ASYMPTOTIC BEHAVIOR OF MULTISCALE STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

MICHAEL RÖCKNER, LONGJIE XIE AND LI YANG

ABSTRACT. In this paper, we study the asymptotic behavior of a semi-linear slow-fast stochastic partial differential equation with singular coefficients. Using the Poisson equation in Hilbert space, we first establish the strong convergence in the averaging principe, which can be viewed as a functional law of large numbers. Then we study the stochastic fluctuations between the original system and its averaged equation. We show that the normalized difference converges weakly to an Ornstein-Uhlenbeck type process, which can be viewed as a functional central limit theorem. Furthermore, rates of convergence both for the strong convergence and the normal deviation are obtained, and these convergence rates are shown not to depend on the regularity of the coefficients in the equation for the fast variable, which coincides with the intuition, since in the limit system the fast component has been totally averaged or homogenized out.

AMS 2010 Mathematics Subject Classification: 60H15, 70K65, 60F05.

Keywords and Phrases: Stochastic partial differential equations; averaging principle; normal deviations; Poisson equation in Hilbert space.

CONTENTS

This work is supported by the DFG through CRC 1283 and NSFC (No. 11701233, 11931004).

1. Introduction

Consider the following fully coupled slow-fast stochastic partial differential equation (SPDE for short) in $H_1 \times H_2$:

$$\left\{\begin{array}{ll} \mathrm{d}X_t^\varepsilon = AX_t^\varepsilon \mathrm{d}t + F(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \mathrm{d}W_t^1, & X_0^\varepsilon = x \in H_1, \\ \mathrm{d}Y_t^\varepsilon = \varepsilon^{-1}BY_t^\varepsilon \mathrm{d}t + \varepsilon^{-1}G(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \varepsilon^{-1/2}\mathrm{d}W_t^2, & Y_0^\varepsilon = y \in H_2, \end{array}\right. \tag{1.1}$$

where H_1, H_2 are two Hilbert spaces, $A : \mathcal{D}(A) \subset H_1 \to H_1$ and $B : \mathcal{D}(B) \subset H_2 \to H_2$ are linear operators, $F : H_1 \times H_2 \to H_1$ and $G : H_1 \times H_2 \to H_2$ are reaction coefficients, W_t^1 and W_t^2 are mutually independent H_1- and H_2-valued $(\mathscr{F}_t)$-Wiener processes both defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with a normal filtration $(\mathscr{F}_t)_{t\geqslant 0}$, and the small parameter $0 < \varepsilon \ll 1$ represents the separation of time scales between the slow process X_t^ε (which can be thought of as the mathematical model for a phenomenon appearing at the natural time scale) and the fast motion Y_t^ε (with time order $1/\varepsilon$, which can be interpreted as the fast environment). Such multi-scale models appear frequently in many real-world dynamical systems. Typical examples include climate weather interactions (see e.g. [32, 34]), macro-molecules (see e.g. [3, 29]), geophysical fluid flows (see e.g. [22]), stochastic volatility in finance (see e.g. [20]), etc. However, it is often too difficult to analyze or simulate the underlying system (1.1) directly due to the two widely separated time scales and the cross interactions between the slow and fast modes. Thus a simplified equation which governs the evolution of the system over a long time scale is highly desirable and is quite important for applications.

It is known that under suitable regularity assumptions on the coefficients, the slow process X_t^ε will converge strongly (in the $L^2(\Omega)$-sense) to the solution of the following reduced equation:

$$\mathrm{d}\bar{X}_t = A\bar{X}_t \mathrm{d}t + \bar{F}(\bar{X}_t)\mathrm{d}t + \mathrm{d}W_t^1, \quad \bar{X}_0 = x \in H_1, \tag{1.2}$$

where the averaged coefficient is given by

$$\bar{F}(x) := \int_{H_2} F(x, y)\mu^x(\mathrm{d}y), \tag{1.3}$$

and $\mu^x(\mathrm{d}y)$ is the unique invariant measure of the process Y_t^x, which is the solution of the frozen equation

$$\mathrm{d}Y_t^x = BY_t^x \mathrm{d}t + G(x, Y_t^x)\mathrm{d}t + \mathrm{d}W_t^2, \quad Y_0^x = y \in H_2. \tag{1.4}$$

The effective system (1.2) then captures the essential dynamics of the system (1.1), which does not depend on the fast variable any more and thus is much simpler than

SPDE (1.1). This theory, known as the averaging principle, was first developed for deterministic systems by Bogoliubov [8], and extended to stochastic differential equations (SDEs for short) by Khasminskii [30]. In the past decades, the averaging principle for systems with a finite number of degrees of freedom has been intensively studied, see e.g. [2, 24, 25, 31, 41] and the references therein. Passing from the finite dimensional to the infinite dimensional setting is more difficult, and the existing results in the literature are relatively few. In [13], Cerrai and Freidlin proved the averaging principle for slow-fast stochastic reaction-diffusion system where there is no noise in the slow equation. Later, Cerrai [10, 12] generalized this result to general reaction-diffusion equations with multiplicative noise and coefficients of polynomial growth , see also [14, 18, 21] for further developments. We also mention that in these results, no rates of convergence in terms of $\varepsilon \to 0$ are provided. But for numerical purposes, it is important to know the rate of convergence of the slow variable to the effective dynamic. The main motivation comes from the well-known Heterogeneous Multi-scale Methods used to approximate the slow component in system (1.1), see e.g. [5, 19, 33]. In this direction, Bréhier [4] first studied the rates of strong convergence for the averaging principle of SPDEs with noise only in the fast motion, and $(\frac{1}{2}\text{-})$-order of convergence is obtained. Extensions to general stochastic reaction-diffusion equations are made in [42], and $\frac{1}{2}$-order of convergence is obtained. For more recent results, we refer the interested readers to the work [7] and the references therein.

The strong convergence in the averaging principle can be viewed as a functional law of large numbers. Once we obtain the validity of the averaging principle, it is natural to go one step further to consider the functional central limit theorem. Namely, to study the small fluctuations of the original system (1.1) around its averaged equation (1.2). To leading order, these fluctuations can be captured by characterizing the asymptotic behavior of the normalized difference

$$Z_t^\varepsilon := \frac{X_t^\varepsilon - \bar{X}_t}{\sqrt{\varepsilon}} \tag{1.5}$$

as ε tends to 0. Under extra assumptions, the deviation process Z_t^ε is known to converge weakly (in the distribution sense) towards a Gaussian process $\bar{Z}_t$, whose covariance can be described explicitly. Such result, also known as the Gaussian approximation, is closely related to the homogenization for solutions of partial differential equations with singularly perturbed terms, which has its own interest in the theory of PDEs, see e.g. [26, 27]. For the study of normal deviations of multi-scale SDEs, we refer the readers to the fundamental paper by Khasminskii [30], see also [36, 37, 40] for further developments. In the infinite dimensional situation, as far as we know, there exist only two papers. Cerrai [11] studied the normal deviations for slow-fast SPDEs in a special case, i.e., a deterministic reaction-diffusion equation with one dimensional space variable perturbed by a fast motion. Later, this was generalized to general stochastic reaction-diffusion equations by Wang and Roberts [42]. In both papers the methods of proof are based on the time discretisation procedure which involve some complicated tightness arguments. We point out that besides having intrinsic interest, the functional central limit theorem is also useful in applications. In particular, we can get the formal asymptotic expansion

$$X_t^\varepsilon \overset{\mathcal{D}}{\approx} \bar{X}_t + \sqrt{\varepsilon}\bar{Z}_t,$$

where $\overset{\mathcal{D}}{\approx}$ means approximate equality of probability distributions. Such expansion has been introduced in the context of stochastic climate models. In physics this is also called the Van Kampen's scheme (see e.g. [1, 28]), which provides better approximations for the original system (1.1).

In the present paper, we shall first establish a stronger convergence result in the averaging principle for SPDE (1.1). More precisely, we show that for any $T > 0, q \geqslant 1$ and $\gamma \in [0, 1/2)$, there exists a constant $C_T > 0$ such that

$$\sup_{t\in[0,T]} \mathbb{E}\|(-A)^\gamma(X_t^\varepsilon - \bar{X}_t)\|^q \leqslant C_T\, \varepsilon^{\frac{q}{2}},$$

see **Theorem 2.2** below. Compared with the existing results in the literature, we assume that the coefficients are only Hölder continuous with respect to the fast variable, and we obtain not only the strong convergence in $L^q(\Omega)$-sense with any $q \geqslant 1$, but also in $\|\cdot\|_{(-A)^\gamma}$ norm with any $\gamma \in [0, 1/2)$, which is particularly interesting for SPDEs in comparison with the finite dimensional setting since A is an unbounded operator. Moreover, the $\frac{1}{2}$-order rate of convergence is also obtained, which is known to be optimal (when $\gamma = 0$). In particular, we show that the convergence rate does not depend on the regularity of the coefficients with respect to the fast variable. This coincides with the intuition, since in the limit equation the fast component has been totally averaged out. We point out that the strong convergence of $(-A)^\gamma X_t^\varepsilon$ to $(-A)^\gamma \bar{X}_t$ will play an important role below in our study of the homogenization behavior for the normalized difference Z_t^ε. Furthermore, the index $\gamma < 1/2$ should be the best possible, see Remark 2.3 for more detailed explanations.

The argument we shall use to establish the above strong convergence is different from those in [4, 10, 12, 13, 14, 18, 21, 42], where the classical Khasminskii's time discretisation procedure is used. Our method is based on the Poisson equation. More precisely, consider the following Poisson equation in the Hilbert space $H_1 \times H_2$:

$$\mathcal{L}_2(x, y)\psi(x, y) = -\phi(x, y), \quad y \in H_2, \tag{1.6}$$

where $\mathcal{L}_2(x, y)$ is an ergodic elliptic operator with respect to the y variable (see (2.6) below), $x \in H_1$ is regarded as a parameter, and $\phi : H_1 \times H_2 \to \mathbb{R}$ is a measurable function. Such kind of equation, i.e., with a parameter and in the whole space (without boundary condition), has been studied only relatively recently and is now realized to be very important in the theory of limit theorems in probability theory and numerical approximation for time-averaging estimators and invariant measures, see e.g. [35, 39]. In the finite dimensional situation, equations of the form (1.6) have been studied in a series of papers by Pardoux and Veretennikov [36, 37, 38], see also [40] and the references therein for further developments. Undoubtedly, extension to the infinite dimensional setting will be more difficult due to the unboundedness of the involved linear operators. In the recent work [7], the author studies the rate of convergence in the averaging principle for slow-fast SPDEs with regular coefficients by assuming the solvability of the corresponding Poisson equation as well as regularity properties of the solutions. In addition, the SPDE considered therein is not fully coupled, i.e., the fast component Y_t^ε does not depend on the slow process X_t^ε, and the two Hilbert spaces H_1, H_2 and the unbounded operators A, B are assumed to be the same, which are used in the whole proof in an essential way. Here, we shall establish the well-posedness of the Poisson equation (1.6) with only Hölder coefficients and in general Hilbert spaces $H_1 \times H_2$, and study the regularity properties of the unique solution with respect to both the y-variable and the parameter x, see **Theorem 3.2** below, which should be of independent interest. Then, we use the Poisson equation to derive a strong fluctuation estimate (see Lemma 4.2) for an integral functional of the slow-fast SPDE (1.1). The strong convergence in the averaging principle with optimal rate of convergence then follows directly. In addition, we also provide a simple way to verify the regularity of the averaged coefficients by using Theorem 3.2 (see Lemma 3.7 below), which is a separate problem that one always encounters in the study of averaging principles, central limit theorems, homogenization and other limit theorems.

Next, we proceed to study the small fluctuations of the slow process X_t^ε around its averaged motion $\bar{X}_t$. Recall that Z_t^ε is defined by (1.5). In view of (1.1) and (1.2), we have

$$\begin{aligned} \mathrm{d}Z_t^\varepsilon &= AZ_t^\varepsilon \mathrm{d}t + \frac{1}{\sqrt{\varepsilon}}\Big[F(X_t^\varepsilon, Y_t^\varepsilon) - \bar{F}(\bar{X}_t)\Big]\mathrm{d}t \\ &= AZ_t^\varepsilon \mathrm{d}t + \frac{1}{\sqrt{\varepsilon}}\Big[\bar{F}(X_t^\varepsilon) - \bar{F}(\bar{X}_t)\Big]\mathrm{d}t + \frac{1}{\sqrt{\varepsilon}}\delta F(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t, \end{aligned} \tag{1.7}$$

where

$$\delta F(x,y) := F(x,y) - \bar{F}(x). \tag{1.8}$$

We demonstrate that Z_t^ε converges weakly to an Ornstein-Uhlenbeck type process $\bar{Z}_t$ which satisfies the following linear SPDE:

$$\mathrm{d}\bar{Z}_t = A\bar{Z}_t\mathrm{d}t + D_x\bar{F}(\bar{X}_t).\bar{Z}_t\mathrm{d}t + \sigma(\bar{X}_t)\mathrm{d}\tilde{W}_t,$$

where $\tilde{W}_t$ is another cylindrical Wiener process which is independent of W_t^1, and the diffusion coefficient σ is Hilbert-Schmidt operator valued and given by (2.8), see **Theorem 2.4** below. Compared with [11, 42], our system (1.1) is more general, and the coefficients are assumed to be only Hölder continuous with respect to the fast variable, and we provide a more precise formula for the new diffusion coefficient σ. Moreover, the arguments we use to prove the above convergence is different from [11, 42], and in addition the rate of convergence is obtained, which does not depends on the regularity of the coefficients with respect to the fast variable.

It turns out that our method to prove the above functional central limit theorem is closely and universally connected with the proof of the strong convergence in the averaging principle. Namely, we shall first use the result on the Poisson equation (1.6) established in Theorem 3.2 to derive some weak fluctuation estimates (see Lemma 5.5) for an integral functional involving the processes $(X_t^\varepsilon, Y_t^\varepsilon)$ and Z_t^ε. Combining with the Kolmogorov equation associated with the process $(\bar{X}_t, \bar{Z}_t)$, we prove the weak convergence of Z_t^ε to $\bar{Z}_t$ directly. Here, we note that the whole system of equations satisfied by $(\bar{X}_t, \bar{Z}_t)$ is an SPDE with multiplicative noise. Even though infinite dimensional Kolmogorov equations with nonlinear diffusion coefficients have been studied very recently in [6], the regularity of the solutions obtained therein are not sufficient for our purpose. Thus, we derive some better regularity for the solution with respect to the z variable (see Theorem 5.1 below), and develop a trick in the proof of Theorem 2.4 to avoid using the regularity for the solution with respect to the x variable. It will be quite easy to capture the structure of the homogenization limit $\bar{Z}_t$ from our proof. Furthermore, our approach can also be adapted to study the normal deviations for other classes of multi-scale SPDEs. We shall address these problems in future works.

The rest of this paper is organized as follows. In Section 2, we introduce some assumptions and state our main results. Section 3 is devoted to study the Poisson equation in Hilbert spaces. Then, we prove the strong convergence result, Theorem 2.2, and the normal deviation result, Theorem 2.4, in Section 4 and Section 5, respectively. Finally, in the Appendix we prove some necessary estimates for the solution of the multiscale system (1.1), which are slight generalizations of the existing results in the literature. Throughout this paper, the letter C with or without subscripts will denote a positive constant, whose value may change in different places, and whose dependence on parameters can be traced from the calculations.

Notations: To end this section, we introduce some notations, which will be used throughout this paper. Let H_1, H_2 and H be three Hilbert spaces endowed with the scalar products $\langle\cdot,\cdot\rangle_1$, $\langle\cdot,\cdot\rangle_2$ and $\langle\cdot,\cdot\rangle_H$, respectively. The corresponding norms will be denoted by $\|\cdot\|_1, \|\cdot\|_2$ and $\|\cdot\|_H$. We use $\mathscr{L}(H_1, H_2)$ to denote the space of all linear and bounded operators from H_1 to H_2. If $H_1 = H_2$, we write $\mathscr{L}(H_1) = \mathscr{L}(H_1, H_1)$ for simplicity. Recall that an operator $Q \in \mathscr{L}(H)$ is called Hilbert-Schmidt if

$$\|Q\|^2_{\mathscr{L}_2(H)} := Tr(QQ^*) < +\infty.$$

We shall denote the space of all Hilbert-Schmidt operators on H by $\mathscr{L}_2(H)$.

For any $x \in H_1$, $y \in H_2$ and $\phi : H_1 \times H_2 \to H$, we say that ϕ is Gâteaux differentiable at x if there exists a $D_x\phi(x,y) \in \mathscr{L}(H_1, H)$ such that for all $h_1 \in H_1$,

$$\lim_{\tau\to 0} \frac{\phi(x+\tau h_1, y) - \phi(x,y)}{\tau} = D_x\phi(x,y).h_1.$$

If in addition

$$\lim_{\|h_1\|_1\to 0} \frac{\|\phi(x+h_1,y) - \phi(x,y) - D_x\phi(x,y).h_1\|_H}{\|h_1\|_1} = 0,$$

ϕ is called Fréchet differentiable at x. Similarly, for any $k \geqslant 2$ we can define the k times Gâteaux and Fréchet derivative of ϕ at x, and we will identify the higher order derivative $D_x^k\phi(x,y)$ with a linear operator in $\mathscr{L}^k(H_1, H) := \mathscr{L}(H_1, \mathscr{L}^{(k-1)}(H_1, H))$, endowed with the operator norm

$$\|D_x^k\phi(x,y)\|_{\mathscr{L}^k(H_1,H)} := \sup_{\|h_1\|_1\leqslant 1, \|h_2\|_1\leqslant 1, \cdots, \|h_k\|_1\leqslant 1, \|h\|_H\leqslant 1} \langle D_x^k\phi(x,y).(h_1,h_2,\cdots,h_k), h\rangle_H.$$

In the same way, we define the Gâteaux and Fréchet derivatives of ϕ with respect to the y variable, and we have $D_y\phi(x,y) \in \mathscr{L}(H_2, H)$, and for $k \geqslant 2$, $D_y^k\phi(x,y) \in \mathscr{L}^k(H_2, H) := \mathscr{L}(H_2, \mathscr{L}^{(k-1)}(H_2, H))$.

We will denote by $L_p^\infty(H_1 \times H_2, H)$ the space of all measurable maps $\phi : H_1 \times H_2 \to H$ with linear growth in x and polynomial growth in y, i.e., there exists a constant $p \geqslant 1$ such that

$$\|\phi\|_{L_p^\infty(H)} := \sup_{(x,y)\in H_1\times H_2} \frac{\|\phi(x,y)\|_H}{1 + \|x\|_1 + \|y\|_2^p} < \infty.$$

For $k \in \mathbb{N}$, the space $C_p^{k,0}(H_1 \times H_2, H)$ contains all maps $\phi \in L_p^\infty(H_1 \times H_2, H)$ which are k times Gâteaux differentiable at any $x \in H_1$ and

$$\|\phi\|_{C_p^{k,0}(H)} := \sup_{(x,y)\in H_1\times H_2} \frac{\sum_{\ell=1}^k \|D_x^\ell\phi(x,y)\|_{\mathscr{L}^\ell(H_1,H)}}{1 + \|y\|_2^p} < \infty.$$

Similarly, the space $C_p^{0,k}(H_1 \times H_2, H)$ consists of all maps $\phi \in L_p^\infty(H_1 \times H_2, H)$ which are k times Gâteaux differentiable at any $y \in H_2$ and

$$\|\phi\|_{C_p^{0,k}(H)} := \sup_{(x,y)\in H_1\times H_2} \frac{\sum_{\ell=1}^k \|D_y^\ell\phi(x,y)\|_{\mathscr{L}^\ell(H_2,H)}}{1 + \|x\|_1 + \|y\|_2^p} < \infty. \tag{1.9}$$

We also introduce the space $\mathbb{C}_p^{0,k}(H_1 \times H_2, H)$ consisting of all maps which are k times Fréchet differentiable at any $y \in H_2$ and satisfies (1.9). For $k, \ell \in \mathbb{N}$, let $C_p^{k,\ell}(H_1 \times H_2, H)$ be the space of all maps satisfying

$$\|\phi\|_{C_p^{k,\ell}(H)} := \|\phi\|_{L_p^\infty(H)} + \|\phi\|_{C_p^{k,0}(H)} + \|\phi\|_{C_p^{0,\ell}(H)} < \infty,$$

and for $\eta \in (0,1)$, we use $C_p^{k,\eta}(H_1 \times H_2, H)$ to denote the subspace of $C_p^{k,0}(H_1 \times H_2, H)$ consisting of all maps such that

$$\|\phi(x,y_1) - \phi(x,y_2)\|_H \leqslant C_0\|y_1 - y_2\|_2^\eta \big(1 + \|x\|_1 + \|y_1\|_2^p + \|y_2\|_2^p\big).$$

When the subscript p is replaced by b in the notations for above spaces, we mean that the map itself and its derivatives are all bounded. When $H = \mathbb{R}$, we will omit the letter H in the above notations for simplicity.

