



**HAL**  
open science

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

Charles-Edouard Bréhier

► **To cite this version:**

Charles-Edouard Bréhier. Orders of convergence in the averaging principle for SPDEs: the case of a stochastically forced slow component. Stochastic Processes and their Applications, 2019, 10.1016/j.spa.2019.09.015 . hal-01896026

**HAL Id: hal-01896026**

**<https://hal.science/hal-01896026>**

Submitted on 15 Oct 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ORDERS OF CONVERGENCE IN THE AVERAGING PRINCIPLE FOR SPDES: THE CASE OF A STOCHASTICALLY FORCED SLOW COMPONENT

CHARLES-EDOUARD BRÉHIER

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.

## 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, 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 [13, 14, 27, 30] 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 [24], and to the following extensions since then, e.g. [25], [26], [29], [35] (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)

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

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 [11], and Section 2. The fast component is given by  $y^\epsilon(t, \cdot) = y(\epsilon^{-1}t, \cdot)$ , and is

---

*Key words and phrases.* Stochastic Partial Differential Equations, Averaging Principle, Poisson equation in infinite dimension, Heterogeneous Multiscale Method, strong and weak error estimates.

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  $x^\epsilon$ .

The averaging principle for such SPDE systems has been proved in a very general framework in [8], [10] for globally Lipschitz coefficients, and later in [9] in the case of non-globally Lipschitz continuous coefficients. In these results, no order of convergence, in terms of  $\epsilon \rightarrow 0$ , 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 [12, 15, 16, 19, 20, 28], in the parabolic SPDE case. See [17, 18], for some parabolic-hyperbolic systems. See [21, 22] 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. First, strong convergence deals with the mean-square error. Second, 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 [24] 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 [25, 26]. 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 [7] 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  $x^\epsilon$ . In fact, using this condition, the techniques of [3] may be applied, essentially with no modification, hence weak order of convergence equal to 1 in [19]. The objective of this manuscript is to obtain similar results with weaker regularity conditions. 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 [24], the Poisson equation technique described for instance in [30] 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  $y^\epsilon$  is not coupled with the slow component  $x^\epsilon$ , and it would be interesting to study the coupled case.

This article is organized as follows. Section 2 is devoted to the introduction of the 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 2.5 and 2.6 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, Theorems 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 shows 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, an application of the main result is presented, for the construction and analysis of efficient numerical schemes, based on Heterogeneous Multiscale Methods.

## 2. SETTING

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

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

This is the abstract formulation of a parabolic, semilinear, SPDE, in the framework of [11]. The stochastic forcing is given by a  $Q$ -Wiener process  $W^Q$ . In addition,  $Y^\epsilon$  is another stochastic process with values in  $L^2$ . In this work, it is assumed that  $Y^\epsilon$  and  $W^Q$  are independent.

We are interested in the regime of a small parameter  $\epsilon$ , and of a timescale separation:  $Y^\epsilon(t) = Y(t\epsilon^{-1})$ . As a consequence,  $X^\epsilon$  is referred to as the slow component, and  $Y^\epsilon$  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  $(w^q(t))_{t \geq 0}$  is a  $q$ -Wiener process, independent of  $W^Q$ . Then  $Y^\epsilon$  solves an equation of the type

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

in which case one has the equality in distribution (but not almost surely) of the processes  $(Y(t\epsilon^{-1}))_{t \geq 0}$  and  $(Y^\epsilon(t))_{t \geq 0}$ . 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  $p \in [2, \infty]$ , 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  $\gamma$ -Radonifying operators from  $L^2$  to  $L^p$ . Recall that a linear operator  $\Psi \in \mathcal{L}(L^2, L^p)$  is a  $\gamma$ -radonifying operator, if the image by  $\Psi$  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

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

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  $\|\Psi\|_{\mathcal{R}(L^2, L^2)}^2 = \text{Tr}(\Psi\Psi^*)$ , where  $\text{Tr}(\cdot)$  is the trace operator, and  $\Psi^*$  is the adjoint of  $\Psi$ .

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

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

Finally, recall the left and right ideal property for  $\gamma$ -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

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

Let  $(W(t))_{t \geq 0}$  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  $\int_0^T \Phi(t) dW^Q(t)$  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

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

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

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

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 [6, 33, 34] for Banach space valued stochastic integrals, and to [11] for the Hilbert space case.

If  $\varphi : L^2 \rightarrow \mathbb{R}$  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

$$\begin{cases} D(A) = H^2(0, 1) \cap H_0^1(0, 1), \\ Ax = \Delta x, \quad \forall x \in D(A), \end{cases}$$

where  $\Delta$  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 [31]) that there exists a complete orthonormal system  $(e_n)_{n \in \mathbb{N}}$  of  $L^2$ , and a non-decreasing sequence  $(\lambda_n)_{n \in \mathbb{N}}$  of positive real numbers such that

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

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 semi-group  $(e^{tA})_{t \geq 0}$ , 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 [31]:

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

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

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

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

Introduce also the kernel function  $K$  associated with the semigroup  $(e^{tA})_{t \geq 0}$ :

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

This kernel satisfies the two following properties:

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

To conclude this section, we state useful calculus inequalities, which are employed in a crucial way in Section 5.

**Proposition 2.1.** *For any  $\alpha \in [0, \frac{1}{2})$ , any  $\kappa \in (0, \frac{1}{2} - \alpha)$ , and any  $p \in [2, \infty)$ , there exists  $C_{\alpha, \kappa, p} \in (0, \infty)$ , such that for all  $x_1, x_2$ ,*

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

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  $C(\phi) \in (0, \infty)$  such that for all  $x, y_1, y_2$ ,

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

For the first inequality, we refer to [7, Section 3.2] and [32]. The second inequality is a straightforward consequence of the first inequality and of a first order Taylor formula:

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

**2.3. Assumptions on  $F$  and  $Q$ .** The coefficient  $F$  in (1) is defined as the Nemytskii operator (see Definition 2.4) associated with a smooth function  $f$  (see Assumption 2.2).

**Assumption 2.2.** *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  $1, \dots, 4$ .*

**Remark 2.3.** *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 2.4.** *For all  $p, q \in [2, \infty)$ , the mapping  $F : L^p \times L^p \rightarrow L^{p \wedge q}$  is defined as the Nemytskii operator, with  $F(x, y) = f(x(\cdot), y(\cdot))$  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,

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

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 2.5 below.

**Assumption 2.5.** *There exists a family of nonnegative real numbers  $(q_n)_{n \in \mathbb{N}}$  and a complete orthonormal system  $(f_n)_{n \in \mathbb{N}}$  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  $p \in [2, \infty)$ ,

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

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

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

where  $(\gamma_n)_{n \in \mathbb{N}}$  is a sequence of independent standard Gaussian random variables. Note that  $W^Q(t) = Q^{\frac{1}{2}} W(t)$  where  $W(t) = \sum_{n \in \mathbb{N}} \gamma_n f_n$  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 (2) does not depend on the time  $T$ , it only depends on  $Q$ , and on the parameters  $\alpha$  and  $p$ .

A sufficient condition for (2) 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  $(Y^\epsilon(t))_{t \geq 0}$  in (1) is defined in terms of an ergodic Markov process  $Y$ , such that, for all  $t \geq 0$ ,

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

**Assumption 2.6.** *The process  $Y = (Y(t))_{t \geq 0}$  is a continuous, ergodic, Markov process on  $H = L^2$ . Its unique invariant probability distribution is denoted by  $\mu$ .*

*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

$$(3) \quad \sup_{t \geq 0} \mathbb{E} |(-A)^\gamma Y(t)|_{L^p}^M \leq C_{\gamma,p,M} (1 + \mathbb{E} |(-A)^\gamma Y(0)|_{L^p}^M),$$

$$(4) \quad \int |(-A)^\gamma y|_{L^p}^M \mu(dy) \leq C_{\gamma,p,M},$$

The following standard notation is used:  $(Y_y(t))_{t \geq 0}$  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 2.7.** Define admissible functions  $\phi : H \rightarrow \mathbb{R}$ , to be such that, for some  $q \in [2, \infty)$ ,  $\phi$  is twice Fréchet differentiable on  $L^q$ , and such that there exists  $C \in (0, \infty)$  such that for all  $y \in H$ ,  $h, h_1, h_2 \in L^q$ ,

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

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

Assume that for any admissible function  $\phi$ , the Poisson equation

$$(5) \quad -\mathbf{L}\psi = \phi - \int \phi d\mu$$

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

$$(6) \quad \psi(y) = \int_0^\infty \mathbb{E} [\phi(Y_y(t)) - \int \phi d\mu] dt.$$

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  $\phi : H \rightarrow \mathbb{R}$  be an admissible function, and assume that there exists  $C(\phi) \in (0, \infty)$ , such that for all  $y_1, y_2 \in H$

$$(7) \quad |\phi(y_2) - \phi(y_1)| \leq C(\phi) (1 + |(-A)^\gamma y_1|_{L^p}^M + |(-A)^\gamma y_2|_{L^p}^M) |(-A)^\gamma (y_2 - y_1)|_{L^p}.$$

Then the solution  $\psi$  of the Poisson equation (5) satisfies, for all  $y \in H$ ,

$$(8) \quad |\psi(y)| \leq C_{\gamma,p,M} C(\phi) (1 + |(-A)^\gamma y|_{L^p}^{M+1}).$$

A sufficient condition to have the estimate (8) satisfied is given by Proposition 2.8.

**Proposition 2.8.** Let Assumption 2.6 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  $y_1, y_2 \in L^p$ ,

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

Then the estimate (8) is satisfied.

*Proof of Proposition 2.8.* By stationnarity,  $\psi$  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 (7) and Assumption 2.6,

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

□

**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  $\epsilon \in (0, 1)$ .

**Proposition 2.9.** *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  $t \in [0, T]$ ,*

$$(9) \quad X^\epsilon(t) = e^{tA}x_0 + \int_0^t e^{(t-s)A}F(X^\epsilon(s), Y^\epsilon(s))ds + \int_0^t e^{(t-s)A}dW^Q(s).$$

*In addition, the following moment estimates are satisfied, uniformly with respect to  $\epsilon \in (0, 1)$ : for any  $T \in (0, \infty)$ ,  $\alpha \in [0, \alpha_{\max})$ ,  $p \geq 2$  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  $|(-A)^\alpha x_0|_{L^p} < \infty$ ,*

$$(10) \quad \sup_{\epsilon \in (0,1)} \sup_{t \in [0,T]} \mathbb{E}[|(-A)^\alpha X^\epsilon(t)|_{L^p}^M] \leq C_{\alpha,p,M}(T) (1 + |(-A)^\alpha x_0|_{L^p}^M + |y_0|_{L^p}^M + M_{\alpha,p}(Q^{\frac{1}{2}}, T)^M).$$

**Remark 2.10.** *Using regularization properties of the semigroup  $(e^{tA})_{t \geq 0}$ , the moment estimates (10) may be replaced with*

$$\sup_{\epsilon \in (0,1)} \mathbb{E}[|(-A)^\alpha X^\epsilon(t)|_{L^p}^M] \leq C_{\alpha,p,M}(T) (1 + t^{-\alpha M} |x_0|_{L^p}^M + |y_0|_{L^p}^M + M_{\alpha,p}(Q^{\frac{1}{2}}, T)^M).$$

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

We conclude this section with a sketch of proof of Proposition 2.9.

*Proof of Proposition 2.9.* The existence and uniqueness of a mild solution (9) of (1) is obtained by a standard fixed point argument, see for instance [11].

The proof of the moment estimates (10) combines the following observations. On the one hand, by regularization properties of the semigroup and Lipschitz continuity of  $F$ ,

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

On the other hand, thanks to (2), see Assumption 2.5, the moment estimate

$$\mathbb{E}|(-A)^\alpha \int_0^t e^{(t-s)A}dW^Q(s)|_{L^p}^2 \leq c_p \int_0^T \|(-A)^\alpha e^{tA}Q^{\frac{1}{2}}\|_{\mathcal{R}(L^2, L^p)}^2 dt = c_p M_{\alpha,p}(Q^{\frac{1}{2}}, T)^2 < \infty$$

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 2.9.  $\square$

### 3. THE AVERAGING PRINCIPLE

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

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

$$(11) \quad \bar{F}(x) = \int F(x, y) d\mu(y) \in L^2,$$

where  $\mu$  is the unique invariant probability distribution of the ergodic process  $Y$ , see Assumption 2.6.

Since  $F$  is the Nemytskii operator associated with a globally Lipschitz continuous function  $f$ , using Assumption 2.6, it is straightforward to check that  $\bar{F}(x) \in L^p$  if  $x \in L^p$ , for all  $p \in [2, \infty)$ .

The first and second order derivatives of the averaged coefficient  $\bar{F}$  satisfy the following estimates.

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

$$(12) \quad \sup_{x \in L^2} |D\bar{F}(x).h|_{L^p} \leq C_p |h|_{L^p}.$$

For all  $p \in [2, \infty)$ , 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  $h_2 \in L^{p_2}$ ,

$$(13) \quad \sup_{x \in L^2} |D^2\bar{F}(x).(h_1, h_2)|_{L^p} \leq C_{p_1, p_2} |h_1|_{L^{p_1}} |h_2|_{L^{p_2}}.$$

*Proof of Proposition 3.2.* Note that for all  $x, y \in L^2$  and  $h \in L^p$ ,  $h_1 \in L^{p_1}$  and  $h_2 \in L^{p_2}$ ,

$$\begin{aligned} D_x F(x, y).h &= \partial_{z_1} f(x, y)h \\ D_x x F(x, y).(h_1, h_2) &= \partial_{z_1}^2 f(x, y)h_1 h_2. \end{aligned}$$

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

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

$$(14) \quad d\bar{X}(t) = A\bar{X}(t)dt + \bar{F}(\bar{X}(t))dt + dW^Q(t).$$

**Proposition 3.3.** Let  $T \in (0, \infty)$ . For any  $x_0 \in H$ , the SPDE (14) admits a unique mild solution, with initial condition  $\bar{X}(0) = x_0$ , such that for all  $t \in [0, T]$ ,

$$(15) \quad \bar{X}(t) = e^{tA}x_0 + \int_0^t e^{(t-s)A}\bar{F}(\bar{X}(s))ds + \int_0^t e^{(t-s)A}dW^Q(s).$$

In addition, the following moment estimates are satisfied, uniformly with respect to  $\epsilon \in (0, 1)$ : for any  $T \in (0, \infty)$ ,  $\alpha \in [0, \alpha_{\max})$ ,  $p \geq 2$  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  $|(-A)^\alpha x_0|_{L^p} < \infty$ ,

$$(16) \quad \sup_{t \in [0, T]} \mathbb{E}[|(-A)^\alpha \bar{X}(t)|_{L^p}^M] \leq C_{\alpha, p, M}(T) (1 + |(-A)^\alpha x_0|_{L^p}^M + M_{\alpha, p}(Q^{\frac{1}{2}}, T)^M).$$

The proof is omitted. Existence and uniqueness of the mild solution follows from the global Lipschitz continuity property of  $\bar{F}$  (thanks to (12), see Proposition 3.2). The moment estimates (16) are proved using the same arguments as in the proof of Proposition 2.9, in particular using (2), see Proposition 2.5.

To conclude this section, the infinitesimal generator associated with the averaged equation (14) is introduced:

$$(17) \quad \bar{\mathcal{L}}\varphi(x) = \langle D_x\varphi(x), Ax + \bar{F}(x) \rangle + \frac{1}{2} \sum_{n \in \mathbb{N}} q_n D_x^2\varphi(x) \cdot (f_n, f_n).$$

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

#### 4. STATEMENTS OF THE MAIN RESULTS

This section is devoted to the statements of the main results of this article, concerning the error in the averaging principle. We exhibit both strong and weak orders of convergence, with respect to  $\epsilon$ . Two situations need to be considered, depending on the regularity properties of the slow and the fast component, more precisely in terms of  $\gamma_{\max}$  (see Assumption 2.6) and  $\alpha_{\max}$  (see Assumption 2.5).

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

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

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

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

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

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

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

*Let  $\alpha_{\max} = \frac{1}{2}(1 - \frac{d}{2}(1 - \frac{2}{\rho}))$ . Moreover, assume that  $\gamma_{\max} \leq \alpha_{\max}$ .*

The definition of the parameter  $\alpha_{\max}$  in Assumption 4.2 is consistent with Assumption 2.5, see Proposition 8.1. Note that  $\alpha_{\max} \in (0, \frac{1}{2}]$ .

**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  $\text{Tr}(Q) < \infty$  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  $\text{Tr}(Q) = \|Q^{\frac{1}{2}}\|_{\mathcal{R}(L^2, L^2)} < \infty$ , 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 [30]). 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 (2). 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  $\kappa \in (0, \gamma_{\max} - \gamma)$ .

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

$$(18) \quad \sup_{t \in [0, T]} (\mathbb{E}|X^\epsilon(t) - \bar{X}(t)|_{L^2}^2)^{\frac{1}{2}} \leq C(T, x_0, y_0) (\text{Tr}(Q) + M_{1-\gamma,8}(Q^{\frac{1}{2}}, T))^{\frac{1}{2}} \epsilon^{\frac{1}{2}}.$$

To state the weak error result, an appropriate notion of admissible test function is used.

**Definition 4.6.** *Let  $\varphi : L^2 \rightarrow \mathbb{R}$ . It is called admissible if the following derivatives of  $\varphi$  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  $h \in L^2$ ,

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

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

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

- 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  $h_1 \in L^{p_1}, h_2 \in L^{p_2}, h_3 \in L^{p_3}$ ,

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

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

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

where  $\tilde{\varphi} : \mathbb{R} \rightarrow \mathbb{R}$  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 admissible 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  $\varphi : L^2 \rightarrow \mathbb{R}$  be an admissible test function.

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

$$(19) \quad \sup_{t \in [0, T]} |\mathbb{E}[\varphi(X^\epsilon(t))] - \mathbb{E}[\varphi(\bar{X}(t))]| \leq C(T, x_0, y_0, \varphi) (\text{Tr}(Q) + M_{1-\gamma,4}(Q^{\frac{1}{2}}, T)) \epsilon.$$

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 of the semigroup  $(e^{tA})_{t \geq 0}$ , and minor modifications in the proofs. However, one could not consider the supremum over time  $t \in [0, T]$  in (18) and (19). 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  $\frac{1}{2}$ , 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.

**4.2. The less regular case.** In this section it is assumed that Assumption 4.2 is satisfied. Let

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

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

**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

$$(20) \quad \sup_{t \in [0, T]} (\mathbb{E}|X^\epsilon(t) - \bar{X}(t)|_{L^2}^2)^{\frac{1}{2}} \leq C_\beta(T, x_0, y_0, Q)\epsilon^\beta.$$

Moreover, let  $\varphi : L^2 \rightarrow \mathbb{R}$  be an admissible 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

$$(21) \quad \sup_{t \in [0, T]} |\mathbb{E}[\varphi(X^\epsilon(t))] - \mathbb{E}[\varphi(\bar{X}(t))]| \leq C_\beta(T, x_0, y_0, Q, \varphi)\epsilon^{2\beta}.$$

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.

## 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  $\Phi(x, \cdot, \theta) : L^2 \rightarrow \mathbb{R}$  as the unique solution of

$$(22) \quad -\mathbf{L}\Phi(x, \cdot, \theta) = \langle F(x, \cdot) - \bar{F}(x), \theta \rangle,$$

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 2.7 is satisfied.

The function  $\Phi$  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 below. 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, Lemma 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 (22).

**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  $\theta \in L^{p'}$ ,*

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

*Proof.* For any fixed  $x \in L^2$  and  $\theta \in L^2 \subset L^{p'}$ , the mapping  $y \mapsto \langle F(x, y) - \bar{F}(x), \theta \rangle$  is an admissible function (with  $q = 4$ ). In addition, using Lipschitz continuity of  $F$ , 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 (7) is satisfied, with the parameters  $\alpha = 0$ ,  $p$ , and  $M = 0$ . By Assumption 2.7, then (8) is satisfied, which concludes the proof of Lemma 5.1.  $\square$

**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*

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

*Proof.* Observe that

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

using the third inequality in Proposition 2.1. This proves that (7) is satisfied, thus (8) follows, and this concludes the proof of Lemma 5.2.  $\square$

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  $x \in L^2$ ,  $y, \theta, h \in L^4$ ,*

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

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

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

where  $\phi_{x, \theta, h}(y) = \langle D_x(F(x, \cdot) - \bar{F}(x)).h, \theta \rangle$ . It is straightforward to check that  $\phi_{x, \theta, h}$  is an admissible function (by Assumption 2.2,  $f$  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

$$\begin{aligned} |\phi_{x, \theta, h}(y_2) - \phi_{x, \theta, h}(y_1)| &= |\langle D_x(F(x, y_2) - F(x, y_1)).h, \theta \rangle| \\ &= |\langle (\partial_{z_1} f(x, y_2) - \partial_{z_1} f(x, y_1))h, \theta \rangle| \\ &= |\langle (\partial_{z_1} f(x, y_2) - \partial_{z_1} f(x, y_1))\theta, h \rangle| \\ &\leq C|y_2 - y_1|_{L^4} \min(|\theta|_{L^4}|h|_{L^2}, |\theta|_{L^2}|h|_{L^4}), \end{aligned}$$

using Hölder inequality and by 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 (7), and consequently (8), are satisfied, for an appropriate choice of the parameters. This concludes the proof of Lemma 5.3.  $\square$

**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*

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

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

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

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

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

thanks to the second inequality of Proposition 2.1. It remains to use Assumption 2.7 to conclude the proof of Lemma 5.3.  $\square$

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  $h_1, h_2 \in L^8$ ,*

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

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

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

where  $\phi_{x, \theta, h_1, h_2}^{(2)}(y) = \langle D_x^2 (F(x, \cdot) - \bar{F}(x)) \cdot (h_1, h_2), \theta \rangle$ . It is straightforward to check that  $\phi_{x, \theta, h_1, h_2}^{(2)}$  is an admissible function (thanks to Assumption 2.2,  $f$  is of class  $\mathcal{C}^4$  with bounded derivatives of order  $1, \dots, 4$ ).

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

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

Thus it remains to apply Assumption 2.7 to conclude the proof of Lemma 5.5.  $\square$

## 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  $\delta F(x, y) = F(x, y) - \bar{F}(x)$ . Thanks to the mild formulations (9) and (15), the following decomposition of the averaging error is obtained:

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

Recall that  $F$  is globally Lipschitz-continuous, thanks to Assumption 2.2. The mean-square error is then bounded from above as follows:

$$\begin{aligned} \mathbb{E}|X^\epsilon(t) - \bar{X}(t)|^2 &\leq CT \int_0^t \mathbb{E}|X^\epsilon(s) - \bar{X}(s)|^2 ds \\ &\quad + 2 \int_0^t \int_s^t \mathbb{E} \left[ \langle e^{(t-s)A} \delta F(\bar{X}(s), Y^\epsilon(s)), e^{(t-r)A} \delta F(\bar{X}(r), Y^\epsilon(r)) \rangle \right] dr ds. \end{aligned}$$

Let  $\theta_{s,t}(r) = e^{(2t-s-r)A} \delta F(\bar{X}(s), Y^\epsilon(s))$ . Observe that  $\partial_r \theta_{s,t}(r) = -A \theta_{s,t}(r)$ . Using the definition (22) of  $\Phi$ , considering the quantity  $\mathbb{E}\Phi(\bar{X}(t), Y^\epsilon(t), \theta_{s,t}(t)) - \mathbb{E}\Phi(\bar{X}(s), Y^\epsilon(s), \theta_{s,t}(s))$ , and applying Itô formula, one obtains

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

where

$$(23) \quad \mathcal{I}_1^\epsilon(s, t) = \epsilon \mathbb{E}\Phi(\bar{X}(s), Y^\epsilon(s), \theta_{s,t}(s)) - \epsilon \mathbb{E}\Phi(\bar{X}(t), Y^\epsilon(t), \theta_{s,t}(t))$$

$$(24) \quad \mathcal{I}_2^\epsilon(s, t) = -\epsilon \int_s^t \mathbb{E} \left[ \Phi(\bar{X}(r), Y^\epsilon(r), A \theta_{s,t}(r)) \right] dr,$$

$$(25) \quad \mathcal{I}_3^\epsilon(s, t) = \epsilon \int_s^t \mathbb{E} \left[ \bar{\mathcal{L}}\Phi(\bar{X}(r), Y^\epsilon(r), \theta_{s,t}(r)) \right] dr,$$

where  $\bar{\mathcal{L}}$  is the infinitesimal generator associated with the averaged equation (14), see (17).

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

$$(26) \quad \mathcal{I}_{3,1}^\epsilon(s, t) = \epsilon \int_s^t \mathbb{E} \langle \bar{F}(\bar{X}(r)), D_x \Phi(\bar{X}(r), Y^\epsilon(r), \theta_{s,t}(r)) \rangle dr$$

$$(27) \quad \mathcal{I}_{3,2}^\epsilon(s, t) = \epsilon \int_s^t \mathbb{E} \langle A \bar{X}(r), D_x \Phi(\bar{X}(r), Y^\epsilon(r), \theta_{s,t}(r)) \rangle dr$$

$$(28) \quad \mathcal{I}_{3,3}^\epsilon(s, t) = \frac{\epsilon}{2} \int_s^t \mathbb{E} \text{Tr} \left( Q D_x^2 \Phi(\bar{X}(r), Y^\epsilon(r), \theta_{s,t}(r)) \right) dr$$

Lemmas 6.1, 6.2 and 6.3 below 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  $\epsilon \in (0, 1)$ ,*

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

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

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

**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  $\epsilon \in (0, 1)$ ,*

$$|\mathcal{I}_3^\epsilon(s, t)| \leq C_{\gamma, \kappa}(T) (t-s)^{-\gamma - \frac{\kappa}{2}} \epsilon (1 + |(-A)^{\gamma + \kappa} x_0|_{L^8}^3 + |(-A)^{\gamma + \kappa} y_0|_{L^8}^3) (1 + |(-A)^{1-\gamma} x_0|_{L^8} + M_{1-\gamma, 8}(Q^{\frac{1}{2}}) + \text{Tr}(Q)).$$



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

*Proof of Lemma 6.1.* For  $r \in \{s, t\}$ , note that  $\mathbb{E}|\theta_{s,t}(r)|_{L^2}^2 \leq C(1 + |x_0|_{L^2}^2 + |y_0|_{L^2}^2)$ , since  $F$  has at most linear growth, and thanks to moment estimates, in Proposition 3.3 and in Assumption 2.6. Thus, thanks to Lemma 5.1.

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

which concludes the proof.  $\square$

*Proof of Lemma 6.2.* For  $r \in (s, t)$ , using Lemma 5.1, and Assumption 2.6,

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

It is straightforward to check that  $\int_s^t \|Ae^{(2t-s-r)A}\|_{\mathcal{L}(H)} dr \leq C(T)\|(-A)^{\frac{1}{2}}e^{(t-s)A}\|_{\mathcal{L}(L^2, L^2)} \leq C(T)(t-s)^{-\frac{1}{2}}$ , thus one obtains

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

which concludes the proof.  $\square$

*Proof of Lemma 6.3.* First, proceeding as in the proof of Lemma 6.1, and using the global Lipschitz continuity of  $\bar{F}$ , it is straightforward to check that

$$|\mathcal{I}_{3,1}^\epsilon(s, t)| \leq C_T\epsilon(1 + |x_0|_{L^2}^3 + |y_0|_{L^2}^3).$$

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

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

Using the conditions on  $\gamma$  and  $\kappa$  above, and the moment estimates, one obtains

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

It remains to deal with the trace term,  $\mathcal{I}_{3,3}^\epsilon$ . Using Lemma 5.5 and Assumption 2.5,

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

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.  $\square$

Note that the assumption that  $\text{Tr}(Q) = \sum_{n \in \mathbb{N}} q_n$  is finite may be removed, using further regularity properties of the second order derivative  $D_x^2 \Phi$ . However, this does not seem to improve the result.

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

$$\begin{aligned} \mathbb{E}|X^\epsilon(t) - \bar{X}(t)|_{L^2}^2 &\leq CT \int_0^t \mathbb{E}|X^\epsilon(s) - \bar{X}(s)|_{L^2}^2 ds + \int_0^t (|\mathcal{I}_1^\epsilon(s, t)| + |\mathcal{I}_2^\epsilon(s, t)| + |\mathcal{I}_3^\epsilon(s, t)|) ds \\ &\leq CT \int_0^t \mathbb{E}|X^\epsilon(s) - \bar{X}(s)|_{L^2}^2 ds + C(T, x_0, y_0) M_{1-\gamma, 8}(Q^{\frac{1}{2}}, T) \epsilon. \end{aligned}$$

It remains to apply Gronwall Lemma to conclude the proof.  $\square$

## 7. PROOF OF THEOREM 4.7

The goal of this section is to provide a proof of Theorem 4.7, *i.e.* that 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  $\bar{u}$  defined below:

$$(29) \quad \bar{u}(t, x) = \mathbb{E}[\varphi(\bar{X}^x(t))].$$

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

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

with initial condition  $\bar{u}(0, \cdot) = \varphi$ , where  $\bar{\mathcal{L}}$  is the infinitesimal generator associated with the averaged equation (14), see (17).

To deal with this infinite dimensional PDE, usually an auxiliary approximation procedure is employed, see for instance [7], 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.

**Proposition 7.1** (Regularity properties of the derivatives of  $\bar{u}$ ). *Let  $\varphi$  be an admissible test function.*

*For all  $\beta < 1$ , there exists  $C_\beta(T) \in (0, \infty)$ , such that for all  $t \in (0, T]$ ,*

$$(30) \quad |D_x \bar{u}(t, x) \cdot (-A)^\beta h| \leq C_\beta(T) t^{-\beta} |h|_{L^2},$$

*and 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  $t \in (0, T]$ ,*

$$(31) \quad |D_x^2 \bar{u}(t, x) \cdot ((-A)^{\beta_1} h_1, (-A)^{\beta_2} h_2)| \leq C_{\beta_1, \beta_2}(T) t^{-\beta_1 - \beta_2} |h_1|_{L^2} |h_2|_{L^2}.$$

*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  $t \in [0, T]$ ,*

$$(32) \quad |D_x^3 \bar{u}(t, x) \cdot (h_1, h_2, h_3)| \leq C_{p_1, p_2, p_3}(T) |h_1|_{L^{p_1}} |h_2|_{L^{p_2}} |h_3|_{L^{p_3}}.$$

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 [7] for a recent overview of this topic and for further results. A sketch of proof is provided at the end of this section.

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

$$(33) \quad v(t, x, y) = \Phi(x, y, D_x \bar{u}(t, x)),$$

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

By construction,  $v(t, x, \cdot)$  is the solution of the Poisson equation (22) with  $\theta = D_x \bar{u}(t, x)$ , *i.e.* one has the fundamental identity

$$(34) \quad -\mathbf{L}v(t, x, y) = \langle F(x, y) - \bar{F}(x), D_x \bar{u}(t, x) \rangle.$$

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

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

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

Applying Itô formula, the weak error is written as

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

thanks to the identity (34). To exploit this formula for the weak error, note that the Itô formula applied with the function  $v$  yields the identity

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

As a consequence, the weak error has the following decomposition

$$(35) \quad \mathbb{E}[\varphi(X^\epsilon(T))] - \mathbb{E}[\varphi(\bar{X}(T))] = \mathcal{J}_1^\epsilon + \mathcal{J}_2^\epsilon + \mathcal{J}_3^\epsilon,$$

where

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

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

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

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  $x_0, y_0 \in H$ ,*

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

**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  $x_0, y_0 \in H$ ,*

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

**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  $x_0, y_0 \in L^8$ ,*

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

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  $t \in [0, T]$ ,  $x, y \in L^2$ ,

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

Combined with Assumption 2.6, this estimate concludes the proof of Lemma 7.2.  $\square$

*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,

$$\begin{aligned} \partial_t v(t, x, y) &= \Phi(x, y, \partial_t D_x \bar{u}(t, x)) \\ &= \Phi(x, y, D_x \partial_t \bar{u}(t, x)) \\ &= \Phi(x, y, D_x (\bar{\mathcal{L}}\bar{u}(t, x))) \\ &= \Phi(x, y, \Theta_1(t, x)) + \Phi(x, y, \Theta_2(t, x)) + \Phi(x, y, \Theta_3(t, x)), \end{aligned}$$

where

$$(36) \quad \langle \Theta_1(t, x), h \rangle = \langle Ah + D\bar{F}(x).h, D_x \bar{u}(t, x) \rangle,$$

$$(37) \quad \langle \Theta_2(t, x), h \rangle = D_x^2 \bar{u}(t, x).(h, Ax + \bar{F}(x)),$$

$$(38) \quad \langle \Theta_3(t, x), h \rangle = \frac{1}{2} \sum_{n \in \mathbb{N}} q_n D_x^3 \bar{u}(t, x).(f_n, f_n, h).$$

Let  $\kappa \in (0, 1)$ . Thanks to (30), one has

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

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

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

Thanks to (31), one has

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

Thus, thanks to Lemma 5.1,

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

Finally, thanks to (32) and Assumption 2.5, one has, for all  $h \in L^4$ ,

$$|\langle \Theta_3(t, x), h \rangle| \leq C(T) \sum_{n \in \mathbb{N}} q_n |f_n|_{L^8}^2 |h|_{L^4} \leq C(T)|h|_{L^4},$$

*i.e.*  $|\Theta_3(t, x)|_{L^{\frac{4}{3}}} = \sup_{h \in L^4, |h|_{L^4} \leq 1} |\langle \Theta_3(t, x), h \rangle| \leq C(T)$ . Thus Lemma 5.1 yields

$$|\Phi(x, y, \Theta_3(t, x))| \leq C(T)(1 + |y|_{L^4}).$$

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

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

This concludes the proof of Lemma 7.3.  $\square$

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

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

Observe that  $|D_x^2 \bar{u}(t, x) \cdot (h, \cdot)|_{L^2} \leq C|h|_{L^2}$ , thanks to (31), with  $\beta_1 = \beta_2 = 0$ . Then, thanks to Lemma 5.1 and Lemma 5.3, one obtains

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

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

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

To treat the second term,  $\mathcal{J}_{3,2}^\epsilon$ , observe that  $|D_x^2 \bar{u}(t, x) \cdot (h, \cdot)|_{L^2} \leq C_\kappa t^{-1+\kappa}|(-A)^{-1+\kappa}h|_{L^2}$ , for all  $\kappa \in (0, 1]$ , thanks to (31). In addition,  $|(-A)^{1-\kappa}D_x \bar{u}(t, x)|_{L^2} \leq C_\kappa t^{-1+\kappa}$ , thanks to (30). Then, thanks to Lemma 5.1 and Lemma 5.4,

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

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

As a consequence

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

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,

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

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

$$\begin{aligned} D_x^2 v(t, x, y) \cdot (h, h) &= \Phi(x, y, D_x^3 \bar{u}(t, x) \cdot (h, h, \cdot)) + 2\langle D_x \Phi(x, y, D_x^2 \bar{u}(t, x) \cdot (h, \cdot)), h \rangle \\ &\quad + D_x^2 \Phi(x, y, D_x \bar{u}(t, x)) \cdot (h, h). \end{aligned}$$

First, observe that  $|D_x^3 \bar{u}(t, x) \cdot (h, h, k)| \leq C|h|_{L^4}^2|k|_{L^2}$ , thanks to (32). Equivalently, this means that  $|D_x^3 \bar{u}(t, x) \cdot (h, h, \cdot)|_{L^2} \leq C|h|_{L^4}^2$ , then Lemma 5.1 yields

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

Second, thanks to Lemma 5.3,

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

Finally, Lemma 5.5 and (30) yield

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

As a consequence, one obtains

$$\begin{aligned} |\mathcal{J}_{3,3}^\epsilon| &= \left| \frac{\epsilon}{2} \int_0^T \sum_{n \in \mathbb{N}} q_n \mathbb{E} [D_x^2 v(T-t, X^\epsilon(t), Y^\epsilon(t)) \cdot (f_n, f_n)] dt \right| \\ &\leq C\epsilon \sum_{n \in \mathbb{N}} q_n |f_n|_{L^4}^2 \int_0^T (1 + \mathbb{E}|Y^\epsilon(t)|_{L^4}) dt \\ &\leq C(T) \text{Tr}(Q) \epsilon (1 + |y_0|_{L^4}), \end{aligned}$$

thanks to Assumption 2.5, and a moment estimate, see Assumption 2.6.

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

$$|\mathcal{J}_3^\epsilon| \leq C_{\gamma, \kappa} \epsilon (1 + |(-A)^{1-\gamma} x_0|_{L^8}^4 + |(-A)^{\gamma+\kappa} y_0|_{L^8}^4).$$

This concludes the proof of Lemma 7.4.  $\square$

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

*Proof.* Thanks to the decomposition (35) of the weak error, it suffices to gather the estimates of Lemmas 7.2, 7.3 and 7.4 to conclude.  $\square$

To conclude this section, we provide a sketch of proof of Proposition 7.1, which states the regularity properties for the spatial derivatives of  $\bar{u}(t, x)$  used above.

*Sketch of proof of Proposition 7.1.* The proof is based on computing the derivatives of  $\bar{u}$  in terms of tangent processes, which are solutions of PDEs with random coefficients (noise is additive in (14)). A mild formulation and regularity properties of the semigroup  $(e^{tA})_{t \geq 0}$  yield the required estimates.

- First-order derivative.

Note that

$$D_x \bar{u}(t, x) \cdot h = \mathbb{E} [D\varphi(\bar{X}^x(t)) \cdot \eta^h(t)],$$

where

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

Thanks to the global Lipschitz continuity of the averaged coefficient  $\bar{F}$ , one obtains, for all  $t \in (0, T]$ ,

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

An application of Gronwall Lemma yields

$$|\eta^h(t)|_{L^2} \leq C_\beta(T) t^{-\beta} |(-A)^{-\beta} h|_{L^2},$$

hence (30)

- Second-order derivative.

Note that

$$\begin{aligned} D_x^2 \bar{u}(t, x) \cdot (h_1, h_2) &= \mathbb{E} [\langle D\varphi(\bar{X}^x(t)), \zeta^{h_1, h_2}(t) \rangle] \\ &\quad + \mathbb{E} [D^2 \varphi(\bar{X}^x(s)) \cdot (\eta^{h_1}(t), \eta^{h_2}(t))], \end{aligned}$$

where

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

First,  $\varphi$  is an admissible test function, thus one obtains

$$\mathbb{E}[D^2\varphi(\bar{X}^x(s)) \cdot (\eta^{h_1}(t), \eta^{h_2}(t))] \leq C\mathbb{E}[|\eta^{h_1}(t)|_{L^2}|\eta^{h_2}(t)|_{L^2}],$$

which is treated using the estimate proved above.

To treat the second term, the inequality  $\|e^{tA}\|_{\mathcal{L}(L^1, L^2)} \leq C_d t^{-\frac{d}{4}}$  is used. This may be proved as follows. First, by a duality argument,  $\|e^{tA}\|_{\mathcal{L}(L^1, L^2)} = \|e^{tA}\|_{\mathcal{L}(L^2, L^\infty)}$ . In addition, by Jensen inequality, one has

$$e^{tA}x(\xi)^2 = \left( \int K(t, \xi, \eta)x(\eta)d\eta \right)^2 \leq \int K(t, \xi, \eta)x(\eta)^2d\eta \leq C_d t^{-\frac{d}{2}}|x|_{L^2}^2,$$

where  $K$  is the kernel associated with the semigroup. As a consequence, one has  $\|e^{tA}\|_{\mathcal{L}(L^2, L^\infty)} \leq C_d t^{-\frac{d}{4}}$ . Note that  $\frac{d}{4} < 1$  if  $d \in \{1, 2, 3\}$ .

Thanks to (12) and (13), and the estimate above, one obtains, with the application of Gronwall Lemma,

$$\begin{aligned} |\zeta^{h_1, h_2}(t)|_{L^2} &\leq C_p \int_0^t |\zeta^{h_1, h_2}(s)|_{L^2} ds + C_{\beta_1, \beta_2} \int_0^t (t-s)^{-\frac{d}{4}} s^{-\beta_1 - \beta_2} ds |(-A)^{-\beta_1} h_1|_{L^2} |(-A)^{-\beta_2} h_2|_{L^2} \\ &\leq C_{\beta_1, \beta_2}(T) |(-A)^{-\beta_1} h_1|_{L^2} |(-A)^{-\beta_2} h_2|_{L^2}. \end{aligned}$$

It is then straightforward to obtain (31).

- Third-order derivative: the proof is omitted, since the computations are similar.

□

## 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

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

**Proposition 8.1.** *Let Assumption 4.2 be satisfied. Then (2) is satisfied: for all  $\alpha \in [0, \alpha_{\max})$  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  $\gamma$ -Radonifying operators,

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

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

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

Recall that  $K$  is the kernel associated with the semigroup  $(e^{tA})_{t \geq 0}$ .

On the one hand, Assumption 2.5 implies that for all  $z \in \mathcal{D}$ ,

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

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

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

using the properties of the kernel  $K$  stated above.

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

$$\left| \sum_{n \in \mathbb{N}} q_n(e^{tA} f_n)^2 \right|_{L^{\frac{p}{2}}} \leq Ct^{-\frac{d}{2\zeta}},$$

thus

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

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

This concludes the proof of Proposition 8.1.  $\square$

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  $\delta > 0$ ,*

$$(39) \quad \text{Tr}(e^{2\delta A} Q) + M_{1-\gamma, p}(e^{\delta A} Q^{\frac{1}{2}}, T) \leq C_{\alpha, \gamma, p}(Q, T) \delta^{\alpha + \gamma - 1}.$$

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

To deal with the trace term, we use the Hölder type inequality for Schatten norms  $\|\cdot\|_{\mathcal{L}_\rho(L^2)}$ , with parameter  $\rho \in [1, \infty]$ , see for instance [23, Corollary D.2.4, Appendix D]. One obtains

$$\text{Tr}(e^{2\delta A} Q) = \|e^{2\delta A} Q\|_{\mathcal{L}_1(L^2)} \leq \|e^{2\delta A}\|_{\mathcal{L}_\zeta(L^2)} \|Q\|_{\mathcal{L}_{\frac{\rho}{2}}(L^2)},$$

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

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

using  $\lambda_n \underset{n \rightarrow \infty}{\sim} c_d n^{\frac{d}{2}}$ . As a consequence,

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

using the definition of  $\alpha_{\max} = \frac{1}{2}(1 - \frac{d}{2}(1 - \frac{2}{\rho})) = \frac{1}{2}(1 - \frac{d}{2\zeta})$ .

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



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  $\delta > 0$  (this parameter will be chosen below),  $X_\delta^\epsilon$  and  $\bar{X}_\delta$  are solutions of

$$(40) \quad \begin{aligned} dX_\delta^\epsilon(t) &= AX_\delta^\epsilon(t)dt + F(X_\delta^\epsilon(t), Y^\epsilon(t))dt + e^{\delta A}dW^Q(t), \\ d\bar{X}_\delta(t) &= A\bar{X}_\delta(t)dt + \bar{F}(\bar{X}_\delta(t))dt + e^{\delta A}dW^Q(t), \end{aligned}$$

with initial conditions  $X_\delta^\epsilon(0) = \bar{X}_\delta = x_0$ .

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 Theorems 4.5 and 4.7 may be applied when considering the auxiliary processes  $X_\delta^\epsilon$  and  $\bar{X}_\delta$  defined by (40).

**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*

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

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

Let  $\varphi$  be an admissible 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  $\delta > 0$ ,

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

and

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

Second, the distances between  $X_\delta^\epsilon$  and  $X^\epsilon$ , and between  $\bar{X}_\delta$  and  $\bar{X}$ , 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  $\varphi$  be an admissible 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  $\delta > 0$ , one has*

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

and

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

*Proof of Lemma 8.3.* This is a straightforward application of Theorems 4.5 and 4.7 combined with Lemma 8.2.  $\square$

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

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

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

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 (29). 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

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

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  $h \in L^2$ ,  $D^2 \overline{u}(t, x) \cdot h \in L^2$  is characterized by

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

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

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

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

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

Finally,  $(-A)^{-1+2\kappa} (I - e^{\delta A})$  is a self-adjoint, compact, linear operator, thus, for  $\alpha \leq \frac{1}{2}$ , one has

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

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

Finally, one obtains

$$|\text{Tr}(D^2 \overline{u}(t-s, \overline{X}_\delta(s)) (e^{\delta A} - I) Q e^{\delta A})| \leq C_\alpha \delta^{2\alpha} (t-s)^{-1+2\kappa},$$

and similarly

$$|\text{Tr}(D^2 \overline{u}(t-s, \overline{X}_\delta(s)) Q (e^{\delta A} - I))| \leq C_\alpha \delta^{2\alpha} (t-s)^{-1+2\kappa}.$$

It is then straightforward to conclude that

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

This concludes the proof of Lemma 8.4.  $\square$

We are now in position to provide the proof of Theorem 4.8, which consists in choosing  $\delta$  in terms of  $\epsilon$  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  $\delta > 0$ ,

$$\begin{aligned} \sup_{t \in [0, T]} (\mathbb{E}|X^\epsilon(t) - \bar{X}(t)|_{L^2}^2)^{\frac{1}{2}} &\leq C_{\alpha, \gamma}(T, x_0, y_0, Q) \left( \delta^{-\frac{1-\alpha-\gamma}{2}} \epsilon^{\frac{1}{2}} + \delta^\alpha \right), \\ \sup_{t \in [0, T]} |\mathbb{E}[\varphi(X^\epsilon(t))] - \mathbb{E}[\varphi(\bar{X}(t))]| &\leq C_{\alpha, \gamma}(T, x_0, y_0, Q, \varphi) \left( \delta^{-(1-\alpha-\gamma)} \epsilon + \delta^{2\alpha} \right). \end{aligned}$$

Choosing  $\delta = \epsilon^{\frac{1}{1+\alpha-\gamma}}$ , with  $\alpha_{\max} - \alpha$  and  $\gamma_{\max} - \gamma$  arbitrarily small then concludes the proof.  $\square$

**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  $p \geq 2$ , 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  $\alpha_{\max} \geq \frac{1}{2}$ , the situation is the same as in Theorem 4.8. On the other hand, if  $\alpha_{\max} < \frac{1}{2}$ , 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, is left for future works.*

## 9. EFFICIENT NUMERICAL APPROXIMATION OF THE SLOW COMPONENT

The goal of this section is to describe a temporal discretization scheme for the slow component  $X^\epsilon$  in (1), which is stable and efficient when  $\epsilon \rightarrow 0$ . Indeed, given a time-step size  $h > 0$ , stability for the discretization of the evolutions of  $X^\epsilon$  and  $Y^\epsilon$  requires to choose  $h$  such that  $h = O(\epsilon)$ . The scheme proposed below is based on Heterogeneous Multiscale Methods, see [4] and references therein. Instead of using a single time-step size  $h > 0$ , two time-step sizes  $\Delta t > 0$  and  $\delta t > 0$  are introduced. The slow component  $X^\epsilon$  is discretized using a macro-scheme, with time-step size  $\Delta t$ : the scheme is constructed such that  $\Delta t$  does not depend on the small parameter  $\epsilon$ . The fast component  $Y^\epsilon$  is discretized using a micro-scheme, with time-step size  $\delta t$ . Since in (1), the fast component  $Y^\epsilon(t) = Y(\epsilon^{-1}t)$  is not coupled with the slow component  $X^\epsilon$ , in this manuscript we can rely on a discretization of the process  $Y$ , with a time-step size  $\tau > 0$ .

The detailed construction of the scheme is presented and discussed in Section 9.1. Convergence results are stated in Section 9.2. The proofs are omitted, since they would be similar to those in [4], where the slow component was not driven by a stochastic forcing.

**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  $\tau > 0$ . Two other integer parameters  $M$  and  $M_a \in \{1, \dots, M\}$  are used, to insert data from the micro-scheme into the macro-scheme, in terms of temporal averages.

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  $(Y_m^\tau)_{m \in \mathbb{N}_0}$  be computed using a numerical integrator  $\Phi^\tau$  for the stochastic process  $(Y(t))_{t \geq 0}$ . It is assumed that this discrete-time process defines an ergodic Markov chain on  $L^2$ , with unique invariant probability distribution denoted by  $\mu^\tau$ .

It is natural, see for instance [4], to assume that the error between  $\mu$  and  $\mu^\tau$  is of the order  $\tau^{2\gamma}$ , for all  $\gamma \in [0, \gamma_{\max})$ , in the following sense: for all functions  $\varphi : L^2 \rightarrow \mathbb{R}$  of class  $\mathcal{C}^2$ , with bounded first and second order derivatives, and all  $\gamma \in [0, \gamma_{\max})$ , there exists  $C_\gamma(\varphi) \in (0, \infty)$  such that

$$(41) \quad \left| \int \varphi d\mu^\tau - \int \varphi d\mu \right| \leq C_\gamma(\varphi) \tau^{2\gamma}.$$

Moreover, define an averaged coefficient  $\bar{F}^\tau$  with respect to the probability distribution  $\mu^\tau$ :

$$(42) \quad \bar{F}^\tau(x) = \int F(x, y) d\mu^\tau(y), \quad \forall x \in L^2.$$

The approximation result above is extended as follows: assume that, for all  $\gamma \in [0, \gamma_{\max})$ , there exists  $C_\gamma \in (0, \infty)$ , such that

$$\sup_{x \in L^2} |\bar{F}(x) - \bar{F}^\tau(x)|_{L^2} \leq C_\gamma \tau^{2\gamma}.$$

The conditions above are satisfied if  $Y$  is the solution of a parabolic semilinear SPDE, driven by additive noise, and  $Y^\tau$  is obtained applying the linear-implicit Euler scheme, under suitable conditions on the nonlinearity in the equation, see [5].

To state convergence results, notation concerning the speed of convergence to equilibrium is introduced. Let  $\rho : (0, \infty) \rightarrow (0, \infty)$  be non-increasing,  $\rho(t) \xrightarrow{t \rightarrow \infty} 0$ , and assume that

$$(43) \quad |\mathbb{E}[\varphi(Y_m^\tau)] - \int \varphi d\mu^\tau| \leq C(\varphi) \rho(m\tau),$$

and that

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

Finally, define the following quantities

$$R_1(M, M_a, \tau) = \frac{1}{M_a} \sum_{m=M-M_a+1}^M \rho(m\tau), \quad R_2(M, M_a, \tau) = \frac{1}{M_a^2} \sum_{M-M_a+1 \leq m_1 < m_2 \leq M} \rho((m_2 - m_1)\tau).$$

Using a Cesaro type argument, for fixed  $\tau > 0$ , if  $M_a \rightarrow \infty$ , then both  $R_1(M, M_a, \tau) \rightarrow 0$  and  $R_2(M, M_a, \tau) \rightarrow 0$ . If  $M - M_a \rightarrow \infty$ , then also  $R_1(M, M_a, \tau) \rightarrow 0$ .

More precisely, if  $\rho(t) = e^{-ct}$  for some  $c > 0$ , note that there exists  $C \in (0, \infty)$  such that for all  $M_a \leq M$  and all  $\tau > 0$ ,

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

9.1.2. *Macro-scheme.* We are now in position to define the macro-scheme. The principle is to approximate  $\bar{X}(t)$  instead of  $X^\epsilon(t)$ , thanks to the averaging principle, and using error estimates given by Theorems 4.5, 4.7 and 4.8. It is thus sufficient to compute an approximation of the averaged coefficient  $\bar{F}$ , using the ergodicity of the micro-scheme and the error estimate (41), and to apply a standard integrator with time-step size  $\Delta t > 0$ .

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

$$(44) \quad X_{n+1} = S_{\Delta t}(X_n + \Delta t \tilde{F}_n + \Delta W_n^Q),$$

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

$$(45) \quad \tilde{F}_n = \frac{1}{M_a} \sum_{m=M-M_a+1}^M F(X_n, Y_{n,m}^\tau),$$

computed as a temporal average, depending on the parameters  $M_a$  and  $M$ .

## 9.2. Convergence of the multiscale scheme (44)-(45).

9.2.1. *Auxiliary schemes.* In order to analyze the multi-scale scheme given by (44)-(45), and to give a clear discussion, several schemes are introduced.

First, applying the same integrator as in (44), *i.e.* the linear implicit Euler scheme, to discretize the averaged SPDE (14), introduce the following scheme,

$$\bar{X}_{n+1} = S_{\Delta t}(\bar{X}_n + \Delta t \bar{F}(\bar{X}_n) + \Delta W_n^Q), \quad \bar{X}_0 = x_0.$$

Second, due to the error in sampling the invariant distribution  $\mu$  using the micro-scheme with time-step size  $\tau > 0$ , define a modified averaged equation

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

and the associated numerical discretization

$$(46) \quad \bar{X}_{n+1}^\tau = S_{\Delta t}(\bar{X}_n^\tau + \Delta t \bar{F}^\tau(\bar{X}_n^\tau) + \Delta W_n^Q), \quad \bar{X}_0^\tau = x_0.$$

Based on the literature concerning the numerical analysis of SPDEs, rates of convergence for these numerical schemes are assumed to be as follows: for all  $T \in (0, \infty)$ , all  $\alpha \in [0, \alpha_{\max})$ , and all test functions  $\varphi$  of class  $\mathcal{C}_b^2$ ,

$$(47) \quad \begin{aligned} \sup_{0 \leq t \leq T} (\mathbb{E}|\bar{X}(n\Delta t) - \bar{X}_n|^2)^{\frac{1}{2}} &\leq C_\alpha(T) \Delta t^{\min(\alpha, \frac{1}{2})}, \\ \sup_{0 \leq t \leq T} |\mathbb{E}[\varphi(\bar{X}(n\Delta t))] - \mathbb{E}[\varphi(\bar{X}_n)]| &\leq C_\alpha(T, \varphi) \Delta t^{\min(2\alpha, 1)}. \end{aligned}$$

9.2.2. *Error estimates.* Proposition 9.1 below states a general convergence result, depending on the parameters  $\Delta t$ ,  $\tau$ ,  $M$  and  $M_a$ .

Let  $\beta_{\max} = \frac{1}{2}$  when Assumption 4.1 is satisfied, and recall that  $\beta_{\max} = \frac{\alpha_{\max}}{1 + \alpha_{\max} - \gamma_{\max}}$  if Assumption 4.2 is satisfied.

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

$$(48) \quad \begin{aligned} \sup_{0 \leq n\Delta t \leq T} (\mathbb{E}|X_n - X^\epsilon(n\Delta t)|_{L^2}^2)^{\frac{1}{2}} &\leq C_{\alpha, \gamma, \beta}(T) (\epsilon^\beta + \Delta t^{\min(\alpha, \frac{1}{2})} + \tau^{2\gamma}) \\ &+ C_{\alpha, \gamma, \beta}(T) \left( \sqrt{R_1(M, M_a, \tau)} + \sqrt{\Delta t} \left( \frac{1}{\sqrt{M_a}} + \sqrt{R_2(M, M_a, \tau)} \right) \right), \end{aligned}$$

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

$$(49) \quad \begin{aligned} \sup_{0 \leq n\Delta t \leq T} |\mathbb{E}[\varphi(X_n)] - \mathbb{E}[\varphi(X^\epsilon(n\Delta t))]| &\leq C_{\alpha, \gamma, \beta}(T, \varphi) (\epsilon^{2\beta} + \Delta t^{\min(2\alpha, 1)} + \tau^{2\gamma}) \\ &+ C_{\alpha, \gamma, \beta}(T, \varphi) \left( R_1(M, M_a, \tau) + \Delta t \left( \frac{1}{M_a} + R_2(M, M_a, \tau) \right) \right). \end{aligned}$$

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

$$(50) \quad \sup_{n \in \mathbb{N}} |\mathbb{E}[\psi(Y_{nM})] - \int \psi d\mu| \leq C_\tau(\psi) (\tau^{2\gamma} + \rho(nM\tau)).$$

In fact, Proposition 9.1 is a straightforward corollary of Proposition 9.2 below, 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 (47) for the macro-scheme applied to the averaged SPDE (14),
- the sampling error (41) between the invariant distributions  $\mu$  and  $\mu^\tau$ , which gives an error estimate  $\sup_{0 \leq n\Delta t \leq T} (\mathbb{E}|\bar{X}_n - \bar{X}_n^\tau|^2)^{\frac{1}{2}} \leq C_\gamma \tau^{2\gamma}$  by a straightforward Gronwall type argument.

Note that (50) is a straightforward consequence of (41) and (43).

**Proposition 9.2.** *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  $1 \leq M_a \leq M$ , one has*

$$(51) \quad \sup_{0 \leq n\Delta t \leq T} (\mathbb{E}|X_n - X_n^\tau|_{L^2}^2)^{\frac{1}{2}} \leq C(T) \left( \sqrt{R_1(M, M_a, \tau)} + \sqrt{\Delta t} \left( \frac{1}{\sqrt{M_a}} + \sqrt{R_2(M, M_a, \tau)} \right) \right),$$

and, for all test functions  $\varphi$  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  $1 \leq M_a \leq M$ , one has

$$(52) \quad \sup_{0 \leq n\Delta t \leq T} |\mathbb{E}[\varphi(X_n)] - \mathbb{E}[\varphi(\bar{X}_n^\tau)]| \leq C(T, \varphi) \left( R_1(M, M_a, \tau) + \Delta t \left( \frac{1}{M_a} + R_2(M, M_a, \tau) \right) \right).$$

Observe that Proposition 9.2 implies the convergence of the macro-scheme (44) to the scheme (46), when  $M_a \rightarrow \infty$ , for any fixed values of  $\Delta t$  and  $\tau$ . 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  $\epsilon \rightarrow 0$ . The limit scheme (46) is not an integrator for the averaged equation (14), but to a modified equation, with a residual depending on the micro time-step size  $\tau$ .

A full analysis of the cost of the multiscale scheme (44)-(45), depending on parameters  $\Delta t$ ,  $\tau$ ,  $M$  and  $M_a$ , requires to balance the error terms in (48) and (49). We refer to [4].

To conclude this section, a sketch of proof of Proposition 9.2 is provided, see [4] for more details.

*Sketch of proof of Proposition 9.2.* To deal with the strong error estimate (51), note that

$$\begin{aligned} X_n - \bar{X}_n^\tau &= \Delta t \sum_{k=0}^{n-1} S_{\Delta t}^{n-k} (\tilde{F}_k - \bar{F}^\tau(\bar{X}_k^\tau)) \\ &= \Delta t \sum_{k=0}^{n-1} S_{\Delta t}^{n-k} (\bar{F}^\tau(X_k) - \bar{F}^\tau(\bar{X}_k^\tau)) + \Delta t \sum_{k=0}^{n-1} S_{\Delta t}^{n-k} (\tilde{F}_k - \bar{F}^\tau(X_k^\tau)). \end{aligned}$$

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

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

On the other hand, a straightforward expansion yields

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

An expansion of the average which defines  $\tilde{F}_k$  yields

$$\begin{aligned} \mathcal{E}_1 &= \frac{\Delta t^2}{M_a^2} \sum_{k=0}^{n-1} \sum_{m=M-M_a+1}^M \mathbb{E}|F(X_k, Y_{k,m}^\tau) - \bar{F}^\tau(X_k)|^2 \\ &\quad + \frac{2\Delta t^2}{M_a^2} \sum_{k=0}^{n-1} \sum_{M-M_a+1 \leq m_1 < m_2 \leq M} \mathbb{E}\langle F(X_k, Y_{k,m_1}^\tau) - \bar{F}^\tau(X_k), F(X_k, Y_{k,m_2}^\tau) - \bar{F}^\tau(X_k) \rangle \\ &\leq C \frac{\Delta t}{M_a} + C \frac{\Delta t}{M_a^2} \sum_{M-M_a+1 \leq m_1 < m_2 \leq M} \rho((m_2 - m_1)\tau), \end{aligned}$$

thanks to a conditioning argument. In addition, using another conditioning argument, one gets

$$|\mathcal{E}_2| \leq C\Delta t^2 \sum_{0 \leq k_1 < k_2 \leq n-1} \frac{1}{M_a} \sum_{m=M-M_a+1}^M \rho(m\tau) \leq CR_1(M, M_a, \tau).$$

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

The treatment of the weak error estimate (52) requires to introduce the auxiliary function  $\bar{u}^\tau$  as follows: for all  $n \in \mathbb{N}_0$  and  $x \in L^2$ ,

$$\bar{u}^\tau(n, x) = \mathbb{E}[\varphi(\bar{X}_n^\tau) | \bar{X}_0^\tau = x].$$

The weak error is then written as a telescoping sum

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

In addition, using Markov property and a second-order Taylor expansion, one obtains

$$\begin{aligned} \mathbb{E}[\bar{u}^\tau(n-k, X_k)] - \mathbb{E}[\bar{u}^\tau(n-k-1, X_{k+1})] &= \mathbb{E}[\bar{u}^\tau(n-k-1, S_{\Delta t}(X_k + \Delta t \bar{F}^\tau(X_k) + \Delta W_k^Q))] \\ &\quad - \mathbb{E}[\bar{u}^\tau(n-k-1, S_{\Delta t}(X_k + \Delta t \tilde{F}_k + \Delta W_k^Q))] \\ &= \Delta t \mathbb{E}[D_x \bar{u}^\tau(n-k-1, S_{\Delta t}(X_k + \Delta t \bar{F}^\tau(X_k) + \Delta W_k^Q)) \cdot (S_{\Delta t} \bar{F}^\tau(X_k) - S_{\Delta t} \tilde{F}_k)] \\ &\quad + O(\Delta t^2) \mathbb{E}|\bar{F}^\tau(X_k) - \tilde{F}_k|^2. \end{aligned}$$

Note that  $O(\Delta t^2) \mathbb{E}|\bar{F}^\tau(X_k) - \tilde{F}_k|^2$  is treated as  $\mathcal{E}_1$  above.

The remaining error term is interpreted in terms of the auxiliary function

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

as

$$\frac{\Delta t}{M_a} \sum_{m=M-M_a+1}^M (\mathbb{E}[\Psi(n-k-1, X_k, Y_{k,m})] - \int \Psi(n-k-1, X_k, \cdot) d\mu^\tau).$$

Studying regularity properties of  $\Psi(n - k - 1, x, \cdot)$ , one then concludes using (43).

Finally,

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

and it remains to sum from  $k = 0$  to  $k = n - 1$  to conclude the proof of Proposition 9.2.  $\square$

#### ACKNOWLEDGMENTS

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, and T. Tangarife. Kinetic theory of jet dynamics in the stochastic barotropic and 2D Navier-Stokes equations. *J. Stat. Phys.*, 153(4):572–625, 2013.
- [2] F. Bouchet, C. Nardini, and T. Tangarife. Stochastic averaging, large deviations and random transitions for the dynamics of 2D and geostrophic turbulent vortices. *Fluid Dyn. Res.*, 46(6):061416, 11, 2014.
- [3] C.-E. Bréhier. Strong and weak orders in averaging for SPDEs. *Stochastic Process. Appl.*, 122(7):2553–2593, 2012.
- [4] C.-E. Bréhier. Analysis of an HMM time-discretization scheme for a system of stochastic PDEs. *SIAM J. Numer. Anal.*, 51(2):1185–1210, 2013.
- [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):1–40, 2014.
- [6] Z. Brzeźniak. On stochastic convolution in Banach spaces and applications. *Stochastics Stochastics Rep.*, 61(3-4):245–295, 1997.
- [7] C.-E. Bréhier and A. Debussche. Kolmogorov equations and weak order analysis for spdes with nonlinear diffusion coefficient. *Journal de Mathématiques Pures et Appliquées*, 2018.
- [8] S. Cerrai. A Khasminskii type averaging principle for stochastic reaction-diffusion equations. *Ann. Appl. Probab.*, 19(3):899–948, 2009.
- [9] S. Cerrai. Averaging principle for systems of reaction-diffusion equations with polynomial nonlinearities perturbed by multiplicative noise. *SIAM J. Math. Anal.*, 43(6):2482–2518, 2011.
- [10] S. Cerrai and M. Freidlin. Averaging principle for a class of stochastic reaction-diffusion equations. *Probab. Theory Related Fields*, 144(1-2):137–177, 2009.
- [11] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.
- [12] Z. Dong, X. Sun, H. Xiao, and J. Zhai. Averaging principle for one dimensional stochastic burgers equation. *Journal of Differential Equations*, 2018.
- [13] J.-P. Fouque, J. Garnier, G. Papanicolaou, and K. Sø Ina. *Wave propagation and time reversal in randomly layered media*, volume 56 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2007.
- [14] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, third edition, 2012. Translated from the 1979 Russian original by Joseph Szücs.
- [15] H. Fu and J. Duan. An averaging principle for two-scale stochastic partial differential equations. *Stoch. Dyn.*, 11(2-3):353–367, 2011.
- [16] H. Fu and J. Liu. Strong convergence in stochastic averaging principle for two time-scales stochastic partial differential equations. *J. Math. Anal. Appl.*, 384(1):70–86, 2011.
- [17] H. Fu, L. Wan, and J. Liu. Strong convergence in averaging principle for stochastic hyperbolic-parabolic equations with two time-scales. *Stochastic Process. Appl.*, 125(8):3255–3279, 2015.
- [18] H. Fu, L. Wan, J. Liu, and X. Liu. Weak order in averaging principle for stochastic wave equation with a fast oscillation. *Stochastic Process. Appl.*, 128(8):2557–2580, 2018.
- [19] H. Fu, L. Wan, J. Liu, and 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, Y. Wang, and J. Liu. Strong convergence rate in averaging principle for stochastic FitzHugh-Nagumo system with two time-scales. *J. Math. Anal. Appl.*, 416(2):609–628, 2014.
- [21] P. Gao. Averaging principle for the higher order nonlinear Schrödinger equation with a random fast oscillation. *J. Stat. Phys.*, 171(5):897–926, 2018.



- [22] P. Gao and Y. Li. Averaging principle for the Schrödinger equations. *Discrete Contin. Dyn. Syst. Ser. B*, 22(6):2147–2168, 2017.
- [23] T. Hytönen, J. van Neerven, M. Veraar, and L. Weis. *Analysis in Banach spaces. Vol. I. Martingales and Littlewood-Paley theory*, volume 63 of *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]*. Springer, Cham, 2016.
- [24] R. Z. Khasminskii. On the principle of averaging the Itô’s stochastic differential equations. *Kybernetika (Prague)*, 4:260–279, 1968.
- [25] R. Z. Khasminskii and G. Yin. On averaging principles: an asymptotic expansion approach. *SIAM J. Math. Anal.*, 35(6):1534–1560, 2004.
- [26] R. Z. Khasminskii and G. Yin. Limit behavior of two-time-scale diffusions revisited. *J. Differential Equations*, 212(1):85–113, 2005.
- [27] C. Kuehn. *Multiple time scale dynamics*, volume 191 of *Applied Mathematical Sciences*. Springer, Cham, 2015.
- [28] S. Li, X. Sun, Y. Xie, and Y. Zhao. Averaging principle for two dimensional stochastic navier-stokes equations. *arXiv preprint arXiv:1810.02282*, 2018.
- [29] D. Liu. Strong convergence of principle of averaging for multiscale stochastic dynamical systems. *Commun. Math. Sci.*, 8(4):999–1020, 2010.
- [30] G. A. Pavliotis and A. M. Stuart. *Multiscale methods*, volume 53 of *Texts in Applied Mathematics*. Springer, New York, 2008. Averaging and homogenization.
- [31] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.
- [32] H. Triebel. *Interpolation theory, function spaces, differential operators*. Johann Ambrosius Barth, Heidelberg, second edition, 1995.
- [33] J. M. A. M. van Neerven, M. C. Veraar, and L. Weis. Stochastic integration in UMD Banach spaces. *Ann. Probab.*, 35(4):1438–1478, 2007.
- [34] J. M. A. M. van Neerven, M. C. Veraar, and L. Weis. Stochastic evolution equations in UMD Banach spaces. *J. Funct. Anal.*, 255(4):940–993, 2008.
- [35] A. Y. Veretennikov. On an averaging principle for systems of stochastic differential equations. *Mat. Sb.*, 181(2):256–268, 1990.

UNIV LYON, CNRS, UNIVERSITÉ CLAUDE BERNARD LYON 1, UMR5208, INSTITUT CAMILLE JORDAN, F-69622 VILLEURBANNE, FRANCE

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