AVERAGING PRINCIPLE AND NORMAL DEVIATIONS FOR MULTISCALE STOCHASTIC SYSTEMS

MICHAEL RÖCKNER AND LONGJIE XIE

ABSTRACT. We study the asymptotic behavior for an inhomogeneous multiscale stochastic dynamical system with non-smooth coefficients. Depending on the averaging regime and the homogenization regime, two strong convergences in the averaging principle of functional law of large numbers type are established. Then we consider the small fluctuations of the system around its average. Nine cases of functional central limit type theorems are obtained. In particular, even though the averaged equation for the original system is the same, the corresponding homogenization limit for the normal deviation can be quite different due to the difference in the interactions between the fast scales and the deviation scales. We provide quite intuitive explanations for each case. Furthermore, sharp rates both for the strong convergences and the functional central limit theorems are obtained, and these convergences are shown to rely only on the regularity of the coefficients of the system with respect to the slow variable, and do not depend on their regularity with respect to the fast variable, which coincide with the intuition since in the limit equations the fast component has been totally averaged or homogenized out.

AMS 2010 Mathematics Subject Classification: 60H10, 60J60, 60F05.

Keywords and Phrases: Multiscale dynamical systems; averaging principle; central limit theorem; homogenization; Zvonkin's transformation.

CONTENTS

This work is supported by the DFG through CRC 1283, the Alexander-von-Humboldt foundation and NSFC (No. 12090011, 12071186, 11931004).

1. Introduction

Consider the following fast-slow stochastic system in $\mathbb{R}^{d_1+d_2}$:

$$\begin{cases} \mathrm{d}X_t^\varepsilon = \varepsilon^{-1}b(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \varepsilon^{-1/2}\sigma(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}W_t^1, & X_0^\varepsilon = x \in \mathbb{R}^{d_1}, \\ \mathrm{d}Y_t^\varepsilon = F(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + G(Y_t^\varepsilon)\mathrm{d}W_t^2, & Y_0^\varepsilon = y \in \mathbb{R}^{d_2}, \end{cases} \tag{1.1}$$

where $d_1, d_2 \in \mathbb{N}$, $b : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}^{d_1}$, $F : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}^{d_2}$, $\sigma : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}^{d_1} \otimes \mathbb{R}^{d_1}$ and $G : \mathbb{R}^{d_2} \to \mathbb{R}^{d_2} \otimes \mathbb{R}^{d_2}$ are Borel measurable functions, W_t^1, W_t^2 are d_1, d_2-dimensional independent standard Brownian motions respectively, both defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$, and the small parameter $0 < \varepsilon \ll 1$ represents the separation of time scales between the fast motion X_t^ε (with time order $1/\varepsilon$) and the slow component Y_t^ε. Such multiscale models appear frequently in many real world dynamical systems. Typical examples include climate weather interactions (see e.g. [37, 45]), intracellular biochemical reactions (see e.g. [4, 31]), geophysical fluid flows (see e.g. [23]), stochastic volatility in finance (see e.g. [20]), etc. We refer the interested readers to the books [42, 52] for a more comprehensive overview. In fact, as mentioned in [56], almost all physical systems have a certain hierarchy in which not all components evolve at the same rate, and a mathematical description for such phenomena can be formulated by a singularly perturbed differential equation with a small parameter such as the one given in (1.1). Usually, the underlying system (1.1) is difficult to deal with due to the widely separated time scales and the cross interactions between the fast and slow 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.

The intuitive idea for deriving such a simplified equation for system (1.1) is based on the observation that during the fast transients, the slow variable remains "constant" and by the time its changes become noticeable, the fast variable has almost reached its "quasi-steady state". Noting that after the natural time scaling $t \mapsto \varepsilon t$, the process $\tilde{X}_t^\varepsilon := X_{\varepsilon t}^\varepsilon$ satisfies

$$\mathrm{d}\tilde{X}_t^\varepsilon = b(\tilde{X}_t^\varepsilon, Y_{\varepsilon t}^\varepsilon)\mathrm{d}t + \sigma(\tilde{X}_t^\varepsilon, Y_{\varepsilon t}^\varepsilon)\mathrm{d}\tilde{W}_t^1, \quad \tilde{X}_0^\varepsilon = x \in \mathbb{R}^{d_1},$$

where $\tilde{W}_t^1 := \varepsilon^{-1/2}W_{\varepsilon t}^1$ is a new Brownian motion. Thus we need to consider the auxiliary process X_t^y which is the solution of the following frozen stochastic differential equation (SDE for short): for fixed $y \in \mathbb{R}^{d_2}$,

$$\mathrm{d}X_t^y = b(X_t^y, y)\mathrm{d}t + \sigma(X_t^y, y)\mathrm{d}W_t^1, \quad X_0^y = x \in \mathbb{R}^{d_1}. \tag{1.2}$$

The re-scaled process $\tilde{X}_t^\varepsilon$ will be asymptotically identical in distribution to X_t^y. Under certain recurrence conditions, the process X_t^y admits a unique invariant measure $\mu^y(\mathrm{d}x)$. Then by averaging the coefficient with respect to parameters in the fast variable, the slow component Y_t^ε in system (1.1) will converge strongly (in the L^2-sense) as $\varepsilon \to 0$ to the solution of the following so-called averaged equation in $\mathbb{R}^{d_2}$:

$$\mathrm{d}\bar{Y}_t = \bar{F}(\bar{Y}_t)\mathrm{d}t + G(\bar{Y}_t)\mathrm{d}W_t^2, \quad \bar{Y}_0 = y \in \mathbb{R}^{d_2}, \tag{1.3}$$

where the new averaged drift coefficient is defined by

$$\bar{F}(y) := \int_{\mathbb{R}^{d_1}} F(x,y)\mu^y(\mathrm{d}x).$$

The effective dynamics (1.3) does not depend on the fast variable any more and thus is much simpler than SDE (1.1). This theory, known as the averaging principle, can be regarded as classical functional law of large numbers and has been intensively studied in both the deterministic ($\sigma = G \equiv 0$) and the stochastic context in the past decades, see e.g. [14, 15, 16, 24, 27, 33, 34, 36, 44] and the references therein, see also [10, 12, 13, 63] for similar results concerning stochastic partial differential equations (SPDEs for short). Note that the diffusion coefficient G in SDE (1.1) does not depend on the fast variable x, otherwise, the strong convergence may not be true (see e.g. [38]). Meanwhile, as a rule the averaging method requires certain regularities of the coefficients of the original system (1.1) to guarantee the above convergence, and we point out that all the aforementioned papers assumed at least local Lipschitz continuity of all the coefficients. For the averaging principle of SDE (1.1) with non-smooth coefficients, we refer to [54, 60].

However, the effective equation (1.3) is valid only in the limiting sense, and the time scale separation is never infinite in reality. For small but positive ε, the slow process Y_t^ε will experience fluctuations around its averaged motion $\bar{Y}_t$. To leading order, these fluctuations can be captured by characterizing the asymptotic behavior of the normalized difference

$$Z_t^\varepsilon := \frac{Y_t^\varepsilon - \bar{Y}_t}{\sqrt{\varepsilon}}$$

as ε tends to 0. Under extra regularity assumptions on the coefficients and when $G \equiv \mathbb{I}_{d_2}$ (the $d_2 \times d_2$ identity matrix), the deviation process Z_t^ε is known to converge weakly (i.e., in distribution) towards an Ornstein-Uhlenbeck type process $\bar{Z}_t$ with $\bar{Z}_t$ satisfying the following (linear) SDE in $\mathbb{R}^{d_2}$:

$$\mathrm{d}\bar{Z}_t = \nabla_y \bar{F}(\bar{Y}_t)\bar{Z}_t\mathrm{d}t + \zeta(\bar{Y}_t)\mathrm{d}\tilde{W}_t, \tag{1.4}$$

where $\bar{Y}_t$ solves the averaged equation (1.3), $\tilde{W}_t$ is another standard Brownian motion, and the new diffusion coefficient is given by

$$\zeta(y) := \sqrt{\int_0^\infty \int_{\mathbb{R}^{d_1}} \mathbb{E}\big[F(X_t^y(x),y) - \bar{F}(y)\big]\big[F(x,y) - \bar{F}(y)\big]^* \mu^y(\mathrm{d}x)\mathrm{d}t}. \tag{1.5}$$

Such result, also known as the Gaussian approximation, is an analogue of the functional central limit theorem of Donsker. We refer the readers to the fundamental paper by Khasminskii [33], see also [30, 39, 40, 47, 57] for further developments and [11, 29] for the corresponding central limit theorem type results for multiscale SPDEs. Besides having intrinsic interest, the central limit theorem is also useful in applications. In particular, we can get the formal asymptotic expansion

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

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 approximation (see e.g. [1]), which provides better approximations for the original system (1.1), see also [2, 38] and the references therein. We mention that other limit theorems for SDE (1.1) have also been widely studied in the literature, see e.g. [5, 8, 17, 53, 58, 61] for the large deviations and [25, 26, 46] for the moderate deviations.

In this paper, we study a broader class of system, i.e., consider the following inhomogeneous multiscale SDE in $\mathbb{R}^{d_1+d_2}$:

$$\begin{cases} \mathrm{d}X_t^\varepsilon = \alpha_\varepsilon^{-2} b(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \beta_\varepsilon^{-1} c(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \alpha_\varepsilon^{-1}\sigma(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}W_t^1, \\ \mathrm{d}Y_t^\varepsilon = F(t, X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \gamma_\varepsilon^{-1} H(t, X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + G(t, Y_t^\varepsilon)\mathrm{d}W_t^2, \\ X_0^\varepsilon = x \in \mathbb{R}^{d_1}, \quad Y_0^\varepsilon = y \in \mathbb{R}^{d_2}, \end{cases} \tag{1.6}$$

where the small parameters $\alpha_\varepsilon, \beta_\varepsilon, \gamma_\varepsilon \to 0$ as $\varepsilon \to 0$, and without loss of generality we assume $\alpha_\varepsilon^2/\beta_\varepsilon \to 0$ as $\varepsilon \to 0$, and conventionally take $\beta_\varepsilon \equiv 1$ when $c \equiv 0$ and $\gamma_\varepsilon \equiv 1$ when $H \equiv 0$. The infinitesimal generator $\mathscr{L}_\varepsilon$ corresponding to system (1.6) has the form

$$\mathscr{L}_\varepsilon := \frac{1}{\alpha_\varepsilon^2}\mathscr{L}_0(x,y) + \frac{1}{\beta_\varepsilon}\mathscr{L}_3(x,y) + \frac{1}{\gamma_\varepsilon}\mathscr{L}_2(t,x,y) + \mathscr{L}_1(t,x,y),$$

where $\mathscr{L}_0(x,y)$ is given by

$$\mathscr{L}_0 := \mathscr{L}_0(x,y) := \sum_{i,j=1}^{d_1} a^{ij}(x,y)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d_1} b^i(x,y)\frac{\partial}{\partial x_i} \tag{1.7}$$

with $a(x,y) := \sigma\sigma^*(x,y)/2$ (where σ^* denotes the transpose of σ), and

$$\begin{aligned} \mathscr{L}_3 &:= \mathscr{L}_3(x,y) := \sum_{i=1}^{d_1} c^i(x,y)\frac{\partial}{\partial x_i}, \\ \mathscr{L}_2 &:= \mathscr{L}_2(t,x,y) := \sum_{i=1}^{d_2} H^i(t,x,y)\frac{\partial}{\partial y_i}, \\ \mathscr{L}_1 &:= \mathscr{L}_1(t,x,y) := \sum_{i,j}^{d_2} \mathcal{G}^{ij}(t,y)\frac{\partial^2}{\partial y_i \partial y_j} + \sum_{i=1}^{d_2} F^i(t,x,y)\frac{\partial}{\partial y_i} \end{aligned} \tag{1.8}$$

with $\mathcal{G}(t,y) = GG^*(t,y)/2$. Note that there exist two time scales in the fast motion X_t^ε and even the slow process Y_t^ε has a fast varying component. This is known to be important, in particular, for applications in the homogenization of second order parabolic and elliptic equations with singularly perturbed terms, which has its own interest in the theory of PDEs, see e.g. [28, 41, 49] and [22, Chapter IV]. The study of such generalized systems (1.6) was first carried out by Papanicolaou, Stroock and Varadhan [48] for a compact state space for the fast component and time-independent coefficients when $\alpha_\varepsilon = \beta_\varepsilon = \gamma_\varepsilon$, see also [3] for a similar result in terms of PDEs. Later, a non-compact homogeneous case with $c \equiv 0$ and $\alpha_\varepsilon = \gamma_\varepsilon$ was studied in [50, 51] by using the method of the martingale problem and in [35] by the asymptotic expansion approach, see also [55] for a more systematical study. However, all these papers concern only weak convergence of the slow process Y_t^ε in SDE (1.6). In [57], the convergence in probability of Y_t^ε and its small fluctuations around the averaged motion were studied in a particular homogeneous case where $\alpha_\varepsilon = \delta/\sqrt{\varepsilon}$, $\beta_\varepsilon = \delta$, $\gamma_\varepsilon = \delta/\varepsilon$ and with small noise perturbations, i.e., with G replaced by $\sqrt{\varepsilon}G$ in SDE (1.6). To the best of our knowledge, no strong convergence and functional central limit theorems type results as well as rates of convergence in terms of $\varepsilon \to 0$ for SDE (1.6) in the general case have been obtained so far.

We shall first study the strong convergence in the averaging principle for SDE (1.6) with non-smooth coefficients. The main result is given by **Theorem 2.1** below. It turns out that depending on the orders how $\alpha_\varepsilon, \beta_\varepsilon, \gamma_\varepsilon$ go to zero, we need to distinguish two different regimes of interactions, which lead to two different asymptotic behaviors for system (1.6) as $\varepsilon \to 0$, i.e.,

$$\begin{cases} \lim\limits_{\varepsilon\to 0} \dfrac{\alpha_\varepsilon}{\gamma_\varepsilon} = 0 \quad \text{and} \quad \lim\limits_{\varepsilon\to 0} \dfrac{\alpha_\varepsilon^2}{\beta_\varepsilon\gamma_\varepsilon} = 0, & \text{Regime 1;} \\ \lim\limits_{\varepsilon\to 0} \dfrac{\alpha_\varepsilon}{\gamma_\varepsilon} = 0 \quad \text{and} \quad \alpha_\varepsilon^2 = \beta_\varepsilon\gamma_\varepsilon, & \text{Regime 2.} \end{cases} \tag{1.9}$$

If α_ε and α_ε^2 go to zero faster than γ_ε and $\beta_\varepsilon\gamma_\varepsilon$ respectively (Regime 1), we show that the averaged equation for system (1.6) coincides with the traditional case which corresponds to $c = H \equiv 0$; whereas if α_ε goes to zero faster than γ_ε, while α_ε^2 and $\beta_\varepsilon\gamma_\varepsilon$ are of the same order (Regime 2) (which means that the term $\beta_\varepsilon^{-1}c(X_t^\varepsilon, Y_t^\varepsilon)$ is varying fast enough), then the averaging effect of term c and the homogenization effect of term H will occur in the effective dynamics. Furthermore, unlike most previous publications (see e.g. [35, 48, 51, 57]), we will mainly focus on the impact of noises on the averaging principle for system (1.6). Namely, we prove that for non-degenerate noises, the averaging principle holds for system (1.6) with only Hölder continuous drifts (the corresponding deterministic system can even be ill-posed under such weak conditions on the coefficients), and the convergence in the averaging principle relies only on the regularities of the coefficients with respect to the slow component (y-variable), and does not depend on their regularities with respect to the fast term (x-variable). This coincides with the intuition, since in the limit equation the fast variable has been totally averaged out. See **Remark 2.2** and **Remark 2.9** below for more detailed explanations and comparisons of our results with the previous literature on the subject.

Our method to prove the strong convergence in Regime 1 and Regime 2 is unified and rather simple as we do not need the classical time discretisation procedure, which is commonly used in the literature to prove the averaging principle (see e.g. [10, 33, 44, 60, 63]). Two ingredients are crucial in our proof: Zvonkin's transformation and the Poisson equation in the whole space. First of all, due to the low regularity of the coefficients, we shall use Zvonkin's argument as in [54, 60] to transform the equations for Y_t^ε and its average into new ones with better coefficients. Then we employ the result of Poisson equation established in [55] to prove the strong convergence for system (1.6). In both regimes, rates of convergence are also obtained as easy by-products of our arguments. The convergence rates are known to be important for the analysis of numerical schemes for multiscale systems, see e.g. [6, 7, 18, 32]. Moreover, it will play a crucial role below for us to study the homogenization behavior for the fluctuations of Y_t^ε around its average in determining the deviation scales, which in turn implies that the strong convergence rates obtained here are optimal.

After the strong convergence of the functional law of large numbers type is established, we then proceed to study the functional central limit theorem for system (1.6). More precisely, we will be interested in the asymptotic behavior for the normal deviations of Y_t^ε from its averaged motion $\bar{Y}_t^k$ ($k=1,2$ which correspond to Regime 1 and Regime 2 in (1.9)), i.e., to identify the limit of the normalized difference

$$Z_t^{k,\varepsilon} := \frac{Y_t^\varepsilon - \bar{Y}_t^k}{\eta_\varepsilon}$$

with proper choice of deviation scale η_ε such that $\eta_\varepsilon \to 0$ as ε tends to 0. It turns out that the asymptotic limit for $Z_t^{k,\varepsilon}$ is strongly linked to the interactions of the fast scales as well as the deviation scale η_ε. Even though the law of large numbers type limit for Y_t^ε is the same, the homogenization behavior in the functional central limit theorems for the deviation process $Z_t^{k,\varepsilon}$ can be quite different. We need to distinguish three main cases:

Case 0: $H \equiv 0$ in SDE (1.6), i.e., there is no homogenization term in the slow equation. Note that even in this case, the system is still more general than the traditional ones due to the existence of the extra term $\beta_\varepsilon^{-1} c(X_t^\varepsilon, Y_t^\varepsilon)$ in the fast motion. We shall show that the limit equation for the deviation process could be given in terms of the drift c and the solution for an auxiliary Poisson equation involving the drift F.

Case 1: $H \neq 0$ with Regime 1 described in (1.9). In this case, we shall show that the averaging effect of the drift c and the homogenization effect of the fast term H may arise in the limit equation for $Z_t^{1,\varepsilon}$, while the effects involving the drift F as in Caes 0 will not appear any more.

Case 2: $H \neq 0$ with Regime 2 described in (1.9). As mentioned before, in this case homogenization has already occurred even in the averaged equation for Y_t^ε (see (2.4) below). Thus we shall show that the second order homogenization involving the drift c and H may arise in the limit equation for $Z_t^{2,\varepsilon}$. Furthermore, the effects involving the drift F as in Caes 0 may occur again.

The main results are given by **Theorem 2.3**, **Theorem 2.5** and **Theorem 2.7**, respectively. Moreover, it is interesting to find that in each case, depending on the choice of the deviation scale we still get a different limit behavior for $Z_t^{k,\varepsilon}$. Several new terms will appear which seem to be never observed before in the literature. In certain regimes the limit equation for $Z_t^{k,\varepsilon}$ can even be given by a linear random ordinary differential equation (ODE for short), while in certain situations an extra Gaussian part appears and the limit equation will be given by a linear SDE. We shall provide some quite intuitive explanations for each respective result, see **Remark 2.4**, **Remark 2.6** and **Remark 2.8** below. In particular, our results lead to a deep understanding of the effects and the interactions between the extra averaging term $\beta_\varepsilon^{-1}c(X_t^\varepsilon, Y_t^\varepsilon)$ and the homogenization term $\gamma_\varepsilon^{-1}H(t, X_t^\varepsilon, Y_t^\varepsilon)$ in system (1.6).

We will prove the functional central limit theorem for system (1.6) in each regime in a very robust and unified way. Our method relies only on the technique of Poisson equation, and neither involves an extra time discretisation procedure, nor martingale problem or tightness arguments (see e.g. [11, 33, 39, 57, 63]) and thus is quite simple. Moreover, the conditions on the coefficients are weaker than in the known results in the literature even in the traditional case (i.e., $c = H \equiv 0$), and rates of convergence are also obtained, which we believe are rather sharp. Furthermore, it will be pretty clear from our approach that which parts should be the leading terms in the fluctuations (whose effects arise in the homogenization procedure), which parts should be the lower order terms (whose fluctuations go to zero eventually) and what the deviation scales η_ε should be in each regime in order to observe non-trivial behavior for the limits.

The rest of this paper proceeds as follows. In Section 2 we state our main results. Section 3 is devoted to the preparation of the main tools that shall be used to prove the results. Then we prove Theorem 2.1 in Section 4, Theorem 2.3 in Section 5, and Theorem 2.5 and Theorem 2.7 in Section 6, respectively. Throughout our paper, we use the following convention: C and c with or without subscripts will denote positive constants, whose values may change in different places, and whose dependence on parameters can be traced from the calculations. Given a function space, the subscript b will stand for boundness, while the subscript p stands for polynomial growth in the x variable. To be more precise, for a function $f(t, x, y, z)$ defined on $\mathbb{R}_+ \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2+d_2}$, by $f \in L_p^\infty := L_p^\infty(\mathbb{R}_+ \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2+d_2})$ we mean there exist constants $C, m > 0$ such that

$$|f(t, x, y, z)| \leqslant C(1 + |x|^m), \quad \forall t > 0, x \in \mathbb{R}^{d_1}, y, z \in \mathbb{R}^{d_2},$$

and $C_p^{\gamma,\delta,\vartheta,\eta} := C_{t,x,y,z}^{\gamma,\delta,\vartheta,\eta}(\mathbb{R}_+ \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2+d_2})$ with $0 < \delta \leqslant 1$ denotes the space of all functions f such that for every fixed $x \in \mathbb{R}^{d_1}$, $f(\cdot, x, \cdot, \cdot) \in C_b^{\gamma,\vartheta,\eta}(\mathbb{R}_+ \times \mathbb{R}^{d_2+d_2})$ and for any $(t, y, z) \in \mathbb{R}_+ \times \mathbb{R}^{d_2+d_2}$ and $x_1, x_2 \in \mathbb{R}^{d_1}$,

$$|f(t, x_1, y, z) - f(t, x_2, y, z)| \leqslant C|x_1 - x_2|^\delta \big(1 + |x_1|^m + |x_2|^m\big).$$

2. Statement of main results

2.1. **Strong convergence: functional law of large numbers.** Let us first introduce some basic assumptions. Throughout this paper, we shall always assume the following non-degeneracy conditions on the diffusion coefficients:

($\mathbf{A}_\sigma$): the coefficient $a = \sigma\sigma^*/2$ is non-degenerate in x uniformly with respect to y, i.e., there exists a $\lambda > 1$ such that for any $y \in \mathbb{R}^{d_2}$,

$$\lambda^{-1}|\xi|^2 \leqslant |\sigma^*(x, y)\xi|^2 \leqslant \lambda|\xi|^2, \quad \forall \xi \in \mathbb{R}^{d_1}.$$

($\mathbf{A}_G$): the coefficient $\mathcal{G} = GG^*/2$ is non-degenerate in y uniformly with respect to t, i.e., there exists a $\lambda > 1$ such that for any $t > 0$,

$$\lambda^{-1}|\xi|^2 \leqslant |G^*(t, y)\xi|^2 \leqslant \lambda|\xi|^2, \quad \forall \xi \in \mathbb{R}^{d_2}.$$

Recall that the frozen equation is given by SDE (1.2). We make the following very weak recurrence assumption on the drift b to ensure the existence of an invariant measure $\mu^y(\mathrm{d}x)$ for X_t^y (cf. [62]):

$(\mathbf{A}_b)$: $\lim_{|x|\to\infty}\sup_y \langle x, b(x,y)\rangle = -\infty.$

Note that the drift c in SDE (1.6) is not involved in the frozen equation. We need the following additional condition on c to ensure the non-explosion of the solution X_t^ε: for $\varepsilon>0$ small enough, it holds that

$$\lim_{|x|\to\infty}\sup_y \langle x, b(x,y)+\varepsilon c(x,y)\rangle = -\infty. \tag{2.1}$$

Concerning the fast term in the slow component of SDE (1.6), it is natural to make the following assumption:

$(\mathbf{A}_H)$: the drift H is centered, i.e.,

$$\int_{\mathbb{R}^{d_1}} H(t,x,y)\mu^y(\mathrm{d}x) = 0, \quad \forall (t,y)\in\mathbb{R}_+\times\mathbb{R}^{d_2}, \tag{2.2}$$

where $\mu^y(\mathrm{d}x)$ is the invariant measure for SDE (1.2).

Under (2.2) and according to Theorem 3.1 below, there exists a unique solution $\Phi(t,x,y)$ to the following Poisson equation in $\mathbb{R}^{d_1}$:

$$\mathscr{L}_0(x,y)\Phi(t,x,y) = -H(t,x,y), \quad x\in\mathbb{R}^{d_1}, \tag{2.3}$$

where $\mathscr{L}_0(x,y)$ is given by (1.7), and $(t,y)\in\mathbb{R}_+\times\mathbb{R}^{d_2}$ are regarded as parameters. We introduce the new averaged drift coefficients by

$$\begin{aligned}\bar{F}_1(t,y) &:= \int_{\mathbb{R}^{d_1}} F(t,x,y)\mu^y(\mathrm{d}x);\\ \bar{F}_2(t,y) &:= \int_{\mathbb{R}^{d_1}} \big[F(t,x,y)+c(x,y)\cdot\nabla_x\Phi(t,x,y)\big]\mu^y(\mathrm{d}x),\end{aligned} \tag{2.4}$$

which correspond to Regime 1 and Regime 2 described in (1.9), respectively. Then the precise formulation of the averaged equation for SDE (1.6) is as follows: for $k=1,2$,

$$\mathrm{d}\bar{Y}_t^k = \bar{F}_k(t,\bar{Y}_t^k)\mathrm{d}t + G(t,\bar{Y}_t^k)\mathrm{d}W_t^2, \quad \bar{Y}_0^k = y\in\mathbb{R}^{d_2}. \tag{2.5}$$

The following is the first main result of this paper.

Theorem 2.1 (Strong convergence)**.** *Let* $(\mathbf{A}_\sigma)$, $(\mathbf{A}_b)$, $(\mathbf{A}_G)$, $(\mathbf{A}_H)$ *and (2.1) hold,* $\delta,\vartheta>0$ *and* $\lim_{\varepsilon\to0}\alpha_\varepsilon^\vartheta/\gamma_\varepsilon = 0$. *Then for any $T>0$ and every $q\geqslant 1$,*

(i) (Regime 1) if $b,\sigma\in C_b^{\delta,\vartheta}$, $F,H\in C_p^{\vartheta/2,\delta,\vartheta}$, $G\in C_b^{\vartheta/2,1}$ *and* $c\in L_p^\infty$, *we have*

$$\sup_{t\in[0,T]}\mathbb{E}|Y_t^\varepsilon-\bar{Y}_t^1|^q \leqslant C_T\Big(\frac{\alpha_\varepsilon^{\vartheta\wedge1}}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon\gamma_\varepsilon}\Big)^q; \tag{2.6}$$

(ii) (Regime 2) if $b,\sigma\in C_b^{\delta,\vartheta}$, $F,H\in C_p^{\vartheta/2,\delta,\vartheta}$, $G\in C_b^{\vartheta/2,1}$ *and* $c\in C_p^{\delta,\vartheta}$, *we have*

$$\sup_{t\in[0,T]}\mathbb{E}|Y_t^\varepsilon-\bar{Y}_t^2|^q \leqslant C_T\Big(\frac{\alpha_\varepsilon^{\vartheta\wedge1}}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^q, \tag{2.7}$$

where for $k=1,2$, $\bar{Y}^k_t$ is the unique strong solution for SDE (2.5), and $C_T>0$ is a constant independent of δ,ε.

Let us list some important comments regarding the above result.

Remark 2.2. *(i) [Non-smooth coefficients]. Note that homogenization occurs in Regime 2 and an additional drift part appears in the limit. In both regimes, we do not make any Lipschitz-type assumptions on the drift coefficients b, c, F and H. We mention that if $\sigma=0$ or $G=0$, the corresponding deterministic system may even be ill-posed for only Hölder continuous coefficients. This reflects the regularization effects of the noises. In fact, under our assumptions we have for every $k=1,2$, $\bar{F}_k\in C_b^{\vartheta/2,\vartheta}$. Thus, the strong well-posedness for system (1.6) and SDE (2.5) follows by* [59] *or* [64, Theorem 1.3].

We also point out that the above results still hold in the small noise perturbation case, i.e., with G replaced by $\lambda_\varepsilon G$, where $\lambda_\varepsilon\to 0$ as $\varepsilon\to 0$. Then we need to assume the coefficients b, c, F and H to be Lipschitz continuous with respect to the y variable in order to ensure the well-posedness of the averaged system. For the sake of simplicity, we do not deal with this setting in the present article.

(ii) [Dependence of convergence]. In both regimes, the convergence rates do not depend on the index δ. This suggests that the convergence in the averaging principle relies only on the regularities of the coefficients in the original system with respect to the time variable and the y (slow) variable, and does not depend on their regularities with respect to the x (fast) variable.

(iii) [Sharp rates]. The traditional result can be viewed as a particular case of Regime 1 by taking $c=H\equiv 0$ (i.e., $\beta_\varepsilon=\gamma_\varepsilon\equiv 1$). In this case, our result implies that the averaging principle holds for SDE (1.6) with a strong convergence rate $\alpha_\varepsilon^\vartheta$ when the coefficients are ϑ-Hölder continuous. This order is known to be optimal when $\vartheta=1$. In the general case and when $\vartheta=1$, estimate (2.6) means that in Regime 1, the slow process Y^ε_t will converge to $\bar{Y}^1_t$ strongly with rate $\frac{\alpha_\varepsilon}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon\gamma_\varepsilon}$, while estimate (2.7) suggests that in Regime 2, Y^ε_t converges to $\bar{Y}^2_t$ strongly with order $\frac{\alpha_\varepsilon}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}$. We shall show that these rates are also optimal by studying the respective functional central limit theorems.

2.2. **Functional central limit theorem: without homogenization.** We first consider SDE (1.6) with $H\equiv 0$, i.e., there is no fast term in the slow component. To avoid confusion of notations, we shall denote by $Y^{0,\varepsilon}_t$ the slow process. More precisely, consider

$$\begin{cases}\mathrm{d}X^\varepsilon_t=\alpha_\varepsilon^{-2}b(X^\varepsilon_t,Y^{0,\varepsilon}_t)\mathrm{d}t+\beta_\varepsilon^{-1}c(X^\varepsilon_t,Y^{0,\varepsilon}_t)\mathrm{d}t+\alpha_\varepsilon^{-1}\sigma(X^\varepsilon_t,Y^{0,\varepsilon}_t)\mathrm{d}W^1_t,\\ \mathrm{d}Y^{0,\varepsilon}_t=F(t,X^\varepsilon_t,Y^{0,\varepsilon}_t)\mathrm{d}t+G(t,Y^{0,\varepsilon}_t)\mathrm{d}W^2_t,\\ X^\varepsilon_0=x\in\mathbb{R}^{d_1},\quad Y^{0,\varepsilon}_0=y\in\mathbb{R}^{d_2}.\end{cases}\tag{2.8}$$

Note that even in this case, the above system is still broader than the traditional ones due to the existence of the extra term $\beta_\varepsilon^{-1}c(X^\varepsilon_t,Y^{0,\varepsilon}_t)$ in the fast equation. According to Theorem 2.1 (i) with $H\equiv 0$ (then $\gamma_\varepsilon\equiv 1$) and $\vartheta=1$, the slow process $Y^{0,\varepsilon}_t$ will converge to $\bar{Y}^1_t$ strongly with a best possible rate $\alpha_\varepsilon+\alpha_\varepsilon^2/\beta_\varepsilon$. We intend to study the small fluctuations of $Y^{0,\varepsilon}_t$ from its average $\bar{Y}^1_t$, i.e., to characterize the asymptotic behavior of the normalized difference

$$Z_t^{0,\varepsilon} := \frac{Y_t^{0,\varepsilon} - \bar{Y}_t^1}{\eta_\varepsilon} \tag{2.9}$$

with proper deviation scale η_ε such that $\eta_\varepsilon \to 0$ as $\varepsilon \to 0$. It turns out that the limit behavior for $Z_t^{0,\varepsilon}$ is strongly linked to the deviation scale η_ε. Formally, if α_ε goes to 0 faster than $\alpha_\varepsilon^2/\beta_\varepsilon$, then the convergence rate of $Y_t^{0,\varepsilon}$ to $\bar{Y}_t^1$ is dominated by $\alpha_\varepsilon^2/\beta_\varepsilon$. Thus one needs time of order $\alpha_\varepsilon^2/\beta_\varepsilon$ to observe non-trivial behavior for $Z_t^{0,\varepsilon}$; while if α_ε is of the same order or lower order than $\alpha_\varepsilon^2/\beta_\varepsilon$, then $Y_t^{0,\varepsilon}$ will converge to $\bar{Y}_t^1$ with rate α_ε and we shall need deviation scale α_ε to observe non-trivial homogenization effects. Consequently, the natural choice of the deviation scale η_ε should be divided into the following three regimes:

$$\begin{cases} \eta_\varepsilon = \dfrac{\alpha_\varepsilon^2}{\beta_\varepsilon} \quad \text{and} \quad \lim\limits_{\varepsilon\to 0} \dfrac{\beta_\varepsilon}{\alpha_\varepsilon} = 0, & \text{Regime 0.1;} \\ \eta_\varepsilon = \alpha_\varepsilon \quad \text{and} \quad \lim\limits_{\varepsilon\to 0} \dfrac{\alpha_\varepsilon}{\beta_\varepsilon} = 0, & \text{Regime 0.2;} \\ \eta_\varepsilon = \alpha_\varepsilon = \beta_\varepsilon, & \text{Regime 0.3.} \end{cases} \tag{2.10}$$

Such choices of η_ε will also appear to be natural from our proof procedure. Let $\Upsilon(t,x,y)$ be the unique solution of the following Poisson equation in $\mathbb{R}^{d_1}$:

$$\mathscr{L}_0(x,y)\Upsilon(t,x,y) = -\big[F(t,x,y) - \bar{F}_1(t,y)\big] := -\tilde{F}(t,x,y), \tag{2.11}$$

where $\mathscr{L}_0(x,y)$ is defined by (1.7), $\bar{F}_1$ is given by (2.4), and $(t,y) \in \mathbb{R}_+ \times \mathbb{R}^{d_2}$ are regarded as parameters. Define

$$\begin{aligned} \overline{c\cdot\nabla_x\Upsilon}(t,y) &:= \int_{\mathbb{R}^{d_1}} c(x,y)\cdot\nabla_x\Upsilon(t,x,y)\mu^y(\mathrm{d}x), \\ \overline{\tilde{F}\cdot\Upsilon^*}(t,y) &:= \int_{\mathbb{R}^{d_1}} \tilde{F}(t,x,y)\cdot\Upsilon^*(t,x,y)\mu^y(\mathrm{d}x). \end{aligned} \tag{2.12}$$

Then the limit processes $\bar{Z}_{\ell,t}^0$ ($\ell = 1,2,3$) for $Z_t^{0,\varepsilon}$ corresponding to Regime 0.1-Regime 0.3 in (2.10) turn out to satisfy the following linear equations:

$$\begin{aligned} \mathrm{d}\bar{Z}_{1,t}^0 &= \nabla_y\bar{F}_1(t,\bar{Y}_t^1)\bar{Z}_{1,t}^0\mathrm{d}t + \nabla_y G(t,\bar{Y}_t^1)\bar{Z}_{1,t}^0\mathrm{d}W_t^2 + \overline{c\cdot\nabla_x\Upsilon}(t,\bar{Y}_t^1)\mathrm{d}t; \\ \mathrm{d}\bar{Z}_{2,t}^0 &= \nabla_y\bar{F}_1(t,\bar{Y}_t^1)\bar{Z}_{2,t}^0\mathrm{d}t + \nabla_y G(t,\bar{Y}_t^1)\bar{Z}_{2,t}^0\mathrm{d}W_t^2 + \sqrt{\overline{\tilde{F}\cdot\Upsilon^*}(t,\bar{Y}_t^1)}\mathrm{d}\tilde{W}_t; \\ \mathrm{d}\bar{Z}_{3,t}^0 &= \nabla_y\bar{F}_1(t,\bar{Y}_t^1)\bar{Z}_{3,t}^0\mathrm{d}t + \nabla_y G(t,\bar{Y}_t^1)\bar{Z}_{3,t}^0\mathrm{d}W_t^2 \\ &\quad + \overline{c\cdot\nabla_x\Upsilon}(t,\bar{Y}_t^1)\mathrm{d}t + \sqrt{\overline{\tilde{F}\cdot\Upsilon^*}(t,\bar{Y}_t^1)}\mathrm{d}\tilde{W}_t, \end{aligned} \tag{2.13}$$

with initial data $\bar{Z}_{\ell,0}^0 = 0$, where $\bar{Y}_t^1$ is the unique strong solution for SDE (2.5) with $k=1$, and $\tilde{W}_t$ is another Brownian motion independent of W_t^2.

Our first functional central limit theorem type result is as follows.

Theorem 2.3 (Central limit theorem)**.** *Let* $(\mathbf{A}_\sigma)$*,* $(\mathbf{A}_b)$*,* $(\mathbf{A}_G)$ *and (2.1) hold,* $0 < \delta, \vartheta \leqslant 1$*. Then for any* $T > 0$ *and every* $\varphi \in C_b^4(\mathbb{R}^{d_2})$*,* *(i) (Regime 0.1) if $b,\sigma\in C_b^{\delta,1+\vartheta}$, $F\in C_p^{(1+\vartheta)/2,\delta,1+\vartheta}$, $G\in C_b^{1/2,1+\vartheta}$ and $c\in C_p^{\delta,\vartheta}$, we have*

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[\varphi(Z_t^{0,\varepsilon})]-\mathbb{E}[\varphi(\bar{Z}_{1,t}^0)]\Big|\leqslant C_T\Big(\frac{\beta_\varepsilon^2}{\alpha_\varepsilon^2}+\frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta}\Big);$$

(ii) (Regime 0.2) if $b,\sigma\in C_b^{\delta,1+\vartheta}$, $F\in C_p^{(1+\vartheta)/2,\delta,1+\vartheta}$, $G\in C_b^{1/2,1+\vartheta}$ and $c\in L_p^\infty$, we have

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[\varphi(Z_t^{0,\varepsilon})]-\mathbb{E}[\varphi(\bar{Z}_{2,t}^0)]\Big|\leqslant C_T\Big(\frac{\alpha_\varepsilon}{\beta_\varepsilon}+\alpha_\varepsilon^\vartheta\Big);$$

(iii) (Regime 0.3) if $b,\sigma\in C_b^{\delta,1+\vartheta}$, $F\in C_p^{(1+\vartheta)/2,\delta,1+\vartheta}$, $G\in C_b^{1/2,1+\vartheta}$ and $c\in C_p^{\delta,\vartheta}$, we have

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[\varphi(Z_t^{0,\varepsilon})]-\mathbb{E}[\varphi(\bar{Z}_{3,t}^0)]\Big|\leqslant C_T\,\alpha_\varepsilon^\vartheta,$$

where for $\ell=1,2,3$, $\bar{Z}_{\ell,t}^0$ satisfy the linear equation (2.13), and $C_T>0$ is a constant independent of δ,ε.

Remark 2.4. *(i) By Theorem 3.1 below, we have*

$$\overline{\tilde{F}\cdot\Upsilon^*}(t,y)=\int_0^\infty\int_{\mathbb{R}^{d_1}}\mathbb{E}\big[\tilde{F}(t,x,y)\tilde{F}^*(t,X_t^y(x),y)\big]\mu^y(\mathrm{d}x)\mathrm{d}t.$$

Thus, in view of (1.4) and (1.5) the classical result can be viewed as a particular case of Regime 0.2 by taking the coefficients to be time-independent, $c\equiv 0$ (then $\beta_\varepsilon\equiv 1$) and $G\equiv\mathbb{I}_{d_2}$. Even in this case, our result is still new in the sense that the conditions on the coefficients are weaker and the rate of convergence (i.e., $\alpha_\varepsilon^\vartheta$) is obtained, which again depends only on the regularity of the coefficients with respect to the slow variable.

(ii) The above result reveals the effect of the extra fast term $\beta_\varepsilon^{-1}c(X_t^\varepsilon,Y_t^{0,\varepsilon})$ in system (2.8): even though it does not play any role in the functional law of large numbers for $Y_t^{0,\varepsilon}$, it does affect the deviations of $Y_t^{0,\varepsilon}$ from $\bar{Y}_t^1$ through the functional central limit theorem. Note that both in Regime 0.2 and Regime 0.3 there exists an additional Gaussian part involving the drift F in the limit equations. In particular, if $G\equiv\mathbb{I}_{d_2}$, the limit is an Ornstein-Uhlenbeck type process. While in Regime 0.1 (i.e., $\beta_\varepsilon/\alpha_\varepsilon\to 0$, which implies the term $\beta_\varepsilon^{-1}c(X_t^\varepsilon,Y_t^{0,\varepsilon})$ is varying fast enough), there exists only a new drift part involving the term c in the limit equation for $\bar{Z}_{1,t}^0$. In particular, if $G\equiv\mathbb{I}_{d_2}$, then $\bar{Z}_{1,t}^0$ will satisfy a linear random ODE.

(iii) Let us give some intuitive explanations for the above result. If $\beta_\varepsilon/\alpha_\varepsilon\to 0$ (Regime 0.1), then $Y_t^{0,\varepsilon}$ will converge to $\bar{Y}_t^1$ with rate $\alpha_\varepsilon^2/\beta_\varepsilon$. This means that the fast term $\beta_\varepsilon^{-1}c(X_t^\varepsilon,Y_t^{0,\varepsilon})$ is the dominant term in the strong convergence. Thus its averaging effect appears in the functional central limit theorem. While if $\alpha_\varepsilon/\beta_\varepsilon\to 0$ (Regime 0.2), then $Y_t^{0,\varepsilon}$ converges to $\bar{Y}_t^1$ with order α_ε (independent of β_ε). This suggests that the term $\beta_\varepsilon^{-1}c(X_t^\varepsilon,Y_t^{0,\varepsilon})$ is of lower order now, whose effect will go to zero eventually in the homogenization procedure. Finally, when $\alpha_\varepsilon=\beta_\varepsilon$ (Regime 0.3), there is a balance between the fluctuations involving $\beta_\varepsilon^{-1}c(X_t^\varepsilon,Y_t^{0,\varepsilon})$ and $F(t,X_t^\varepsilon,Y_t^{0,\varepsilon})$, and thus the effects of both terms can be observed in the limit equation.

2.3. **Functional central limit theorem: homogenization case.** In this subsection, we consider SDE (1.6) with $H \neq 0$, i.e., there exists a fast varying term even in the slow component. According to Theorem 2.1, the averaged equation for Y_t^ε can be divided into two cases: Regime 1 and Regime 2. We proceed to identify the asymptotic limit for the normalized difference in each regimes: for $k = 1, 2$,

$$Z_t^{k,\varepsilon} := \frac{Y_t^\varepsilon - \bar{Y}_t^k}{\eta_\varepsilon}$$

with suitable deviation scale η_ε such that $\eta_\varepsilon \to 0$ as $\varepsilon \to 0$.

Let us first consider Regime 1 in (1.9). Note that in this case, SDE (1.6) has the same averaged equation as system (2.8). According to Theorem 2.1 (i) with $\vartheta = 1$, the process Y_t^ε will converge to $\bar{Y}_t^1$ strongly with the best possible rate $\alpha_\varepsilon/\gamma_\varepsilon + \alpha_\varepsilon^2/(\beta_\varepsilon\gamma_\varepsilon)$. Thus by the same formal discussions as before, we expect that the natural choice of the deviation scale η_ε in order to observe non-trivial homogenization behavior for $Z_t^{1,\varepsilon}$ should be divided into the following three regimes:

$$\begin{cases} \eta_\varepsilon = \dfrac{\alpha_\varepsilon^2}{\beta_\varepsilon\gamma_\varepsilon} \quad \text{and} \quad \displaystyle\lim_{\varepsilon\to 0}\frac{\beta_\varepsilon}{\alpha_\varepsilon} = 0, & \text{Regime 1.1;} \\ \eta_\varepsilon = \dfrac{\alpha_\varepsilon}{\gamma_\varepsilon} \quad \text{and} \quad \displaystyle\lim_{\varepsilon\to 0}\frac{\alpha_\varepsilon}{\beta_\varepsilon} = 0, & \text{Regime 1.2;} \\ \eta_\varepsilon = \dfrac{\alpha_\varepsilon}{\gamma_\varepsilon} \quad \text{and} \quad \alpha_\varepsilon = \beta_\varepsilon, & \text{Regime 1.3.} \end{cases} \tag{2.14}$$

Recall that Φ is the unique solution to the Poisson equation (2.3), and define

$$\overline{c \cdot \nabla_x \Phi}(t, y) := \int_{\mathbb{R}^{d_1}} c(x, y) \cdot \nabla_x \Phi(t, x, y)\mu^y(\mathrm{d}x); \tag{2.15}$$

$$\overline{H \cdot \Phi^*}(t, y) := \int_{\mathbb{R}^{d_1}} H(t, x, y) \cdot \Phi^*(t, x, y)\mu^y(\mathrm{d}x). \tag{2.16}$$

Then the limit processes $\bar{Z}_{\ell,t}^1$ ($\ell = 1, 2, 3$) for $Z_t^{1,\varepsilon}$ corresponding to Regime 1.1-Regime 1.3 in (2.14) shall be given by:

$$\begin{aligned} \mathrm{d}\bar{Z}_{1,t}^1 &= \nabla_y \bar{F}_1(t, \bar{Y}_t^1)\bar{Z}_{1,t}^1 \mathrm{d}t + \nabla_y G(t, \bar{Y}_t^1)\bar{Z}_{1,t}^1 \mathrm{d}W_t^2 + \overline{c \cdot \nabla_x \Phi}(t, \bar{Y}_t^1)\mathrm{d}t; \\ \mathrm{d}\bar{Z}_{2,t}^1 &= \nabla_y \bar{F}_1(t, \bar{Y}_t^1)\bar{Z}_{2,t}^1 \mathrm{d}t + \nabla_y G(t, \bar{Y}_t^1)\bar{Z}_{2,t}^1 \mathrm{d}W_t^2 + \sqrt{\overline{H \cdot \Phi^*}(t, \bar{Y}_t^1)}\mathrm{d}\tilde{W}_t; \\ \mathrm{d}\bar{Z}_{3,t}^1 &= \nabla_y \bar{F}_1(t, \bar{Y}_t^1)\bar{Z}_{3,t}^1 \mathrm{d}t + \nabla_y G(t, \bar{Y}_t^1)\bar{Z}_{3,t}^1 \mathrm{d}W_t^2 \\ &\quad + \overline{c \cdot \nabla_x \Phi}(t, \bar{Y}_t^1)\mathrm{d}t + \sqrt{\overline{H \cdot \Phi^*}(t, \bar{Y}_t^1)}\mathrm{d}\tilde{W}_t, \end{aligned} \tag{2.17}$$

with initial data $\bar{Z}_{\ell,0}^1 = 0$ ($\ell = 1, 2, 3$), where $\tilde{W}_t$ is another Brownian motion independent of W_t^2.

The following is our second main result for the functional central limit theorems.

Theorem 2.5 (Central limit theorem: Regime 1)**.** *Let* $(\mathbf{A}_\sigma)$, $(\mathbf{A}_b)$, $(\mathbf{A}_G)$, $(\mathbf{A}_H)$ *and* (2.1) *hold,* $0 < \delta, \vartheta \leqslant 1$. *Then for any* $T > 0$ *and every* $\varphi \in C_b^4(\mathbb{R}^{d_2})$, *(i) (Regime 1.1) if $b,\sigma\in C_b^{\delta,1+\vartheta}$, $F,H\in C_p^{(1+\vartheta)/2,\delta,1+\vartheta}$, $G\in C_b^{1/2,1+\vartheta}$ and $c\in C_p^{\delta,\vartheta}$, we have*

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[\varphi(Z_t^{1,\varepsilon})]-\mathbb{E}[\varphi(\bar{Z}_{1,t}^1)]\Big|\leqslant C_T\Big(\gamma_\varepsilon+\frac{\beta_\varepsilon^2}{\alpha_\varepsilon^2}+\frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta\gamma_\varepsilon^\vartheta}\Big);$$

(ii) (Regime 1.2) if $b,\sigma\in C_b^{\delta,1+\vartheta}$, $F,H\in C_p^{(1+\vartheta)/2,\delta,1+\vartheta}$, $G\in C_b^{1/2,1+\vartheta}$ and $c\in L_p^\infty$, we have

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[\varphi(Z_t^{1,\varepsilon})]-\mathbb{E}[\varphi(\bar{Z}_{2,t}^1)]\Big|\leqslant C_T\Big(\gamma_\varepsilon+\frac{\alpha_\varepsilon}{\beta_\varepsilon}+\frac{\alpha_\varepsilon^{\vartheta}}{\gamma_\varepsilon^\vartheta}\Big);$$

(iii) (Regime 1.3) if $b,\sigma\in C_b^{\delta,1+\vartheta}$, $F,H\in C_p^{(1+\vartheta)/2,\delta,1+\vartheta}$, $G\in C_b^{1/2,1+\vartheta}$ and $c\in C_p^{\delta,\vartheta}$, we have

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[\varphi(Z_t^{1,\varepsilon})]-\mathbb{E}[\varphi(\bar{Z}_{3,t}^1)]\Big|\leqslant C_T\Big(\gamma_\varepsilon+\frac{\alpha_\varepsilon^{\vartheta}}{\gamma_\varepsilon^\vartheta}\Big),$$

where for $\ell=1,2,3$, $\bar{Z}_{\ell,t}^1$ satisfy the linear equation (2.17), and $C_T>0$ is a constant independent of δ,ε.

Remark 2.6. *(i) Compared with SDE (2.13) and Theorem 2.3, the homogenization effect of the drift F never appears in SDE (2.17). In fact, all the corresponding terms involving F (through the Poisson equation (2.11)) in SDE (2.13) are now replaced by the drift H (through the Poisson equation (2.3)). This is intuitively natural since in this case the fluctuations will be dominated by the fast component $\gamma_\varepsilon^{-1}H(t,X_t^\varepsilon,Y_t^\varepsilon)$, and the term $F(t,X_t^\varepsilon,Y_t^\varepsilon)$ is only of lower order now, thus its effect in the fluctuations will go to zero eventually in each regime.*

(ii) In particular, if $G\equiv\mathbb{I}_{d_2}$ and $F\equiv0$ then Theorem 2.1 asserts that Y_t^ε converges strongly to $\bar{Y}_t^1=y+W_t^2$. Thus the deviation process is given by

$$Z_t^{1,\varepsilon}=z+\frac{1}{\eta_\varepsilon\gamma_\varepsilon}\int_0^t H(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s,$$

which is an inhomogeneous integral functional of the Markov process $(X_t^\varepsilon,Y_t^\varepsilon)$. Theorem 2.5 provides the limit for $Z_t^{1,\varepsilon}$ in each regime, which is of independent interest (see e.g. [9, 25]*).*

(iii) Let us give more intuitive explanations for Regime 1.1 and Regime 1.2. Under Regime 1.1, we have $\alpha_\varepsilon/\gamma_\varepsilon$ goes to 0 faster than $\alpha_\varepsilon^2/(\beta_\varepsilon\gamma_\varepsilon)$, and the process Y_t^ε will converge to $\bar{Y}_t^1$ with rate $\alpha_\varepsilon^2/(\beta_\varepsilon\gamma_\varepsilon)$. Thus the fast term $\beta_\varepsilon^{-1}c(X_t^\varepsilon,Y_t^\varepsilon)$ is the leading term in the convergence and its averaging effect appears in the limit equation of $\bar{Z}_{1,t}^1$. While in Regime 1.2, we have $\alpha_\varepsilon^2/(\beta_\varepsilon\gamma_\varepsilon)$ goes to 0 faster than $\alpha_\varepsilon/\gamma_\varepsilon$, and the process Y_t^ε converges to $\bar{Y}_t^1$ with order $\alpha_\varepsilon/\gamma_\varepsilon$ (independent of β_ε). This implies that the term $\beta_\varepsilon^{-1}c(X_t^\varepsilon,Y_t^\varepsilon)$ is of lower order now, whereas $\gamma_\varepsilon^{-1}H(t,X_t^\varepsilon,Y_t^\varepsilon)$ is the leading term and its homogenization effect appears in the limit equation of $\bar{Z}_{2,t}^1$.

Now we consider Regime 2 in (1.9), where homogenization already occurs even in the functional law of large numbers. Recall that $\alpha_\varepsilon^2=\beta_\varepsilon\gamma_\varepsilon$ in this case. In particular, we shall always have $\lim_{\varepsilon\to0}\alpha_\varepsilon/\beta_\varepsilon=\infty$. According to Theorem 2.1 (ii) with $\vartheta=1$, the process

Y_t^ε will converge to $\bar{Y}_t^2$ strongly with the best possible rate $\alpha_\varepsilon/\gamma_\varepsilon + \alpha_\varepsilon^2/\beta_\varepsilon$. Thus the natural choice of the derivation scale η_ε in order to observe non-trivial homogenization behavior for $Z_t^{2,\varepsilon}$ should be divided into the following three regimes:

$$\begin{cases} \eta_\varepsilon = \dfrac{\alpha_\varepsilon^2}{\beta_\varepsilon} \quad \text{and} \quad \lim\limits_{\varepsilon\to 0} \dfrac{\beta_\varepsilon}{\alpha_\varepsilon\gamma_\varepsilon} = 0, & \text{Regime 2.1;} \\ \eta_\varepsilon = \dfrac{\alpha_\varepsilon}{\gamma_\varepsilon} \quad \text{and} \quad \lim\limits_{\varepsilon\to 0} \dfrac{\alpha_\varepsilon\gamma_\varepsilon}{\beta_\varepsilon} = 0, & \text{Regime 2.2;} \\ \eta_\varepsilon = \dfrac{\alpha_\varepsilon}{\gamma_\varepsilon} = \dfrac{\alpha_\varepsilon^2}{\beta_\varepsilon}, & \text{Regime 2.3.} \end{cases} \tag{2.18}$$

Recall that Φ is the unique solution to the Poisson equation (2.3), and $\overline{c\cdot\nabla_x\Phi}$ is defined by (2.15). Let Ψ solves the following Poisson equation:

$$\mathscr{L}_0(x,y)\Psi(t,x,y) = -\big[c(x,y)\cdot\nabla_x\Phi(t,x,y) - \overline{c\cdot\nabla_x\Phi}(t,y)\big], \tag{2.19}$$

where $\mathscr{L}_0(x,y)$ is defined by (1.7), and $(t,y)\in\mathbb{R}_+\times\mathbb{R}^{d_2}$ are regarded as parameters. Define

$$\overline{c\cdot\nabla_x\Psi}(t,y) := \int_{\mathbb{R}^{d_1}} c(x,y)\cdot\nabla_x\Psi(t,x,y)\mu^y(\mathrm{d}x).$$

Then the limit processes $\bar{Z}_{\ell,t}^2$ ($\ell = 1,2,3$) for $Z_t^{2,\varepsilon}$ corresponding to Regime 2.1-Regime 2.3 in (2.18) shall be given by:

$$\begin{aligned} \mathrm{d}\bar{Z}_{1,t}^2 &= \nabla_y\bar{F}_2(t,\bar{Y}_t^2)\bar{Z}_{1,t}^2\mathrm{d}t + \nabla_y G(t,\bar{Y}_t^2)\bar{Z}_{1,t}^2\mathrm{d}W_t^2 \\ &\quad + \big(\overline{c\cdot\nabla_x\Upsilon} + \overline{c\cdot\nabla_x\Psi}\big)(t,\bar{Y}_t^2)\mathrm{d}t; \\ \mathrm{d}\bar{Z}_{2,t}^2 &= \nabla_y\bar{F}_2(t,\bar{Y}_t^2)\bar{Z}_{2,t}^2\mathrm{d}t + \nabla_y G(t,\bar{Y}_t^2)\bar{Z}_{2,t}^2\mathrm{d}W_t^2 + \sqrt{\overline{H\cdot\Phi^*}(t,\bar{Y}_t^2)}\mathrm{d}\tilde{W}_t; \\ \mathrm{d}\bar{Z}_{3,t}^2 &= \nabla_y\bar{F}_2(t,\bar{Y}_t^2)\bar{Z}_{3,t}^2\mathrm{d}t + \nabla_y G(t,\bar{Y}_t^2)\bar{Z}_{3,t}^2\mathrm{d}W_t^2 \\ &\quad + \big(\overline{c\cdot\nabla_x\Upsilon} + \overline{c\cdot\nabla_x\Psi}\big)(t,\bar{Y}_t^2)\mathrm{d}t + \sqrt{\overline{H\cdot\Phi^*}(t,\bar{Y}_t^2)}\mathrm{d}\tilde{W}_t \end{aligned} \tag{2.20}$$

with initial data $\bar{Z}_{\ell,0}^2 = 0$, where $\overline{c\cdot\nabla_x\Upsilon}$ and $\overline{H\cdot\Phi^*}$ are defined by (2.12) and (2.16), respectively, $\bar{Y}_t^2$ is the unique strong solution of SDE (2.5) with $k=2$, and $\tilde{W}_t$ is another Brownian motion independent of W_t^2.

Our main result in this case is as follows.

Theorem 2.7 (Central limit theorem: Regime 2)**.** *Let* $(\mathbf{A}_\sigma)$, $(\mathbf{A}_b)$, $(\mathbf{A}_G)$, $(\mathbf{A}_H)$ *and* (2.1) *hold,* $0<\delta,\vartheta\leqslant 1$. *Assume that* $b,\sigma\in C_b^{\delta,1+\vartheta}$, $c\in C_p^{\delta,1+\vartheta}$, $F,H\in C_p^{(1+\vartheta)/2,\delta,1+\vartheta}$ *and* $G\in C_b^{1/2,1+\vartheta}$. *Then for any* $T>0$ *and every* $\varphi\in C_b^4(\mathbb{R}^{d_2})$,

(i) (Regime 2.1) we have

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[\varphi(Z_t^{2,\varepsilon})] - \mathbb{E}[\varphi(\bar{Z}_{1,t}^2)]\Big| \leqslant C_T\Big(\frac{\beta_\varepsilon^2}{\alpha_\varepsilon^2\gamma_\varepsilon^2} + \frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta}\Big);$$

(ii) (Regime 2.2) we have

$$\sup_{t\in[0,T]}\Big|\mathbb{E}[\varphi(Z_t^{2,\varepsilon})] - \mathbb{E}[\varphi(\bar{Z}_{2,t}^2)]\Big| \leqslant C_T\Big(\frac{\alpha_\varepsilon\gamma_\varepsilon}{\beta_\varepsilon} + \frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon^\vartheta}\Big);$$

(iii) (Regime 2.3) we have

$$\sup_{t\in[0,T]}\left|\mathbb{E}[\varphi(Z_t^{2,\varepsilon})]-\mathbb{E}[\varphi(\bar{Z}_{3,t}^2)]\right|\leqslant C_T\frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon^\vartheta},$$

where for $\ell=1,2,3$, $\bar{Z}_{\ell,t}^2$ satisfy the linear equation (2.20), and $C_T>0$ is a constant independent of δ,ε.

Remark 2.8. *(i) Since homogenization already occurs in the averaged equation of $\bar{Y}_t^2$, it is natural to expect that the second order homogenization will appear in the functional central limit theorem. This is exactly characterized through the function Ψ in Regime 2.1 and Regime 2.3 which solves the Poisson equation (2.19). In Regime 2.2, the slow process Y_t^ε will converge to $\bar{Y}_t^2$ with rate $\alpha_\varepsilon/\gamma_\varepsilon$ (independent of β_ε), which means that the term $\gamma_\varepsilon^{-1}H(t,X_t^\varepsilon,Y_t^\varepsilon)$ is the only leading term, and thus the averaging effect involving $\beta_\varepsilon^{-1}c(X_t^\varepsilon,Y_t^\varepsilon)$ does not arise.*

(ii) Note that the same homogenization behaviors involving the drift F as in SDE (2.13) appear again in Regime 2.1 and Regime 2.3. This implies that the component $\gamma_\varepsilon^{-1}H(t,X_t^\varepsilon,Y_t^\varepsilon)$ is of the same order as the drift term $F(t,X_t^\varepsilon,Y_t^\varepsilon)$ in the fluctuations, which should not be a contradiction since in Regime 2, the fast varying of $\gamma_\varepsilon^{-1}H(t,X_t^\varepsilon,Y_t^\varepsilon)$ has already been homogenized out in the averaged equation.

(iii) It is interesting to note that in order to observe all non-trivial behaviors of every term simultaneously, we need to take $\gamma_\varepsilon=\alpha_\varepsilon^{1/2}$ and $\beta_\varepsilon=\alpha_\varepsilon^{3/2}$ (Regime 2.3) to balance the averaging effect of $\beta_\varepsilon^{-1}c(X_t^\varepsilon,Y_t^\varepsilon)$ and the homogenization effect of $\gamma_\varepsilon^{-1}H(t,X_t^\varepsilon,Y_t^\varepsilon)$.

Finally, we mention that regime specific analysis for system (1.6) is also done in [55, 57]. We list the following comparisons with our results.

Remark 2.9. *(i) In [57], the convergence in probability and the central limit theorem for system (1.6) was studied in a homogeneous case where $\alpha_\varepsilon=\delta/\sqrt{\varepsilon}$, $\beta_\varepsilon=\delta$, $\gamma_\varepsilon=\delta/\varepsilon$ and with small noise perturbations, i.e., with G replaced by $\sqrt{\varepsilon}G$. Thus, the corresponding results in [57] can be seen as a particular case of Regime 2 and Regimes 2.1-2.3 of this paper. Moreover, we handle non-smooth coefficients, which is much more general than the case studied in [57], and we obtain optimal rates of convergence. Note that all the above rates of convergence do not depend on the regularity of the coefficients with respect to the fast variable. This reflects that the slow process is the main term in the limiting procedure of the multi-scale system, which coincides with the intuition since the fast component has been totally averaged or homogenized out in the limit equation.*

The author in [57] also considered the cases where $\gamma_\varepsilon\to\gamma\in(0,\infty)$ as $\varepsilon\to0$. In this case, the part $\gamma_\varepsilon^{-1}H(t,X_t^\varepsilon,Y_t^\varepsilon)$ is no longer a fast term and can be handled by using the same arguments as for the drift part F, which is sightly easier.

(ii) In [55], only convergence in distribution for system (1.6) was studied. In Theorem 2.1, we provide the strong convergence in the pathwise sense. Due to the low regularity conditions on the coefficients, Zvonkin's transformation (see Lemma 4.4 below) is needed to handle the non-smooth coefficients. Furthermore, the central limit results in this paper provide more delicate characterisation for the interactions between the averaging term $\beta_\varepsilon^{-1}c(X_t^\varepsilon,Y_t^\varepsilon)$ and the homogenization term $\gamma_\varepsilon^{-1}H(t,X_t^\varepsilon,Y_t^\varepsilon)$ in system (1.6).

Notations: Since we shall prove the main results in a quite unified way, we introduce some notations here for brevity. Let $\bar{\mathscr{L}}_k$ be the infinitesimal generator for $\bar{Y}_t^k$, i.e., for $k = 1, 2$,

$$\bar{\mathscr{L}}_k := \bar{\mathscr{L}}_k(t,y) := \sum_{i,j=1}^{d_2} \mathcal{G}^{ij}(t,y)\frac{\partial^2}{\partial y_i \partial y_j} + \sum_{i=1}^{d_2} \bar{F}_k^i(t,y)\frac{\partial}{\partial y_i}, \tag{2.21}$$

where $\bar{F}_k$ are defined by (2.4). Note that the averaged process for $Y_t^{0,\varepsilon}$ in SDE (2.8) is also given by $\bar{Y}_t^1$, we let $\bar{Y}_t^0 := \bar{Y}_t^1$ and $\bar{\mathscr{L}}_0 := \bar{\mathscr{L}}_1$ for consistency. We also introduce

$$\bar{\mathscr{L}}_3^0 := \bar{\mathscr{L}}_3^0(t,y,z) := \sum_{i=1}^{d_2} \left(\overline{c \cdot \nabla_x \Upsilon}\right)^i(t,y)\frac{\partial}{\partial z_i}, \tag{2.22}$$

$$\bar{\mathscr{L}}_3^1 := \bar{\mathscr{L}}_3^1(t,y,z) := \sum_{i=1}^{d_2} \left(\overline{c \cdot \nabla_x \Phi}\right)^i(t,y)\frac{\partial}{\partial z_i}, \tag{2.23}$$

$$\bar{\mathscr{L}}_3^2 := \bar{\mathscr{L}}_3^2(t,y,z) := \sum_{i=1}^{d_2} \left(\overline{c \cdot \nabla_x \Psi}\right)^i(t,y)\frac{\partial}{\partial z_i}, \tag{2.24}$$

$$\bar{\mathscr{L}}_4^1 := \bar{\mathscr{L}}_4^1(t,y,z) := \sum_{i,j=1}^{d_2} \left(\overline{\tilde{F} \cdot \Upsilon^*}\right)^{ij}(t,y)\frac{\partial^2}{\partial z_i \partial z_j}, \tag{2.25}$$

$$\bar{\mathscr{L}}_4^2 := \bar{\mathscr{L}}_4^2(t,y,z) := \sum_{i,j=1}^{d_2} \left(\overline{H \cdot \Phi^*}\right)^{ij}(t,y)\frac{\partial^2}{\partial z_i \partial z_j}, \tag{2.26}$$

and for $k = 1, 2$,

$$\begin{aligned} \bar{\mathscr{L}}_5^k := \bar{\mathscr{L}}_5^k(t,y,z) := &\sum_{i=1}^{d_2} \left(\nabla_y \bar{F}_k(t,y)z\right)^i \frac{\partial}{\partial z_i} + \frac{1}{2}\sum_{i,j=1}^{d_2} \left(G(t,y)[\nabla_y G(t,y)z]^*\right)^{ij} \frac{\partial^2}{\partial y_i \partial z_j} \\ &+ \frac{1}{2}\sum_{i,j=1}^{d_2} \left([\nabla_y G(t,y)z][\nabla_y G(t,y)z]^*\right)^{ij} \frac{\partial^2}{\partial z_i \partial z_j}. \end{aligned} \tag{2.27}$$

Then the infinitesimal generator of $(\bar{Y}_t^k, \bar{Z}_{\ell,t}^k)$ can be written as $\bar{\mathscr{L}}_k + \bar{\mathcal{L}}_\ell^k$, where for $k = 0, 1, 2$ and $\ell = 1, 2, 3$, the operator $\bar{\mathcal{L}}_\ell^k$ are defined as follows:
corresponding to $\bar{Z}_{\ell,t}^0$ ($\ell = 1, 2, 3$) in SDE (2.13),

$$\bar{\mathcal{L}}_1^0 := \bar{\mathscr{L}}_5^1 + \bar{\mathscr{L}}_3^0, \quad \bar{\mathcal{L}}_2^0 := \bar{\mathscr{L}}_5^1 + \bar{\mathscr{L}}_4^1, \quad \bar{\mathcal{L}}_3^0 := \bar{\mathscr{L}}_5^1 + \bar{\mathscr{L}}_3^0 + \bar{\mathscr{L}}_4^1; \tag{2.28}$$

corresponding to $\bar{Z}_{\ell,t}^1$ ($\ell = 1, 2, 3$) in SDE (2.17),

$$\bar{\mathcal{L}}_1^1 := \bar{\mathscr{L}}_5^1 + \bar{\mathscr{L}}_3^1, \quad \bar{\mathcal{L}}_2^1 := \bar{\mathscr{L}}_5^1 + \bar{\mathscr{L}}_4^2, \quad \bar{\mathcal{L}}_3^1 := \bar{\mathscr{L}}_5^1 + \bar{\mathscr{L}}_3^1 + \bar{\mathscr{L}}_4^2; \tag{2.29}$$

and corresponding to $\bar{Z}_{\ell,t}^2$ ($\ell = 1, 2, 3$) in SDE (2.20),

$$\bar{\mathcal{L}}_1^2 := \bar{\mathscr{L}}_5^2 + \bar{\mathscr{L}}_3^0 + \bar{\mathscr{L}}_3^2, \quad \bar{\mathcal{L}}_2^2 := \bar{\mathscr{L}}_5^2 + \bar{\mathscr{L}}_4^2, \quad \bar{\mathcal{L}}_3^2 := \bar{\mathscr{L}}_5^2 + \bar{\mathscr{L}}_3^0 + \bar{\mathscr{L}}_3^2 + \bar{\mathscr{L}}_4^2. \tag{2.30}$$

3. Poisson equation and Cauchy problem

This section collects the main tools that we shall use to prove our results. As mentioned in the introduction, the Poisson equation will play an important role in the proof both for the strong convergence in the averaging principle and the central limit theorems. Let us first recall some results in this direction.

Consider the following Poisson equation in $\mathbb{R}^{d_1}$:

$$\mathscr{L}_0(x,y)u(x,y) = -f(x,y), \tag{3.1}$$

where $\mathscr{L}_0(x,y)$ is defined by (1.7), and $y \in \mathbb{R}^{d_2}$ is regarded as a parameter. Note that $\mathscr{L}_0(x,y)$ is the infinitesimal generator of X_t^y given by (1.2). To ensure the well-posedness of equation (3.1), we need to make the following centering condition on f:

$$\int_{\mathbb{R}^{d_1}} f(x,y)\mu^y(\mathrm{d}x) = 0, \quad \forall y \in \mathbb{R}^{d_2}, \tag{3.2}$$

where $\mu^y(\mathrm{d}x)$ is the invariant measure for X_t^y. The following result was proved in [55, Theorem 2.1], which will be used frequently below.

Theorem 3.1. *Let* $(\mathbf{A}_\sigma)$ *and* $(\mathbf{A}_b)$ *hold. Assume that* $a, b \in C_b^{\delta,\vartheta}$ *with* $0 < \delta \leqslant 1$ *and* $\vartheta \geqslant 0$. *Then for every function* $f \in C_p^{\delta,\vartheta}$ *satisfying (3.2), there exists a unique solution* $u \in C_p^{2+\delta,\vartheta}$ *to equation (3.1) satisfying (3.2) which is given by*

$$u(x,y) = \int_0^\infty \mathbb{E}f(X_t^y(x),y)\mathrm{d}t.$$

Moreover, there exists a constant $m > 0$ *such that:*
(i) for any $x \in \mathbb{R}^{d_1}$ *and* $y \in \mathbb{R}^{d_2}$,

$$|u(x,y)| + |\nabla_x u(x,y)| + |\nabla_x^2 u(x,y)| \leqslant C_0(1+|x|^m), \tag{3.3}$$

where $C_0 > 0$ *depends only on* d_1, d_2 *and* $\|a\|_{C_b^{\delta,0}}, \|b\|_{C_b^{\delta,0}}, [f]_{C_p^{\delta,0}}$;
(ii) for any $x \in \mathbb{R}^{d_1}$,

$$\|u(x,\cdot)\|_{C_b^\vartheta} \leqslant C_1(1+|x|^m), \tag{3.4}$$

where $C_1 > 0$ *depends on* d_1, d_2 *and* $\|a\|_{C_b^{\delta,\vartheta}}, \|b\|_{C_b^{\delta,\vartheta}}, [f]_{C_p^{\delta,\vartheta}}$.

We will need to use Itô's formula for the solution of the Poisson equation with the fast and slow components in SDE (1.6) plugged in for both variables, say $u(X_t^\varepsilon, Y_t^\varepsilon)$, which in turn requires at least two derivatives of u with respect to the x variable as well as the y variable. In view of estimate (3.3), the derivatives with respect to x are not a problem since we can get them for free by virtue of the uniform ellipticity property of the operator. However, due to our low regularities of the coefficients with respect to the y variable (only Hölder continuous) and taking into account (3.4), we cannot get the desired two derivatives for $u(x,\cdot)$ directly. To overcome this problem, we use some mollification arguments.

Let $\rho_1 : \mathbb{R} \to [0,1]$ and $\rho_2 : \mathbb{R}^{d_2} \to [0,1]$ be two smooth radial convolution kernel functions such that $\int_{\mathbb{R}} \rho_1(r)\mathrm{d}r = \int_{\mathbb{R}^{d_2}} \rho_2(y)\mathrm{d}y = 1$, and for any $k \geqslant 1$, there exist constants $C_k > 0$ such that $|\nabla^k \rho_1(r)| \leqslant C_k \rho_1(r)$ and $|\nabla^k \rho_2(y)| \leqslant C_k \rho_2(y)$. For every $n \in \mathbb{N}^*$, set

$$\rho_1^n(r) := n^2 \rho_1(n^2 r) \quad \text{and} \quad \rho_2^n(y) := n^{d_2} \rho_2(ny).$$

Given a function $f(t,x,y,z)$, we define the mollifying approximations of f in t and y variables by

$$f_n(t,x,y,z) := f * \rho_2^n * \rho_1^n := \int_{\mathbb{R}^{d_2+1}} f(t-s,x,y-y',z)\rho_2^n(y')\rho_1^n(s)\mathrm{d}y'\mathrm{d}s. \tag{3.5}$$

The following easy result can be proved similarly as in [55, Lemma 4.1], we omit the details.

Lemma 3.2. *Let $f \in C_p^{\vartheta/2,0,\vartheta,0}$ with $0 < \vartheta \leqslant 2$ and define f_n by (3.5). Then we have*

$$\|f(\cdot,x,\cdot,\cdot) - f_n(\cdot,x,\cdot,\cdot)\|_\infty \leqslant C_0 n^{-\vartheta}(1+|x|^m), \tag{3.6}$$

$$\|\nabla_y f_n(\cdot,x,\cdot,\cdot)\|_\infty \leqslant C_0 n^{1-(\vartheta\wedge 1)}(1+|x|^m), \tag{3.7}$$

and

$$\|\partial_t f_n(\cdot,x,\cdot,\cdot)\|_\infty + \|\nabla_y^2 f_n(\cdot,x,\cdot,\cdot)\|_\infty \leqslant C_0 n^{2-\vartheta}(1+|x|^m), \tag{3.8}$$

where $C_0 > 0$ is a constant independent of n.

Given a function $h(t,x,y)$, we denote its average with respect to the measure $\mu^y(\mathrm{d}x)$ by $\bar{h}(t,y)$, i.e.,

$$\bar{h}(t,y) := \int_{\mathbb{R}^{d_1}} h(t,x,y)\mu^y(\mathrm{d}x). \tag{3.9}$$

The following result specifies the regularity of an averaged function, which explains the assumptions we made on the coefficients in our main results.

Lemma 3.3. *Let $(\mathbf{A}_\sigma)$ and $(\mathbf{A}_b)$ hold. Assume that $a, b \in C_b^{\delta,\vartheta}$ with $0 < \delta \leqslant 1$ and $\vartheta \geqslant 0$. Then for every $h \in C_p^{\vartheta/2,\delta,\vartheta}$, we have $\bar{h} \in C_b^{\vartheta/2,\vartheta}$. In particular,*

(i) under conditions in Theorem 2.1, we have for every $k = 1,2$, $\bar{F}_k \in C_b^{\vartheta/2,\vartheta}$;

(ii) under conditions in Theorem 2.3, Theorem 2.5 and Theorem 2.7, we have for every $k = 1,2$,

$$\nabla_y \bar{F}_k, \overline{c\cdot\nabla_x\Upsilon}, \overline{\tilde{F}\cdot\Upsilon^*}, \overline{c\cdot\nabla_x\Phi}, \overline{H\cdot\Phi^*}, \overline{c\cdot\nabla_x\Psi} \in C_b^{\vartheta/2,\vartheta}.$$

Proof. The assertion that $\bar{h} \in C_b^{\vartheta/2,\vartheta}$ was proved in [55, Lemma 3.2]. Then, under the assumptions in Theorem 2.1 (Regime 1), the conclusion that $\bar{F}_1 \in C_b^{\vartheta/2,\vartheta}$ follows directly. Recall that Φ solves (2.3). By the assumptions in Theorem 2.1 (Regime 2) and Theorem 3.1, we have $\Phi \in C_p^{\vartheta/2,2+\delta,\vartheta}$. This together with the condition that $c \in C_p^{\delta,\vartheta}$ implies that $c\cdot\nabla_x\Phi \in C_p^{\vartheta/2,2+\delta,\vartheta}$, which in turn yields $\bar{F}_2 \in C_b^{\vartheta/2,\vartheta}$. (ii) can be proved by the same arguments, so we omit the details. ☐

Another main tool we will use to prove the functional central limit theorems is the Cauchy problem corresponding to the limit dynamics $(\bar{Y}_t^k, \bar{Z}_{\ell,t}^k)$, $k = 0,1,2$ and $\ell = 1,2,3$. Note that the processes $\bar{Y}_t^k$ depend on the initial value y, while $\bar{Z}_{\ell,t}^k$ depend on y but with initial value 0. Below, we shall write $\bar{Y}_t^k(y)$ when we want to stress the dependence on the initial value, and use $\bar{Z}^k_{\ell,t}(y,z)$ to denote processes $\bar{Z}^k_{\ell,t}$ with initial point $z\in\mathbb{R}^{d_2}$. Fix a $T>0$ below, consider the following Cauchy problem on $[0,T]\times\mathbb{R}^{d_2}\times\mathbb{R}^{d_2}$:

$$\begin{cases} \partial_t u^k_\ell(t,y,z)-(\bar{\mathscr{L}}_k+\bar{\mathcal{L}}^k_\ell)u^k_\ell(t,y,z)=0, \quad t\in(0,T], \\ u^k_\ell(0,y,z)=\varphi(z), \end{cases} \tag{3.10}$$

where $\bar{\mathscr{L}}_k$ and $\bar{\mathcal{L}}^k_\ell$ are defined by (2.21), (2.28), (2.29) and (2.30), respectively. We have the following result.

Theorem 3.4. *For every $k=0,1,2$, $\ell=1,2,3$ and $\varphi\in C^4_b(\mathbb{R}^{d_2})$, there exists a unique solution $u^k_\ell\in C_b^{(2+\vartheta)/2,2+\vartheta,4}$ to equation (3.10) which is given by*

$$u^k_\ell(t,y,z)=\mathbb{E}\varphi\big(\bar{Z}^k_{\ell,t}(y,z)\big). \tag{3.11}$$

Proof. We only prove the assertion for $k=0$ and $\ell=1$. Although this is not the most general one, we choose this case since it carries the key difficulties. For simplicity, we shall write u instead of u^0_1 for the solution. Without loss of generality, we may assume that the coefficients are smooth, and focus on proving the a priori estimates for u. Since $\bar{\mathscr{L}}_0+\bar{\mathcal{L}}^0_1$ is the generator of the Markov process $(\bar{Y}^0_t,\bar{Z}^0_{1,t})$, it is well known that the solution for (3.10) will be given by (3.11). By the assumption $\varphi\in C^4_b(\mathbb{R}^{d_2})$, the fact that $\bar{Y}^0_t$ does not depend on z, and since $\bar{Z}^0_{1,t}(y,z)$ satisfies the linear equation (2.13), it is easily checked that for every $(t,y)\in\mathbb{R}_+\times\mathbb{R}^{d_2}$, we have $u(t,y,\cdot)\in C^4_b(\mathbb{R}^{d_2})$, and for $i=1,\cdots,4$,

$$\|\nabla^i_z u(t,y,\cdot)\|_\infty\leqslant C_0\|\varphi\|_{C^4_b},$$

where $C_0>0$ depends only on $\|\nabla_y\bar{F}_1\|_\infty$ and $\|\nabla_y G\|_\infty$. It remains to prove that for every $z\in\mathbb{R}^{d_2}$, $u(\cdot,\cdot,z)\in C_b^{(2+\vartheta)/2,2+\vartheta}$. To this end, we rewrite equation (3.10) as

$$\partial_t u(t,y,z)-\bar{\mathscr{L}}_0 u(t,y,z)=\bar{\mathcal{L}}^0_1 u(t,y,z).$$

By regarding z as a parameter in the above equation, and recalling that $\mathcal{G}$ is uniformly elliptic, it suffices to show that

$$\nabla_z u(\cdot,\cdot,z),\nabla^2_z u(\cdot,\cdot,z)\in C_b^{\vartheta/2,\vartheta}.$$

Then the conclusion follows by classical PDE's result, see e.g. [43, Chapter IV, Section 5]. For any $y_1,y_2\in\mathbb{R}^{d_2}$, we write

$$\begin{aligned}\big|\nabla_z u(t,y_1,z)-\nabla_z u(t,y_2,z)\big| &\leqslant \|\nabla_z\varphi\|_\infty\mathbb{E}\big|\nabla_z\bar{Z}^0_{1,t}(y_1,z)-\nabla_z\bar{Z}^0_{1,t}(y_2,z)\big| \\ &\quad+\mathbb{E}\Big(\big|\nabla_z\varphi\big(\bar{Z}^0_{1,t}(y_1,z)\big)-\nabla_z\varphi\big(\bar{Z}^0_{1,t}(y_2,z)\big)\big|\cdot|\nabla_z\bar{Z}^0_{1,t}(y_2,z)|\Big) \\ &\leqslant C_1\mathbb{E}\Big(\big|\nabla_z\bar{Z}^0_{1,t}(y_1,z)-\nabla_z\bar{Z}^0_{1,t}(y_2,z)\big|^2+\big|\bar{Z}^0_{1,t}(y_1,z)-\bar{Z}^0_{1,t}(y_2,z)\big|^2\Big)^{1/2}.\end{aligned}$$

Note that

$$\mathrm{d}\nabla_z\bar{Z}^0_{1,t}=\nabla_y\bar{F}_1(t,\bar{Y}^1_t)\nabla_z\bar{Z}^0_{1,t}\mathrm{d}t+\nabla_y G(t,\bar{Y}^1_t)\nabla_z\bar{Z}^0_{1,t}\mathrm{d}W^2_t.$$

Thus by the fact that $\nabla_y\bar{F}_k(t,\cdot),\nabla_y G(t,\cdot)\in C^\vartheta_b$, we deduce that

$$\mathbb{E}\big|\nabla_z\bar{Z}^0_{1,t}(y_1,z)-\nabla_z\bar{Z}^0_{1,t}(y_2,z)\big|^2\leqslant C_2\mathbb{E}\left(\int_0^t\big|\nabla_z\bar{Z}^0_{1,s}(y_1,z)-\nabla_z\bar{Z}^0_{1,s}(y_2,z)\big|^2\mathrm{d}s\right)$$

$$
\begin{aligned}
&+ C_2\mathbb{E}\bigg(\int_0^t \big|\nabla_y \bar{F}_1(s,\bar{Y}_s^1(y_1)) - \nabla_y \bar{F}_1(s,\bar{Y}_s^1(y_2))\big|^2 \\
&\qquad\qquad + \big|\nabla_y \bar{G}(s,\bar{Y}_s^1(y_1)) - \nabla_y \bar{G}(s,\bar{Y}_s^1(y_2))\big|^2 \mathrm{d}s\bigg) \\
&\leqslant C_2\mathbb{E}\left(\int_0^t \big|\nabla_z \bar{Z}_{1,s}^0(y_1,z) - \nabla_z \bar{Z}_{1,s}^0(y_2,z)\big|^2 \mathrm{d}s\right) \\
&\quad + C_2\mathbb{E}\bigg(\int_0^t \big|\bar{Y}_s^1(y_1) - \bar{Y}_s^1(y_2)\big|^{2\vartheta} \mathrm{d}s\bigg).
\end{aligned} \tag{3.12}
$$

It is well known that $y \mapsto \bar{Y}_t^1(y)$ is a C^1-diffeomorphism, i.e., for every $t \in [0,T]$,

$$
\mathbb{E}\big|\bar{Y}_t^1(y_1) - \bar{Y}_t^1(y_2)\big| \leqslant C_T|y_1 - y_2|.
$$

Taking this back into (3.12) and by Gronwall's inequality, we get

$$
\mathbb{E}\big|\nabla_z \bar{Z}_{1,t}^0(y_1,z) - \nabla_z \bar{Z}_{1,t}^0(y_2,z)\big|^2 \leqslant C_3|y_1 - y_2|^{2\vartheta}.
$$

By the same arguments and more easily, we also have

$$
\mathbb{E}\big|\bar{Z}_{1,t}^0(y_1,z) - \bar{Z}_{1,t}^0(y_2,z)\big|^2 \leqslant C_4|y_1 - y_2|^{2\vartheta},
$$

which in turn implies that

$$
\big|\nabla_z u(t,y_1,z) - \nabla_z u(t,y_2,z)\big| \leqslant C_5|y_1 - y_2|^{\vartheta}.
$$

The corresponding regularity for $\nabla_z u$ with respect to t variable and for $\nabla_z^2 u$ can be proved similarly. □

4. Strong convergence in the averaging principle

Using the technique of Poisson equation, we shall first derive some fluctuation estimates in Subsection 4.1. Then we prove the strong convergence in the averaging principle of SDE (1.6) in Subsection 4.2 by Zvonkin's transformation.

4.1. **Strong fluctuation estimates.**

Given a function $h(t,x,y)$, recall that $\bar{h}(t,y)$ is defined by (3.9). It is easy to see that $f(t,x,y) := h(t,x,y) - \bar{h}(t,y)$ satisfies the centering condition, i.e.,

$$
\int_{\mathbb{R}^{d_1}} f(t,x,y)\mu^y(\mathrm{d}x) = 0, \quad \forall (t,y) \in \mathbb{R}_+ \times \mathbb{R}^{d_2}. \tag{4.1}
$$

The following result gives an estimate for the fluctuations between $h(s,X_s^\varepsilon,Y_s^\varepsilon)$ and $\bar{h}(s,Y_s^\varepsilon)$ over the time interval $[0,t]$.

Lemma 4.1. *Let* $(\mathbf{A}_\sigma)$, $(\mathbf{A}_b)$ *and* (2.1) *hold. Assume that* $b,\sigma \in C_b^{\delta,\vartheta}$ *with* $0 < \delta,\vartheta \leqslant 2$ *and* $c,F,H,G \in L_p^\infty$. *Then for every* $f \in C_p^{\vartheta/2,\delta,\vartheta}$ *satisfying* (4.1) *and any* $q \geqslant 2$, *we have*

$$
\mathbb{E}\left|\int_0^t f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s\right|^q \leqslant C_t\Big(\alpha_\varepsilon^{\vartheta\wedge 1} + \frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^q, \tag{4.2}
$$

where $C_t > 0$ *is a constant independent of* δ,ε.

Remark 4.2. *We call (4.2) a strong fluctuation estimate because we take the absolute value for the integral over* $[0,t]$*. Compared with Lemma 5.1 and Lemma 6.1 below, we shall see that the involved martingale part will be one of the leading terms in the control of error bounds in this case, and this is the main reason why the power* $\vartheta\wedge 1$ *appears on the right hand side of (4.2).*

Proof. By the assumptions that $f\in C_p^{\vartheta/2,\delta,\vartheta}$ satisfying (4.1), $b,\sigma\in C_b^{\delta,\vartheta}$ and according to Theorem 3.1, there exists a unique solution $\Phi^f(t,x,y)\in C_p^{\vartheta/2,2+\delta,\vartheta}$ to the following Poisson equation in $\mathbb{R}^{d_1}$:

$$\mathscr{L}_0(x,y)\Phi^f(t,x,y)=-f(t,x,y), \tag{4.3}$$

where $(t,y)\in\mathbb{R}_+\times\mathbb{R}^{d_2}$ are regarded as parameters. Let Φ_n^f be the mollifyer of Φ^f defined as in (3.5) (which does not depend on the z-variable here). Using Itô's formula, we have

$$\begin{aligned}\Phi_n^f(t,X_t^\varepsilon,Y_t^\varepsilon)&=\Phi_n^f(0,x,y)+\int_0^t\Big(\partial_s+\mathscr{L}_1+\gamma_\varepsilon^{-1}\mathscr{L}_2+\beta_\varepsilon^{-1}\mathscr{L}_3\Big)\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s\\&\quad+\frac{1}{\alpha_\varepsilon^2}\int_0^t\mathscr{L}_0\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s+\frac{1}{\alpha_\varepsilon}M_n^1(t)+M_n^2(t),\end{aligned}$$

where $\mathscr{L}_1,\mathscr{L}_2$ and $\mathscr{L}_3$ are given by (1.8), and for $i=1,2$, $M_n^i(t)$ are martingales defined by

$$\begin{aligned}M_n^1(t)&:=\int_0^t\nabla_x\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon)\sigma(X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}W_s^1,\\M_n^2(t)&:=\int_0^t\nabla_y\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon)G(s,Y_s^\varepsilon)\mathrm{d}W_s^2.\end{aligned}$$

By (4.3) this in turn yields that

$$\begin{aligned}\int_0^t f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s&=\alpha_\varepsilon^2\Phi_n^f(0,x,y)-\alpha_\varepsilon^2\Phi_n^f(t,X_t^\varepsilon,Y_t^\varepsilon)+\alpha_\varepsilon M_n^1(t)+\alpha_\varepsilon^2M_n^2(t)\\&\quad+\alpha_\varepsilon^2\int_0^t\big(\partial_s+\mathscr{L}_1\big)\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s\\&\quad+\frac{\alpha_\varepsilon^2}{\gamma_\varepsilon}\int_0^t\mathscr{L}_2\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\int_0^t\mathscr{L}_3\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s\\&\quad+\int_0^t\big(\mathscr{L}_0\Phi_n^f-\mathscr{L}_0\Phi^f\big)(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s.\end{aligned} \tag{4.4}$$

As a result, we have for any $q\geqslant 2$,

$$\begin{aligned}\mathcal{Q}(\varepsilon)&:=\mathbb{E}\left|\int_0^t f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s\right|^q\leqslant C_q\Big[\alpha_\varepsilon^{2q}\Big(|\Phi_n^f(0,x,y)|^q+\mathbb{E}|\Phi_n^f(t,X_t^\varepsilon,Y_t^\varepsilon)|^q\Big)\\&\quad+\alpha_\varepsilon^q\,\mathbb{E}|M_n^1(t)|^q+\alpha_\varepsilon^{2q}\,\mathbb{E}|M_n^2(t)|^q\Big]+C_q\alpha_\varepsilon^{2q}\,\mathbb{E}\left|\int_0^t\big(\partial_s+\mathscr{L}_1\big)\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s\right|^q\\&\quad+C_q\frac{\alpha_\varepsilon^{2q}}{\gamma_\varepsilon^q}\mathbb{E}\left|\int_0^t\mathscr{L}_2\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s\right|^q+C_q\frac{\alpha_\varepsilon^{2q}}{\beta_\varepsilon^q}\mathbb{E}\left|\int_0^t\mathscr{L}_3\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s\right|^q\end{aligned}$$

$$+ C_q\,\mathbb{E}\left|\int_0^t \big(\mathscr{L}_0\Phi_n^f - \mathscr{L}_0\Phi^f\big)(s, X_s^\varepsilon, Y_s^\varepsilon)\mathrm{d}s\right|^q =: \sum_{i=1}^5 \mathcal{Q}_i(\varepsilon).$$

Under $(\mathbf{A}_\sigma)$ and (2.1), it follows from [62, Lemma 1] (see also [51, Lemma 1] or [55]) that for any $m > 0$, there exists a constant $C_0 > 0$ such that

$$\sup_{\varepsilon\in(0,1/2)} \mathbb{E}|X_t^\varepsilon|^m \leqslant C_0(1+|x|^{m+2}). \tag{4.5}$$

Thus by (3.3) we have that there exist constants $m > 0$ and $C_1 > 0$ independent of n such that

$$|\Phi_n^f(0,x,y)|^q + \mathbb{E}|\Phi_n^f(t, X_t^\varepsilon, Y_t^\varepsilon)|^q \leqslant C_1\mathbb{E}\big(1+|X_t^\varepsilon|^{mq}\big) < \infty.$$

At the same time, by Hölder's inequality,

$$\mathbb{E}|M_n^1(t)|^q \leqslant C_1\mathbb{E}\left(\int_0^t \Big(1+|X_s^\varepsilon|^{2m}\Big)\mathrm{d}s\right)^{q/2} \leqslant C_1\mathbb{E}\left(\int_0^t \Big(1+|X_s^\varepsilon|^{mq}\Big)\mathrm{d}s\right) < \infty,$$

and in view of (3.7),

$$\mathbb{E}|M_n^2(t)|^q \leqslant C_1 n^{q(1-(\vartheta\wedge 1))}\mathbb{E}\left(\int_0^t \Big(1+|X_s^\varepsilon|^{mq}\Big)\mathrm{d}s\right) \leqslant C_1 n^{q(1-(\vartheta\wedge 1))}.$$

Consequently, we get

$$\mathcal{Q}_1(\varepsilon) \leqslant C_1\Big(\alpha_\varepsilon^q + \alpha_\varepsilon^{2q} n^{q(1-(\vartheta\wedge 1))}\Big).$$

To control the second term, by (3.8) and the assumptions that $F, G \in L_p^\infty$, we deduce that

$$\begin{aligned}\|\big(\partial_s + \mathscr{L}_1\big)\Phi_n^f(\cdot,x,\cdot)\|_\infty &\leqslant C_2\big(1+|x|^m\big)\\ &\quad\times\Big(\|\partial_s\Phi_n^f(\cdot,x,\cdot)\|_\infty + \sum_{\ell=1,2}\big\|\nabla_y^\ell\Phi_n^f(\cdot,x,\cdot)\big\|_\infty\Big) \leqslant C_2 n^{2-\vartheta}(1+|x|^{2m}).\end{aligned}$$

Taking into account (4.5) yields

$$\mathcal{Q}_2(\varepsilon) \leqslant C_2\alpha_\varepsilon^{2q} n^{q(2-\vartheta)}\mathbb{E}\left(\int_0^t \big(1+|X_s^\varepsilon|^{2mq}\big)\mathrm{d}s\right) \leqslant C_2\alpha_\varepsilon^{2q} n^{q(2-\vartheta)}.$$

Using the same argument as above, one can check that

$$\|\mathscr{L}_2\Phi_n^f(\cdot,x,\cdot)\|_\infty \leqslant C_3(1+|x|^m)\|\nabla_y\Phi_n^f(\cdot,x,\cdot)\|_\infty \leqslant C_3 n^{1-(\vartheta\wedge 1)}(1+|x|^{2m}).$$

Thus we have

$$\mathcal{Q}_3(\varepsilon) \leqslant C_3\frac{\alpha_\varepsilon^{2q}}{\gamma_\varepsilon^q} n^{q(1-(\vartheta\wedge 1))}\mathbb{E}\left(\int_0^t \big(1+|X_s^\varepsilon|^{2mq}\big)\mathrm{d}s\right) \leqslant C_3\frac{\alpha_\varepsilon^{2q}}{\gamma_\varepsilon^q} n^{q(1-(\vartheta\wedge 1))}.$$

Furthermore, by the assumption that $c \in L_p^\infty$, it follows directly that

$$\mathcal{Q}_4(\varepsilon) \leqslant C_4\frac{\alpha_\varepsilon^{2q}}{\beta_\varepsilon^q}\mathbb{E}\left(\int_0^t \big(1+|X_s^\varepsilon|^{2mq}\big)\mathrm{d}s\right) \leqslant C_4\frac{\alpha_\varepsilon^{2q}}{\beta_\varepsilon^q}.$$

Finally, since $\nabla_x^2\Phi^f \in C_p^{\vartheta/2,\delta,\vartheta}$ and due to the fact that

$$\nabla_x^2(\Phi_n^f) = (\nabla_x^2\Phi^f) * \rho_1^n * \rho_2^n,$$

we derive by (3.6) that

$$\begin{aligned}\mathcal{Q}_5(\varepsilon) &\leqslant C_5\,\mathbb{E}\left(\int_0^t \sum_{\ell=1,2}\big\|(\nabla_x^\ell \Phi_n^f-\nabla_x^\ell\Phi^f)(\cdot,X_s^\varepsilon,\cdot)\big\|_\infty^q\mathrm{d}s\right)\\ &\leqslant C_5 n^{-q\vartheta}\mathbb{E}\left(\int_0^t\big(1+|X_s^\varepsilon|^{mq}\big)\mathrm{d}s\right)\leqslant C_5 n^{-q\vartheta}.\end{aligned}$$

Combining the above computations, we arrive at

$$\mathcal{Q}(\varepsilon)\leqslant C_6\Big(\alpha_\varepsilon+\alpha_\varepsilon^2 n^{1-(\vartheta\wedge 1)}+\alpha_\varepsilon^2 n^{2-\vartheta}+\frac{\alpha_\varepsilon^2}{\gamma_\varepsilon}n^{1-(\vartheta\wedge1)}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}+n^{-\vartheta}\Big)^q.$$

Taking $n=\alpha_\varepsilon^{-1}$, we thus get

$$\mathcal{Q}(\varepsilon)\leqslant C_7\Big(\alpha_\varepsilon+\alpha_\varepsilon^\vartheta+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^q\leqslant C_7\Big(\alpha_\varepsilon^{\vartheta\wedge1}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^q,$$

and the proof is finished. ☐

The above result can be regraded as a law of large numbers type fluctuation estimate. Now, we derive a central limit type fluctuation estimate for the integral of $f(s,X_s^\varepsilon,Y_s^\varepsilon)$ over the time interval $[0,t]$. This will depend on the two regimes described in (1.9). Recall that Φ^f is the solution to the Poisson equation (4.3). For simplify, we set

$$\overline{c\cdot\nabla_x\Phi^f}(t,y):=\int_{\mathbb{R}^{d_1}}c(x,y)\cdot\nabla_x\Phi^f(t,x,y)\mu^y(\mathrm{d}x).$$

The following result will play an important role below.

Lemma 4.3. *Let* $(\mathbf{A}_\sigma)$, $(\mathbf{A}_b)$ *and (2.1) hold. Assume that* $b,\sigma\in C_b^{\delta,\vartheta}$ *with* $0<\delta,\vartheta\leqslant 2$, $F,H,G\in L_p^\infty$ *and* $\lim_{\varepsilon\to0}\alpha_\varepsilon^\vartheta/\gamma_\varepsilon=0$. *Then for every* $f\in C_p^{\vartheta/2,\delta,\vartheta}$ *satisfying (4.1) and any* $q\geqslant 2$, *the following hold:*
(i) (Regime 1) if $c\in L_p^\infty$, *we have*

$$\mathbb{E}\left|\frac{1}{\gamma_\varepsilon}\int_0^t f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s\right|^q\leqslant C_t\Big(\frac{\alpha_\varepsilon^{\vartheta\wedge1}}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon\gamma_\varepsilon}\Big)^q;$$

(ii) (Regime 2) if $c\in C_p^{\delta,\vartheta}$, *we have*

$$\mathbb{E}\left|\frac{1}{\gamma_\varepsilon}\int_0^t f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s-\int_0^t\overline{c\cdot\nabla_x\Phi^f}(s,Y_s^\varepsilon)\mathrm{d}s\right|^q\leqslant C_t\Big(\frac{\alpha_\varepsilon^{\vartheta\wedge1}}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^q,$$

where $C_t>0$ *is a constant independent of* δ,ε.

Proof. We provide the proof for the two regimes separately.
(i) (Regime 1) This case follows by Lemma 4.1 directly since no homogenization occurs. In fact, we have

$$\mathbb{E}\left|\frac{1}{\gamma_\varepsilon}\int_0^t f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s\right|^q=\frac{1}{\gamma_\varepsilon^q}\mathbb{E}\left|\int_0^t f(s,X_s^\varepsilon,Y_s^\varepsilon)\mathrm{d}s\right|^q\leqslant C_1\frac{1}{\gamma_\varepsilon^q}\Big(\alpha_\varepsilon^{\vartheta\wedge1}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^q,$$

which in turn yields the desired result.

(ii) (Regime 2) Recall that $\alpha_\varepsilon^2 = \beta_\varepsilon \gamma_\varepsilon$ in this case. As in the proof of Lemma 4.1, by (4.4) we have

$$\begin{aligned}
\hat{\mathcal{Q}}(\varepsilon) &:= \mathbb{E}\left|\frac{1}{\gamma_\varepsilon}\int_0^t f(s, X_s^\varepsilon, Y_s^\varepsilon)\mathrm{d}s - \int_0^t \overline{c\cdot\nabla_x\Phi^f}(s, Y_s^\varepsilon)\mathrm{d}s\right|^q \\
&\leqslant C_2\left[\frac{\alpha_\varepsilon^{2q}}{\gamma_\varepsilon^q}\Big(|\Phi_n^f(0,x,y)|^q + \mathbb{E}|\Phi_n^f(t, X_t^\varepsilon, Y_t^\varepsilon)|^q\Big) + \frac{\alpha_\varepsilon^q}{\gamma_\varepsilon^q}\mathbb{E}|M_n^1(t)|^q + \frac{\alpha_\varepsilon^{2q}}{\gamma_\varepsilon^q}\mathbb{E}|M_n^2(t)|^q\right] \\
&\quad + C_2\frac{\alpha_\varepsilon^{2q}}{\gamma_\varepsilon^q}\mathbb{E}\left|\int_0^t \big(\partial_s + \mathscr{L}_1\big)\Phi_n^f(s, X_s^\varepsilon, Y_s^\varepsilon)\mathrm{d}s\right|^q + C_2\frac{\alpha_\varepsilon^{2q}}{\gamma_\varepsilon^{2q}}\mathbb{E}\left|\int_0^t \mathscr{L}_2\Phi_n^f(s, X_s^\varepsilon, Y_s^\varepsilon)\mathrm{d}s\right|^q \\
&\quad + C_2\frac{1}{\gamma_\varepsilon^q}\mathbb{E}\left|\int_0^t \big(\mathscr{L}_0\Phi_n^f - \mathscr{L}_0\Phi^f\big)(s, X_s^\varepsilon, Y_s^\varepsilon)\mathrm{d}s\right|^q \\
&\quad + C_2\mathbb{E}\left|\int_0^t \mathscr{L}_3\Phi_n^f(s, X_s^\varepsilon, Y_s^\varepsilon) - \mathscr{L}_3\Phi^f(s, X_s^\varepsilon, Y_s^\varepsilon)\mathrm{d}s\right|^q \\
&\quad + C_2\mathbb{E}\left|\int_0^t \mathscr{L}_3\Phi^f(s, X_s^\varepsilon, Y_s^\varepsilon) - \overline{c\cdot\nabla_x\Phi^f}(s, Y_s^\varepsilon)\mathrm{d}s\right|^q =: \sum_{i=1}^6 \hat{\mathcal{Q}}_i(\varepsilon),
\end{aligned}$$

where Φ_n^f is the mollifyer of Φ^f defined as in (3.5). Following exactly the same arguments as in the proof of Lemma 4.1, one can check that

$$\sum_{i=1}^5 \hat{\mathcal{Q}}_i(\varepsilon) \leqslant C_3\Big(\frac{\alpha_\varepsilon^q}{\gamma_\varepsilon^q} + \frac{\alpha_\varepsilon^{2q}}{\gamma_\varepsilon^q}n^{q(2-\vartheta)} + \frac{\alpha_\varepsilon^{2q}}{\gamma_\varepsilon^{2q}}n^{q(1-(\vartheta\wedge 1))} + \frac{1}{\gamma_\varepsilon^q}n^{-q\vartheta}\Big).$$

Taking $n = \alpha_\varepsilon^{-1}$, we further get

$$\sum_{i=1}^5 \hat{\mathcal{Q}}_i(\varepsilon) \leqslant C_4\Big(\frac{\alpha_\varepsilon^q}{\gamma_\varepsilon^q} + \frac{\alpha_\varepsilon^{q\vartheta}}{\gamma_\varepsilon^q}\Big) \leqslant C_4\frac{\alpha_\varepsilon^{q(\vartheta\wedge 1)}}{\gamma_\varepsilon^q}.$$

For the last term, note that by definition we have

$$\mathscr{L}_3\Phi^f(t,x,y) - \overline{c\cdot\nabla_x\Phi^f}(t,y) = c(x,y)\cdot\nabla_x\Phi^f(t,x,y) - \overline{c\cdot\nabla_x\Phi^f}(t,y),$$

which satisfies the centering condition (4.1). Furthermore, since $c \in C_p^{\delta,\vartheta}$, $\Phi^f \in C_p^{\vartheta/2, 2+\delta, \vartheta}$ and by Lemma 3.3, we have $c\cdot\nabla_x\Phi^f - \overline{c\cdot\nabla_x\Phi^f} \in C_p^{\vartheta/2,\delta,\vartheta}$. As a direct result of Lemma 4.1, we get

$$\hat{\mathcal{Q}}_6(\varepsilon) \leqslant C_5\Big(\alpha_\varepsilon^{\vartheta\wedge 1} + \frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^q.$$

Consequently, we arrive at

$$\hat{\mathcal{Q}}(\varepsilon) \leqslant C_6\Big(\frac{\alpha_\varepsilon^{\vartheta\wedge 1}}{\gamma_\varepsilon} + \alpha_\varepsilon^{\vartheta\wedge 1} + \frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^q \leqslant C_6\Big(\frac{\alpha_\varepsilon^{\vartheta\wedge 1}}{\gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^q,$$

and the proof is finished. □

4.2. **Proof of Theorem 2.1.** Throughout this subsection, we fix $T > 0$ and always assume that the conditions in Theorem 2.1 hold. Recall that $\bar{Y}_t^k$ $(k = 1, 2)$ are given by (2.5). Due to our low regularity assumptions on the drift coefficients of system (1.6), it seems to be not possible to prove the strong convergence of Y_t^ε to $\bar{Y}_t^k$ directly. For this reason, we shall use Zvonkin's argument to transform the equations for Y_t^ε and $\bar{Y}_t$ into new ones.

Consider the following backward PDE on $[0,T] \times \mathbb{R}^{d_2}$: for $k = 1, 2$,

$$\begin{cases} \partial_t v_k(t,y) + \bar{\mathscr{L}}_k v_k(t,y) + \bar{F}_k(t,y) = 0, \quad t \in [0,T), \\ v_k(T,y) = 0, \end{cases} \tag{4.6}$$

where $\bar{\mathscr{L}}_k$ are defined by (2.21). By Lemma 3.3, we have for every $k = 1, 2$, $\bar{F}_k \in C_b^{\vartheta/2,\vartheta}$ with $0 < \vartheta \leqslant 1$. Thus under $(\mathbf{A}_G)$, it is well known that there exits a unique solution $v_k \in L^\infty\big([0,T]; C_b^{2+\vartheta}(\mathbb{R}^{d_2})\big) \cap C_b^{1+\vartheta/2}\big([0,T]; L^\infty(\mathbb{R}^{d_2})\big)$ for equation (4.6), see e.g. [43, Chapter IV, Section 5]. Moreover, we can choose T small enough so that for any $0 < t \leqslant T$ (see e.g. [19, Theorem 2] or [21, Theorem 2]),

$$|\nabla_y v_k(t,y)| \leqslant 1/2, \quad \forall y \in \mathbb{R}^{d_2}.$$

Define the transformation function by

$$\Gamma_k(t,y) := y + v_k(t,y).$$

Then the map $y \mapsto \Gamma_k(t,y)$ is a C^1-diffeomorphism and for every $t \in (0,T]$ and $y \in \mathbb{R}^{d_2}$,

$$1/2 \leqslant |\nabla_y \Gamma_k(t,y)| \leqslant 3/2. \tag{4.7}$$

Let us define two new processes by

$$\bar{V}_t^k := \Gamma_k(t, \bar{Y}_t^k) \quad \text{and} \quad V_t^{k,\varepsilon} := \Gamma_k(t, Y_t^\varepsilon). \tag{4.8}$$

We have the following result.

Lemma 4.4 (Zvonkin's transformation)**.** *Let $\bar{V}_t^k$ and $V_t^{k,\varepsilon}$ be defined by (4.8). Then we have for $k = 1, 2$,*

$$\mathrm{d}\bar{V}_t^k = G(t, \bar{Y}_t^k)\nabla_y \Gamma_k(t, \bar{Y}_t^k)\mathrm{d}W_t^2, \qquad V_0^k = \Gamma_k(0,y) \tag{4.9}$$

and

$$\begin{aligned} \mathrm{d}V_t^{k,\varepsilon} = \big[F(t, X_t^\varepsilon, Y_t^\varepsilon) - \bar{F}_k(t, Y_t^\varepsilon)\big]\nabla_y \Gamma_k(t, Y_t^\varepsilon)\mathrm{d}t + \gamma_\varepsilon^{-1} H(t, X_t^\varepsilon, Y_t^\varepsilon)\nabla_y \Gamma_k(t, Y_t^\varepsilon)\mathrm{d}t \\ + G(t, Y_t^\varepsilon)\nabla_y \Gamma_k(t, Y_t^\varepsilon)\mathrm{d}W_t^2, \qquad V_0^\varepsilon = \Gamma_k(0,y). \end{aligned} \tag{4.10}$$

Proof. Using Itô's formula, we have for $k = 1, 2$,

$$\begin{aligned} v_k(t, Y_t^\varepsilon) &= v_k(0,y) + \int_0^t \big(\partial_s + \mathscr{L}_1\big) v_k(s, Y_s^\varepsilon)\mathrm{d}s + \frac{1}{\gamma_\varepsilon}\int_0^t \mathscr{L}_2 v_k(s, Y_s^\varepsilon)\mathrm{d}s \\ &\quad + \int_0^t G(s, Y_s^\varepsilon)\nabla_y v_k(s, Y_s^\varepsilon)\mathrm{d}W_s^2 \\ &= v_k(0,y) + \int_0^t \big(\partial_s + \bar{\mathscr{L}}_k\big) v_k(s, Y_s^\varepsilon)\mathrm{d}s \\ &\quad + \int_0^t \big[F(s, X_s^\varepsilon, Y_s^\varepsilon) - \bar{F}_k(s, Y_s^\varepsilon)\big]\nabla_y v_k(s, Y_s^\varepsilon)\mathrm{d}s \end{aligned}$$

$$+\frac{1}{\gamma_\varepsilon}\int_0^t \mathscr{L}_2 v_k(s,Y_s^\varepsilon)\mathrm{d}s+\int_0^t G(s,Y_s^\varepsilon)\nabla_y v_k(s,Y_s^\varepsilon)\mathrm{d}W_s^2,$$

where $\mathscr{L}_1$ and $\mathscr{L}_2$ are given by (1.8). Taking into account (4.6), we further get

$$\begin{aligned}v_k(t,Y_t^\varepsilon)=v_k(0,y)-&\int_0^t \bar{F}_k(s,Y_s^\varepsilon)\mathrm{d}s+\int_0^t\big[F(s,X_s^\varepsilon,Y_s^\varepsilon)-\bar{F}_k(s,Y_s^\varepsilon)\big]\nabla_y v_k(s,Y_s^\varepsilon)\mathrm{d}s\\&+\frac{1}{\gamma_\varepsilon}\int_0^t \mathscr{L}_2 v_k(s,Y_s^\varepsilon)\mathrm{d}s+\int_0^t G(s,Y_s^\varepsilon)\nabla_y v_k(s,Y_s^\varepsilon)\mathrm{d}W_s^2,\end{aligned}$$

This together with the equation for Y_t^ε and the fact that $\nabla_y\Gamma_k(t,y)=\mathbb{I}_{d_2}+\nabla_y v_k(t,y)$ yields (4.10). The proof of (4.9) is easier and follows by the same argument. □

Now, we are in the position to give:

Proof of Theorem 2.1. Let us first assume that $T>0$ is sufficiently small so that (4.7) holds. By the definition (4.8), we have for any $t\in[0,T]$, $k=1,2$ and $q\geqslant 2$ that

$$\mathbb{E}\big|Y_t^\varepsilon-\bar{Y}_t^k\big|^q\leqslant C_q\,\mathbb{E}\big|V_t^{k,\varepsilon}-\bar{V}_t^k\big|^q. \tag{4.11}$$

In view of (4.9) and (4.10), we may write

$$\begin{aligned}V_t^{k,\varepsilon}-\bar{V}_t^k=&\int_0^t\big[G(s,Y_s^\varepsilon)\nabla_y\Gamma_k(s,Y_s^\varepsilon)-G(s,\bar{Y}_s^k)\nabla_y\Gamma_k(s,\bar{Y}_s^k)\big]\mathrm{d}W_s^2\\&+\int_0^t\big[F(s,X_s^\varepsilon,Y_s^\varepsilon)-\bar{F}_k(s,Y_s^\varepsilon)\big]\nabla_y\Gamma_k(s,Y_s^\varepsilon)\mathrm{d}s\\&+\frac{1}{\gamma_\varepsilon}\int_0^t H(s,X_s^\varepsilon,Y_s^\varepsilon)\nabla_y\Gamma_k(s,Y_s^\varepsilon)\mathrm{d}s.\end{aligned}$$

Taking expectation of both sides of the above equality, we get that there exists a constant $C_0>0$ such that

$$\begin{aligned}\mathbb{E}|V_t^{k,\varepsilon}-\bar{V}_t^k|^q\leqslant{}& C_0\mathbb{E}\left(\int_0^t\big|G(s,Y_s^\varepsilon)\nabla_y\Gamma_k(s,Y_s^\varepsilon)-G(s,\bar{Y}_s^k)\nabla_y\Gamma_k(s,\bar{Y}_s^k)\big|^2\mathrm{d}s\right)^{q/2}\\&+C_0\mathbb{E}\bigg|\int_0^t\big[F(s,X_s^\varepsilon,Y_s^\varepsilon)-\bar{F}_k(s,Y_s^\varepsilon)\big]\nabla_y\Gamma_k(s,Y_s^\varepsilon)\mathrm{d}s\\&\qquad+\frac{1}{\gamma_\varepsilon}\int_0^t H(s,X_s^\varepsilon,Y_s^\varepsilon)\nabla_y\Gamma_k(s,Y_s^\varepsilon)\mathrm{d}s\bigg|^q=:\mathscr{Q}_1(t,\varepsilon)+\mathscr{Q}_2(t,\varepsilon).\end{aligned}$$

Note that for each $k=1,2$,

$$y\mapsto G(t,\cdot)\nabla_y\Gamma_k(t,\cdot)\in C_b^1(\mathbb{R}^{d_2}).$$

Thus we have

$$\mathscr{Q}_1(t,\varepsilon)\leqslant C_1\mathbb{E}\left(\int_0^t|Y_s^\varepsilon-\bar{Y}_s^k|^2\mathrm{d}s\right)^{q/2}\leqslant C_t\,\mathbb{E}\left(\int_0^t|Y_s^\varepsilon-\bar{Y}_s^k|^q\mathrm{d}s\right), \tag{4.12}$$

where $C_t>0$ is a constant independent of ε. Below, we proceed to control the second term according to Regime 1 and Regime 2 in (1.9) separately.

Regime 1 ($k=1$). In this case, note that the function $[F(t,x,y)-\bar{F}_1(t,y)]\nabla_y\Gamma_1(t,y)$ satisfies the centering condition (4.1) and belongs to $C_p^{\vartheta/2,\delta,\vartheta}$. Thus by Lemma 4.1 we have

$$\mathbb{E}\bigg|\int_0^t\big[F(s,X_s^\varepsilon,Y_s^\varepsilon)-\bar{F}_1(s,Y_s^\varepsilon)\big]\nabla_y\Gamma_1(s,Y_s^\varepsilon)\mathrm{d}s\bigg|^q\leqslant C_2\Big(\alpha_\varepsilon^{\vartheta\wedge1}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^q.$$

At the same time, thanks to assumption $(\mathbf{A}_H)$, the function $H(t,x,y)\nabla_y\Gamma_1(t,y)$ also satisfies the centering condition (4.1). It then follows by Lemma 4.3 (i) that

$$\mathbb{E}\bigg|\frac{1}{\gamma_\varepsilon}\int_0^t H(s,X_s^\varepsilon,Y_s^\varepsilon)\nabla_y\Gamma_1(s,Y_s^\varepsilon)\mathrm{d}s\bigg|^q\leqslant C_3\Big(\frac{\alpha_\varepsilon^{\vartheta\wedge1}}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon\gamma_\varepsilon}\Big)^q.$$

Combining the above estimates, we get

$$\mathscr{Q}_2(t,\varepsilon)\leqslant C_4\Big(\frac{\alpha_\varepsilon^{\vartheta\wedge1}}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon\gamma_\varepsilon}\Big)^q.$$

Now, in view of (4.11) and (4.12), we arrive at

$$\mathbb{E}\big|Y_t^\varepsilon-\bar{Y}_t^1\big|^q\leqslant C_5\mathbb{E}\left(\int_0^t|Y_s^\varepsilon-\bar{Y}_s^1|^q\mathrm{d}s\right)+C_6\Big(\frac{\alpha_\varepsilon^{\vartheta\wedge1}}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon\gamma_\varepsilon}\Big)^q,$$

which in turn yields the desired result by Gronwall's inequality. For general $T>0$, the result can be proved by induction and analogous arguments.

Regime 2 ($k=2$). In this case, recall that we have $\bar{F}_2(t,y)=\bar{F}_1(t,y)+\overline{c\cdot\nabla_x\Phi}(t,y)$, where Φ solves the Poisson equation (2.3) and $\overline{c\cdot\nabla_x\Phi}$ is given by (2.15). Thus, we deduce that

$$\begin{aligned}\mathscr{Q}_2(t,\varepsilon)&\leqslant C_2\mathbb{E}\bigg|\int_0^t\big[F(s,X_s^\varepsilon,Y_s^\varepsilon)-\bar{F}_1(s,Y_s^\varepsilon)\big]\nabla_y\Gamma_2(s,Y_s^\varepsilon)\mathrm{d}s\bigg|^q\\&\quad+C_2\mathbb{E}\bigg|\frac{1}{\gamma_\varepsilon}\int_0^t H(s,X_s^\varepsilon,Y_s^\varepsilon)\nabla_y\Gamma_2(s,Y_s^\varepsilon)\mathrm{d}s-\int_0^t\overline{c\cdot\nabla_x\Phi}(s,Y_s^\varepsilon)\nabla_y\Gamma_2(s,Y_s^\varepsilon)\mathrm{d}s\bigg|^q\\&=:\mathscr{Q}_{21}(t,\varepsilon)+\mathscr{Q}_{22}(t,\varepsilon).\end{aligned}$$

Following exactly the same arguments as above, we get

$$\mathscr{Q}_{21}(t,\varepsilon)\leqslant C_3\Big(\alpha_\varepsilon^{\vartheta\wedge1}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^q.$$

On the other hand, let $\tilde{\Phi}(t,x,y):=\Phi(t,x,y)\cdot\nabla_y\Gamma_2(t,y)$. Then one can check that

$$\mathscr{L}_0(x,y)\tilde{\Phi}(t,x,y)=-H(t,x,y)\cdot\nabla_y\Gamma_2(t,y)$$

and

$$\overline{c\cdot\nabla_x\tilde{\Phi}}(t,y):=\int_{\mathbb{R}^{d_1}}c(x,y)\cdot\nabla_x\tilde{\Phi}(t,x,y)\mu^y(\mathrm{d}x)=\overline{c\cdot\nabla_x\Phi}(t,y)\cdot\nabla_y\Gamma_2(t,y).$$

Consequently, by Lemma 4.3 (ii) we have

$$\mathscr{Q}_{22}(t,\varepsilon)\leqslant C_4\Big(\frac{\alpha_\varepsilon^{\vartheta\wedge1}}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^q.$$

Combining the above computations, we arrive at

$$\mathbb{E}\big|Y_t^\varepsilon - \bar{Y}_t^2\big|^q \leqslant C_5\mathbb{E}\left(\int_0^t |Y_s^\varepsilon - \bar{Y}_s^2|^q \mathrm{d}s\right) + C_6\Big(\frac{\alpha_\varepsilon^{\vartheta\wedge 1}}{\gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^q,$$

which in turn yields the desired result by Gronwall's inequality. Hence the whole proof is finished. □

5. Central limit theorem without homogenization

In this section, we study the central limit theorem for SDE (2.8). We shall first derive some weak fluctuation estimates in Subsection 5.1. Then we prove Theorem 2.3 in Subsection 5.2.

5.1. **Weak fluctuation estimate (i).**

Recall that $Y_t^{0,\varepsilon}$ converges strongly to $\bar{Y}_t^1$, and $Z_t^{0,\varepsilon}$, $\bar{Z}_{\ell,t}^0$ ($\ell = 1,2,3$) are defined by (2.9), (2.13), respectively. To prove the weak convergence of $Z_t^{0,\varepsilon}$ to $\bar{Z}_{\ell,t}^0$, we shall view the process $(X_t^\varepsilon, Y_t^{0,\varepsilon}, \bar{Y}_t^1, Z_t^{0,\varepsilon}, \bar{Z}_{\ell,t}^0)$ as a whole system, i.e., we consider

$$\begin{cases} \mathrm{d}X_t^\varepsilon = \alpha_\varepsilon^{-2} b(X_t^\varepsilon, Y_t^{0,\varepsilon})\mathrm{d}t + \beta_\varepsilon^{-1} c(X_t^\varepsilon, Y_t^{0,\varepsilon})\mathrm{d}t + \alpha_\varepsilon^{-1}\sigma(X_t^\varepsilon, Y_t^{0,\varepsilon})\mathrm{d}W_t^1, & X_0 = x,\\ \mathrm{d}Y_t^{0,\varepsilon} = F(t, X_t^\varepsilon, Y_t^{0,\varepsilon})\mathrm{d}t + G(t, Y_t^{0,\varepsilon})\mathrm{d}W_t^2, & Y_0^{0,\varepsilon} = y,\\ \mathrm{d}\bar{Y}_t^1 = \bar{F}_1(t, \bar{Y}_t^1)\mathrm{d}t + G(t, \bar{Y}_t^1)\mathrm{d}W_t^2, & \bar{Y}_0^1 = y,\\ \mathrm{d}Z_t^{0,\varepsilon} = \dfrac{1}{\eta_\varepsilon}\Big[F(t, X_t^\varepsilon, Y_t^{0,\varepsilon}) - \bar{F}_1(t, \bar{Y}_t^1)\Big]\mathrm{d}t + \dfrac{1}{\eta_\varepsilon}\Big[G(t, Y_t^{0,\varepsilon}) - G(t, \bar{Y}_t^1)\Big]\mathrm{d}W_t^2, & Z_0^{0,\varepsilon} = 0,\\ \mathrm{d}\bar{Z}_{\ell,t}^0 = \varrho_\ell^0(t, \bar{Y}_t^1, \bar{Z}_{\ell,t}^0)\mathrm{d}t + \chi_\ell^0(t, \bar{Y}_t^1, \bar{Z}_{\ell,t}^0)\mathrm{d}W_t^2, & \bar{Z}_{\ell,0}^1 = 0, \end{cases}$$

where for $\ell = 1,2,3$, η_ε is given in (2.10), and $\varrho_\ell^0(t,y,z)$ and $\chi_\ell^0(t,y,z)$ denote the drift and diffusion coefficients for $\bar{Z}_{\ell,t}^0$ in SDE (2.13), respectively. We write ϱ_ℓ^0 and χ_ℓ^0 here just for simplicity, and we shall not use them below.

Note that by definition, we can write

$$\begin{aligned} \mathrm{d}Z_t^{0,\varepsilon} &= \frac{1}{\eta_\varepsilon}\Big[F(t, X_t^\varepsilon, Y_t^{0,\varepsilon}) - \bar{F}_1(t, Y_t^{0,\varepsilon})\Big]\mathrm{d}t\\ &\quad + \left(\frac{1}{\eta_\varepsilon}\Big[\bar{F}_1(t, Y_t^{0,\varepsilon}) - \bar{F}_1(t, \bar{Y}_t^1)\Big]\mathrm{d}t + \frac{1}{\eta_\varepsilon}\Big[G(t, Y_t^{0,\varepsilon}) - G(t, \bar{Y}_t^1)\Big]\mathrm{d}W_t^2\right). \end{aligned}$$

To shorten the notation, we define

$$\mathscr{L}_4^1 := \mathscr{L}_4^1(t,x,y,z) := \sum_{i=1}^{d_2}\big[F^i(t,x,y) - \bar{F}_1^i(t,y)\big]\frac{\partial}{\partial z_i}, \tag{5.1}$$

and for $k=1,2$,

$$
\begin{aligned}
\mathscr{L}_5^{k,\varepsilon} := \mathscr{L}_5^{k,\varepsilon}(t,y,\bar{y},z) &:= \eta_\varepsilon^{-1}\sum_{i=1}^{d_2}\big[\bar{F}_k^i(t,y)-\bar{F}_k^i(t,\bar{y})\big]\frac{\partial}{\partial z_i} \\
&\quad + \eta_\varepsilon^{-2}/2\sum_{i,j=1}^{d_2}\Big(\big[G(t,y)-G(t,\bar{y})\big]\big[G(t,y)-G(t,\bar{y})\big]^*\Big)^{ij}\frac{\partial^2}{\partial z_i\partial z_j} \qquad (5.2)\\
&\quad + \eta_\varepsilon^{-1}/2\sum_{i,j=1}^{d_2}\Big(G(t,y)\big[G(t,y)-G(t,\bar{y})\big]^*\Big)^{ij}\frac{\partial^2}{\partial y_i\partial z_j}.
\end{aligned}
$$

Let $f(t,x,y,z)$ be a function satisfying the centering condition, i.e.,

$$
\int_{\mathbb{R}^{d_1}} f(t,x,y,z)\mu^y(\mathrm{d}x)=0, \quad \forall (t,y,z)\in\mathbb{R}_+\times\mathbb{R}^{d_2}\times\mathbb{R}^{d_2}. \tag{5.3}
$$

Let $\Phi^f(t,x,y,z)$ denote the unique solution to the following Poisson equation:

$$
\mathscr{L}_0(x,y)\Phi^f(t,x,y,z)=-f(t,x,y,z), \tag{5.4}
$$

where $(t,y,z)\in\mathbb{R}_+\times\mathbb{R}^{d_2}\times\mathbb{R}^{d_2}$ are regarded as parameters. We have the following fluctuation estimate for the process $(X_t^\varepsilon, Y_t^{0,\varepsilon}, Z_t^{0,\varepsilon})$.

Lemma 5.1. *Let* $(\mathbf{A}_\sigma)$, $(\mathbf{A}_b)$, *(2.1) hold and* $0<\delta,\vartheta\leqslant 2$. *Assume that* $b,\sigma\in C_b^{\delta,1\vee\vartheta}$, $c\in L_p^\infty$, $F\in C_p^{0,\delta,1}$ *and* $G\in C_b^{0,1}$. *Then for every* $f\in C_p^{\vartheta/2,\delta,\vartheta,2}$ *satisfying (5.3), we have*

$$
\begin{aligned}
\mathbb{E}\left(\int_0^t f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right) &\leqslant C_t\bigg[\alpha_\varepsilon^\vartheta+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\mathbb{E}\left(\int_0^t \mathscr{L}_3\Phi^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\\
&\qquad +\frac{\alpha_\varepsilon^2}{\eta_\varepsilon}\mathbb{E}\left(\int_0^t \mathscr{L}_4^1\Phi^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\bigg].
\end{aligned} \tag{5.5}
$$

where $\mathscr{L}_3$ *and* $\mathscr{L}_4^1$ *are given by (1.8) and (5.1), respectively, and* $C_t>0$ *is a constant independent of* δ,ε.

Remark 5.2. *(i) Note that under the above assumptions and according to Theorem 3.1, we in fact have* $\Phi^f\in C_p^{\vartheta/2,2+\delta,\vartheta,2}$. *Thus we get*

$$
\begin{aligned}
&\mathbb{E}\left(\int_0^t \mathscr{L}_3\Phi^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)+\mathbb{E}\left(\int_0^t \mathscr{L}_4^1\Phi^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\\
&\leqslant C_0\mathbb{E}\left(\int_0^t \big(1+|X_s^\varepsilon|^{2m}\big)\mathrm{d}s\right)<\infty,
\end{aligned}
$$

which in turn yields that

$$
\mathbb{E}\left(\int_0^t f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\leqslant C_t\Big(\alpha_\varepsilon^\vartheta+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon}\Big). \tag{5.6}
$$

However, the homogenization effects of the last two terms in (5.5) will appear when we study the central limit theorems, so we just keep them on the right hand side for later use.

(ii) Compared with Lemma 4.1, we do not take absolute value for the integral over $[0,t]$ on the left hand side of (5.5). We shall see that the involved martingale part will not play any role in the control of the error bound in this case.

Proof. Let Φ_n^f be the mollifyer of Φ^f defined as in (3.5). Using Itô's formula, we have

$$\begin{aligned}
\Phi_n^f(t,X_t^\varepsilon,Y_t^{0,\varepsilon},Z_t^{0,\varepsilon}) &= \Phi_n^f(0,x,y,0)+\frac{1}{\alpha_\varepsilon}\tilde{M}_n^1(t)+\tilde{M}_n^2(t)+\frac{1}{\eta_\varepsilon}\tilde{M}_n^3(t)\\
&\quad+\int_0^t\big(\partial_s+\mathscr{L}_1+\mathscr{L}_5^{1,\varepsilon}\big)\Phi_n^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\\
&\quad+\frac{1}{\alpha_\varepsilon^2}\int_0^t\mathscr{L}_0\Phi_n^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s+\frac{1}{\beta_\varepsilon}\int_0^t\mathscr{L}_3\Phi_n^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\\
&\quad+\frac{1}{\eta_\varepsilon}\int_0^t\mathscr{L}_4^1\Phi_n^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s.
\end{aligned}$$

where $\mathscr{L}_1$ and $\mathscr{L}_5^{1,\varepsilon}$ are given by (1.8) and (5.2), respectively, and $\tilde{M}_n^1(t)$, $\tilde{M}_n^2(t)$, $\tilde{M}_n^3(t)$ are martingales defined by

$$\begin{aligned}
\tilde{M}_n^1(t)&:=\int_0^t\sigma(X_s^\varepsilon,Y_s^{0,\varepsilon})\nabla_x\Phi_n^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}W_s^1,\\
\tilde{M}_n^2(t)&:=\int_0^t G(s,Y_s^{0,\varepsilon})\nabla_y\Phi_n^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}W_s^2,\\
\tilde{M}_n^3(t)&:=\int_0^t\big[G(s,Y_s^{0,\varepsilon})-G(s,\bar{Y}_s^1)\big]\nabla_z\Phi_n^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}W_s^2.
\end{aligned}$$

Taking expectation and in view of (5.4), we have

$$\begin{aligned}
\mathcal{V}(\varepsilon)&:=\mathbb{E}\bigg(\int_0^t f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\bigg)=\alpha_\varepsilon^2\,\mathbb{E}\big[\Phi_n^f(0,x,y,0)-\Phi_n^f(t,X_t^\varepsilon,Y_t^{0,\varepsilon},Z_t^{0,\varepsilon})\big]\\
&\quad+\alpha_\varepsilon^2\mathbb{E}\left(\int_0^t\big(\partial_s+\mathscr{L}_1+\mathscr{L}_5^{1,\varepsilon}\big)\Phi_n^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\\
&\quad+\mathbb{E}\left(\int_0^t\big[\mathscr{L}_0\Phi_n^f-\mathscr{L}_0\Phi^f\big](s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\\
&\quad+\bigg[\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\mathbb{E}\left(\int_0^t\big(\mathscr{L}_3\Phi_n^f-\mathscr{L}_3\Phi^f\big)(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\\
&\quad+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon}\mathbb{E}\left(\int_0^t\big(\mathscr{L}_4^1\Phi_n^f-\mathscr{L}_4^1\Phi^f\big)(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\bigg]\\
&\quad+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\mathbb{E}\left(\int_0^t\mathscr{L}_3\Phi^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\\
&\quad+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon}\mathbb{E}\left(\int_0^t\mathscr{L}_4^1\Phi^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)=:\sum_{i=1}^6\mathcal{V}_i(\varepsilon).
\end{aligned}$$

Let us handle the term involving $\mathscr{L}_5^{1,\varepsilon}$. Due to the assumptions that $b,\sigma\in C_b^{\delta,1\vee\vartheta}$, $F\in C_p^{0,\delta,1}$, and by Lemma 3.3, we have $\bar{F}_1\in C_b^{0,1}$. This together with the condition

$G \in C_b^{0,1}$ and the mean value theorem yields that for some $m>0$ and $C_1>0$,

$$
\begin{aligned}
&\mathbb{E}\left(\int_0^t \mathscr{L}_5^{1,\varepsilon}\Phi_n^f(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right) \leqslant C_1\|\nabla_z^2\Phi^f\|_{L_p^\infty}\mathbb{E}\bigg(\int_0^t \big(1+|X_s^\varepsilon|^m\big)\\
&\qquad\qquad\qquad\times\left[\frac{|\bar F_1(s,Y_s^{0,\varepsilon})-\bar F_1(s,\bar Y_s^1)|}{\eta_\varepsilon}+\frac{|G(s,Y_s^{0,\varepsilon})-G(s,\bar Y_s^1)|^2}{\eta_\varepsilon^2}\right]\mathrm{d}s\bigg)\\
&\quad + C_1\|\nabla_y\nabla_z\Phi_n^f\|_{L_p^\infty}\mathbb{E}\bigg(\int_0^t \big(1+|X_s^\varepsilon|^m\big)\left[\frac{|G(s,Y_s^{0,\varepsilon})-G(s,\bar Y_s^1)|}{\eta_\varepsilon}\right]\mathrm{d}s\bigg)\\
&\leqslant C_1 n^{1-(\vartheta\wedge 1)}\mathbb{E}\left(\int_0^t \big(1+|X_s^\varepsilon|^m\big)\big(1+|Z_s^{0,\varepsilon}|^2\big)\mathrm{d}s\right)\leqslant C_1 n^{1-(\vartheta\wedge 1)}. \qquad (5.7)
\end{aligned}
$$

Then we may argue as in the proof of Lemma 4.1 to get that

$$
\sum_{i=1}^4 \mathcal{V}_i(\varepsilon)\leqslant C_2\Big(\alpha_\varepsilon^2+\alpha_\varepsilon^2 n^{1-(\vartheta\wedge 1)}+\alpha_\varepsilon^2 n^{2-\vartheta}+n^{-\vartheta}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}n^{-\vartheta}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon}n^{-\vartheta}\Big).
$$

Taking $n=\alpha_\varepsilon^{-1}$, and noticing that $\alpha_\varepsilon^2/\beta_\varepsilon\to 0$ and $\alpha_\varepsilon^2/\eta_\varepsilon\to 0$ as $\varepsilon\to 0$, we get the desired result. $\square$

5.2. **Proof of Theorem 2.3.** We are now in the position to give:

Proof of Theorem 2.3. For every $\varphi\in C_b^4(\mathbb{R}^{d_2})$ and $\ell=1,2,3$, let u_ℓ^0 be the solution to the Cauchy problem (3.10) on $[0,T]\times\mathbb{R}^{d_2}\times\mathbb{R}^{d_2}$ with $k=0$, and define

$$
\tilde u_\ell^0(t,y,z):=u_\ell^0(T-t,y,z),\quad t\in[0,T].
$$

Then for any $y\in\mathbb{R}^{d_2}$, we have $\tilde u_\ell^0(T,y,z)\equiv\varphi(z)$ and $\tilde u_\ell^0(0,y,0)=\mathbb{E}\varphi(\bar Z_{\ell,T}^0)$. As a result, for $\ell=1,2,3$ we get that

$$
\mathcal{R}_\ell^0(\varepsilon):=\mathbb{E}\varphi(Z_T^{0,\varepsilon})-\mathbb{E}\varphi(\bar Z_{\ell,T}^0)=\mathbb{E}\tilde u_\ell^0(T,Y_T^{0,\varepsilon},Z_T^{0,\varepsilon})-\tilde u_\ell^0(0,y,0).
$$

According to Theorem 3.4 and by Itô's formula,

$$
\begin{aligned}
\tilde u_\ell^0(T,Y_T^{0,\varepsilon},Z_T^{0,\varepsilon})&=\tilde u_\ell^0(0,y,0)+\int_0^T\Big(\partial_s+\mathscr{L}_1+\mathscr{L}_5^{1,\varepsilon}\Big)\tilde u_\ell^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\\
&\quad+\frac{1}{\eta_\varepsilon}\int_0^T\mathscr{L}_4^1\tilde u_\ell^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s+M_{\ell,T}^1+\frac{1}{\eta_\varepsilon}M_{\ell,T}^2,
\end{aligned}
$$

where $M_{\ell,t}^1$ and $M_{\ell,t}^2$ are martingales given by

$$
\begin{aligned}
M_{\ell,t}^1&:=\int_0^t G(s,Y_s^{0,\varepsilon})\nabla_y\tilde u_\ell^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}W_s^2,\\
M_{\ell,t}^2&:=\int_0^t\big[G(s,Y_s^{0,\varepsilon})-G(s,\bar Y_s^1)\big]\nabla_z\tilde u_\ell^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}W_s^2.
\end{aligned}
$$

Thus we further have

$$
\begin{aligned}
\mathcal{R}_\ell^0(\varepsilon)&\leqslant\mathbb{E}\left(\int_0^T\Big[\partial_s+\mathscr{L}_1+\mathscr{L}_5^{1,\varepsilon}\Big]\tilde u_\ell^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\\
&\quad+\frac{1}{\eta_\varepsilon}\mathbb{E}\left(\int_0^T\mathscr{L}_4^1\tilde u_\ell^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right).
\end{aligned}
$$

Note that by definition, the functions

$$\mathscr{L}_4^1\tilde{u}_\ell^0(t,y,z)=\big[F(t,x,y)-\bar{F}_1(t,y)\big]\nabla_z\tilde{u}_\ell^0(t,y,z)$$

satisfy the centering condition (5.3). Recall that $\Upsilon(t,x,y)$ solves the Poisson equation (2.11). Define for $\ell=1,2,3$,

$$\tilde{\Upsilon}_\ell^0(t,x,y,z):=\Upsilon(t,x,y)\nabla_z\tilde{u}_\ell^0(t,y,z).$$

Then

$$\mathscr{L}_0(x,y)\tilde{\Upsilon}_\ell^0(t,x,y,z)=-\big[F(t,x,y)-\bar{F}_1(t,y)\big]\nabla_z\tilde{u}_\ell^0(t,y,z).$$

Furthermore, by Lemma 3.3 we have $\bar{F}_1\in C_b^{(1+\vartheta)/2,1+\vartheta}$, which in turn implies that $[F-\bar{F}_1]\nabla_z\tilde{u}_\ell^0\in C_p^{(1+\vartheta)/2,\delta,1+\vartheta,2}$. Consequently, it follows by Lemma 5.1 that

$$\begin{aligned}
&\frac{1}{\eta_\varepsilon}\mathbb{E}\left(\int_0^T\mathscr{L}_4^1\tilde{u}_\ell^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\\
&\leqslant C_T\Bigg[\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon}\mathbb{E}\left(\int_0^T\mathscr{L}_3\tilde{\Upsilon}_\ell^0(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\\
&\quad+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2}\mathbb{E}\left(\int_0^T\mathscr{L}_4^1\tilde{\Upsilon}_\ell^0(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\Bigg].
\end{aligned}$$

Thus we arrive at

$$\begin{aligned}
\mathcal{R}_\ell^0(\varepsilon)\leqslant C_T\Bigg[&\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon}+\mathbb{E}\left(\int_0^T\Big[\partial_s+\mathscr{L}_1+\mathscr{L}_5^{1,\varepsilon}\Big]\tilde{u}_\ell^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\\
&+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon}\mathbb{E}\left(\int_0^T\mathscr{L}_3\tilde{\Upsilon}_\ell^0(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\\
&+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2}\mathbb{E}\left(\int_0^T\mathscr{L}_4^1\tilde{\Upsilon}_\ell^0(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\Bigg]=:C_T\Bigg[\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon}+\sum_{i=1}^3\mathscr{C}_i(\varepsilon)\Bigg].
\end{aligned}$$

Below, we consider $\ell=1,2,3$ separately, which correspond to Regime 0.1-Regime 0.3 in the conclusion of Theorem 2.3, respectively.

Case $\ell=1$. (Regime 0.1 in (2.10)). Note that we have

$$\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon}=1\quad\text{and}\quad\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2}=\frac{\beta_\varepsilon^2}{\alpha_\varepsilon^2}\to0\quad\text{as}\quad\varepsilon\to0.$$

Thus we deduce

$$\mathscr{C}_3(\varepsilon)\leqslant C_T\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2}\mathbb{E}\left(\int_0^T\big(1+|X_s^\varepsilon|^{2m}\big)\mathrm{d}s\right)\leqslant C_T\frac{\beta_\varepsilon^2}{\alpha_\varepsilon^2}.$$

Furthermore, recall that we have $\bar{\mathcal{L}}_1^0=\bar{\mathscr{L}}_3^0+\bar{\mathscr{L}}_5^1$ in this case with $\bar{\mathscr{L}}_3^0$ and $\bar{\mathscr{L}}_5^1$ given by (2.22) and (2.27), respectively. As a result,

$$\partial_t\tilde{u}_1^0+\big(\bar{\mathscr{L}}_1+\bar{\mathscr{L}}_3^0+\bar{\mathscr{L}}_5^1\big)\tilde{u}_1^0=0.$$

This in turn yields that

$$\begin{aligned}\mathscr{C}_1(\varepsilon)+\mathscr{C}_2(\varepsilon)=&\,\mathbb{E}\left(\int_0^T\big(\mathscr{L}_1-\bar{\mathscr{L}}_1\big)\tilde{u}_1^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\\&+\mathbb{E}\left(\int_0^T\mathscr{L}_3\tilde{\Upsilon}_1^0(s,X_s^\varepsilon,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})-\bar{\mathscr{L}}_3^0\tilde{u}_1^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)\\&+\mathbb{E}\left(\int_0^T\big(\mathscr{L}_5^{1,\varepsilon}-\bar{\mathscr{L}}_5^1\big)\tilde{u}_1^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right)=:\sum_{i=1}^3\mathscr{U}_i(\varepsilon).\end{aligned}$$

Note that the function $\big[\mathscr{L}_1(t,x,y)-\bar{\mathscr{L}}_1(t,y)\big]\tilde{u}_1^0(t,y,z)$ satisfies the centering condition (5.3) and belongs to $C_p^{\vartheta/2,\delta,\vartheta,2}$. As a direct result of the estimate (5.6), we have

$$\mathscr{U}_1(\varepsilon)\leqslant C_T\Big(\alpha_\varepsilon^\vartheta+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon}\Big).$$

Similarly, by definition,

$$\begin{aligned}&\mathscr{L}_3\tilde{\Upsilon}_1^0(t,x,y,z)-\bar{\mathscr{L}}_3^0\tilde{u}_1^0(t,y,z)\\&=\big[c(x,y)\nabla_x\Upsilon(t,x,y)-\overline{c\cdot\nabla_x\Upsilon}(t,y)\big]\nabla_z\tilde{u}_1^0(t,y,z),\end{aligned}$$

which satisfies (5.3) and belongs to $C_p^{\vartheta/2,\delta,\vartheta,2}$ due to the assumption that $c\in C_p^{\delta,\vartheta}$. Consequently, we also have

$$\mathscr{U}_2(\varepsilon)\leqslant C_T\Big(\alpha_\varepsilon^\vartheta+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon}\Big).$$

Finally, by the fact that $\bar{F}(t,\cdot),G(t,\cdot)\in C_b^{1+\vartheta}$ and the mean value theorem, we deduce that

$$\begin{aligned}\mathscr{U}_3(\varepsilon)\leqslant&\,\mathbb{E}\bigg(\int_0^T\bigg[\frac{\bar{F}_1(s,Y_s^{0,\varepsilon})-\bar{F}_1(s,\bar{Y}_s^1)}{\eta_\varepsilon}-\nabla_y\bar{F}_1(s,Y_s^{0,\varepsilon})Z_s^{0,\varepsilon}\bigg]\nabla_z\tilde{u}_1^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\bigg)\\&+\mathbb{E}\bigg(\int_0^T\bigg[\frac{[G(s,Y_s^{0,\varepsilon})-G(s,\bar{Y}_s^1)][G(s,Y_s^{0,\varepsilon})-G(s,\bar{Y}_s^1)]^*}{\eta_\varepsilon^2}\\&\qquad-\big[\nabla_yG(s,Y_s^{0,\varepsilon})Z_s^{0,\varepsilon}\big]\big[\nabla_yG(s,Y_s^{0,\varepsilon})Z_s^{0,\varepsilon}\big]^*\bigg]\nabla_z^2\tilde{u}_1^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\bigg)\\&+\mathbb{E}\bigg(\int_0^TG(s,Y_s^{0,\varepsilon})\bigg[\frac{[G(s,Y_s^{0,\varepsilon})-G(s,\bar{Y}_s^1)]}{\eta_\varepsilon}\\&\qquad-\nabla_yG(s,Y_s^{0,\varepsilon})Z_s^{0,\varepsilon}\bigg]\nabla_y\nabla_z\tilde{u}_1^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\bigg)\\\leqslant&\,C_T\mathbb{E}\left(\int_0^T\big|Y_s^{0,\varepsilon}-\bar{Y}_s^1\big|^\vartheta\big(1+|Z_s^{0,\varepsilon}|^2\big)\mathrm{d}s\right)\leqslant C_T\Big(\alpha_\varepsilon+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\Big)^\vartheta,\end{aligned}$$

where in the last inequality we also used Hölder's inequality and estimate (2.6). Based on the above estimates, we arrive at

$$\mathcal{R}_1^0(\varepsilon)\leqslant C_T\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon}+\frac{\beta_\varepsilon^2}{\alpha_\varepsilon^2}+\alpha_\varepsilon^\vartheta+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}+\frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta}\Big)\leqslant C_T\Big(\frac{\beta_\varepsilon^2}{\alpha_\varepsilon^2}+\frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta}\Big).$$

Case $\ell = 2$. (Regime 0.2 in (2.10)). Note that we have

$$\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2} = 1 \quad \text{and} \quad \frac{\alpha_\varepsilon^2}{\eta_\varepsilon \beta_\varepsilon} = \frac{\alpha_\varepsilon}{\beta_\varepsilon} \to 0 \quad \text{as} \quad \varepsilon \to 0.$$

Thus by the assumption that $c \in L_p^\infty$,

$$\mathscr{C}_2(\varepsilon) \leqslant C_T \frac{\alpha_\varepsilon^2}{\eta_\varepsilon \beta_\varepsilon} \mathbb{E}\left(\int_0^T \left(1 + |X_s^\varepsilon|^{2m}\right) \mathrm{d}s\right) \leqslant C_T \frac{\alpha_\varepsilon}{\beta_\varepsilon}.$$

Furthermore, recall that we have $\bar{\mathcal{L}}_2^0 = \bar{\mathscr{L}}_4^1 + \bar{\mathscr{L}}_5^1$ in this case with $\bar{\mathscr{L}}_4^1$ given by (2.25). It follows that

$$\partial_t \tilde{u}_2^0 + \left(\bar{\mathscr{L}}_1 + \bar{\mathscr{L}}_4^1 + \bar{\mathscr{L}}_5^1\right) \tilde{u}_2^0 = 0.$$

Consequently, we get

$$\begin{aligned}
\mathscr{C}_1(\varepsilon) + \mathscr{C}_3(\varepsilon) \leqslant\ & \mathbb{E}\left(\int_0^T \left(\mathscr{L}_1 - \bar{\mathscr{L}}_1\right) \tilde{u}_2^0(s, Y_s^{0,\varepsilon}, Z_s^{0,\varepsilon}) \mathrm{d}s\right) \\
& + \mathbb{E}\left(\int_0^T \mathscr{L}_4^1 \tilde{\Upsilon}_2^0(s, X_s^\varepsilon, Y_s^{0,\varepsilon}, Z_s^{0,\varepsilon}) - \bar{\mathscr{L}}_4^1 \tilde{u}_2^0(s, Y_s^{0,\varepsilon}, Z_s^{0,\varepsilon}) \mathrm{d}s\right) \\
& + \mathbb{E}\left(\int_0^T \left(\mathscr{L}_5^{1,\varepsilon} - \bar{\mathscr{L}}_5^1\right) \tilde{u}_2^0(s, Y_s^{0,\varepsilon}, Z_s^{0,\varepsilon}) \mathrm{d}s\right).
\end{aligned}$$

Note that by definition,

$$\begin{aligned}
& \mathscr{L}_4^1 \tilde{\Upsilon}_2^0(t, x, y, z) - \bar{\mathscr{L}}_4^1 \tilde{u}_2^0(t, y, z) \\
& = \left[\tilde{F}(t, x, y) \cdot \Upsilon^*(t, x, y) - \overline{\tilde{F} \cdot \Upsilon^*}(t, y)\right] \nabla_z^2 \tilde{u}_2^0(t, y, z).
\end{aligned}$$

Following exactly the same arguments as above, we get

$$\mathscr{C}_1(\varepsilon) + \mathscr{C}_3(\varepsilon) \leqslant C_T \left(\alpha_\varepsilon^\vartheta + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon} + \frac{\alpha_\varepsilon^{2^\vartheta}}{\beta_\varepsilon^\vartheta}\right) \leqslant C_T \left(\alpha_\varepsilon + \frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\right)^\vartheta.$$

Based on the above estimates, we arrive at

$$\mathcal{R}_2^0(\varepsilon) \leqslant C_T \left(\alpha_\varepsilon^\vartheta + \frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta} + \frac{\alpha_\varepsilon}{\beta_\varepsilon}\right) \leqslant C_T \left(\alpha_\varepsilon^\vartheta + \frac{\alpha_\varepsilon}{\beta_\varepsilon}\right).$$

Case $\ell = 3$. (Regime 0.3 in (2.10)). In this case, we have

$$\partial_t \tilde{u}_3^0 + \left(\bar{\mathscr{L}}_1 + \bar{\mathscr{L}}_5^1 + \bar{\mathscr{L}}_3^0 + \bar{\mathscr{L}}_4^1\right) \tilde{u}_3^0 = 0.$$

Thus we have

$$\begin{aligned}
\sum_{i=1}^3 \mathscr{C}_i(\varepsilon) \leqslant\ & \mathbb{E}\left(\int_0^T \left(\mathscr{L}_1 - \bar{\mathscr{L}}_1\right) \tilde{u}_3^0(s, Y_s^{0,\varepsilon}, Z_s^{0,\varepsilon}) \mathrm{d}s\right) \\
& + \mathbb{E}\left(\int_0^T \mathscr{L}_3 \tilde{\Upsilon}_3^0(s, X_s^\varepsilon, Y_s^{0,\varepsilon}, Z_s^{0,\varepsilon}) - \bar{\mathscr{L}}_3^0 \tilde{u}_3^0(s, Y_s^\varepsilon, Z_s^{0,\varepsilon}) \mathrm{d}s\right) \\
& + \mathbb{E}\left(\int_0^T \mathscr{L}_4^1 \tilde{\Upsilon}_3^0(s, X_s^\varepsilon, Y_s^{0,\varepsilon}, Z_s^{0,\varepsilon}) - \bar{\mathscr{L}}_4^1 \tilde{u}_3^0(s, Y_s^\varepsilon, Z_s^{0,\varepsilon}) \mathrm{d}s\right)
\end{aligned}$$

$$
+\mathbb{E}\left(\int_0^T (\mathscr{L}_5^{1,\varepsilon}-\bar{\mathscr{L}}_5^1)\tilde{u}_3^0(s,Y_s^{0,\varepsilon},Z_s^{0,\varepsilon})\mathrm{d}s\right).
$$

By combining the above two cases we deduce that

$$
\mathcal{R}_3^0(\varepsilon)\leqslant C_T\Big(\alpha_\varepsilon^\vartheta+\frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon}\Big)\leqslant C_T\alpha_\varepsilon^\vartheta,
$$

and the whole proof is finished. □

6. Central limit theorem with homogenization

In this section, we study the functional central limit theorem for SDE (1.6) by following the same procedure as in Section 5. We first derive some weak type fluctuation estimates in Subsection 6.1. Then we prove Theorem 2.5 and Theorem 2.7 in Subsection 6.2 and Subsection 6.3, respectively. We shall mainly focus on the differences in regard to the proof of Theorem 2.3.

6.1. **Weak fluctuation estimate (ii).** Recall that depending on Regime 1 and Regime 2 described in (1.9), the slow process Y_t^ε converges strongly to $\bar{Y}_t^1$ and $\bar{Y}_t^2$, respectively. As before, to prove the weak convergence of $Z_t^{k,\varepsilon}$ to $\bar{Z}_{\ell,t}^k$ ($k=1,2$ and $\ell=1,2,3$), we will view the process $(X_t^\varepsilon,Y_t^\varepsilon,\bar{Y}_t^k,Z_t^{k,\varepsilon},\bar{Z}_{\ell,t}^k)$ as a whole system, i.e., we consider

$$
\begin{cases}
\mathrm{d}X_t^\varepsilon=\alpha_\varepsilon^{-2}b(X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}t+\beta_\varepsilon^{-1}c(X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}t+\alpha_\varepsilon^{-1}\sigma(X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}W_t^1, & X_0=x,\\
\mathrm{d}Y_t^\varepsilon=F(t,X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}t+\gamma_\varepsilon^{-1}H(t,X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}t+G(t,Y_t^\varepsilon)\mathrm{d}W_t^2, & Y_0=y,\\
\mathrm{d}\bar{Y}_t^k=\bar{F}_k(t,\bar{Y}_t^k)\mathrm{d}t+G(t,\bar{Y}_t^k)\mathrm{d}W_t^2, & \bar{Y}_0^k=y,\\
\mathrm{d}Z_t^{k,\varepsilon}=\dfrac{1}{\eta_\varepsilon}\Big[F(t,X_t^\varepsilon,Y_t^\varepsilon)-\bar{F}_k(t,\bar{Y}_t^k)\Big]\mathrm{d}t+\dfrac{1}{\eta_\varepsilon\gamma_\varepsilon}H(t,X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}t & \\
\qquad\qquad+\dfrac{1}{\eta_\varepsilon}\Big[G(t,Y_t^\varepsilon)-G(t,\bar{Y}_t^k)\Big]\mathrm{d}W_t^2, & Z_0^{k,\varepsilon}=0,\\
\mathrm{d}\bar{Z}_{\ell,t}^k=\varrho_\ell^k(t,\bar{Y}_t^k,\bar{Z}_{\ell,t}^k)\mathrm{d}t+\chi_\ell^k(t,\bar{Y}_t^k,\bar{Z}_{\ell,t}^k)\mathrm{d}W_t^2 & \bar{Z}_{\ell,0}^k=0,
\end{cases}
\tag{6.1}
$$

where for $k=1,2$ and $\ell=1,2,3$, η_ε are given by (2.14) and (2.18), ