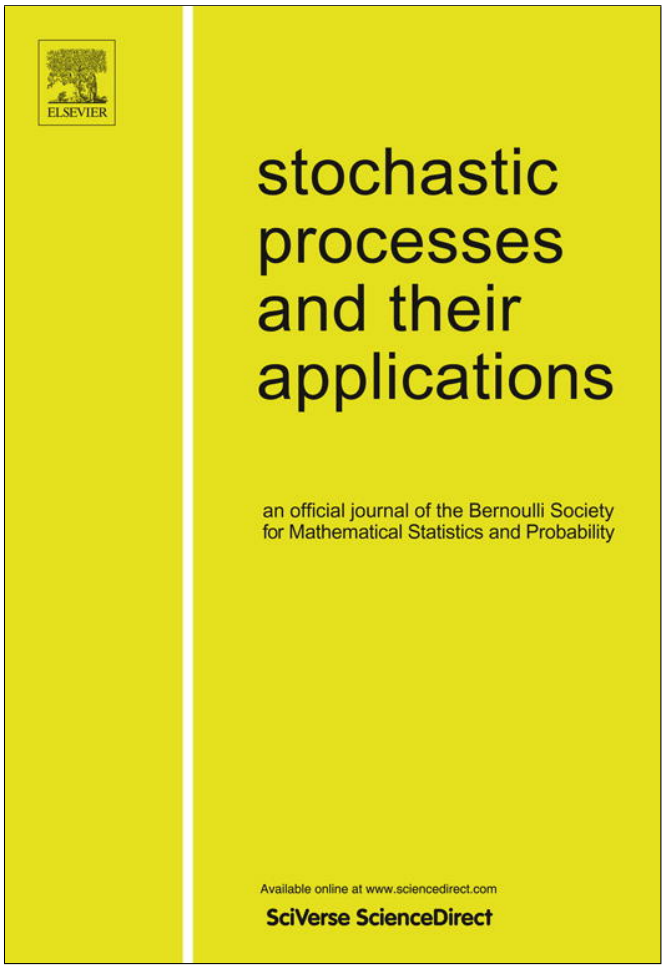
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Available online at www.sciencedirect.com
SciVerse ScienceDirect
Stochastic Processes and their Applications 122 (2012) 2553–2593
stochastic processes and their applications
www.elsevier.com/locate/spa

Strong and weak orders in averaging for SPDEs

Charles-Edouard Bréhier*

ENS Cachan Bretagne - IRMAR, Université Rennes 1, Avenue Robert Schumann, F-35170 Bruz, France

Received 6 April 2011; received in revised form 20 January 2012; accepted 12 April 2012
Available online 23 April 2012

Abstract

We show an averaging result for a system of stochastic evolution equations of parabolic type with slow and fast time scales. We derive explicit bounds for the approximation error with respect to the small parameter defining the fast time scale. We prove that the slow component of the solution of the system converges towards the solution of the averaged equation with an order of convergence 1/2 in a strong sense – approximation of trajectories – and 1 in a weak sense – approximation of laws. These orders turn out to be the same as for the SDE case.
© 2012 Elsevier B.V. All rights reserved.

MSC: 60H15; 70K65; 70K70

Keywords: Stochastic partial differential equations; Averaging principle; Strong and weak approximations

1. Introduction

In this paper, we consider a randomly-perturbed system of reaction–diffusion equations that can be written as

$$\begin{aligned}\frac{\partial x^\epsilon(t,\xi)}{\partial t} &= \frac{\partial^2 x^\epsilon(t,\xi)}{\partial \xi^2} + f(\xi, x^\epsilon(t,\xi), y^\epsilon(t,\xi)),\\ \frac{\partial y^\epsilon(t,\xi)}{\partial t} &= \frac{1}{\epsilon}\frac{\partial^2 y^\epsilon(t,\xi)}{\partial \xi^2} + \frac{1}{\epsilon} g(\xi, x^\epsilon(t,\xi), y^\epsilon(t,\xi)) + \frac{1}{\sqrt{\epsilon}}\frac{\partial \omega(t,\xi)}{\partial t},\end{aligned} \tag{1.1}$$

* Tel.: +33 6 76 24 02 35.
E-mail address: charles-edouard.brehier@bretagne.ens-cachan.fr.

0304-4149/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.spa.2012.04.007

Author's personal copy

for $t \geq 0, \xi \in (0,1)$, with initial conditions $x^\epsilon(0,\xi) = x(\xi)$ and $y^\epsilon(0,\xi) = y(\xi)$, and Dirichlet boundary conditions $x^\epsilon(t,0) = x^\epsilon(t,1) = 0$ and $y^\epsilon(t,0) = y^\epsilon(t,1) = 0$. The stochastic perturbation $\frac{\partial \omega(t,\xi)}{\partial t}$ is a space–time white noise and $\epsilon > 0$ is a small parameter.

Such a system presents a specific structure: while the variations of the first component *a priori* depend on the slow time t, the second component evolves with respect to the fast time $\frac{t}{\epsilon}$. These two natural time scales are coupled through the nonlinear terms in the two equations.

In this setting, the main idea of the *averaging principle*, see for instance [11], is to study the behaviour of the system when ϵ tends to 0 by exhibiting a limit equation – the so-called averaged equation – for the slow component x^ϵ, and to prove the convergence of x^ϵ towards the solution of this averaged equation. Here, we show two approximation results – see Theorems 1.1 and 1.2 – and give explicit order of convergence with respect to ϵ.

The averaged equation comes from the asymptotic behaviour of the fast equation. Heuristically, when $t > 0$ and $\epsilon \to 0$, the fast time $\frac{t}{\epsilon}$ goes to $+\infty$, so that we expect the solution of the fast equation to be quickly close to a stochastic equilibrium (and this is the case under the dissipativity assumptions made in this paper), and that we can replace $y^\epsilon(t,\xi)$ in the slow equation with some stationary – in the stochastic sense – process, which leads to the definition of averaged coefficients in the slow equation.

To give precise results, it is convenient to look at the equations in an abstract setting, where system (1.1) can be rewritten

$$\begin{aligned} dX^\epsilon(t) &= \left(AX^\epsilon(t) + F(X^\epsilon(t), Y^\epsilon(t))\right) dt \\ dY^\epsilon(t) &= \frac{1}{\epsilon}\left(BY^\epsilon(t) + G(X^\epsilon(t), Y^\epsilon(t))\right) dt + \frac{1}{\sqrt{\epsilon}} dW(t), \end{aligned} \tag{1.2}$$

with initial conditions given by $X^\epsilon(0) = x \in H$, $Y^\epsilon(0) = y \in H$, where H is the Hilbert space $L^2(0,1)$, and W is a cylindrical Wiener process on H—see Section 2.1.2. In the case of system (1.1), the definitions of A and B are given in Example 2.2, and the definitions of F and G as Nemytskii operators are given in the second part of Example 2.10. Nevertheless the abstract setting allows for more general equations, and in the sequel we only work with system (1.2).

When ϵ tends to 0, the slow component X^ϵ is approximated by the process $\overline{X}$, which follows the deterministic evolution equation

$$d\overline{X}(t) = (A\overline{X}(t) + \overline{F}(\overline{X}(t)))dt, \tag{1.3}$$

with the initial condition $\overline{X}(0) = x$, where the nonlinear coefficient $\overline{F}$ is obtained via an averaging procedure—explained in detail in Section 2.2.

In this article, we analyse the error between $X^\epsilon(t)$ and $\overline{X}(t)$, with two different criteria. We focus on the order of convergence, i.e. we bound the error by $C\epsilon^\Lambda$, where C is a constant and Λ is the order, which gives an idea of the speed of convergence of $X^\epsilon(t)$ towards $\overline{X}(t)$. As a result, we control the error made when $X^\epsilon(t)$ is approximated by $\overline{X}(t)$. For instance, the order of convergence is crucial for the analysis of numerical schemes used to approximate the slow component X^ϵ. In a forthcoming work, we extend a numerical scheme for SDEs analysed in [19] for systems of SPDEs satisfying the same structure assumptions as system (1.1). This scheme – called the Heterogeneous Multiscale Method – is deeply based on the averaging principle: instead of computing X^ϵ, we approximate $\overline{X}$—and we can control the error we make. Moreover, the nonlinear averaged coefficient $\overline{F}$ is never explicitly calculated in the scheme, but only approximated by using numerical approximations of the values of the fast component at large times. The theorems we prove here allow to analyse the convergence of such a scheme with the

Author's personal copy

same kind of criteria, and without knowing the order of convergence in the averaging principle it would not be possible to control the error made in the numerical approximation.

The two main theorems give bounds on the error between X^ϵ and $\overline{X}$; they need different dissipativity conditions (SD) and (WD) which determine how the fast equation converges to its equilibrium, as explained below.

First, when Assumption 2.8 holds, the error can be estimated in a strong sense, where trajectories of the processes are compared at a given time t:

Theorem 1.1 (*Strong-Order*). *Assume (SD). For any* $0 < r < 1/2$, $T > 0$, $x \in H$, $y \in H$, *there exists* $C = C(T, r, x, y) > 0$– *depending also on the constants of the problem – such that for any* $\epsilon > 0$ *and* $0 \leq t \leq T$

$$\mathbb{E}|X^\epsilon(t) - \overline{X}(t)|_H \leq C\epsilon^{1/2-r}. \tag{1.4}$$

The error can also be estimated in a weak sense, where we are interested in the distance between the laws of the processes at a given time t; then only Assumption 2.9 is necessary, since we only need consequences of dissipativity at the level of the transition semi-group, instead of trajectories:

Theorem 1.2 (*Weak-Order*). *Assume (WD). For any* $0 < r < 1, T > 0, 0 < \theta \leq 1$, $x \in D(-A)^\theta$, $y \in H, \phi \in \mathcal{C}_b^2(H)$, *there exists* $C > 0$, *depending on* $r, T, \phi, |x|_{(-A)^\theta}, |y|$ *and the constants of the problem, such that for any* $\epsilon > 0$ *and* $t \leq T$

$$|\mathbb{E}[\phi(X^\epsilon(t))] - \mathbb{E}[\phi(\overline{X}(t))]| \leq C\epsilon^{1-r}. \tag{1.5}$$

The domains $D(-A)^\theta$ are usually the classical Sobolev spaces $H^{2\theta}$ with respect to the eigenbasis of A—see Definition 2.3. We remark that for the first theorem no regularity is needed for the initial condition – i.e. we can take $\theta = 0$ – while we require $\theta > 0$ for the second one; this is explained in the proof of Theorem 1.2. We need to take a small parameter $r > 0$, which can be as small as possible, but different from 0. This is an effect of the infinite dimensional setting.

As a consequence, we can say that the strong order in averaging is $1/2$, while the weak order is 1. It is a general fact that the weak order is greater than the strong order (since test functions ϕ in the theorem are Lipschitz continuous), but it is worth proving that there is a gap; this fact was known for SDEs see [11,12], but had not been proved yet for SPDEs.

The strong convergence Theorem 1.1 is proved when the fast equation satisfies a strict dissipativity assumption: for any $x \in H$, the function $G(x, .)$ is Lipschitz continuous, with constant L_g – independent of x – satisfying the following condition:

$$L_g < \mu,$$

where μ is the smallest eigenvalue of the linear operator $-B$. Thanks to this assumption, we can easily analyse the asymptotic behaviour of the fast equation with frozen slow component; we can identify a unique invariant probability measure – depending on x – and show some exponential convergence to equilibrium. More precisely, we control in a strong sense the difference between two solutions of this fast equation starting from different initial conditions, and driven by the same noise W: under the previous assumption, the ergodicity comes from properties of the deterministic equation only.

Author's personal copy

The weak convergence Theorem 1.2 needs a weaker dissipativity Assumption 2.9, which yields the same ergodicity properties—unique invariant probability measure, exponential convergence to equilibrium – but with different arguments: the asymptotic behaviour of the transition semi-group can be analysed, thanks to the non-degeneracy of the noise – leading to a Strong Feller Property. A coupling method – adapted from the study of Markov processes, like in [13] or [16] – implies that the laws – instead of trajectories – of the fast process issued from two different initial conditions are exponentially closed. We refer to Section 2.2 for a precise result, and to [7] for a detailed proof. It seems that for the first time an averaging result is obtained for SPDEs under a weak dissipativity condition.

Notice that we have assumed that the slow equation has no white noise term $dw(t)$; as a consequence, the averaged equation is a deterministic parabolic partial differential equation. Considering a more general situation with some additive noise terms in the slow equation, independent of the noise in the fast equation, we could still prove in a similar way a strong order result, the only changes being time regularity of solutions. We would obtain order $1/5$, which can also be compared with the order $1/3$ obtained for SDEs. But if we introduce noise in the slow equation, the method we used to prove the weak order theorem becomes more complicated, and we have not extended the result to this situation so far.

In the case of stochastic differential equations, averaging results are already well-known—see for instance [10,11]. Convergence in law or in probability of X^ϵ to $\overline{X}$ in the space $\mathcal{C}([0,T],H)$ can be shown by different techniques: by using a Khasminskii technique based on a subdivision of the interval $[0,T]$(see [19,15]); a Poisson equation (see [18]); the method of perturbed test functions and of a martingale problem approach (see [9]); or an asymptotic expansion of the solutions of Kolmogorov equations (see [19,12]).

As far as stochastic partial differential equations are concerned, in [4] both the Khasminskii technique and a martingale problem approach are used; in [5] a modified Poisson equation is the essential tool. Then convergence in law or in probability of X^ϵ to $\overline{X}$ in the space $\mathcal{C}([0,T],H)$ (the space of continuous functions from $[0,T]$ to H) is proved; but order of convergence was never given.

Our proof of Theorem 1.1 relies on the Khasminskii technique already known for SDEs: we introduce an auxiliary process for which the slow component of the fast variable is frozen on small intervals of a subdivision. We use Hölder regularity of order $1-r$ in time of the slow component, for which we do not need $\theta > 0$.

To prove Theorem 1.2, we adapt the method of finding an expansion with respect to ϵ of the solutions of the Kolmogorov equations related to our system. This seems to be the first time that such a method is used to prove an averaging result for SPDEs. New technical difficulties due to infinite dimension arise: we use unbounded linear operators and nonsmooth nonlinear coefficients, and the Kolmogorov equations are more difficult to use. For these reasons, we use a reduction to finite dimension technique, keeping in mind that bounds must be independent of dimension, so that precise estimates are needed for each term appearing in the expansion. We interpret the necessity of $\theta > 0$ with a singularity which needs to be integrable.

In Section 2, we set the notations and give some results on the fast equation, allowing to define the averaged equation; we also precise the assumptions needed to prove the theorems. Then in Section 3, we prove the strong-order result. In Section 4, we give the details of the method for proving the weak-order result. Finally in Section 5 and in the Appendix, we prove all the necessary estimates.

Author's personal copy

2. Preliminaries

2.1. Assumptions and notations

2.1.1. Test functions

To study weak convergence, we use test functions ϕ in the space $\mathcal{C}_b^2(H, \mathbb{R})$ of functions from H to $\mathbb{R}$ that are twice continuously differentiable, with first and second order bounded derivatives.

In the sequel, we often identify the first derivative $D\phi(x) \in \mathcal{L}(H, \mathbb{R})$ with the gradient in H, and the second derivative $D^2\phi(x)$ with a linear operator on H, via the formulae:

$$\langle D\phi(x), h\rangle = D\phi(x) \cdot h \quad \text{for every } h \in H$$
$$\langle D^2\phi(x) \cdot h, k\rangle = D^2\phi(x) \cdot (h, k) \quad \text{for every } h, k \in H.$$

2.1.2. Stochastic integration in Hilbert spaces

In this section, we recall the definition of the cylindrical Wiener process and of stochastic integral on a separable Hilbert space H (its norm is denoted by $|.|_H$ or just $|.|$). For more details, see [6].

We first fix a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. A cylindrical Wiener process on H is defined with two elements:

- a complete orthonormal system of H, denoted by $(q_i)_{i\in I}$, where I is a subset of $\mathbb{N}$;
- a family $(\beta_i)_{i\in I}$ of independent real Wiener processes with respect to the filtration $((\mathcal{F}_t)_{t\geq 0})$:

$$W(t) = \sum_{i\in I} \beta_i(t) q_i. \tag{2.1}$$

When I is a finite set, we recover the usual definition of Wiener processes in the finite dimensional space $\mathbb{R}^{|I|}$. However the subject here is the study of some Stochastic Partial Differential Equations, so that in the sequel the underlying Hilbert space H is infinite dimensional; for instance when $H = L^2(0, 1)$, an example of complete orthonormal system is $(q_k) = (\sin(k.))_{k\geq 1}$—see Example 2.2.

A fundamental remark is that the series in (2.1) does not converge in H; but if a linear operator $\Psi : H \to K$ is Hilbert–Schmidt, then $\Psi W(t)$ converges in $L^2(\Omega, H)$ for any $t \geq 0$.

We recall that a linear operator $\Psi : H \to K$ is said to be Hilbert–Schmidt when

$$|\Psi|^2_{\mathcal{L}_2(H,K)} := \sum_{k=0}^{+\infty} |\Psi(q_k)|^2_K < +\infty,$$

where the definition is independent of the choice of the orthonormal basis (q_k) of H. The space of Hilbert–Schmidt operators from H to K is denoted $\mathcal{L}_2(H, K)$; endowed with the norm $|.|_{\mathcal{L}_2(H,K)}$ it is an Hilbert space.

The stochastic integral $\int_0^t \Psi(s)dW(s)$ is defined in K for predictable processes Ψ with values in $\mathcal{L}_2(H, K)$ such that $\int_0^t |\Psi(s)|^2_{\mathcal{L}_2(H,K)}ds < +\infty$ a.s; moreover when $\Psi \in L^2(\Omega \times [0, t]; \mathcal{L}_2(H, K))$, the following two properties hold:

$$\mathbb{E}\left|\int_0^t \Psi(s)dW(s)\right|^2_K = \mathbb{E}\int_0^t |\Psi(s)|^2_{\mathcal{L}_2(H,K)}ds, \quad \text{(Itô isometry)},$$
$$\mathbb{E}\int_0^t \Psi(s)dW(s) = 0.$$

A generalization of Itô formula also holds—see [6].

Author's personal copy

For instance, if $v = \sum_{k\in\mathbb{N}} v_k q_k \in H$, we can define

$$\langle W(t), v\rangle = \int_0^t \langle v, dW(s)\rangle = \sum_{k\in\mathbb{N}} \beta_k(t) v_k;$$

we then have the following space–time white noise property

$$\mathbb{E}\langle W(t), v_1\rangle\langle W(s), v_2\rangle = t \wedge s \langle v_1, v_2\rangle.$$

Therefore to be able to integrate a process with respect to W requires some strong properties on the integrand; in our SPDE setting, the Hilbert–Schmidt properties follow from the assumptions made on the linear coefficients of the equations.

2.1.3. Assumptions on the linear operators

We have to specify some properties of the linear operators A and B coming into the definition of system (1.2); we assume that the linear parts are of parabolic type, with space variable $\xi \in (0, 1)$.

We assume that A and B are unbounded linear operators, with domains $D(A)$ and $D(B)$, which satisfy the following assumptions:

Assumption 2.1. (1) We assume that $(e_k)_{k\in\mathbb{N}}$ and $(f_k)_{k\in\mathbb{N}}$ are orthonormal basis of H, and $(\lambda_k)_{k\in\mathbb{N}}$ and $(\mu_k)_{k\in\mathbb{N}}$ are non-decreasing sequences of real positive numbers such that:

$$Ae_k = -\lambda_k e_k \quad \text{for all } k \in \mathbb{N}$$
$$Bf_k = -\mu_k f_k \quad \text{for all } k \in \mathbb{N}.$$

We use the notations $\lambda := \lambda_0 > 0$ and $\mu := \mu_0 > 0$ for the smallest eigenvalues of A and B.

(2) The sequences (λ_k) and (μ_k) go to $+\infty$; moreover we have some control of the behaviour of (μ_k) given by:

$$\sum_{k=0}^{+\infty} \frac{1}{\mu_k^\alpha} < +\infty \Leftrightarrow \alpha > 1/2.$$

Example 2.2. $A = B = \frac{d^2}{dx^2}$, with domain $H^2(0, 1) \cap H_0^1(0, 1) \in L^2(0, 1)$ – homogeneous Dirichlet boundary conditions: in that case $\lambda_k = \mu_k = \pi^2 k^2$, and $e_k(\xi) = f_k(\xi) = \sqrt{2}\sin(k\pi\xi)$ – see [2].

In the abstract setting, powers of $-A$ and $-B$, with their domains can be easily defined:

Definition 2.3. For $a, b \in [0, 1]$, we define the operators $(-A)^a$ and $(-B)^b$ by

$$(-A)^a x = \sum_{k=0}^{\infty} \lambda_k^a x_k e_k \in H,$$
$$(-B)^b y = \sum_{k=0}^{\infty} \mu_k^b y_k f_k \in H,$$

with domains

$$D(-A)^a = \left\{ x = \sum_{k=0}^{+\infty} x_k e_k \in H;\ |x|^2_{(-A)^a} := \sum_{k=0}^{+\infty} (\lambda_k)^{2a} |x_k|^2 < +\infty \right\};$$

Author's personal copy

$$D(-B)^b = \left\{ y = \sum_{k=0}^{+\infty} y_k f_k \in H, |y|^2_{(-B)^b} := \sum_{k=0}^{+\infty} (\mu_k)^{2b} |y_k|^2 < +\infty \right\}.$$

The domains $D(-A)^a$ are related to Sobolev spaces $H^{2a}(0,1)$: therefore when x belongs to a space $D(-A)^a$, the exponent a represents some regularity of the function x.

The semi-groups $(e^{tA})_{t\geq 0}$ and $(e^{tB})_{t\geq 0}$ can be defined by the Hille–Yosida Theorem (see [2]). We use the following spectral formulae: if $x = \sum_{k=0}^{+\infty} x_k e_k \in H$ and $y = \sum_{k=0}^{+\infty} y_k f_k \in H$, then for any $t \geq 0$

$$e^{tA}x = \sum_{k=0}^{+\infty} e^{-\lambda_k t} x_k e_k \quad \text{and} \quad e^{tB}y = \sum_{k=0}^{+\infty} e^{-\mu_k t} y_k f_k.$$

For any $t \geq 0$, e^{tA} and e^{tB} are continuous linear operators in H, with respective operator norms $e^{-\lambda t}$ and $e^{-\mu t}$. The semi-group (e^{tA}) is used to define the solution $Z(t) = e^{tA}z$ of the linear Cauchy problem

$$\frac{dZ(t)}{dt} = AZ(t) \quad \text{with } Z(0) = z.$$

To define solutions of more general PDEs of parabolic type, we use mild formulation, and Duhamel principle.

These semi-groups enjoy some smoothing properties that we often use in this work. Basically we need the following properties, which are easily proved using the above spectral properties. We write them for A, but they also hold with B.

Proposition 2.4. *Under Assumption 2.1, for any $\sigma \in [0,1]$, there exists $C_\sigma > 0$ such that we have:*

(1) *for any $t > 0$ and $x \in H$*

$$|e^{tA}x|_{(-A)^\sigma} \leq C_\sigma t^{-\sigma} e^{-\frac{\lambda}{2}t} |x|_H.$$

(2) *for any $0 < s < t$ and $x \in H$*

$$|e^{tA}x - e^{sA}x|_H \leq C_\sigma \frac{(t-s)^\sigma}{s^\sigma} e^{-\frac{\lambda}{2}s} |x|_H.$$

(3) *for any $0 < s < t$ and $x \in D(-A)^\sigma$*

$$|e^{tA}x - e^{sA}x|_H \leq C_\sigma (t-s)^\sigma e^{-\frac{\lambda}{2}s} |x|_{(-A)^\sigma}.$$

Under the previous assumptions on the linear coefficients, it is easy to show that the following stochastic integral is well-defined in H, for any $t \geq 0$:

$$W^B(t) = \int_0^t e^{(t-s)B} dW(s). \tag{2.2}$$

It is called a stochastic convolution, and it is the unique mild solution of

$$dZ(t) = BZ(t)dt + dW(t) \quad \text{with } Z(0) = 0.$$

Under the second condition of Assumption 2.1, there exists $\delta > 0$ such that for any $t > 0$ we have $\int_0^t \frac{1}{s^\delta} |e^{sB}|^2_{\mathcal{L}_2(H)} ds < +\infty$; it can then be proved that W^B has continuous trajectories – via the *factorization method*, see [6] – and that for any $1 \leq p < +\infty$ $\sup_{t\geq 0} \mathbb{E}|W^B(t)|^p_H < +\infty$.

Author's personal copy

2.1.4. Assumptions on the nonlinear coefficients

We now give the Assumptions on the nonlinear coefficients $F, G : H \times H \to H$. First, we need some regularity properties:

Assumption 2.5. We assume that there exists $0 \leq \eta < \frac{1}{2}$ and a constant C such that the following directional derivatives are well-defined and controlled:

- For any $x, y \in H$ and $h \in H$, $|D_x F(x,y)\cdot h| \leq C|h|_H$ and $|D_y F(x,y)\cdot h| \leq C|h|_H$.
- For any $x, y \in H, h \in H, k \in D(-A)^\eta$, $|D^2_{xx} F(x,y)\cdot(h,k)| \leq C|h|_H|k|_{(-A)^\eta}$.
- For any $x, y \in H, h \in H, k \in D(-B)^\eta$, $|D^2_{yy} F(x,y)\cdot(h,k)| \leq C|h|_H|k|_{(-B)^\eta}$.
- For any $x, y \in H, h \in H, k \in D(-B)^\eta$, $|D^2_{xy} F(x,y)\cdot(h,k)| \leq C|h|_H|k|_{(-B)^\eta}$.
- For any $x, y \in H, h \in D(-A)^\eta, k \in H$, $|D^2_{xy} F(x,y)\cdot(h,k)| \leq C|h|_{(-A)^\eta}|k|_H$.

We moreover assume that F is bounded.

Remark 2.6. We warn the reader that constants may vary from line to line during the proofs, and that in order to use lighter notations we usually forget to mention dependence on the parameters. We use the generic notation C for such constants.

We assume that the fast equation is a gradient system: for any x the nonlinear coefficient $G(x,.)$ is the derivative of some potential U. We also assume regularity assumptions as for F.

Assumption 2.7. The function G is defined through $G(x,y) = \nabla_y U(x,y)$, for some potential $U : H \times H \to \mathbb{R}$. Moreover we assume that G is bounded, and that the regularity assumptions given in the Assumption 2.5 are also satisfied for G.

Finally, we need to assume some dissipativity of the fast equation. Assumption 2.8 is necessary to prove Theorem 1.1, while Assumption 2.9 is weaker and is sufficient to prove Theorem 1.2.

Assumption 2.8 (*Strict Dissipativity*)**.** Let L_g denote the Lipschitz constant of G with respect to its second variable; then

$$L_g < \mu, \tag{SD}$$

where μ is defined in Assumption 2.1.

Assumption 2.9 (*Weak Dissipativity*)**.** There exist $c > 0$ and $C > 0$ such that for any $x \in H$ and $y \in D(B)$

$$\langle By + G(x,y), y\rangle \leq -c|y|^2 + C. \tag{WD}$$

Indeed, according to Assumptions 2.1 and 2.7, the weak dissipativity Assumption 2.9 is always satisfied, while strict dissipativity requires a condition on the Lipschitz constant of G.

Example 2.10. We give some fundamental examples of nonlinearities for which the previous assumptions are satisfied:

- Functions $F, G : H \times H \to H$ of class $\mathcal{C}^2$, bounded and with bounded derivatives, such that $G(x,y) = \nabla_y U(x,y)$ and satisfying (SD) fit in the framework, with the choice $\eta = 0$.
- Functions F and G can be *Nemytskii* operators: let $f : (0,1) \times \mathbb{R}^2 \to \mathbb{R}$ be a measurable function such that for almost every $\xi \in (0,1)$ $f(\xi,.)$ is twice continuously differentiable,

Author's personal copy

bounded and with uniformly bounded derivatives. Then F is defined for every $x, y \in H = L^2(0,1)$ by

$$F(x,y)(\xi) = f(\xi, x(\xi), y(\xi)).$$

For G, we assume that there exists a function g with the same properties as f above, such that $G(x,y)(\xi) = g(\xi, x(\xi), y(\xi))$. The strict dissipativity assumption (SD) is then satisfied when

$$\sup_{\xi\in(a,b),x\in\mathbb{R},y\in\mathbb{R}} \left|\frac{\partial g}{\partial y}(\xi,x,y)\right| < \mu.$$

The conditions in Assumption 2.5 are then satisfied for F and G as soon as there exists $\eta < 1/2$ such that $D(-A)^\eta$ and $D(-B)^\eta$ are continuously embedded into $L^\infty(0,1)$—it is the case for A and B given in Example 2.2, with $\eta > 1/4$.

We remark that under Assumption 2.7, if we define $U_0(x,y) = \int_0^1 < G(x,sy), y > ds$, we have $U_0(x,y) = U(x,y) - U(x,0)$; therefore U_0 is another potential for G, and is the only one such that for any $x \in H$ we have $U_0(x,0) = 0$. In the sequel, it is therefore not restrictive to assume $U = U_0$.

Now we can define solutions of system (1.2); under Assumptions 2.1, 2.5 and 2.7, we notice that the nonlinearities F and G are Lipschitz continuous, and the following proposition is classical—see [6]:

Proposition 2.11. *For every $\epsilon > 0, T > 0, x \in H, y \in H$, system (1.2) admits a unique mild solution $(X^\epsilon, Y^\epsilon) \in (L^2(\Omega, \mathcal{C}([0,T],H)))^2$:*

$$\begin{aligned} X^\epsilon(t) &= e^{tA}x + \int_0^t e^{(t-s)A}F(X^\epsilon(s),Y^\epsilon(s))ds \\ Y^\epsilon(t) &= e^{\frac{t}{\epsilon}B}y + \frac{1}{\epsilon}\int_0^t e^{\frac{(t-s)}{\epsilon}B}G(X^\epsilon(s),Y^\epsilon(s))ds + \frac{1}{\sqrt{\epsilon}}\int_0^t e^{\frac{(t-s)}{\epsilon}B}dW(s). \end{aligned} \tag{2.3}$$

In other words, system (1.2) is well-posed for any $\epsilon > 0$, on any finite time interval $[0,T]$.

Some properties – bounds on moments, space and time regularity, differentiability with respect to the parameters – of X^ϵ and Y^ϵ are given in the Appendix.

2.2. Known results about the fast equation and the averaged equation

In this section, we just recall without proof the main results on the fast equation with frozen slow component and on the averaged equation, defined below. Proofs can be found in [5] for the strict dissipative case, and the extension to the weakly dissipative situation relies on arguments explained below.

If $x \in H$, we define an equation on the fast variable where the slow variable is fixed and equal to x:

$$\begin{aligned} dY_x(t,y) &= (BY_x(t,y) + G(x,Y_x(t,y)))dt + dW(t), \\ Y_x(0,y) &= y. \end{aligned} \tag{2.4}$$

This equation admits a unique mild solution, defined on $[0,+\infty[$.

Since Y^ϵ is involved at time $t > 0$, heuristically we need to analyse the properties of $Y_x(\frac{t}{\epsilon}, y)$, with $\epsilon \to 0$, and by a change of time we need to understand the asymptotic behaviour of $Y_x(.,y)$ when time goes to infinity.

Author's personal copy

Under the strict dissipativity Assumption 2.8, we obtain a contractivity of trajectories issued from different initial conditions and driven by the same noise:

Proposition 2.12. *With (SD), for any $t \geq 0$, $x, y_1, y_2 \in H$ we have*

$$|Y_x(t, y_1) - Y_x(t, y_2)|_H \leq e^{-\frac{(\mu - L_g)}{2}t} |y_1 - y_2|_H.$$

Under the weak dissipativity Assumption 2.9, we obtain such an exponential convergence result for the laws instead of trajectories. The proof of this result is not straightforward, and can be found in [7].

Proposition 2.13. *With (WD), there exist $c > 0, C > 0$ such that for any bounded test function ϕ, any $t \geq 0$ and any $y_1, y_2 \in H$*

$$|\mathbb{E}\phi(Y(t, y_1)) - \mathbb{E}\phi(Y(t, y_2))| \leq C\|\phi\|_\infty (1 + |y_1|^2 + |y_2|^2) e^{-ct}. \tag{2.5}$$

The idea of coupling relies on the following formula: if ν_1 and ν_2 are two probability measures on a state space S, their total variation distance satisfies

$$d_{TV}(\nu_1, \nu_2) = \inf \{\mathbb{P}(X_1 \neq X_2)\},$$

which is an infimum over random variables (X_1, X_2) defined on a same probability space, and such that $X_1 \sim \nu_1$ and $X_2 \sim \nu_2$.

The principle is to define a coupling $(Z_1(t, y_1, y_2), Z_2(t, y_1, y_2))_{t \geq 0}$ for the processes $(Y(t, y_1))_{t \geq 0}$ and $(Y(t, y_2))_{t \geq 0}$ such that the coupling time $\mathcal{T}$ of Z_1 and Z_2 – i.e. the first time the processes are equal – has an exponentially decreasing tail.

This technique was first used in the study of the asymptotic behaviour of Markov chains – see [1,8,14,17] – and was later adapted for SDEs and more recently for SPDEs—see for instance [13,16].

As a consequence, we can show that there exists a unique invariant probability measure associated with Y_x, and that the convergence to equilibrium is exponentially fast.

First, let $\nu = \mathcal{N}(0, (-B)^{-1}/2)$ be the centered Gaussian probability measure on H with the covariance operator $(-B)^{-1}/2$—which is positive and trace-class, thanks to Assumption 2.1.

Then μ^x defined by

$$\mu^x(dy) = \frac{1}{Z(x)} e^{2U(x,y)} \nu(dy), \tag{2.6}$$

where $Z(x) \in]0, +\infty[$ is a normalization constant, is the unique probability invariant measure associated to Y_x. This expression comes from the gradient structure of equation (2.4), given in Assumption 2.7.

Second, under both dissipativity assumptions, the convergence to equilibrium is exponential in the following sense:

Proposition 2.14. *If we assume (SD) or (WD), there exist constants $C, c > 0$ such that for any bounded function $\phi : H \to \mathbb{R}$ or $\phi : H \to H$, $t \geq 0$ and $x, y \in H$ we have*

$$\left| \mathbb{E}\phi(Y_x(t, y)) - \int_H \phi(z) \mu^x(dz) \right| \leq C\|\phi\|_\infty (1 + |y|_H^2) e^{-ct}.$$

Under the strict dissipativity Assumption 2.8, this is a consequence of Proposition 2.12 – see Theorem 3.5 and Remark 3.6 of [5]; under the weak dissipativity Assumption 2.9,

Author's personal copy

Proposition 2.14 is a consequence of Proposition 2.13 and of the properties of the invariant measures μ^x – which have finite moments of any order, uniformly bounded with respect to x.

Now we define the averaged equation. First we define the averaged nonlinear coefficient $\overline{F}$:

Definition 2.15. For any $x \in H$,

$$\overline{F}(x) = \int_H F(x, y)\mu^x(dy). \tag{2.7}$$

Using Assumptions 2.5 and 2.7 and the expression of μ^x, we can easily prove the following properties on $\overline{F}$:

Proposition 2.16. *There exists $0 \leq \eta < 1$ and a constant C such that the following directional derivatives of $\overline{F}$ are well-defined and controlled:*

- *For any $x \in H, h \in H$, $|D\overline{F}(x) \cdot h| \leq C|h|_H$.*
- *For any $x \in H, h \in H, k \in D(-A)^\eta$, $|D^2\overline{F}(x) \cdot (h, k)| \leq C|h|_H|k|_{(-A)^\eta}$.*

Moreover, $\overline{F}$ is bounded and Lipschitz continuous.

Remark 2.17. Even when F and G are Nemytskii operators, $\overline{F}$ is not such an operator in general.

Then the averaged equation – see (1.3) in the introduction – can be defined:

$$d\overline{X}(t) = (A\overline{X}(t) + \overline{F}(\overline{X}(t)))dt,$$

with initial condition $\overline{X}(0) = x \in H$. For any $T > 0$, this deterministic equation admits a unique mild solution $\overline{X} \in \mathcal{C}([0, T], H)$.

3. Proof of the strong-order result

The main idea – inspired by the work on SDEs of Khasminskii in [11] – of the proof of Theorem 1.1 is the construction of auxiliary processes $(\tilde{X}^\epsilon, \tilde{Y}^\epsilon)$ for any ϵ, for which the analysis is simpler.

In this section, we assume that dissipativity of the fast equation is strict: we have (SD).

