Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications xxx (xxxx) xxx
stochastic processes and their applications
www.elsevier.com/locate/spa

Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component

Charles-Edouard Bréhier

Univ Lyon, CNRS, Université Claude Bernard Lyon 1, UMR5208, Institut Camille Jordan, F-69622 Villeurbanne, France

Received 27 October 2018; received in revised form 27 June 2019; accepted 29 September 2019
Available online xxxx

Abstract

This article is devoted to the analysis of semilinear, parabolic, Stochastic Partial Differential Equations, with slow and fast time scales. Asymptotically, an averaging principle holds: the slow component converges to the solution of another semilinear, parabolic, SPDE, where the nonlinearity is averaged with respect to the invariant distribution of the fast process.

We exhibit orders of convergence, in both strong and weak senses, in two relevant situations, depending on the spatial regularity of the fast process and on the covariance of the Wiener noise in the slow equation. In a very regular case, strong and weak orders are equal to $\frac{1}{2}$ and 1. In a less regular case, the weak order is also twice the strong order.

This study extends previous results concerning weak rates of convergence, where either no stochastic forcing term was included in the slow equation, or the covariance of the noise was extremely regular.

An efficient numerical scheme, based on Heterogeneous Multiscale Methods, is briefly discussed.

Keywords: Stochastic partial differential equations; Averaging principle; Poisson equation in infinite dimension; Heterogeneous multiscale method; Strong and weak error estimates

1. Introduction

Systems with multiple time scales, and possibly stochastic forcing terms, appear in all fields of modern science, at fundamental and applied levels, for instance in physics, chemistry,

E-mail address: brehier@math.univ-lyon1.fr.

https://doi.org/10.1016/j.spa.2019.09.015
0304-4149/

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

biology, engineering, etc... Understanding how properties at micro-scales transfer to macro-scales, and the design of efficient numerical schemes, are still challenging issues. In the mathematical literature, powerful limiting procedures have been developed, e.g. averaging and homogenization techniques are available. We refer for instance to [14,15,30,36] for monographs devoted to the study of multiscale stochastic systems.

The averaging principle can be interpreted as a law of large numbers, in cases where a slow component is driven by an equation with coefficients depending on a fast component, which is an ergodic stochastic process: when the separation of time scales goes to infinity, the slow component converges to the solution of an averaged equation, where coefficient have been averaged out with respect to some invariant probability distribution for the fast component. For results concerning the averaging principle for Stochastic Differential Equations (SDEs), we refer to the pioneering work [26], and to the following extensions since then, e.g. [27,28,32,39] (this list is not exhaustive).

In this article, the aim is to study the averaging principle for a class of parabolic, semilinear, Stochastic Partial Differential Equations (SPDEs)

$$\begin{aligned}
\frac{\partial x^\epsilon(t,\xi)}{\partial t} &= \Delta x^\epsilon(t,\xi) + f\big(x^\epsilon(t,\xi), y^\epsilon(t,\xi)\big) + \dot{W}(t,\xi), \quad t\geq 0,\ \xi\in\mathcal{D},\\
x^\epsilon(t,\cdot)_{|\partial\mathcal{D}} &= 0,\ t\geq 0,\\
x^\epsilon(0,\cdot) &= x_0,\ y^\epsilon(0,\cdot)=y_0,
\end{aligned}$$

where the domain is $\mathcal{D}=(0,1)^d$, for some $d\in\{1,2,3\}$, and $\dot{W}$ is a Gaussian noise which is white in time, and white or correlated in space. For a mathematically precise formulation, see (1), in the framework of [12], and Section 2. The fast component is given by $y^\epsilon(t,\cdot) = y(\epsilon^{-1}t,\cdot)$, where y is assumed to be an ergodic process, independent of the Wiener process W which appears in the slow equation, *i.e.* the equation for the slow component x^ϵ.

The averaging principle for such SPDE systems has been proved in a very general framework in [9,11] for globally Lipschitz coefficients, and later in [10] in the case of non-globally Lipschitz continuous coefficients. In these results, no order of convergence, in terms of $\epsilon\to 0$, is provided. The first study of orders of convergence for the averaging principle in the SPDE case, was performed by the author in [3]. The main motivation for studying the rates of convergence is the construction of efficient numerical schemes, based on the Heterogeneous Multiscale Methods, see [4] and references therein.

In recent years, many works have been devoted to the study of the averaging principle for different classes of SPDEs. For instance, see [13,16,17,19,21,31], in the parabolic SPDE case. See [18,20], for some parabolic–hyperbolic systems. See [22,23] in the Schrödinger equations case. Finally, see [1,2], where stochastic fluid mechanics equations are considered, with motivations coming from physics.

As is usual when dealing with stochastic equations, orders of convergence are understood in two senses. On the one hand, strong convergence deals with the mean-square error. On the other hand, weak convergence is related to convergence in distribution, considering sufficiently smooth test functions. If the averaging principle for SDEs (with globally Lipschitz continuous coefficients) is considered, the strong order of convergence is $\frac{1}{2}$, whereas the weak order is 1, and these results are optimal in general. The technique of [26] is perfectly suited to prove strong convergence results, whereas to study weak convergence, approaches based on asymptotic expansions of solutions of Kolmogorov equations are very efficient, see [27,28]. The generalization to SPDEs, where there is no Wiener noise in the slow equation, has been considered in [3]. If a stochastic forcing term is present, much less is known. Indeed, for

SPDEs, *i.e.* for infinite dimensional stochastic equations, the analysis of the order of weak convergence and of the Kolmogorov equations, is notoriously challenging, we refer for instance to [6] for a recent contribution, and the discussions and references therein.

The aim of this manuscript is to study the weak order of convergence in the averaging principle, for semilinear, parabolic, SPDEs, with a stochastic forcing in the slow equation. Note that this question has been recently investigated in [19] with a very strong regularity condition on the covariance of the noise, which implies a high spatial regularity of the process x^ϵ. In fact, using this condition, the techniques of [3] may be applied, essentially with no modification, hence the weak order of convergence equal to 1 obtained in [19]. The objective of this article is to weaken the regularity conditions required to obtain strong and weak rates of convergence in the averaging principle. The main finding is that a trade-off between regularity properties of the slow and the fast components is at play. First, we prove that the strong (resp. weak) order of convergence is 1/2 (resp. 1), under an appropriate condition (the *very regular case*), which is in general much weaker than the assumption in [19]. Second, we weaken the condition (the *less regular case*), and exhibit appropriate strong and weak orders of convergence depending on the regularity properties of the slow and fast component. In that case, as expected, the weak order is twice the strong order, however, whether these results are optimal is not known, indeed the proof is based on an approximation argument and may not be optimal. For statements of the main results, see Theorems 4.5, 4.7 and 4.8.

Even if the main motivation of this work is the analysis of the weak order of convergence, a detailed analysis of the strong order of convergence is also provided, for two reasons. First, it allows us to check that the weak order is twice the strong order, as expected. Second, the technique of proof is different from the one used in previous publications on the strong convergence in the averaging principle for SPDEs, such as [3] instead of employing the technique introduced by Khasminskii in [26], the Poisson equation technique described for instance in [36] is generalized to a situation where mild solutions of SPDEs are considered.

Several relevant questions are left for future works. For instance, in this manuscript, it is assumed that the fast component y^ϵ is not coupled with the slow component x^ϵ, and it would be interesting to study the coupled case.

This article is organized as follows. Section 2 is devoted to introducing the necessary functional analysis framework and to stating precise assumptions. Regularity parameters $\alpha_{\max}$ and $\gamma_{\max}$, which are used to define the very regular and less regular cases, are introduced in Assumptions 2.6 and 2.7 respectively. Section 3 presents the averaged equation.

The main results of this article are stated and discussed in Section 4. First, in the very regular case, see Assumption 4.1 and Section 4.1, Theorem 4.5 and 4.7 state that the strong (resp. weak) order is equal to $\frac{1}{2}$ (resp. 1). Second, in the less regular case, see Assumption 4.2 and Section 4.2, Theorem 4.8 show that the strong order is (at least) $\beta_{\max} = \frac{\alpha_{\max}}{1+\alpha_{\max}-\gamma_{\max}} \leq \frac{1}{2}$ and the weak order is (at least) $2\beta_{\max}$.

Auxiliary but fundamental and nontrivial regularity results concerning a family of Poisson equations are studied in Section 5.

Proofs of the main results are provided in Sections 6, 7 and 8.

Finally, in Section 9, a class of temporal discretization numerical schemes, based on Heterogeneous Multiscale Methods, see [4], is presented. The goal of the schemes is to approximate the slow component x^ϵ, and to avoid conditions on the time step size h of the type $h = \mathrm{o}(\epsilon)$, which arise when discretizing naively the slow–fast system, and are prohibitive when $\epsilon \to 0$. The guideline of the method is the averaging principle (and the error estimates proved in this article are crucial for a full error analysis): one may approximate the solution

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

of the averaged equation, and estimate the (unknown) averaged equation using a numerical scheme for the fast process, with different and independent time step sizes at the slow and fast scales. Section 9 presents the construction of the method and error estimates. In particular, in certain regimes, one recovers an averaging principle at the discrete-time level. Section 9 may be skipped by readers which are interested only in the averaging principle for the SPDEs.

2. Setting

The objective of this section is to state precise assumptions, and to derive moment estimates (uniform in ϵ), for the following Stochastic Evolution Equation

$$dX^{\epsilon}(t) = AX^{\epsilon}(t)dt + F\big(X^{\epsilon}(t), Y^{\epsilon}(t)\big)dt + dW^{Q}(t). \tag{1}$$

This is the abstract formulation of a parabolic, semilinear, SPDE, in the framework of [12]. The stochastic forcing is given by a Q-Wiener process W^Q. In addition, Y^ϵ is another stochastic process with values in L^2. In this work, it is assumed that Y^ϵ and W^Q are independent.

We are interested in the regime of a small parameter ϵ, and of a timescale separation: $Y^\epsilon(t) = Y(t\epsilon^{-1})$. As a consequence, X^ϵ is referred to as the slow component, and Y^ϵ as the fast component.

For instance, the process Y may be the solution of an equation of the type

$$dY(t) = AY(t)dt + G(Y(t))dt + dw^q(t),$$

where $\big(w^q(t)\big)_{t\geq 0}$ is a q-Wiener process, independent of W^Q. Then Y^ϵ solves an equation of the type

$$dY^{\epsilon}(t) = \frac{1}{\epsilon}\big(AY^{\epsilon}(t) + G(Y^{\epsilon}(t))\big)dt + \frac{1}{\sqrt{\epsilon}}dw^q(t),$$

in which case one has the equality in distribution (but not almost surely) of the processes $\big(Y(t\epsilon^{-1})\big)_{t\geq 0}$ and $\big(Y^\epsilon(t)\big)_{t\geq 0}$. The assumption that Y is independent of W^Q means that G does not depend on the slow component, thus the fast evolution is not coupled with the slow evolution.

Considering the coupled situation, where G depends also on the slow component, would substantially modify some computations below. However, the treatment of the uncoupled case, considered in this manuscript, already requires the use of original and nontrivial arguments. The objective of this manuscript is to exhibit these arguments, in the simplest nontrivial framework. The treatment of the coupled case is left for future work.

2.1. Notation

Let $\mathcal{D} = (0,1)^d$, with dimension $d \in \{1,2,3\}$, denote a domain. For any $p \in [1,\infty]$, let $L^p = L^p(\mathcal{D})$, and denote by $|\cdot|_{L^p}$ the associated L^p-norm. When $p = 2$, $H = L^2$ is a separable, infinite dimensional, Hilbert space, with norm $|\cdot|_H = |\cdot|_{L^2}$, and inner product denoted by $\langle \cdot,\cdot\rangle$.

For any $p, q \in [2,\infty)$, let $\mathcal{L}(L^p, L^q)$ denote the space of bounded linear operators from L^p to L^q. The associated norm is denoted by $\|\cdot\|_{\mathcal{L}(L^p,L^q)}$.

For $p \in [2,\infty)$, let $\mathcal{R}(L^2, L^p) \subset \mathcal{L}(L^2, L^p)$ denote the space of γ-Radonifying operators from L^2 to L^p. Recall that a linear operator $\Psi \in \mathcal{L}(L^2, L^p)$ is a γ-radonifying operator, if the image by Ψ of the canonical Gaussian distribution on L^2 extends to a Borel probability measure on L^p. The space $\mathcal{R}(L^2,L^p)$ is equipped with the norm $\|\cdot\|_{\mathcal{R}(L^2,L^p)}$ defined by

$$\|\Psi\|^2_{\mathcal{R}(L^2,L^p)}=\tilde{\mathbb{E}}\Big|\sum_{n\in\mathbb{N}}\gamma_n\Psi f_n\Big|^2_{L^p},$$

where $(\gamma_n)_{n\in\mathbb{N}}$ is any sequence of independent standard (mean 0 and variance 1) Gaussian random variables, defined on a probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}})$, with expectation operator denoted by $\tilde{\mathbb{E}}$, and $(f_n)_{n\in\mathbb{N}}$ is any complete orthonormal system of L^2. When $p=2$, $\mathcal{R}(L^2,L^2)=\mathcal{L}_2(L^2)$ is the space of Hilbert–Schmidt operators on L^2, and $\|\Psi\|^2_{\mathcal{R}(L^2,L^2)}=\mathrm{Tr}(\Psi\Psi^\star)$, where $\mathrm{Tr}(\cdot)$ is the trace operator, and $\Psi^\star$ is the adjoint of Ψ.

Note that, for any $p\in[2,\infty)$, there exists $c_p\in(0,\infty)$ such that for any $\Psi\in\mathcal{R}(L^2,L^p)$,

$$\|\Psi\|^2_{\mathcal{R}(L^2,L^p)}\le c_p\Big|\sum_{n\in\mathbb{N}}(\Psi f_n)^2\Big|_{L^{\frac{p}{2}}}.$$

Finally, recall the left and right ideal property for γ-Radonifying operators: for all $p,q\in[2,\infty)$, for all operators $L_1\in\mathcal{L}(L^p,L^q)$, $\Psi\in\mathcal{R}(L^2,L^p)$ and $L_2\in\mathcal{L}(L^2,L^2)$, then $L_1\Psi L_2\in\mathcal{R}(L^2,L^q)$, and

$$\|L_1\Psi L_2\|_{\mathcal{R}(L^2,L^q)}\le\|L_1\|_{\mathcal{L}(L^p,L^q)}\|\Psi\|_{\mathcal{R}(L^2,L^p)}\|L_2\|_{\mathcal{L}(L^2,L^2)}.$$

Let $\big(W(t)\big)_{t\ge0}$ denote a cylindrical Wiener process defined on L^2, on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. For any $T\in(0,\infty)$ and $p\in[2,\infty)$, the L^p-valued Itô integral $\int_0^T\Phi(t)dW(t)$ is defined for predictable processes $\Phi\in L^2(\Omega\times(0,T),\mathcal{R}(L^2,L^p))$. Moreover, there exists $c_p\in(0,\infty)$, such that

$$\mathbb{E}\Big[\Big|\int_0^T\Phi(t)dW(t)\Big|^2_{L^p}\Big]\le c_p\int_0^T\mathbb{E}\|\Phi(t)\|^2_{\mathcal{R}(L^2,L^p)}dt.$$

In the case $p=2$, the inequality above is replaced by the following Itô isometry property:

$$\mathbb{E}\Big[\Big|\int_0^T\Phi(t)dW(t)\Big|^2_{L^2}\Big]=\int_0^T\mathbb{E}\|\Phi(t)\|^2_{\mathcal{R}(L^2,L^2)}dt.$$

Higher order moments of stochastic integrals are estimated using Burkholder–Davis–Gundy type inequalities.

For statements, proofs, and generalizations, of the results above, we refer for instance to [7,34,35] for Banach space valued stochastic integrals, and to [12] for the Hilbert space case.

If $\varphi:L^2\to\mathbb{R}$ is a function of class $\mathcal{C}^1$, its first order derivative $D\varphi(x)\in\mathcal{L}(L^2,\mathbb{R})$ may be identified with a element of L^2, thanks to Riesz Theorem: as a consequence, for all $x,h\in L^2$, we write $D\varphi(x).h=\langle D\varphi(x),h\rangle$.

2.2. The linear operator A

Let A denote the unbounded linear operator on $H=L^2$, with

$$\begin{cases}D(A)=H^2(\mathcal{D})\cap H^1_0(\mathcal{D}),\\ Ax=\Delta x,\ \forall\ x\in D(A),\end{cases}$$

where Δ is the Laplace differential operator in dimension d. The domain is chosen in order to consider homogeneous Dirichlet boundary conditions in evolution equations. It is a standard result (see for instance [37]) that there exist a complete orthonormal system $\big(e_n\big)_{n\in\mathbb{N}}$ of L^2, and a non-decreasing sequence $\big(\lambda_n\big)_{n\in\mathbb{N}}$ of positive real numbers such that

$$Ae_n = -\lambda_n e_n,\ \forall\, n\in\mathbb{N}\ ,\quad \lambda_n \underset{n\to\infty}{\sim} c_d n^{\frac{2}{d}}\quad ,\quad \sup_{n\in\mathbb{N}} |e_n|_{L^\infty} < \infty. \tag{2}$$

The operator A can also be considered as an unbounded linear operator on L^p, for all $p\in[2,\infty)$, in a consistent way as p varies. The linear operator A generates an analytic semigroup $\big(e^{tA}\big)_{t\geq 0}$, on L^p for $p\in[2,\infty)$. For $\alpha\in(0,1)$, the linear operators $(-A)^{-\alpha}$ and $(-A)^{\alpha}$ are constructed in a standard way, see for instance [37, Chapter 2, Section 2.6]:

$$(-A)^{-\alpha} = \frac{\sin(\pi\alpha)}{\pi}\int_0^\infty t^{-\alpha}(tI-A)^{-1}dt\ ,\ (-A)^{\alpha} = \frac{\sin(\pi\alpha)}{\pi}\int_0^\infty t^{\alpha-1}(-A)(tI-A)^{-1}dt,$$

where $(-A)^\alpha$ is defined as an unbounded linear operator on L^p. In the case $p=2$, note that

$$(-A)^{-\alpha}x = \sum_{i\in\mathbb{N}^\star}\lambda_i^{-\alpha}\langle x,e_i\rangle e_i,\quad x\in H,$$

$$(-A)^{\alpha}x = \sum_{i\in\mathbb{N}^\star}\lambda_i^{\alpha}\langle x,e_i\rangle e_i,\quad x\in D_2\big((-A)^\alpha\big) = \Big\{x\in H;\ \sum_{i=1}^\infty \lambda_i^{2\alpha}\langle x,e_i\rangle^2<\infty\Big\}.$$

Since $\big(e^{tA}\big)_{t\geq0}$ is an analytic semigroup on L^p, one has the following regularization property: for all $p\in[2,\infty)$ and $\alpha\in[0,1]$, there exists $C_{p,\alpha}\in(0,\infty)$ such that for all $x\in L^p$, one has

$$|(-A)^\alpha e^{tA}x|_{L^p}\leq C_{p,\alpha}t^{-\alpha}|x|_{L^p}, \tag{3}$$

see [37, Chapter 2, Theorem 6.13].

Introduce also the kernel function K associated with the semigroup $\big(e^{tA}\big)_{t\geq0}$:

$$e^{tA}\varphi(\xi) = \int_{\mathcal{D}} K(t,\xi,\eta)\varphi(\eta)d\eta. \tag{4}$$

The kernel function K satisfies the two following properties: there exist $c, C\in(0,\infty)$, such that for all $t>0$ and $\xi,\eta\in\mathcal{D}$, one has

$$0\leq K(t,\xi,\eta)\leq Ct^{-\frac{d}{2}}\exp\big(-ct^{-1}|\xi-\eta|^2\big)\ ,\quad \int_{\mathcal{D}}K(t,\xi,\eta)d\eta\leq 1. \tag{5}$$

To conclude this section, useful calculus inequalities are stated. They are employed in a crucial way in Section 5.

Proposition 2.1. *For any $\alpha\in[0,\frac12)$, any $\kappa\in(0,\frac12-\alpha)$, and any $p\in[2,\infty)$, there exists $C_{\alpha,\kappa,p}\in(0,\infty)$, such that for all $x_1,x_2\in L^{2p}$,*

$$|(-A)^\alpha(x_1x_2)|_{L^p}\leq C_{\alpha,\kappa,p}|(-A)^{\alpha+\kappa}x_1|_{L^{2p}}|(-A)^{\alpha+\kappa}x_2|_{L^{2p}}.$$

Moreover, let $\phi:(z_1,z_2)\in\mathbb{R}\times\mathbb{R}\mapsto\phi(z_1,z_2)\in\mathbb{R}$ be a Lipschitz continuous function of class $\mathcal{C}^1$. Then there exists $C_{\alpha,\kappa,p}(\phi)\in(0,\infty)$ such that for all $x,y_1,y_2\in L^{2p}$,

$$\begin{aligned}|(-A)^\alpha\big(\phi(x,y_2)-\phi(x,y_1)\big)|_{L^p}\leq C_{\alpha,\kappa,p}(\phi)\big(1+|(-A)^{\alpha+\kappa}x|_{L^{2p}}\\ +\sum_{j=1,2}|(-A)^{\alpha+\kappa}y_j|_{L^{2p}}\big)\big|(-A)^{\alpha+\kappa}(y_2-y_1)\big|_{L^{2p}}.\end{aligned}$$

Remark 2.2. In the statement of Proposition 2.1 (and in the whole article), the following convention is used to simplify the presentation: one mentions that $x\in L^p$, whereas $|(-A)^\alpha x|_{L^p}$, for some p and $\alpha > 0$, appears on the right-hand side. It should be understood that the inequality is valid only if $|(-A)^\alpha x|_{L^p} < \infty$, whereas if $|(-A)^\alpha x|_{L^p} = \infty$ then the inequality provides no information.

For the first inequality, we refer to [6, Section 3.2] and [38]. The second inequality is a straightforward consequence of the first inequality and of a first order Taylor formula:

$$\phi(x, y_2) - \phi(x, y_2) = \int_0^1 \partial_{z_2}\phi(x, \lambda y_2 + (1-\lambda)y_1)(y_2 - y_1)d\lambda.$$

Finally, the following inequalities are used below: for every $\kappa > 0$, there exists $C_\kappa \in (0, \infty)$ such that for all $x \in D((-A)^{\frac{d}{4}+\kappa})$,

$$|x|_{L^\infty} \le C_\kappa |(-A)^{\frac{d}{4}+\kappa} x|_{L^2} , \quad |(-A)^{-\frac{d}{4}-\kappa} x|_{L^2} \le C_\kappa |x|_{L^1}. \tag{6}$$

The first inequality is interpreted as a Sobolev inequality, and is a straightforward consequence of the properties (2) of eigenvalues λ_n and eigenfunctions e_n of A:

$$|x|_{L^\infty} \le \sup_{n\in\mathbb{N}} |e_n|_{L^\infty} \sum_{n\in\mathbb{N}} \frac{|\langle x, e_n\rangle|\lambda_n^{\frac{d}{4}+\kappa}}{\lambda_n^{\frac{d}{4}+\kappa}} \le \sup_{n\in\mathbb{N}} |e_n|_{L^\infty} \Big(\sum_{n\in\mathbb{N}} \frac{1}{\lambda_n^{\frac{d}{2}+2\kappa}}\Big)^{\frac{1}{2}} |(-A)^{\frac{d}{4}+\kappa} x|_{L^2}.$$

The second inequality is obtained by a duality argument. As a consequence, one obtains the following inequality: for all $x \in L^1$,

$$|e^{tA} x|_{L^2} \le C_\kappa t^{-\frac{d}{4}-\kappa} |x|_{L^1}. \tag{7}$$

2.3. Assumptions on F and Q

The coefficient F in (1) is defined as the Nemytskii operator (see Definition 2.5) associated with a smooth function f (see Assumption 2.3).

Assumption 2.3. Assume that $f : (z_1, z_2) \in \mathbb{R} \times \mathbb{R} \mapsto f(z_1, z_2) \in \mathbb{R}$ is a function of class $\mathcal{C}^4$, with bounded derivatives of order $1, \ldots, 4$.

Remark 2.4. In the calculations below, quantitative estimates will only depend on the bounds on the derivatives of f of order 1, 2, 3. Existence of the fourth order derivative is only employed to justify some calculations in Section 5.

Definition 2.5. For all $p, q \in [2, \infty)$, the mapping $F : L^p \times L^q \to L^{p\wedge q}$ is defined as the Nemytskii operator, with $F(x, y) = f\big(x(\cdot), y(\cdot)\big)$ for all $x \in L^p$, $y \in L^q$.

Note that the definition of F is consistent when parameters p and q vary.

Observe that, for any $p, q \in [2, \infty)$, and fixed $y \in L^q$, then the mapping $x \in L^p \mapsto F(x, y) \in L^{p\wedge q}$ is globally Lipschitz continuous, uniformly in $y \in L^q$, and in p, q. More precisely,

$$\mathrm{Lip}\big(F(\cdot, y)\big) \le \sup_{(z_1,z_2)\in\mathbb{R}^2} |\partial_{z_1} f(z_1, z_2)|.$$

In addition, if $y \in L^q$ and $x_1, x_2 \in L^p$, then one has $F(x_2, y) - F(x_1, y) \in L^p$, with $|F(x_2, y) - F(x_1, y)|_{L^p} \le \sup_{(z_1,z_2)\in\mathbb{R}^2} |\partial_{z_1} f(z_1, z_2)|\ |x_2 - x_1|_{L^p}$.

Nemytskii operators, in general, are not Fréchet differentiable. In addition, second and higher order derivatives are defined only in certain directions (consistently with Hölder inequality).

Gâteaux derivatives of F can be computed as follows: for all $x, y \in L^2$ and $h \in L^p$, $h_1 \in L^{p_1}$, $h_2 \in L^{p_2}$ and $h_3 \in L^{p_3}$,

$$\begin{aligned} &D_x F(x, y).h = \partial_{z_1} f(x, y)h \in L^p, \\ &D_{xx} F(x, y).(h_1, h_2) = \partial_{z_1}^2 f(x, y)h_1 h_2 \in L^{q_2}, \\ &D_{xxx} F(x, y).(h_1, h_2, h_3) = \partial_{z_1}^3 f(x, y)h_1 h_2 h_3 \in L^{q_3}, \end{aligned} \tag{8}$$

with $\frac{1}{q_2} = \frac{1}{p_1} + \frac{1}{p_2}$ and $\frac{1}{q_3} = \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3}$.

The stochastic perturbation in the slow component of (1) is given by a Q-Wiener process. The covariance operator Q is a bounded, self-adjoint, operator on L^2, and satisfies Assumption 2.6.

Assumption 2.6. There exists a family of nonnegative real numbers $\big(q_n\big)_{n\in\mathbb{N}}$ and a complete orthonormal system $\big(f_n\big)_{n\in\mathbb{N}}$ of H, such that $\sup_{n\in\mathbb{N}} q_n < \infty$, and

$$Q = \sum_{n\in\mathbb{N}} q_n \langle f_n, \cdot \rangle f_n.$$

Let $Q^{\frac{1}{2}}$ be defined as

$$Q^{\frac{1}{2}} = \sum_{n\in\mathbb{N}} \sqrt{q_n} \langle f_n, \cdot \rangle f_n.$$

It is assumed that $f_n \in L^\infty$ for all $n \in \mathbb{N}$, and that

$$\sup_{n\in\mathbb{N}} |f_n|_{L^\infty} < \infty.$$

Finally, assume that there exists $\alpha_{\max} \in (0, 1]$ such that, for all $\alpha \in [0, \alpha_{\max})$ and $p \in [2, \infty)$,

$$M_{\alpha,p}(Q^{\frac{1}{2}}, T) = \Big(\int_0^T \|(-A)^\alpha e^{tA} Q^{\frac{1}{2}}\|^2_{\mathcal{R}(L^2,L^p)} dt\Big)^{\frac{1}{2}} < \infty. \tag{9}$$

The Q-Wiener process W^Q is then defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, as follows:

$$W^Q(t) = \sum_{n\in\mathbb{N}} \sqrt{q_n} \gamma_n(t) f_n,$$

where $\big(\gamma_n\big)_{n\in\mathbb{N}}$ is a sequence of independent real-valued Wiener processes. Note that $W^Q(t) = Q^{\frac{1}{2}} W(t)$ where $W(t) = \sum_{n\in\mathbb{N}} \gamma_n(t) f_n$ is a cylindrical Wiener process.

Note that, for all $T \in (0, \infty)$, $M_{\alpha,p}(Q^{\frac{1}{2}}, T) < \infty$ if and only if $M_{\alpha,p}(Q^{\frac{1}{2}}, 1) < \infty$. Thus the condition expressed in (9) does not depend on the time T, it only depends on Q, and on the parameters α and p.

A sufficient condition for (9) to hold is the following: for all $\alpha \in [0, \alpha_{\max})$ and $p \in [2, \infty)$, $\|(-A)^{\alpha-\frac{1}{2}} Q^{\frac{1}{2}}\|_{\mathcal{R}(L^2,L^p)} < \infty$. For another class of sufficient conditions, see Assumption 4.2 and Proposition 8.1.

2.4. Assumptions on the fast process

The fast process $\big(Y^\epsilon(t)\big)_{t\geq 0}$ in (1) is defined in terms of an ergodic Markov process Y, such that, for all $t \geq 0$,

$$Y^\epsilon(t) = Y(\frac{t}{\epsilon}).$$

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

Assumption 2.7. The process $Y = \big(Y(t)\big)_{t\geq 0}$ is a continuous, ergodic, Markov process on $H = L^2$. Its unique invariant probability distribution is denoted by μ.

Moreover, it is assumed that Y and the Q-Wiener process W^Q are independent.

Finally, there exists a parameter $\gamma_{\max} \in (0, \frac{1}{2}]$, such that the following estimates are satisfied: for all $\gamma \in [0, \gamma_{\max})$, all $p \in [2, \infty)$, and $M \in \mathbb{N}$, there exists $C_{\gamma,p,M} \in (0, \infty)$ such that

$$\sup_{t\geq 0} \mathbb{E}|(-A)^\gamma Y(t)|_{L^p}^M \leq C_{\gamma,p,M}(1 + \mathbb{E}|(-A)^\gamma Y(0)|_{L^p}^M), \tag{10}$$

$$\int |(-A)^\gamma y|_{L^p}^M \mu(dy) \leq C_{\gamma,p,M}, \tag{11}$$

The following standard notation is used: $\big(Y_y(t)\big)_{t\geq 0}$ denotes the Markov process with initial condition $Y(0) = y$.

Another key assumption concerning the fast process deals with solvability of Poisson equations, and on regularity properties of the solutions.

Assumption 2.8. Define admissible functions $\phi : H \to \mathbb{R}$, to be such that, for some $q \in [2, \infty)$, ϕ is twice Fréchet differentiable on L^q, and such that there exists $C \in (0, \infty)$ such that for all $y \in H$, $h, h_1, h_2 \in L^q$,

$$|D\phi(y).h| \leq C|h|_{L^q} \quad , \; |D^2\phi(y).(h_1, h_2)| \leq C|h_1|_{L^q}|h_2|_{L^q}.$$

Let $\mathbf{L}$ be the infinitesimal generator of the Markov process Y.

Assume that for any admissible function ϕ, the Poisson equation

$$-\mathbf{L}\psi = \phi - \int \phi d\mu \tag{12}$$

admits a unique solution such that $\int \psi d\mu = 0$, and that this solution is given by

$$\psi(y) = \int_0^\infty \mathbb{E}\big[\phi(Y_y(t)) - \int \phi d\mu\big]dt. \tag{13}$$

Moreover, for all $\gamma \in [0, \gamma_{\max})$, $p \in [2, \infty)$ and $M \in \mathbb{N}_0$, assume that there exists $C_{\gamma,p,M} \in (0, \infty)$ such that the following property is satisfied. Let $\phi : H \to \mathbb{R}$ be an admissible function, and assume that there exists $C(\phi) \in (0, \infty)$, such that for all $y_1, y_2 \in H$

$$|\phi(y_2) - \phi(y_1)| \leq C(\phi)(1 + |(-A)^\gamma y_1|_{L^p}^M + |(-A)^\gamma y_2|_{L^p}^M)|(-A)^\gamma (y_2 - y_1)|_{L^p}. \tag{14}$$

Then the solution ψ of the Poisson equation (12) satisfies, for all $y \in H$,

$$|\psi(y)| \leq C_{\gamma,p,M} C(\phi)(1 + |(-A)^\gamma y|_{L^p}^{M+1}). \tag{15}$$

A sufficient condition to have the estimate (15) satisfied is given by Proposition 2.9.

Proposition 2.9. *Let Assumption 2.7 be satisfied. Assume that for all $\gamma \in [0, \gamma_{\max})$, $p \in [2, \infty)$ and $M \in \mathbb{N}_0$, there exists $C_{\gamma,p,M} \in (0, \infty)$ such that for all $y_1, y_2 \in L^p$,*

$$\int_0^\infty \big(\mathbb{E}|(-A)^\gamma (Y_{y_2}(t) - Y_{y_1}(t))|_{L^p}^M\big)^{\frac{1}{M}} dt \leq C_{\gamma,p,M}|(-A)^\gamma (y_2 - y_1)|_{L^p}.$$

Then the estimate (15) is satisfied.

Proof of Proposition 2.9. By stationarity, ψ is written as follows:

$$\psi(y) = \int \int_0^\infty \mathbb{E}[\phi(Y^y(t)) - \phi(Y^z(t))]dt\mu(dz).$$

Then, using (14) and Assumption 2.7,

$$\begin{aligned}|\psi(y)| &\le C(\phi)\int\int_0^\infty \big(\mathbb{E}|(-A)^\gamma(Y^y(t)-Y^z(t))|_{L^p}^2\big)^{\frac12}\big(1+(\mathbb{E}|(-A)^\gamma Y^y(t)|_{L^p}^{2M})^{\frac12}\\
&\quad+(\mathbb{E}|(-A)^\gamma Y^z(t)|_{L^p}^{2M})^{\frac12}\big)dt\mu(dz)\\
&\le C_{\gamma,p,M}(\phi)\int\int_0^\infty \big(\mathbb{E}|(-A)^\gamma(Y^y(t)-Y^z(t))|_{L^p}^2\big)^{\frac12}dt\big(1+|(-A)^\gamma y|_{L^p}^M\\
&\quad+|(-A)^\gamma z|_{L^p}^M\big)\mu(dz)\\
&\le C_{\gamma,p,M}(\phi)\int |(-A)^\gamma(y-z)|_{L^p}\big(1+|(-A)^\gamma y|_{L^p}^M+|(-A)^\gamma z|_{L^p}^M\big)\mu(dz)\\
&\le C_{\gamma,p,M}\big(1+|(-A)^\gamma y|_{L^p}^{M+1}\big).\quad\square\end{aligned}$$

2.5. Well-posedness and moment estimates

We are now in position to state (and give a sketch of proof of) a well-posedness result for (1), for arbitrary $\epsilon > 0$. Without loss of generality, it is assumed that $\epsilon \in (0, 1)$.

We also state moment estimates for $X^\epsilon(t)$. These estimates are uniform with respect to the parameter $\epsilon \in (0, 1)$. Recall that $\alpha_{\max}$ is defined in Assumption 2.6 and that it is assumed that $\alpha_{\max} > 0$.

Proposition 2.10. *Let $T \in (0,\infty)$. For any $x_0 \in H$, any $y_0 \in H$, and any $\epsilon \in (0, 1)$, the SPDE (1) admits a unique mild solution, with initial conditions $X^\epsilon(0) = x_0$, $Y^\epsilon(0) = y_0$, such that for all $t \in [0, T]$,*

$$X^\epsilon(t) = e^{tA}x_0 + \int_0^t e^{(t-s)A}F(X^\epsilon(s), Y^\epsilon(s))ds + \int_0^t e^{(t-s)A}dW^Q(s). \tag{16}$$

In addition, the following moment estimates are satisfied, uniformly with respect to $\epsilon \in (0, 1)$: for any $T \in (0,\infty)$, $\alpha \in [0, \alpha_{\max})$, $p \ge 2$ and $M \in \mathbb{N}$, there exists $C_{\alpha,p,M}(T) \in (0,\infty)$, such that for all $x_0, y_0 \in L^p$, such that $|(-A)^\alpha x_0|_{L^p} < \infty$,

$$\sup_{\epsilon\in(0,1)}\sup_{t\in[0,T]}\mathbb{E}\big[|(-A)^\alpha X^\epsilon(t)|_{L^p}^M\big] \le C_{\alpha,p,M}(T)\big(1+|(-A)^\alpha x_0|_{L^p}^M+|y_0|_{L^p}^M+M_{\alpha,p}(Q^{\frac12},T)^M\big). \tag{17}$$

Remark 2.11. Using the regularization property (3) of the semigroup $\big(e^{tA}\big)_{t\ge0}$, the moment estimates (17) may be replaced with

$$\sup_{\epsilon\in(0,1)}\mathbb{E}\big[|(-A)^\alpha X^\epsilon(t)|_{L^p}^M\big] \le C_{\alpha,p,M}(T)\big(1+t^{-\alpha M}|x_0|_{L^p}^M+|y_0|_{L^p}^M+M_{\alpha,p}(Q^{\frac12},T)^M\big).$$

Therefore the regularity assumption on the initial condition x_0 may be relaxed.

We conclude this section with a sketch of proof of Proposition 2.10.

Proof of Proposition 2.10. The existence and uniqueness of a mild solution (16) of (1) is obtained by a standard fixed point argument, see for instance [12].

The proof of the moment estimates (17) combines the following observations. On the one hand, owing to (3) and to Lipschitz continuity of F,

$$\Big|(-A)^\alpha\int_0^t e^{(t-s)A}F(X^\epsilon(s), Y^\epsilon(s))ds\Big|_{L^p} \le C\int_0^t (t-s)^{-\alpha}\big(1+|X^\epsilon(s)|_{L^p}+|Y^\epsilon(s)|_{L^p}\big)ds.$$

On the other hand, owing to (9), see Assumption 2.6, the moment estimate

$$\mathbb{E}|(-A)^{\alpha}\int_0^t e^{(t-s)A}dW^Q(s)|_{L^p}^2 \le c_p\int_0^T \|(-A)^{\alpha}e^{tA}Q^{\frac12}\|_{\mathcal{R}(L^2,L^p)}^2 dt$$
$$= c_p M_{\alpha,p}(Q^{\frac12},T)^2 < \infty$$

for the stochastic convolution, is easily obtained, in the case $M = 2$. Higher order moments are estimated using a Burkholder–Davis–Gundy type inequality.

The case $\alpha = 0$ is treated using Gronwall inequality, and then the case $\alpha \in (0, \alpha_{\max})$ follows from the estimates above.

This concludes the sketch of proof of Proposition 2.10. □

3. The averaged equation

Let us first define the so-called averaged coefficient $\overline{F}$.

Definition 3.1. For any $x \in L^2$, define

$$\overline{F}(x) = \int F(x, y)d\mu(y) \in L^2, \tag{18}$$

where μ is the unique invariant probability distribution of the ergodic process Y, see Assumption 2.7.

Since F is the Nemytskii operator associated with a globally Lipschitz continuous function f, using Assumption 2.7, it is straightforward to check that $\overline{F}(x) \in L^p$ if $x \in L^p$, for all $p \in [2,\infty)$.

The derivatives of the averaged coefficient $\overline{F}$ satisfy the following estimates.

Proposition 3.2. *For all $p \in [2,\infty)$, there exists $C_p \in (0,\infty)$ such that for all $h \in L^p$*

$$\sup_{x\in L^2} |D\overline{F}(x).h|_{L^p} \le C_p|h|_{L^p}. \tag{19}$$

For all $p \in [1,\infty)$, and $p_1, p_2 \in [2,\infty)$, such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$, there exists $C_{p_1,p_2} \in (0,\infty)$ such that for all $h_1 \in L^{p_1}$ and $h_2 \in L^{p_2}$,

$$\sup_{x\in L^2} |D^2\overline{F}(x).(h_1,h_2)|_{L^p} \le C_{p_1,p_2}|h_1|_{L^{p_1}}|h_2|_{L^{p_2}}. \tag{20}$$

For all $p \in [1,\infty)$, and $p_1, p_2, p_3 \in [2,\infty)$, such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3}$, there exists $C_{p_1,p_2,p_3} \in (0,\infty)$ such that for all $h_1 \in L^{p_1}$, $h_2 \in L^{p_2}$ and $h_3 \in L^{p_3}$,

$$\sup_{x\in L^2} |D^3\overline{F}(x).(h_1,h_2,h_3)|_{L^p} \le C_{p_1,p_2,p_3}|h_1|_{L^{p_1}}|h_2|_{L^{p_2}}|h_3|_{L^{p_3}}. \tag{21}$$

Proof of Proposition 3.2. The derivatives of F are given by (8).

The conclusion follows using boundedness of the derivatives of f, the Hölder inequality, and integrating with respect to $d\mu(y)$. □

We are now in position to state a well-posedness result for the averaged equation:

$$d\overline{X}(t) = A\overline{X}(t)dt + \overline{F}(\overline{X}(t))dt + dW^Q(t). \tag{22}$$

Proposition 3.3. *Let $T \in (0,\infty)$. For any $x_0 \in H$, the SPDE (22) admits a unique mild solution, with initial condition $\overline{X}(0) = x_0$, such that for all $t \in [0,T]$,*

$$\overline{X}(t) = e^{tA}x_0 + \int_0^t e^{(t-s)A}\overline{F}(\overline{X}(s))ds + \int_0^t e^{(t-s)A}dW^Q(s). \tag{23}$$

In addition, the following moment estimates are satisfied: for any $T \in (0,\infty)$, $\alpha \in [0,\alpha_{\max})$ (where $\alpha_{\max}$ is defined in Assumption 2.6), $p \geq 2$ and $M \in \mathbb{N}$, there exists $C_{\alpha,p,M}(T) \in (0,\infty)$, such that for all $x_0, y_0 \in L^p$, such that $|(-A)^\alpha x_0|_{L^p} < \infty$,

$$\sup_{t\in[0,T]} \mathbb{E}\big[|(-A)^\alpha \overline{X}(t)|_{L^p}^M\big] \leq C_{\alpha,p,M}(T)\big(1 + |(-A)^\alpha x_0|_{L^p}^M + M_{\alpha,p}(Q^{\frac{1}{2}},T)^M\big). \tag{24}$$

The proof is omitted. Existence and uniqueness of the mild solution follows from the global Lipschitz continuity property of $\overline{F}$ (thanks to (19), see Proposition 3.2). The moment estimates (24) are proved using the same arguments as in the proof of Proposition 2.10, in particular using (9), see Assumption 2.6.

To conclude this section, introduce the infinitesimal generator associated with the averaged equation (22):

$$\overline{\mathcal{L}}\varphi(x) = \langle D_x\varphi(x), Ax + \overline{F}(x)\rangle + \frac{1}{2}\sum_{n\in\mathbb{N}} q_n D_x^2\varphi(x).\big(f_n, f_n\big). \tag{25}$$

This definition makes sense for sufficiently regular functions $\varphi : L^2 \to \mathbb{R}$.

4. Statements of the main results

The main results of this article, concerning the error in the averaging principle, are stated in this section. We exhibit both strong and weak orders of convergence, with respect to ϵ. Two situations need to be considered, depending on the regularity properties of the slow and the fast component, more precisely in terms of $\alpha_{\max}$ (defined in Assumption 2.6) and of $\gamma_{\max}$ (defined in Assumption 2.7).

Let us introduce Assumption 4.1 (resp. Assumption 4.2) which defines the *very regular* case (resp. the *less regular* case).

Assumption 4.1. The parameters $\alpha_{\max}$ and $\gamma_{\max}$ satisfy the condition

$$\alpha_{\max} + \gamma_{\max} > 1.$$

Moreover, assume that $\mathrm{Tr}(Q) = \sum_{n\in\mathbb{N}} q_n < \infty$.

Assumption 4.2. Assume that $\big(\sum_{n\in\mathbb{N}} q_n^{\frac{\varrho}{2}}\big)^{\frac{2}{\varrho}} < \infty$, for some

$$\varrho \in \begin{cases} [2,\infty], & d = 1, \\ [2, \frac{d}{d-2}), & d \in \{2,3\}, \end{cases}$$

with usual conventions if $\varrho = \infty$ or if $d = 2$.

In this case, let

$$\alpha_{\max} = \frac{1}{2}\big(1 - \frac{d}{2}(1 - \frac{2}{\varrho})\big). \tag{26}$$

Moreover, assume that $\gamma_{\max} \leq \alpha_{\max}$.

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

The definition of the parameter $\alpha_{\max}$ in Assumption 4.2 is consistent with Assumption 2.6, see Proposition 8.1. Note that in this case, one has $\alpha_{\max} \in (0, \frac{1}{2}]$.

Remark 4.3. If Assumption 4.2 is satisfied, the condition $\alpha_{\max} \geq \gamma_{\max}$ is not restrictive. Indeed, in practice, when the regularity $\alpha_{\max}$ is given, one may always replace $\gamma_{\max}$ with $\min(\alpha_{\max}, \gamma_{\max})$ without extra assumption.

Remark 4.4. The condition $\mathrm{Tr}(Q) < \infty$ in Assumption 4.1 is not very restrictive. For instance, if $\alpha_{\max}$ is characterized by the property that for all $\alpha < \alpha_{\max}$ and all $p \in [2, \infty)$, $\|(-A)^{\alpha-\frac{1}{2}} Q^{\frac{1}{2}}\|_{\mathcal{R}(L^2,L^p)} < \infty$, the condition is satisfied since $\mathrm{Tr}(Q) = \|Q^{\frac{1}{2}}\|_{\mathcal{R}(L^2,L^2)} < \infty$, with $\alpha = \frac{1}{2} < \alpha_{\max}$.

First, in the very regular case (see Section 4.1), the strong (resp. weak) order of convergence is equal to $\frac{1}{2}$ (resp. 1). This coincides with the orders of convergence obtained in [3], where no stochastic perturbation is acting in the slow component, *i.e.* $Q = 0$ in (1). The weak order 1 also essentially coincides with the result from [19], where it is assumed that $\alpha_{\max} = 1$ and $\gamma_{\max} = 0$. Moreover, this also coincides with the orders obtained in the case of SDEs (see for instance [36]). In particular, these values are optimal in general.

Second, in the less regular case (see Section 4.2), Assumption 4.2 is satisfied, and it is proved that the strong (resp. weak) order of convergence is equal to $\frac{\alpha_{\max}}{1+\alpha_{\max}-\gamma_{\max}}$ (resp. $\frac{2\alpha_{\max}}{1+\alpha_{\max}-\gamma_{\max}}$). The proof is based on the application of the result in the regular case for a well-chosen approximate problem, with modified covariance operator Q. It is not known whether these strong and weak orders of convergence are optimal. On the one hand, observe the orders of convergence are maximal when $\gamma_{\max} = \alpha_{\max}$, in which case the strong and weak orders are $\alpha_{\max}$ and $2\alpha_{\max}$ respectively, hence are clearly related to the spatial and temporal regularity of the processes. On the other hand, when $\gamma_{\max}$ is arbitrarily small, the strong and weak orders of convergence are $\frac{\alpha_{\max}}{1+\alpha_{\max}}$ and $\frac{2\alpha_{\max}}{1+\alpha_{\max}}$ respectively. The application of the standard Khasminskii strategy would also lead to a strong order of convergence equal to $\frac{\alpha_{\max}}{1+\alpha_{\max}}$, see [3]. As a consequence, the additional use of regularity properties of the fast process in the analysis allows us to get improved orders of convergence.

4.1. The very regular case

In this section, it is assumed that Assumption 4.1 is satisfied. As a consequence, there exists $\gamma \in (1 - \alpha_{\max}, \gamma_{\max})$, such that $M_{1-\gamma,p}(Q^{\frac{1}{2}}, T) < \infty$ for all $p \in [2, \infty)$ and all $T \in (0, \infty)$, see (9). In particular,

$$M_{1-\gamma,8}(Q^{\frac{1}{2}}, T) < \infty,$$

for all $T \in (0, \infty)$,.

We are now in position to provide precise statements of the results, concerning the order of convergence of the averaging error, in the regular case.

Theorem 4.5. *Let Assumption 4.1 be satisfied. Let $T \in (0, \infty)$, and assume that the initial conditions x_0, y_0, satisfy*

$$|(-A)^{1-\gamma} x_0|_{L^8} + |(-A)^{\gamma+\kappa} y_0|_{L^8} < \infty,$$

for some $\gamma \in (1 - \alpha_{\max}, \gamma_{\max})$ and some $\kappa \in (0, \gamma_{\max} - \gamma)$.

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

Then there exists $C(T,x_0,y_0)\in(0,\infty)$ *such that for all* $\epsilon\in(0,1)$,

$$\sup_{t\in[0,T]}\big(\mathbb{E}|X^\epsilon(t)-\overline{X}(t)|_{L^2}^2\big)^{\frac12}\le C(T,x_0,y_0)\big(\mathrm{Tr}(Q)+M_{1-\gamma,8}(Q^{\frac12},T)\big)^{\frac12}\epsilon^{\frac12}. \tag{27}$$

To state the weak error result, an appropriate notion of nice test functions is used.

Definition 4.6. Let $\varphi:L^2\to\mathbb{R}$. It is called a nice test function if the following derivatives of φ exist and are continuous, and if the estimates below are satisfied.

- There exists $C\in(0,\infty)$ such that for all $x\in L^2$ and $h\in L^2$,

  $$|D\varphi(x).h|\le C|h|_{L^2},$$

  and, for all $x\in L^2$, $h_1,h_2\in L^2$,

  $$|D^2\varphi(x).(h_1,h_2)|\le C|h_1|_{L^2}|h_2|_{L^2}.$$

- For every $p_1,p_2,p_3\in[2,\infty)$ such that $1=\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p_3}$, there exists $C_{p_1,p_2,p_3}\in(0,\infty)$ such that for all $x\in L^2$, and $h_1\in L^{p_1}$, $h_2\in L^{p_2}$, $h_3\in L^{p_3}$,

  $$|D^3\varphi(x).(h_1,h_2,h_3)|\le C_{p_1,p_2,p_3}|h_1|_{L^{p_1}}|h_2|_{L^{p_2}}|h_3|_{L^{p_3}}.$$

For instance, a function $\varphi:L^2\to\mathbb{R}$ of class $\mathcal{C}^3$, with bounded derivatives of order $1,2,3$, is a nice test function. Other nice test functions are constructed as follows:

$$\varphi(x)=\langle\omega,\tilde\varphi(x)\rangle,$$

where $\tilde\varphi:\mathbb{R}\to\mathbb{R}$ is of class $\mathcal{C}_b^3$, and $\omega\in L^\infty$ is a weight function.

The weak error is estimated for the class of nice test functions introduced above.

Theorem 4.7. *Let Assumption 4.1 be satisfied. Let* $T\in(0,\infty)$, *and assume that the initial conditions* x_0, y_0, *satisfy*

$$|(-A)^{1-\gamma}x_0|_{L^8}+|(-A)^{\gamma+\kappa}y_0|_{L^8}<\infty,$$

for some $\gamma\in(1-\alpha_{\max},\gamma_{\max})$ *and some* $\kappa\in(0,\gamma_{\max}-\gamma)$. *Let* $\varphi:L^2\to\mathbb{R}$ *be a nice test function.*

There exists $C(T,x_0,y_0,\varphi)\in(0,\infty)$ *such that for all* $\epsilon\in(0,1)$,

$$\sup_{t\in[0,T]}|\mathbb{E}[\varphi(X^\epsilon(t))]-\mathbb{E}[\varphi(\overline{X}(t))]|\le C(T,x_0,y_0,\varphi)\big(\mathrm{Tr}(Q)+M_{1-\gamma,4}(Q^{\frac12},T)\big)\epsilon. \tag{28}$$

The apparently strong conditions imposed on the initial conditions x_0 and y_0 may be weakened using standard arguments, thanks to the regularization properties (3) of the semigroup $\big(e^{tA}\big)_{t\ge0}$, and minor modifications in the proofs. However, one could not consider the supremum over time $t\in[0,T]$ in (27) and (28). In addition, assuming that the initial conditions possess nice spatial regularity properties allows us to focus on the most important issues solved in this manuscript.

Note that the strong order of convergence is equal to $\frac12$, whereas the weak order of convergence is equal to 1. As explained above, these values are optimal.

The proofs of Theorems 4.5 and 4.7 are postponed to Sections 6 and 7 respectively.

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

4.2. The less regular case

In this section it is assumed that Assumption 4.2 is satisfied, in particular $\alpha_{\max}$ is defined by (26) and $\gamma_{\max} \le \alpha_{\max}$. Let

$$\beta_{\max} = \frac{\alpha_{\max}}{1 + \alpha_{\max} - \gamma_{\max}},$$

and observe that $\beta_{\max} \le \frac{1}{2}$.

Theorem 4.8. *Let Assumption 4.2 be satisfied. Let $T \in (0, \infty)$, and assume that the initial conditions x_0, y_0, satisfy*

$$|(-A)^{1-\gamma} x_0|_{L^8} + |(-A)^{\gamma+\kappa} y_0|_{L^8} < \infty,$$

for some $\gamma \in (0, \gamma_{\max})$ and some $\kappa \in (0, \alpha_{\max} - \gamma)$.

For any $\beta \in [0, \beta_{\max})$, there exists $C_\beta(T, x_0, y_0, Q) \in (0, \infty)$ such that for all $\epsilon \in (0, 1)$, the strong error is estimated by

$$\sup_{t\in[0,T]} \left(\mathbb{E}|X^\epsilon(t) - \overline{X}(t)|_{L^2}^2\right)^{\frac{1}{2}} \le C_\beta(T, x_0, y_0, Q)\epsilon^\beta. \tag{29}$$

Moreover, let $\varphi : L^2 \to \mathbb{R}$ be a nice test function. For any $\beta \in [0, \beta_{\max})$, there exists $C_\beta(T, x_0, y_0, Q, \varphi) \in (0, \infty)$ such that for all $\epsilon \in (0, 1)$, the weak error is estimated by

$$\sup_{t\in[0,T]} |\mathbb{E}[\varphi(X^\epsilon(t))] - \mathbb{E}[\varphi(\overline{X}(t))]| \le C_\beta(T, x_0, y_0, Q, \varphi)\epsilon^{2\beta}. \tag{30}$$

Note that in Theorem 4.8, the weak order is equal to twice the strong order, as discussed above. The proof of Theorem 4.8 is postponed to Section 8.

Remark 4.9. Whether the supremum in the strong error estimates (27) and (29) can be put inside the expectation is not clear. This question is left open for future work.

5. Auxiliary regularity results for solutions of the Poisson equation

This section is devoted to the analysis of the Poisson equation below: for any $x \in L^2$ and $\theta \in L^2$, define $\Phi(x, \cdot, \theta) : L^2 \to \mathbb{R}$ as the unique solution of

$$-\mathbf{L}\Phi(x, \cdot, \theta) = \langle F(x, \cdot) - \overline{F}(x), \theta\rangle, \tag{31}$$

with the condition $\int \Phi(x, \cdot, \theta)d\mu = 0$. Observe that $\theta \mapsto \Phi(x, y, \theta)$ is a (possibly unbounded) linear mapping.

Recall that $\mathbf{L}$ is the generator of the Markov process Y. It is assumed that Assumption 2.8 is satisfied.

The function Φ plays a key role in the analysis of the error in the averaging principle, both in the strong and in the weak senses. It is straightforward to obtain estimates on $\Phi(x, y, \theta)$, on $D_x\Phi(x, y, \theta).h$ and on $D_x^2\Phi(x, y, \theta).(h_1, h_2)$, in terms on L^p norms of $x, y, \theta, h, h_1, h_2$ (for well-chosen p), see Lemmas 5.1, 5.3 and 5.5. The main original results in this manuscript are estimates of $\Phi(x, y, \theta)$ in terms of $|(-A)^{-\gamma}\theta|_{L^p}$ (see Lemma 5.2), and of $D_x\Phi(x, y, \theta).h$ in terms of $|(-A)^{-\gamma}h|_{L^p}$ (see Lemma 5.4), for positive $\gamma \in (0, \gamma_{\max})$. These two results are specific to the analysis of the averaging principle for parabolic SPDEs, and they allow us to exhibit the trade-off between the regularity properties of the slow and fast processes in the identification of the strong and weak orders of convergence discussed above. These results are consequences of Proposition 2.1.

First, Lemmas 5.1 and 5.2 deal with estimates of $\Phi(x,y,\theta)$. In particular, note that Lemma 5.1 implies the well-posedness of (31).

Lemma 5.1. *Let $p \in [2,\infty)$ and $p' = \frac{p}{p-1} \in (1,2]$. There exists $C_p \in (0,\infty)$, such that for all $x \in L^2$, $y \in L^p$ and all $\theta \in L^{p'}$,*

$$|\Phi(x,y,\theta)| \leq C_q(1+|y|_{L^p})|\theta|_{L^{p'}}.$$

Proof. For any fixed $x \in L^2$ and $\theta \in L^2 \subset L^{p'}$, the mapping $y \mapsto \langle F(x,y) - \overline{F}(x), \theta\rangle$ is an admissible function (with $q = 4$). In addition, using Lipschitz continuity of F (and boundedness of the derivatives of f, see Assumption 2.3), one has the estimate

$$|\langle F(x,y_2) - F(x,y_1), \theta\rangle| \leq |F(x,y_2) - F(x,y_1)|_{L^p}|\theta|_{L^{p'}} \leq C|y_2 - y_1|_{L^p}|\theta|_{L^{p'}}.$$

This proves that (14) is satisfied, with the parameters $\alpha = 0$, p, and $M = 0$. By Assumption 2.8, then (15) is satisfied, which concludes the proof of Lemma 5.1. □

Lemma 5.2. *Let $\gamma \in (0,\gamma_{\max})$. For all $\kappa \in (0,\gamma_{\max} - \gamma)$, there exists $C_{\gamma,\kappa} \in (0,\infty)$ such that for all $x, y \in L^4$ and $\theta \in L^2$, then*

$$|\Phi(x,y,\theta)| \leq C_{\gamma,\kappa}\big(1 + |(-A)^{\gamma+\kappa}x|_{L^4}^2 + |(-A)^{\gamma+\kappa}y|_{L^4}^2\big)|(-A)^{-\gamma}\theta|_{L^2}.$$

Proof. Observe that

$$\begin{aligned}|\langle F(x,y_2) - F(x,y_2), \theta\rangle| &\leq C_\gamma|(-A)^{\gamma}(F(x,y_2) - F(x,y_1))|_{L^2}|(-A)^{-\gamma}\theta|_{L^2}\\ &\leq C_\gamma\big(1 + |(-A)^{\gamma+\kappa}x|_{L^4} + |(-A)^{\gamma+\kappa}y_1|_{L^4} + |(-A)^{\gamma+\kappa}y_2|_{L^4}\big)\\ &\qquad |(-A)^{\gamma+\kappa}(y_2 - y_1)|_{L^4}|(-A)^{-\gamma}\theta|_{L^2},\end{aligned}$$

using the second inequality in Proposition 2.1. This proves that (14) is satisfied, thus (15) follows, and this concludes the proof of Lemma 5.2. □

Lemmas 5.3 and 5.4 deal with the first order derivative of $\Phi(x,y,\theta)$ with respect to x.

Lemma 5.3. *There exists $C \in (0,\infty)$, such that for all $x \in L^2$, $y, \theta, h \in L^4$,*

$$|\langle D_x\Phi(x,y,\theta), h\rangle| \leq C(1+|y|_{L^4})\min\Big(|\theta|_{L^4}|h|_{L^2}, |\theta|_{L^2}|h|_{L^4}\Big).$$

Moreover, for all $x \in L^2$, $y, h \in L^8$, $\theta \in L^{\frac{4}{3}}$, one has

$$|\langle D_x\Phi(x,y,\theta), h\rangle| \leq C(1+|y|_{L^8})|\theta|_{L^{\frac{4}{3}}}|h|_{L^8}.$$

Proof. For all $x, h \in L^2$, $\theta \in L^4$, the function $y \mapsto \langle D_x\Phi(x,y,\theta), h\rangle$ solves the Poisson equation

$$-\mathbf{L}\big(D_x\Phi(x,\cdot,\theta).h\big) = \phi_{x,\theta,h}$$

where $\phi_{x,\theta,h}(y) = \langle D_x\big(F(x,\cdot) - \overline{F}(x)\big).h, \theta\rangle$. It is straightforward to check that $\phi_{x,\theta,h}$ is an admissible function (by Assumption 2.3, f is of class $\mathcal{C}^3$ with bounded derivatives), with $q = 8$.

Let $x, h \in L^2$ and $\theta \in L^4$, then for all $y_1, y_2 \in L^4$, one has

$$\begin{aligned}|\phi_{x,\theta,h}(y_2) - \phi_{x,\theta,h}(y_1)| &= \big|\langle D_x\big(F(x,y_2) - F(x,y_1)\big).h, \theta\rangle\big|\\ &= \big|\langle\big(\partial_{z_1}f(x,y_2) - \partial_{z_1}f(x,y_1)\big)h, \theta\rangle\big|\end{aligned}$$

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

$$= \left|\langle\big(\partial_{z_1} f(x,y_2) - \partial_{z_1} f(x,y_1)\big)\theta, h\rangle\right|$$
$$\leq C|y_2 - y_1|_{L^4} \min\Big(|\theta|_{L^4}|h|_{L^2}, |\theta|_{L^2}|h|_{L^4}\Big),$$

using the Hölder inequality and by the boundedness of the first-order partial derivative $\partial_{z_1} f(z_1, z_2)$.

Alternatively,

$$|\phi_{x,\theta,h}(y_2) - \phi_{x,\theta,h}(y_1)| \leq C|y_2 - y_1|_{L^8}|\theta|_{L^{\frac{4}{3}}}|h|_{L^8}.$$

Thus (14), and consequently (15), are satisfied, for an appropriate choice of the parameters. This concludes the proof of Lemma 5.3. □

Lemma 5.4. *Let $\gamma \in (0, \gamma_{\max})$. For all $\kappa \in (0, \gamma_{\max} - \gamma)$, there exists $C_{\gamma,\kappa} \in (0, \infty)$ such that for all $x, y, \theta \in H$, then*

$$|\langle D_x \Phi(x,y,\theta), h\rangle| \leq C_{\gamma,\kappa}\big(1 + |(-A)^{\gamma+\kappa}x|_{L^8}^2 + |(-A)^{\gamma+\kappa}y|_{L^8}^2\big)$$
$$\min\Big(|(-A)^{\gamma+\frac{\kappa}{2}}\theta|_{L^4}|(-A)^{-\gamma}h|_{L^2}, |(-A)^{\gamma+\frac{\kappa}{2}}\theta|_{L^2}|(-A)^{-\gamma}h|_{L^4}\Big).$$

Proof. Let x, θ, h be fixed. Proceeding as in the proof of Lemma 5.3, for all $y_1, y_2 \in H$,

$$|\phi_{x,\theta,h}(y_2) - \phi_{x,\theta,h}(y_1)| = \left|\langle\big(\partial_x f(x,y_2) - \partial_x f(x,y_1)\big)h, \theta\rangle\right|$$
$$= \left|\langle\big(\partial_x f(x,y_2) - \partial_x f(x,y_1)\big)\theta, h\rangle\right|,$$

thus, thanks to Hölder inequality and to the first inequality in Proposition 2.1, it is sufficient to consider

$$|(-A)^{\gamma+\frac{\kappa}{2}}\big(\partial_x f(x,y_2) - \partial_x f(x,y_1)\big)|_{L^4}$$
$$\leq C_{\gamma,\kappa}(1 + |(-A)^{\gamma+\kappa}y_1|_{L^8} + |(-A)^{\gamma+\kappa}y_2|_{L^8})|(-A)^{\gamma+\kappa}(y_2 - y_1)|_{L^8},$$

thanks to the second inequality of Proposition 2.1. It remains to use Assumption 2.8 to conclude the proof of Lemma 5.3. □

Finally, it remains to state and prove a result, Lemma 5.5, concerning the second order derivative.

Lemma 5.5. *There exists $C \in (0,\infty)$ such that for all $x \in L^2$ $y, \theta \in L^4$ and $h_1, h_2 \in L^8$,*

$$|D_x^2 \Phi(x,y,\theta).(h_1,h_2)| \leq C(1 + |y|_{L^4}) \min\Big(|\theta|_{L^4}|h_1|_{L^4}|h_2|_{L^4}, |\theta|_{L^2}|h_1|_{L^8}|h_2|_{L^8}\Big).$$

Proof. For all x, θ, h_1, h_2, the function $y \mapsto D_x^2 \Phi(x,y,\theta).(h_1,h_2)$ solves the Poisson equation

$$-\mathbf{L}\big(D_x^2 \Phi(x,\cdot,\theta).(h_1,h_2)\big) = \phi^{(2)}_{x,\theta,h_1,h_2}$$

where $\phi^{(2)}_{x,\theta,h_1,h_2}(y) = \langle D_x^2\big(F(x,\cdot) - \overline{F}(x)\big).(h_1,h_2), \theta\rangle$. It is straightforward to check that $\phi^{(2)}_{x,\theta,h}$ is an admissible function (thanks to Assumption 2.3, f is of class $\mathcal{C}^4$ with bounded derivatives of order $1,\ldots,4$).

For all $y_1, y_2 \in H$, using the boundedness of the third-order derivative $\partial_x^{(3)} f$ and the Hölder inequality, one obtains

$$|\phi^{(2)}_{x,\theta,h_1,h_2}(y_2) - \phi^{(2)}_{x,\theta,h_1,h_2}(y_1)| \leq C|y_2 - y_1|_{L^4} \min\Big(|\theta|_{L^4}|h_1|_{L^4}|h_2|_{L^4}, |\theta|_{L^2}|h_1|_{L^8}|h_2|_{L^8}\Big).$$

Thus it remains to apply Assumption 2.8 to conclude the proof of Lemma 5.5. □

6. Proof of Theorem 4.5

The goal of this section is to provide a proof of Theorem 4.5, *i.e.* that under Assumption 4.1, the strong order of convergence in the averaging principle is equal to $\frac{1}{2}$.

Let $T \in (0,\infty)$. Thanks to Assumption 4.1, let also $\gamma \in (1-\alpha_{\max}, \gamma_{\max})$, $\kappa \in (0, \gamma_{\max}-\gamma)$, and let the initial conditions x_0 and y_0 satisfy $|(-A)^{1-\gamma}x_0|_{L^8} + |(-A)^{\gamma+\kappa}y_0|_{L^8} < \infty$.

Introduce the auxiliary function $\delta F(x,y) = F(x,y) - \overline{F}(x)$. Thanks to the mild formulations (16) and (23), the following decomposition of the averaging error is obtained:

$$X^\epsilon(t) - \overline{X}(t) = \int_0^t e^{(t-s)A}\big(F(X^\epsilon(s),Y^\epsilon(s)) - F(\overline{X}(s),Y^\epsilon(s))\big)ds + \int_0^t e^{(t-s)A}\delta F(\overline{X}(s),Y^\epsilon(s))ds.$$

Recall that F is globally Lipschitz-continuous, owing to Assumption 2.3. The mean-square error is then bounded from above as follows:

$$\begin{aligned}
\mathbb{E}\big|X^\epsilon(t) - \overline{X}(t)\big|^2 &\le 2\mathbb{E}\Big[\Big|\int_0^t e^{(t-s)A}\big(F(X^\epsilon(s),Y^\epsilon(s)) - F(\overline{X}(s),Y^\epsilon(s))\big)ds\Big|^2\Big] \\
&\quad + 2\mathbb{E}\Big[\Big|\int_0^t e^{(t-s)A}\delta F(\overline{X}(s),Y^\epsilon(s))ds\Big|^2\Big] \\
&\le 2T\int_0^t \mathbb{E}\big|e^{(t-s)A}\big(F(X^\epsilon(s),Y^\epsilon(s)) - F(\overline{X}(s),Y^\epsilon(s))\big)\big|^2 ds \\
&\quad + 2\mathbb{E}\Big[\langle\int_0^t e^{(t-s)A}\delta F(\overline{X}(s),Y^\epsilon(s))ds, \int_0^t e^{(t-r)A}\delta F(\overline{X}(r),Y^\epsilon(r))dr\rangle\Big] \\
&\le CT\int_0^t \mathbb{E}\big|X^\epsilon(s) - \overline{X}(s)\big|^2 ds \\
&\quad + 4\int_0^t\int_s^t \mathbb{E}\Big[\langle e^{(t-s)A}\delta F(\overline{X}(s),Y^\epsilon(s)), e^{(t-r)A}\delta F(\overline{X}(r),Y^\epsilon(r))\rangle\Big]drds.
\end{aligned}$$

Let $\theta_{s,t}(r) = e^{(2t-s-r)A}\delta F(\overline{X}(s),Y^\epsilon(s))$. Observe that $\partial_r\theta_{s,t}(r) = -A\theta_{s,t}(r)$. Using the definition (31) of Φ, considering the quantity $\mathbb{E}\Phi(\overline{X}(t),Y^\epsilon(t),\theta_{s,t}(t)) - \mathbb{E}\Phi(\overline{X}(s),Y^\epsilon(s),\theta_{s,t}(s))$, and applying the Itô formula, one obtains

$$\begin{aligned}
\int_s^t \mathbb{E}\Big[\langle e^{(t-s)A}\delta F(\overline{X}(s),Y^\epsilon(s)), e^{(t-r)A}\delta F(\overline{X}(r),Y^\epsilon(r))\rangle\Big]dr
&= \int_s^t \mathbb{E}\big[-\mathbf{L}\Phi\big(\overline{X}(r),Y^\epsilon(r),\theta_{s,t}(r)\big)\big]dr \\
&= \mathcal{I}_1^\epsilon(s,t) + \mathcal{I}_2^\epsilon(s,t) + \mathcal{I}_3^\epsilon(s,t),
\end{aligned}$$

where

$$\mathcal{I}_1^\epsilon(s,t) = \epsilon\mathbb{E}\Phi(\overline{X}(s),Y^\epsilon(s),\theta_{s,t}(s)) - \epsilon\mathbb{E}\Phi(\overline{X}(t),Y^\epsilon(t),\theta_{s,t}(t)), \tag{32}$$

$$\mathcal{I}_2^\epsilon(s,t) = -\epsilon\int_s^t \mathbb{E}\Big[\Phi\big(\overline{X}(r),Y^\epsilon(r),A\theta_{s,t}(r)\big)\Big]dr, \tag{33}$$

$$\mathcal{I}_3^\epsilon(s,t) = \epsilon\int_s^t \mathbb{E}\Big[\overline{\mathcal{L}}\Phi\big(\overline{X}(r),Y^\epsilon(r),\theta_{s,t}(r)\big)\Big]dr, \tag{34}$$

where $\overline{\mathcal{L}}$ is the infinitesimal generator associated with the averaged equation (22), see (25).

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

For future use, a more detailed decomposition of the third term is introduced: $\mathcal{I}_3^\epsilon(s,t) = \mathcal{I}_{3,1}^\epsilon(s,t) + \mathcal{I}_{3,2}^\epsilon(s,t) + \mathcal{I}_{3,3}^\epsilon(s,t)$, with

$$\mathcal{I}_{3,1}^\epsilon(s,t) = \epsilon \int_s^t \mathbb{E}\langle \overline{F}(\overline{X}(r)), D_x \Phi\big(\overline{X}(r), Y^\epsilon(r), \theta_{s,t}(r)\big)\rangle dr \tag{35}$$

$$\mathcal{I}_{3,2}^\epsilon(s,t) = \epsilon \int_s^t \mathbb{E}\langle A\overline{X}(r), D_x \Phi\big(\overline{X}(r), Y^\epsilon(r), \theta_{s,t}(r)\big)\rangle dr \tag{36}$$

$$\mathcal{I}_{3,3}^\epsilon(s,t) = \frac{\epsilon}{2} \int_s^t \mathbb{E}\mathrm{Tr}\Big(QD_x^2 \Phi\big(\overline{X}(r), Y^\epsilon(r), \theta_{s,t}(r)\big)\Big) dr \tag{37}$$

Lemmas 6.1, 6.2 and 6.3 state the necessary estimates in order to conclude the analysis of the strong error. Observe that Assumption 4.1 is only used effectively in Lemma 6.3.

Lemma 6.1. *There exists $C(T) \in (0,\infty)$, such that, for all $\epsilon \in (0,1)$,*

$$\sup_{0\le s\le t\le T} |\mathcal{I}_1^\epsilon(s,t)| \le C(T)\epsilon(1 + |x_0|_{L^2}^2 + |y_0|_{L^2}^2).$$

Lemma 6.2. *There exists $C(T) \in (0,\infty)$, such that, for all $\epsilon \in (0,1)$,*

$$\sup_{0\le s<t\le T} (t-s)^{\frac{1}{2}} |\mathcal{I}_2^\epsilon(s,t)| \le C(T)\epsilon(1 + |x_0|_{L^2}^2 + |y_0|_{L^2}^2).$$

Lemma 6.3. *Let Assumption 4.1 be satisfied, and let $\gamma \in (1-\alpha_{\max}, \gamma_{\max})$ and $\kappa \in (0,\gamma)$. There exists $C_{\gamma,\kappa}(T) \in (0,\infty)$ such that, for all $\epsilon \in (0,1)$,*

$$\begin{aligned}\sup_{0\le s<t\le T} & (t-s)^{\gamma+\frac{\kappa}{2}} |\mathcal{I}_3^\epsilon(s,t)| \\ & \le C_{\gamma,\kappa}(T)\epsilon\big(1 + |(-A)^{\gamma+\kappa}x_0|_{L^8}^3 + |(-A)^{\gamma+\kappa}y_0|_{L^8}^3\big) \\ & \quad \times \big(1 + |(-A)^{1-\gamma}x_0|_{L^8} + M_{1-\gamma,8}(Q^{\frac{1}{2}}) + \mathrm{Tr}(Q)\big).\end{aligned}$$

The proofs of the three auxiliary lemmas above are provided below, then the proof of Theorem 4.5 is concluded.

Proof of Lemma 6.1. For $r \in \{s,t\}$, note that $\mathbb{E}|\theta_{s,t}(r)|_{L^p}^2 \le C_p(1 + |x_0|_{L^p}^2 + |y_0|_{L^p}^2)$, for all $p \in [2,\infty)$ since F has at most linear growth, and thanks to moment estimates, in Proposition 3.3 and in Assumption 2.7. Thus, thanks to Lemma 5.1.

$$\begin{aligned}\mathbb{E}|\Phi(\overline{X}(r), Y^\epsilon(r), \theta_{s,t}(r))| &\le C(1 + (\mathbb{E}|Y^\epsilon(r)|_{L^2}^2)^{\frac{1}{2}})(\mathbb{E}|\theta_{s,t}(r)|_{L^2}^2)^{\frac{1}{2}} \\ &\le C(1 + |y_0|_{L^2})(1 + |x_0|_{L^2} + |y_0|_{L^2}),\end{aligned}$$

which concludes the proof. □

Proof of Lemma 6.2. For $r \in (s,t)$, using Lemma 5.1, and Assumption 2.7,

$$\begin{aligned}\mathbb{E}|\Phi\big(\overline{X}(r), Y^\epsilon(r), A\theta_{s,t}(r)\big)| &\le C(1 + |y_0|_{L^2})(\mathbb{E}|A\theta_{s,t}(r)|_{L^2}^2)^{\frac{1}{2}} \\ &\le C\|Ae^{(2t-s-r)A}\|_{\mathcal{L}(L^2,L^2)}(1 + |x_0|_{L^2}^2 + |y_0|_{L^2}^2).\end{aligned}$$

Using twice the regularization property (3) of the semigroup, one has

$$\begin{aligned}\int_s^t \|Ae^{(2t-s-r)A}\|_{\mathcal{L}(L^2,L^2)} dr &\le \int_0^T \|(-A)^{\frac{1}{2}}e^{(t-r)A}\|_{\mathcal{L}(L^2,L^2)} dr \|(-A)^{\frac{1}{2}}e^{(t-s)A}\|_{\mathcal{L}(L^2,L^2)} \\ &\le C(T)(t-s)^{-\frac{1}{2}}.\end{aligned}$$

Thus one obtains

$$(t-s)^{\frac{1}{2}}|\mathcal{I}_2^\epsilon(s,t)| \le C(T)\epsilon(1+|x_0|_{L^2}^2+|y_0|_{L^2}^2),$$

which concludes the proof. □

Proof of Lemma 6.3. First, Lemma 5.3 yields

$$\begin{aligned}
|\mathcal{I}_{3,1}^\epsilon(s,t)| &\le \epsilon\int_s^t \mathbb{E}\big[(1+|Y^\epsilon(r)|_{L^4})|\overline{F}(\overline{X}(r))|_{L^4}|\theta_{s,t}(r)|_{L^2}\big]dr\\
&\le C_T\epsilon(1+|y_0|_{L^4})(1+|x_0|_{L^4})\big(\mathbb{E}|\theta_{s,t}(r)|_{L^2}^2\big)^{\frac{1}{2}}\\
&\le C_T\epsilon(1+|x_0|_{L^4}^2+|y_0|_{L^4}^2),
\end{aligned}$$

using the moment estimates, in Proposition 3.3 and in Assumption 2.7, and the estimate given in the proof of Lemma 6.1.

Second, let $\gamma\in[0,\gamma_{\max})$, and $\kappa\in(\gamma_{\max}-\gamma)$. Thanks to Lemma 5.4, and Hölder inequality, one obtains

$$\begin{aligned}
|\mathcal{I}_{3,2}^\epsilon(s,t)| &\le C\epsilon\int_s^t (\mathbb{E}|(-A)^{1-\gamma}\overline{X}(r)|_{L^2}^4)^{\frac{1}{4}}(\mathbb{E}|(-A)^{\gamma+\frac{\kappa}{2}}\theta_{s,t}(r)|_{L^4}^2)^{\frac{1}{2}}\\
&\qquad\big(1+(\mathbb{E}|(-A)^{\gamma+\kappa}\overline{X}(r)|_{L^8}^8)^{\frac{1}{4}}+(\mathbb{E}|(-A)^{\gamma+\kappa}Y^\epsilon(r)|_{L^8}^8)^{\frac{1}{4}}\big)dr\\
&\le C\epsilon(t-s)^{-\gamma-\frac{\kappa}{2}}(1+|x_0|_{L^4}+|y_0|_{L^4})\sup_{r\in[0,T]}(\mathbb{E}|(-A)^{1-\gamma}\overline{X}(r)|_{L^2}^4)^{\frac{1}{4}}\\
&\qquad\big(1+|(-A)^{\gamma+\kappa}y_0|_{L^8}^2+\sup_{r\in[0,T]}(\mathbb{E}|(-A)^{\gamma+\kappa}\overline{X}(r)|_{L^8}^8)^{\frac{1}{4}}\big).
\end{aligned}$$

Observe that the conditions on γ and κ above imply that $\gamma+\kappa\le 1-\gamma$. Using the moment estimates in Proposition 3.3 and in Assumption 2.7, one obtains

$$\begin{aligned}
|\mathcal{I}_{3,2}^\epsilon(s,t)| &\le C_{\gamma,\kappa,T}\epsilon(t-s)^{-\gamma-\frac{\kappa}{2}}(1+|(-A)^{1-\gamma}x_0|_{L^8}^3+|(-A)^{\gamma+\kappa}y_0|_{L^8}^3)\\
&\quad\times\big(1+|(-A)^{1-\gamma}x_0|_{L^8}+M_{1-\gamma,8}(Q^{\frac{1}{2}})\big).
\end{aligned}$$

It remains to deal with the trace term, $\mathcal{I}_{3,3}^\epsilon$. Using Lemma 5.5 and Assumption 2.6,

$$\begin{aligned}
|\mathcal{I}_{3,3}^\epsilon(s,t)| &\le C\epsilon\int_s^t\sum_{n\in\mathbb{N}}q_n|D_x^2\Phi(\overline{X}(r),Y^\epsilon(r),\theta_{s,t}(r)).(f_n,f_n)|dr\\
&\le C\epsilon\sum_{n\in\mathbb{N}}q_n|f_n|_{L^4}^2\int_s^t(\mathbb{E}|\theta_{s,t}(r)|_{L^4}^2)^{\frac{1}{2}}(1+(\mathbb{E}|Y^\epsilon(r)|_{L^4}^2)^{\frac{1}{2}})dr\\
&\le C\epsilon\mathrm{Tr}(Q)(1+|x_0|_{L^4}^2+|y_0|_{L^4}^2).
\end{aligned}$$

Gathering the estimates on $|\mathcal{I}_{3,1}^\epsilon(s,t)|$, $|\mathcal{I}_{3,2}^\epsilon(s,t)|$ and $|\mathcal{I}_{3,3}^\epsilon(s,t)|$ then concludes the proof of Lemma 6.3. □

Remark 6.4. The assumption that $\mathrm{Tr}(Q)=\sum_{n\in\mathbb{N}}q_n$ is finite may be removed, using further regularity properties of the second order derivative $D_x^2\Phi$.

We are now in position to provide the proof of Theorem 4.5.

Proof of Theorem 4.5. Gathering estimates from Lemmas 6.1, 6.2 and 6.3, gives

$$\begin{aligned}\mathbb{E}\big|X^\epsilon(t)-\overline{X}(t)\big|_{L^2}^2 &\le CT\int_0^t \mathbb{E}\big|X^\epsilon(s)-\overline{X}(s)\big|_{L^2}^2 ds + \int_0^t \Big(|\mathcal{I}_1^\epsilon(s,t)| + |\mathcal{I}_2^\epsilon(s,t)| \\ &\quad + |\mathcal{I}_3^\epsilon(s,t)| \Big) ds \\ &\le CT\int_0^t \mathbb{E}\big|X^\epsilon(s)-\overline{X}(s)\big|_{L^2}^2 ds + C(T,x_0,y_0) \\ &\quad \times \big(\mathrm{Tr}(Q) + M_{1-\gamma,8}(Q^{\frac12},T)\big)\epsilon.\end{aligned}$$

It remains to apply the Gronwall Lemma to conclude the proof. □

7. Proof of Theorem 4.7

The goal of this section is to provide a proof of Theorem 4.7, *i.e.* that in the very regular case, under Assumption 4.1, the weak order of convergence in the averaging principle is equal to 1.

Let $T \in (0,\infty)$. Thanks to Assumption 4.1, let also $\gamma \in (1-\alpha_{\max}, \gamma_{\max})$, $\kappa \in (0, \gamma_{\max}-\gamma)$, and let the initial conditions x_0 and y_0 satisfy $|(-A)^{1-\gamma}x_0|_{L^8} + |(-A)^{\gamma+\kappa}y_0|_{L^8} < \infty$.

A key tool in the analysis is the function $\overline{u}$ defined below:

$$\overline{u}(t,x) = \mathbb{E}[\varphi(\overline{X}^x(t))]. \tag{38}$$

Note that $\overline{u}$ is the solution of the Kolmogorov equation

$$\partial_t \overline{u} = \overline{\mathcal{L}}\overline{u},$$

with initial condition $\overline{u}(0,\cdot) = \varphi$, where $\overline{\mathcal{L}}$ is the infinitesimal generator associated with the averaged equation (22), see (25).

To deal with this infinite dimensional PDE, usually an auxiliary approximation procedure is employed, see for instance [6], in order to justify the computations. To simplify notation, this is omitted in this manuscript.

The regularity properties stated in Proposition 7.1 play a fundamental role in the analysis of the weak error below. When the auxiliary approximation procedure mentioned above is explicitly written, the upper bounds are meant to be independent of the approximation parameter.

Proposition 7.1 (*Regularity Properties of the Derivatives of $\overline{u}$*)**.** *Let φ be a nice test function.*

For all $\beta \in [0,1)$, there exists $C_\beta(T) \in (0,\infty)$, such that for all $t \in (0,T]$, and $x,h \in L^2$,

$$|D_x\overline{u}(t,x).(-A)^\beta h| \le C_\beta(T)t^{-\beta}|h|_{L^2}. \tag{39}$$

For all $\beta_1,\beta_2 \in [0,1)$ such that $\beta_1+\beta_2<1$, there exists $C_{\beta_1,\beta_2}(T) \in (0,\infty)$, such that for all $t \in (0,T]$, and all $x,h_1,h_2 \in L^2$,

$$\Big|D_x^2\overline{u}(t,x).\big((-A)^{\beta_1}h_1,(-A)^{\beta_2}h_2\big)\Big| \le C_{\beta_1,\beta_2}(T)t^{-\beta_1-\beta_2}|h_1|_{L^2}|h_2|_{L^2}. \tag{40}$$

In addition, for $p_1,p_2,p_3 \in [2,\infty)$ such that $1 = \frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p_3}$, there exists $C_{p_1,p_2,p_3} \in (0,\infty)$, such that for all $t \in [0,T]$, and all $x \in L^2$, $h_1 \in L^{p_1}$, $h_2 \in L^{p_2}$ and $h_3 \in L^{p_3}$,

$$|D_x^3\overline{u}(t,x).(h_1,h_2,h_3)| \le C_{p_1,p_2,p_3}(T)|h_1|_{L^{p_1}}|h_2|_{L^{p_2}}|h_3|_{L^{p_3}}. \tag{41}$$

Regularity properties for infinite dimensional Kolmogorov equations, as stated in Proposition 7.1, are now a classical tool in the analysis of parabolic SPDEs. We refer to [8] for a reference textbook and to [6] for a recent overview of this topic and for further results.

Proof of Proposition 7.1. The proof is based on computing the derivatives of $\overline{u}$ in terms of tangent processes, which are solutions of PDEs with random coefficients (noise is additive in (22)). See for instance [8] where such computations are justified in a general setting.

- The first-order derivative is expressed as follows:

$$D_x\overline{u}(t,x).h = \mathbb{E}\big[\langle D\varphi(\overline{X}^x(t)), \eta^h(t)\rangle\big],$$

where the process $\big(\eta^h(t)\big)_{t\geq 0}$ solves the linear evolution equation

$$d\eta^h(t) = A\eta^h(t)dt + D\overline{F}(\overline{X}^x(t))dt$$

with initial condition $\eta^h(0) = h$. The unique solution of this equation is rewritten in a mild form as follows: for all $t \geq 0$,

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

The operator $\overline{F}$ is globally Lipschitz continuous due to Proposition 3.2, and using the regularization property (3) of the semigroup, one obtains, for all $p \in [2,\infty)$ and for all $t \in (0,T]$,

$$|\eta^h(t)|_{L^p} \leq C_{p,\beta}t^{-\beta}|(-A)^{-\beta}h|_{L^2} + \int_0^t |\eta^h(s)|_{L^2}ds.$$

Applying the Gronwall Lemma yields

$$|\eta^h(t)|_{L^p} \leq C_{p,\beta}(T)t^{-\beta}|(-A)^{-\beta}h|_{L^p}.$$

Since φ is a nice test function, one obtains

$$|D_x\overline{u}(t,x).h| \leq C\mathbb{E}[|\eta^h(t)|_{L^2}] \leq C_\beta(T)t^{-\beta}|(-A)^{-\beta}h|_{L^2},$$

hence (39).

- The second-order derivative is expressed as follows:

$$D_x^2\overline{u}(t,x).(h_1,h_2) = \mathbb{E}\big[\langle D\varphi(\overline{X}^x(t)), \zeta^{h_1,h_2}(t)\rangle\big] + \mathbb{E}\big[D^2\varphi(\overline{X}^x(s)).(\eta^{h_1}(t),\eta^{h_2}(t))\big],$$

where the process $\big(\zeta^{h_1,h_2}(t)\big)_{t\geq 0}$ solves the linear equation

$$d\zeta^{h_1,h_2}(t) = A\zeta^{h_1,h_2}(t)dt + D\overline{F}(\overline{X}^x(t)).\zeta^{h_1,h_2}(t)dt + D^2\overline{F}(\overline{X}^x(t)).(\eta^{h_1}(t),\eta^{h_2}(t))dt,$$

with initial condition $\zeta^{h_1,h_2}(0) = 0$. The unique solution of this equation is rewritten in a mild form as follows: for all $t \geq 0$,

$$\zeta^{h_1,h_2}(t) = \int_0^t e^{(t-s)A} D\overline{F}(\overline{X}^x(s)).\zeta^{h_1,h_2}(s)ds + \int_0^t e^{(t-s)A} D^2\overline{F}(\overline{X}^x(s)).(\eta^{h_1}(s),\eta^{h_2}(s))ds.$$

First, φ being a nice test function, one obtains

$$\mathbb{E}\big[D^2\varphi(\overline{X}^x(s)).(\eta^{h_1}(t),\eta^{h_2}(t))\big] \leq C\mathbb{E}\big[|\eta^{h_1}(t)|_{L^2}|\eta^{h_2}(t)|_{L^2}\big] \leq C_{\beta_1,\beta_2}(T)t^{-\beta_1-\beta_2}|(-A)^{-\beta_1}h_1|_{L^2}|(-A)^{-\beta_2}h_2|_{L^2}.$$

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

Second, thanks to (19) and (20), and using the inequality (7), one obtains, for $\kappa \in (0, 1-\frac{d}{4})$,

$$\begin{aligned}
|\zeta^{h_1,h_2}(t)|_{L^2} &\le C\int_0^t |\zeta^{h_1,h_2}(s)|_{L^2}ds \\
&\quad + \int_0^t (t-s)^{-\frac{d}{4}-\kappa}|D^2\overline{F}(\overline{X}^x(s)).(\eta^{h_1}(s),\eta^{h_2}(s))|_{L^1}ds \\
&\le C\int_0^t |\zeta^{h_1,h_2}(s)|_{L^2}ds \\
&\quad + C_{\beta_1,\beta_2}(T)\int_0^t (t-s)^{-\frac{d}{4}-\kappa}s^{-\beta_1-\beta_2}ds|(-A)^{-\beta_1}h_1|_{L^2}|(-A)^{-\beta_2}h_2|_{L^2} \\
&\le C\int_0^t |\zeta^{h_1,h_2}(s)|_{L^2}ds + C_{\beta_1,\beta_2}(T)|(-A)^{-\beta_1}h_1|_{L^2}|(-A)^{-\beta_2}h_2|_{L^2} \\
&\le C_{\beta_1,\beta_2}(T)|(-A)^{-\beta_1}h_1|_{L^2}|(-A)^{-\beta_2}h_2|_{L^2},
\end{aligned}$$

using the Gronwall Lemma to obtain the last inequality. Thus, since φ is a nice test function, one obtains

$$|\mathbb{E}\big[\langle D\varphi(\overline{X}^x(t)),\zeta^{h_1,h_2}(t)\rangle\big]| \le C_{\beta_1,\beta_2}(T)|(-A)^{-\beta_1}h_1|_{L^2}|(-A)^{-\beta_2}h_2|_{L^2}$$

Gathering the estimates finally yields (40).

- The third-order derivative is expressed as follows:

$$\begin{aligned}
D_x^3\overline{u}(t,x).(h_1,h_2,h_3) &= \mathbb{E}\big[\langle D\varphi(\overline{X}^x(t)),\xi^{h_1,h_2,h_3}(t)\rangle\big] \\
&\quad + \sum_{\sigma\in\mathfrak{S}_3} c_\sigma \mathbb{E}\big[D^2\varphi(\overline{X}^x(t)).(\zeta^{h_{\sigma(1)},h_{\sigma(2)}}(t),\eta^{h_{\sigma(3)}}(t))\big] \\
&\quad + \mathbb{E}\big[D^3\varphi(\overline{X}^x(t)).(\eta^{h_1}(t),\eta^{h_2}(t),\eta^{h_3}(t))\big],
\end{aligned}$$

where $\mathfrak{S}_3$ is the set of permutations of $\{1,2,3\}$, $c_\sigma \in \{0,1\}$, and the process $\big(\xi^{h_1,h_2,h_3}(t)\big)_{t\ge 0}$ solves the linear equation

$$\begin{aligned}
d\xi^{h_1,h_2,h_3}(t) &= A\xi^{h_1,h_2,h_3}(t)dt + D\overline{F}(\overline{X}^x(t)).\xi^{h_1,h_2,h_3}(t)dt \\
&\quad + \sum_{\sigma\in\mathfrak{S}_3} c_\sigma D^2\overline{F}(\overline{X}^x(t)).(\zeta^{h_{\sigma(1)},h_{\sigma(2)}}(t),\eta^{h_{\sigma(3)}}(t))dt \\
&\quad + D^3\overline{F}(\overline{X}^x(t)).(\eta^{h_1}(t),\eta^{h_2}(t),\eta^{h_3}(t))dt
\end{aligned}$$

with initial condition $\xi^{h_1,h_2,h_3}(0) = 0$. The unique solution of this equation is rewritten in a mild form as follows: for all $t \ge 0$,

$$\begin{aligned}
\xi^{h_1,h_2,h_3}(t) &= \int_0^t e^{(t-s)A}D\overline{F}(\overline{X}^x(s)).\xi^{h_1,h_2,h_3}(s)ds \\
&\quad + \sum_{\sigma\in\mathfrak{S}_3} c_\sigma \int_0^t e^{(t-s)A}D^2\overline{F}(\overline{X}^x(s)).(\zeta^{h_{\sigma(1)},h_{\sigma(2)}}(s),\eta^{h_{\sigma(3)}}(s))ds \\
&\quad + \int_0^t e^{(t-s)A}D^3\overline{F}(\overline{X}^x(s)).(\eta^{h_1}(s),\eta^{h_2}(s),\eta^{h_3}(s))ds.
\end{aligned}$$

Since φ is a nice test function, using the estimates above for $\eta^h(t)$ and $\zeta^{h_1,h_2}(t)$, one obtains

$$\begin{aligned}\left|\mathbb{E}\left[D^2\varphi(\overline{X}^x(t)).(\zeta^{h_{\sigma(1)},h_{\sigma(2)}}(t),\eta^{h_{\sigma(3)}}(t))\right]\right| &\le \mathbb{E}\left[|\zeta^{h_{\sigma(1)},h_{\sigma(2)}}(t)|_{L^2}|\eta^{h_{\sigma(3)}}(t)|_{L^2}\right]\\ &\le C(T)\mathbb{E}|h_1|_{L^2}|h_2|_{L^2}|h_3|_{L^2}\\ \left|\mathbb{E}\left[D^3\varphi(\overline{X}^x(t)).(\eta^{h_1}(t),\eta^{h_2}(t),\eta^{h_3}(t))\right]\right| &\le C\mathbb{E}\left[|\eta^{h_1}(t)|_{L^{p_1}}|\eta^{h_2}(t)|_{L^{p_2}}|\eta^{h_3}(t)|_{L^{p_3}}\right]\\ &\le C(T)|h_1|_{L^{p_1}}|h_2|_{L^{p_2}}|h_3|_{L^{p_3}}.\end{aligned}$$

In addition, one has

$$\left|\mathbb{E}\left[\langle D\varphi(\overline{X}^x(t)),\xi^{h_1,h_2,h_3}(t)\rangle\right]\right| \le C\mathbb{E}[|\xi^{h_1,h_2,h_3}(t)|_{L^2}],$$

where, using the mild formulation, the estimates from Proposition 3.2 on derivatives of $\overline{F}$, and the inequality (7), one has

$$\begin{aligned}|\xi^{h_1,h_2,h_3}(t)|_{L^2} \le C\int_0^t |\xi^{h_1,h_2,h_3}(s)|_{L^2}ds\\ &+ C_\kappa \sum_{\sigma\in\mathfrak{S}_3} c_\sigma \int_0^t (t-s)^{-\frac{d}{4}-\kappa}|\zeta^{h_{\sigma(1)},h_{\sigma(2)}}(s)|_{L^2}|\eta^{h_{\sigma(3)}}(s)|_{L^2}ds\\ &+ C_\kappa \int_0^t (t-s)^{-\frac{d}{4}-\kappa}|\eta^{h_1}(s)|_{L^{p_1}}|\eta^{h_2}(s)|_{L^{p_2}}|\eta^{h_3}(s)|_{L^{p_3}}ds\\ &\le C_\kappa \int_0^t |\xi^{h_1,h_2,h_3}(s)|_{L^2}ds + C|h_1|_{L^2}|h_2|_{L^2}|h_3|_{L^2}\\ &+ C|h_1|_{L^{p_1}}|h_2|_{L^{p_2}}|h_3|_{L^{p_3}},\end{aligned}$$

with $1=\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p_3}$. Applying Gronwall Lemma then yields (using the condition $p_1,p_2,p_3\ge 2$)

$$|\xi^{h_1,h_2,h_3}(t)|_{L^2} \le C|h_1|_{L^{p_1}}|h_2|_{L^{p_2}}|h_3|_{L^{p_3}},$$

and gathering the estimates then concludes the proof of (41). □

For the analysis of the averaging error, in the weak sense, the fundamental object is the auxiliary function v defined by

$$v(t,x,y)=\Phi\big(x,y,D_x\overline{u}(t,x)\big), \tag{42}$$

where the first order derivative $D_x\overline{u}(t,x)$ is interpreted as an element of L^2.

By construction, $v(t,x,\cdot)$ is the solution of the Poisson equation (31) with $\theta=D_x\overline{u}(t,x)$, *i.e.* one has the fundamental identity

$$-\mathbf{L}v(t,x,y)=\langle F(x,y)-\overline{F}(x),D_x\overline{u}(t,x)\rangle. \tag{43}$$

For all $y\in L^2$, denote by $\mathcal{L}_y$ the infinitesimal generator given by

$$\mathcal{L}_y\varphi(x)=\langle D_x\varphi(x),Ax+F(x,y)\rangle+\frac{1}{2}\sum_{n\in\mathbb{N}}q_nD_x^2\varphi(x,y).\big(f_n,f_n\big),$$

for functions $\varphi: x\in L^2\mapsto\varphi(x)\in\mathbb{R}$, depending only on the slow variable x.

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

Applying the Itô formula, the weak error is written as

$$\begin{aligned}\mathbb{E}[\varphi(X^\epsilon(T))]-\mathbb{E}[\varphi(\overline{X}(T))]&=\mathbb{E}[\overline{u}(0,X^\epsilon(T))]-\mathbb{E}[\overline{u}(T,X^\epsilon(0))]\\&=\int_0^T\mathbb{E}\big[-\partial_t\overline{u}(T-t,X^\epsilon(t))+\mathcal{L}_{Y^\epsilon(t)}\overline{u}(T-t,X^\epsilon(t))\big]dt\\&=\int_0^T\mathbb{E}\big[(\mathcal{L}_{Y^\epsilon(t)}-\overline{\mathcal{L}})\overline{u}(T-t,X^\epsilon(t))\big]dt\\&=\int_0^T\mathbb{E}\Big[\,\langle F(X^\epsilon(t),Y^\epsilon(t))\\&\quad-\overline{F}(X^\epsilon(t)),D_x\overline{u}(T-t,X^\epsilon(t))\rangle\,\Big]\,dt\\&=\int_0^T\mathbb{E}\big[-\mathbf{L}v(T-t,X^\epsilon(t),Y^\epsilon(t))\big]dt,\end{aligned}$$

thanks to the identity (43). To exploit this formula for the weak error, note that the Itô formula applied with the function v yields the identity

$$\begin{aligned}\mathbb{E}[v(0,X^\epsilon(T),Y^\epsilon(T))]&=\mathbb{E}[v(T,X^\epsilon(0),Y^\epsilon(0))]\\&\quad+\int_0^T\mathbb{E}\Big[\Big(\mathcal{L}_{Y^\epsilon(t)}+\frac{1}{\epsilon}\mathbf{L}-\partial_t\Big)v(T-t,X^\epsilon(t),Y^\epsilon(t))\Big]dt.\end{aligned}$$

As a consequence, the weak error has the following decomposition

$$\mathbb{E}[\varphi(X^\epsilon(T))]-\mathbb{E}[\varphi(\overline{X}(T))]=\mathcal{J}_1^\epsilon+\mathcal{J}_2^\epsilon+\mathcal{J}_3^\epsilon,\tag{44}$$

where

$$\begin{aligned}\mathcal{J}_1^\epsilon&=\epsilon\big(\mathbb{E}[v(T,X^\epsilon(0),Y^\epsilon(0))]-\mathbb{E}[v(0,X^\epsilon(T),Y^\epsilon(T))]\big)\\\mathcal{J}_2^\epsilon&=-\epsilon\int_0^T\mathbb{E}\big[\partial_t v(T-t,X^\epsilon(t),Y^\epsilon(t))\big]dt\\\mathcal{J}_3^\epsilon&=\epsilon\int_0^T\mathbb{E}\big[\mathcal{L}_{Y^\epsilon(t)}v(T-t,X^\epsilon(t),Y^\epsilon(t))\big]dt,\end{aligned}$$

and the third expression is decomposed as $\mathcal{J}_3^\epsilon=\mathcal{J}_{3,1}^\epsilon+\mathcal{J}_{3,2}^\epsilon+\mathcal{J}_{3,3}^\epsilon$, where

$$\begin{aligned}\mathcal{J}_{3,1}^\epsilon&=\epsilon\int_0^T\mathbb{E}\big[\langle F(X^\epsilon(t),Y^\epsilon(t)),D_xv(T-t,X^\epsilon(t),Y^\epsilon(t))\rangle\big]dt\\\mathcal{J}_{3,2}^\epsilon&=\epsilon\int_0^T\mathbb{E}\big[\langle AX^\epsilon(t),D_xv(T-t,X^\epsilon(t),Y^\epsilon(t))\rangle\big]dt\\\mathcal{J}_{3,3}^\epsilon&=\frac{\epsilon}{2}\int_0^T\mathbb{E}\big[\sum_{n\in\mathbb{N}}q_nD_x^2v(T-t,X^\epsilon(t),Y^\epsilon(t)).\big(f_n,f_n\big)\big]dt.\end{aligned}$$

Theorem 4.7 is then a straightforward consequence of the three auxiliary results stated below.

Lemma 7.2. *There exists $C(T)\in(0,\infty)$, such that, for all $\epsilon\in(0,1)$, and all $x_0,y_0\in H$,*

$$|\mathcal{J}_1^\epsilon|\le C(T)\epsilon(1+|y_0|_{L^2}).$$

Lemma 7.3. *Let $\kappa\in(0,\gamma_{\max})$. There exists $C_\kappa(T)\in(0,\infty)$, such that, for all $\epsilon\in(0,1)$, and all $x_0,y_0\in H$,*

$$|\mathcal{J}_2^\epsilon|\le C_\kappa(T)\epsilon(1+\mathrm{Tr}(Q))\big(1+|(-A)^{2\kappa}x_0|_{L^4}^2+|(-A)^{2\kappa}y_0|_{L^4}^2\big).$$

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

Lemma 7.4. *Let $\gamma \in (1-\alpha_{\max}, \gamma_{\max})$ and $\kappa \in (0, \gamma_{\max}-\gamma)$. There exists $C_{\gamma,\kappa}(T) \in (0,\infty)$ such that, for all $\epsilon \in (0,1)$, and all $x_0, y_0 \in L^8$,*

$$|\mathcal{J}_3^\epsilon| \le C_{\gamma,\kappa}(T)\epsilon\big(1+|(-A)^{\gamma+\kappa}x_0|_{L^8}^2+|(-A)^{\gamma+\kappa}y_0|_{L^8}^2\big) \\ \times \big(1+|(-A)^{1-\gamma}x_0|_{L^4}+\mathrm{Tr}(Q)+M_{\alpha,4}(Q^{\frac{1}{2}},T)\big).$$

Note that Assumption 4.1 is only required in Lemma 7.4.

Proof of Lemma 7.2. Thanks to Lemma 5.1 and Proposition 7.1, for all $t \in [0,T]$, $x, y \in L^2$,

$$|v(t,x,y)| = |\Phi(x,y,D_x\overline{u}(t,x))| \le C(1+|y|)|D_x\overline{u}(t,x)|_{L^2} \le C(T,\varphi)(1+|y|_{L^2}).$$

Combined with Assumption 2.7, this estimate concludes the proof of Lemma 7.2. □

Proof of Lemma 7.3. Since the mapping $\theta \in H \mapsto \Phi(x,y,\theta)$ is a continuous linear mapping (thanks to Lemma 5.1), one has the following expression (justified below)

$$\partial_t v(t,x,y) = \Phi\big(x,y,\partial_t D_x\overline{u}(t,x)\big) \\ = \Phi\big(x,y,\Theta_1(t,x)\big)+\Phi\big(x,y,\Theta_2(t,x)\big)+\Phi\big(x,y,\Theta_3(t,x)\big), \tag{45}$$

where

$$\langle\Theta_1(t,x),h\rangle = \langle Ah+D\overline{F}(x).h, D_x\overline{u}(t,x)\rangle, \tag{46}$$

$$\langle\Theta_2(t,x),h\rangle = D_x^2\overline{u}(t,x).(h, Ax+\overline{F}(x)), \tag{47}$$

$$\langle\Theta_3(t,x),h\rangle = \frac{1}{2}\sum_{n\in\mathbb{N}} q_n D_x^3\overline{u}(t,x).(f_n,f_n,h). \tag{48}$$

Indeed, the quantity $\overline{U}(t,x,h) = D_x\overline{u}(t,x).h$ can be expressed as

$$\overline{U}(t,x,h) = \mathbb{E}\big[\langle D_x\varphi(\overline{X}^x(s)), \eta^{x,h}(s)\rangle\big],$$

with $d\eta^{x,h}(t) = A\eta^{x,h}(t)dt + D\overline{F}(\overline{X}^x(t)).\eta^{x,h}(t)dt$ and $\eta^{x,h}(0) = h$ (where dependence with respect to x is indicated), see the proof of Proposition 7.1. Consider the auxiliary Markov process defined by $\tilde{X}^{x,h}(t) = \big(\overline{X}^x(t), \eta^{x,h}(t)\big)$, with generator given by

$$\tilde{\mathcal{L}}\phi(x,h) = \overline{\mathcal{L}}\phi(x,h) + \langle Ah + D\overline{F}(x).h, D_h\phi(x,h)\rangle.$$

Then applying the Itô formula yields the identity

$$\partial_t\big(D_x\overline{u}(t,x).h\big) = \partial_t\overline{U}(t,x,h) \\ = \tilde{\mathcal{L}}\overline{U}(t,x,h) \\ = \langle Ah+D\overline{F}(x).h, D_h\overline{U}(t,x,h)\rangle \\ + D_x^2\overline{u}(t,x).(h,Ax+\overline{F}(x)) + \frac{1}{2}\sum_{n\in\mathbb{N}} q_n D_x^3\overline{u}(t,x).(f_n,f_n,h),$$

and it remains to observe that $D_h\overline{U}(t,x,h) = D_x\overline{u}(t,x)$ to conclude the derivation of (45).

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

Let $\kappa \in (0,1)$. Thanks to (39), one has

$$|\langle \Theta_1(t,x),h\rangle| \le C_\kappa(T)t^{-1+\kappa}|(-A)^\kappa h|_{L^2},$$

which implies $|(-A)^{-\kappa}\Theta_1(t,x)|_{L^2} \le C_\kappa(T)t^{-1+\kappa}$. Thus, thanks to Lemma 5.2,

$$\big|\Phi\big(x,y,\Theta_1(t,x)\big)\big| \le C_\kappa(T)t^{-1+\kappa}\big(1+|(-A)^{2\kappa}x|_{L^4}^2+|(-A)^{2\kappa}y|_{L^4}^2\big).$$

Thanks to (40), one has

$$|\Theta_2(t,x)|_{L^2} = \sup_{h\in L^2, |h|_{L^2}\le 1} |\langle \Theta_2(t,x),h\rangle| \le C_\kappa(T)t^{-1+\kappa}(1+|(-A)^\kappa x|_{L^2}).$$

Thus, thanks to Lemma 5.1,

$$\big|\Phi\big(x,y,\Theta_2(t,x)\big)\big| \le C_\kappa(T)t^{-1+\kappa}(1+|(-A)^\kappa x|_{L^2}^2+|y|_{L^2}^2).$$

Finally, thanks to (41) and Assumption 2.6, one has, for all $h\in L^2$,

$$|\langle \Theta_3(t,x),h\rangle| \le C(T)\sum_{n\in\mathbb{N}} q_n|f_n|_{L^4}^2|h|_{L^2} \le C(T)\mathrm{Tr}(Q)|h|_{L^2},$$

i.e. $|\Theta_3(t,x)|_{L^2} = \sup_{h\in L^2,|h|_{L^2}\le 1} |\langle \Theta_3(t,x),h\rangle| \le C(T)\mathrm{Tr}(Q)$. Thus Lemma 5.1 yields

$$\big|\Phi\big(x,y,\Theta_3(t,x)\big)\big| \le C(T)\mathrm{Tr}(Q)(1+|y|_{L^2}).$$

Gathering the above estimates then yields, if $2\kappa < \gamma_{\max}$,

$$\begin{aligned}|\mathcal{J}_2^\epsilon| &\le C_\kappa(T)\epsilon\int_0^T (1+t^{-1+\kappa})\mathbb{E}\big(1+|(-A)^{2\kappa}X^\epsilon(t)|_{L^4}^2+|(-A)^{2\kappa}Y^\epsilon(t)|_{L^4}^2\big)dt\\ &\le C_\kappa(T)\epsilon T^\kappa(1+\mathrm{Tr}(Q))\big(1+|(-A)^{2\kappa}x_0|_{L^4}^2+|(-A)^{2\kappa}y_0|_{L^4}^2\big).\end{aligned}$$

This concludes the proof of Lemma 7.3. □

Proof of Lemma 7.4. Note that the first-order derivative of v with respect to x satisfies the following identity:

$$\langle D_x v(t,x,y),h\rangle = \Phi\big(x,y,D_x^2\overline{u}(t,x).(h,\cdot)\big) + \langle D_x\Phi\big(x,y,D_x\overline{u}(t,x)\big),h\rangle.$$

Observe that $|D_x^2\overline{u}(t,x).(h,\cdot)|_{L^2} \le C|h|_{L^2}$, thanks to (40), with $\beta_1=\beta_2=0$. Then, thanks to Lemmas 5.1 and 5.3, one obtains

$$\begin{aligned}|\langle D_x v(t,x,y),h\rangle| &\le C(1+|y|_{L^2})|D_x^2\overline{u}(t,x).(h,\cdot)|_{L^2} + C(1+|y|_{L^4})|D_x\overline{u}(t,x)|_{L^2}|h|_{L^4}\\ &\le C(1+|y|_{L^4})|h|_{L^4}.\end{aligned}$$

Since F has at most linear growth, using moment estimates then yields

$$\mathbb{E}|\mathcal{J}_{3,1}^\epsilon| \le C(T)\epsilon(1+|x_0|_{L^4}^2+|y_0|_{L^4}^2).$$

To treat the second term, $\mathcal{J}_{3,2}^\epsilon$, observe that $|D_x^2\overline{u}(t,x).(h,\cdot)|_{L^2} \le C_\kappa t^{-1+\kappa}|(-A)^{-1+\kappa}h|_{L^2}$, for all $\kappa\in(0,1]$, thanks to (40). In addition, $|(-A)^{1-\kappa}D_x\overline{u}(t,x)|_{L^2} \le C_\kappa t^{-1+\kappa}$, thanks to (39). Then, thanks to Lemmas 5.1 and 5.4,

$$\begin{aligned}|\langle D_x v(t,x,y),h\rangle| &\le C_\kappa(1+|y|_{L^2})t^{-1+\kappa}|(-A)^{-1+\kappa}h|_{L^2}\\ &\quad + C_{\gamma,\kappa}\big(1+|(-A)^{\gamma+\kappa}x|_{L^8}^2+|(-A)^{\gamma+\kappa}y|_{L^8}^2\big)t^{-\gamma-\frac{\kappa}{2}}|(-A)^{-\gamma}h|_{L^4},\end{aligned}$$

where $\gamma<\gamma_{\max}$ and $\kappa\in(0,\gamma_{\max}-\gamma)$.

As a consequence

$$|\mathcal{J}_{3,2}^{\epsilon}| \leq C_\kappa \epsilon \int_0^T \big(1+\mathbb{E}|Y^\epsilon(t)|_{L^2}^2\big)^{\frac{1}{2}} \big(\mathbb{E}|(-A)^\kappa X^\epsilon(t)|_{L^2}^2\big)^{\frac{1}{2}} (T-t)^{-1+\kappa} dt$$
$$+ C_{\gamma,\kappa}\epsilon \int_0^T \big(1+\mathbb{E}|(-A)^{\gamma+\kappa} X^\epsilon(t)|_{L^8}^4 + |(-A)^{\gamma+\kappa} Y^\epsilon(t)|_{L^8}^4\big)^{\frac{1}{2}}$$
$$\times \big(\mathbb{E}|(-A)^{1-\gamma} X^\epsilon(t)|_{L^4}^2\big)^{\frac{1}{2}} (T-t)^{-\gamma-\frac{\kappa}{2}} dt.$$

It remains to use the condition that $\gamma \in (1-\alpha_{\max}, \gamma_{\max})$, thanks to Assumption 4.1. Note that $\gamma + \kappa \leq \gamma_{\max} \leq \frac{1}{2} \leq 1-\gamma$. Finally, thanks to moment estimates,

$$|\mathcal{J}_{3,2}^{\epsilon}| \leq C_{\gamma,\kappa}\epsilon\big(1+|(-A)^{\gamma+\kappa}x_0|_{L^8}^2+|(-A)^{\gamma+\kappa}y_0|_{L^8}^2\big)\big(1+|(-A)^{1-\gamma}x_0|_{L^4}+M_{1-\gamma,4}(Q^{\frac{1}{2}},T)\big).$$

It remains to deal with the third term, $\mathcal{J}_{3,3}^{\epsilon}$. Note that the second-order derivative of v with respect to x satisfies the identity

$$D_x^2 v(t,x,y).(h,h) = \Phi\big(x,y,D_x^3\overline{u}(t,x).(h,h,\cdot)\big) + 2\langle D_x\Phi\big(x,y,D_x^2\overline{u}(t,x).(h,\cdot)\big),h\rangle$$
$$+ D_x^2\Phi\big(x,y,D_x\overline{u}(t,x)\big).(h,h).$$

First, observe that $|D_x^3\overline{u}(t,x).(h,h,k)| \leq C|h|_{L^4}^2|k|_{L^2}$, thanks to (41). Equivalently, this means that $|D_x^3\overline{u}(t,x).(h,h,\cdot)|_{L^2} \leq C|h|_{L^4}^2$, then Lemma 5.1 yields

$$\big|\Phi\big(x,y,D_x^3\overline{u}(t,x).(h,h,\cdot)\big)\big| \leq C(1+|y|_{L^2})|h|_{L^4}^2.$$

Second, thanks to Lemma 5.3,

$$\big|\langle D_x\Phi\big(x,y,D_x^2\overline{u}(t,x).(h,\cdot)\big),h\rangle\big| \leq C(1+|y|_{L^4})|h|_{L^4}|D_x^2\overline{u}(t,x).(h,\cdot)|_{L^2}$$
$$\leq C(1+|y|_{L^4})|h|_{L^4}^2.$$

Finally, Lemma 5.5 and (39) yield

$$\big|D_x^2\Phi\big(x,y,D_x\overline{u}(t,x)\big).(h,h)\big| \leq C(1+|y|_{L^4})|D_x\overline{u}(t,x)|_{L^2}|h|_{L^8}^2 \leq C(1+|y|_{L^4})|h|_{L^8}^2.$$

As a consequence, one obtains

$$|\mathcal{J}_{3,3}^{\epsilon}| = \Big|\frac{\epsilon}{2}\int_0^T \sum_{n\in\mathbb{N}} q_n \mathbb{E}\big[D_x^2 v(T-t,X^\epsilon(t),Y^\epsilon(t)).(f_n,f_n)\big]dt\Big|$$
$$\leq C\epsilon \sum_{n\in\mathbb{N}} q_n|f_n|_{L^8}^2 \int_0^T \big(1+\mathbb{E}|Y^\epsilon(t)|_{L^4}\big)dt$$
$$\leq C(T)\mathrm{Tr}(Q)\epsilon(1+|y_0|_{L^4}),$$

thanks to Assumption 2.6, and a moment estimate, see Assumption 2.7.

Gathering the estimates for $\mathcal{J}_{3,1}^{\epsilon}$, $\mathcal{J}_{3,2}^{\epsilon}$ and $\mathcal{J}_{3,3}^{\epsilon}$, one obtains

$$|\mathcal{J}_{3}^{\epsilon}| \leq C_{\gamma,\kappa}\epsilon\big(1+|(-A)^{1-\gamma}x_0|_{L^8}^4+|(-A)^{\gamma+\kappa}y_0|_{L^8}^4\big)$$
$$\times \big(1+|(-A)^{1-\gamma}x_0|_{L^4}+\mathrm{Tr}(Q)+M_{1-\gamma,4}(Q^{\frac{1}{2}},T)\big).$$

This concludes the proof of Lemma 7.4. □

We are now in position to conclude the proof of Theorem 4.7.

Proof. Thanks to the decomposition (44) of the weak error, it suffices to gather the estimates of Lemmas 7.2, 7.3 and 7.4 to conclude. □

8. Proof of Theorem 4.8

This section is devoted to the proof of Theorem 4.8. Let Assumption 4.2 be satisfied. First, let us justify the definition of

$$\alpha_{\max} = \frac{1}{2}\big(1 - \frac{d}{2}(1 - \frac{2}{\varrho})\big).$$

Proposition 8.1. *Let Assumption 4.2 be satisfied. Then (9) is satisfied: for all $\alpha \in [0, \alpha_{\max})$, all $T \in (0, \infty)$, and all $p \geq 2$,*

$$M_{\alpha,p}(Q^{\frac{1}{2}}, T) < \infty.$$

Proof of Proposition 8.1. Let $\varsigma = \frac{\varrho}{\varrho-2} > 1$, and note that $1 = \frac{2}{\varrho} + \frac{1}{\varsigma}$.

Using the ideal property for γ-Radonifying operators and the regularizing properties (3) of the semigroup,

$$\begin{aligned}\int_0^T \|e^{tA}(-A)^\alpha Q^{\frac{1}{2}}\|^2_{\mathcal{R}(L^2,L^p)} dt &\leq \int_0^T \|e^{\frac{t}{2}A}(-A)^\alpha\|^2_{\mathcal{L}(L^p,L^p)} \|e^{\frac{t}{2}A} Q^{\frac{1}{2}}\|^2_{\mathcal{R}(L^2,L^p)} dt \\ &\leq C_{\alpha,p} \int_0^T t^{-2\alpha} \Big| \sum_{n\in\mathbb{N}} q_n (e^{tA} f_n)^2 \Big|_{L^{\frac{p}{2}}} dt.\end{aligned}$$

Using the Hölder inequality, for all $\xi \in \mathcal{D}$, and all $t > 0$,

$$\begin{aligned}\sum_{n\in\mathbb{N}} q_n \big(e^{tA} f_n\big)^2(\xi) &\leq \big(\sum_{n\in\mathbb{N}} q_n^{\frac{\varrho}{2}}\big)^{\frac{2}{\varrho}} \Big(\sum_{n\in\mathbb{N}} \big(e^{tA} f_n\big)^{2\varsigma}(\xi)\Big)^{\frac{1}{\varsigma}} \\ &\leq C(Q)\Big(\sup_{k\in\mathbb{N}} \big(e^{tA} f_k\big)^{\frac{2(\varsigma-1)}{\varsigma}}(\xi)\Big)\Big(\sum_{n\in\mathbb{N}} \big(e^{tA} f_n\big)^2(\xi)\Big)^{\frac{1}{\varsigma}}.\end{aligned}$$

On the one hand, Assumption 2.6 and the property (5) of the kernel K imply that for all $\xi \in \mathcal{D}$,

$$\sup_{k\in\mathbb{N}} |e^{tA} f_k(\xi)| \leq \int_{\mathcal{D}} K(t, \xi, \cdot) \sup_{k\in\mathbb{N}} |f_k|_{L^\infty} \leq C.$$

On the other hand, $\big(f_n\big)_{n\in\mathbb{N}}$ is a complete orthonormal system of L^2, hence

$$\begin{aligned}\sum_{n\in\mathbb{N}} \big(e^{tA} f_n\big)^2(\xi) &= \sum_{n\in\mathbb{N}} \langle K(t, \xi, \cdot), f_n\rangle^2 = |K(t, \xi, \cdot)|^2_{L^2} \\ &= \int_{\mathcal{D}} K(t, \xi, \eta)^2 d\eta \\ &\leq C t^{-\frac{d}{2}} \int_{\mathcal{D}} K(t, \xi, \eta) d\eta = C t^{-\frac{d}{2}},\end{aligned}$$

using the properties (5) of the kernel K.

Finally, for all $t > 0$ and all $z \in \mathcal{D}$, one obtains

$$\Big| \sum_{n\in\mathbb{N}} q_n (e^{tA} f_n)^2 \Big|_{L^{\frac{p}{2}}} \leq C t^{-\frac{d}{2\varsigma}},$$

thus

$$\int_0^T \|e^{tA}(-A)^\alpha Q^{\frac{1}{2}}\|^2_{\mathcal{R}(L^2,L^p)} dt \leq C \int_0^T t^{-2\alpha - \frac{d}{2\varsigma}} dt.$$

It remains to check that $2\alpha - \frac{d}{2\varsigma} = 2\alpha - \frac{d}{2}(1-\frac{2}{\varrho}) < 1$ for $\alpha < \alpha_{\max} = \frac{1}{2}\big(1 - \frac{d}{2}(1-\frac{2}{\varrho})\big)$.

This concludes the proof of Proposition 8.1. □

The approximation argument is based on the following estimate.

Lemma 8.2. *Let Assumption 4.2 be satisfied. For all $\alpha \in [0, \alpha_{\max})$, $\gamma \in [0, \gamma_{\max})$, such that $\alpha \geq \gamma$ and $\alpha + \gamma \leq 1$, all $T \in (0, \infty)$ and $p \geq 2$, there exists $C_{\alpha,\gamma,p}(Q, T) \in (0, \infty)$, such that for all $\delta \in (0, 1)$,*

$$\mathrm{Tr}(e^{2\delta A} Q) + M_{1-\gamma,p}(e^{\delta A} Q^{\frac{1}{2}}, T) \leq C_{\alpha,\gamma,p}(Q, T)\delta^{\alpha+\gamma-1}. \tag{49}$$

Proof of Lemma 8.2. First, note that

$$M_{1-\gamma,p}(e^{\delta A} Q^{\frac{1}{2}}, T) \leq \|(-A)^{1-\gamma-\alpha} e^{\delta A}\|_{\mathcal{L}(L^p,L^p)} M_{\alpha,p}(Q^{\frac{1}{2}}, T),$$

and that $\|(-A)^{1-\gamma-\alpha} e^{\delta A}\|_{\mathcal{L}(L^p,L^p)} \leq C_{\alpha,\gamma}\delta^{\gamma+\alpha-1}$, in the regime $\alpha + \gamma \leq 1$.

To deal with the trace term, we use the Hölder type inequality for Schatten norms $\|\cdot\|_{\mathcal{L}_\varrho(L^2)}$, with parameter $\varrho \in [1, \infty]$, see for instance [24, Corollary D.2.4, Appendix D]. One obtains

$$\mathrm{Tr}(e^{2\delta A} Q) = \|e^{2\delta A} Q\|_{\mathcal{L}_1(L^2)} \leq \|e^{2\delta A}\|_{\mathcal{L}_\varsigma(L^2)} \|Q\|_{\mathcal{L}_{\frac{\varrho}{2}}(L^2)},$$

where $1 = \frac{2}{\varrho} + \frac{1}{\varsigma}$. By assumption, $\|Q\|_{\mathcal{L}_{\frac{\varrho}{2}}(L^2)} < \infty$. In addition,

$$\|e^{2\delta A}\|^{\varsigma}_{\mathcal{L}_\varsigma(L^2)} \leq \sum_{n\in\mathbb{N}} |e^{2\delta A} e_n|^{\varsigma}_{L^2} \leq \sum_{n\in\mathbb{N}} e^{-2\varsigma\delta\lambda_n} \leq C_\varsigma \delta^{-\frac{d}{2}},$$

using $\lambda_n \sim_{n\to\infty} c_d n^{\frac{2}{d}}$. Indeed, assume that $\delta \in [\frac{1}{N^{2/d}}, \frac{1}{(N-1)^{2/d}}]$, with $N \in \mathbb{N}$, $N \geq 2$. Then, for some $c > 0$, a comparison argument between a Riemann sum and an integral yields

$$\begin{aligned}
\sum_{n\in\mathbb{N}} e^{-2\varsigma\delta\lambda_n} &\leq \sum_{n\in\mathbb{N}} e^{-c\frac{n^{2/d}}{N^{2/d}}} = N \times \frac{1}{N}\sum_{n\in\mathbb{N}} e^{-c\frac{n^{2/d}}{N^{2/d}}} \\
&\leq N \int_0^{+\infty} e^{-cx^{2/d}} dx \leq C\frac{1}{\delta^{d/2}},
\end{aligned}$$

where in the last inequality the condition $N - 1 \leq \delta^{-\frac{d}{2}}$ is used. Since N is arbitrary, this concludes the argument. As a consequence,

$$\mathrm{Tr}(e^{2\delta A} Q) \leq C_\varsigma \delta^{-\frac{d}{2\varsigma}} = C_\varsigma \delta^{2\alpha_{\max}-1} \leq C_\varsigma \delta^{2\alpha-1},$$

using the definition of $\alpha_{\max} = \frac{1}{2}\big(1 - \frac{d}{2}(1-\frac{2}{\varrho})\big) = \frac{1}{2}\big(1 - \frac{d}{2\varsigma}\big)$.

Finally, one concludes using $2\alpha - 1 \leq \alpha + \gamma - 1 \leq 0$. □

The result of Lemma 8.2 motivates the introduction of the following auxiliary SPDE problems, where $Q^{\frac{1}{2}}$ is replaced by $e^{\delta A} Q^{\frac{1}{2}}$. For all $\delta \in (0, 1)$ (this parameter will be chosen below), X^ϵ_δ and $\overline{X}_\delta$ are solutions of

$$\begin{aligned}
dX^\epsilon_\delta(t) &= AX^\epsilon_\delta(t)dt + F\big(X^\epsilon_\delta(t), Y^\epsilon(t)\big)dt + e^{\delta A} dW^Q(t), \\
d\overline{X}_\delta(t) &= A\overline{X}_\delta(t)dt + \overline{F}(\overline{X}(t))dt + e^{\delta A} dW^Q(t),
\end{aligned} \tag{50}$$

with initial conditions $X^\epsilon_\delta(0) = \overline{X}_\delta = x_0$.

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

Then Theorem 4.8 follows from Lemmas 8.3 and 8.4 stated below.

First, thanks to Lemma 8.2, the strong and weak convergence results, with orders $\frac{1}{2}$ and 1, from Theorem 4.5 and 4.7 may be applied when considering the auxiliary processes X_δ^ϵ and $\overline{X}_\delta$ defined by (50).

Lemma 8.3. *Let Assumption 4.2 be satisfied. Let $T \in (0,\infty)$, and assume that the initial conditions x_0, y_0, satisfy*

$$|(-A)^{1-\gamma}x_0|_{L^8} + |(-A)^{\gamma+\kappa}y_0|_{L^8} < \infty,$$

with $\gamma \in (0,\gamma_{\max})$ and $\kappa \in (0,\gamma_{\max}-\gamma)$.

Let φ be a nice test function.

For all $\alpha \in (0,\alpha_{\max})$, there exist $C_{\alpha,\gamma}(T,x_0,y_0,Q) \in (0,\infty)$ and $C_{\alpha,\gamma}(T,x_0,y_0,Q,\varphi) \in (0,\infty)$, such that for all $\epsilon \in (0,1)$ and $\delta \in (0,1)$,

$$\sup_{t\in[0,T]} \big(\mathbb{E}|X_\delta^\epsilon(t) - \overline{X}_\delta(t)|_{L^2}^2\big)^{\frac{1}{2}} \le C_{\alpha,\gamma}(T,x_0,y_0,Q)\delta^{-\frac{1-\alpha-\gamma}{2}}\epsilon^{\frac{1}{2}}$$

and

$$\sup_{t\in[0,T]} |\mathbb{E}[\varphi(X_\delta^\epsilon(t))] - \mathbb{E}[\varphi(\overline{X}_\delta(t))]| \le C_{\alpha,\gamma}(T,x_0,y_0,Q,\varphi)\delta^{-(1-\alpha-\gamma)}\epsilon.$$

Second, the distances between X_δ^ϵ and X^ϵ, and between $\overline{X}_\delta$ and $\overline{X}$, are estimated in the following result, using standard arguments.

Lemma 8.4. *Let Assumption 4.2 be satisfied. Let $T \in (0,\infty)$, and assume that $x_0 \in L^2$ and $y_0 \in L^2$. Let φ be a nice test function. Let $\alpha \in [0,\alpha_{\max})$. There exist $C_\alpha(T,x_0,y_0,Q) \in (0,\infty)$ and $C_\alpha(T,x_0,y_0,Q,\varphi) \in (0,\infty)$ such that for all $\epsilon \in (0,1)$ and $\delta \in (0,1)$, one has*

$$\sup_{t\in[0,T]} \big(\mathbb{E}|X_\delta^\epsilon(t) - X^\epsilon(t)|_{L^2}^2\big)^{\frac{1}{2}} + \sup_{t\in[0,T]} \big(\mathbb{E}|\overline{X}_\delta(t) - \overline{X}(t)|_{L^2}^2\big)^{\frac{1}{2}} \le C_\alpha(T,x_0,y_0,Q)\delta^\alpha.$$

and

$$\sup_{t\in[0,T]} |\mathbb{E}[\varphi(X_\delta^\epsilon(t))] - \mathbb{E}[\varphi(X^\epsilon(t))]| + \sup_{t\in[0,T]} |\mathbb{E}[\varphi(\overline{X}_\delta(t))] - \mathbb{E}[\varphi(\overline{X}(t))]| \\ \le C_\alpha(T,x_0,y_0,Q,\varphi)\delta^{2\alpha}.$$

Proof of Lemma 8.3. This is a straightforward application of Theorems 4.5 and 4.7 combined with Lemma 8.2. □

Proof of Lemma 8.4. Consider first the estimates of the strong error. Since the nonlinear operators F and $\overline{F}$ are globally Lipschitz continuous, it is sufficient to prove the following estimate:

$$\begin{aligned}\mathbb{E}\Big|\int_0^t e^{(t-s)A}\big(e^{\delta A}-I\big)dW^Q(s)\Big|_{L^2}^2 ds &= \int_0^t \|e^{sA}\big(e^{\delta A}-I\big)Q^{\frac{1}{2}}\|_{\mathcal{R}(L^2,L^2)}^2 ds \\ &\le \|(e^{\delta A}-I)(-A)^{-\alpha}\|_{\mathcal{L}(L^2,L^2)}^2 \\ &\quad \times \int_0^t \|e^{sA}(-A)^\alpha Q^{\frac{1}{2}}\|_{\mathcal{R}(L^2,L^2)}^2 ds \\ &\le C_\alpha \delta^{2\alpha} M_{\alpha,2}(Q,T)^2,\end{aligned}$$

and the strong error estimates are straightforward consequences of the Gronwall Lemma.

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

It remains to prove the estimates of the weak error. Since the argument is the same for both estimates, we only deal with the second one. Note that

$$\mathbb{E}[\varphi(\overline{X}_\delta(t))] - \mathbb{E}[\varphi(\overline{X}(t))] = \mathbb{E}[\overline{u}(0, \overline{X}_\delta(t))] - \mathbb{E}[\overline{u}(t, \overline{X}_\delta(0))],$$

where $\overline{u}$ is defined by the expression (38). Observe that, even if Assumption 4.2 is satisfied instead of Assumption 4.1, the regularity estimates on spatial derivatives of $\overline{u}$ stated in Proposition 7.1 remain valid without modification.

Using Itô formula, one obtains

$$\begin{aligned}
&\mathbb{E}[\varphi(\overline{X}_\delta(t))] - \mathbb{E}[\varphi(\overline{X}(t))] \\
&\quad = \mathbb{E}\int_0^t \sum_{n\in\mathbb{N}} q_n \Big(D^2\overline{u}(t-s, \overline{X}_\delta(s)).(e^{\delta A} f_n, e^{\delta A} f_n) \\
&\qquad - D^2\overline{u}(t-s, \overline{X}_\delta(s)).(f_n, f_n) \Big) ds \\
&\quad = \mathbb{E}\int_0^t \Big(\mathrm{Tr}\big(D^2\overline{u}(t-s, \overline{X}_\delta(s))e^{\delta A} Q e^{\delta A}\big) - \mathrm{Tr}\big(D^2\overline{u}(t-s, \overline{X}_\delta(s))Q\big)\Big) ds \\
&\quad = \mathbb{E}\int_0^t \mathrm{Tr}\big(D^2\overline{u}(t-s, \overline{X}_\delta(s))\big(e^{\delta A} - I\big) Q e^{\delta A}\big) ds \\
&\qquad + \mathbb{E}\int_0^t \mathrm{Tr}\big(D^2\overline{u}(t-s, \overline{X}_\delta(s)) Q \big(e^{\delta A} - I\big)\big) ds,
\end{aligned}$$

where $D^2\overline{u}(t, x)$ is interpreted as a bounded, self-adjoint, linear operator from L^2 to L^2, instead of a symmetric, bilinear form on L^2, using Riesz Theorem: for all $h \in L^2$, $D^2\overline{u}(t, x).h \in L^2$ is characterized by

$$\langle D^2\overline{u}(t, x)h, \cdot\rangle = D^2\overline{u}(t, x).(h, \cdot).$$

Let $\alpha \in (0, \alpha_{\max})$ and $\kappa \in (0, \alpha_{\max} - \alpha)$. Then, using the Hölder type inequality for Schatten norms, for all $0 \le s < t \le T$,

$$\begin{aligned}
\Big|\mathrm{Tr}\big(D^2\overline{u}(t-s, \overline{X}_\delta(s))\big(e^{\delta A} - I\big) Q e^{\delta A}\big)\Big| &= \|D^2\overline{u}(t-s, \overline{X}_\delta(s))\big(e^{\delta A} - I\big) Q e^{\delta A}\|_{\mathcal{L}_1(L^2)} \\
&\le \|D^2\overline{u}(t-s, \overline{X}_\delta(s))(-A)^{1-2\kappa}\|_{\mathcal{L}_\infty(L^2)} \|(-A)^{-1+2\kappa}(I - e^{\delta A})\|_{\mathcal{L}_\varsigma(L^2)} \|Q\|_{\mathcal{L}_{\frac{\varrho}{2}}(L^2)},
\end{aligned}$$

where $1 = \frac{2}{\varrho} + \frac{1}{\varsigma}$. By assumption, one has $\|Q\|_{\mathcal{L}_{\frac{\varrho}{2}}} < \infty$. In addition, thanks to Proposition 7.1, one has

$$\begin{aligned}
\|D^2\overline{u}(t-s, \overline{X}_\delta(s))(-A)^{1-2\kappa}\|_{\mathcal{L}_\infty(L^2)} &= \|D^2\overline{u}(t-s, \overline{X}_\delta(s))(-A)^{1-2\kappa}\|_{\mathcal{L}(L^2, L^2)} \\
&\le C_\kappa (t-s)^{-1+2\kappa}.
\end{aligned}$$

Finally, $(-A)^{-1+2\kappa}(I - e^{\delta A})$ is a self-adjoint, compact, linear operator, thus, for $\alpha \le \frac{1}{2}$, using the standard inequality $0 \le 1 - e^{-x} \le 2^{1-\alpha}x^\alpha$ for all $x \in [0, \infty)$, with $x = \lambda_n\delta$, one has

$$\begin{aligned}
\|(-A)^{-1+2\kappa}(I - e^{\delta A})\|_{\mathcal{L}_\varsigma(L^2)}^\varsigma &= \sum_{n\in\mathbb{N}} \lambda_n^{-(1-2\kappa)\varsigma}(1 - e^{-\delta\lambda_n})^\varsigma \\
&\le C_\alpha \delta^{2\alpha\varsigma} \sum_{n\in\mathbb{N}} \lambda_n^{-(1-2\kappa-2\alpha)\varsigma}.
\end{aligned}$$

Finally, with the condition $\alpha + \kappa < \alpha_{\max} = \frac{1}{2}(1 - \frac{d}{2\varsigma})$, one has $(1 - 2\kappa - 2\alpha)\varsigma > \frac{d}{2}$, thus (using $\lambda_n \sim c_d n^{\frac{2}{d}}$) one has $\sum_{n\in\mathbb{N}} \lambda_n^{-(1-2\kappa-2\alpha)\varsigma} < \infty$.

Finally, one obtains

$$\left|\mathrm{Tr}\Big(D^2\overline{u}(t-s,\overline{X}_\delta(s))\big(e^{\delta A}-I\big)Qe^{\delta A}\Big)\right| \le C_\alpha \delta^{2\alpha}(t-s)^{-1+2\kappa},$$

and similarly

$$\left|\mathrm{Tr}\Big(D^2\overline{u}(t-s,\overline{X}_\delta(s))Q\big(e^{\delta A}-I\big)\Big)\right| \le C_\alpha \delta^{2\alpha}(t-s)^{-1+2\kappa}.$$

It is then straightforward to conclude that

$$\left|\mathbb{E}[\varphi(\overline{X}_\delta(t))]-\mathbb{E}[\varphi(\overline{X}(t))]\right| \le C_\alpha \delta^{2\alpha}.$$

This concludes the proof of Lemma 8.4. □

We are now in position to provide the proof of Theorem 4.8, which consists in choosing δ in terms of ϵ to maximize the order of convergence.

Proof of Theorem 4.8. Thanks to the strong and weak error estimates from Lemmas 8.3 and 8.4, one obtains, for all $\epsilon \in (0,1)$ and $\delta \in (0,1)$,

$$\sup_{t\in[0,T]} \left(\mathbb{E}|X^\epsilon(t)-\overline{X}(t)|_{L^2}^2\right)^{\frac{1}{2}} \le C_{\alpha,\gamma}(T,x_0,y_0,Q)\Big(\delta^{-\frac{1-\alpha-\gamma}{2}}\epsilon^{\frac{1}{2}}+\delta^\alpha\Big),$$

$$\sup_{t\in[0,T]} |\mathbb{E}[\varphi(X^\epsilon(t))]-\mathbb{E}[\varphi(\overline{X}(t))]| \le C_{\alpha,\gamma}(T,x_0,y_0,Q,\varphi)\Big(\delta^{-(1-\alpha-\gamma)}\epsilon+\delta^{2\alpha}\Big).$$

It remains to choose the regularization parameter δ in terms of ϵ, in order to minimize the right-hand sides in the estimates above. Choosing

$$\delta = \epsilon^{\frac{1}{1+\alpha-\gamma}},$$

the two error terms are equal. Since $\alpha_{\max}-\alpha$ and $\gamma_{\max}-\gamma$ are arbitrarily small, the orders of convergence can be rewritten in the form given by Theorem 4.8, in terms of the parameter $\beta_{\max}$ and this concludes the proof. □

Remark 8.5. Let us replace Assumption 4.2 by the following condition: $\alpha_{\max} \in [0,1)$ is such that for all $\alpha \in [0,\alpha_{\max})$ and all $p \ge 2$, one has $\|(-A)^{\alpha-\frac{1}{2}}Q^{\frac{1}{2}}\|_{\mathcal{R}(L^2,L^p)} < \infty$. Then the results of this section can be generalized as follows, using similar techniques. Lemma 8.3 holds true, whereas Lemma 8.4 needs to be modified: the strong error remains bounded by $C_\alpha\delta^\alpha$, and the weak error is bounded by $C_\alpha\delta^{\min(1,2\alpha)}$. On the one hand, if $\alpha_{\max} \le \frac{1}{2}$, the situation is the same as in Theorem 4.8. On the other hand, if $\alpha_{\max} \ge \frac{1}{2}$, the strong and the weak rates one obtains using the approximation approach considered above, are $\frac{\alpha_{\max}}{1+\alpha_{\max}-\gamma_{\max}}$ and $\frac{1}{2-\alpha_{\max}-\gamma_{\max}}$ respectively. This statement and the approach are not satisfactory in this case since the weak order is not equal to twice the strong order anymore. Whether this issue can be fixed, and whether the rates of convergence given above are optimal, are left for future works.

9. Efficient numerical approximation of the slow component

The objective of this section is to describe a multiscale numerical scheme (for temporal discretization) for the approximation of the slow component X^ϵ, in the regime $\epsilon \to 0$. This task is challenging and crucial: indeed a direct discretization for the system $\big(X^\epsilon(t),Y^\epsilon(t)\big)_{t\ge 0}$ requires to choose a time step size h satisfying a condition $h = \mathrm{o}(\epsilon)$, since $Y^\epsilon(t) = Y(t/\epsilon)$. Such a condition is prohibitive, and to circumvent this issue one may benefit from the averaging principle: when $\epsilon \to 0$, it is more appropriate to approximate the solution $\overline{X}$ of the averaged equation, for which the time step size Δt is not constrained anymore by the value of ϵ. However, at this stage, in general the averaged coefficient $\overline{F}$ is not known. Using for instance the following interpretation of $\overline{F}(x)$,

$$\overline{F}(x) = \int F(x, y)d\mu(y) = \lim_{T\to\infty} \frac{1}{T}\int_0^T F(x, Y(t))dt,$$

it suffices to estimate the averaged coefficient $\overline{F}(x)$ using a numerical scheme for Y with time step size τ (or equivalently a numerical scheme for Y^ϵ with time step size $\delta = \tau\epsilon$).

The resulting scheme belongs to the class of Heterogeneous Multiscale Methods, see [4] and references therein. It consists of a macro-scheme combined with the use of a micro-scheme at each step. It is crucial to note that the two time step sizes Δt and τ may be chosen independently when using such strategies. The multiscale scheme is closely related to the averaging principle. Indeed, the analysis of the full error of the scheme requires as a preliminary step to estimate the error in the averaging principle, using the main results of this article given in Section 4. Finally, choosing some parameters of the scheme in an appropriate way is shown to reproduce an averaging principle at the discretized level.

The construction of the multiscale scheme is presented in Section 9.1. An analysis of the convergence of the scheme (using the results of Section 4) is provided in Section 9.2.

The setting is different from the earlier work [4]. On the one hand, it is simpler since the fast component does not depend on the slow component. On the other hand, contrary to [4], the slow component is driven by a Wiener process which changes orders of convergence of the macro scheme and complicates the analysis. In addition, regularity of the processes X and Y depends on the parameter $\alpha_{\max}$ and $\gamma_{\max}$.

9.1. Construction of the scheme

As explained above, the main parameters of the multiscale scheme are the macro-time step size $\Delta t > 0$ and the micro-time step size $\tau > 0$. In addition, two other integer parameters M and $M_a \in \{1, \dots, M\}$ are used, in order to estimate the averaged coefficient in terms of temporal averages along the micro-scheme, at each step of the macro-scheme.

In this section, to avoid cumbersome notation, precise regularity conditions, and dependence in error estimates, on the initial conditions x_0, y_0 are not indicated.

9.1.1. Micro-scheme

Let $\big(Y_m^\tau\big)_{m\in\mathbb{N}_0}$ be computed using a numerical scheme Φ^τ for the stochastic process $\big(Y(t)\big)_{t\geq 0}$. If this process is solution of a SPDE of the type $dY(t) = AY(t)dt + G(Y(t))dt + dw^q(t)$, this means that the scheme is of the form $Y_{m+1} = \Phi^\tau\big(Y_m, w^q((m+1)\tau) - w^q(m\tau)\big)$, where the mapping Φ^τ depends on the chosen scheme.

First, it is assumed that, for all (sufficiently small) values of $\tau > 0$, the discrete-time process $\big(Y_m^\tau\big)_{m\in\mathbb{N}_0}$ is an ergodic Markov chain. Let μ^τ denote its unique invariant probability distribution.

Even if the temporal discretization preserves the ergodicity of the process, in general μ^τ is not equal to μ. To estimate the error between μ^τ and μ, let $\vartheta_{\max} > 0$ and assume that the following estimate is satisfied: for all functions $\varphi : L^2 \to \mathbb{R}$ of class $\mathcal{C}^2$, with bounded first and second order derivatives, and for every $\vartheta \in (0, \vartheta_{\max})$, there exists $C_\vartheta(\varphi) \in (0, \infty)$ such that

$$|\int \varphi d\mu^\tau - \int \varphi d\mu| \leq C_\theta(\varphi)\tau^\vartheta. \tag{51}$$

For instance, assume that the fast component is solution of a SPDE of the type $dY(t) = AY(t)dt + G(Y(t))dt + dw^q(t)$, driven by a q-Wiener process. Then the parameter $\gamma_{\max}$ arises from assumptions on the covariance operator q, and one checks that this parameter describes the spatial and temporal regularity of the process Y: almost surely, the trajectories are γ-Hölder continuous, for all $\gamma \in (0, \gamma_{\max})$. In such situations, a standard numerical scheme has a strong order equal to $\gamma_{\max}$ and a weak order equal to $2\gamma_{\max}$, see for instance [6] or [29]. Thus in this case, it is legitimate to assume that $\vartheta = 2\gamma_{\max}$. See [5] for a full analysis of the approximation of the invariant distribution when $\gamma_{\max} = \frac{1}{4}$ and $d = 1$, *i.e.* when the fast process is driven by space–time white noise.

For a given time step size $\tau > 0$, introduce the averaged coefficient $\overline{F}^\tau$ with respect to the probability distribution μ^τ:

$$\overline{F}^\tau(x) = \int F(x, y)d\mu^\tau(y), \ \ \forall x \in L^2. \tag{52}$$

The mapping $\overline{F}^\tau$ is globally Lipschitz continuous, uniformly with respect to $\tau > 0$, owing to Assumption 2.3. More generally, the estimates of Proposition 3.2 are also satisfied by $\overline{F}^\tau$, uniformly with respect to $\tau > 0$.

The error estimate (51) is generalized as follows: assume that, for all $\vartheta \in [0, \vartheta_{\max})$ and $\kappa \in (0, 1 - \frac{d}{4})$, there exists $C_{\vartheta,\kappa} \in (0, \infty)$, such that

$$\sup_{x \in L^2} \left|(-A)^{-\frac{d}{4}-\kappa}\left(\overline{F}(x) - \overline{F}^\tau(x)\right)\right|_{L^2} \leq C_{\vartheta,\kappa}\tau^\vartheta. \tag{53}$$

To obtain (53) from the error estimate (51), it suffices to use a duality argument, and to apply (51) with the test function $\varphi = \langle h, (-A)^{-\frac{d}{4}-\kappa}(\overline{F} - \overline{F}^\tau)\rangle$, which is sufficiently regular using (7) and regularity properties of $D\overline{F}^\tau$.

To state convergence results, some notation concerning the speed of convergence to equilibrium is introduced. Let $\rho : (0, \infty) \to (0, \infty)$ be a non-increasing function, such that $\rho(t) \to_{t\to\infty} 0$. It is assumed that for every globally Lipschitz function $\varphi : L^2 \to \mathbb{R}$, there exists $C(\varphi) \in (0, \infty)$ such that

$$|\mathbb{E}[\varphi(Y_m^\tau)] - \int \varphi d\mu^\tau| \leq C(\varphi)\rho(m\tau), \tag{54}$$

and that

$$\sup_{x \in L^2} \left|\mathbb{E}[F(x, Y_m^\tau)] - \overline{F}^\tau(x)\right|_{L^2} \leq C\rho(m\tau). \tag{55}$$

The error estimates stated below depend on the following quantities,

$$\begin{aligned} R_1(M, M_a, \tau) = \frac{1}{M_a} \sum_{m=M-M_a+1}^{M} \rho(m\tau)\ , \ \ R_2(M, M_a, \tau) \\ = \frac{1}{M_a^2} \sum_{M-M_a+1 \leq m_1 < m_2 \leq M} \rho((m_2 - m_1)\tau), \end{aligned} \tag{56}$$

which depend on the auxiliary parameters M_a (number of terms in the averages) and M (maximal value of m in the average).

If it straightforward to check that, for fixed τ, one has $R_1(M, M_a, \tau) \to 0$ and $R_2(M, M_a, \tau) \to 0$ if $M_a \to \infty$. In addition, $R_1(M, M_a, \tau) \to 0$ if $M - M_a \to \infty$. More precise estimates are derived if the convergence is exponentially fast: when $\rho(t) = e^{-ct}$ for some $c > 0$, there exists $C \in (0, \infty)$ such that for all $M_a \leq M$ and all $\tau > 0$, one has

$$R_1(M, M_a, \tau) \leq \frac{Ce^{-c(M-M_a+1)\tau}}{M_a\tau + 1} \;,\; R_2(M, M_a, \tau) \leq \frac{C}{M_a\tau + 1}.$$

9.1.2. Macro-scheme

We are now in position to define the macro-scheme, and the resulting multiscale scheme. The guideline is to approximate the solution $\overline{X}(t)$ of the averaged problem, instead of $X^\epsilon(t)$, owing to the averaging principle, and to estimate the error by Theorems 4.5, 4.7 and 4.8.

Set $Y^\tau_{n,m} = Y^\tau_{nM+m}$ for all $n \in \mathbb{N}$, and $m \in \{0, \ldots, M\}$. The macro-scheme is based on the linear implicit Euler scheme: define

$$X_{n+1} = S_{\Delta t}\big(X_n + \Delta t \tilde{F}_n + \Delta W^Q_n\big), \tag{57}$$

where $X_0 = x_0$, $S_{\Delta t} = (I - \Delta t A)^{-1}$, $\Delta W^Q_n = W^Q\big((n+1)\Delta t\big) - W^Q\big(n\Delta t\big)$ are Wiener increments, and with the following approximation of $\overline{F}(X_n)$,

$$\tilde{F}_n = \frac{1}{M_a} \sum_{m=M-M_a+1}^{M} F(X_n, Y^\tau_{n,m}), \tag{58}$$

computed as a temporal average along the micro-scheme, depending on the parameters M_a and M.

9.1.3. Auxiliary processes

In order to analyze the multiscale scheme given by (57)–(58), and to give a clear discussion, two auxiliary processes are introduced.

First, applying the same integrator as in (57), *i.e.* the linear implicit Euler scheme, to discretize the averaged SPDE (22), set

$$\overline{X}_{n+1} = S_{\Delta t}\big(\overline{X}_n + \Delta t \overline{F}(\overline{X}_n) + \Delta W^Q_n\big), \qquad \overline{X}_0 = x_0.$$

If the averaged coefficient $\overline{F}$ was known, this scheme may be used directly, without the need to consider the multiscale scheme (indeed, the role of the micro-scheme is to provide an approximation of the averaged coefficient).

The following strong and weak error estimates for the discretization of the averaged equation are assumed to be satisfied: for all $T \in (0, \infty)$, all $\alpha \in [0, \alpha_{\max})$, and all test functions φ of class $\mathcal{C}^2_b$,

$$\begin{aligned} \sup_{0\leq t\leq T} \big(\mathbb{E}|\overline{X}(n\Delta t) - \overline{X}_n|^2\big)^{\frac{1}{2}} &\leq C_\alpha(T)\Delta t^{\min(\alpha,\frac{1}{2})}, \\ \sup_{0\leq t\leq T} \big|\mathbb{E}[\varphi(\overline{X}(n\Delta t))] - \mathbb{E}[\varphi(\overline{X}_n)]\big| &\leq C_\alpha(T,\varphi)\Delta t^{\min(2\alpha,1)}. \end{aligned} \tag{59}$$

Indeed, observe that the trajectories of the process $\overline{X}$ are almost surely α-Hölder continuous, for all $\alpha \in (0, \min(\alpha, \frac{1}{2}))$. When applying a simple numerical scheme such as the linear implicit Euler scheme, the strong and weak orders of convergence are directly related to the Hölder regularity of trajectories, see for instance [25,29,33] for references on numerical methods for

SPDEs. The fact that the weak order is twice the strong order is expected, see for instance [6] and references therein.

Second, recall that the invariant distribution μ^τ of the discrete-time process is not equal to μ in general. Using the averaged coefficient $\overline{F}^\tau$ defined by (52), set

$$d\overline{X}^\tau(t) = A\overline{X}^\tau(t)dt + \overline{F}^\tau(\overline{X}^\tau(t))dt + dW^Q(t), \ \ \overline{X}^\tau(0) = x_0,$$

and the associated numerical discretization

$$\overline{X}_{n+1}^\tau = S_{\Delta t}\big(\overline{X}_n^\tau + \Delta t\overline{F}^\tau(\overline{X}_n^\tau) + \Delta W_n^Q\big), \qquad \overline{X}_0^\tau = x_0, \tag{60}$$

which also uses the same integrator as in (57).

9.2. Convergence of the multiscale scheme (57) –(58)

9.2.1. Error estimates

Proposition 9.1 states a general convergence result, depending on the parameters Δt, τ, M and M_a. To simplify the analysis, it is assumed that F is bounded, however this restriction may be removed by proving moment estimates for the numerical scheme.

Let $\beta_{\max} = \frac{1}{2}$ when Assumption 4.1 is satisfied, whereas $\beta_{\max} = \frac{\alpha_{\max}}{1+\alpha_{\max}-\gamma_{\max}}$ when Assumption 4.2 is satisfied, where the parameters $\alpha_{\max}$ and $\gamma_{\max}$ are defined in Asssumptions 2.6 and 2.7. Recall that in many cases, the order of convergence for the approximation of the invariant distribution μ using the micro-scheme (see (51)) can be chosen equal to $\vartheta_{\max} = 2\gamma_{\max}$, however results are stated for a general value of $\vartheta_{\max}$.

Proposition 9.1. *For all $T \in (0, \infty)$, all $\alpha \in [0, \alpha_{\max})$, $\vartheta \in [0, \vartheta_{\max})$, and $\beta \in [0, \beta_{\max})$, there exists $C_{\alpha,\vartheta,\beta}(T) \in (0, \infty)$ such that the strong error is of size*

$$\begin{aligned}\sup_{0\le n\Delta t\le T} \big(\mathbb{E}|X_n - X^\epsilon(n\Delta t)|_{L^2}^2\big)^{\frac{1}{2}} &\le C_{\alpha,\vartheta,\beta}(T)\big(\epsilon^\beta + \Delta t^{\min(\alpha,\frac{1}{2})} + \tau^\vartheta\big)\\ &+ C_{\alpha,\vartheta,\beta}(T)\Bigg(\sqrt{R_1(M, M_a, \tau)}\\ &+ \sqrt{\Delta t}\Big(\frac{1}{\sqrt{M_a}} + \sqrt{R_2(M, M_a, \tau)}\Big)\Bigg).\end{aligned} \tag{61}$$

In addition, for all test functions φ of class $\mathcal{C}_b^2$, there exists $C_{\alpha,\vartheta,\beta}(T, \varphi) \in (0, \infty)$ such that the weak error is of size

$$\begin{aligned}\sup_{0\le n\Delta t\le T} \big|\mathbb{E}[\varphi(X_n)] - \mathbb{E}[\varphi(X^\epsilon(n\Delta t))]\big| &\le C_{\alpha,\vartheta,\beta}(T, \varphi)\big(\epsilon^{2\beta} + \Delta t^{\min(2\alpha,1)} + \tau^\vartheta\big)\\ &+ C_{\alpha,\vartheta,\beta}(T, \varphi)\Bigg(R_1(M, M_a, \tau)\\ &+ \Delta t\Big(\frac{1}{M_a} + R_2(M, M_a, \tau)\Big)\Bigg).\end{aligned} \tag{62}$$

Finally, for all test functions ψ of class $\mathcal{C}_b^2$, there exists $C_\vartheta(\psi) \in (0, \infty)$, such that,

$$\sup_{n\in\mathbb{N}} |\mathbb{E}[\psi(Y_{nM})] - \int \psi d\mu| \le C_\vartheta(\psi)\Big(\tau^\vartheta + \rho(nM\tau)\Big). \tag{63}$$

In fact, Proposition 9.1 is a straightforward corollary of Proposition 9.2, combined with results stated above:

- Theorems 4.5 and 4.7, or Theorem 4.8, to deal with the error in the averaging principle, which are the main results of this article,

- strong and weak error estimates (59) for the macro-scheme applied to the averaged SPDE (22),
- the estimate (53) (sampling error between the invariant distributions μ and μ^τ), which gives an error estimate $\sup_{0\le n\Delta t\le T}\big(\mathbb{E}|\overline{X}_n-\overline{X}_n^\tau|^2\big)^{\frac12}\le C_\vartheta\tau^\vartheta$ using a Gronwall type argument.

Note that (63) is a straightforward consequence of (51) and (54).

Proposition 9.2. *One has the following strong and weak error estimates.*

(1) For all $T\in(0,\infty)$, there exists $C(T)\in(0,\infty)$ such that, for all $\Delta t\in(0,1)$, $\tau\in(0,1)$, and $1\le M_a\le M$, one has

$$\sup_{0\le n\Delta t\le T}\big(\mathbb{E}|X_n-\overline{X}_n^\tau|_{L^2}^2\big)^{\frac12}\le C(T)\Big(\sqrt{R_1(M,M_a,\tau)}+\sqrt{\Delta t}\Big(\frac{1}{\sqrt{M_a}}+\sqrt{R_2(M,M_a,\tau)}\Big)\Big). \tag{64}$$

(2) For all test functions φ of class $\mathcal{C}_b^2$, there exists $C(T,\varphi)\in(0,\infty)$ such that, for all $\Delta t\in(0,1)$, $\tau\in(0,1)$, and $1\le M_a\le M$, one has

$$\sup_{0\le n\Delta t\le T}\big|\mathbb{E}[\varphi(X_n)]-\mathbb{E}[\varphi(\overline{X}_n^\tau)]\big|\le C(T,\varphi)\Big(R_1(M,M_a,\tau)+\Delta t\Big(\frac{1}{M_a}+R_2(M,M_a,\tau)\Big)\Big). \tag{65}$$

Observe that Proposition 9.2 implies the convergence of the macro-scheme (57) to the scheme (60), when $M_a\to\infty$, for any fixed values of Δt and τ. Note that to respect time-scales in (1), it is appropriate to choose parameters such that $M\tau=\epsilon^{-1}\Delta t$, and also $M_a\tau=\tilde{M}_a\epsilon^{-1}\Delta t$, thus the convergence property stated above may be interpreted as arising from taking the limit $\epsilon\to0$, and as an averaging principle at the discrete-time level. The limit scheme (60) is not an integrator for the averaged equation (22), but to a modified equation, with a residual depending on the micro time-step size τ.

We refer to [4] for a full analysis of the cost of the multiscale scheme (57)–(58).

9.2.2. Proof of Proposition 9.2

In order to prove the weak error estimate (65), an auxiliary function is introduced: for all $n\in\mathbb{N}_0$ and $x\in L^2$,

$$\overline{u}^\tau(n,x)=\mathbb{E}[\varphi(\overline{X}_n^\tau)\big|\overline{X}_0^\tau=x], \tag{66}$$

where the Markov chain $\big(\overline{X}_n^\tau\big)_{n\ge0}$ is given by (60). The function $\overline{u}^\tau$ is a discrete-time version of the function $\overline{u}$ studied in Proposition 7.1. Similarly, the spatial derivatives of $\overline{u}^\tau$ satisfy estimates given in Lemma 9.3.

Lemma 9.3. *Let $\varphi:L^2\to\mathbb{R}$ be a nice test function, and let $T\in(0,\infty)$. There exists $C(T)\in(0,\infty)$, such that for all $\tau>0$, all $\Delta t\in(0,T)$, and all $x,h,h_1,h_2\in L^2$, one has*

$$\sup_{n\Delta t\le T}|D_x\overline{u}^\tau(n,x).h|\le C(T)|h|_{L^2}\ ,\qquad \sup_{n\Delta t\le T}|D_x^2\overline{u}^\tau(n,x).(h_1,h_2)|\le C(T)|h_1|_{L^2}|h_2|_{L^2}.$$

The arguments in the proof are similar to those employed for the proof of Proposition 7.1, but are somehow simpler since no regularization effect is required in the remainder of the analysis.

Proof. The first and second order spatial derivatives of $\overline{u}^\tau$ are expressed as follows: for all $n \in \mathbb{N}$ and $x, h, h_1, h_2 \in H$, one has

$$D_x\overline{u}^\tau(n,x).h = \mathbb{E}\big[D\varphi(\overline{X}_n^\tau).\eta_n^{h,\tau}\big]$$
$$D_x^2\overline{u}^\tau(n,x).(h_1,h_2) = \mathbb{E}\big[D\varphi(\overline{X}_n^\tau).\zeta_n^{h_1,h_2,\tau}\big] + \mathbb{E}\big[D^2\varphi(\overline{X}_n^\tau).(\eta_n^{h_1,\tau},\eta_n^{h_2,\tau})\big],$$

where the process $\big(\eta_n^{h,\tau}\big)_{n\geq 0}$ and $\big(\zeta_n^{h,\tau}\big)_{n\geq 0}$ are solutions of

$$\eta_{n+1}^{h,\tau} = S_{\Delta t}\big(\eta_n^{h,\tau} + \Delta t D\overline{F}^\tau(\overline{X}_n^\tau).\eta_n^{h,\tau}\big),$$
$$\zeta_{n+1}^{h_1,h_2,\tau} = S_{\Delta t}\big(\zeta_n^{h_1,h_2,\tau} + \Delta t D\overline{F}^\tau(\overline{X}_n^\tau).\zeta_n^{h_1,h_2,\tau} + \Delta t D^2\overline{F}^\tau(\overline{X}_n^\tau).(\eta_n^{h_1,\tau},\eta_n^{h_2,\tau})\big),$$

with initial conditions $\eta_0^{h,\tau} = h$ and $\zeta_0^{h_1,h_2,\tau} = 0$. By recursion, mild formulations are obtained:

$$\eta_n^{h,\tau} = S_{\Delta t}^n h + \Delta t\sum_{k=0}^{n-1} S_{\Delta t}^{n-k} D\overline{F}^\tau(\overline{X}_k^\tau).\eta_k^{h,\tau},$$
$$\zeta_n^{h_1,h_2,\tau} = \Delta t\sum_{k=0}^{n-1} S_{\Delta t}^{n-k} D\overline{F}^\tau(\overline{X}_k^\tau).\zeta_k^{h,\tau} + \Delta t\sum_{k=0}^{n-1} S_{\Delta t}^{n-k} D^2\overline{F}^\tau(\overline{X}_k^\tau).(\eta_k^{h_1,\tau},\eta_k^{h_2,\tau}).$$

The following inequalities are used: for all $\alpha \in [0,1)$, there exists $C_\alpha \in (0,\infty)$ such that for all $n \in \mathbb{N}$,

$$|(-A)^\alpha S_{\Delta t}^n x|_{L^2} \leq C_\alpha (n\Delta t)^{-\alpha}|x|_{L^2},$$

see for instance [6, Lemma 5.2]. Combined with the inequality (7) and regularity properties of $\overline{F}^\tau$ (which are the same as for $\overline{F}$ given in Proposition 3.2), one obtains

$$|\eta_n^{h,\tau}|_{L^2} \leq C|h| + C\Delta t\sum_{k=0}^{n-1}|\eta_k^{h,\tau}|_{L^2},$$
$$|\zeta_n^{h_1,h_2,\tau}|_{L^2} \leq C\Delta t\sum_{k=0}^{n-1}|\zeta_k^{h_1,h_2,\tau}|_{L^2} + C_\kappa\Delta t\sum_{k=0}^{n-1}((n-k-1)\Delta t)^{-\frac{d}{4}-\kappa}|\eta_k^{h_1,\tau}|_{L^2}|\eta_k^{h_2,\tau}|_{L^2},$$

with $\kappa \in (0, 1-\frac{d}{4})$. Using a discrete Gronwall Lemma, one obtains successively, when $n\Delta t \leq T$, the inequalities

$$|\eta_n^{h,\tau}|_{L^2} \leq C(T)|h|_{L^2}\ , \quad |\zeta_n^{h_1,h_2,\tau}|_{L^2} \leq C(T)|h_1|_{L^2}|h_2|_{L^2},$$

and using the assumption that φ is a nice test function then concludes the proof. □

We are now in position to prove Proposition 9.2.

Proof of Proposition 9.2. Let us first establish the strong error estimate (64). For all $n \geq 0$, one has the equality

$$X_n - \overline{X}_n^\tau = \Delta t\sum_{k=0}^{n-1} S_{\Delta t}^{n-k}\big(\tilde{F}_k - \overline{F}^\tau(\overline{X}_k^\tau)\big)$$
$$= \Delta t\sum_{k=0}^{n-1} S_{\Delta t}^{n-k}\big(\overline{F}^\tau(X_k) - \overline{F}^\tau(\overline{X}_k^\tau)\big) + \Delta t\sum_{k=0}^{n-1} S_{\Delta t}^{n-k}\big(\tilde{F}_k - \overline{F}^\tau(X_k)\big).$$

On the one hand, thanks to the Lipschitz continuity of $\overline{F}^{\tau}$, one has

$$\Big(\mathbb{E}\Big|\Delta t\sum_{k=0}^{n-1}S_{\Delta t}^{n-k}\big(\overline{F}^{\tau}(X_k)-\overline{F}^{\tau}(\overline{X}_k^{\tau})\big)\Big|_{L^2}^2\Big)^{\frac{1}{2}}\le C\Delta t\sum_{k=0}^{n-1}\big(\mathbb{E}|X_k-\overline{X}_k^{\tau}|_{L^2}^2\big)^{\frac{1}{2}}.$$

On the other hand, writing $|\cdot|_{L^2}^2=\langle\cdot,\cdot\rangle$, and expanding the sums by bilinearity of the inner product, one obtains

$$\begin{aligned}\mathbb{E}\Big|\Delta t\sum_{k=0}^{n-1}S_{\Delta t}^{n-k}\big(\tilde{F}_k-\overline{F}^{\tau}(X_k)\big)\Big|_{L^2}^2&=\Delta t^2\sum_{k_1,k_2=0}^{n-1}\mathbb{E}\langle S_{\Delta t}^{n-k_1}\big(\tilde{F}_{k_1}-\overline{F}^{\tau}(X_{k_1})\big),\\&\quad S_{\Delta t}^{n-k_2}\big(\tilde{F}_{k_2}-\overline{F}^{\tau}(X_{k_2})\big)\rangle\\&\le\Delta t^2\sum_{k=0}^{n-1}\mathbb{E}|\tilde{F}_k-\overline{F}^{\tau}(X_k)|_{L^2}^2\\&\quad+2\Delta t^2\sum_{0\le k_1<k_2\le n-1}\mathbb{E}\langle S_{\Delta t}^{n-k_1}\big(\tilde{F}_{k_1}-\overline{F}^{\tau}(X_{k_1})\big),S_{\Delta t}^{n-k_2}\big(\tilde{F}_{k_2}-\overline{F}^{\tau}(X_{k_2})\big)\rangle\\&=\mathcal{E}_1+\mathcal{E}_2.\end{aligned}$$

The error term $\mathcal{E}_1$ is treated as follows. Using the definition (58) of $\tilde{F}_k$ as an average, and using the same type of expansion as above, one gets

$$\begin{aligned}\mathbb{E}|\tilde{F}_k-\overline{F}^{\tau}(X_k)|_{L^2}^2&\le\frac{1}{M_a^2}\sum_{m=M-M_a+1}^{M}\mathbb{E}|F(X_k,Y_{k,m}^{\tau})-\overline{F}^{\tau}(X_k)|^2\\&\quad+\frac{2}{M_a^2}\sum_{M-M_a+1\le m_1<m_2\le M}\mathbb{E}\langle F(X_k,Y_{k,m_1}^{\tau})-\overline{F}^{\tau}(X_k),F(X_k,Y_{k,m_2}^{\tau})\\&\quad-\overline{F}^{\tau}(X_k)\rangle.\end{aligned}$$

On the one hand, the assumption that F is bounded yields the inequality

$$\frac{1}{M_a^2}\sum_{m=M-M_a+1}^{M}\mathbb{E}|F(X_k,Y_{k,m}^{\tau})-\overline{F}^{\tau}(X_k)|^2\le\frac{C}{M_a}.$$

On the other hand, owing to a conditioning argument, combined with boundedness of F and the estimate (55), one obtains, when $m_1<m_2$, the estimate

$$\big|\mathbb{E}\langle F(X_k,Y_{k,m_1}^{\tau})-\overline{F}^{\tau}(X_k),F(X_k,Y_{k,m_2}^{\tau})-\overline{F}^{\tau}(X_k)\rangle\big|\le C\rho((m_2-m_1)\tau).$$

Gathering the estimates and using (56), one obtains

$$\mathbb{E}|\tilde{F}_k-\overline{F}^{\tau}(X_k)|_{L^2}^2\le C\Big(\frac{1}{M_a}+R_2(M,M_a,\tau)\Big),$$

and

$$\mathcal{E}_1\le C\Delta t\Big(\frac{1}{M_a}+R_2(M,M_a,\tau)\Big).$$

The error term $\mathcal{E}_2$ is treated as follows. Using again a conditioning argument, combined with boundedness of F and the estimate (55), one obtains, when $k_1<k_2$, the estimate

$$\big|\mathbb{E}\langle S_{\Delta t}^{n-k_1}\big(\tilde{F}_{k_1}-\overline{F}^{\tau}(X_{k_1})\big),S_{\Delta t}^{n-k_2}\big(\tilde{F}_{k_2}-\overline{F}^{\tau}(X_{k_2})\big)\rangle\big|\le\frac{C}{M_a}\sum_{m=M-M_a+1}^{M}\rho(m\tau),$$

hence (using (56))

$$|\mathcal{E}_2| \le C R_1(M, M_a, \tau).$$

Gathering the estimates yields

$$\big(\mathbb{E}|X_n - X_n^\tau|_{L^2}^2\big)^{\frac{1}{2}} \le C\Delta t \sum_{k=0}^{n-1} \big(\mathbb{E}|X_k - \overline{X}_k^\tau|_{L^2}^2\big)^{\frac{1}{2}} + C\Big(\frac{\Delta t}{M_a} + \Delta t R_2(M, M_a, \tau) + R_1(M, M_a, \tau)\Big)^{\frac{1}{2}}.$$

It remains to apply a discrete Gronwall Lemma to conclude the proof of the strong error estimate.

It remains to establish the weak error estimate (65).

The first and fundamental step in the analysis is to express the weak error in terms of the auxiliary function $\overline{u}^\tau$ defined by (66): one has

$$\begin{aligned}\mathbb{E}[\varphi(X_n)] - \mathbb{E}[\varphi(\overline{X}_n^\tau)] &= \mathbb{E}[\overline{u}^\tau(0, X_n)] - \mathbb{E}[\overline{u}^\tau(n, X_0)] \\ &= \sum_{k=0}^{n-1} \big(\mathbb{E}[\overline{u}^\tau(n-k-1, X_{k+1})] - \mathbb{E}[\overline{u}^\tau(n-k, X_k)]\big),\end{aligned}$$

by a telescoping sum argument. Using the definitions (60) and (57) of the schemes, and the Markov property, one obtains

$$\begin{aligned}\mathbb{E}[\overline{u}^\tau(n-k-1, X_{k+1})] &= \mathbb{E}[\overline{u}^\tau(n-k-1, S_{\Delta t}\big(X_k + \Delta t \tilde{F}_k + \Delta W_k^Q\big))], \\ \mathbb{E}[\overline{u}^\tau(n-k, X_k)] &= \mathbb{E}[\overline{u}^\tau(n-k-1, S_{\Delta t}\big(X_k + \Delta t \overline{F}^\tau(X_k) + \Delta W_k^Q\big))].\end{aligned}$$

A second-order Taylor expansion, combined with Lemma 9.3, then implies that

$$\begin{aligned}&\mathbb{E}[\overline{u}^\tau(n-k-1, X_{k+1})] - \mathbb{E}[\overline{u}^\tau(n-k, X_k)] \\ &\quad= \Delta t \mathbb{E}\Big[\, \langle D_x \overline{u}^\tau\big(n-k-1, S_{\Delta t}(X_k + \Delta t \overline{F}^\tau(X_k) + \Delta W_k^Q)\big), \\ &\qquad S_{\Delta t}\big(\tilde{F}_k - \overline{F}^\tau(X_k)\big)\rangle \,\Big] \\ &\qquad + \mathrm{O}(\Delta t^2)\mathbb{E}|\tilde{F}_k - \overline{F}^\tau(X_k)|_{L^2}^2 \\ &\quad= \frac{\Delta t}{M_a} \sum_{m=M-M_a+1}^{M} \Big(\, \mathbb{E}[\,\Psi^{\tau,\Delta t}(n-k-1, X_k, Y_{k,m}^\tau)] \\ &\qquad - \int \Psi(n-k-1, X_k, \cdot) d\mu^\tau \,\Big) \\ &\qquad + \mathrm{O}(\Delta t^2)\mathbb{E}|\tilde{F}_k - \overline{F}^\tau(X_k)|_{L^2}^2,\end{aligned}$$

with the auxiliary function $\Psi^{\tau,\Delta t}$ defined by:

$$\Psi^{\tau,\Delta t}(k, x, y) = \mathbb{E}\big[D_x \overline{u}^\tau\big(k, S_{\Delta t}(x + \Delta t \overline{F}^\tau(x) + \Delta W_0^Q)\big).(S_{\Delta t} F(x, y))\big],$$

where the Wiener increment ΔW_0^Q is independent of the process $\big(Y_k^\tau\big)_{k\ge 0}$.

In order to use (54), it suffices to check that the first and order derivative of $\Psi^{\tau,\Delta t}$ with respect to y satisfy the following estimates: for all $T \in (0, \infty)$ and $\kappa \in (0, 1 - \frac{d}{4})$, there exists $C_\kappa(T) \in (0, \infty)$, such that for all $x, y, h \in L^2$, all $\tau, \Delta t > 0$, one has

$$\sup_{k\Delta t \le T} |D_y \Psi^{\tau,\Delta t}(k, x, y).h| \le C_\kappa(T)|h|_{L^2}. \tag{67}$$

This claim is proved as follows: one has the identity

$$D_y \Psi^{\tau,\Delta t}(k,x,y).h = \mathbb{E}\big[\langle D_x \overline{u}^\tau\big(k, S_{\Delta t}(x + \Delta t \overline{F}^\tau(x) + \Delta W_0^Q)\big)., S_{\Delta t} D_y F(x,y).h\rangle\big],$$

and Lemma 9.3 and Assumption 2.3 imply the estimate

$$|D_y \Psi^{\tau,\Delta t}(k,x,y).h| \leq C(T)|D_y F(x,y).h|_{L^2} \leq C(T)|h|_{L^2},$$

which concludes the proof of (67). As a consequence, using (54) yields

$$\Big|\mathbb{E}[\Psi^{\tau,\Delta t}(n-k-1, X_k, Y_{k,m}^\tau)] - \int \Psi(n-k-1, X_k, \cdot) d\mu^\tau\Big| \leq C\rho(m\tau).$$

In addition, the following error estimate has been proved above:

$$\mathbb{E}|\tilde{F}_k - \overline{F}^\tau(X_k)|_{L^2}^2 \leq C\big(\frac{1}{M_a} + R_2(M, M_a, \tau)\big).$$

Finally, one obtains

$$\Big|\mathbb{E}[\overline{u}^\tau(n-k-1, X_{k+1})] - \mathbb{E}[\overline{u}^\tau(n-k, X_k)]\Big| \leq C\Delta t R_1(M, M_a, \tau) + C\Delta t^2\big(\frac{1}{M_a} + R_2(M, M_a, \tau)\big),$$

and taking the sum for $k \in \{0, \dots, n-1\}$ yields

$$\Big|\mathbb{E}[\varphi(X_n)] - \mathbb{E}[\varphi(\overline{X}_n^\tau)]\Big| \leq C R_1(M, M_a, \tau) + C\Delta t\big(\frac{1}{M_a} + R_2(M, M_a, \tau)\big),$$

which concludes the proof of (65). □

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The author would like to thank the two anonymous referees for their very careful reading of the initial submission and for their suggestions which improved the quality of the article.

The author would like to thank A. Debussche for discussions during the preparation of this manuscript. This work was partially supported by the project BORDS (ANR-16-CE40-0027-01) operated by the French National Research Agency (ANR).

References

[1] F. Bouchet, C. Nardini, T. Tangarife, Kinetic theory of jet dynamics in the stochastic barotropic and 2D Navier-Stokes equations, J. Stat. Phys. 153 (4) (2013) 572–625.

[2] F. Bouchet, C. Nardini, T. Tangarife, Stochastic averaging, large deviations and random transitions for the dynamics of 2D and geostrophic turbulent vortices, Fluid Dyn. Res. 46 (6) (2014) 061416, 11.

[3] C.-E. Bréhier, Strong and weak orders in averaging for SPDEs, Stochastic Process. Appl. 122 (7) (2012) 2553–2593.

[4] C.-E. Bréhier, Analysis of an HMM time-discretization scheme for a system of stochastic PDEs, SIAM J. Numer. Anal. 51 (2) (2013) 1185–1210.

[5] C.-E. Bréhier, Approximation of the invariant measure with an Euler scheme for stochastic PDEs driven by space–time white noise, Potential Anal. 40 (1) (2014) 1–40.

[6] C.-E. Bréhier, A. Debussche, Kolmogorov equations and weak order analysis for spdes with nonlinear diffusion coefficient, J. Math. Pures Appl. (2018).

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

[7] Z. Brzeźniak, On stochastic convolution in Banach spaces and applications, Stoch. Stoch. Rep. 61 (3–4) (1997) 245–295.
[8] S. Cerrai, Second Order PDE's in Finite and Infinite Dimension, in: Lecture Notes in Mathematics, vol. 1762, Springer-Verlag, Berlin, 2001, A probabilistic approach.
[9] S. Cerrai, A Khasminskii type averaging principle for stochastic reaction–diffusion equations, Ann. Appl. Probab. 19 (3) (2009) 899–948.
[10] S. Cerrai, Averaging principle for systems of reaction–diffusion equations with polynomial nonlinearities perturbed by multiplicative noise, SIAM J. Math. Anal. 43 (6) (2011) 2482–2518.
[11] S. Cerrai, M. Freidlin, Averaging principle for a class of stochastic reaction–diffusion equations, Probab. Theory Related Fields 144 (1–2) (2009) 137–177.
[12] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, second ed., in: Encyclopedia of Mathematics and its Applications, vol. 152, Cambridge University Press, Cambridge, 2014.
[13] Z. Dong, X. Sun, H. Xiao, J. Zhai, Averaging principle for one dimensional stochastic burgers equation, J. Differential Equations (2018).
[14] J.-P. Fouque, J. Garnier, G. Papanicolaou, K. Sølna, Wave Propagation and Time Reversal in Randomly Layered Media, in: Stochastic Modelling and Applied Probability, vol. 56, Springer, New York, 2007.
[15] M.I. Freidlin, A.D. Wentzell, Random Perturbations of Dynamical Systems, third ed., in: Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 260, Springer, Heidelberg, 2012, Translated from the 1979 Russian original by Joseph Szücs.
[16] H. Fu, J. Duan, An averaging principle for two-scale stochastic partial differential equations, Stoch. Dyn. 11 (2–3) (2011) 353–367.
[17] H. Fu, J. Liu, Strong convergence in stochastic averaging principle for two time-scales stochastic partial differential equations, J. Math. Anal. Appl. 384 (1) (2011) 70–86.
[18] H. Fu, L. Wan, J. Liu, Strong convergence in averaging principle for stochastic hyperbolic-parabolic equations with two time-scales, Stochastic Process. Appl. 125 (8) (2015) 3255–3279.
[19] H. Fu, L. Wan, J. Liu, X. Liu, Weak order in averaging principle for two-time-scale stochastic partial differential equations. arXiv preprint arXiv:1802.00903, 2018.
[20] H. Fu, L. Wan, J. Liu, X. Liu, Weak order in averaging principle for stochastic wave equation with a fast oscillation, Stochastic Process. Appl. 128 (8) (2018) 2557–2580.
[21] H. Fu, L. Wan, Y. Wang, J. Liu, Strong convergence rate in averaging principle for stochastic FitzHugh-Nagumo system with two time-scales, J. Math. Anal. Appl. 416 (2) (2014) 609–628.
[22] P. Gao, Averaging principle for the higher order nonlinear Schrödinger equation with a random fast oscillation, J. Stat. Phys. 171 (5) (2018) 897–926.
[23] P. Gao, Y. Li, Averaging principle for the Schrödinger equations, Discrete Contin. Dyn. Syst. Ser. B 22 (6) (2017) 2147–2168.
[24] T. Hytönen, J. van Neerven, M. Veraar, L. Weis, Analysis in Banach Spaces. Vol. I. Martingales and Littlewood-Paley Theory, in: Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 63, Springer, Cham, 2016.
[25] A. Jentzen, P.E. Kloeden, Taylor Approximations for Stochastic Partial Differential Equations, in: CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 83, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.
[26] R.Z. Khasminskii, On the principle of averaging the Itô's stochastic differential equations, Kybernetika (Prague) 4 (1968) 260–279.
[27] R.Z. Khasminskii, G. Yin, On averaging principles: an asymptotic expansion approach, SIAM J. Math. Anal. 35 (6) (2004) 1534–1560.
[28] R.Z. Khasminskii, G. Yin, Limit behavior of two-time-scale diffusions revisited, J. Differential Equations 212 (1) (2005) 85–113.
[29] R. Kruse, Strong and Weak Approximation of Semilinear Stochastic Evolution Equations, in: Lecture Notes in Mathematics, vol. 2093, Springer, Cham, 2014.
[30] C. Kuehn, Multiple Time Scale Dynamics, in: Applied Mathematical Sciences, vol. 191, Springer, Cham, 2015.
[31] S. Li, X. Sun, Y. Xie, Y. Zhao, Averaging principle for two dimensional stochastic navier-stokes equations. arXiv preprint arXiv:1810.02282, 2018.
[32] D. Liu, Strong convergence of principle of averaging for multiscale stochastic dynamical systems, Commun. Math. Sci. 8 (4) (2010) 999–1020.

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.

[33] G.J. Lord, C.E. Powell, T. Shardlow, An Introduction to Computational Stochastic PDEs, in: Cambridge Texts in Applied Mathematics, Cambridge University Press, New York, 2014.

[34] J.M.A.M. van Neerven, M.C. Veraar, L. Weis, Stochastic integration in UMD Banach spaces, Ann. Probab. 35 (4) (2007) 1438–1478.

[35] J.M.A.M. van Neerven, M.C. Veraar, L. Weis, Stochastic evolution equations in UMD Banach spaces, J. Funct. Anal. 255 (4) (2008) 940–993.

[36] G.A. Pavliotis, A.M. Stuart, Multiscale Methods, in: Texts in Applied Mathematics, vol. 53, Springer, New York, 2008, Averaging and homogenization.

[37] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, in: Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983.

[38] H. Triebel, Interpolation Theory, Function Spaces, Differential Operators, second ed., Johann Ambrosius Barth, Heidelberg, 1995.

[39] A.Y. Veretennikov, On an averaging principle for systems of stochastic differential equations, Mat. Sb. 181 (2) (1990) 256–268.

Please cite this article as: C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.09.015.