2. Statement of the main results

2.1. **Assumptions.** For $i = 1, 2$, let $\{e_{i,n}\}_{n\in\mathbb{N}}$ be a complete orthonormal basis of H_i. We assume that the two unbounded linear operators A and B, with domains $\mathcal{D}(A)$ and $\mathcal{D}(B)$, satisfy the following condition:

(A1): There exist non-decreasing sequences of real positive numbers $\{\alpha_n\}_{n\in\mathbb{N}}$ and $\{\beta_n\}_{n\in\mathbb{N}}$
such that

$$Ae_{1,n} = -\alpha_n e_{1,n}, \quad Be_{2,n} = -\beta_n e_{2,n}, \quad \forall n \in \mathbb{N}.$$

In this setting, the powers of $-A$ and $-B$ can be easily defined as follows: for any $\theta \in [0,1]$,

$$(-A)^\theta x := \sum_{n\in\mathbb{N}} \alpha_n^\theta \langle x, e_{1,n}\rangle_1 e_{1,n} \quad \text{and} \quad (-B)^\theta y := \sum_{n\in\mathbb{N}} \beta_n^\theta \langle y, e_{2,n}\rangle_2 e_{2,n},$$

with domains

$$\mathcal{D}((-A)^\theta) := \Big\{x \in H_1 : \|x\|_{(-A)^\theta}^2 := \sum_{n\in\mathbb{N}} \alpha_n^{2\theta}\langle x, e_{1,n}\rangle_1^2 < \infty\Big\}$$

and

$$\mathcal{D}((-B)^\theta) := \Big\{y \in H_2 : \|y\|_{(-B)^\theta}^2 := \sum_{n\in\mathbb{N}} \beta_n^{2\theta}\langle y, e_{2,n}\rangle_2^2 < \infty\Big\}.$$

Moreover, the corresponding semigroups $\{e^{tA}\}_{t\geqslant 0}$ and $\{e^{tB}\}_{t\geqslant 0}$ can be defined through the following spectral formulas: for any $t \geqslant 0$, $x \in H_1$ and $y \in H_2$,

$$e^{tA}x := \sum_{n\in\mathbb{N}} e^{-\alpha_n t}\langle x, e_{1,n}\rangle_1\, e_{1,n} \quad \text{and} \quad e^{tB}y := \sum_{n\in\mathbb{N}} e^{-\beta_n t}\langle y, e_{2,n}\rangle_2\, e_{2,n}.$$

The following regularization properties for these semigroups are more or less standard. We write them for e^{tA}, but they also hold for e^{tB}.

Proposition 2.1. *Let $\gamma \in [0,1]$ and $\theta \in [0,\gamma]$. We have:*

(i) For any $t > 0$ and $x \in \mathcal{D}((-A)^\theta)$,

$$\|e^{tA}x\|_{(-A)^\gamma} \leqslant C_{\gamma,\theta} t^{-\gamma+\theta} e^{-\frac{\alpha_1}{2}t}\|x\|_{(-A)^\theta}; \tag{2.1}$$

(ii) For any $0 \leqslant s \leqslant t$ and $x \in H_1$,

$$\|e^{tA}x - e^{sA}x\|_1 \leqslant C_\gamma (t-s)^\gamma \|e^{sA}x\|_{(-A)^\gamma}; \tag{2.2}$$

(iii) For any $0 < s \leqslant t$ and $x \in \mathcal{D}((-A)^\theta)$,

$$\|e^{tA}x - e^{sA}x\|_1 \leqslant C_{\gamma,\theta} \frac{(t-s)^\gamma}{s^{\gamma-\theta}} e^{-\frac{\alpha_1}{2}s} \|x\|_{(-A)^\theta}, \tag{2.3}$$

where α_1 is the smallest eigenvalue of A, and $C_\gamma, C_{\gamma,\theta} > 0$ are constants.

Proof. For any $t > 0$ and $x \in \mathcal{D}((-A)^\theta)$, we have

$$\begin{aligned}
\|(-A)^\gamma e^{tA}x\|_1^2 &= \Big\| \sum_{n\in\mathbb{N}} \alpha_n^\gamma e^{-\alpha_n t} \langle x, e_{1,n}\rangle_1 e_{1,n} \Big\|_1^2 \\
&= \sum_{n\in\mathbb{N}} \alpha_n^{2(\gamma-\theta)} e^{-2\alpha_n t} \alpha_n^{2\theta} |\langle x, e_{1,n}\rangle_1|^2 \\
&\leqslant C_{\gamma,\theta} t^{-2(\gamma-\theta)} e^{-\alpha_1 t/2} \sum_{n\in\mathbb{N}} \alpha_n^{2\theta} |\langle x, e_{1,n}\rangle_1|^2 = C_{\theta,\gamma} t^{-2(\gamma-\theta)} e^{-\alpha_1 t/2} \|x\|_{(-A)^\theta}^2,
\end{aligned}$$

which yields (2.1). To show estimate (2.2), by the basic inequality that $1 - e^{-\alpha t} \leqslant C_{\alpha,\gamma} t^\gamma$ with $\gamma \in [0,1], \alpha > 0$, we deduce that for any $0 \leqslant s \leqslant t$ and $x \in H_1$,

$$\begin{aligned}
\|e^{tA}x - e^{sA}x\|_1^2 &= \Big\| \sum_{n\in\mathbb{N}} (e^{-\alpha_n t} - e^{-\alpha_n s}) \langle x, e_{1,n}\rangle_1 e_{1,n} \Big\|_1^2 \\
&\leqslant C_\gamma \sum_{n\in\mathbb{N}} \alpha_n^{2\gamma} (t-s)^{2\gamma} e^{-2\alpha_n s} |\langle x, e_{1,n}\rangle_1|^2 \\
&= C_\gamma (t-s)^{2\gamma} |\langle (-A)^\gamma e^{sA}x, e_{1,n}\rangle_1|^2 \\
&= C_\gamma (t-s)^{2\gamma} \|e^{sA}x\|_{(-A)^\gamma}^2.
\end{aligned}$$

Combining (2.1) and (2.2), we immediately get (2.3). □

For $i = 1, 2$, let Q_i be two linear self-adjoint operators on H_i with positive eigenvalues $\{\lambda_{i,n}\}_{n\in\mathbb{N}}$, i.e.,

$$Q_i e_{i,n} = \lambda_{i,n} e_{i,n}, \quad \forall n \in \mathbb{N}.$$

Let $W_t^i, i = 1, 2$, be H_i-valued Q_i-Wiener processes both defined on a complete filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$. Then it is known that W_t^i can be written as

$$W_t^i = \sum_{n\in\mathbb{N}} \sqrt{\lambda_{i,n}} \beta_{i,n}(t) e_{i,n},$$

where $\{\beta_{i,n}\}_{n\in\mathbb{N}}$ are mutual independent real-valued Brownian motions. We shall assume that:

(A2): For $i = 1, 2$,

$$Tr(Q_i) := \sum_{n\in\mathbb{N}} \lambda_{i,n} < +\infty \quad \text{and} \quad Tr((-A)Q_1) < +\infty.$$

2.2. **Main results.** The first main result of this paper is about the strong convergence in the averaging principle for SPDE (1.1).

Theorem 2.2 (Strong convergence). *Let $T > 0$, $x \in \mathcal{D}((-A)^\theta)$ and $y \in \mathcal{D}((-B)^\theta)$ with $\theta > 0$. Assume that (**A1**) and (**A2**) hold, $F \in C_p^{2,\eta}(H_1 \times H_2, H_1)$ and $G \in C_b^{2,\eta}(H_1 \times H_2, H_2)$ with $\eta > 0$. Then for any $q \geqslant 1$ and $\gamma \in [0, \theta \wedge 1/2)$, we have*

$$\sup_{t\in[0,T]} \mathbb{E}\|X_t^\varepsilon - \bar{X}_t\|_{(-A)^\gamma}^q \leqslant C_1\, \varepsilon^{\frac{q}{2}}, \tag{2.4}$$

where $\bar{X}_t$ solves equation (1.2), and $C_1 = C(T, x, y) > 0$ is a constant independent of η and ε.

To compare our result with previous work in the literature, we make the following comments:

Remark 2.3. *(i) When $\gamma = 0$ in (2.4), the $1/2$-order rate of convergence in the $L^2(\Omega)$-sense is known to be optimal, which is the same as in the SDE case. However, the convergence in $\|\cdot\|_{(-A)^\gamma}$ norm seems to have never been studied before. This is particularly interesting for SPDEs since A is in general an unbounded operator.*

(ii) Note that the coefficients are assumed to be only η-Hölder continuous with respect to the fast variable, and the convergence rate does not dependent on η. This indicates that the convergence in the averaging principle does not depend on the regularity of the coefficients with respect to the fast variable, which coincides with the intuition, since in the limit equation the fast component has been totally averaged out.

(iii) Let us explain why $\gamma < 1/2$ should be the best possible. In fact, the main reason is that the processes X_t^ε and Y_t^ε are only γ-Hölder continuous with respect to the time variable with $\gamma < 1/2$. From another point of view, for Z_t^ε given by (1.5), estimate (2.4) means that for every $t \geqslant 0$, we have

$$\sup_{\varepsilon\in(0,1)} \mathbb{E}\|(-A)^\gamma Z_t^\varepsilon\|_1^2 < \infty.$$

But by Theorem 2.4 below, we have that Z_t^ε converges to $\bar{Z}_t$ with $\bar{Z}_t$ satisfying (2.7). Through straightforward computations we find that $\mathbb{E}\|(-A)^\gamma \bar{Z}_t\|_1^2 < \infty$ only when $\gamma < 1/2$.

Recall that Z_t^ε is defined by (1.5). To study the homogenization for Z_t^ε, we need to consider the following Poisson equation:

$$\mathcal{L}_2(x, y)\Psi(x, y) = -\delta F(x, y), \tag{2.5}$$

where δF is given by (1.8), and $\mathcal{L}_2(x,y)$ is defined by

$$\mathcal{L}_2\varphi(x,y) := \mathcal{L}_2(x,y)\varphi(x,y) := \langle By + G(x,y), D_y\varphi(x,y)\rangle_2 + \frac{1}{2}Tr\big[D_y^2\varphi(x,y)Q_2\big], \quad \forall \varphi \in C_p^{0,2}(H_1 \times H_2). \tag{2.6}$$

According to Theorem 3.2 and Remark 3.3 below, there exists a unique solution Ψ to equation (2.5). It turns out that the limit $\bar{Z}_t$ of Z_t^ε satisfies the following linear equation:

$$\mathrm{d}\bar{Z}_t = A\bar{Z}_t\mathrm{d}t + D_x\bar{F}(\bar{X}_t).\bar{Z}_t\mathrm{d}t + \sigma(\bar{X}_t)\mathrm{d}\tilde{W}_t, \quad \bar{Z}_0 = 0, \tag{2.7}$$

where $\tilde{W}_t$ is a cylindrical Wiener process in H_1 which is independent of W_t^1, and $\sigma : H_1 \to \mathscr{L}(H_1)$ satisfies

$$\frac{1}{2}\sigma(x)\sigma^*(x) = \overline{\delta F \otimes \Psi}(x) := \int_{H_2} \big[\delta F(x,y) \otimes \Psi(x,y)\big]\mu^x(\mathrm{d}y). \tag{2.8}$$

The following is the second main result of this paper.

Theorem 2.4 (Normal deviations)**.** *Let $T > 0$, $x \in \mathcal{D}((-A)^\theta)$ and $y \in \mathcal{D}((-B)^\theta)$ with $\theta > 0$. Assume that (**A1**) and (**A2**) hold, $F \in C_p^{2,\eta}(H_1 \times H_2, H_1)$ and $G \in C_b^{2,\eta}(H_1 \times H_2, H_2)$ with $\eta > 0$. Then for any $\gamma \in (0, 1/2)$ and $\varphi \in \mathbb{C}_b^4(H_1)$, we have*

$$\sup_{t\in[0,T]} \big|\mathbb{E}[\varphi(Z_t^\varepsilon)] - \mathbb{E}[\varphi(\bar{Z}_t)]\big| \leqslant C_2\, \varepsilon^{\frac{1}{2}-\gamma},$$

where $C_2 = C(T, x, y, \varphi) > 0$ is a constant independent of η and ε.

Remark 2.5. *Note that we claim that $\tilde{W}_t$ in (2.7) is independent of W_t^1. The advantage of formula (2.8) is that we can study the regularity properties of σ directly by using the result of the Poisson equation established in Theorem 3.2 below. Furthermore, one can check that $\sigma(x)$ is a Hilbert-Schmidt operator. In fact, by Theorem 3.2 we have*

$$\begin{aligned}
\|\sigma(x)\|^2_{\mathscr{L}_2(H_1)} &= \sum_{n\in\mathbb{N}} \langle \sigma(x)\sigma^*(x)e_{1,n}, e_{1,n}\rangle_1 \\
&= 2\sum_{n\in\mathbb{N}} \langle \int_{H_2} \big[\delta F(x,y) \otimes \Psi(x,y)\big]\mu^x(\mathrm{d}y)e_{1,n}, e_{1,n}\rangle_1 \\
&= 2\int_{H_2} \langle \delta F(x,y), \Psi(x,y)\rangle_1 \mu^x(\mathrm{d}y) \\
&\leqslant C_0 \int_{H_2} (1 + \|y\|_2^{2p})\mu^x(\mathrm{d}y) < \infty.
\end{aligned}$$

Thus, the stochastic integral part in (2.7) is well-defined.

3. Poisson Equation in Hilbert space

Consider the following Poisson equation in the infinite dimensional Hilbert space H_2:

$$\mathcal{L}_2(x,y)\psi(x,y) = -\phi(x,y), \tag{3.1}$$

where $\mathcal{L}_2(x,y)$ is defined by (2.6), $x \in H_1$ is regarded as a parameter, and $\phi : H_1 \times H_2 \to \mathbb{R}$ is a Borel-measurable function. Recall that $Y_t^x(y)$ satisfies the frozen equation (1.4) and $\mu^x(\mathrm{d}y)$ is the invariant measure of $Y_t^x(y)$ (see Lemma 3.4 below). Since we are considering (3.1) on the whole space and not on a compact subset, it is necessary to make the following "centering" assumption on ϕ:

$$\int_{H_2} \phi(x,y)\mu^x(\mathrm{d}y) = 0, \quad \forall x \in H_1. \tag{3.2}$$

Such kind of assumption is also natural and analogous to the centering condition in the standard central limit theorem, see e.g. [36, 37].

We first introduce the following definition of solutions for equation (3.1).

Definition 3.1. *A measurable function $\psi : H_1 \times H_2 \to \mathbb{R}$ is said to be a classical solution to equation (3.1) if:*

(i) the function $\psi(x,y) \in C_p^{0,2}(H_1 \times H_2)$ and for any $(x,y) \in H_1 \times H_2$, the operator $D_y^2\psi(x,y) \in \mathscr{L}(H_2)$;

(ii) for any $x \in H_1$ and $y \in \mathcal{D}(B)$, the function ψ satisfies equation (3.1).

The main aim of this section is to prove the following result.

Theorem 3.2. *Let $\eta > 0$ and $k = 0, 1, 2$. Assume that* **(A1)**, **(A2)** *hold, and $G \in C_b^{k,\eta}(H_1 \times H_2, H_2)$. Then for every $\phi \in C_p^{k,\eta}(H_1 \times H_2)$ satisfying (3.2), there exists a unique classical solution $\psi \in C_p^{k,0}(H_1 \times H_2) \cap \mathbb{C}_p^{0,2}(H_1 \times H_2)$ to equation (3.1) satisfying (3.2), which is given by*

$$\psi(x,y) = \int_0^\infty \mathbb{E}\big[\phi(x, Y_t^x(y))\big]\mathrm{d}t, \tag{3.3}$$

where $Y_t^x(y)$ satisfies the frozen equation (1.4).

Remark 3.3. *We can also solve the Poisson equation (3.1) for Hilbert space valued function $\tilde{\phi}$, i.e., $\tilde{\phi} : H_1 \times H_2 \to H$ with H being another Hilbert space. In fact, let $\{e_n\}_{n\in\mathbb{N}}$ be the orthonormal basis of H, and define*

$$\phi_n(x,y) := \langle \tilde{\phi}(x,y), e_n \rangle_H.$$

Then for each $n \in \mathbb{N}$, we have $\phi_n : H_1 \times H_2 \to \mathbb{R}$, and thus there exists a solution $\psi_n : H_1 \times H_2 \to \mathbb{R}$ to the equation (3.1) with ϕ replaced by ϕ_n. Define a H-valued function by

$$\tilde{\psi}(x,y) := \sum_{n\in\mathbb{N}} \psi_n(x,y)e_n = \int_0^\infty \mathbb{E}[\tilde{\phi}(x, Y_t^x(y))]\mathrm{d}t.$$

Then $\tilde{\psi}$ solves

$$\mathcal{L}_2(x,y)\tilde{\psi}(x,y) = -\tilde{\phi}(x,y).$$

3.1. **Properties of the frozen transition semigroup.** Given $\phi : H_1 \times H_2 \to \mathbb{R}$, let

$$T_t\phi(x,y) := \mathbb{E}\big[\phi(x, Y_t^x(y))\big],$$

In view of (3.3), we need to study the behavior of $T_t\phi$ as well as its first and second order derivatives with respect to the y variable both near $t = 0$ and as $t \to \infty$. Let us first collect the following estimates for $Y_t^x(y)$.

Lemma 3.4. *Assume* **(A1)**, **(A2)** *hold, and that $G \in C_b^{0,\eta}(H_1 \times H_2, H_2)$. Then there exists a unique mild solution $Y_t^x(y)$ to the equation (1.4). Moreover, we have:*

(i) For any $t \geqslant 0$ and $q \geqslant 1$, there exist constants $C_q, \lambda > 0$ such that

$$\mathbb{E}\|Y_t^x(y)\|_2^q \leqslant C_q\big(1 + e^{-\lambda t}\|y\|_2^q\big); \tag{3.4}$$

(ii) $Y_t^x(y)$ is strong Feller and irreducible;

(iii) There exist constants $C_0, \lambda > 0$ such that for any $t \geqslant 0$ and every $\phi \in L_p^\infty(H_1 \times H_2)$,

$$\left|T_t\phi(x,y) - \int_{H_2} \phi(x,z)\mu^x(\mathrm{d}z)\right| \leqslant C_0\|\phi\|_{L_p^\infty}(1 + \|x\|_1 + \|y\|_2^p)e^{-\lambda t}. \tag{3.5}$$

Proof. The existence and uniqueness of solutions to SPDE (1.4) with Hölder continuous coefficients follows from [17, Theorem 7]. We only need to verify that the assumptions 4, 5, 6 in [17] hold. To this end, let

$$Q_t := \int_0^t e^{sB}Q_2e^{sB^*}\mathrm{d}s \quad \text{and} \quad \Lambda_t = Q_t^{-1/2}e^{tB}.$$

Then we have

$$Tr(Q_t) = \sum_{n\in\mathbb{N}} \frac{\lambda_{2,n}}{2\beta_n}(1 - e^{-2\beta_n t}) \leqslant \sum_{n\in\mathbb{N}} \frac{\lambda_{2,n}}{2\beta_1} \leqslant C_0 Tr(Q_2) < +\infty,$$

and

$$\|\Lambda_t\|^2_{\mathscr{L}(H_2)} = \sup_{n\in\mathbb{N}} \frac{2\beta_n}{\lambda_{2,n}(e^{2\beta_n t} - 1)} \leqslant C_1 t^{-1},$$

which imply the desired result. Meanwhile, estimate (3.4) can be proved by following the same argument as in [9, Theorem 7.3], and the conclusions in (ii) follow by [15, Theorem 4 and Proposition 4]. Furthermore, for any $t \in [0,T]$ one can check that there exists a $\theta > 0$ such that

$$\mathbb{E}\|Y_t^x(y)\|_{(-B)^\theta} \leqslant C_T\big(1 + \|y\|_2^p\big).$$

For any $r, R > 0$, let $B_r := \{y \in H_2 : \|y\|_2 \leqslant r\}$ and $K = \{y \in H_2 : \|y\|_{(-B)^\theta} \leqslant R\}$. Then we have that for R large enough,

$$\inf_{y\in B_r} \mathbb{P}(Y_T^x(y) \in K) = \inf_{y\in B_r} \mathbb{P}(\|Y_T^x(y)\|_{(-B)^\theta} \leqslant R) = 1 - \sup_{y\in B_r} \mathbb{P}(\|Y_T^x(y)\|_{(-B)^\theta} > R)$$

$$\geqslant 1 - \sup_{y\in B_r} \frac{\mathbb{E}\|Y_T^{x,y}\|_{(-B)^\theta}}{R} \geqslant 1 - \frac{C_T(1+r^p)}{R} > 0.$$

Thus, estimate (3.5) follows by [23, Theorem 2.5]. □

Let P_t be the Ornstein-Uhlenbeck semigroup defined by

$$P_t\phi(x,y) := \mathbb{E}\big[\phi(x, R_t(y))\big],$$

where

$$\mathrm{d}R_t = BR_t\mathrm{d}t + \mathrm{d}W_t^2, \quad R_0 = y \in H_2.$$

The following result was proved by [17, Theorem 4] if ϕ is bounded and measurable. In view of (3.4), we can generalize it to ϕ with polynomial growth by following exactly the same argument. We omit the details here.

Lemma 3.5. *Assume* **(A1)**, **(A2)** *hold. Then for every $\phi \in L_p^\infty(H_1\times H_2)$ and $t \in (0,T]$, we have $P_t\phi(x,y) \in \mathbb{C}_p^{0,2}(H_1 \times H_2)$. Moreover,*

$$\|D_y P_t\phi(x,y)\|_2 \leqslant C_T \frac{1}{\sqrt{t}} \|\phi\|_{L_p^\infty}(1 + \|x\|_1 + \|y\|_2^p), \tag{3.6}$$

and for any $\eta \in [0,1]$,

$$\|D_y^2 P_t\phi(x,y)\|_{\mathscr{L}(H_2)} \leqslant C_T \frac{1}{t^{1-\eta/2}} \|\phi\|_{C_p^{0,\eta}}(1 + \|x\|_1 + \|y\|_2^p), \tag{3.7}$$

where $C_T > 0$ is a constant.

Based on Lemma 3.5, we have the following result.

Lemma 3.6. *Assume* **(A1)**, **(A2)** *hold, and that $G \in C_b^{0,\eta}(H_1 \times H_2, H_2)$. Then for every $\phi \in L_p^\infty(H_1 \times H_2)$ satisfying (3.2), we have $T_t\phi(x,y) \in \mathbb{C}_p^{0,1}(H_1 \times H_2)$ with*

$$|T_t\phi(x,y)| \leqslant C_0\|\phi\|_{L_p^\infty}(1 + \|x\|_1 + \|y\|_2^p)e^{-\lambda t} \tag{3.8}$$

and

$$\|D_y T_t\phi(x,y)\|_2 \leqslant C_0 \frac{1}{\sqrt{t}\wedge 1} \|\phi\|_{L_p^\infty}(1 + \|x\|_1 + \|y\|_2^p)e^{-\lambda t}, \tag{3.9}$$

where $C_0, \lambda > 0$ are constants independent of t. If we further assume that $\phi \in C_p^{0,\eta}(H_1 \times H_2)$ with $\eta \in (0,1)$, then $T_t\phi(x,y) \in \mathbb{C}_p^{0,2}(H_1 \times H_2)$ and

$$\|D_y^2 T_t\phi(x,y)\|_{\mathscr{L}(H_2)} \leqslant C_0 \frac{1}{t^{1-\eta/2}\wedge 1} \|\phi\|_{C_p^{0,\eta}}(1 + \|x\|_1 + \|y\|_2^p)e^{-\lambda t}. \tag{3.10}$$

Proof. Estimate (3.8) follows by (3.5) directly. The assertions that $T_t\phi(x,y) \in \mathbb{C}_p^{0,1}(H_1 \times H_2)$ and $T_t\phi(x,y) \in \mathbb{C}_p^{0,2}(H_1 \times H_2)$ can be obtained as in [17, Theorem 5]. Let us focus on the a-priori estimates (3.9) and (3.10). By Duhamel's formula (see e.g. [17, (16)]), for any $t > 0$ we have

$$T_t\phi(x,y) = P_t\phi(x,y) + \int_0^t P_{t-s}\langle G, D_yT_s\phi\rangle_2(x,y)\mathrm{d}s.$$

In view of (3.6) and by the assumption that G is bounded, we have for every $t \in (0,T]$,

$$\|D_yT_t\phi(x,y)\|_2 \leqslant C_0(1+\|x\|_1+\|y\|_2^p)\left(\frac{1}{\sqrt{t}}\|\phi\|_{L_p^\infty} + \int_0^t \frac{1}{\sqrt{t-s}}\|D_yT_s\phi(x,y)\|_2\mathrm{d}s\right).$$

By Gronwall's inequality we obtain

$$\|D_yT_t\phi(x,y)\|_2 \leqslant C_0\frac{1}{\sqrt{t}}\|\phi\|_{L_p^\infty}(1+\|x\|_1+\|y\|_2^p), \tag{3.11}$$

which means that (3.9) is true for $t \leqslant 2$. For $t > 2$, by the Markov property we have

$$T_t\phi(x,y) = \mathbb{E}\big[T_{t-1}\phi(x, Y_1^x(y))\big].$$

Using (3.8) and (3.11) with $t = 1$ and ϕ replaced by $T_{t-1}\phi$, we deduce that

$$\begin{aligned}\|D_yT_t\phi(x,y)\|_2 &\leqslant C_1\|T_{t-1}\phi\|_{L_p^\infty}(1+\|x\|_1+\|y\|_2^p)\\ &\leqslant C_1e^{-\lambda(t-1)}\|\phi\|_{L_p^\infty}(1+\|x\|_1+\|y\|_2^p).\end{aligned}$$

To prove (3.10), we first note that by (3.6), (3.7) and interpolation, we have that for any $\eta \in (0,1)$,

$$\|D_yP_t\phi\|_{C_p^{0,\eta}} \leqslant C_0\frac{1}{t^{(1+\eta)/2}}\|\phi\|_{L_p^\infty}.$$

Thus, we derive that

$$\begin{aligned}\|D_yT_t\phi\|_{C_p^{0,\eta}} &\leqslant C_0\left(\frac{1}{t^{(1+\eta)/2}}\|\phi\|_{L_p^\infty} + \int_0^t \frac{1}{(t-s)^{(1+\eta)/2}}\|D_yT_s\phi\|_{L_p^\infty}\mathrm{d}s\right)\\ &\leqslant C_0\frac{1}{t^{(1+\eta)/2}\wedge 1}\|\phi\|_{L_p^\infty}.\end{aligned}$$

Combining this with (3.7) and the assumption that $G \in C_b^{0,\eta}(H_1 \times H_2)$, we get for $k_1, k_2 \in H_2$,

$$|D_y^2T_t\phi(x,y).(k_1,k_2)| \leqslant C_2\frac{1}{t^{(2-\eta)/2}\wedge 1}\|\phi\|_{C_p^{0,\eta}}(1+\|x\|_1+\|y\|_2^p)\|k_1\|_2\|k_2\|_2,$$

which means that (3.10) holds for $t \leqslant 2$. Following the same ideas as above, we obtain that (3.10) holds for $t > 2$. □

3.2. **Proof of Theorem 3.2.**

Proof of Theorem 3.2. We divide the proof into three steps.

Step 1. Let $\psi(x,y)$ be defined by (3.3). We first prove that $\psi \in \mathbb{C}_p^{0,2}(H_1 \times H_2)$. In fact, for every $\phi \in L_p^\infty(H_1 \times H_2)$ satisfying (3.2), by (3.8) we deduce that

$$
\begin{aligned}
|\psi(x,y)| &\leqslant \int_0^\infty |T_t\phi(x,y)|\mathrm{d}t \leqslant C_0\|\phi\|_{L_p^\infty} \int_0^\infty (1+\|x\|_1+\|y\|_2^p)e^{-\lambda t}\mathrm{d}t \\
&\leqslant C_0\|\phi\|_{L_p^\infty}(1+\|x\|_1+\|y\|_2^p),
\end{aligned} \tag{3.12}
$$

and by (3.9) we have for every $k_1 \in H_2$,

$$
\begin{aligned}
|\langle D_y\psi(x,y), k_1\rangle_2| &\leqslant \int_0^\infty |\langle D_y T_t\phi(x,y), k_1\rangle_2|\mathrm{d}t \\
&\leqslant C_1\|\phi\|_{L_p^\infty} \int_0^\infty \frac{1}{\sqrt{t}\wedge 1}(1+\|x\|_1+\|y\|_2^p)\|k_1\|_2\, e^{-\lambda t}\mathrm{d}t \\
&\leqslant C_1\|\phi\|_{L_p^\infty}(1+\|x\|_1+\|y\|_2^p)\|k_1\|_2.
\end{aligned} \tag{3.13}
$$

Furthermore, by the dominated convergence theorem we deduce that

$$
\begin{aligned}
&\lim_{\|k_1\|_2\to 0} \frac{|\psi(x,y+k_1)-\psi(x,y)-\langle D_y\psi(x,y),k_1\rangle_2|}{\|k_1\|_2} \\
&\leqslant \lim_{\|k_1\|_2\to 0} \int_0^\infty \frac{|T_t\phi(x,y+k_1)-T_t\phi(x,y)-\langle D_yT_t\phi(x,y),k_1\rangle_2|}{\|k_1\|_2}\mathrm{d}t = 0.
\end{aligned}
$$

Similarly, by using (3.10) we can prove that $\psi \in \mathbb{C}_p^{0,2}(H_1 \times H_2)$ and for every $k_1, k_2 \in H_2$,

$$
|D_y^2\psi(x,y).(k_1,k_2)| \leqslant C_2\|\phi\|_{C_p^{0,\eta}}(1+\|x\|_1+\|y\|_2^p)\|k_1\|_2\|k_2\|_2. \tag{3.14}
$$

Here, we remark that the control of ψ and $D_y\psi$ depends only on the $\|\cdot\|_{L_p^\infty}$-norm of the function ϕ. In addition, by Fubini's theorem and the property of the invariant measure, we have

$$
\begin{aligned}
\int_{H_2} \psi(x,y)\mu^x(\mathrm{d}y) &= \int_0^\infty\int_{H_2} T_t\phi(x,y)\mu^x(\mathrm{d}y)\mathrm{d}t \\
&= \int_0^\infty\int_{H_2} \phi(x,y)\mu^x(\mathrm{d}y)\mathrm{d}t = 0.
\end{aligned}
$$

Thus, the assertion that ψ is the unique solution for equation (3.1) follows by Itô's formula, see e.g. [4, Lemma 4.3].

Step 2. When $k=1$ in the assumptions, we prove that $\psi(x,y) \in C_p^{1,0}(H_1 \times H_2)$. In fact, for every $h_1 \in H_1$ and $\tau > 0$, we have

$$
\mathcal{L}_2(x,y)\frac{\psi(x+\tau h_1,y)-\psi(x,y)}{\tau} = -\frac{\phi(x+\tau h_1,y)-\phi(x,y)}{\tau}
$$

$$- \frac{\langle G(x+\tau h_1, y) - G(x,y), D_y\psi(x+\tau h_1, y)\rangle_2}{\tau} =: -\phi_{\tau,h_1}(x,y).$$

By the assumptions on ϕ and G, and using estimates (3.13) and (3.14), one can check that $\phi_{\tau,h_1}(x,y) \in C_p^{0,\eta}(H_1 \times H_2)$. We claim that

$$\int_{H_2} \phi_{\tau,h_1}(x,y)\mu^x(\mathrm{d}y) = 0, \quad \forall \tau > 0, x, h_1 \in H_1. \tag{3.15}$$

Then, according to Step 1, we obtain that for every $\tau > 0$,

$$\frac{\psi(x+\tau h_1, y) - \psi(x,y)}{\tau} = \int_0^\infty \mathbb{E}\phi_{\tau,h_1}(x, Y_t^x(y))\mathrm{d}t. \tag{3.16}$$

Note that

$$\lim_{\tau\to 0} \phi_{\tau,h_1}(x,y) = \langle D_x\phi(x,y), h_1\rangle_1 + \langle D_x G(x,y).h_1, D_y\psi(x,y)\rangle_2 =: \phi_{h_1}(x,y).$$

Using the assumption that $G \in C_b^{1,0}(H_1 \times H_2)$ and (3.13), we find that

$$\begin{aligned} |\phi_{h_1}(x,y)| &\leqslant \|D_x\phi(x,y)\|_1\|h_1\|_1 + \|D_xG(x,y)\|_{\mathscr{L}(H_1,H_2)}\|h_1\|_1\|D_y\psi(x,y)\|_2 \\ &\leqslant C_3(1 + \|x\|_1 + \|y\|_2^p)\|h_1\|_1. \end{aligned} \tag{3.17}$$

Thus, by the dominated convergence theorem, we obtain

$$\int_{H_2} \phi_{h_1}(x,y)\mu^x(\mathrm{d}y) = 0, \quad \forall x, h_1 \in H_1.$$

Combining this with (3.8), we have

$$\begin{aligned} |T_t\phi_{\tau,h_1}(x,y)| &\leqslant C_4\|\phi_{\tau,h_1}\|_{L_p^\infty}(1 + \|x\|_1 + \|y\|_2^p)e^{-\lambda t} \\ &\leqslant C_4\big(1 + \|\phi_{h_1}\|_{L_p^\infty}\big)(1 + \|x\|_1 + \|y\|_2^p)e^{-\lambda t}. \end{aligned}$$

As a result, taking the limit $\tau \to 0$ on both sides of (3.16) we get

$$\langle D_x\psi(x,y), h_1\rangle_1 = \int_0^\infty \mathbb{E}\phi_{h_1}(x, Y_s^x(y))\mathrm{d}s. \tag{3.18}$$

It remains to prove (3.15). To this end, for every $n \in \mathbb{N}$, let $H_2^n := span\{e_{2,k}; 1 \leqslant k \leqslant n\}$ and denote the orthogonal projection of H_2 onto H_2^n by P_2^n. We introduce the following approximation of system (1.4):

$$\mathrm{d}Y_t^{x,n} = B_nY_t^{x,n}\mathrm{d}t + G_n(x, Y_t^{x,n})\mathrm{d}t + P_2^n\mathrm{d}W_t^2, \quad Y_0^{x,n} = P_2^n y \in H_2^n. \tag{3.19}$$

where for $(x,y) \in H_1 \times H_2$,

$$B_n y := BP_2^n y \quad \text{and} \quad G_n(x,y) := P_2^nG(x, P_2^n y). \tag{3.20}$$

It is easy to check that G_n is uniformly bounded with respect to n. Thus the solution to equation (3.19) has the same long-time behavior as the one to equation (1.4). Let $\mu_n^x(\mathrm{d}y)$ be the invariant measure for $Y_t^{x,n}$, and define $\mathcal{L}_2^n(x,y)$ by

$$\begin{aligned}\mathcal{L}_2^n(x,y)\varphi(x,y) := \langle B_n y + G_n(x,y), D_y\varphi(x,y)\rangle_2 \\ + \frac{1}{2}Tr\big[D_y^2\varphi(x,y)Q_{2,n}\big], \quad \forall \varphi \in C_p^{0,2}(H_1\times H_2^n),\end{aligned}$$

where $Q_{2,n} := Q_2P_2^n$. Consider the Poisson equation corresponding to (3.19):

$$\mathcal{L}_2^n(x,y)\psi^n(x,y) = -\phi(x,P_2^n y) =: -\phi^n(x,y). \tag{3.21}$$

As in Step 1, the unique solution is given by

$$\psi^n(x,y) = \int_0^\infty \mathbb{E}\big[\phi^n(x,Y_t^{x,n}(y))\big]\mathrm{d}t.$$

According to [4, Subsection 4.1] (see also [13, Section 6]), we have for every $x\in H_1$ and $y \in H_2$,

$$\lim_{n\to\infty}\psi^n(x,y) = \psi(x,y), \quad \lim_{n\to\infty}\langle D_y\psi^n(x,y),k_2\rangle_2 = \langle D_y\psi(x,y),k_2\rangle_2. \tag{3.22}$$

For every $\tau>0$ and $h_1\in H_1$, define

$$\begin{aligned}\phi^n_{\tau,h_1}(x,y) := \big[\phi^n(x+\tau h_1,y)-\phi^n(x,y)\big] \\ + \langle G_n(x+\tau h_1,y) - G_n(x,y), D_y\psi^n(x+\tau h_1,y)\rangle_2.\end{aligned}$$

Since (3.21) is an equation in finite dimensions, according to [40, Lemma 3.2] we have

$$\int_{H_2^n}\phi^n_{\tau,h_1}(x,y)\mu_n^x(\mathrm{d}y) = 0.$$

Using estimates (3.17), (3.22), the formula above [4, (4.4)] and taking the limit $n\to\infty$ on both sides of the above equality, we obtain (3.15).

Step 3. When $k=2$ in the assumptions, we prove that $\psi(x,y)\in C_p^{2,0}(H_1\times H_2)$. In this case, we mainly focus on the a-priori estimate. The specific procedure can be done as in Step 2. In view of (3.18) and according to the results in Step 1, we can conclude that $\langle D_x\psi(x,y),h_1\rangle_1 \in \mathbb{C}_p^{0,2}(H_1\times H_2)$ with

$$\begin{aligned}|\langle D_x\psi(x,y),h_1\rangle_1| &\leqslant C_5\|\phi_{h_1}\|_{L_p^\infty}(1+\|x\|_1+\|y\|_2^p) \\ &\leqslant C_5(1+\|x\|_1+\|y\|_2^p)\|h_1\|_1,\end{aligned}$$

and

$$\begin{aligned}|D_yD_x\psi(x,y).(h_1,k)| &\leqslant C_5\|\phi_{h_1}\|_{L_p^\infty}(1+\|x\|_1+\|y\|_2^p)\|k\|_2 \\ &\leqslant C_5(1+\|x\|_1+\|y\|_2^p)\|h_1\|_1\|k\|_2.\end{aligned} \tag{3.23}$$

Moreover, we have

$$\mathcal{L}_2(x,y)\langle D_x\psi(x,y),h_1\rangle_1 = -\phi_{h_1}(x,y), \tag{3.24}$$

Since $\langle D_x\psi(x,y), h_1\rangle_1$ is a classical solution, by taking derivative with respect to the x variable on both sides of the equation, we have that for any $h_1, h_2 \in H_1$,

$$\begin{aligned}\mathcal{L}_2(x,y)(D_x^2\psi(x,y).(h_1,h_2)) = & -D_x^2\phi(x,y).(h_1,h_2)\\ & - 2D_yD_x\psi(x,y).(h_2, D_xG(x,y).h_1)\\ & - \langle D_x^2G(x,y).(h_1,h_2), D_y\psi(x,y)\rangle_2 =: -\phi_{h_1,h_2}(x,y).\end{aligned}$$

By the assumption that $G \in C_b^{2,0}(H_1 \times H_2)$ and (3.23), we get

$$\begin{aligned}|\phi_{h_1,h_2}(x,y)| \leqslant & \|D_x^2\phi(x,y)\|_{\mathscr{L}(H_1\times H_1)}\|h_1\|_1\|h_2\|_1 + 2|D_yD_x\psi(x,y).(h_2, D_xG(x,y).h_1)|\\ & + |\langle D_x^2G(x,y).(h_1,h_2), D_y\psi(x,y)\rangle_2|\\ \leqslant & C_6(1+\|x\|_1+\|y\|_2^p)\|h_1\|_1\|h_2\|_1.\end{aligned}$$

Furthermore, by using [40, Lemma 3.2] again and the same approximation argument as in Step 2, we have that $\phi_{h_1,h_2}(x,y)$ satisfies the centering condition

$$\int_{H_2}\phi_{h_1,h_2}(x,y)\mu^x(\mathrm{d}y) = 0. \tag{3.25}$$

Thus, in view of (3.12) and (3.13) we can get that $(D_x^2\psi(x,y).(h_1,h_2)) \in \mathbb{C}_p^{0,2}(H_1\times H_2)$ with

$$\begin{aligned}|D_x^2\psi(x,y).(h_1,h_2)| \leqslant & C_7\|\phi_{h_1,h_2}\|_{L_p^\infty}(1+\|x\|_1+\|y\|_2^p)\\ \leqslant & C_7(1+\|x\|_1+\|y\|_2^p)\|h_1\|_1\|h_2\|_1.\end{aligned}$$

The proof is finished. □

Given a function $F(x,y)$, recall that $\bar{F}$ is defined by (1.3). Usually, it is not so easy to study the regularity of the averaged function, which contains a separate problem connected with the smoothness of the invariant measure $\mu^x(\mathrm{d}y)$. Here, we provide a simple method by using Theorem 3.2.

Lemma 3.7. *Assume that $F \in C_p^{1,\eta}(H_1\times H_2)$ with $\eta > 0$, and let $\Psi(x,y)$ solve the Poisson equation (2.5). Then for any $h_1 \in H_1$, we have*

$$D_x\bar{F}(x).h_1 = \int_{H_2}\Big[D_xF(x,y).h_1 + \langle D_xG(x,y).h_1, D_y\Psi(x,y)\rangle_2\Big]\mu^x(\mathrm{d}y). \tag{3.26}$$

Furthermore, assume that $F \in C_p^{2,\eta}(H_1\times H_2)$, then we have for any $h_1, h_2 \in H_1$,

$$\begin{aligned}D_x^2\bar{F}(x).(h_1,h_2) = \int_{H_2}\Big[& D_x^2F(x,y).(h_1,h_2) + 2D_yD_x\Psi(x,y).(h_2, D_xG(x,y).h_1)\\ & + \langle D_x^2G(x,y).(h_1,h_2), D_y\Psi(x,y)\rangle_2\Big]\mu^x(\mathrm{d}y).\end{aligned} \tag{3.27}$$

In particular, we have

$$\|D_x\bar{F}(x).h_1\|_1 \leqslant C_0\|h_1\|_1 \quad \textit{and} \quad \|D_x^2\bar{F}(x).(h_1,h_2)\|_1 \leqslant C_0\|h_1\|_1\|h_2\|_1, \tag{3.28}$$

where $C_0 > 0$ is a constant.

Remark 3.8. *The interesting point in formula (3.26) (and also in (3.27)) lies in that the regularity of the averaged function $\bar{F}$ with respect to the x-variable is transferred to the regularity of the solution Ψ with respect to the y-variable. Since Ψ is the solution to the corresponding Poisson equation, we can get the required regularity for free by the uniform ellipticity property of the generator $\mathcal{L}_2(x,y)$ as been proven in Theorem 3.2.*

Proof. Recall that

$$\mathcal{L}_2(x,y)\Psi(x,y) = -\delta F(x,y) = -(F(x,y) - \bar{F}(x)).$$

Note that δF satisfies the centering condition (3.2). As in the proof of (3.24), we have

$$\mathcal{L}_2(x,y)D_x\Psi(x,y).h_1 = -D_x\delta F(x,y).h_1 - \langle D_xG(x,y).h_1, D_y\Psi(x,y)\rangle_2.$$

Moreover, we have

$$\int_{H_2}\Big[D_x\delta F(x,y).h_1 + \langle D_xG(x,y).h_1, D_y\Psi(x,y)\rangle_2\Big]\mu^x(\mathrm{d}y) = 0.$$

Note that

$$\int_{H_2} D_x\bar{F}(x).h_1\mu^x(\mathrm{d}y) = D_x\bar{F}(x).h_1,$$

hence we get (3.26). Similarly, as in (3.25) we have

$$\int_{H_2}\Big[D_x^2\delta F(x,y).(h_1,h_2) + 2D_yD_x\Psi(x,y).(h_2, D_xG(x,y).h_1) \\ + \langle D_x^2G(x,y).(h_1,h_2), D_y\Psi(x,y)\rangle_2\Big]\mu^x(\mathrm{d}y) = 0,$$

which in turn yields (3.27). Finally, due to the fact that for any $p \geqslant 1$,

$$\int_{H_2}\big(1 + \|y\|_2\big)^p\mu^x(\mathrm{d}y) < \infty,$$

we get estimate (3.28). □

4. Strong convergence in the averaging principle

4.1. **Galerkin approximation.** Itô's formula will be used frequently below in the proof of the main result. However, due to the presence of unbounded operators in the equation, we can not apply Itô's formula for SPDE (1.1) directly. For this reason, we introduce the following Galerkin approximation scheme.

For $n \in \mathbb{N}$, let $H_1^n := span\{e_{1,k}; 1 \leqslant k \leqslant n\}$ and denote the orthogonal projection of H_1 onto H_1^n by P_1^n. Recall that $G_n(x,y)$ is defined by (3.20), and for $(x,y) \in H_1^n \times H_2^n$, define $F_n(x,y) := P_1^n F(x,y)$. We reduce the infinite dimensional system (1.1) to the following finite dimensional system in $H_1^n \times H_2^n$:

$$\begin{cases} \mathrm{d}X_t^{n,\varepsilon} = AX_t^{n,\varepsilon}\mathrm{d}t + F_n(X_t^{n,\varepsilon}, Y_t^{n,\varepsilon})\mathrm{d}t + P_1^n \mathrm{d}W_t^1, \\ \mathrm{d}Y_t^{n,\varepsilon} = \varepsilon^{-1}BY_t^{n,\varepsilon}\mathrm{d}t + \varepsilon^{-1}G_n(X_t^{n,\varepsilon}, Y_t^{n,\varepsilon})\mathrm{d}t + \varepsilon^{-1/2}P_2^n \mathrm{d}W_t^2, \end{cases} \tag{4.1}$$

with initial values $X_0^{n,\varepsilon} = x^n := P_1^n x \in H_1^n$ and $Y_0^{n,\varepsilon} = y^n := P_2^n y \in H_2^n$. It is easy to check that F_n and G_n satisfy the same conditions as F and G with bounds which are uniform with respect to n. Thus the equation (4.1) is well-posed in $H_1^n \times H_2^n$. The corresponding averaged equation for system (4.1) can be formulated as

$$\mathrm{d}\bar{X}_t^n = A\bar{X}_t^n\mathrm{d}t + \bar{F}_n(\bar{X}_t^n)\mathrm{d}t + P_1^n\mathrm{d}W_t^1, \quad \bar{X}_0^n = x^n \in H_1^n, \tag{4.2}$$

where $\bar{F}_n(x)$ is defined by

$$\bar{F}_n(x) := \int_{H_2^n} F_n(x,y)\mu_n^x(\mathrm{d}y). \tag{4.3}$$

Note that $\bar{X}_t^n$ is not the Galerkin approximation of $\bar{X}_t$.

The following result states the convergence of the finite dimensional system (4.1) to the initial equation (1.1).

Lemma 4.1. *Let $T > 0$, $x \in \mathcal{D}((-A)^\theta)$ and $y \in \mathcal{D}((-B)^\theta)$ with $\theta \in [0,1]$. Then for every $q \geqslant 1$ and $\gamma \in [0,\theta]$, we have for every $t \in [0,T]$,*

$$\lim_{n\to\infty}\Big(\mathbb{E}\|(-A)^\gamma(X_t^\varepsilon - X_t^{n,\varepsilon})\|_1^q + \mathbb{E}\|Y_t^\varepsilon - Y_t^{n,\varepsilon}\|_2^q + \mathbb{E}\|(-A)^\gamma(\bar{X}_t - \bar{X}_t^n)\|_1^q\Big) = 0. \tag{4.4}$$

Proof. When $\theta = 0$ (and thus $\gamma = 0$), (4.4) was proven in [4, Lemma 4.2]. For general $\gamma \in (0,\theta]$, by Lemma 6.4 we know that $X_t^\varepsilon \in \mathcal{D}((-A)^\gamma)$. As a result, we have

$$\mathbb{E}\Big(\|(-A)^\gamma(X_t^\varepsilon - X_t^{n,\varepsilon})\|_1^q\Big) = \mathbb{E}\left[\left(\sum_{k=n+1}^\infty \alpha_k^{2\gamma}\langle X_t^\varepsilon, e_{1,k}\rangle_1^2\right)^{q/2}\right] \to 0 \text{ as } n\to\infty.$$

Furthermore, by Lemma 6.5 we deduce that

$$\begin{aligned}\mathbb{E}\Big(\|(-A)^\gamma(\bar{X}_t - \bar{X}_t^n)\|_1^q\Big) \leqslant{}& \mathbb{E}\left(\int_0^t \|(-A)^\gamma e^{(t-s)A}[\bar{F}(\bar{X}_s) - \bar{F}_n(\bar{X}_s)]\|_1\mathrm{d}s\right)^q \\ &+ \mathbb{E}\left\|\int_0^t (-A)^\gamma e^{(t-s)A}(I - P_1^n)\mathrm{d}W_s\right\|_1^q \\ &+ \mathbb{E}\left(\int_0^t \|(-A)^\gamma e^{(t-s)A}[\bar{F}_n(\bar{X}_s) - \bar{F}_n(\bar{X}_s^n)]\|_1\mathrm{d}s\right)^q.\end{aligned}$$

Since $\|\bar{F}_n - \bar{F}\|_1 \to 0$ as $n \to \infty$ (see e.g. [4, (4.4)]), the first two terms go to 0 as $n \to \infty$ by the dominated convergence theorem. For the last term, we have

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}\left(\int_0^t \|(-A)^\gamma e^{(t-s)A}[\bar{F}_n(\bar{X}_s) - \bar{F}_n(X_s^n)]\|_1 \mathrm{d}s\right)^q \\
&\leqslant \lim_{n\to\infty} C_1\, \mathbb{E}\left(\int_0^t (t-s)^{-\gamma}\|\bar{X}_s - \bar{X}_s^n\|_1 \mathrm{d}s\right)^q = 0,
\end{aligned}$$

which in turn yields the desired result. □

4.2. **Proof of Theorem 2.2.** Let $T > 0$, $x \in \mathcal{D}((-A)^\theta)$ and $y \in \mathcal{D}((-B)^\theta)$ with $\theta \in [0,1]$. Note that for any $t \in [0,T]$, $q \geqslant 1$ and $\gamma \in [0, \theta \wedge 1/2)$, we have

$$\begin{aligned}
\mathbb{E}\|(-A)^\gamma(X_t^\varepsilon - \bar{X}_t)\|_1^q &\leqslant \mathbb{E}\|(-A)^\gamma(X_t^\varepsilon - X_t^{n,\varepsilon})\|_1^q \\
&\quad + \mathbb{E}\|(-A)^\gamma(X_t^{n,\varepsilon} - \bar{X}_t^n)\|_1^q + \mathbb{E}\|(-A)^\gamma(\bar{X}_t^n - \bar{X}_t)\|_1^q.
\end{aligned}$$

By Lemma 4.1, the first and the last terms on the right-hand side of the above inequality converge to 0 as $n \to \infty$. Therefore, in order to prove Theorem 2.2, we only need to show that

$$\sup_{t\in[0,T]} \mathbb{E}\|(-A)^\gamma(X_t^{n,\varepsilon} - \bar{X}_t^n)\|_1^q \leqslant C_T\, \varepsilon^{q/2}, \tag{4.5}$$

where $C_T > 0$ **is a constant independent of** n. In the following subsection, we shall only work with the approximation system (4.1), and prove bounds that are uniform with respect to the dimension. But in order to simplify the notations, we omit the index n. In particular, for $i = 1, 2$, the spaces H_i^n are denoted by H_i.

Define

$$\begin{aligned}
\mathcal{L}_1\varphi(x,y) := \mathcal{L}_1(x,y)\varphi(x,y) &:= \langle Ax + F(x,y), D_x\varphi(x,y)\rangle_1 \\
&\quad + \frac{1}{2} Tr\big[D_x^2\varphi(x,y)Q_1\big], \quad \forall \varphi \in C_p^{2,0}(H_1 \times H_2).
\end{aligned} \tag{4.6}$$

We first establish the following strong fluctuation estimate for an appropriate integral functional of $(X_r^\varepsilon, Y_r^\varepsilon)$ over the time interval $[s,t]$, which will play an important role in proving (4.5).

Lemma 4.2 (Strong fluctuation estimate)**.** *Let $T, \theta > 0$, $x \in \mathcal{D}((-A)^\theta)$ and $y \in \mathcal{D}((-B)^\theta)$. Assume that (**A1**) and (**A2**) hold, $F \in C_p^{2,\eta}(H_1 \times H_2, H_1)$ and $G \in C_b^{2,\eta}(H_1 \times H_2, H_2)$ with $\eta > 0$. Then for any $\gamma \in [0, \theta \wedge 1/2)$, $q \geqslant 1$, $0 \leqslant s \leqslant t \leqslant T$ and $\tilde{\phi} \in C_p^{2,\eta}(H_1 \times H_2, H_1)$ satisfying (3.2), we have*

$$\mathbb{E}\left\|\int_s^t (-A)^\gamma e^{(t-r)A}\tilde{\phi}(X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r\right\|_1^q \leqslant C_{q,\gamma,T}(t-s)^{(\theta-\gamma)q}\,\varepsilon^{q/2}, \tag{4.7}$$

where $C_{q,\gamma,T} > 0$ is a constant.

Proof. Let $\tilde{\psi}$ solve the Poisson equation

$$\mathcal{L}_2(x,y)\tilde{\psi}(x,y) = -\tilde{\phi}(x,y),$$

and define

$$\tilde{\psi}_{t,\gamma}(r,x,y) := (-A)^{\gamma} e^{(t-r)A}\tilde{\psi}(x,y). \tag{4.8}$$

Since $\mathcal{L}_2$ is an operator with respect to the y-variable, one can check that

$$\mathcal{L}_2(x,y)\tilde{\psi}_{t,\gamma}(r,x,y) = -(-A)^{\gamma} e^{(t-r)A}\tilde{\phi}(x,y). \tag{4.9}$$

According to Theorem 3.2, we know that $\tilde{\psi} \in C_p^{2,0}(H_1 \times H_2, H_1) \cap \mathbb{C}_p^{0,2}(H_1 \times H_2, H_1)$. Applying Itô's formula to $\tilde{\psi}_{t,\gamma}(t, X_t^\varepsilon, Y_t^\varepsilon)$ we get

$$\begin{aligned}\tilde{\psi}_{t,\gamma}(t, X_t^\varepsilon, Y_t^\varepsilon) &= \tilde{\psi}_{t,\gamma}(s, X_s^\varepsilon, Y_s^\varepsilon) + \int_s^t (\partial_r + \mathcal{L}_1)\tilde{\psi}_{t,\gamma}(r, X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r \\ &\quad + \frac{1}{\varepsilon}\int_s^t \mathcal{L}_2\tilde{\psi}_{t,\gamma}(r, X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r + M_{t,s}^1 + \frac{1}{\sqrt{\varepsilon}} M_{t,s}^2, \end{aligned} \tag{4.10}$$

where $M_{t,s}^1$ and $M_{t,s}^2$ are defined by

$$M_{t,s}^1 := \int_s^t D_x\tilde{\psi}_{t,\gamma}(r, X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}W_r^1 \quad \text{and} \quad M_{t,s}^2 := \int_s^t D_y\tilde{\psi}_{t,\gamma}(r, X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}W_r^2.$$

Multiplying both sides of (4.10) by ε and using (4.9), we obtain

$$\begin{aligned}\int_s^t (-A)^{\gamma} e^{(t-r)A}\tilde{\phi}(X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r &= -\int_s^t \mathcal{L}_2\tilde{\psi}_{t,\gamma}(r, X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r \\ &= \varepsilon\left[\tilde{\psi}_{t,\gamma}(s, X_s^\varepsilon, Y_s^\varepsilon) - \tilde{\psi}_{t,\gamma}(t, X_t^\varepsilon, Y_t^\varepsilon)\right] \\ &\quad + \varepsilon\int_s^t (\partial_r + \mathcal{L}_1)\tilde{\psi}_{t,\gamma}(r, X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r + \varepsilon\, M_{t,s}^1 + \sqrt{\varepsilon}\, M_{t,s}^2.\end{aligned}$$

Note that

$$\begin{aligned}\int_s^t \partial_r\tilde{\psi}_{t,\gamma}(r, X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r &= \int_s^t \partial_r\tilde{\psi}_{t,\gamma}(r, X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}r \\ &\quad + \int_s^t \partial_r\left[\tilde{\psi}_{t,\gamma}(r, X_r^\varepsilon, Y_r^\varepsilon) - \tilde{\psi}_{t,\gamma}(r, X_t^\varepsilon, Y_t^\varepsilon)\right]\mathrm{d}r \\ &= \tilde{\psi}_{t,\gamma}(t, X_t^\varepsilon, Y_t^\varepsilon) - \tilde{\psi}_{t,\gamma}(s, X_t^\varepsilon, Y_t^\varepsilon) \\ &\quad + \int_s^t \partial_r\left[\tilde{\psi}_{t,\gamma}(r, X_r^\varepsilon, Y_r^\varepsilon) - \tilde{\psi}_{t,\gamma}(r, X_t^\varepsilon, Y_t^\varepsilon)\right]\mathrm{d}r,\end{aligned}$$

and that

$$\partial_r\tilde{\psi}_{t,\gamma}(r,x,y) = (-A)^{1+\gamma} e^{(t-r)A}\tilde{\psi}(x,y).$$

As a result, we further get

$$\int_s^t (-A)^\gamma e^{(t-r)A}\tilde{\phi}(X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r = \varepsilon\,(-A)^\gamma e^{(t-s)A}\big[\tilde{\psi}(X_s^\varepsilon, Y_s^\varepsilon) - \tilde{\psi}(X_t^\varepsilon, Y_t^\varepsilon)\big] + \varepsilon \int_s^t (-A)^{1+\gamma} e^{(t-r)A}\Big(\tilde{\psi}(X_r^\varepsilon, Y_r^\varepsilon) - \tilde{\psi}(X_t^\varepsilon, Y_t^\varepsilon)\Big)\,\mathrm{d}r + \varepsilon \int_s^t \mathcal{L}_1\tilde{\psi}_{t,\gamma}(r, X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r + \varepsilon\, M_{t,s}^1 + \sqrt{\varepsilon}\, M_{t,s}^2.$$

Thus for any $0 \leqslant s \leqslant t \leqslant T$ and $q \geqslant 1$, we deduce that

$$\begin{aligned}
&\mathbb{E}\left\|\int_s^t (-A)^\gamma e^{(t-r)A}\tilde{\phi}(X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r\right\|_1^q \\
&\leqslant C_0\bigg(\varepsilon^q\,\mathbb{E}\big\|(-A)^\gamma e^{(t-s)A}\big[\tilde{\psi}(X_s^\varepsilon, Y_s^\varepsilon) - \tilde{\psi}(X_t^\varepsilon, Y_t^\varepsilon)\big]\big\|_1^q \\
&\quad + \varepsilon^q\,\mathbb{E}\left\|\int_s^t (-A)^{1+\gamma} e^{(t-r)A}\Big(\tilde{\psi}(X_r^\varepsilon, Y_r^\varepsilon) - \tilde{\psi}(X_t^\varepsilon, Y_t^\varepsilon)\Big)\,\mathrm{d}r\right\|_1^q \\
&\quad + \varepsilon^q\,\mathbb{E}\left\|\int_s^t \mathcal{L}_1\tilde{\psi}_{t,\gamma}(r, X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r\right\|_1^q + \varepsilon^q\,\mathbb{E}\|M_{t,s}^1\|_1^q + \varepsilon^{q/2}\,\mathbb{E}\|M_{t,s}^2\|_1^q\bigg) =: \sum_{i=1}^5 \mathscr{J}_i(t,s,\varepsilon).
\end{aligned}$$

For the first term, by Lemmas 6.1, 6.2, 6.3 below and the fact that $\theta < 1/2$, we have

$$\begin{aligned}
\mathscr{J}_1(t,s,\varepsilon) &\leqslant C_1\,\varepsilon^q\,(t-s)^{-\gamma q}\Big(\mathbb{E}\big(1 + \|X_t^\varepsilon\|_1 + \|X_s^\varepsilon\|_1 + \|Y_t^\varepsilon\|_2^p + \|Y_s^\varepsilon\|_2^p\big)^{2q}\Big)^{1/2} \\
&\qquad \cdot\Big(\mathbb{E}\|X_t^\varepsilon - X_s^\varepsilon\|_1^{2q} + \mathbb{E}\|Y_t^\varepsilon - Y_s^\varepsilon\|_2^{2q}\Big)^{1/2} \\
&\leqslant C_1\,(t-s)^{(\theta-\gamma)q}\varepsilon^{(1-\theta)q} \leqslant C_1\,(t-s)^{(\theta-\gamma)q}\varepsilon^{q/2}.
\end{aligned}$$

Similarly, by Minkowski's inequality we also have

$$\begin{aligned}
\mathscr{J}_2(t,s,\varepsilon) &\leqslant C_2\,\varepsilon^q\left(\int_s^t (t-r)^{-1-\gamma}\left[\Big(\mathbb{E}\big[\|X_t^\varepsilon - X_r^\varepsilon\|_1^{2q}\big]\Big)^{1/2q} + \Big(\mathbb{E}\big[\|Y_t^\varepsilon - Y_r^\varepsilon\|_2^{2q}\big]\Big)^{1/2q}\right]\mathrm{d}r\right)^q \\
&\leqslant C_2\,\varepsilon^q\left(\int_s^t (t-r)^{-1-\gamma}\frac{(t-s)^\theta}{\varepsilon^\theta}\mathrm{d}r\right)^q \\
&\leqslant C_2\,(t-s)^{(\theta-\gamma)q}\varepsilon^{(1-\theta)q} \leqslant C_2\,(t-s)^{(\theta-\gamma)q}\varepsilon^{q/2}.
\end{aligned}$$

To control the third term, by definitions (4.6), (4.8) and Theorem 3.2, one can check that

$$\|\mathcal{L}_1\tilde{\psi}_{t,\gamma}(r,x,y)\|_1 \leqslant C_3\,(t-r)^{-\gamma}\big(1 + \|Ax\|_1 + \|y\|_2^p\big)(1 + \|x\|_1 + \|y\|_2^p),$$

which in turn yields by Minkowski's inequality and Lemma 6.4 that for $\gamma' \in (0,1/2)$,

$$\begin{aligned}\mathscr{J}_3(t,s,\varepsilon) &\leqslant C_3\,\varepsilon^{(1-\gamma')q}\left(\int_s^t (t-r)^{-\gamma} r^{(\theta-1)}\mathrm{d}r\right)^q\\ &\leqslant C_3\,(t-s)^{(\theta-\gamma)q}\varepsilon^{(1-\gamma')q} \leqslant C_3\,(t-s)^{(\theta-\gamma)q}\varepsilon^{q/2}.\end{aligned}$$

As for $\mathscr{I}_4(t,s,\varepsilon)$, by Burkholder-Davis-Gundy's inequality and the assumption (**A2**), we have

$$\begin{aligned}\mathscr{J}_4(t,s,\varepsilon) &\leqslant C_4\,\varepsilon^q\left(\int_s^t \mathbb{E}\|(-A)^\gamma e^{(t-r)A} D_x\tilde{\psi}(X_r^\varepsilon,Y_r^\varepsilon)Q_1^{1/2}\|^2_{\mathscr{L}_2(H_1)}\mathrm{d}r\right)^{q/2}\\ &\leqslant C_4\,\varepsilon^q\left(\int_s^t \Big(1+\mathbb{E}\|X_r^\varepsilon\|_1^2+\mathbb{E}\|Y_r^\varepsilon\|_2^{2p}\Big)\|(-A)^\gamma e^{(t-r)A}Q_1^{1/2}\|^2_{\mathscr{L}_2(H_1)}\mathrm{d}r\right)^{q/2}\\ &\leqslant C_4\,(t-s)^{(1/2-\gamma)q}\varepsilon^q \leqslant C_4\,(t-s)^{(\theta-\gamma)q}\varepsilon^q,\end{aligned}$$

and similarly one can check that

$$\mathscr{J}_5(t,s,\varepsilon) \leqslant C_5\,(t-s)^{(1/2-\gamma)q}\varepsilon^{q/2} \leqslant C_5\,(t-s)^{(\theta-\gamma)q}\varepsilon^{q/2}.$$

Combining the above computations, we get the desired estimate. □

Now, we are in the position to give:

Proof of estimate (4.5). Fix $T>0$ below. In view of (6.1) and (6.5), we have for every $t\in[0,T]$ and $\gamma\in[0,\theta\wedge 1/2)$,

$$\begin{aligned}(-A)^\gamma(X_t^\varepsilon-\bar{X}_t) &= \int_0^t (-A)^\gamma e^{(t-s)A}\big[\bar{F}(X_s^\varepsilon)-\bar{F}(\bar{X}_s)\big]\mathrm{d}s\\ &\quad+\int_0^t (-A)^\gamma e^{(t-s)A}\delta F(X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s,\end{aligned}$$

where δF is defined by (1.8). Thus for any $q\geqslant 1$, we have

$$\begin{aligned}\mathbb{E}\|(-A)^\gamma(X_t^\varepsilon-\bar{X}_t)\|_1^q &\leqslant C_q\,\mathbb{E}\left\|\int_0^t (-A)^\gamma e^{(t-s)A}\big[\bar{F}(X_s^\varepsilon)-\bar{F}(\bar{X}_s)\big]\mathrm{d}s\right\|_1^q\\ &\quad+C_q\,\mathbb{E}\left\|\int_0^t (-A)^\gamma e^{(t-s)A}\delta F(X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s\right\|_1^q =: \mathscr{I}_1(t,\varepsilon)+\mathscr{I}_2(t,\varepsilon).\end{aligned}$$

By Lemma 3.7 and Minkowski's inequality, we deduce that

$$\begin{aligned}\mathscr{I}_1(t,\varepsilon) &\leqslant C_1\,\mathbb{E}\left(\int_0^t (t-s)^{-\gamma}\|\bar{F}(X_s^\varepsilon)-\bar{F}(\bar{X}_s)\|_1\mathrm{d}s\right)^q\\ &\leqslant C_1\,\left(\int_0^t (t-s)^{-\gamma}\Big(\mathbb{E}\|X_s^\varepsilon-\bar{X}_s\|_1^q\Big)^{1/q}\mathrm{d}s\right)^q.\end{aligned}$$

For the second term, note that $\delta F(x,y)$ satisfies the centering condition (3.2). As a result, it follows by Lemma 4.2 directly that

$$\mathscr{I}_2(t,\varepsilon) \leqslant C_2\,\varepsilon^{q/2}.$$

Thus we arrive at

$$\mathbb{E}\|(-A)^{\gamma}(X_t^{\varepsilon}-\bar{X}_t)\|_1^q \leqslant C_3\,\varepsilon^{q/2} + C_3\left(\int_0^t (t-s)^{-\gamma}\Big(\mathbb{E}\|X_s^{\varepsilon}-\bar{X}_s\|_1^q\Big)^{1/q}\mathrm{d}s\right)^q. \quad (4.11)$$

Letting $\gamma = 0$, by Gronwall's inequality we obtain

$$\sup_{t\in[0,T]}\mathbb{E}\|X_t^{\varepsilon}-\bar{X}_t\|_1^q \leqslant C_4\,\varepsilon^{q/2}.$$

Taking this back into (4.11), we get the desired result. □

5. Normal deviations

5.1. **Kolmogorov equation.** Recall that $\bar{X}_t$ and $\bar{Z}_t$ satisfy the equations (1.2) and (2.7), respectively. We write a system of equations for the process $(\bar{X}_t, \bar{Z}_t)$ as follows:

$$\begin{cases} \mathrm{d}\bar{X}_t = A\bar{X}_t\mathrm{d}t + \bar{F}(\bar{X}_t)\mathrm{d}t + \mathrm{d}W_t^1, & \bar{X}_0 = x, \\ \mathrm{d}\bar{Z}_t = A\bar{Z}_t\mathrm{d}t + D_x\bar{F}(\bar{X}_t).\bar{Z}_t\mathrm{d}t + \sigma(\bar{X}_t)\mathrm{d}\tilde{W}_t, & \bar{Z}_0 = 0. \end{cases}$$

Note that the processes $\bar{X}_t$ and $\bar{Z}_t$ depend on the initial value x. Below, we shall write $\bar{X}_t(x)$ when we want to stress its dependence on the initial value, and use $\bar{Z}_t(x,z)$ to denote the process $\bar{Z}_t$ with initial point $\bar{Z}_0 = z \in H_1$. Let $\bar{\mathcal{L}}$ be the formal infinitesimal generator of the Markov process $(\bar{X}_t, \bar{Z}_t)$, i.e.,

$$\bar{\mathcal{L}} := \bar{\mathcal{L}}_1 + \bar{\mathcal{L}}_3,$$

where $\bar{\mathcal{L}}_1$ and $\bar{\mathcal{L}}_3$ are defined by

$$\bar{\mathcal{L}}_1\varphi(x) := \bar{\mathcal{L}}_1(x)\varphi(x) := \langle Ax + \bar{F}(x), D_x\varphi(x)\rangle_1 + \frac{1}{2}Tr[D_x^2\varphi(x)Q_1], \quad \forall\varphi \in C_p^2(H_1), \quad (5.1)$$

$$\bar{\mathcal{L}}_3\varphi(z) := \bar{\mathcal{L}}_3(x,z)\varphi(z) := \langle Az + D_x\bar{F}(x)z, D_z\varphi(z)\rangle_1 + \frac{1}{2}Tr\big[D_z^2\varphi(z)\sigma(x)\sigma^*(x)\big], \quad \forall\varphi \in C_p^2(H_1). \quad (5.2)$$

Fix $T > 0$, consider the following Cauchy problem on $[0,T]\times H_1 \times H_1$:

$$\begin{cases} \partial_t\bar{u}(t,x,z) = \bar{\mathcal{L}}\,\bar{u}(t,x,z), & t \in (0,T], \\ \bar{u}(0,x,z) = \varphi(z), \end{cases} \quad (5.3)$$

where $\varphi : H_1 \to \mathbb{R}$ is measurable. We have the following result, which will be used below to prove the weak convergence of Z_t^{ε} to $\bar{Z}_t$.

Theorem 5.1. *For every $\varphi \in \mathbb{C}_b^4(H_1)$, there exists a solution $\bar{u} \in C_b^{1,2,4}([0,T]\times H_1 \times H_1)$ to the equation (5.3) which is given by*

$$\bar{u}(t,x,z) = \mathbb{E}\big[\varphi(\bar{Z}_t(x,z))\big]. \tag{5.4}$$

Moreover, we have:

(i) For any $t \in (0,T]$, $x,z \in H_1$ and $h \in \mathcal{D}((-A)^\beta)$ with $\beta \in [0,1]$,

$$|D_z\bar{u}(t,x,z).(-A)^\beta h| \leqslant C_1\, t^{-\beta}\|h\|_1; \tag{5.5}$$

(ii) For any $t \in (0,T]$, $x,z \in H_1$, $h_1 \in \mathcal{D}((-A)^{\beta_1})$ and $h_2 \in \mathcal{D}((-A)^{\beta_2})$ with $\beta_1,\beta_2 \in [0,1]$,

$$|D_z^2\bar{u}(t,x,z).((-A)^{\beta_1}h_1,(-A)^{\beta_2}h_2)| \leqslant C_2\, t^{-\beta_1-\beta_2}\|h_1\|_1\|h_2\|_1, \tag{5.6}$$

and for any $x,z,h_2 \in H_1$ and $h_1 \in \mathcal{D}((-A)^\beta)$ with $\beta \in [0,1]$,

$$|D_xD_z\bar{u}(t,x,z).((-A)^\beta h_1,h_2)| \leqslant C_2\, t^{-\beta}\|h_1\|_1\|h_2\|_1; \tag{5.7}$$

(iii) For any $t \in (0,T]$, $x,z \in H_1$, $h_1 \in \mathcal{D}((-A)^{\beta_1}), h_2 \in \mathcal{D}((-A)^{\beta_2})$ and $h_3 \in \mathcal{D}((-A)^{\beta_3})$ with $\beta_1,\beta_2,\beta_3 \in [0,1]$,

$$|D_z^3\bar{u}(t,x,z).((-A)^{\beta_1}h_1,(-A)^{\beta_2}h_2,(-A)^{\beta_3}h_3)| \leqslant C_3\, t^{-\beta_1-\beta_2-\beta_3}\|h_1\|_1\|h_2\|_1\|h_3\|_1, \tag{5.8}$$

and for any $x,z,h_3 \in H_1$, $h_1 \in \mathcal{D}((-A)^{\beta_1})$ and $h_2 \in \mathcal{D}((-A)^{\beta_2})$ with $\beta_1,\beta_2 \in [0,1]$,

$$|D_xD_z^2\bar{u}(t,x,z).((-A)^{\beta_1}h_1,(-A)^{\beta_2}h_2,h_3)| \leqslant C_3\, t^{-\beta_1-\beta_2}\|h_1\|_1\|h_2\|_1\|h_3\|_1; \tag{5.9}$$

(iv) For any $t \in (0,T]$, $x \in \mathcal{D}(-A)$ and $z,h \in H_1$,

$$|\partial_tD_z\bar{u}(t,x,z).h| \leqslant C_4\left(t^{-1}(1+\|z\|_1) + \|Ax\|_1 + \|x\|_1^2\right)\|h\|_1; \tag{5.10}$$

(v) For any $t \in (0,T]$, $x \in \mathcal{D}(-A)$ and $z,h_1,h_2 \in H_1$,

$$|\partial_tD_z^2\bar{u}(t,x,z).(h_1,h_2)| \leqslant C_5\left(t^{-1}(1+\|z\|_1) + \|Ax\|_1 + \|x\|_1^2\right)\|h_1\|_1\|h_2\|_1, \tag{5.11}$$

and for any $x \in \mathcal{D}(-A), z,h_1 \in H_1$ and $h_2 \in \mathcal{D}((-A))$,

$$\begin{aligned}|\partial_tD_xD_z\bar{u}(t,x,z).(h_1,h_2)| \leqslant\, & C_5\left(t^{-1}(1+\|z\|_1) + \|Ax\|_1 + \|x\|_1^2\right)\|h_1\|_1\|h_2\|_1 \\ & + C_5\,\|h_1\|_1\|Ah_2\|_1;\end{aligned} \tag{5.12}$$

where C_i, $i = 1,\cdots,5$, are positive constants.

Remark 5.2. *The estimates in (i)-(iii) have been studied in* [7, Proposition 7.1] *when the diffusion coefficient is a constant and in* [6, Theorem 4.2, Theorem 4.3 and Proposition 4.5] *for general nonlinear diffusion coefficients. However, the index β in (5.5), (5.7) and (5.9), β_1,β_2,β_3 in (5.6) and (5.8) are restricted to* $[0,1)$, *which is not sufficient for us to use below. The key observation here is that the equation (2.7) satisfied by $\bar{Z}_t$ is a linear one, and we do not involve estimates for $D_x\bar{u}$ and $D_x^2\bar{u}$. Thus some new techniques are needed in the proof of Theorem 2.4 to avoid using these estimates.*

Proof. (i)-(iii). By using the same argument as in [16, Theorem 13], we can prove that $\bar{u}$ defined by (5.4) is a solution to the equation (5.3). Moreover, $\bar{u}$ has bounded Gatêaux derivatives with respect to the x-variable up to order 2 and with respect to the z-variable up to order 4, see also [7, Section 7] and [6, Section 4]. Furthermore, in view of (5.4) we deduce that for any $\beta \in [0,1]$,

$$D_z\bar{u}(t,x,z).(-A)^\beta h = \mathbb{E}\Big[\langle \varphi'(\bar{Z}_t(x,z)), D_z\bar{Z}_t(x,z).(-A)^\beta h\rangle_1\Big].$$

Since $\bar{Z}_t$ satisfies (2.7), we thus have

$$\mathrm{d}(D_z\bar{Z}_t(x,z).(-A)^\beta h) = \big(A + D_x\bar{F}(\bar{X}_t)\big).(D_z\bar{Z}_t(x,z).(-A)^\beta h)\mathrm{d}t,$$

and the initial value is given by $D_z\bar{Z}_0(x,z).(-A)^\beta h = (-A)^\beta h$. As a result,

$$\|D_z\bar{Z}_t(x,z).(-A)^\beta h\|_1 \leqslant C_0\, t^{-\beta}\|h\|_1,$$

which in turn yields (5.5). Estimates (5.6)-(5.9) can be proved similarly, hence we omit the details here.

(iv) To prove estimate (5.10), by (5.3) we note that for any $h \in H_1$,

$$\partial_t D_z\bar{u}(t,x,z).h = D_z\partial_t\bar{u}(t,x,z).h = D_z(\bar{\mathcal{L}}_1 + \bar{\mathcal{L}}_3)\bar{u}(t,x,z).h. \tag{5.13}$$

By definition (5.1) we have

$$\begin{aligned} D_z\bar{\mathcal{L}}_1\bar{u}(t,x,z).h &= D_zD_x\bar{u}(t,x,z).(Ax + \bar{F}(x), h) \\ &\quad + \frac{1}{2}\sum_{n=1}^\infty \lambda_{1,n}D_zD_x^2\bar{u}(t,x,z).(e_{1,n}, e_{1,n}, h), \end{aligned} \tag{5.14}$$

which implies that

$$|D_z\bar{\mathcal{L}}_1\bar{u}(t,x,z).h| \leqslant C_1(1 + \|Ax\|_1)\|h\|_1. \tag{5.15}$$

Similarly, by definition (5.2) we have

$$\begin{aligned} D_z\bar{\mathcal{L}}_3\bar{u}(t,x,z).h &= \langle Ah + D_x\bar{F}(x).h, D_z\bar{u}(t,x,z)\rangle_1 \\ &\quad + D_z^2\bar{u}(t,x,z).(Az + D_x\bar{F}(x).z, h) \\ &\quad + \frac{1}{2}\sum_{n=1}^\infty D_z^3\bar{u}(t,x,z).(\sigma(x)e_{1,n}, \sigma(x)e_{1,n}, h), \end{aligned} \tag{5.16}$$

which together with (5.5) and (5.6) yields that

$$|D_z\bar{\mathcal{L}}_3\bar{u}(t,x,z).h| \leqslant C_2\, t^{-1}(1 + \|z\|_1)\|h\|_1 + (1 + \|x\|_1^2)\|h\|_1. \tag{5.17}$$

Combining (5.13), (5.15) and (5.17), we obtain (5.10).

(v) In view of (5.13), (5.14) and (5.16), we note that for any $h_1, h_2 \in H_1$,

$$\begin{aligned} \partial_t D_z^2\bar{u}(t,x,z).(h_1,h_2) &= D_z^2D_x\bar{u}(t,x,z).(Ax + \bar{F}(x), h_1, h_2) \\ &\quad + D_z^2\bar{u}(t,x,z).(Ah_1 + D_x\bar{F}(x).h_1, h_2) \end{aligned}$$

$$
\begin{aligned}
&+ D_z^2\bar{u}(t,x,z).(Ah_2 + D_x\bar{F}(x).h_2, h_1)\\
&+ D_z^3\bar{u}(t,x,z).(Az + D_x\bar{F}(x).z, h_1, h_2)\\
&+ \frac{1}{2}\sum_{n=1}^{\infty}\lambda_{1,n}D_z^2D_x^2\bar{u}(t,x,z).(e_{1,n}, e_{1,n}, h_1, h_2)\\
&+ \frac{1}{2}\sum_{n=1}^{\infty}D_z^4\bar{u}(t,x,z).(\sigma(x)e_{1,n}, \sigma(x)e_{1,n}, h_1, h_2).
\end{aligned}
$$

Using (5.6), (5.8) and Lemma 3.7, one can check that

$$
|\partial_t D_z^2\bar{u}(t,x,z).(h_1,h_2)| \leqslant C_3\,(t^{-1} + \|Ax\|_1 + \|x\|_1^2 + t^{-1}\|z\|_1)\|h_1\|_1\|h_2\|_1,
$$

which means that (5.11) holds. Finally, we have

$$
\begin{aligned}
\partial_t D_xD_z\bar{u}(t,x,z).(h_1,h_2) &= D_xD_zD_x\bar{u}(t,x,z).(Ax + \bar{F}(x), h_1, h_2)\\
&+ D_zD_x\bar{u}(t,x,z).(Ah_2 + D_x\bar{F}(x).h_2, h_1)\\
&+ D_xD_z\bar{u}(t,x,z).(Ah_1 + D_x\bar{F}(x).h_1, h_2)\\
&+ \langle D_x^2\bar{F}(x).(h_1,h_2), D_z\bar{u}(t,x,z)\rangle_1 + D_z^2\bar{u}(t,x,z).\big(D_x^2\bar{F}(x).(z,h_2), h_1\big)\\
&+ D_xD_z^2\bar{u}(t,x,z).(Az + D_x\bar{F}(x).z, h_1, h_2)\\
&+ \frac{1}{2}\sum_{n=1}^{\infty}\lambda_{1,n}D_xD_zD_x^2\bar{u}(t,x,z).(e_{1,n}, e_{1,n}, h_1, h_2)\\
&+ \frac{1}{2}\sum_{n=1}^{\infty}D_xD_z^3\bar{u}(t,x,z).(\sigma(x)e_{1,n}, \sigma(x)e_{1,n}, h_1, h_2)\\
&+ \sum_{n=1}^{\infty}D_z^3\bar{u}(t,x,z).((D_x\sigma(x).h_2)e_{1,n}, \sigma(x)e_{1,n}, h_1).
\end{aligned}
$$

Using (5.7), (5.9) and Lemma 3.7 we obtain

$$
\begin{aligned}
|\partial_t D_xD_z\bar{u}(t,x,z).(h_1,h_2)| \leqslant C_4\Big(t^{-1} + \|Ax\|_1 + \|x\|_1^2 + t^{-1}\|z\|_1\Big)\|h_1\|_1\|h_2\|_1\\
+ C_4\|h_1\|_1\|Ah_2\|_1,
\end{aligned}
$$

which yields (5.12). □

5.2. **Estimates for Z_t^ε.** Recall that Z_t^ε satisfies (1.7). In particular, we have

$$
Z_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}}\int_0^t e^{(t-s)A}[\bar{F}(X_s^\varepsilon) - \bar{F}(\bar{X}_s)]\mathrm{d}s + \frac{1}{\sqrt{\varepsilon}}\int_0^t e^{(t-s)A}\delta F(X_s^\varepsilon, Y_s^\varepsilon)\mathrm{d}s. \tag{5.18}
$$

Furthermore, by Theorem 2.2 we get that for any $q \geqslant 1$, $x \in \mathcal{D}((-A)^\theta)$, $y \in \mathcal{D}((-B)^\theta)$ with $\theta > 0$ and $\gamma \in [0, \theta \wedge 1/2)$,

$$\mathbb{E}\|(-A)^\gamma Z_t^\varepsilon\|_1^q < \infty. \tag{5.19}$$

We shall need the following regularity property of Z_t^ε with respect to the time variable.

Lemma 5.3. *Let $T > 0$, $x \in \mathcal{D}((-A)^\theta)$ and $y \in \mathcal{D}((-B)^\theta)$ with $\theta \in (0, 1]$. Assume that (**A1**) and (**A2**) hold, $F \in C_p^{2,\eta}(H_1 \times H_2, H_1)$ and $G \in C_b^{2,\eta}(H_1 \times H_2, H_2)$ with $\eta > 0$. Then for any $q \geqslant 1$, $0 \leqslant s \leqslant t \leqslant T$ and $\vartheta \in (0, \theta)$, there exists a constant $C_{q,T} > 0$ such that*

$$\mathbb{E}\|Z_t^\varepsilon - Z_s^\varepsilon\|_1^q \leqslant C_{q,T}(t-s)^{q\vartheta}.$$

Proof. By (5.18), we have

$$\begin{aligned}
Z_t^\varepsilon - Z_s^\varepsilon &= \frac{1}{\sqrt{\varepsilon}}\int_s^t e^{(t-r)A}(\bar{F}(X_r^\varepsilon) - \bar{F}(\bar{X}_r))\mathrm{d}r \\
&\quad + \big(e^{(t-s)A} - I\big)\frac{1}{\sqrt{\varepsilon}}\int_0^s e^{(s-r)A}(\bar{F}(X_r^\varepsilon) - \bar{F}(\bar{X}_r))\mathrm{d}r \\
&\quad + \frac{1}{\sqrt{\varepsilon}}\int_s^t e^{(t-r)A}\delta F(X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r \\
&\quad + \big(e^{(t-s)A} - I\big)\frac{1}{\sqrt{\varepsilon}}\int_0^s e^{(s-r)A}\delta F(X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r =: \sum_{i=1}^4 \mathscr{Z}_i(t,s).
\end{aligned}$$

Using Minkowski's inequality and Theorem 2.2 with $\gamma = 0$, we get that

$$\mathbb{E}\|\mathscr{Z}_1(t,s)\|_1^q \leqslant C_1\left(\frac{1}{\sqrt{\varepsilon}}\int_s^t \big(\mathbb{E}\|X_r^\varepsilon - \bar{X}_r\|_1^q\big)^{1/q}\mathrm{d}r\right)^q \leqslant C_1\,(t-s)^q.$$

Furthermore, by Proposition 2.1 (ii) we have that for any $\vartheta \in (0,1)$,

$$\begin{aligned}
\mathbb{E}\|\mathscr{Z}_2(t,s)\|_1^q &\leqslant C_2\,(t-s)^{q\vartheta}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^s \Big(\mathbb{E}\|(-A)^\vartheta e^{(s-r)A}(\bar{F}(X_r^\varepsilon) - \bar{F}(\bar{X}_r))\|_1^q\Big)^{1/q}\mathrm{d}r\right)^q \\
&\leqslant C_2\,(t-s)^{q\vartheta}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^s (s-r)^{-\vartheta}\big(\mathbb{E}\|X_r^\varepsilon - \bar{X}_r\|_1^q\big)^{1/q}\mathrm{d}r\right)^q \leqslant C_2\,(t-s)^{q\vartheta}.
\end{aligned}$$

Note that $\delta F(x,y)$ satisfies the centering condition (3.2). As a direct consequence of the fluctuation estimate (4.7), we obtain that

$$\mathbb{E}\|\mathscr{Z}_3(t,s)\|_1^q \leqslant C_3\,(t-s)^{q\theta}.$$

Finally, by making use of Proposition 2.1 (ii) and (4.7) again, we have for any $\vartheta \in (0,\theta)$,

$$\mathbb{E}\|\mathscr{Z}_4(t,s)\|_1^q \leqslant C_4\,(t-s)^{q\vartheta}\,\mathbb{E}\left\|\frac{1}{\sqrt{\varepsilon}}\int_0^s (-A)^\vartheta e^{(s-r)A}\delta F(X_r^\varepsilon, Y_r^\varepsilon)\mathrm{d}r\right\|_1^q$$

$$\leqslant C_4\,(t-s)^{q\vartheta}.$$

Combining the above computations, we get the desired result. □

As in Section 4, to prove Theorem 2.4 we also need to reduce the infinite dimensional problem to a finite dimensional one by the Galerkin approximation. Recall that $X_t^{n,\varepsilon}$ and $\bar{X}_t^n$ are defined by (4.1) and (4.2), respectively. Define

$$Z_t^{n,\varepsilon} := \frac{X_t^{n,\varepsilon}-\bar{X}_t^n}{\sqrt{\varepsilon}}.$$

Then we have

$$\mathrm{d}Z_t^{n,\varepsilon} = AZ_t^{n,\varepsilon}\mathrm{d}t+\varepsilon^{-1/2}[\bar{F}_n(X_t^{n,\varepsilon})-\bar{F}_n(\bar{X}_t^n)]\mathrm{d}t+\varepsilon^{-1/2}\delta F_n(X_t^{n,\varepsilon},Y_t^{n,\varepsilon})\mathrm{d}t,$$

where $\bar{F}_n$ is given by (4.3), and $\delta F_n(x,y):=F_n(x,y)-\bar{F}_n(x)$. Let $\bar{Z}_t^n$ satisfy the following linear equation:

$$\mathrm{d}\bar{Z}_t^n = A\bar{Z}_t^n\mathrm{d}t+D_x\bar{F}_n(\bar{X}_t^n).\bar{Z}_t^n\mathrm{d}t+P_1^n\sigma(\bar{X}_t^n)\mathrm{d}\tilde{W}_t, \tag{5.20}$$

where $\tilde{W}_t$ is a cylindrical Wiener process in H_1, and $\sigma(x)$ is defined by (2.8). We have the following approximation result.

Lemma 5.4. *For every $\varepsilon>0$ and $x,z\in H_1$, we have*

$$\lim_{n\to\infty}\mathbb{E}\Big(\|Z_t^{\varepsilon}-Z_t^{n,\varepsilon}\|_1+\|\bar{Z}_t-\bar{Z}_t^n\|_1\Big)=0. \tag{5.21}$$

Proof. By the definition of $Z_t^{\varepsilon}, Z_t^{\varepsilon,n}$ and Lemma 4.1, we have

$$\lim_{n\to\infty}\mathbb{E}\|Z_t^{\varepsilon}-Z_t^{n,\varepsilon}\|_1\leqslant\lim_{n\to\infty}\frac{1}{\sqrt{\varepsilon}}\Big(\mathbb{E}\|X_t^{\varepsilon}-X_t^{n,\varepsilon}\|_1+\mathbb{E}\|\bar{X}_t-\bar{X}_t^n\|_1\Big)=0.$$

Furthermore, in view of (2.7) and (5.20), we have

$$\begin{aligned}\bar{Z}_t-\bar{Z}_t^n &= \int_0^t e^{(t-s)A}[D_x\bar{F}(\bar{X}_s).\bar{Z}_s-D_x\bar{F}(\bar{X}_s^n).\bar{Z}_s]\mathrm{d}s\\ &\quad+\int_0^t e^{(t-s)A}[D_x\bar{F}(\bar{X}_s^n).\bar{Z}_s-D_x\bar{F}_n(\bar{X}_s^n).\bar{Z}_s^n]\mathrm{d}s\\ &\quad+\int_0^t e^{(t-s)A}[\sigma(\bar{X}_s)-\sigma(\bar{X}_s^n)]P_1^n\mathrm{d}\tilde{W}_s.\end{aligned}$$

Thus we deduce that

$$\begin{aligned}\mathbb{E}\|\bar{Z}_t-\bar{Z}_t^n\|_1^2 &\leqslant C_1\left(\int_0^t\left(\mathbb{E}\|\bar{X}_s-\bar{X}_s^n\|_1^4\right)^{1/4}\left(\mathbb{E}\|\bar{Z}_s\|_1^4\right)^{1/4}\mathrm{d}s\right)^2\\ &\quad+C_1\int_0^t\left(\mathbb{E}\|\bar{X}_s-\bar{X}_s^n\|_1^2+\|D_x\bar{F}-D_x\bar{F}_n\|^2_{\mathscr{L}(H_1)}\right)\mathrm{d}s+C_1\int_0^t\mathbb{E}\|\bar{Z}_s-\bar{Z}_s^n\|_1^2\mathrm{d}s.\end{aligned}$$

By Gronwall's inequality we obtain

$$\mathbb{E}\|\bar{Z}_t - \bar{Z}_t^n\|_1^2 \leqslant C_2\, e^{C_2\, t} \left(\int_0^t \left(\mathbb{E}\|\bar{X}_s - \bar{X}_s^n\|_1^2 + \|D_x\bar{F} - D_x\bar{F}_n\|^2_{\mathscr{L}(H_1)} \right) \mathrm{d}s \right),$$

which yields the desired result. □

5.3. **Proof of Theorem 2.4.** For any $T > 0$ and $\varphi \in \mathbb{C}_b^4(H_1)$, we have for $t \in [0, T]$,

$$\begin{aligned} \left| \mathbb{E}[\varphi(Z_t^\varepsilon)] - \mathbb{E}[\varphi(\bar{Z}_t)] \right| &\leqslant |\mathbb{E}[\varphi(Z_t^\varepsilon)] - \mathbb{E}[\varphi(Z_t^{n,\varepsilon})]| \\ &\quad + \left| \mathbb{E}[\varphi(Z_t^{n,\varepsilon})] - \mathbb{E}[\varphi(\bar{Z}_t^n)] \right| + \left| \mathbb{E}[\varphi(\bar{Z}_t^n)] - \mathbb{E}[\varphi(\bar{Z}_t)] \right|. \end{aligned} \tag{5.22}$$

By making use of (5.21), the first and the last terms on the right-hand of (5.22) converge to 0 as $n \to \infty$. Therefore, in order to prove Theorem 2.4, we only need to show that for any $\gamma \in (0, 1/2)$,

$$\sup_{t\in[0,T]} \left| \mathbb{E}[\varphi(Z_t^{n,\varepsilon})] - \mathbb{E}[\varphi(\bar{Z}_t^n)] \right| \leqslant C_T\, \varepsilon^{\frac{1}{2}-\gamma}, \tag{5.23}$$

where $C_T > 0$ **is a constant independent of** n. As before, we shall only work with the approximation system in the following subsection, and proceed to prove bounds that are uniform with respect to the dimension. To simplify the notations, we shall omit the index n as before.

Fix $T > 0$, and let

$$\begin{aligned} \mathcal{L}_3\varphi(z) &:= \mathcal{L}_3(x, y, \bar{x}, z)\varphi(z) := \langle Az, D_z\varphi(z)\rangle_1 \\ &\quad + \frac{1}{\sqrt{\varepsilon}}\langle \bar{F}(x) - \bar{F}(\bar{x}), D_z\varphi(z)\rangle_1 + \frac{1}{\sqrt{\varepsilon}}\langle \delta F(x, y), D_z\varphi(z)\rangle_1, \quad \varphi \in C_p^1(H_1). \end{aligned} \tag{5.24}$$

We call a function $\phi(t, x, y, \bar{x}, z)$ defined on $[0, T] \times H_1 \times H_2 \times H_1 \times H_1$ admissible, if it is centered, i.e.,

$$\int_{H_2} \phi(t, x, y, \bar{x}, z)\mu^x(\mathrm{d}y) = 0, \quad \forall t > 0, x, \bar{x}, z \in H_1, \tag{5.25}$$

and the following conditions hold:

(H): for any $t \in [0, T)$, $x, z \in H_1$, $y \in H_2$, $\bar{x} \in \mathcal{D}(-A)$ and $h_1, h_2 \in H_1$,

$$\begin{aligned} &|\partial_t\phi(t, x, y, \bar{x}, z)| + |D_x\partial_t\phi(t, x, y, \bar{x}, z).h_1| + |D_z\partial_t\phi(t, x, y, \bar{x}, z).h_2| \\ &\leqslant C_0\,(T - t)^{-1}(1 + \|A\bar{x}\|_1 + \|\bar{x}\|_1^2 + \|z\|_1)(1 + \|x\|_1 + \|y\|_2^p)(\|h_1\|_1 + \|h_2\|_1), \end{aligned} \tag{5.26}$$

and for any $h_3 \in \mathcal{D}((-A))$,

$$\begin{aligned} |D_{\bar{x}}\partial_t\phi(t, x, y, \bar{x}, z).h_3| &\leqslant C_0\Big((T - t)^{-1}(1 + \|A\bar{x}\|_1 + \|\bar{x}\|_1^2 + \|z\|_1)\|h_3\|_1 + \|Ah_3\|_1\Big) \\ &\quad \times (1 + \|x\|_1 + \|y\|_2^p), \end{aligned} \tag{5.27}$$

and for any $h \in \mathcal{D}((-A)^\vartheta)$ with $\vartheta \in [0, 1]$,

$$|D_z\phi(t, x, y, \bar{x}, z).(-A)^\vartheta h| \leqslant C_0\,(T - t)^{-\vartheta}(1 + \|x\|_1 + \|y\|_2^p)\|h\|_1. \tag{5.28}$$

Given an admissible function $\phi(t,x,y,\bar{x},z) \in C_p^{1,2,\eta,2,2}([0,T]\times H_1\times H_2\times H_1\times H_1)$ with $\eta>0$, let $\psi(t,x,y,\bar{x},z)$ solve the following Poisson equation:

$$\mathcal{L}_2(x,y)\psi(t,x,y,\bar{x},z) = -\phi(t,x,y,\bar{x},z), \tag{5.29}$$

and define

$$\overline{\delta F\cdot\nabla_z\psi}(t,x,\bar{x},z) := \int_{H_2}\nabla_z\psi(t,x,y,\bar{x},z).\delta F(x,y)\mu^x(\mathrm{d}y).$$

The following weak fluctuation estimates for an integral functional of $(X_t^\varepsilon, Y_t^\varepsilon, \bar{X}_t, Z_t^\varepsilon)$ will play an important role in proving (5.23). Compared with Lemma 4.2, extra efforts are needed to control the time singularity in the integral.

Lemma 5.5 (Weak fluctuation estimates)**.** *Let $T,\theta>0$, $x\in\mathcal{D}((-A)^\theta)$ and $y\in\mathcal{D}((-B)^\theta)$. Assume that (**A1**), (**A2**) hold, $F\in C_p^{2,\eta}(H_1\times H_2, H_1)$ and $G\in C_b^{2,\eta}(H_1\times H_2,H_2)$. Then for every admissible function $\phi\in C_p^{1,2,\eta,2,2}([0,T]\times H_1\times H_2\times H_1\times H_1)$ with $\eta>0$, $t\in[0,T]$ and $\gamma\in(0,1/2)$, we have*

$$\mathbb{E}\left(\int_0^t \phi(t,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\mathrm{d}s\right) \leqslant C_T\,\varepsilon^{\frac{1}{2}}, \tag{5.30}$$

and

$$\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^t \phi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\mathrm{d}s - \int_0^T \overline{\delta F\cdot\nabla_z\psi}(s,X_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\mathrm{d}s\right) \leqslant C_T\,\varepsilon^{\frac{1}{2}-\gamma}, \tag{5.31}$$

where $C_T>0$ is a constant.

Proof. We divide the proof into two steps.

Step 1. We prove estimate (5.30). By Theorem 3.2, we have that $\psi\in C_p^{1,2,2,2,2}([0,T]\times H_1\times H_2\times H_1\times H_1)$. Thus we can apply Itô's formula to $\psi(t,X_t^\varepsilon,Y_t^\varepsilon,\bar{X}_t,Z_t^\varepsilon)$ to derive that

$$\begin{aligned}
&\mathbb{E}[\psi(t,X_t^\varepsilon,Y_t^\varepsilon,\bar{X}_t,Z_t^\varepsilon)] \\
&= \psi(0,x,y,x,0) + \mathbb{E}\left(\int_0^t(\partial_s+\mathcal{L}_1+\bar{\mathcal{L}}_1+\mathcal{L}_3)\psi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\mathrm{d}s\right) \\
&\quad + \frac{1}{\varepsilon}\mathbb{E}\left(\int_0^t\mathcal{L}_2\psi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\mathrm{d}s\right),
\end{aligned}$$

where $\mathcal{L}_1,\mathcal{L}_2,\bar{\mathcal{L}}_1$ and $\mathcal{L}_3$ are defined by (4.6), (2.6), (5.1) and (5.24), respectively. Multiplying both sides of the above equality by ε and taking into account (5.29), we obtain

$$\begin{aligned}
&\mathbb{E}\left(\int_0^t\phi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\mathrm{d}s\right) \\
&= \varepsilon\,\mathbb{E}\big[\psi(0,x,y,x,0)-\psi(t,X_t^\varepsilon,Y_t^\varepsilon,\bar{X}_t,Z_t^\varepsilon)\big]
\end{aligned}$$

$$
\begin{aligned}
&+\varepsilon\,\mathbb{E}\left(\int_0^t(\partial_s+\bar{\mathcal{L}}_1+\mathcal{L}_1+\mathcal{L}_3)\psi(s,X^\varepsilon_s,Y^\varepsilon_s,\bar{X}_s,Z^\varepsilon_s)\mathrm{d}s\right)\\
=\;&\varepsilon\,\mathbb{E}\big[\psi(0,x,y,x,0)-\psi(0,X^\varepsilon_t,Y^\varepsilon_t,\bar{X}_t,Z^\varepsilon_t)\big]\\
&+\varepsilon\,\mathbb{E}\left(\int_0^t\big[\partial_s\psi(s,X^\varepsilon_s,Y^\varepsilon_s,\bar{X}_s,Z^\varepsilon_s)-\partial_s\psi(s,X^\varepsilon_t,Y^\varepsilon_t,\bar{X}_t,Z^\varepsilon_t)\big]\mathrm{d}s\right)\\
&+\varepsilon\,\mathbb{E}\left(\int_0^t(\mathcal{L}_1+\bar{\mathcal{L}}_1)\psi(s,X^\varepsilon_s,Y^\varepsilon_s,\bar{X}_s,Z^\varepsilon_s)\mathrm{d}s\right)\\
&+\varepsilon\,\mathbb{E}\left(\int_0^t\mathcal{L}_3\psi(s,X^\varepsilon_s,Y^\varepsilon_s,\bar{X}_s,Z^\varepsilon_s)\mathrm{d}s\right)=:\sum_{i=1}^4\mathscr{O}_i(t,\varepsilon).
\end{aligned}
$$

By making use of Theorem 3.2 and Lemma 6.1, we have

$$
\mathscr{O}_1(t,\varepsilon)\leqslant C_1\,\varepsilon\,\mathbb{E}\big(1+\|X^\varepsilon_t\|_1+\|Y^\varepsilon_t\|_2^p\big)\leqslant C_1\,\varepsilon.
$$

For the second term, since ϕ satisfies (5.26) and (5.27), and $t,x,\bar{x},z$ all are parameters in equation (5.29), by Theorem 3.2 we get that ψ satisfies (5.26) and (5.27) too, which together with Lemmas 6.2, 6.3, 6.5, 5.3 and Hölder's inequality implies that for any $\gamma\in(0,1/2)$,

$$
\begin{aligned}
\mathscr{O}_2(t,\varepsilon)\leqslant\;& C_2\,\varepsilon\int_0^t(t-s)^{-1}\Big(\mathbb{E}(\|\bar{X}_s-\bar{X}_t\|_1^3+\|X^\varepsilon_s-X^\varepsilon_t\|_1^3\\
&\qquad+\|Y^\varepsilon_s-Y^\varepsilon_t\|_2^3+\|Z^\varepsilon_s-Z^\varepsilon_t\|_1^3)\Big)^{1/3}\mathrm{d}s\\
&+C_2\,\varepsilon\int_0^t\Big(\mathbb{E}(\|A\bar{X}_s\|^2+\|A\bar{X}_t\|_1^2)\Big)^{1/2}\mathrm{d}s\\
\leqslant\;& C_2\,\varepsilon^{1-\gamma}\leqslant C_2\,\varepsilon^{1/2}.
\end{aligned}
$$

To treat the third term, since for each $t\in[0,T]$, $\psi(t,\cdot,\cdot,\cdot,\cdot)\in C_p^{2,2,2,2}(H_1\times H_1\times H_2\times H_1)$, we have

$$
\begin{aligned}
&\|(\mathcal{L}_1+\bar{\mathcal{L}}_1)\psi(t,x,y,\bar{x},z)\|_1\\
&\leqslant|\langle Ax+F(x,y),D_x\psi(t,x,y,\bar{x},z)\rangle_1|+\frac{1}{2}Tr(Q_1)\|D^2_x\psi(t,x,y,\bar{x},z)\|_{\mathscr{L}(H_1\times H_1,\mathbb{R})}\\
&\quad+|\langle A\bar{x}+\bar{F}(\bar{x}),D_{\bar{x}}\psi(t,x,y,\bar{x},z)\rangle_1|+\frac{1}{2}Tr(Q_1)\|D^2_{\bar{x}}\psi(t,x,y,\bar{x},z)\|_{\mathscr{L}(H_1\times H_1,\mathbb{R})}\\
&\leqslant C_3\big(1+\|A\bar{x}\|_1+\|Ax\|_1+\|y\|_2^p)(1+\|x\|_1+\|y\|_2^p\big).
\end{aligned}
$$

Thus by Lemmas 6.1, 6.4 and 6.5, we have

$$
\mathscr{O}_3(t,\varepsilon)\leqslant C_3\,\varepsilon^{1-\gamma}\leqslant C_3\,\varepsilon^{1/2}.
$$

For the last term, we write

$$\begin{aligned}\mathscr{O}_4(t,\varepsilon) &= \varepsilon\,\mathbb{E}\left(\int_0^t \langle AZ_s^\varepsilon, D_z\psi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\rangle_1 \mathrm{d}s\right)\\ &\quad+\sqrt{\varepsilon}\,\mathbb{E}\left(\int_0^t \langle \bar{F}(X_s^\varepsilon)-\bar{F}(\bar{X}_s), D_z\psi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\rangle_1 \mathrm{d}s\right)\\ &\quad+\sqrt{\varepsilon}\,\mathbb{E}\left(\int_0^t \langle \delta F(X_s^\varepsilon,Y_s^\varepsilon), D_z\psi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\rangle_1 \mathrm{d}s\right) =: \sum_{i=1}^3 \mathscr{O}_{4,i}(t,\varepsilon).\end{aligned}$$

In view of (5.28), Theorem 3.2 and (5.19), we have for $\gamma\in(0,1/2\wedge\theta)$,

$$\mathscr{O}_{4,1}(t,\varepsilon)\leqslant C_4\,\varepsilon\int_0^t (t-s)^{-1+\gamma}\big(\mathbb{E}\|(-A)^\gamma Z_s^\varepsilon\|_1^2\big)^{1/2}\mathrm{d}s\leqslant C_4\,\varepsilon.$$

Furthermore, it is easy to see that

$$\mathscr{O}_{4,2}(t,\varepsilon)+\mathscr{O}_{4,3}(t,\varepsilon)\leqslant C_4\sqrt{\varepsilon}\int_0^t \big(1+\mathbb{E}\|X_s^\varepsilon\|_1^2+\mathbb{E}\|\bar{X}_s\|_1^2+\mathbb{E}\|Y_s^\varepsilon\|_2^{2p}\big)\mathrm{d}s\leqslant C_4\sqrt{\varepsilon}.$$

Combining the above computations, we get the desired result.

Step 2. We proceed to prove estimate (5.31). By following exactly the same argument as in the proof of Step 1, we obtain that for any $\gamma\in(0,1/2)$,

$$\begin{aligned}&\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^t \phi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\mathrm{d}s\right)\\ &=\sqrt{\varepsilon}\,\mathbb{E}\big[\psi(0,x,y,x,0)-\psi(0,X_t^\varepsilon,Y_t^\varepsilon,\bar{X}_t,Z_t^\varepsilon)\big]\\ &+\sqrt{\varepsilon}\,\mathbb{E}\left(\int_0^t \big(\partial_t\psi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)-\partial_t\psi(s,X_t^\varepsilon,Y_t^\varepsilon,\bar{X}_t,Z_t^\varepsilon)\big)\,\mathrm{d}s\right)\\ &+\sqrt{\varepsilon}\,\mathbb{E}\left(\int_0^t (\bar{\mathcal{L}}_1+\mathcal{L}_1)\psi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\mathrm{d}s\right)+\sqrt{\varepsilon}\,\mathbb{E}\left(\int_0^t \mathcal{L}_3\psi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\mathrm{d}s\right)\\ &\leqslant C_1\,\varepsilon^{1/2-\gamma}+\sqrt{\varepsilon}\,\mathbb{E}\left(\int_0^t \mathcal{L}_3\psi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\mathrm{d}s\right).\end{aligned}$$

Now, for the last term we write

$$\begin{aligned}&\sqrt{\varepsilon}\,\mathbb{E}\left(\int_0^t \mathcal{L}_3\psi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\mathrm{d}s\right)\\ &=\sqrt{\varepsilon}\,\mathbb{E}\left(\int_0^t \langle AZ_s^\varepsilon, D_z\psi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\rangle_1\mathrm{d}s\right)\\ &\quad+\mathbb{E}\left(\int_0^t \langle \bar{F}(X_s^\varepsilon)-\bar{F}(\bar{X}_s), D_z\psi(s,X_s^\varepsilon,Y_s^\varepsilon,\bar{X}_s,Z_s^\varepsilon)\rangle_1\mathrm{d}s\right)\end{aligned}$$

$$+\mathbb{E}\left(\int_0^t \langle \delta F(X_s^\varepsilon, Y_s^\varepsilon), D_z\psi(s, X_s^\varepsilon, Y_s^\varepsilon, \bar{X}_s, Z_s^\varepsilon)\rangle_1 \mathrm{d}s\right) =: \sum_{i=1}^3 \tilde{\mathscr{O}}_i(t,\varepsilon).$$

We argue as for $\mathscr{O}_{4,1}(t,\varepsilon)$ to get that

$$\tilde{\mathscr{O}}_1(t,\varepsilon) \leqslant C_2\,\varepsilon^{1/2}.$$

Using Theorem 2.2, we further have

$$\tilde{\mathscr{O}}_2(t,\varepsilon) \leqslant C_3\int_0^t \big(1+\mathbb{E}\|X_s^\varepsilon\|_1^2+\mathbb{E}\|Y_s^\varepsilon\|_2^{2p}\big)^{1/2}\big(\mathbb{E}\|X_s^\varepsilon-\bar{X}_s\|_1^2\big)^{1/2}\mathrm{d}s \leqslant C_3\,\varepsilon^{1/2}.$$

Consequently, we obtain that for any $\gamma\in(0,1/2)$,

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^t \phi(s, X_s^\varepsilon, Y_s^\varepsilon, \bar{X}_s, Z_s^\varepsilon)\mathrm{d}s - \int_0^t \overline{\delta F\cdot\nabla_z\psi}(s, X_s^\varepsilon, \bar{X}_s, Z_s^\varepsilon)\mathrm{d}s\right)\\
&\leqslant C_4\,\varepsilon^{1/2-\gamma} + \mathbb{E}\bigg(\int_0^t \Big(\langle \delta F(X_s^\varepsilon, Y_s^\varepsilon), D_z\psi(s, X_s^\varepsilon, Y_s^\varepsilon, \bar{X}_s, Z_s^\varepsilon)\rangle_1\\
&\qquad\qquad - \overline{\delta F\cdot\nabla_z\psi}(s, X_s^\varepsilon, \bar{X}_s, Z_s^\varepsilon)\Big)\mathrm{d}s\bigg).
\end{aligned}$$

Note that by the definition of $\overline{\delta F\cdot\nabla_z\psi}$, the function

$$\tilde{\phi}(t,x,y,\bar{x},z) := \langle \delta F(x,y), D_z\psi(t,x,y,\bar{x},z)\rangle_1 - \overline{\delta F\cdot\nabla_z\psi}(t,x,\bar{x},z)$$

satisfies the centering condition (5.25) and assumption **(H)**. Thus, using (5.30) directly, we obtain

$$\begin{aligned}
\mathbb{E}\bigg(\int_0^t \Big(&\langle \delta F(X_s^\varepsilon, Y_s^\varepsilon), D_z\psi(s, X_s^\varepsilon, Y_s^\varepsilon, \bar{X}_s, Z_s^\varepsilon)\rangle_1\\
&- \overline{\delta F\cdot\nabla_z\psi}(s, X_s^\varepsilon, \bar{X}_s, Z_s^\varepsilon)\Big)\mathrm{d}s\bigg) \leqslant C_4\,\varepsilon^{1/2}.
\end{aligned}$$

Combining the above computations, we get the desired result. □

Now, we are in the position to give:

Proof of estimate (5.23). Fix $T>0$ below. Let $\bar{u}(t,x,z)$ be the solution of the Cauchy problem (5.3). For $t\in[0,T]$, define

$$\tilde{u}(t,x,z) := \bar{u}(T-t,x,z).$$

Then it is easy to check that

$$\tilde{u}(0,x,0) = \bar{u}(T,x,0) = \mathbb{E}[\varphi(\bar{Z}_T)] \quad \text{and} \quad \tilde{u}(T,x,z) = \bar{u}(0,x,z) = \varphi(z).$$

As a result, by Itô's formula and (5.3) we deduce that

$$\mathbb{E}[\varphi(Z_T^\varepsilon)] - \mathbb{E}[\varphi(\bar{Z}_T)] = \mathbb{E}[\tilde{u}(T, \bar{X}_T, Z_T^\varepsilon) - \tilde{u}(0,x,0)]$$

$$
\begin{aligned}
&= \mathbb{E}\left(\int_0^T \left(\partial_t + \bar{\mathcal{L}}_1 + \mathcal{L}_3\right)\tilde{u}(t,\bar{X}_t,Z_t^\varepsilon)\mathrm{d}t\right)\\
&= \mathbb{E}\left(\int_0^T (\mathcal{L}_3 - \bar{\mathcal{L}}_3)\tilde{u}(t,\bar{X}_t,Z_t^\varepsilon)\mathrm{d}t\right)\\
&= \mathbb{E}\left(\int_0^T \left\langle \frac{\bar{F}(X_t^\varepsilon)-\bar{F}(\bar{X}_t)}{\sqrt{\varepsilon}} - D_x\bar{F}(\bar{X}_t).Z_t^\varepsilon, D_z\tilde{u}(t,\bar{X}_t,Z_t^\varepsilon))\right\rangle_1 \mathrm{d}t\right)\\
&\quad + \frac{1}{2}\mathbb{E}\left(\int_0^T Tr\Big(D_z^2\tilde{u}(t,\bar{X}_t,Z_t^\varepsilon)\big[\sigma(X_t^\varepsilon)\sigma^*(X_t^\varepsilon)-\sigma(\bar{X}_t)\sigma^*(\bar{X}_t)\big]\Big)\mathrm{d}t\right)\\
&\quad + \Bigg[\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^T \langle \delta F(X_t^\varepsilon,Y_t^\varepsilon), D_z\tilde{u}(t,\bar{X}_t,Z_t^\varepsilon))\rangle_1 \mathrm{d}t\right)\\
&\qquad\qquad - \frac{1}{2}\mathbb{E}\left(\int_0^T Tr(D_z^2\tilde{u}(t,\bar{X}_t,Z_t^\varepsilon)\sigma(X_t^\varepsilon)\sigma^*(X_t^\varepsilon))\mathrm{d}t\right)\Bigg] =: \sum_{i=1}^3 \mathscr{N}_i(T,\varepsilon).
\end{aligned}
$$

By the mean value theorem, Hölder's inequality, (5.5), (5.19) and Theorem 2.2, we deduce that for some $\vartheta \in (0,1)$,

$$
\begin{aligned}
\mathscr{N}_1(T,\varepsilon) &\leqslant \mathbb{E}\left(\int_0^T \Big|\big\langle [D_x\bar{F}(X_t^\varepsilon + \vartheta(X_t^\varepsilon - \bar{X}_t)) - D_x\bar{F}(\bar{X}_t)].Z_t^\varepsilon, D_z\tilde{u}(t,\bar{X}_t,Z_t^\varepsilon)\big\rangle_1\Big|\mathrm{d}t\right)\\
&\leqslant C_1\int_0^T \big(\mathbb{E}\|X_t^\varepsilon - \bar{X}_t\|_1^2\big)^{1/2}\big(\mathbb{E}\|Z_t^\varepsilon\|_1^2\big)^{1/2}\mathrm{d}t \leqslant C_1\,\varepsilon^{1/2}.
\end{aligned}
$$

Furthermore, let $\mathscr{U}_{t,\bar{x},z}(x) := Tr(D_z^2\tilde{u}(t,\bar{x},z)\sigma(x)\sigma^*(x))$. Then we have that for every $h \in H_1$,

$$
|D_x\mathscr{U}_{t,\bar{x},z}(x).h| \leqslant C_2(1+\|x\|_1^2)\|h\|_1,
$$

which together with Theorem 2.2, Lemmas 6.1 and 6.5 yields that

$$
\mathscr{N}_2(T,\varepsilon) \leqslant C_2\int_0^T \big(1+\mathbb{E}\|X_t^\varepsilon\|_1^4 + \mathbb{E}\|\bar{X}_t\|_1^4\big)^{1/2}\big(\mathbb{E}\|X_t^\varepsilon - \bar{X}_t\|_1^2\big)^{1/2} \leqslant C_2\,\varepsilon^{1/2}.
$$

It remains to control the last term $\mathscr{N}_3(T,\varepsilon)$. For this purpose, recall that Ψ solves the Poisson equation (2.5), and define

$$
\Phi(t,x,y,\bar{x},z) := \langle \Psi(x,y), D_z\tilde{u}(t,\bar{x},z)\rangle_1.
$$

Since $\mathcal{L}_2$ is an operator with respect to the y variable, one can check that Φ solves the following Poisson equation:

$$
\mathcal{L}_2(x,y)\Phi(t,x,y,\bar{x},z) = -\langle \delta F(x,y), D_z\tilde{u}(t,\bar{x},z)\rangle_1 =: -\phi(t,x,y,\bar{x},z).
$$

It is obvious that ϕ satisfies the centering condition (5.25). Furthermore, in view of (5.6), (5.10), (5.11) and (5.12), we have that for any $t \in [0,T)$, $x,z \in H_1$, $y \in H_2$, $\bar{x} \in \mathcal{D}(-A)$

and $h_1, h_2 \in H_1$,

$$
\begin{aligned}
&|\partial_t \phi(t,x,y,\bar{x},z)| + |D_x \partial_t \phi(t,x,y,\bar{x},z).h_1| + |D_z \partial_t \phi(t,x,y,\bar{x},z).h_2| \\
&\leqslant |\langle \delta F(x,y), \partial_t D_z \bar{u}(T-t,\bar{x},z)\rangle_1| + |\langle D_x \delta F(x,y).h_1, \partial_t D_z \bar{u}(T-t,\bar{x},z)\rangle_1| \\
&\quad + |\partial_t D_z^2 \bar{u}(T-t,\bar{x},z).(\delta F(x,y), h_2)| \\
&\leqslant C_3\,(T-t)^{-1}(\|h_1\|_1 + \|h_2\|_1)(1 + \|A\bar{x}\|_1 + \|\bar{x}\|_1^2 + \|z\|_1)(1 + \|x\|_1 + \|y\|_2^p),
\end{aligned}
$$

and for any $h_3 \in \mathcal{D}(-A)$,

$$
\begin{aligned}
&|D_{\bar{x}} \partial_t \phi(t,x,y,\bar{x},z).h_3| = \partial_t D_{\bar{x}} D_z \bar{u}(T-t,\bar{x},z).(\delta F(x,y), h_3) \\
&\leqslant C_3 \Big((T-t)^{-1}(1 + \|A\bar{x}\|_1 + \|\bar{x}\|_1^2 + \|z\|_1)\|h_3\|_1 + \|Ah_3\|_1 \Big)(1 + \|x\|_1 + \|y\|_2^p),
\end{aligned}
$$

and for any $h \in \mathcal{D}((-A)^\vartheta)$ with $\vartheta \in [0,1]$,

$$
\begin{aligned}
|D_z \phi(t,x,y,\bar{x},z).(-A)^\vartheta h| &= D_z^2 \bar{u}(T-t,\bar{x},z).(\delta F(x,y), (-A)^\vartheta h) \\
&\leqslant C_3\,(T-t)^{-\vartheta}(1 + \|x\|_1 + \|y\|_2^p)\|h\|_1.
\end{aligned}
$$

Furthermore, by the definition of σ in (2.8), we have

$$
\begin{aligned}
&\overline{\delta F \cdot \nabla_z \Phi}(t,x,\bar{x},z) = \int_{H_2} D_z \Phi(t,x,y,\bar{x},z).\delta F(x,y)\mu^x(\mathrm{d}y) \\
&= \int_{H_2} D_z^2 \tilde{u}(t,\bar{x},z).(\Psi(x,y), \delta F(x,y))\mu^x(\mathrm{d}y) = \frac{1}{2} Tr(D_z^2 \tilde{u}(t,\bar{x},z)\sigma(x)\sigma^*(x)).
\end{aligned}
$$

Thus, it follows by (5.31) directly that for any $\gamma \in (0,1/2)$,

$$
\mathscr{N}_3(T,\varepsilon) \leqslant C_3\,\varepsilon^{1/2-\gamma}.
$$

Combining the above computations, we get the desired result. □

6. Appendix

Throughout this section, we assume that (**A1**), (**A2**) hold, $F \in C_p^{1,\eta}(H_1 \times H_2, H_1)$ and $G \in C_b^{1,\eta}(H_1 \times H_2, H_2)$ with $\eta > 0$. We have the following result.

Lemma 6.1. *For any $(x,y) \in H_1 \times H_2$, there exists a unique mild solution for the equation (1.1), i.e., for every $t \geqslant 0$,*

$$
\begin{cases}
X_t^\varepsilon = e^{tA}x + \displaystyle\int_0^t e^{(t-s)A}F(X_s^\varepsilon, Y_s^\varepsilon)\mathrm{d}s + \int_0^t e^{(t-s)A}\mathrm{d}W_s^1, \\
Y_t^\varepsilon = e^{\frac{t}{\varepsilon}B}y + \varepsilon^{-1}\displaystyle\int_0^t e^{\frac{t-s}{\varepsilon}B}G(X_s^\varepsilon, Y_s^\varepsilon)\mathrm{d}s + \varepsilon^{-1/2}\int_0^t e^{\frac{t-s}{\varepsilon}B}\mathrm{d}W_s^2.
\end{cases}
\tag{6.1}
$$

Moreover, for any $T > 0$, $q \geqslant 1$ and $x \in \mathcal{D}((-A)^\theta)$ with $\theta \in [0,1)$, we have

$$
\sup_{\varepsilon \in (0,1)} \sup_{t \in [0,T]} \mathbb{E}\|(-A)^\theta X_t^\varepsilon\|_1^q \leqslant C_{\theta,q,T}\left(1 + \|x\|_{(-A)^\theta}^q + \|y\|_2^{pq}\right) \tag{6.2}
$$

and

$$\sup_{\varepsilon\in(0,1)}\sup_{t\in[0,T]}\mathbb{E}\|Y_t^\varepsilon\|_2^q\leqslant C_{q,T}\,(1+\|y\|_2^q),\tag{6.3}$$

where $C_{\theta,q,T}, C_{q,T}>0$ are constants.

Proof. The well-posedness of SPDE (1.1) with Hölder continuous coefficients follows from [17, Theorem 7]. Furthermore, estimate (6.2) can be proved similarly as in [7, Proposition 2.10] or [10, Proposition 4.3], and estimate (6.3) can be proved as in [10, Proposition 4.2]. We omit the details here. □

Note that estimate (6.2) holds only for $\theta<1$. In order to get the estimate for $\theta=1$, we need some extra regularity results for X_t^ε and Y_t^ε with respect to the time variable. The following two results extend [10, Proposition 4.4] and [4, Proposition A.4], respectively.

Lemma 6.2. *Let $T>0$, $\gamma\in[0,1]$, $x\in\mathcal{D}((-A)^\theta)$ with $\theta\in[0,\gamma]$ and $y\in H_2$. Then for every $q\geqslant 1$ and $0<s\leqslant t\leqslant T$, we have*

$$\Big(\mathbb{E}\|X_t^\varepsilon-X_s^\varepsilon\|_1^q\Big)^{\frac{1}{q}}\leqslant C_{\theta,\gamma,q,T}\left(\frac{(t-s)^\gamma}{s^{\gamma-\theta}}e^{-\frac{\alpha_1}{2}s}\|x\|_{(-A)^\theta}+(t-s)^{\frac{1}{2}}\big(1+\|x\|_1+\|y\|_2^p\big)\right),$$

where $C_{\theta,\gamma,q,T}>0$ is a constant.

Proof. In view of (6.1), we have

$$\begin{aligned}X_t^\varepsilon-X_s^\varepsilon&=(e^{tA}-e^{sA})x+\int_s^t e^{(t-r)A}F(X_r^\varepsilon,Y_r^\varepsilon)\mathrm{d}r\\&\quad+\int_0^s(e^{(t-r)A}-e^{(s-r)A})F(X_r^\varepsilon,Y_r^\varepsilon))\mathrm{d}r\\&\quad+\int_s^t e^{(t-r)A}\mathrm{d}W_r^1+\int_0^s(e^{(t-r)A}-e^{(s-r)A})\mathrm{d}W_r^1=:\sum_{i=1}^5\mathscr{X}_i(t,s).\end{aligned}\tag{6.4}$$

Below, we estimate each term on the right hand side of (6.4) separately. For the first term, by Proposition 2.1 (iii) we easily get

$$\|\mathscr{X}_1(t,s)\|_1\leqslant C_1\,\frac{(t-s)^\gamma}{s^{\gamma-\theta}}e^{-\frac{\alpha_1}{2}s}\|x\|_{(-A)^\theta}.$$

For the second term, by Minkowski's inequality and Lemma 6.1, we deduce that

$$\begin{aligned}\mathbb{E}\|\mathscr{X}_2(t,s)\|_1^q&\leqslant\left(\int_s^t\Big(\mathbb{E}\|e^{(t-r)A}F(X_r^\varepsilon,Y_r^\varepsilon)\|_1^q\Big)^{1/q}\mathrm{d}r\right)^q\\&\leqslant C_2\,(t-s)^q\big(1+\|x\|_1^q+\|y\|_2^{pq}\big).\end{aligned}$$

Similarly, using Proposition 2.1 (ii), Lemma 6.1 and Minkowski's inequality again, we have

$$\begin{aligned}\mathbb{E}\|\mathscr{X}_3(t,s)\|_1^q &\leqslant \left(\int_0^s \left(\mathbb{E}\|(e^{(t-s)A}-I)e^{(s-r)A}F(X_r^\varepsilon,Y_r^\varepsilon)\|_1^q\right)^{1/q}\mathrm{d}r\right)^q\\ &\leqslant C_3\,(t-s)^{q/2}\left(\int_0^s \left(\mathbb{E}\|(-A)^{1/2}e^{(s-r)A}F(X_r^\varepsilon,Y_r^\varepsilon)\|_1^q\right)^{1/q}\mathrm{d}r\right)^q\\ &\leqslant C_3\,(t-s)^{q/2}\big(1+\|x\|_1^q+\|y\|_2^{pq}\big).\end{aligned}$$

Using Burkholder-Davis-Gundy's inequality and the assumption **(A2)**, we further get

$$\mathbb{E}\|\mathscr{X}_4(t,s)\|_1^q \leqslant C_4\left(\int_s^t \|e^{(t-r)A}Q_1^{1/2}\|^2_{\mathscr{L}_2(H_1)}\mathrm{d}r\right)^{q/2} \leqslant C_4\,(t-s)^{q/2},$$

and

$$\mathbb{E}\|\mathscr{X}_5(t,s)\|_1^q \leqslant C_5\,(t-s)^{q/2}\left(\int_0^s \|(-A)^{1/2}e^{(s-r)A}Q_1^{1/2}\|^2_{\mathscr{L}_2(H_1)}\mathrm{d}r\right)^{q/2} \leqslant C_5\,(t-s)^{q/2}.$$

Combining the above computations, we get the desired result. □

Lemma 6.3. *Let $T>0$, $\gamma\in[0,1/2]$, $x\in H_1$ and $y\in\mathcal{D}((-B)^\theta)$ with $\theta\in[0,\gamma]$. Then for every $q\geqslant 1$ and $0<s\leqslant t\leqslant T$, we have*

$$\big(\mathbb{E}\|Y_t^\varepsilon-Y_s^\varepsilon\|_2^q\big)^{\frac{1}{q}} \leqslant C_{\theta,\gamma,q,T}\left(\frac{(t-s)^\gamma}{s^{\gamma-\theta}\varepsilon^\theta}e^{-\frac{\mu_1}{2\varepsilon}s}\|y\|_{(-B)^\theta}+\frac{(t-s)^\gamma}{\varepsilon^\gamma}\right),$$

where $C_{\theta,\gamma,q,T}>0$ is a constant.

Proof. In view of (6.1), we have

$$\begin{aligned}Y_t^\varepsilon-Y_s^\varepsilon &= (e^{\frac{t}{\varepsilon}B}-e^{\frac{s}{\varepsilon}B})y+\frac{1}{\varepsilon}\int_s^t e^{\frac{(t-r)}{\varepsilon}B}G(X_r^\varepsilon,Y_r^\varepsilon)\mathrm{d}r\\ &\quad+\frac{1}{\varepsilon}\int_0^s (e^{\frac{(t-r)}{\varepsilon}B}-e^{\frac{(s-r)}{\varepsilon}B})G(X_r^\varepsilon,Y_r^\varepsilon))\mathrm{d}r\\ &\quad+\frac{1}{\sqrt{\varepsilon}}\int_s^t e^{\frac{(t-r)}{\varepsilon}B}\mathrm{d}W_r^2+\frac{1}{\sqrt{\varepsilon}}\int_0^s (e^{\frac{(t-r)}{\varepsilon}B}-e^{\frac{(s-r)}{\varepsilon}B})\mathrm{d}W_r^2 =: \sum_{i=1}^5\mathscr{Y}_i(t,s)\end{aligned}$$

In exactly the same way as in the proof of Lemma 6.2, we deduce that

$$\|\mathscr{Y}_1(t,s)\|_2 \leqslant C_1\,\frac{(t-s)^\gamma}{s^{\gamma-\theta}\varepsilon^\theta}e^{-\frac{\mu_1}{2\varepsilon}s}\|y\|_{(-B)^\theta},$$

and for any $\gamma\in[0,1/2]$,

$$\mathbb{E}\|\mathscr{Y}_2(t,s)\|_2^q \leqslant C_2\left(\frac{1}{\varepsilon}\int_s^t e^{\frac{-\mu_1}{2\varepsilon}(t-r)}\Big(\mathbb{E}\|G(X_r^\varepsilon,Y_r^\varepsilon)\|_2^q\Big)^{1/q}\mathrm{d}r\right)^q \leqslant C_2\,\frac{(t-s)^{\gamma q}}{\varepsilon^{\gamma q}},$$

and

$$\begin{aligned}
\mathbb{E}\|\mathscr{Y}_3(t,s)\|_1^q &\leqslant \left(\frac{1}{\varepsilon}\int_0^s \left(\mathbb{E}\|(e^{\frac{(t-s)}{\varepsilon}B)}-I)e^{\frac{(s-r)}{\varepsilon}B}G(X_r^\varepsilon,Y_r^\varepsilon)\|_2^q\right)^{1/q}\mathrm{d}r\right)^q \\
&\leqslant C_3\,\frac{(t-s)^{\gamma q}}{\varepsilon^{\gamma q}}\left(\int_0^{s/\varepsilon}\|(-B)^\gamma e^{rB}\|_{\mathscr{L}(H_2)}\mathrm{d}r\right)^q \\
&\leqslant C_3\,\frac{(t-s)^{\gamma q}}{\varepsilon^{\gamma q}}\left(\int_0^{s/\varepsilon} r^{-\gamma}e^{\frac{-\mu_1}{2}r}\mathrm{d}r\right)^q \leqslant C_3\,\frac{(t-s)^{\gamma q}}{\varepsilon^{\gamma q}}.
\end{aligned}$$

To control the last two terms, by Burkholder-Davis-Gundy's inequality and the assumption (**A2**), we deduce that for any $\gamma\in[0,1]$,

$$\begin{aligned}
\mathbb{E}\|\mathscr{Y}_4(t,s)\|_1^q &\leqslant C_4\left(\frac{1}{\varepsilon}\int_s^t \|e^{\frac{(t-r)}{\varepsilon}B}Q_2^{1/2}\|_{\mathscr{L}_2(H_2)}^2\mathrm{d}r\right)^{\frac{q}{2}} \\
&\leqslant C_4\left(\sum_{n=1}^\infty \lambda_{2,n}\mu_n^{\gamma-1}\frac{(t-s)^\gamma}{\varepsilon^\gamma}\right)^{q/2} \leqslant C_4\,\frac{(t-s)^{\frac{\gamma q}{2}}}{\varepsilon^{\frac{\gamma q}{2}}},
\end{aligned}$$

and for $\gamma\in[0,1/2]$,

$$\begin{aligned}
\mathbb{E}\|\mathscr{Y}_5(t,s)\|_1^q &\leqslant C_5\,\frac{(t-s)^{\gamma q}}{\varepsilon^{\gamma q}}\left(\frac{1}{\varepsilon}\int_0^s \|(-B)^\gamma e^{\frac{(s-r)}{\varepsilon}B}Q_2^{1/2}\|_{\mathscr{L}_2(H_2)}^2\mathrm{d}r\right)^{q/2} \\
&\leqslant C_5\,\frac{(t-s)^{\gamma q}}{\varepsilon^{\gamma q}}\left(\sum_{n=1}^\infty \lambda_{2,n}\mu_n^{2\gamma-1}\right)^{q/2} \leqslant C_5\,\frac{(t-s)^{\gamma q}}{\varepsilon^{\gamma q}}.
\end{aligned}$$

Combining the above computations, we get the desired result. □

Now we can prove the following moment estimate.

Lemma 6.4. *Let $T>0$, $x\in\mathcal{D}((-A)^\theta)$ with $\theta\in[0,1]$ and $y\in H_2$. Then for any $q\geqslant 1$, $\gamma>0$ and $0\leqslant t\leqslant T$, we have*

$$\left(\mathbb{E}\|AX_t^\varepsilon\|_1^q\right)^{1/q}\leqslant C_{\theta,\gamma,q,T}\Big(t^{(\theta-1)}+\varepsilon^{-\gamma}\Big)\left(1+\|x\|_{(-A)^\theta}^2+\|y\|_2^{2p}\right),$$

where $C_{\theta,\gamma,q,T}>0$ is a constant.

Proof. We have

$$\begin{aligned}
AX_t^\varepsilon &= Ae^{tA}x+\int_0^t Ae^{(t-s)A}F(X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}s \\
&\quad+\int_0^t Ae^{(t-s)A}\big(F(X_s^\varepsilon,Y_s^\varepsilon)-F(X_t^\varepsilon,Y_t^\varepsilon)\big)\mathrm{d}s+\int_0^t Ae^{(t-s)A}\mathrm{d}W_s^1 =: \sum_{i=1}^4\mathscr{X}_i(t,\varepsilon)
\end{aligned}$$

By Proposition 2.1 (i), we easily see that

$$\|\mathscr{X}_1(t,\varepsilon)\|_1 \leqslant C_1\, t^{(\theta-1)}\|x\|_{(-A)^\theta}.$$

For the second term, note that

$$\int_0^t Ae^{(t-s)A}F(X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}s = \int_0^t \partial_t e^{(t-s)A}F(X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}s = -(e^{tA}-I)F(X_t^\varepsilon,Y_t^\varepsilon),$$

hence we deduce that

$$\mathbb{E}\|\mathscr{X}_2(t,\varepsilon)\|_1^q \leqslant C_2\,(1+\mathbb{E}\|X_t^\varepsilon\|_1^q+\mathbb{E}\|Y_t^\varepsilon\|_2^{pq}) \leqslant C_2\,\big(1+\|x\|_1^q+\|y\|_2^{pq}\big).$$

Furthermore, by applying Lemmas 6.2 and 6.3 with $\theta=0$, we get that for any $\gamma\in(0,1/2]$,

$$\begin{aligned}\mathbb{E}\|\mathscr{X}_3(t,\varepsilon)\|_1^q \leqslant C_3\,(1+\|x\|_1^q+\|y\|_2^{pq})\Bigg(\int_0^t (t-s)^{-1}\bigg[\Big(\mathbb{E}\big[\|X_t^\varepsilon-X_s^\varepsilon\|_1^{2q}\big]\Big)^{1/2q}\\ +\Big(\mathbb{E}\big[\|Y_t^\varepsilon-Y_s^\varepsilon\|_2^{2\eta q}\big]\Big)^{1/2q}\bigg]\mathrm{d}s\Bigg)^q\\ \leqslant C_3\,\big(1+\|x\|_1^{2q}+\|y\|_2^{2pq}\big)\bigg(\int_0^t (t-s)^{\eta\gamma-1}\Big(\frac{1}{s^{\eta\gamma}}+\frac{1}{\varepsilon^{\eta\gamma}}\Big)\mathrm{d}s\bigg)^q\\ \leqslant C_3\,\varepsilon^{-\gamma q}\big(1+\|x\|_1^{2q}+\|y\|_2^{2pq}\big).\end{aligned}$$

Finally, by Burkholder-Davis-Gundy's inequality and assumption **(A2)**, we have

$$\mathbb{E}\|\mathscr{X}_4(t,\varepsilon)\|_1^q \leqslant C_4\left(\int_0^t \|Ae^{(t-s)A}Q_1^{1/2}\|^2_{\mathscr{L}_2(H_1)}\mathrm{d}s\right)^{q/2} \leqslant C_4.$$

The conclusion follows by the above estimates. □

The following results for the averaged equation can be proved as in Lemmas 6.1, 6.2 and 6.4, so we omit the details here.

Lemma 6.5. *For $x\in H_1$, the averaged equation (1.2) has a unique mild solution, i.e., for all $t>0$,*

$$\bar{X}_t = e^{tA}x + \int_0^t e^{(t-s)A}\bar{F}(\bar{X}_s)\mathrm{d}s + \int_0^t e^{(t-s)A}\mathrm{d}W_s^1. \tag{6.5}$$

In addition, we have:

(i) For any $q\geqslant 1$ and $x\in\mathcal{D}(-A)^\theta$ with $\theta\in[0,1)$,

$$\sup_{t\in[0,T]}\mathbb{E}\|(-A)^\theta\bar{X}_t\|_1^q \leqslant C_{\theta,q,T}(1+\|x\|^q_{(-A)^\theta});$$

(ii) For any $q\geqslant 1$, $\theta\in[0,1]$ and $0\leqslant t\leqslant T$,

$$(\mathbb{E}\|A\bar{X}_t\|_1^q)^{1/q} \leqslant C_{\theta,q,T}(1+t^{\theta-1}\|x\|_{(-A)^\theta});$$

(iii) For any $q \geqslant 1$, $\gamma \in [0,1]$, $x \in \mathcal{D}((-A)^\theta)$ with $\theta \in [0,\gamma]$ and $0 < s \leqslant t \leqslant T$,

$$\Big(\mathbb{E}\|\bar{X}_t - \bar{X}_s\|_1^q\Big)^{\frac{1}{q}} \leqslant C_{\theta,\gamma,q,T}\left(\frac{(t-s)^\gamma}{s^{\gamma-\theta}} e^{-\frac{\alpha_1}{2}s}\|x\|_{(-A)^\theta} + (t-s)^{\frac{1}{2}}\big(1+\|x\|_1\big)\right),$$

where $C_{\theta,q,T}, C_{\theta,\gamma,q,T} > 0$ are constants.

References

[1] Arnold L.: Hasslemann's program revisited: The analysis of stochasticity in deterministic climate models. *Stochastic Climate Models*, Progress in Probability book series, Springer, **49** (2001), 141–157.

[2] Bakhtin V. and Kifer Y.: Diffusion approximation for slow motion in fully coupled averaging. *Probab. Theory Related Fields*, **129 (2)** (2004), 157–181.

[3] Ball K., Kurtz T. G., Rempala G. and Popovic L.: Asymptotic analysis of multiscale approximations to reaction networks. *Ann. Appl. Prob.*, **16** (2005), 1925–1961.

[4] Bréhier C. E.: Strong and weak orders in averaging for SPDEs. *Stoch. Process. Appl.*, **122** (2012), 2553–2593.

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

[6] Bréhier C. E. and Debussche A.: Kolmogorov equations and weak order analysis for SPDEs with nonlinear diffusion coefficient. *J. Math. Pures Appl.*, **119** (2018), 193–254.

[7] Bréhier C. E.: Orders of convergence in the averaging principle for SPDEs: the case of a stochastically forced slow component. *Stoch. Proc. Appl.*, **130 (6)** (2020), 3325–3368.

[8] Bogoliubov N. N. and Mitropolsky Y. A.: Asymptotic methods in the theory of non-linear oscillations. Gordon and Breach Science Publishers, New York, 1961.

[9] Cerrai S.: Asymptotic behavior of systems of stochastic partial differential equations with multiplicative noise. Lecture notes in Pure and Applied Math., **245** (2006), 61–75.

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

[11] Cerrai S.: Normal deviations from the averaged motion for some reaction-diffusion equations with fast oscillating perturbation. *J. Math. Pures Appl.*, **91** (2009), 614–647.

[12] Cerrai S.: Averaging principle for systems of reaction-diffusion equations with polynomial nonlinearities perturbed by multiplicative noise. *SIAM J. Math. Anal.*, **43** (2011) 2482–2518.

[13] Cerrai S. and Freidlin M.: Averaging principle for stochastic reaction-diffusion equations. *Probab. Theory Related Fields*, **144** (2009), 137–177.

[14] Cerrai S. and Lunardi A.: Averaging principle for non-autonomous slow-fast systems of stochastic reaction-diffusion equations: the almost periodic case. *SIAM J. Math. Anal.*, **49** (2017), 2843–2884.

[15] Chojnowska-Michalik A. and Goldys B.: Existence, uniqueness and invariant measures for stochastic semilinear equations on Hilbert spaces. *Probab. Theory Related Fields*, **102** (1995), 331–356.

[16] Da Prato G.: Kolmogorov equations for stochastic PDEs with multiplicative noise. *Stoch. Anal. Appl.* **2** (2007), 235–263.

[17] Da Prato G. and Flandoli F.: Pathwise uniqueness for a class of SDEs in Hilbert spaces and applications. *J. Funct. Anal.*, **259** (2010), 243–267.

[18] Dong Z., Sun X., Xiao H. and Zhai J.: Averaging principle for one dimensional stochastic Burgers equation. *J. Diff. Equ.*, **265** (2018), 4749–4797.

[19] E W., Liu D. and Vanden-Eijnden E.: Analysis of multiscale methods for stochastic differential equations. *Comm. Pure Appl. Math.*, **58** (2005), 1544–1585.

[20] Feng J., Fouque, J. P. and Kumar R.: Small-time asymptotics for fast mean reverting stochastic volatility models. *Ann. Appl. Probab.*, **22 (4)** (2012), 1541–1575.

[21] Fu H., Wan L. and Liu J.: Strong convergence in averaging principle for stochastic hyperbolic-parabolic equations with two time-scales. *Stoch. Process. Appl.*, **125 (8)** (2015), 3255–3279.

[22] Gao H. and Duan J.: Dynamics of quasi-geostrophic fluid motion with rapidly oscillating Coriolis force. *Nonl. Anal.: Real World Appl.*, **4** (2003), 127–138.

[23] Goldys B. and Maslowski B.: Exponential ergodicity for stochastic reaction-diffusion equations, Stochastic partial differential equations and applications-VII, 115–131, Lect. Notes Pure Appl. Math., 245, Chapman Hall/CRC, Boca Raton, FL, 2006.

[24] Gonzales-Gargate I. I. and Ruffino P. R.: An averaging principle for diffusions in foliated spaces. *Ann. Probab.*, **44** (2016), 567–588.

[25] Hairer M. and Li X. M.: Averaging dynamics driven by fractional Brownian motion. https://arxiv.org/pdf/1902.11251.pdf.

[26] Hairer M. and Pardoux E.: Homogenization of periodic linear degenerate PDEs. *J. Func. Anal.*, **255** (2008), 2462–2487.

[27] Hairer M. and Pardoux E.: Fluctuations around a homogenised semilinear random PDE. https://arxiv.org/abs/1911.02865.

[28] Hasselmann K.: Stochastic climate models part I. *Theory. Tellus*, **28 (6)** (1976), 473–485.

[29] Kang H. W. and Kurtz T. G.: Separation of time-scales and model reduction for stochastic reaction networks. *Ann. Appl. Probab.*, **23 (2)** (2013), 529–583.

[30] Khasminskii R. Z.: On stochastic processes defined by differential equations with a small parameter. *Theory Probab. Appl.*, **11** (1966), 211–228.

[31] Khasminskii R. Z. and Yin G.: On averaging principles: An asymptotic expansion approach. *SIAM J. Math. Anal.*, **35 (6)** (2004), 1534–1560.

[32] Kifer Y.: Averaging and climate models. *Stochastic Climate Models*, **49** (2001), 171–188.

[33] Konda V. R. and Tsitsiklis J. N.: Convergence rate of linear two-time-scale stochastic approximation. *Ann. Appl. Probab.*, **14** (2004), 796–819.

[34] Majda A., Timofeyev I. and Vanden-Eijnden E.: A mathematical framework for stochastic climate models. *Comm. Pure Appl. Math.*, **54** (2001), 891–974.

[35] Mattingly J. C., Stuart A. M. and Tretyakov M. V.: Convergence of numerical time-averaging and stationary measures via Poisson equations. *SIAM J. Numer. Anal.*, **48** (2010), 552–577.

[36] Pardoux E. and Veretennikov A. Yu.: On the Poisson equation and diffusion approximation. I. *Ann. Probab.*, **29** (2001), 1061–1085.

[37] Pardoux E. and Veretennikov A. Yu.: On the Poisson equation and diffusion approximation 2. *Ann. Probab.*, **31** (2003), 1166–1192.

[38] Pardoux E. and Veretennikov A. Yu.: On the Poisson equation and diffusion approximation 3. *Ann. Probab.*, **33** (2005), 1111–1133.

[39] Pagés G. and Panloup F.: Ergodic approximation of the distribution of a stationary diffusion: rate of convergence. *Ann. Appl. Probab.*, **22** (2012), 1059–1100.

[40] Röckner M. and Xie L.: Diffusion approximation for fully coupled stochastic differential equations. https://bibos.math.uni-bielefeld.de/preprints/19-09-560.pdf.

[41] Veretennikov A. Yu.: On the averaging principle for systems of stochastic differential equations. *Math. USSR Sborn.*, **69** (1991), 271–284.

[42] Wang W. and Roberts A. J.: Average and deviation for slow-fast stochastic partial differential equations. *J. Diff. Equ.*, **253** (2012), 1265–1286.

Michael Röckner: Fakultät für Mathematik, Universität Bielefeld, D-33501 Bielefeld, Germany & Academy of Mathematics and Systems Science, Chinese Academy of Sciences (CAS), Beijing, 100190, P.R.China, Email: roeckner@math.uni-bielefeld.de

Longjie Xie: School of Mathematics and Statistics & Research Institute of Mathematical Science, Jiangsu Normal University, Xuzhou, Jiangsu 221000, P.R.China, Email: longjiexie@jsnu.edu.cn

Li Yang: Department of Mathematics, Shandong University, Jinan, Shandong 250100, P.R.China, Email: llyang@mail.sdu.edu.cn