Let $T > 0$, $x \in H$, $y \in H$ and $\epsilon > 0$ be fixed. We introduce the parameter

$$\delta = \delta(\epsilon) = \sqrt{\epsilon} \tag{3.1}$$

to define a subdivision of $[0, T]$. We also fix $r > 0$.

We define $\tilde{X}^\epsilon$ and $\tilde{Y}^\epsilon$ via a mild formulation: for any $0 \leq t \leq T$

$$\begin{aligned}
\tilde{Y}^\epsilon(t) &= e^{\frac{t}{\epsilon}B}y + \frac{1}{\epsilon}\int_0^t e^{\frac{(t-s)}{\epsilon}B}G(X^\epsilon(\lfloor\frac{s}{\delta}\rfloor\delta), \tilde{Y}^\epsilon(s))ds + \frac{1}{\sqrt{\epsilon}}\int_0^t e^{\frac{(t-s)}{\epsilon}B}dW(s),\\
\tilde{X}^\epsilon(t) &= e^{tA}x + \int_0^t e^{(t-s)A}F(X^\epsilon(\lfloor\frac{s}{\delta}\rfloor\delta), \tilde{Y}^\epsilon(s))ds,
\end{aligned} \tag{3.2}$$

where $\lfloor . \rfloor$ denotes the integer part function.

$\tilde{X}^\epsilon$ and $\tilde{Y}^\epsilon$ are continuous processes and they satisfy $\tilde{X}^\epsilon(0) = x = X^\epsilon(0)$ and $\tilde{Y}^\epsilon(0) = y = Y^\epsilon(0)$. Moreover, on any subinterval $[k\delta, (k+1)\delta]$, with $0 \leq k \leq N := \lfloor\frac{T}{\delta}\rfloor$, we have

$$\begin{aligned}
d\tilde{X}^\epsilon(t) &= (A\tilde{X}^\epsilon(t) + F(X^\epsilon(k\delta), \tilde{Y}^\epsilon(t)))dt,\\
d\tilde{Y}^\epsilon(t) &= \frac{1}{\epsilon}(B\tilde{Y}^\epsilon(t) + G(X^\epsilon(k\delta), \tilde{Y}^\epsilon(t)))dt + \frac{1}{\sqrt{\epsilon}}dW(t).
\end{aligned} \tag{3.3}$$

Author's personal copy

We remark that on such subintervals the fast component $\tilde{Y}^\epsilon$ does not depend on the slow component $\tilde{X}^\epsilon$, but only on the value of X^ϵ at the first point of the interval.

Since F and G are supposed to be bounded, we easily see that for any $t \geq 0$

$$\mathbb{E}|\tilde{X}^\epsilon(t)|_H^2 \leq C(1+|x|_H^2) \quad \text{and} \quad \mathbb{E}|\tilde{Y}^\epsilon(t)|_H^2 \leq C(1+|y|_H^2).$$

We show the following lemmas:

Lemma 3.1. *There exists $C>0$ such that for any $0 \leq t \leq T$ and any $\epsilon > 0$*

$$\mathbb{E}|X^\epsilon(t) - \tilde{X}^\epsilon(t)| \leq C\left(\delta^{1-r} + \frac{\epsilon}{\delta}\right).$$

Lemma 3.2. *There exists $C>0$ such that for any $0 \leq t \leq T$ and any $\epsilon > 0$*

$$\mathbb{E}|\tilde{X}^\epsilon(t) - \overline{X}(t)| \leq C\left(\epsilon(1+\delta^{-r})\left(1+\frac{1}{1-e^{-c\frac{\delta}{\epsilon}}}\right)\right)^{1/2} + C\left(\delta^{1-r}+\frac{\epsilon}{\delta}\right).$$

The constant C above depends on r, T, x, y, but not on t, ϵ or $\delta(\epsilon)$.

Lemma 3.1 explains why we can replace $\tilde{X}^\epsilon$ by X^ϵ, and Lemma 3.2 gives an estimate of the distance between $\tilde{X}^\epsilon(t)$ and $\overline{X}(t)$. With the choice of $\delta(\epsilon) = \sqrt{\epsilon}$ given by (3.1), the proof of Theorem 1.1 is straightforward.

Proof of Lemma 3.1.

- Estimate of $Y^\epsilon - \tilde{Y}^\epsilon$.

  We define $\rho^\epsilon(t) = Y^\epsilon(t) - \tilde{Y}^\epsilon(t)$ for any $0 \leq t \leq T$.

  We fix $k \geq 0$; then for any $t \in [k\delta, (k+1)\delta]$ we have

  $$d\rho^\epsilon(t) = \frac{1}{\epsilon}B\rho^\epsilon(t)dt + \frac{1}{\epsilon}(G(X^\epsilon(t), Y^\epsilon(t)) - G(X^\epsilon(k\delta), \tilde{Y}^\epsilon(t)))dt.$$

  Using a mild formulation and Gronwall lemma, for any $t \in [k\delta, (k+1)\delta]$ we have

  $$\mathbb{E}|\rho^\epsilon(t)| \leq e^{-\frac{\mu-L_g}{\epsilon}(t-k\delta)}\mathbb{E}|\rho^\epsilon(k\delta)| + \frac{C}{\epsilon}\int_{k\delta}^t e^{-\frac{\mu-L_g}{\epsilon}(t-s)}\mathbb{E}|X^\epsilon(s) - X^\epsilon(k\delta)|ds.$$

  Since $\mathbb{E}|Y^\epsilon(t)| \leq C(1+|y|)$ and $\mathbb{E}|\tilde{Y}^\epsilon(t)| \leq C(1+|y|)$, we have the same bound on ρ^ϵ.

  We can integrate the previous inequality over the interval $t \in [k\delta, (k+1)\delta]$ and get

  $$\begin{aligned}
  &\int_{k\delta}^{(k+1)\delta} \mathbb{E}|\rho^\epsilon(t)|dt \\
  &\leq C\int_{k\delta}^{(k+1)\delta} e^{-\frac{\mu-L_g}{\epsilon}(t-k\delta)}dt + \frac{C}{\epsilon}\int_{k\delta}^{(k+1)\delta}\int_{k\delta}^t e^{-\frac{\mu-L_g}{\epsilon}(t-s)}\mathbb{E}|X^\epsilon(s) - X^\epsilon(k\delta)|dsdt \\
  &\leq C\frac{\epsilon}{\mu-L_g} + C\int_{k\delta}^{(k+1)\delta}\mathbb{E}|X^\epsilon(s) - X^\epsilon(k\delta)|\int_s^{(k+1)\delta}\frac{1}{\epsilon}e^{-\frac{\mu-L_g}{\epsilon}(t-s)}dtds \\
  &\leq C\epsilon + C\int_{k\delta}^{(k+1)\delta}\mathbb{E}|X^\epsilon(s) - X^\epsilon(k\delta)|ds.
  \end{aligned}$$

  We recall that strict dissipativity $\mu - L_g > 0$ holds, thanks to Assumption 2.8.

Author's personal copy

It remains to take the sum over $k \in \left\{0, \dots, \lfloor \frac{t}{\delta} \rfloor\right\}$, where $t \leq T$; using Proposition A.3 of the Appendix, we then obtain

$$\int_0^t \mathbb{E}|\rho^\epsilon(s)|ds \leq C(r,T)\left(\frac{\epsilon}{\delta} + \delta^{1-r}\right).$$

- Estimate of $X^\epsilon - \tilde{X}^\epsilon$.

  We have for any $0 \leq t \leq T$

$$\begin{aligned}
|X^\epsilon(t) - \tilde{X}^\epsilon(t)| &= \left|\int_0^t e^{(t-s)A}(F(X^\epsilon(s), Y^\epsilon(s)) - F(X^\epsilon(\lfloor \tfrac{s}{\delta} \rfloor \delta), \tilde{Y}(s)))ds\right| \\
&\leq \int_0^t \left|F(X^\epsilon(s), Y^\epsilon(s)) - F\left(X^\epsilon\left(\lfloor \tfrac{s}{\delta} \rfloor \delta\right), \tilde{Y}(s)\right)\right|^2 ds \\
&\leq C\int_0^T \left(\left|X^\epsilon(s) - X^\epsilon\left(\lfloor \tfrac{s}{\delta} \rfloor \delta\right)\right| + |Y^\epsilon(s) - \tilde{Y}^\epsilon(s)|\right) ds.
\end{aligned}$$

  Using the previous estimate and the regularity result from Proposition A.3, we obtain for any $0 \leq t \leq T$

$$\mathbb{E}|X^\epsilon(t) - \tilde{X}^\epsilon(t)| \leq C\left(\delta^{1-r} + \frac{\epsilon}{\delta}\right). \quad \square \tag{3.4}$$

Proof of Lemma 3.2. We introduce the following decomposition, for any $0 \leq t \leq T$:

$$\begin{aligned}
\tilde{X}^\epsilon(t) - \overline{X}(t) &= \int_0^t e^{(t-s)A}\left(F\left(X^\epsilon\left(\lfloor \tfrac{s}{\delta} \rfloor \delta\right), \tilde{Y}^\epsilon(s)\right) - \overline{F}(\overline{X}(s))\right) ds \\
&= \int_0^t e^{(t-s)A}\left(F\left(X^\epsilon\left(\lfloor \tfrac{s}{\delta} \rfloor \delta\right), \tilde{Y}^\epsilon(s)\right) - \overline{F}\left(X^\epsilon\left(\lfloor \tfrac{s}{\delta} \rfloor \delta\right)\right)\right) ds \\
&\quad + \int_0^t e^{(t-s)A}\left(\overline{F}\left(X^\epsilon\left(\lfloor \tfrac{s}{\delta} \rfloor \delta\right)\right) - \overline{F}(X^\epsilon(s))\right) ds \\
&\quad + \int_0^t e^{(t-s)A}(\overline{F}(X^\epsilon(s)) - \overline{F}(\tilde{X}^\epsilon(s)))ds \\
&\quad + \int_0^t e^{(t-s)A}(\overline{F}(\tilde{X}^\epsilon(s)) - \overline{F}(\overline{X}(s)))ds \\
&= I_1(t) + I_2(t) + I_3(t) + I_4(t).
\end{aligned}$$

According to Proposition 2.16, $\overline{F}$ is Lipschitz continuous; thanks to Proposition A.3 and Lemma 3.1, we then show that for any $0 \leq t \leq T$

$$\begin{aligned}
\mathbb{E}|I_2(t)| &\leq C\int_0^T \mathbb{E}\left|X^\epsilon\left(\lfloor \tfrac{s}{\delta} \rfloor \delta\right) - X^\epsilon(s)\right| ds \leq C\delta^{1-r} \\
\mathbb{E}|I_3(t)| &\leq C\int_0^T |X^\epsilon(s) - \tilde{X}^\epsilon(s)|ds \leq C\left(\delta^{1-r} + \frac{\epsilon}{\delta}\right) \\
\mathbb{E}|I_4(t)| &\leq CT\int_0^t \mathbb{E}|\tilde{X}^\epsilon(r) - \overline{X}(r)|ds.
\end{aligned}$$

The I_4 term is treated via the Gronwall Lemma.

Author's personal copy

It remains to focus on the I_1 term. It is fundamental to look at $\mathbb{E}|I_1(t)|^2$ and not only $\mathbb{E}|I_1(t)|$ in order to obtain the best estimate leading to order $1/2$. For that, we use the subdivision of $[0, T]$ and we expand the scalar product in H: we have for any $0 \leq t \leq T$

$$\begin{aligned}
|I_1(t)|^2 &= \left| \int_0^t e^{(t-s)A} \left(F\left(X^\epsilon\left(\lfloor \frac{s}{\delta} \rfloor \delta \right), \tilde{Y}^\epsilon(s) \right) - \overline{F}\left(X^\epsilon\left(\lfloor \frac{s}{\delta} \rfloor \delta \right) \right) \right) ds \right|^2 \\
&= \left| \sum_{k=0}^{\lfloor \frac{t}{\delta} \rfloor} \int_{k\delta}^{(k+1)\delta \wedge t} e^{(t-s)A} (F(X^\epsilon(k\delta), \tilde{Y}^\epsilon(s)) - \overline{F}(X^\epsilon(k\delta))) ds \right|^2 \\
&= A_1(t) + A_2(t),
\end{aligned}$$

where

$$A_1(t) := \sum_{k=0}^{\lfloor \frac{t}{\delta} \rfloor} \left| \int_{k\delta}^{(k+1)\delta \wedge t} e^{(t-s)A} (F(X^\epsilon(k\delta), \tilde{Y}^\epsilon(s)) - \overline{F}(X^\epsilon(k\delta))) ds \right|^2 \tag{3.5}$$

and

$$\begin{aligned}
A_2(t) := 2 \sum_{0 \leq i < j \leq \lfloor \frac{t}{\delta} \rfloor} \Bigg\langle & \int_{i\delta}^{(i+1)\delta \wedge t} e^{(t-s)A} (F(X^\epsilon(i\delta), \tilde{Y}^\epsilon(s)) - \overline{F}(X^\epsilon(i\delta))) ds, \\
& \times \int_{j\delta}^{(j+1)\delta \wedge t} e^{(t-s)A} (F(X^\epsilon(j\delta), \tilde{Y}^\epsilon(s)) - \overline{F}(X^\epsilon(j\delta))) ds \Bigg\rangle.
\end{aligned} \tag{3.6}$$

We claim that $\mathbb{E}A_1(t) \leq C\epsilon$ and $\mathbb{E}A_2(t) \leq C\epsilon(1 + \delta^{-r})(1 + \frac{1}{1-e^{-c\frac{\delta}{\epsilon}}})$, where $C = C(T, r, \theta, x, y)$. Using Gronwall Lemma, we get the result.

- We first prove the estimate on $\mathbb{E}A_1(t)$. We use conditional expectation with respect to $\mathcal{F}_s$. Then for any $0 \leq k \leq \lfloor \frac{t}{\delta} \rfloor$, using some symmetry for variables s and σ,

$$\begin{aligned}
&\mathbb{E} \left| \int_{k\delta}^{(k+1)\delta \wedge t} e^{(t-s)A} (F(X^\epsilon(k\delta), \tilde{Y}^\epsilon(s)) - \overline{F}(X^\epsilon(k\delta))) ds \right|^2 \\
&= 2\mathbb{E} \int_{k\delta}^{(k+1)\delta \wedge t} ds \int_s^{(k+1)\delta \wedge t} d\sigma \langle e^{(t-s)A} (F(X^\epsilon(k\delta), \tilde{Y}^\epsilon(s)) - \overline{F}(X^\epsilon(k\delta))), \\
&\quad \times e^{(t-\sigma)A} (F(X^\epsilon(k\delta), \tilde{Y}^\epsilon(\sigma)) - \overline{F}(X^\epsilon(k\delta))) \rangle \\
&= 2\mathbb{E} \int_{k\delta}^{(k+1)\delta \wedge t} ds \int_s^{(k+1)\delta \wedge t} d\sigma \langle e^{(t-s)A} (F(X^\epsilon(k\delta), \tilde{Y}^\epsilon(s)) - \overline{F}(X^\epsilon(k\delta))), \\
&\quad \times e^{(t-\sigma)A} \mathbb{E}[F(X^\epsilon(k\delta), \tilde{Y}^\epsilon(\sigma)) - \overline{F}(X^\epsilon(k\delta)) | \mathcal{F}_s] \rangle \\
&\leq 2 \int_{k\delta}^{(k+1)\delta} \int_s^{(k+1)\delta} \mathbb{E}(|F(X^\epsilon(k\delta), \tilde{Y}^\epsilon(s)) - \overline{F}(X^\epsilon(k\delta))| \, |\mathbb{E}[F(X^\epsilon(k\delta), \tilde{Y}^\epsilon(\sigma)) | \mathcal{F}_s] \\
&\quad - \overline{F}(X^\epsilon(k\delta))|).
\end{aligned}$$

We now define the auxiliary function $\tilde{F}$. Propositions C.1 and C.2 – see the Appendix – give important properties of this function: we have some exponential control with respect to time t of uniform and Lipschitz bounds with respect to x.

Author's personal copy

Definition 3.3. For any $(x, y) \in H^2$ and $t \geq 0$

$$\tilde{F}(x, y, t) = \mathbb{E}F(x, Y_x(t, y)) - \overline{F}(x). \tag{3.7}$$

Since F is bounded, we can control the first factor in the integral; for the second factor, we use the definition of $\tilde{F}$, the Markov property and Proposition C.1 to see that

$$\begin{aligned}
&\mathbb{E}\left|\int_{k\delta}^{(k+1)\delta\wedge t} e^{(t-s)A}(F(X^\epsilon(k\delta), \tilde{Y}^\epsilon(s)) - \overline{F}(X^\epsilon(k\delta)))ds\right|^2 \\
&\quad \leq C\int_{k\delta}^{(k+1)\delta} ds \int_s^{(k+1)\delta} d\sigma \mathbb{E}\left|\tilde{F}\left(X^\epsilon(k\delta), \tilde{Y}^\epsilon(s), \frac{\sigma - s}{\epsilon}\right)\right| \\
&\quad \leq C\int_{k\delta}^{(k+1)\delta} \int_s^{(k+1)\delta} e^{-c\frac{\sigma-s}{\epsilon}} d\sigma ds \leq C\delta\epsilon.
\end{aligned}$$

Therefore we get $\mathbb{E}A_1(t) \leq C\epsilon$.

- Estimate of $\mathbb{E}A_2(t)$.

We have to introduce the following auxiliary processes, which generalize $\tilde{Y}^\epsilon$. $(Z_i^\epsilon(s))_{s\geq i\delta}$, where $i \in \{0, \dots, N\}$ is defined by:

$$\begin{aligned}
dZ_i^\epsilon(s) &= \frac{1}{\epsilon}(BZ_i^\epsilon(s) + G(X^\epsilon(i\delta), Z_i^\epsilon(s)))ds + \frac{1}{\sqrt{\epsilon}}dW(s) \\
Z_i^\epsilon(i\delta) &= \tilde{Y}^\epsilon(i\delta).
\end{aligned} \tag{3.8}$$

It is then clear that for $i\delta \leq s \leq (i+1)\delta$ we have $Z_i^\epsilon(s) = \tilde{Y}^\epsilon(s)$, and that $Z_k^\epsilon((k+1)\delta) = \tilde{Y}^\epsilon((k+1)\delta) = Z_{k+1}^\epsilon((k+1)\delta)$. Moreover the processes (Z_i^ϵ) are uniformly bounded with respect to i and ϵ.

It is then possible to rewrite the integrands appearing in the expression of A_2: when $i\delta \leq s \leq (i+1)\delta \leq j\delta \leq \tau \leq (j+1)\delta$,

$$\begin{aligned}
&\left|\mathbb{E}\left\langle e^{(t-s)A}(F(X^\epsilon(i\delta), \tilde{Y}^\epsilon(s)) - \overline{F}(X^\epsilon(i\delta))), e^{(t-\tau)A}(F(X^\epsilon(j\delta), \tilde{Y}^\epsilon(\tau)) - \overline{F}(X^\epsilon(j\delta)))\right\rangle\right| \\
&\quad = \Big|\mathbb{E}\Big\langle e^{(t-s)A}(F(X^\epsilon(i\delta), \tilde{Y}^\epsilon(s)) - \overline{F}(X^\epsilon(i\delta))), e^{(t-\tau)A}\mathbb{E}[F(X^\epsilon(j\delta), \tilde{Y}^\epsilon(\tau)) \\
&\qquad - \overline{F}(X^\epsilon(j\delta))|\mathcal{F}_{(i+1)\delta}]\Big\rangle\Big| \\
&\quad \leq C\mathbb{E}|\mathbb{E}[F(X^\epsilon(j\delta), \tilde{Y}^\epsilon(\tau)) - \overline{F}(X^\epsilon(j\delta))|\mathcal{F}_{(i+1)\delta}]|,
\end{aligned}$$

since F is bounded. We have $\tilde{Y}^\epsilon(\tau) = Z_j^\epsilon(\tau)$; however since $i < j$ we can use conditional expectation with respect to $\mathcal{F}_{(i+1)\delta}$ instead of $\mathcal{F}_{j\delta}$ in order to get a better estimate. We therefore propose the following decomposition

$$\begin{aligned}
&\mathbb{E}|\mathbb{E}[F(X^\epsilon(j\delta), \tilde{Y}^\epsilon(\tau)) - \overline{F}(X^\epsilon(j\delta))|\mathcal{F}_{(i+1)\delta}]| \\
&\quad \leq C\mathbb{E}|\mathbb{E}[F(X^\epsilon((i+1)\delta), Z_{i+1}^\epsilon(\tau)) - \overline{F}(X^\epsilon((i+1)\delta))|\mathcal{F}_{(i+1)\delta}]| \\
&\qquad + C\mathbb{E}\Big|\mathbb{E}\Big[(F(X^\epsilon(j\delta), Z_j^\epsilon(\tau)) - \overline{F}(X^\epsilon(j\delta))) - (F(X^\epsilon((i+1)\delta), Z_{i+1}^\epsilon(\tau)) \\
&\qquad - \overline{F}(X^\epsilon((i+1)\delta)))|\mathcal{F}_{(i+1)\delta}\Big]\Big| \\
&\quad =: B_1 + B_2.
\end{aligned}$$

Author's personal copy

(1) First, using Markov property we have for any $j\delta \le \tau \le (j+1)\delta$

$$\begin{aligned} B_1 &= C\mathbb{E}\left|\mathbb{E}\left[F(X^\epsilon((i+1)\delta), Z^\epsilon_{i+1}(\tau)) - \overline{F}(X^\epsilon((i+1)\delta))|\mathcal{F}_{(i+1)\delta}\right]\right| \\ &= C\mathbb{E}\left|\mathbb{E}\left[\tilde{F}\left(X^\epsilon((i+1)\delta), Z^\epsilon_{(i+1)}((i+1)\delta), \frac{\tau-(i+1)\delta}{\epsilon}\right)\middle|\mathcal{F}_{(i+1)\delta}\right]\right| \\ &\le Ce^{-c\frac{\tau-(i+1)\delta}{\epsilon}}, \end{aligned}$$

thanks to Proposition C.1.

(2) B_2 can be rewritten using a telescoping sum, and some conditional expectation

$$\begin{aligned} B_2 &= \mathbb{E}\left|\mathbb{E}\left(\sum_{k=i+1}^{j-1}\mathbb{E}\left[[F(X^\epsilon((k+1)\delta), Z^\epsilon_{k+1}(\tau)) - \overline{F}(X^\epsilon((k+1)\delta))]\right.\right.\right. \\ &\quad \left.\left.\left. - [F(X^\epsilon(k\delta), Z^\epsilon_k(\tau)) - \overline{F}(X^\epsilon(k\delta))]|\mathcal{F}_{k\delta}\right]|\mathcal{F}_{(i+1)\delta}\right)\right| \\ &\le \sum_{k=i+1}^{j-1}\mathbb{E}\left|\mathbb{E}\left[[F(X^\epsilon((k+1)\delta), Z^\epsilon_{k+1}(\tau)) - \overline{F}(X^\epsilon((k+1)\delta))]\right.\right. \\ &\quad \left.\left. - [F(X^\epsilon(k\delta), Z^\epsilon_k(\tau)) - \overline{F}(X^\epsilon(k\delta))]|\mathcal{F}_{k\delta}\right]\right|. \end{aligned}$$

Thanks to Markov property, we obtain

$$\begin{aligned} &\mathbb{E}\left[F(X^\epsilon((k+1)\delta), Z^\epsilon_{k+1}(\tau)) - \overline{F}(X^\epsilon((k+1)\delta))|\mathcal{F}_{k\delta}\right] \\ &\quad = \mathbb{E}\left[\tilde{F}(X^\epsilon((k+1)\delta), \tilde{Y}^\epsilon((k+1)\delta), \tau-(k+1)\delta, \epsilon)|\mathcal{F}_{k\delta}\right] \end{aligned}$$

and

$$\begin{aligned} &\mathbb{E}\left[F(X^\epsilon(k\delta), Z^\epsilon_k(\tau)) - \overline{F}(X^\epsilon(k\delta))|\mathcal{F}_{k\delta}\right] \\ &\quad = \mathbb{E}\left[\tilde{F}(X^\epsilon(k\delta), \tilde{Y}^\epsilon((k+1)\delta), \tau-(k+1)\delta, \epsilon)|\mathcal{F}_{k\delta}\right]. \end{aligned}$$

Using the exponential decrease in time of the Lipschitz constant of $\tilde{F}$ with respect to x given by Proposition C.2 in the Appendix, we have

$$\begin{aligned} B_2 &\le \sum_{k=i+1}^{j-1}\mathbb{E}\left|\mathbb{E}\left[\tilde{F}(X^\epsilon((k+1)\delta), \tilde{Y}^\epsilon((k+1)\delta), \tau-(k+1)\delta, \epsilon)|\mathcal{F}_{k\delta}\right]\right. \\ &\quad \left. - \mathbb{E}\left[\tilde{F}(X^\epsilon(k\delta), \tilde{Y}^\epsilon((k+1)\delta), \tau-(k+1)\delta, \epsilon)|\mathcal{F}_{k\delta}\right]\right| \\ &\le C\sum_{k=i+1}^{j-1} e^{-c\frac{\tau-(k+1)\delta}{\epsilon}}\mathbb{E}|X^\epsilon((k+1)\delta) - X^\epsilon(k\delta)|\left(1+\frac{\epsilon^\eta}{(\tau-(k+1)\delta)^\eta}\right) \\ &\le C\left(1+\frac{\epsilon^\eta}{(\tau-j\delta)^\eta}\right)\sum_{k=i+1}^{j-1} e^{-c\frac{\tau-(k+1)\delta}{\epsilon}}\frac{\delta^{1-r}}{(k\delta)^{1-r}} \\ &\le C\left(1+\frac{\epsilon^\eta}{(\tau-j\delta)^\eta}\right)\frac{\delta^{1-r}}{((i+1)\delta)^{1-r}}\frac{e^{-c\frac{\tau-j\delta}{\epsilon}}}{1-e^{-c\frac{\delta}{\epsilon}}}, \end{aligned}$$

by using the regularity proved in Proposition A.3.

Author's personal copy

(3) We are now able to conclude, by using the estimates on B_1 and B_2: we have for any $0 \le t \le T$

$$\begin{aligned}
\mathbb{E}A_2(t) \le\ & C \sum_{0\le i<j\le \lfloor \frac{t}{\delta}\rfloor} \int_{i\delta}^{(i+1)\delta\wedge t} ds \int_{j\delta}^{(j+1)\delta\wedge t} d\tau e^{-c\frac{\tau-(i+1)\delta}{\epsilon}} \\
& + C \sum_{0\le i<j\le \lfloor \frac{t}{\delta}\rfloor} \int_{i\delta}^{(i+1)\delta\wedge t} ds \int_{j\delta}^{(j+1)\delta\wedge t} d\tau \\
& \times \frac{\delta^{1-r}}{((i+1)\delta)^{1-r}} \frac{e^{-c\frac{\tau-j\delta}{\epsilon}}}{1-e^{-c\frac{\delta}{\epsilon}}} \left(1+\frac{\epsilon^\eta}{(\tau-j\delta)^\eta}\right) \\
\le\ & C \sum_{0\le i<j\le \lfloor \frac{T}{\delta}\rfloor} \epsilon\delta e^{-c\frac{(j-i-1)\delta}{\epsilon}} + C\delta^{1-r}\epsilon \\
& \times \frac{1}{1-e^{-c\frac{\delta}{\epsilon}}} \sum_{0\le i<j\le \lfloor \frac{T}{\delta}\rfloor} \frac{\delta^{1-r}}{((i+1)\delta)^{1-r}} \\
\le\ & C(T)\epsilon(1+\delta^{-r})\left(1+\frac{1}{1-e^{-c\frac{\delta}{\epsilon}}}\right). \quad \square
\end{aligned}$$

4. Proof of the weak-order result

The proof of Theorem 1.2 relies on an expansion of the solution of the Kolmogorov equation associated with the stochastic system (1.2) with respect to the small parameter ϵ; the zero-order term corresponds to the averaged equation, and we control the first-order term to get the result.

When working with SDEs, this strategy can be entirely followed; nevertheless in the case of SPDEs, the Kolmogorov equations involve the unbounded operators A and B, and quantities like $\mathrm{Tr}(D^2_{yy}u)$, and this leads to technical problems—see [3] or [5].

The idea of the proof is to reduce the infinite dimensional problem to a finite dimensional one by Galerkin approximation; we apply the method in this finite dimensional setting, and we prove bounds that are uniform with respect to the dimension. We also show that taking the limit when dimension goes to infinity is meaningful and gives the desired result.

The key element in the construction of the expansion mentioned above is given in Lemma 4.3: a Poisson equation can be solved under ergodicity conditions on the fast equation. We notice that Assumption 2.9 is sufficient, since we can analyse the problem through the asymptotic properties of the transition semi-group of the fast equation, instead of trajectories.

First, we explain how we reduce the problem to a finite dimensional one—see Section 4.1; then we explain the method in this setting and show which expressions must be controlled—see Section 4.2; finally we prove the estimates.

4.1. Reduction to a finite dimensional problem

We use Galerkin approximations based on the orthonormal basis (e_k) and (f_k) of H, given by Assumption 2.1. We define the subspaces

$$H_N^{(1)} = \mathrm{span}\,\{e_k;\, 0 \le k \le N-1\} \quad \text{and} \quad H_N^{(2)} = \mathrm{span}\,\{f_k;\, 0 \le k \le N-1\}.$$

We denote by $P_N^{(1)} \in \mathcal{L}(H)$ – resp. $P_N^{(2)}$ – the orthogonal projection of H onto $H_N^{(1)}$—resp. $H_N^{(2)}$.

Author's personal copy

Then we define for $x \in H_N^{(1)}$ and $y \in H_N^{(2)}$

$$\begin{aligned} F_N(x,y) &= P_N^{(1)}(F(x,y)) \\ G_N(x,y) &= P_N^{(2)}(G(x,y)) \\ U_N(x,y) &= U(x,y). \end{aligned}$$

On $H_N^{(1)} \times H_N^{(2)}$ coefficients F_N and G_N are of class $\mathcal{C}^2$. The function U_N is of class $\mathcal{C}^3$ with respect to y and of class $\mathcal{C}^2$ with respect to x, and we have $D_y U_N(x,y) = G_N(x,y)$ for any $(x,y) \in H_N^{(1)} \times H_N^{(2)}$.

Moreover we have bounds on F_N, G_N, U_N and on their derivatives which are uniform with respect to N and satisfy bound like in Assumption 2.5. In particular we still have the weak dissipativity condition with G replaced by G_N.

We can define the following approximation of system (1.2)

$$\begin{aligned} dX_N^\epsilon(t) &= (AX_N^\epsilon(t) + F_N(X_N^\epsilon(t), Y_N^\epsilon(t)))dt \\ dY_N^\epsilon(t) &= \frac{1}{\epsilon}(BY_N^\epsilon(t) + G_N(X_N^\epsilon(t), Y_N^\epsilon(t)))dt + \frac{1}{\sqrt{\epsilon}} dW_N(t), \end{aligned} \tag{4.1}$$

with initial conditions $X_N^\epsilon(0) = P_N^{(1)} x \in H_N^{(1)}$, $Y_N^\epsilon(0) = P_N^{(2)} y \in H_N^{(2)}$.

Since $H_N^{(1)}$ (resp. $H_N^{(2)}$) is a stable subspace of A (resp. B), this system is well-posed in $H_N^{(1)} \times H_N^{(2)}$.

We have by definition $W_N(t) = P_N^{(2)} W(t)$; on $H_N^{(2)}$ it is a N-dimensional Brownian motion.

Below we explain that system (4.1) defines a good approximation of the initial problem (1.2). Moreover, we can check that the structure of the problem remains the same, with bounds independent of the dimension.

First we describe the ergodic properties, and in particular the relations between the invariant measures associated with the associated fast equations with frozen slow component. In Section 2.2, we have defined $\nu = \mathcal{N}(0, \frac{(-B)^{-1}}{2})$ and μ^x – see (2.6); we can do the same for the finite dimensional fast equation with frozen slow component $x \in H$ – and not only for $x \in H_N^{(1)}$:

$$\begin{aligned} dY_{x,N}(t,y) &= (BY_{x,N}(t,y) + P_N^{(2)} G(x, P_N^{(2)} Y_{x,N}(t,y)))dt + dW_N(t) \\ Y_{x,N}(0,y) &= P_N^{(2)} y. \end{aligned} \tag{4.2}$$

Let ν_N be the unique centered Gaussian probability measure on $H_N^{(2)}$ having for covariance operator the induced matrix from $\frac{(B)^{-1}}{2}$ on the subspace $H_N^{(2)}$. We can then build μ_N^x the unique – since we have strict dissipativity – invariant probability measure associated to (4.2): we naturally extend the definition of U_N to $H \times H_N^{(2)}$, by $U_N(x,y) = U(x,y)$, and we define

$$\mu_N^x(dy) = \frac{1}{Z_N(x)} e^{2U_N(x,y)} \nu_N(dy), \tag{4.3}$$

where $Z_N(x) \in]0, +\infty[$ is a normalization constant.

As ν_N is the image measure of ν by $P_N^{(2)}$, we can use the following change of variables formula for suitable test functions Φ:

Author's personal copy

$$\int_{H_N^{(2)}} \Phi(y)\mu_N^x(dy) = \frac{1}{Z_N(x)} \int_{H_N^{(2)}} \Phi(y) e^{2U(x,y)} \nu_N(dy)$$
$$= \frac{1}{Z_N(x)} \int_H \Phi(P_N^{(2)} y) e^{2U(x, P_N^{(2)} y)} \nu(dy).$$

This formula leads to convergence properties: first when $N \to +\infty$ we have

$$Z_N(x) \to Z(x) \quad \text{for any } x \in H.$$

Moreover if $\Phi : H \to H$ is a continuous function such that for any $y \in H$, $|\Phi(y)| \leq c(1 + |y|)$, then

$$\int_{H_N^{(2)}} \Phi(y)\mu_N^x(dy) \to \int_H \Phi(y)\mu^x(dy).$$

For any $x \in H$ we define the averaged coefficient associated with the finite dimensional problem (4.2)

$$\overline{F_N}(x) = \int_{H_N^{(2)}} P_N^{(1)} F(x, y)\mu_N^x(dy), \tag{4.4}$$

and the new averaged equation

$$d\overline{X_N}(t) = (A\overline{X_N}(t) + \overline{F_N}(\overline{X_N}(t)))dt, \tag{4.5}$$

with initial condition $\overline{X_N}(0) = P_N^{(1)} x \in H_N^{(1)}$.

We notice that $\overline{F_N}$ is bounded, and of class $\mathcal{C}^2$ with bounded derivatives; moreover it satisfies the properties described in Proposition 2.16, with constants independent of N.

Remark 4.1. Making the Galerkin projection and then averaging the coefficient with μ_N^x is not the same as averaging the coefficient with μ^x and then making the Galerkin projection. As a consequence $\overline{X_N}$ is not naturally defined by a Galerkin approximation from $\overline{X}$.

The following lemma gives the convergence of the finite dimensional approximations to the initial problem:

Lemma 4.2. (1) *For any fixed $\epsilon > 0$, $t \geq 0$, and any $x \in H$, $y \in H$, we have when $N \to +\infty$*

$$\mathbb{E}|X^\epsilon(t) - X_N^\epsilon(t)|^2 + \mathbb{E}|Y^\epsilon(t) - Y_N^\epsilon(t)|^2 \to 0.$$

(2) *For any $t \geq 0$, $x \in H$, we have when $N \to +\infty$*

$$|\overline{X}(t) - \overline{X_N}(t)| \to 0.$$

It remains to define an approximated test function: for any $x \in H_N^{(1)}$, we define $\phi_N(x) = \phi(x)$.

We can now show how the initial problem can be reduced to a finite dimensional one—provided we are able to give estimates uniform to the dimension: for any initial conditions $x, y \in H$ and $0 \leq t \leq T$,

$$\mathbb{E}[\phi(X^\epsilon(t))] - \mathbb{E}[\phi(\overline{X}(t))] = \mathbb{E}[\phi(X^\epsilon(t))] - \mathbb{E}[\phi(X_N^\epsilon(t))]$$
$$+ \mathbb{E}[\phi_N(X_N^\epsilon(t))] - \mathbb{E}[\phi_N(\overline{X_N}(t))]$$
$$+ \mathbb{E}[\phi(\overline{X_N}(t))] - \mathbb{E}[\phi(\overline{X}(t))].$$

The first and the last terms converge to 0 when $N \to +\infty$, according to Lemma 4.2; we later control the central term with an expression independent of N. Taking the limit as dimension goes

Author's personal copy

to infinity then gives the result. So we have to control

$$\mathbb{E}[\phi_N(X_N^\epsilon(t))] - \mathbb{E}[\phi_N(\overline{X_N}(t))]. \tag{4.6}$$

From now on, we only work with the approximations to obtain estimates, but in order to simplify the notations, we forget the index N—since bounds are uniform with respect to N. The spaces $H_N^{(i)}$ are denoted by $H^{(i)}$ in the next sections.

4.2. The asymptotic expansion

We define the following differential operators: for functions of class $\mathcal{C}^2$ $\psi : H^{(1)} \times H^{(2)} \to \mathbb{R}$, for any $(x, y) \in H^{(1)} \times H^{(2)}$

$$\begin{aligned}
L_1\psi(x, y) &= \langle By + G(x, y), D_y\psi(x, y)\rangle + \frac{1}{2}\mathrm{Tr}(D_{yy}^2\psi(x, y)) \\
L_2\psi(x, y) &= \langle Ax + F(x, y), D_x\psi(x, y)\rangle \\
L^\epsilon &= \frac{1}{\epsilon}L_1 + L_2.
\end{aligned}$$

We also define for $\psi : H^{(1)} \to \mathbb{R}$ of class $\mathcal{C}^1$ $\overline{L}\psi(x) = \langle Ax + \overline{F}(x), D_x\psi(x)\rangle$.
We define the following functions u^ϵ and $\overline{u}$: for $x \in H^{(1)}$, $y \in H^{(2)}$ and $t \geq 0$

$$\begin{aligned}
u^\epsilon(t, x, y) &= \mathbb{E}[\phi(X^\epsilon(t, x, y))] \\
\overline{u}(t, x) &= \phi(\overline{X}(t, x)),
\end{aligned} \tag{4.7}$$

where we have mentioned explicitly the dependence on the initial conditions x, y in X^ϵ and $\overline{X}$.

Since the test function ϕ is of class $\mathcal{C}_b^2$, u^ϵ and $\overline{u}$ are of class $\mathcal{C}^1$ with respect to t and of class $\mathcal{C}_b^2$ with respect to x, y; we also know that u^ϵ and $\overline{u}$ are solutions of the following Kolmogorov equations:

$$\begin{aligned}
&\frac{\partial u^\epsilon}{\partial t}(t, x) = L^\epsilon u^\epsilon(t, x) \\
&u^\epsilon(0, x) = \phi(x);
\end{aligned} \tag{4.8}$$

$$\begin{aligned}
&\frac{\partial \overline{u}}{\partial t}(t, x, y) = \overline{L}\overline{u}(t, x, y) \\
&\overline{u}(0, x, y) = \phi(x).
\end{aligned} \tag{4.9}$$

We remark that the second equation is a linear transport equation with no diffusion term. We then rewrite the expression we want to study (see (4.6)):

$$\mathbb{E}[\phi(X^\epsilon(T, x, y))] - \mathbb{E}[\phi(\overline{X}(T, x))] = u^\epsilon(T, x, y) - \overline{u}(T, x). \tag{4.10}$$

Our strategy is to look for an expansion of u^ϵ with respect to the small parameter ϵ:

$$u^\epsilon = u_0 + \epsilon u_1 + v^\epsilon, \tag{4.11}$$

where v^ϵ is a residual term, while u_0 and u_1 are smooth and are constructed below.

The identification with respect to the powers of ϵ gives the following equations:

$$L_1 u_0 = 0, \tag{4.12}$$

$$\frac{\partial u_0}{\partial t} = L_1 u_1 + L_2 u_0. \tag{4.13}$$

Author's personal copy

The operator L_1 satisfies the following property on the solutions of Poisson equations:

Lemma 4.3. *We fix $x \in H$.*

- *If Ψ is a bounded continuous function such that $\int_H \Psi(y)\mu^x(dy) = 0$, then if Φ is a function of class $\mathcal{C}^2$ satisfying $L_1\Phi = -\Psi$ then for any $y \in H$ we have*
$$\Phi(y) = \int_H \Phi\mu^x + \int_0^{+\infty} \mathbb{E}[\Psi(Y_x(s, y))]ds.$$
- *Moreover if Ψ is of class $\mathcal{C}_b^2$, then Φ defined by*
$$\Phi(y) = \int_0^{+\infty} \mathbb{E}[\Psi(Y_x(s, y))]ds$$
is of class $\mathcal{C}^2$, satisfies $L_1\Phi = -\Psi$, and there exists a constant C– independent on N– such that for any $y \in H$ we have
$$|\Phi(y)| \leq C(1 + |y|^2)\|\Psi\|_\infty.$$

Proof. The first part of the lemma is an easy consequence of Itô formula and of Eq. (2.5), after integration with respect to y_2 under μ^x. To prove the second part of the lemma, we first see that for any fixed $s \in \mathbb{R}^+$ the function $y \mapsto \mathbb{E}[\Psi(Y_x(s, y))] =: v_x(s, y)$ is of class $\mathcal{C}^2$. To be able to exchange integration in s and derivation with respect to y, we need to prove an estimate of the first and the second derivatives which is integrable with respect to s. The derivatives of $Y_x(s, y)$ with respect to the initial condition satisfy the following equations—to simplify notations we do not write dependence in x in those derivatives:

$$\begin{aligned}
&\frac{d\tilde{\eta}^{h,y}(s)}{ds} = B\tilde{\eta}^{h,y}(s) + D_yG(x, Y_x(s, y)).\tilde{\eta}^{h,y}(s, y)\\
&\tilde{\eta}^{h,y}(0) = h,
\end{aligned}$$

and

$$\begin{aligned}
\frac{d\tilde{\zeta}^{h,k,y}(s)}{ds} &= B\tilde{\zeta}^{h,k,y}(s) + D_yG(x, Y_x(s, y)).\tilde{\zeta}^{h,k,y}(s)\\
&\quad + D^2_{yy}G(x, Y_x(s, y)) \cdot (\tilde{\eta}^{h,y}(s), \tilde{\eta}^{k,y}(s))\\
\tilde{\zeta}^{h,k,y}(0) &= 0.
\end{aligned}$$

Without any further dissipativity assumption than (WD), we only get bounds on finite time intervals like $[0, 1]$: there exists $C > 0$ such that for any $y, h, k \in H$ and $0 \leq s \leq 1$

$$\begin{aligned}
&|D_yv_x(s, y) \cdot h| \leq C|h|\\
&|D^2_{yy}v_x(s, y) \cdot (h, k)| \leq C|h|\,|k|.
\end{aligned}$$

However, using the estimate (2.5) and a Bismut–Elworthy–Li formula, we can indeed prove some exponential convergence with respect to s of the derivatives $D_yv_x(s, y)$ and $D^2_{yy}v_x(s, y)$.

Let Ψ^0 be a function such that $|\Psi^0(y)| \leq C(\Psi^0)(1 + |y|^2)$ for any $y \in H$. If we define $v^0_x(s, y) := \mathbb{E}\,\Psi^0(Y_x(s, y))$ for any $y \in H, h, k \in H$, we get for the first order derivative

$$D_yv^0_x(s, y) \cdot h = \frac{1}{s}\mathbb{E}\left[\int_0^s \langle\tilde{\eta}^{h,y}(\sigma), dW(\sigma)\rangle\,\Psi^0(Y_x(s, y))\right]$$

Author's personal copy

$$= \frac{2}{s}\mathbb{E}\left[\int_0^{s/2} \langle \tilde{\eta}^{h,y}(\sigma), dW(\sigma)\rangle v_x^0(s/2, Y_x(s/2, y))\right], \tag{4.14}$$

with the observation that $v_x^0(s, y) = \mathbb{E}v_x^0(s/2, Y_x(s/2, y))$ thanks to the Markov property; the second order derivative satisfies

$$\begin{aligned} D_{yy}^2 v_x^0(s, y)\cdot(h, k) = {} & \frac{2}{s}\mathbb{E}\left[\int_0^{s/2} \tilde{\zeta}^{h,k,y}(\sigma), dW(\sigma) > v_x^0(s/2, Y(s/2, y))\right] \\ & + \frac{2}{s}\mathbb{E}\left[\int_0^{s/2} \langle \tilde{\eta}^{h,y}(\sigma), dW(\sigma)\rangle \right. \\ & \left. \times\ D_y v_x^0(s/2, Y(s/2))\cdot \tilde{\eta}^{k,y}(s/2)\right]. \end{aligned} \tag{4.15}$$

Since Y_x can be controlled in a L^2 norm according to Proposition A.1 in the Appendix, we see that there exists $C > 0$ such that for any $0 < s \le 1$, $y \in H$, $h, k \in H$

$$\begin{aligned} |D_y v_x^0(s, y)\cdot h| &\le \frac{C}{\sqrt{s}} C(\Psi^0)(1 + |y|^2)|h|, \\ |D_{yy}^2 v_x^0(s, y)\cdot(h, k)|\,| &\le \frac{C}{s} C(\Psi^0)(1 + |y|^2)|h|\,|k|. \end{aligned} \tag{4.16}$$

Now when $s \ge 1$ the Markov property implies that $v_x(s, y) = \mathbb{E}v_x(s - 1, Y_x(1, y))$, and choosing $y_1 = y$ and by integrating with respect to $\mu^x(dy_2)$ in (2.5) we have

$$|v_x(s - 1, y)| \le Ce^{-c(s-1)}(1 + |y|^2).$$

By (4.16) at time 1, we obtain for $s \ge 1$

$$\begin{aligned} |D_y v_x(s, y)\cdot h| &\le Ce^{-c(s-1)}(1 + |y|^2)|h| \\ |D_{yy}^2 v_x(s, y)\cdot(h, k)| &\le Ce^{-c(s-1)}(1 + |y|^2)|h|\,|k|. \end{aligned}$$

Moreover we have a uniform control when $0 \le s \le 1$, so that with a change of constants we get the result. □

As a consequence we see from (4.12) that u_0 is independent of y; we then write $u_0(t, x, y) = u_0(t, x)$. We also choose the initial condition $u_0(0, x) = \phi(x)$. The second equation (4.13) then yields

$$\begin{aligned} \frac{\partial u_0}{\partial t}(t, x) &= \int_{H^{(2)}} \frac{\partial u_0}{\partial t}(t, x)\mu^x(dy) \\ &= \int_{H^{(2)}} L_1 u_1(t, x, y)\mu^x(dy) + \int_{H^{(2)}} L_2 u_0(t, x)\mu^x(dy) \\ &= \left\langle Au_0(t, x) + \int_{H^{(2)}} F(x, y)\mu^x(dy), D_x u_0(t, x)\right\rangle \\ &= \overline{L}u_0(t, x). \end{aligned}$$

u_0 and $\overline{u}$ are solutions of the same evolution equation, with the same initial condition; we can then conclude that $u_0 = \overline{u}$.

Author's personal copy

Then the second equation can be transformed: $\overline{L}u_0 = L_1u_1 + L_2u_0$. We then obtain an equation on u_1:

$$L_1u_1(t,x,y) = \langle \overline{F}(x) - F(x,y), D_xu_0(t,x)\rangle =: -\chi(t,x,y), \tag{4.17}$$

where χ is of class $\mathcal{C}_b^2$ with respect to y, and satisfies for any $t \geq 0$ and $x \in H^{(1)}$

$$\int_{H^{(2)}} \chi(t,x,y)\mu^x(dy) = 0.$$

Thanks to Lemma 4.3, we thus obtain the following solution to equation (4.17)

$$u_1(t,x,y) = \int_0^{+\infty} \mathbb{E}[\chi(t,x,Y_x(s,y))]ds. \tag{4.18}$$

Moreover we are able to show regularity of u_1 with respect to t and x, y.

The remainder $v^\epsilon = u^\epsilon - u_0 - \epsilon u_1$ satisfies

$$\left(\partial_t - \frac{1}{\epsilon}L_1 - L_2\right)v^\epsilon = \epsilon\left(L_2u_1 - \frac{\partial u_1}{\partial t}\right). \tag{4.19}$$

Due to non-integrability in 0 of some bounds below, we introduce a parameter $\rho(\epsilon) = \epsilon^{1/\theta} \leq \epsilon$ (since $0 < \theta \leq 1$); it satisfies $\rho(\epsilon) \to 0$ when $\epsilon \to 0$.

Using a variation of constant formula, we obtain

$$\begin{aligned} v^\epsilon(T,x,y) = {} & \mathbb{E}[v^\epsilon(\rho(\epsilon), X^\epsilon(T-\rho(\epsilon),x,y), Y^\epsilon(T-\rho(\epsilon),x,y))] \\ & + \epsilon\mathbb{E}\left[\int_{\rho(\epsilon)}^T \left(L_2u_1 - \frac{\partial u_1}{\partial t}\right)(t, X^\epsilon(T-t,x,y), Y^\epsilon(T-t,x,y))dt\right]. \end{aligned}$$

By (4.11), and since $u_0 = \overline{u}$, we then have

$$\begin{aligned} & u^\epsilon(T,x,y) - \overline{u}(T,x,y) \\ & \quad = \epsilon u_1(T,x,y) \\ & \qquad + \mathbb{E}[v^\epsilon(\rho(\epsilon), X^\epsilon(T-\rho(\epsilon),x,y), Y^\epsilon(T-\rho(\epsilon),x,y))] \\ & \qquad + \epsilon\mathbb{E}\left[\int_{\rho(\epsilon)}^T \left(L_2u_1 - \frac{\partial u_1}{\partial t}\right)(t, X^\epsilon(T-t,x,y), Y^\epsilon(T-t,x,y))dt\right]. \end{aligned} \tag{4.20}$$

The following estimates are proved below:

Lemma 4.4. *There exists a constant C such that for any $0 < t \leq T$, $x, y \in H$,*

$$\begin{aligned} & |u_1(t,x,y)| \leq C(1 + |x| + |y|) \\ & \left|\frac{\partial u_1}{\partial t}(t,x,y)\right| \leq C\left(1 + \frac{1}{t}\right)(1 + |x| + |y|)^2 \\ & |L_2u_1(t,x,y)| \leq C(1 + |x| + |y|)(1 + |Ax|). \end{aligned}$$

Using estimates on X^ϵ and Y^ϵ proved in the Appendix (see Propositions A.1 and A.5) the first and the last expressions of (4.20) are bounded by

$$C\epsilon(1 + |x| + |y|) + C\epsilon^{1-r/2}(1 + |\log(\rho(\epsilon))|)(1 + |x|_{(-A)^\theta} + |y|)^2,$$

which is dominated by $C\epsilon^{1-r}$ with the choice of $\rho(\epsilon)$ given above.

Author's personal copy

We notice that the Assumption $\theta > 0$ is essential to control the part involving $|Ax|$.

We now explain how the central term of (4.20) is controlled; for that we estimate for any $x, y \in H$

$$\begin{aligned} v^\epsilon(\rho(\epsilon), x, y) &= u^\epsilon(\rho(\epsilon), x, y) - u_0(\rho(\epsilon), x) - \epsilon u_1(\rho(\epsilon), x, y) \\ &= -\epsilon u_1(\rho(\epsilon), x, y) + [u^\epsilon(\rho(\epsilon), x, y) - u^\epsilon(0, x, y)] \\ &\quad - [u_0(\rho(\epsilon), x) - u_0(0, x)], \end{aligned} \tag{4.21}$$

since the initial condition ϕ is the same for u^ϵ and $\overline{u}$.

Using Lemma 4.4, the first term above is easily controlled by $C\epsilon(1 + |x| + |y|)$. We now use another method to control the two other terms.

First, we use the definition (4.7) of $\overline{u} = u_0$ to write

$$\begin{aligned} |u_0(\rho(\epsilon), x) - u_0(0, x)| &= \left| \int_0^{\rho(\epsilon)} \frac{\partial}{\partial t} u_0(t, x) dt \right| \\ &= \left| \int_0^{\rho(\epsilon)} \frac{\partial}{\partial t} \phi(\overline{X}(t, x)) dt \right| \\ &= \left| \int_0^{\rho(\epsilon)} D\phi(\overline{X}(t, x)).\frac{d}{dt}\overline{X}(t, x) dt \right| \\ &\leq C \int_0^{\rho(\epsilon)} \left| \frac{d}{dt}\overline{X}(t, x) \right| dt. \end{aligned}$$

By definition of $\overline{X}$ (see (1.3)), and using Proposition B.3, we get for any $t > 0$

$$\left| \frac{d}{dt}\overline{X}(t, x) \right| \leq C(1 + t^{\theta-1})(1 + |x|_{(-A)^\theta}).$$

As a consequence, since $\theta > 0$, we get

$$|u_0(\rho(\epsilon), x) - u_0(0, x)| \leq C\left(\rho(\epsilon) + \frac{\rho(\epsilon)^\theta}{\theta}\right)(1 + |x|_{(-A)^\theta}).$$

The other expression is controlled in the same way; it is important to notice that the assumption that ϕ only depends on the slow variable x is fundamental in this estimate.

$$\begin{aligned} |u^\epsilon(\rho(\epsilon), x, y) - u^\epsilon(0, x, y)| &= \left| \int_0^{\rho(\epsilon)} \frac{\partial}{\partial t} u^\epsilon(t, x, y) dt \right| \\ &= \left| \int_0^{\rho(\epsilon)} \frac{\partial}{\partial t} \mathbb{E}[\phi(X^\epsilon(t, x, y))] dt \right| \\ &= \left| \int_0^{\rho(\epsilon)} \mathbb{E}\left[D\phi(X^\epsilon(t, x, y)).\frac{d}{dt}X^\epsilon(t, x, y) \right] dt \right| \\ &\leq C \int_0^{\rho(\epsilon)} (\mathbb{E}|AX^\epsilon(t, x, y)| + 1) dt. \end{aligned}$$

Author's personal copy

We now use the estimate on $\mathbb{E}|AX^\epsilon(t,x,y)|$ of Proposition A.5 in the Appendix, and we obtain

$$|u^\epsilon(\rho(\epsilon),x,y)-u^\epsilon(0,x,y)| \le C\left(\rho(\epsilon)+\frac{\rho(\epsilon)^\theta}{\theta}+\epsilon^{-r/2}\rho(\epsilon)\right)(1+|x|_{(-A)^\theta}+|y|).$$

Then by (4.21), and using $\rho(\epsilon)=\epsilon^{1/\theta}\le\epsilon$, we get

$$|v^\epsilon(\rho(\epsilon),x,y)| \le C\epsilon^{1-r/2}(1+|x|_{(-A)^\theta}+|y|).$$

Then thanks to (4.20) and to Proposition A.2, we get for $\epsilon\le 1$

$$|u^\epsilon(T,x,y)-\overline{u}(T,x,y)| \le C\epsilon^{1-r}.$$

As explained at the end of Section 4.1, we have indeed proved a bound on (4.6). It is now enough to notice that the above constant C is independent of dimension N, and to let N go to $+\infty$, and Theorem 1.2 follows.

5. Proof of Lemma 4.4

We use results gathered in the Appendices A and B.

5.1. Estimate of u_1

Since u_1 is defined by (4.18), by using Lemma 4.3 we have

$$|u_1(t,x,y)| \le C(1+|y|^2)\|y\mapsto\chi(t,x,y)\|_\infty.$$

According to (4.17), we indeed have for any $y\in H^{(2)}$

$$\chi(t,x,y)=\langle F(x,y)-\overline{F}(x),D_xu_0(t,x)\rangle,$$

and therefore we just have to bound $|D_xu_0(t,x)|$, thanks to the following lemma:

Lemma 5.1. *For any $T\in]0,+\infty[$, there exists $C_0>0$ such that for any $0\le t\le T$ and $x\in H^{(1)}$*

$$|D_xu_0(t,x)|_H \le C_T\sup_{z\in H}|D\phi(z)|_H.$$

Proof. u_0 is the solution of the equation

$$\begin{aligned}&\frac{\partial u_0}{\partial t}(t,x)=\langle Ax+\overline{F}(x),D_xu_0(t,x)\rangle\\&u_0(0,x)=\phi(x).\end{aligned}\tag{5.1}$$

We have a representation formula $u_0(t,x)=\phi(\overline{X}(t,x))$, where $\overline{X}$ is solution of (1.3).
We can differentiate (1.3) with respect to the initial condition x, and we have for any $h\in H^{(1)}$

$$D_xu_0(t,x)\cdot h=D\phi(\overline{X}(t,x))\cdot\eta^h(t,x),$$

where $\eta^h(t,x)$ is the derivative of $\overline{X}$ with respect to x in direction h, and is solution of the variational equation (B.1) in the Appendix.

Using Proposition B.4, we get $|D_xu_0(t,x)\cdot h|\le C_T\sup_{z\in H}|D\phi(z)|\,|h|$, and taking the supremum over h gives the result. □

Therefore, we obtain the first estimate of Lemma 4.4.

Author's personal copy

5.2. Estimate of $\frac{\partial u_1}{\partial t}$

First we check that

$$\int_{H^{(2)}} \frac{\partial \chi}{\partial t}(t,x,z)\mu^x(dz) = \frac{\partial}{\partial t}\int_{H^{(2)}} \chi(t,x,z)\mu^x(dz) = 0.$$

By definition (4.18) of u_1, it is easy to show that we can differentiate with respect to t, and that

$$\frac{\partial u_1}{\partial t}(t,x,y) = \int_0^{+\infty} \mathbb{E}\left[\frac{\partial \chi}{\partial t}(t,x,Y_x(s,y))\right] ds. \tag{5.2}$$

We then obtain

$$\left|\frac{\partial u_1}{\partial t}(t,x,y)\right| \leq C(1+|y|^2)\left\|y \mapsto \frac{\partial \chi}{\partial t}(t,x,y)\right\|_\infty.$$

Since by (4.17) we have

$$\frac{\partial \chi}{\partial t}(t,x,y) = \left\langle F(x,y) - \overline{F}(x), \frac{\partial}{\partial t} D_x u_0(t,x)\right\rangle,$$

we just need to control $|\frac{\partial}{\partial t} D_x u_0(t,x)|$:

Lemma 5.2. *For any $T > 0$, there exists $C_T > 0$ such that for any $0 < t \leq T$, $x \in H^{(1)}$ and $h \in H^{(1)}$ we have*

$$\left|\frac{\partial}{\partial t} D_x u_0(t,x)\cdot h\right| \leq C(1+|x|_H)(1+t^{-1})|h|.$$

Proof. For any $h \in H^{(1)}$, we have

$$\frac{\partial}{\partial t}(D_x u_0(t,x)\cdot h) = D^2\phi(\overline{X}(t,x))\left(\eta^h(t,x), \frac{d}{dt}\overline{X}(t,x)\right) + D\phi(\overline{X}(t,x))\cdot \frac{d}{dt}\eta^h(t,x).$$

(1) Thanks to Proposition B.1, we have $|\eta^h(t,x)| \leq C(1+|x|)$ for any $t \geq 0$.
Moreover $\frac{d}{dt}\overline{X}(t,x) = A\overline{X}(t,x) + \overline{F}(\overline{X}(t,x))$.
On the one hand, $\overline{F}$ is bounded; on the other hand, thanks to Proposition B.3 we have

$$|A\overline{X}(t,x)|_H \leq C_\theta(1+t^{-1})(1+|x|_H).$$

Therefore

$$\left|\frac{d}{dt}\overline{X}(t,x)\right|_H \leq C(1+t^{-1})(1+|x|_H).$$

(2) It remains to control

$$\left|\frac{d}{dt}\eta^h(t,x)\right| = |A\eta^h(t,x) + D\overline{F}(\overline{X}(t,x))\cdot\eta^h(t,x)|.$$

Since $\overline{F}$ is Lipschitz continuous, and using Proposition B.1, we get an estimate of the second term.
Moreover Proposition B.6 gives

$$|A\eta^h(t,x)| \leq C(t^{-1}+1)(1+|x|_H)|h|.$$

Author's personal copy

Therefore

$$\left|\frac{d}{dt}\eta^h(t,x)\right| \leq C(t^{-1}+1)(1+|x|_H)|h|.$$

(3) We then have for any $h \in H^{(1)}$

$$\left|\frac{\partial}{\partial t}(D_x u_0(t,x)\cdot h)\right| \leq C(t^{-1}+1)(1+|x|_H)|h|. \quad \square$$

We then obtain the second estimate of Lemma 4.4.

5.3. Estimate of L_2u_1

To prove Lemma 4.4, it remains to control the part involving L_2u_1.

By definition of L_2, we have

$$L_2u_1(t,x,y) = \langle Ax + F(x,y), D_x u_1(t,x,y)\rangle. \tag{5.3}$$

Therefore we have to estimate $|D_x u_1(t,x,y)|$. We explain how $D_x u_1(t,x,y)\cdot h$ can be calculated for any $h \in H^{(1)}$.

Recall that u_1 defined by (4.18) satisfies

$$L_1u_1(t,x,y) = \langle \overline{F}(x) - F(x,y), D_x u_0(t,x)\rangle = -\chi(t,x,y),$$

where we explicitly write the dependence of the operator L_1 in the two variables x and y.

We fix $t \geq 0$, $x \in H^{(1)}$, $y \in H^{(2)}$, and $h \in H^{(1)}$. Then for any $\xi \neq 0$ we have

$$\begin{aligned}
&L_1(x,y)\frac{u_1(t,x+\xi h,y)-u_1(t,x,y)}{\xi}\\
&\quad= -\frac{\chi(t,x+\xi h,y)-\chi(t,s,y)}{\xi}\\
&\qquad-\left\langle\frac{G(x+\xi h,y)-G(x,y)}{\xi}, D_y u_1(t,x+\xi h,y)\right\rangle\\
&\quad=: -\Gamma(t,x,y,h,\xi),
\end{aligned}$$

where Γ is regular with respect to y; therefore by using Lemma 4.3 we get

$$\begin{aligned}
&\frac{u_1(t,x+\xi h,y)-u_1(t,x,y)}{\xi} - \int_{H^{(2)}}\frac{u_1(t,x+\xi h,y)-u_1(t,x,y)}{\xi}\mu^x(dy)\\
&\quad= \int_0^{+\infty}\mathbb{E}[\Gamma(t,x,Y_x(s,y),h,\xi)]ds.
\end{aligned}$$

We want to take the limit when $\xi \to 0$, in order to prove that we can differentiate, and to obtain an expression that we are able to control.

First, we notice that for any t,x,y we have $\int_{H^{(2)}} u_1(t,x,y)\mu^x(dy) = 0$; so we can write that

$$\begin{aligned}
&\int_{H^{(2)}}\frac{u_1(t,x+\xi h,y)-u_1(t,x,y)}{\xi}\mu^x(dy)\\
&\quad= -\int_{H^{(2)}} u_1(t,x+\xi h,y)\frac{V(x+\xi h,y)-V(x,y)}{\xi}\nu(dy),
\end{aligned}$$

where $V(x,y) := \frac{1}{Z(x)}e^{2U(x,y)}$ (so that we have $\mu^x(dy) = V(x,y)\nu(dy)$).

Author's personal copy

When $\xi \to 0$, we obtain

$$\int_{H^{(2)}} \frac{u_1(t, x+\xi h, y) - u_1(t, x, y)}{\xi} \mu^x(dy) \to \int_{H^{(2)}} u_1(t, x, y) D_x V(x, y) \cdot h \nu(dy) = \int_{H^{(2)}} u_1(t, x, y) H(x, y) \cdot h V(x, y) \nu(dy),$$

where $H(x, y) = 2D_x U(x, y) - 2\int_{H^{(2)}} D_x U(x, z)\mu^x(dz)$.

Moreover $|\int_{H^{(2)}} u_1(t, x, y) H(x, y) \cdot h V(x, y) \nu(dy)| \leq C(1 + |x|)|h|$.

Second, we look at the part involving Γ: we notice that when $\xi \to 0$,

$$\Gamma(t, x, y, h, \xi) \to \Theta(t, x, y) \cdot h,$$

where

$$\Theta(t, x, y) \cdot h = D_x \chi(t, x, y) \cdot h + \langle D_x G(x, y) \cdot h, D_y u_1(t, x, y) \rangle. \tag{5.4}$$

Below, we prove the following estimate on the function Θ:

Lemma 5.3. *There exists a constant C such that for any $x \in H^{(1)}$, $t \geq 0$, $h \in H^{(1)}$ we have for any $y \in H^{(2)}$*

$$|\Theta(t, x, y) \cdot h| \leq C(1 + |y|^2)|h|.$$

We notice that for any t, x, ξ, h we have by definition of Γ $\int_{H^{(2)}} \Gamma(t, x, y, \xi, h)\mu^x(dy) = 0$; then using the bound of the previous lemma and the dominated convergence theorem we obtain $\int_{H^{(2)}} \Theta(t, x, y) \cdot h \mu^x(dy) = 0$ for any $x \in H^{(1)}, t \geq 0, h \in H^{(1)}$. Using this result, Proposition 2.13 – with integration with respect to $\mu^x(dy_2)$ – and the estimate in the previous lemma, we then see that u_1 can be differentiated with respect to x, and that the following formula holds:

$$D_x u_1(t, x, y) \cdot h = \int_{H^{(2)}} u_1(t, x, y) H(x, y) \cdot h V(x, y) \nu(dy) + \int_0^{+\infty} \mathbb{E}[\Theta(t, x, Y_x(s, y)) \cdot h] ds; \tag{5.5}$$

According to Lemma 5.3, we do not know whether Θ is a bounded function, but we only know that it has quadratic growth. However, the result of Proposition 2.13 can easily be extended to such function.

Now we obtain that

$$|D_x u_1(t, x, y) \cdot h| \leq C(1 + |y|^2)|h|$$

and therefore (see (5.3))

$$|L_2 u_1(t, x, y)| \leq C(1 + |y|^2)(1 + |Ax|),$$

which is the third estimate of Lemma 4.4.

It remains to prove Lemma 5.3.

We fix $t \geq 0$, $x \in H^{(1)}$, $h \in H^{(1)}$, and $y, y' \in H^{(2)}$.

Author's personal copy

- On the one hand, χ being defined by (4.17), we have

$$D_x\chi(t,x,y)\cdot h = \langle D_xF(x,y)\cdot h, D_xu_0(t,x)\rangle + D^2_{xx}u_0(t,x)\cdot(h,F(x,y)).$$

Using the boundedness of the first derivative of F, and Lemma 5.1, we easily have

$$|\langle D_xF(x,y)\cdot h, D_xu_0(t,x)\rangle| \le C|h|.$$

The other part can be controlled thanks to the following lemma:

Lemma 5.4. *For any $0 \le t \le T$, $x \in H^{(1)}$, $h,k \in H^{(1)}$, we have*

$$|D^2_{xx}u_0(t,x)\cdot(h,k)| \le C(T,\phi)|h|_H|k|_H.$$

Proof. We have

$$u_0(t,x) = \phi(\overline{X}(t,x))$$
$$D_xu_0(t,x)\cdot h = D\phi(\overline{X}(t,x))\cdot(D_x\overline{X}(t,x)\cdot h)$$

and

$$D^2_{xx}u_0(t,x)\cdot(h,k) = D^2\phi(\overline{X}(t,x))(D_x\overline{X}(t,x)\cdot h, D_x\overline{X}(t,x).k) + D\phi(\overline{X}(t,x))\cdot(D^2_{xx}\overline{X}(t,x)\cdot(h,k)).$$

Using Proposition B.4, we control $\eta^h(t,x) = D_x\overline{X}(t,x)\cdot h$; moreover we notice that the second derivative $\xi^{h,k}(t,x) := D^2_{xx}\overline{X}(t,x)\cdot(h,k)$ satisfies Eq. (B.2); using Proposition B.7, we get the result. □

Therefore $|D_x\chi(t,x,y)\cdot h| \le C|h|$.

- On the other hand,

$$|\langle D_xG(x,y)\cdot h, D_yu_1(t,x,y)\rangle| \le C|h|\,|D_yu_1(t,x,y)|,$$

But we have proved in Lemma 4.3 how to control the derivatives of u_1 with respect to y: we obtain

$$|D_yu_1(t,x,y)\cdot h| \le C(1+|y|^2)|h|.$$

Therefore we have

$$|\langle D_xG(x,y)\cdot h, D_yu_1(t,x,y)\rangle| \le C(1+|y|^2)|h|,$$

and now the result is easily obtained:

$$|\Theta(t,x,y)\cdot h| \le C(1+|y|^2)|h|.$$

Appendix A. Properties of (X^ϵ, Y^ϵ)

The results of this section only require Assumptions 2.1, 2.5 and 2.7; in particular no dissipativity is assumed.

The first important property is the control of moments of any order:

Proposition A.1. *For any $1 \le p < +\infty$, there exists $c_p > 0$ such that for any $(x,y) \in H^2$, $t \ge 0$ and $\epsilon > 0$*

$$\mathbb{E}[|X^\epsilon(t)|^p_H] \le c_p(1+e^{-\lambda t}|x|^p_H) \quad \text{and} \quad \mathbb{E}[|Y^\epsilon(t)|^p_H] \le c_p(1+e^{-\mu t}|y|^p_H).$$

Author's personal copy

We can also give bounds on the moments with respect to $|\,.\,|_{(-A)^a}$ and $|\,.\,|_{(-B)^b}$ norms, for $0 < a < 1$ and $0 < b < 1/4$ (the case $a = 1$ is treated in Proposition A.5).

Proposition A.2. *For any $p \geq 1$, $a \in (0,1)$, $b \in (0,1/4)$, there exists $C_{p,a,b} > 0$ such that for any $x \in D(-A)^a$ and $y \in D(-B)^b$, we have:*

$$\mathbb{E}|X^\epsilon(t,x,y)|^p_{(-A)^a} \leq C_p(1+|x|^p_{(-A)^a}) \quad \text{and}$$
$$\mathbb{E}|Y^\epsilon(t,x,y)|^p_{(-B)^b} \leq C_p(1+|y|^p_{(-B)^b}).$$

We now give some regularity estimates of X^ϵ and Y^ϵ in the time variable. We do not assume any regularity assumption on x or y; as a consequence, we obtain singularities at the origin, which are integrable.

Proposition A.3. *For any $0 < r < 1$, there exists $C_r > 0$ such that for any $x, y \in H$, for any $0 < s \leq t$ and $\epsilon > 0$ we have*

$$\left(\mathbb{E}|X^\epsilon(t) - X^\epsilon(s)|^2_H\right)^{1/2} \leq C_r|t-s|^{1-r}\left(1+\frac{1}{s^{1-r}}\right)(1+|x|_H).$$

Proof. If we fix $0 < s \leq t, x, y \in H$, we have

$$X^\epsilon(t) - X^\epsilon(s) = e^{tA}x - e^{sA}x + \int_0^t e^{(t-\sigma)A}F(X^\epsilon(\sigma), Y^\epsilon(\sigma))d\sigma - \int_0^s e^{(s-\sigma)A}F(X^\epsilon(\sigma), Y^\epsilon(\sigma))d\sigma.$$

For the first term, if $x = \sum_{k=0}^{+\infty} x_k e_k$, we can use Proposition 2.4 to get

$$|e^{tA}x - e^{sA}x|_H \leq C_r\frac{(t-s)^{1-r}}{s^{1-r}}|x|_H.$$

For the second term, we use the following decomposition:

$$\int_0^t e^{(t-\sigma)A}F(X^\epsilon(\sigma), Y^\epsilon(\sigma))d\sigma - \int_0^s e^{(s-\sigma)A}F(X^\epsilon(\sigma), Y^\epsilon(\sigma))d\sigma$$
$$= \int_s^t e^{(t-\sigma)A}F(X^\epsilon(\sigma), Y^\epsilon(\sigma))d\sigma + \int_0^s (e^{(t-\sigma)A} - e^{(s-\sigma)A})F(X^\epsilon(\sigma), Y^\epsilon(\sigma))d\sigma.$$

First, by the Cauchy–Schwarz inequality, we have

$$\mathbb{E}\left|\int_s^t e^{(t-\sigma)A}F(X^\epsilon(\sigma), Y^\epsilon(\sigma))d\sigma\right|^2_H \leq (t-s)\mathbb{E}\int_s^t |e^{(t-\sigma)A}F(X^\epsilon(\sigma), Y^\epsilon(\sigma))|_H d\sigma \leq C(t-s)^2,$$

since F is assumed to be bounded.

Second, we use the second inequality of Proposition 2.4 to control the last expression:

$$\mathbb{E}\left|\int_0^s e^{(s-\sigma)A}(e^{(t-s)A} - I)F(X^\epsilon(\sigma), Y^\epsilon(\sigma))d\sigma\right|^2_H$$
$$\leq \mathbb{E}\left[\int_0^s \left|\left(e^{(t-\sigma)A} - e^{(s-\sigma)A}\right)F(X^\epsilon(\sigma), Y^\epsilon(\sigma))d\sigma\right|_H\right]^2$$

Author's personal copy

$$\leq C_r^2(t-s)^{2(1-r)}\mathbb{E}\left(\int_0^s \frac{e^{-\frac{\lambda}{2}(s-\sigma)}}{(s-\sigma)^{1-r}}|F(X^\epsilon(\sigma),Y^\epsilon(\sigma))|_H d\sigma\right)^2$$
$$\leq C_r^2(t-s)^{2(1-r)},$$

since $\int_0^{+\infty}\frac{e^{-\frac{\lambda}{2}s}}{s^{1-r}}ds<+\infty$. □

Proposition A.4. *For any $0<r<1/4$, there exists a constant C_r such that if $x,y\in H$, then for any $0<s<t$ and $\epsilon>0$*

$$\mathbb{E}|Y^\epsilon(t)-Y^\epsilon(s)|^2\leq C(1+|x|_H^2+|y|_H^2)\left[\left(\frac{t-s}{s}\right)^{2r}+\left(\frac{t-s}{\epsilon}\right)^{2r}\right].$$

Proof. • For any $0<s<t$,

$$Y^\epsilon(t)-Y^\epsilon(s)=\left(e^{\frac{t}{\epsilon}B}-e^{\frac{s}{\epsilon}B}\right)y+\frac{1}{\epsilon}\int_s^t e^{\frac{(t-\sigma)}{\epsilon}B}G(X^\epsilon(\sigma),Y^\epsilon(\sigma))d\sigma$$
$$+\frac{1}{\epsilon}\int_0^s\left(e^{\frac{(t-\sigma)}{\epsilon}B}-e^{\frac{(s-\sigma)}{\epsilon}B}\right)G(X^\epsilon(\sigma),Y^\epsilon(\sigma))d\sigma$$
$$+W^{\epsilon,B}(t)-W^{\epsilon,B}(s),$$

where $W^{\epsilon,B}(r)=\frac{1}{\sqrt{\epsilon}}\int_0^r e^{\frac{(r-\sigma)}{\epsilon}B}dW(\sigma)$. We remark that only the last expression cannot bounded almost surely (since we assume that G is bounded).

- For the first term, using the second inequality of Proposition 2.4, we have for any $0<s<t$

$$\left|\left(e^{\frac{t}{\epsilon}B}-e^{\frac{s}{\epsilon}B}\right)y\right|_H\leq C_r\left(\frac{t-s}{s}\right)^r|y|_H.$$

- For the second term, we have for any $0<s<t$

$$\left|\frac{1}{\epsilon}\int_s^t e^{\frac{(t-\sigma)}{\epsilon}B}G(X^\epsilon(\sigma),Y^\epsilon(\sigma))d\sigma\right|\leq\frac{1}{\epsilon}\int_s^t|e^{\frac{(t-\sigma)}{\epsilon}B}|_{\mathcal{L}(H)}\|G\|_\infty d\sigma$$
$$\leq\frac{C}{\epsilon}\int_s^t e^{-\mu(t-\sigma)/\epsilon}d\sigma$$
$$\leq C\int_0^{(t-s)/\epsilon}e^{-\mu\sigma}d\sigma$$
$$\leq C\frac{(t-s)^r}{\epsilon^r}.$$

- For the third term, we use the second estimate of Proposition 2.4, and we have for any $0<s<t$

$$\left|\frac{1}{\epsilon}\int_0^s\left(e^{\frac{(t-\sigma)}{\epsilon}B}-e^{\frac{(s-\sigma)}{\epsilon}B}\right)G(X^\epsilon(\sigma),Y^\epsilon(\sigma))d\sigma\right|$$
$$\leq\frac{1}{\epsilon}\int_0^s|e^{\frac{(t-\sigma)}{\epsilon}B}-e^{\frac{(s-\sigma)}{\epsilon}B}|_{\mathcal{L}(H)}\|G\|_\infty d\sigma$$
$$\leq\frac{C_r}{\epsilon}\int_0^s\frac{(t-s)^r}{(s-\sigma)^r}e^{-\frac{\mu(s-\sigma)}{2\epsilon}}d\sigma$$
$$\leq C_r\frac{(t-s)^r}{\epsilon^r}\int_0^{+\infty}\frac{1}{\sigma^r}e^{-\frac{\mu\sigma}{2}}d\sigma.$$

Author's personal copy

- For the fourth term, we have for any $0 < s < t$,

$$\begin{aligned}
\mathbb{E}|W^{\epsilon,B}(t) - W^{\epsilon,B}(s)|^2 &= \mathbb{E}\left|\frac{1}{\sqrt{\epsilon}}\int_s^t e^{(t-\sigma)B/\epsilon}dW(\sigma)\right. \\
&\quad \left. + \frac{1}{\sqrt{\epsilon}}\int_0^s (e^{(t-\sigma)B/\epsilon} - e^{(s-\sigma)B/\epsilon})dW(\sigma)\right|^2 \\
&= \mathbb{E}\left|\frac{1}{\sqrt{\epsilon}}\int_s^t e^{(t-\sigma)B/\epsilon}dW(\sigma)\right|^2 \\
&\quad + \mathbb{E}\left|\frac{1}{\sqrt{\epsilon}}\int_0^s (e^{(t-\sigma)B/\epsilon} - e^{(s-\sigma)B/\epsilon})dW(\sigma)\right|^2 \\
&= \frac{1}{\epsilon}\int_s^t |e^{(t-\sigma)B/\epsilon}|^2_{\mathcal{L}_2(H)}d\sigma \\
&\quad + \frac{1}{\epsilon}\int_0^t |e^{(t-\sigma)B/\epsilon} - e^{(s-\sigma)B/\epsilon}|^2_{\mathcal{L}_2(H)}d\sigma.
\end{aligned}$$

On the one hand,

$$\begin{aligned}
\frac{1}{\epsilon}\int_s^t |e^{(t-\sigma)B/\epsilon}|^2_{\mathcal{L}_2(H)}d\sigma &= \frac{1}{\epsilon}\int_s^t \sum_{k=0}^{+\infty} e^{-2(t-\sigma)\mu_k/\epsilon}d\sigma \\
&= \sum_{k=0}^{+\infty}\int_0^{(t-s)/\epsilon} e^{-2\sigma\mu_k}d\sigma \\
&= \sum_{k=0}^{+\infty}\frac{1}{2\mu_k}(1 - e^{-2\mu_k(t-s)/\epsilon}) \\
&\le C_\zeta \sum_{k=0}^{+\infty}\frac{\mu_k^{2r}}{\mu_k}\left(\frac{t-s}{\epsilon}\right)^{2r},
\end{aligned}$$

and we know (by Assumption 2.1) that the above sum is finite if and only if $r < 1/4$; on the other hand,

$$\begin{aligned}
\frac{1}{\epsilon}\int_0^s |e^{(t-\sigma)B/\epsilon} - e^{(s-\sigma)B/\epsilon}|^2_{\mathcal{L}_2(H)}d\sigma &= \frac{1}{\epsilon}\int_0^s \sum_{k=0}^{+\infty} e^{-2(s-\sigma)\mu_k/\epsilon}(1 - e^{-(t-s)\mu_k/\epsilon})^2 d\sigma \\
&\le C_\zeta \sum_{k=0}^{+\infty}\left(\frac{t-s}{\epsilon}\right)^{2r}\frac{\mu_k^{2r}}{\mu_k}(1 - e^{-2s\mu_k}). \quad \square
\end{aligned}$$

Finally the following Proposition gives a control for AX^ϵ. We assume $\theta > 0$, even if the proof is valid for $\theta = 0$.

Proposition A.5. *For any $0 < r < 1$, there exists a constant C_r such that if $x \in D((-A)^\theta)$ and $y \in H$, then for any $t > 0$ and $\epsilon > 0$*

$$(\mathbb{E}[|AX^\epsilon(t)|^2])^{1/2} \le C_r(1 + t^{\theta-1})|x|_{(-A)^\theta} + C_r(1 + \epsilon^{-\frac{r}{2}})(1 + |x|_H + |y|_H).$$

Author's personal copy

Proof. We remark that in Lemma 4.4 of [5] $\mathbb{E}|AX^\epsilon(t)|$ is controlled, but the same approach gives the result for $\mathbb{E}|AX^\epsilon(t)|^2$. We have

$$\begin{aligned} X^\epsilon(t) &= e^{tA}x + \int_0^t e^{(t-s)A}F(X^\epsilon(s), Y^\epsilon(s))ds \\ &= e^{tA}x + \int_0^t e^{(t-s)A}F(X^\epsilon(t), Y^\epsilon(t))ds \\ &\quad + \int_0^t e^{(t-s)A}\left(F(X^\epsilon(s), Y^\epsilon(s)) - F(X^\epsilon(t), Y^\epsilon(t))\right)ds. \end{aligned}$$

For the first term, we have for any $t > 0$

$$|Ae^{tA}x| \le Ct^{\theta-1}|x|_{(-A)^\theta}.$$

For the second term, we have

$$|A\int_0^t e^{(t-s)A}F(X^\epsilon(t), Y^\epsilon(t))ds| = |(e^{tA} - I)F(X^\epsilon(t), Y^\epsilon(t))| \le C.$$

For the third term, we have

$$\begin{aligned} &\left|A\int_0^t e^{(t-s)A}\left(F(X^\epsilon(s), Y^\epsilon(s)) - F(X^\epsilon(t), Y^\epsilon(t))\right)ds\right| \\ &\quad \le \int_0^t \frac{Ce^{-\frac{\lambda}{2}(t-s)}}{t-s}\left(|X^\epsilon(s) - X^\epsilon(t)| + |Y^\epsilon(s) - Y^\epsilon(t)|\right)ds. \end{aligned}$$

Using Minkowski inequality, we get

$$\begin{aligned} &\mathbb{E}\left(\int_0^t \frac{Ce^{-\frac{\lambda}{2}(t-s)}}{t-s}|X^\epsilon(s) - X^\epsilon(t)|ds\right)^2 \\ &\quad \le \left(\int_0^t \frac{Ce^{-\frac{\lambda}{2}(t-s)}}{t-s}(\mathbb{E}|X^\epsilon(t) - X^\epsilon(s)|^2)^{1/2}ds\right)^2 \\ &\mathbb{E}\left(\int_0^t \frac{Ce^{-\frac{\lambda}{2}(t-s)}}{t-s}|Y^\epsilon(s) - Y^\epsilon(t)|ds\right)^2 \\ &\quad \le \left(\int_0^t \frac{Ce^{-\frac{\lambda}{2}(t-s)}}{t-s}(\mathbb{E}|Y^\epsilon(t) - Y^\epsilon(s)|^2)^{1/2}ds\right)^2. \end{aligned}$$

Using Propositions A.3 and A.4, we obtain a regularity result which gives convergent integrals. It is then easy to conclude. □

Appendix B. Properties of $\overline{X}$

Again the results of this section only require Assumptions 2.1, 2.5 and 2.7; in particular no dissipativity is assumed.

Recall that $\overline{X}(t, x)$ is defined via (1.3).

Proposition B.1. *There exists $C > 0$ such that for any $x \in H$ and any $t \ge 0$*

$$|\overline{X}(t, x)| \le C(1 + e^{-\lambda t}|x|).$$

Author's personal copy

Proof. We use the mild representation formula: for any $t \geq 0$ and $x \in H$,

$$\begin{aligned}|\overline{X}(t,x)| &= \left|e^{tA}x + \int_0^t e^{(t-s)A}\overline{F}(\overline{X}(s,x))ds\right| \\ &\leq e^{-\lambda t}x + \int_0^t e^{-\lambda(t-s)}|\overline{F}(\overline{X}(s,x))|ds \\ &\leq C(1+e^{-\lambda t}|x|),\end{aligned}$$

since $\overline{F}$ is bounded. □

Proposition B.2. *For any $0 < r < 1$ and $0 < \theta \leq 1$, there exists $C_r > 0$ such that for any $x \in H$, for any $0 < s \leq t$, we have*

$$|\overline{X}(t,x) - \overline{X}(s,x)| \leq C_r|t-s|^{1-r}\left(1+\frac{1}{s^{1-r}}\right)(1+|x|_H).$$

Proof. • If $0 \leq s < t \leq T$, we can write

$$\begin{aligned}\overline{X}(t,x) - \overline{X}(s,x) = (e^{tA}-e^{sA})x &+ \int_s^t e^{(t-\sigma)A}\overline{F}(\overline{X}(\sigma,x))d\sigma \\ &+ \int_0^s (e^{(t-\sigma)A}-e^{(s-\sigma)A})\overline{F}(\overline{X}(\sigma,x))d\sigma.\end{aligned}$$

- For the first term, it is easy to see that $|(e^{tA}-e^{sA})x|_H \leq C_r|t-s|^{1-r}\left(1+\frac{1}{s^{1-r}}\right)|x|_H$.
- For the second term, since $\overline{F}$ is bounded we have $|\int_s^t e^{(t-\sigma)A}\overline{F}(\overline{X}(\sigma,x))d\sigma|_H \leq C(t-s)$.
- For the third term, we have

$$\begin{aligned}&\left|\int_0^s (e^{(t-\sigma)A}-e^{(s-\sigma)A})\overline{F}(\overline{X}(\sigma,x))d\sigma\right|_H \\ &\quad\leq C_r\int_0^s \frac{e^{-\frac{\lambda}{2}(s-\sigma)}}{(s-\sigma)^{1-r}}(t-s)^{1-r}|\overline{F}(\overline{X}(\sigma,x))|_H d\sigma \\ &\quad\leq C_r(t-s)^{1-r}, \quad \square\end{aligned}$$

Proposition B.3. *For any $0 < \theta \leq 1$, there exists $C(\theta) > 0$ such that if $x \in D(-A)^\theta$, then for any $t > 0$*

$$|A\overline{X}(t,x)|_H \leq C_\theta(1+t^{\theta-1})(1+|x|_{(-A)^\theta}).$$

Proof. We first write that for any $t \geq 0$

$$\begin{aligned}\overline{X}(t,x) &= e^{tA}x + \int_0^t e^{(t-s)A}\overline{F}(\overline{X}(s,x))ds \\ &= e^{tA}x + \int_0^t e^{(t-s)A}\overline{F}(\overline{X}(t,x))ds + \int_0^t e^{(t-s)A}(\overline{F}(\overline{X}(s,x)) - \overline{F}(\overline{X}(t,x)))ds.\end{aligned}$$

We have $|Ae^{tA}x|_H \leq C|x|_{(-A)^\theta}t^{\theta-1}$.

Author's personal copy

For the second term, we have

$$\left|A\int_0^t e^{(t-s)A}\overline{F}(\overline{X}(t,x))ds\right|_H = |(e^{tA}-I)\overline{F}(\overline{X}(t,x))|_H \le |\overline{F}(\overline{X}(t,x))|_H \le C.$$

The third term can be controlled by

$$\left|A\int_0^t e^{(t-s)A}(\overline{F}(\overline{X}(s,x))-\overline{F}(\overline{X}(t,x)))ds\right| \le C\int_0^t \frac{e^{-c(t-s)}}{t-s}|\overline{X}(t,x)-\overline{X}(s,x)|_H ds.$$

In order to get a convergent integral, we use the regularity result of $\overline{X}$ proved in Proposition B.2; therefore we obtain the result. □

The next three Propositions deal with $\eta^h(t,x)$ the derivative of $\overline{X}(t,x)$ with respect to x in direction $h\in H$, at time t: it is the solution of

$$\begin{aligned}&\frac{d\eta^h(t,x)}{dt} = A\eta^h(t,x)+D\overline{F}(\overline{X}(t,x))\cdot\eta^h(t,x)\\ &\eta^h(0,x)=h.\end{aligned} \tag{B.1}$$

Notice that we have to consider a finite horizon $T>0$.

Proposition B.4. *For any $T>0$, there exists $C_T>0$ such that for any $x\in H, h\in H$ and $0<t\le T$*

$$|\eta^h(t,x)|\le C_T|h|$$
$$|\eta^h(t,x)|_{(-A)^\eta}\le C_T\left(1+\frac{1}{t^\eta}\right)|h|.$$

Proof. We use that A is a negative operator to prove say that for any $t\ge 0$

$$\frac{1}{2}\frac{d|\eta^h(t,x)|^2}{dt} = \langle A\eta^h(t,x),\eta^h(t,x)\rangle + \langle D\overline{F}(\overline{X}(t,x))\cdot\eta^h(t,x),\eta^h(t,x)\rangle \le [\overline{F}]_{\mathrm{Lip}}|\eta^h(t,x)|^2 \le C|\eta^h(t,x)|^2.$$

Gronwall Lemma then yields the first estimate. The second one is proved by using the mild formulation for $\eta^h(t,x)$:

$$\eta^h(t,x)=e^{tA}h+\int_0^t e^{(t-s)A}D\overline{F}(\overline{X}(s,x))\cdot\eta^h(s,x)ds;$$

thanks to the previous estimate, the integral is bounded by a constant, while $|e^{tA}h|_{(-A)^\eta}\le\frac{C}{t^\eta}|h|_H$ (see Proposition 2.4). □

Proposition B.5. *For any $T>0, 0<r<1$, there exists $C_{T,r}>0$ such that for any $x\in H, h\in H$ and $0<s\le t\le T$*

$$|\eta^h(t,x)-\eta^h(s,x)|\le C_{T,r}(t-s)^{1-r}\left(1+\frac{1}{s^{1-r}}\right)|h|.$$

Author's personal copy

Proof. • For $0 \leq s < t \leq T$ we can write that

$$\eta^h(t,x) - \eta^h(s,x) = (e^{tA} - e^{sA})h + \int_s^t e^{(t-\sigma)A} D\overline{F}(\overline{X}(\sigma,x)) \cdot \eta^h(\sigma,x)d\sigma + \int_0^s (e^{(t-\sigma)A} - e^{(s-\sigma)A}) D\overline{F}(\overline{X}(\sigma,x)) \cdot \eta^h(\sigma,x)d\sigma.$$

- For the first term, we can see that $|(e^{tA} - e^{sA})h| \leq C_r \frac{(t-s)^{1-r}}{s^{1-r}}|h|$.
- For the second term, we simply have

$$\left| \int_s^t e^{(t-\sigma)A} D\overline{F}(\overline{X}(\sigma,x)) \cdot \eta^h(\sigma,x)d\sigma \right| \leq C(t-s)|h|_H.$$

- For the third term,

$$\begin{aligned} &\left| \int_0^s (e^{(t-\sigma)A} - e^{(s-\sigma)A}) D\overline{F}(\overline{X}(\sigma,x)) \cdot \eta^h(\sigma,x)d\sigma \right| \\ &\quad \leq C_\delta \int_0^s \frac{(t-s)^{1-r}}{(s-\sigma)^{1-r}} |D\overline{F}(\overline{X}(\sigma,x)) \cdot \eta^h(\sigma,x)|_H d\sigma \\ &\quad \leq C_{r,T}(t-s)^{1-r}|h|_H. \quad \square \end{aligned}$$

Proposition B.6. *For any $T > 0$, there exists C_T such that for any $x \in H$, $h \in H$ and $0 < t \leq T$*

$$|A\eta^h(t,x)| \leq C_T(t^{-1}+1)(1+|x|)|h|.$$

Proof. For any $t \geq 0$,

$$\begin{aligned} \eta^h(t,x) &= e^{tA}h + \int_0^t e^{(t-s)A} D\overline{F}(\overline{X}(s,x)) \cdot \eta^h(s,x)ds \\ &= e^{tA}h + \int_0^t e^{(t-s)A} D\overline{F}(\overline{X}(t,x)) \cdot \eta^h(t,x)ds \\ &\quad + \int_0^t e^{(t-s)A} (D\overline{F}(\overline{X}(s,x)).\eta^h(s,x) - D\overline{F}(\overline{X}(t,x)) \cdot \eta^h(t,x))ds. \end{aligned}$$

For the first term, we have $|Ae^{tA}h|_H \leq Ct^{-1}|h|$.
For the second term,

$$\begin{aligned} \left| A \int_0^t e^{(t-s)A} D\overline{F}(\overline{X}(t,x)) \cdot \eta^h(t,x)ds \right|_H &= |(e^{tA} - I) D\overline{F}(\overline{X}(t,x)) \cdot \eta^h(t,x)|_H \\ &\leq 2|D\overline{F}(\overline{X}(t,x)) \cdot \eta^h(t,x)|_H \\ &\leq C|\eta^h(t,x)|_H \\ &\leq C|h|_H. \end{aligned}$$

For the third term, we have

$$\begin{aligned} &\left| A \int_0^t e^{(t-s)A} (D\overline{F}(\overline{X}(s,x)) \cdot \eta^h(s,x) - D\overline{F}(\overline{X}(t,x)) \cdot \eta^h(t,x))ds \right|_H \\ &\quad \leq \int_0^t \frac{C}{t-s} |D\overline{F}(\overline{X}(t,x)) \cdot \eta^h(t,x) - D\overline{F}(\overline{X}(s,x)) \cdot \eta^h(s,x)|_H ds. \end{aligned}$$

Author's personal copy

To get a convergent integral, we need to show some regularity property.
For any $0 \leq s < t \leq T$,

$$\begin{aligned} &D\overline{F}(\overline{X}(t,x)) \cdot \eta^h(t,x) - D\overline{F}(\overline{X}(s,x)).\eta^h(s,x) \\ &\quad = [D\overline{F}(\overline{X}(t,x)) - D\overline{F}(\overline{X}(s,x))] \cdot \eta^h(t,x) \\ &\qquad + D\overline{F}(\overline{X}(s,x)) \cdot (\eta^h(t,x) - \eta^h(s,x)). \end{aligned}$$

On the one hand, using Proposition 2.16 on the regularity of $\overline{F}$, we have

$$\begin{aligned} &|[D\overline{F}(\overline{X}(t,x)) - D\overline{F}(\overline{X}(s,x))] \cdot \eta^h(t,x)| \\ &\quad \leq C|\overline{X}(t,x) - \overline{X}(s,x)|\,|\eta^h(t,x)|_{(-A)^\eta} \\ &\quad \leq C(1+|x|)(t-s)^r \left(1 + \frac{1}{s^r}\right)\left(1 + \frac{1}{s^\eta}\right)|h|_H, \end{aligned}$$

thanks to Propositions B.2 and B.4. Here r must satisfy $r > 0$ and $\eta + r < 1$.
On the other hand, using Proposition B.5,

$$\begin{aligned} |D\overline{F}(\overline{X}(s,x)) \cdot (\eta^h(t,x) - \eta^h(s,x))|_H &\leq C|\eta^h(t,x) - \eta^h(s,x)|_H \\ &\leq C|h|\,|t-s|^{1-r}\left(1 + \frac{1}{s^{1-r}}\right). \end{aligned}$$

By integration, we then obtain the result. □

Finally we focus on $\xi^{h,k}(t,x)$ the second derivative of $\overline{X}(t,x)$ with respect to x in directions $h, k \in H$, at time t: it is solution of

$$\begin{aligned} \frac{d\xi^{h,k}(t,x)}{dt} &= A\xi^{h,k}(t,x) + D_x\overline{F}(\overline{X}(t,x)) \cdot (\xi^{h,k}(t,x)) \\ &\quad + D^2_{xx}\overline{F}(\overline{X}(t,x)) \cdot (\eta^h(t,x), \eta^k(t,x)). \end{aligned} \tag{B.2}$$

Proposition B.7. *For any $T > 0$, there exists $C_T > 0$ such that for any $x \in H, h, k \in H$ and $0 \leq t \leq T$*

$$|\xi^{h,k}(t,x)| \leq C_T|h|\,|k|.$$

Proof. We have—since A is negative, and using the estimates of Proposition 2.16:

$$\begin{aligned} \frac{1}{2}\frac{d|\xi^{h,k}(t,x)|^2}{dt} &\leq |D_x\overline{F}(\overline{X}(t,x))|\,|\xi^{h,k}(t,x)|^2 + C|\eta^h(t,x)|\,|\eta^k(t,x)|_{(-A)^\eta}|\xi^{h,k}(t,x)| \\ &\leq C|\xi^{h,k}(t,x)|^2 + C|\eta^h(t,x)|^2|\eta^k(t,x)|^2_{(-A)^\eta}, \end{aligned}$$

Using Proposition B.4, the Assumption $\eta < \frac{1}{2}$, and the Gronwall Lemma, we get the result. □

Appendix C. Properties of the auxiliary function $\tilde{F}$

The results of this section are used only for the proof of the strong convergence Theorem 1.1. Here we need the strict dissipativity Assumption 2.8.

In the proof of Lemma 3.2, we need to use an auxiliary function $\tilde{F}$—see Definition 3.3.

Thanks to Proposition 2.14, we get:

Author's personal copy

Proposition C.1. *There exists $c > 0, C > 0$ such that for any $(x, y) \in H^2$ and $t \geq 0$,*

$$|\tilde{F}(x, y, t)| \leq Ce^{-ct}(1 + |x|_H + |y|_H).$$

We also need the following estimate on the Lipschitz constant of $\tilde{F}$ with respect to x, which depends on the regularity assumptions made on F and G—see Assumptions 2.5 and 2.7:

Proposition C.2. *There exists $c > 0, C > 0$, such that for any $x_1, x_2, y \in H$ and $t \geq 0$*

$$|\tilde{F}(x_1, y, t) - \tilde{F}(x_2, y, t)| \leq C(1 + |y|)e^{-ct}\left(1 + \frac{1}{t^\eta}\right)|x_1 - x_2|.$$

Proof. For any $t_0 > 0$, we define the following function:

$$\tilde{F}_{t_0}(x, y, t) = \hat{F}(x, y, t) - \hat{F}(x, y, t + t_0),$$

where $\hat{F}(x, y, t) := \mathbb{E}F(x, Y_x(t, y))$.
We claim that it satisfies the following properties:

- $\tilde{F}_{t_0}(x, y, t) \to \tilde{F}(x, y, t)$ when $t_0 \to +\infty$.
- For any t_0, for any x, y, t and any h, $\tilde{F}_{t_0}$ is differentiable with respect to x at (x, y, t) and in direction $h \in H$.
- We have $|D_x\tilde{F}_{t_0}(x, y, t) \cdot h| \leq Ce^{-ct}\left(1 + \frac{1}{t^\eta}\right)(1 + |y|)|h|$, C being independent of t_0.

The first two ones are obvious, thanks to regularity properties of F; moreover as soon as we have the third property, the proof of the Proposition can be finished as follows: if we fix x_1, x_2, y, t, h, then for any $t_0 > 0$

$$|\tilde{F}_{t_0}(x_1, y, t) - \tilde{F}_{t_0}(x_2, y, t)| \leq Ce^{-ct}\left(1 + \frac{1}{t^\eta}\right)(1 + |y|)|x_1 - x_2|.$$

Letting $t_0 \to +\infty$, we get

$$|\tilde{F}(x_1, y, t) - \tilde{F}(x_2, y, t)| \leq Ce^{-ct}\left(1 + \frac{1}{t^\eta}\right)(1 + |y|)|x_1 - x_2|.$$

It remains to estimate $|D_x\tilde{F}_{t_0}(x, y, t) \cdot h|$ for any $h \in H$.
First we notice that thanks to the Markov property we have

$$\begin{aligned}\tilde{F}_{t_0}(x, y, t) &= \hat{F}(x, y, t) - \hat{F}(x, y, t + t_0)\\ &= \hat{F}(x, y, t) - \mathbb{E}F(x, Y_x(t + t_0, y))\\ &= \hat{F}(x, y, t) - \mathbb{E}\hat{F}(x, Y_x(t_0, y), t).\end{aligned}$$

Therefore we have for any h

$$\begin{aligned}D_x\tilde{F}_{t_0}(x, y, t) \cdot h &= D_x\hat{F}(x, y, t) \cdot h - \mathbb{E}D_x\left(\hat{F}(x, Y_x(t_0, y), t)\right) \cdot h\\ &= D_x\hat{F}(x, y, t) \cdot h - \mathbb{E}D_x\hat{F}(x, Y_x(t_0, y), t) \cdot h\\ &\quad - \mathbb{E}D_y\hat{F}(x, Y_x(t_0, y), t) \cdot (D_xY_x(t_0, y) \cdot h).\end{aligned}$$

Then we see that we have to analyse

$$D_x\hat{F}(x, y, t) \cdot h - D_x\hat{F}(x, z, t) \cdot h$$

Author's personal copy

and

$$D_y\hat{F}(x,y,t).$$

- For any $y, z \in H$, we have
$$\begin{aligned}|\hat{F}(x,y,t)-\hat{F}(x,z,t)| &= |\mathbb{E}F(x,Y_x(t,y))-\mathbb{E}F(x,Y_x(t,z))|\\ &\le C\mathbb{E}|Y_t^x(y)-Y_t^x(z)|\\ &\le Ce^{-ct}|y-z|,\end{aligned}$$
and we deduce that $|D_y\hat{F}(x,y,t).k| \le Ce^{-ct}|k|$.
- Moreover we know that $U_t^{x,h}(y) = D_xY_x(t,y)\cdot h$ is solution of
$$\begin{aligned}dU_t^{x,h}(y) &= \left(BU_t^{x,h}(y)+D_xG(x,Y_x(t,y))\cdot h+D_yG(x,Y_x(t,y)).U_t^{x,h}(y)\right)dt\\ U_0^{x,h}(y) &= 0.\end{aligned}$$
We deduce the following property: $|U_t^{x,h}(y)| \le C|h|$ a.s.
As a consequence $|\mathbb{E}D_y\hat{F}(x,Y_x(t_0,y),t)\cdot(D_xY_x(t_0,y)\cdot h)| \le Ce^{-ct}|h|$.
- Now we take x, y, z, t, h, and we compute
$$\begin{aligned}&D_x\hat{F}(x,y,t)\cdot h - D_x\hat{F}(x,z,t)\cdot h\\ &\quad= \mathbb{E}\left(D_xF(x,Y_x(t,y))\cdot h - D_xF(x,Y_x(t,z))\cdot h\right)\\ &\qquad+\mathbb{E}\left(D_yF(x,Y_x(t,y)).U_t^{x,h}(y)-D_yF(x,Y_x(t,z)).U_t^{x,h}(z)\right)\\ &\quad= \mathbb{E}\left(D_xF(x,Y_x(t,y))\cdot h - D_xF(x,Y_x(t,z))\cdot h\right)\\ &\qquad+\mathbb{E}\left([D_yF(x,Y_x(t,y))-D_yF(x,Y_x(t,z))].U_t^{x,h}(y)\right)\\ &\qquad+\mathbb{E}\left(D_yF(x,Y_x(t,z))\cdot(U_t^{x,h}(y)-U_t^{x,h}(z))\right)\end{aligned}$$

First, we have
$$\begin{aligned}&|\mathbb{E}\left(D_xF(x,Y_x(t,y))\cdot h - D_xF(x,Y_x(t,z))\cdot h\right)|\\ &\quad\le \mathbb{E}|D_xF(x,Y_x(t,y))\cdot h - D_xF(x,Y_x(t,z))\cdot h|\\ &\quad\le C|h|_H\mathbb{E}|Y_x(t,y)-Y_x(t,z)|_{(-B)^\eta}\\ &\quad\le C|h|_He^{-ct}\left(1+\frac{1}{t^\eta}\right)|y-z|_H,\end{aligned}$$
using the regularity assumptions on F (see (2.5)), and using the following estimate:
$$\mathbb{E}|Y_x(t,y)-Y_x(t,z)|_{(-B)^\eta} \le Ce^{-ct}\left(1+\frac{1}{t^\eta}\right)|y-z|_H,$$
for some $c > 0$.

Second,
$$\begin{aligned}&|\mathbb{E}([D_yF(x,Y_x(t,y))-D_yF(x,Y_x(t,z))]\cdot U_t^{x,h}(y))|\\ &\quad\le \mathbb{E}|[D_yF(x,Y_x(t,y))-D_yF(x,Y_x(t,z))]\cdot U_t^{x,h}(y)|\\ &\quad\le C\mathbb{E}|U_t^{x,h}(y)|_H|Y_x(t,y)-Y_x(t,z)|_{(-B)^\eta}\\ &\quad\le Ce^{-ct}\left(1+\frac{1}{t^\eta}\right)|h|_H|y-z|_H.\end{aligned}$$

Author's personal copy

Third,

$$|\mathbb{E}\left(D_y F(x, Y_x(t,z)) \cdot (U_t^{x,h}(y) - U_t^{x,h}(z))\right)|$$
$$\leq E|D_y F(x, Y_x(t,z)) \cdot (U_t^{x,h}(y) - U_t^{x,h}(z))|$$
$$\leq \mathbb{E}|U_t^{x,h}(y) - U_t^{x,h}(z)|_H;$$

It remains to look at $|U_t^{x,h}(y) - U_t^{x,h}(z)|_H$; we indeed have

$$|U_t^{x,h}(y) - U_t^{x,h}(z)|^2 \leq C|h|_H^2 |y-z|_H^2 e^{-c_0 t},$$

where $c_0 > 0$.

We use these inequalities with $z := Y_x(t_0, y)$; recalling that for any t_0

$$\mathbb{E}|Y_x(t_0, y)| \leq C(1 + |y|),$$

we get

$$|D_x \tilde{F}_{t_0}(x, y, t) \cdot h| \leq C(1+|y|)e^{-ct}\left(1 + \frac{1}{t^\eta}\right)|h|. \quad \square$$

References

[1] P. Brémaud, Markov Chains. Gibbs Fields, Monte Carlo Simulation, and Queues, in: Texts in Applied Mathematics, Springer, New York, NY, 1999, p. xviii. p. 444.

[2] H. Brézis, Functional Analysis. Theory and Applications, in: Analyse Fonctionnelle. Théorie et Applications (Collection Mathématiques Appliquées pour la Maîtrise), Masson, Paris, 1994, p. 248.

[3] S. Cerrai, Second Order PDE's in Finite and Infinite Dimension, in: Lecture Notes in Mathematics, vol. 1762, Springer, Berlin, 2001, p. ix. p. 330.

[4] S. Cerrai, A Khasminskii type averaging principle for stochastic reaction–diffusion equations, Ann. Appl. Probab. 19 (3) (2009) 899–948.

[5] S. Cerrai, M. Freidlin, Averaging principle for a class of stochastic reaction–diffusion equations, Probab. Theory Related Fields 144 (1–2) (2009) 137–177.

[6] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, in: Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992, p. xviii. p. 454.

[7] A. Debussche, Y. Hu, G. Tessitore, Ergodic BSDEs under weak dissipative assumptions, Stoch. Proc. Appl. 121 (3) (2011) 407–426.

[8] W. Doeblin, Expose de la théorie des chaînes simples constantes de Markoff à un nombre fini d'etats, Rev. Math. Union Interbalkan 2 (1938) 77–105.

[9] J.-P. Fouque, J. Garnier, G. Papanicolaou, K. Solna, Wave Propagation and Time Reversal in Randomly Layered Media, in: Stochastic Modelling and Applied Probability, vol. 56, Springer, New York, NY, 2007, p. xx. p. 612.

[10] M.I. Freidlin, A.D. Wentzell, Random Perturbations of Dynamical Systems, second ed., in: Grundlehren der Mathematischen Wissenschaften, vol. 260, Springer, New York, NY, 1998, p. xi. Transl. from the Russian by Joseph Szuecs, p. 430.

[11] R.Z. Khasminskii, On an averaging principle for Itô stochastic differential equations, Kibernetica (4) (1968) 260–279.

[12] R.Z. Khasminskii, G. Yin, Limit behavior of two-time-scale diffusions revisited, J. Differential Equations 212 (1) (2005) 85–113.

[13] S. Kuksin, A. Shirikyan, A coupling approach to randomly forced nonlinear PDE's. I, Comm. Math. Phys. 221 (2) (2001) 351–366.

[14] E.T. Lindvall, Lectures on the Coupling Method, in: Wiley Series in Probability and Mathematical Statistics (Applied Probability and Statistics), Wiley, New York, NY, 1992, p. 272.

[15] D. Liu, Strong convergence of principle of averaging for multiscale stochastic dynamical systems, Commun. Math. Sci. 8 (4) (2010) 999–1020.

[16] J.C. Mattingly, Exponential convergence for the stochastically forced Navier–Stokes equations and other partially dissipative dynamics, Comm. Math. Phys. 230 (3) (2002) 421–462.

Author's personal copy

[17] S. Meyn, R.L. Tweedie, Markov chains and stochastic stability, in: Peter W. Glynn (Ed.), Cambridge Mathematical Library, second ed., Cambridge University Press, Cambridge, 2009, p. xviii. Prologue by, p. 594.
[18] G.A. Pavliotis, A.M. Stuart, Multiscale Methods. Averaging and Homogenization, in: Texts in Applied Mathematics, vol. 53, Springer, New York, NY, 2008, p. xviii. p. 307.
[19] E. Weinan, D. Liu, E. Vanden-Eijnden, Analysis of multiscale methods for stochastic differential equations, Comm. Pure Appl. Math. 58 (11) (2005) 1544–1585.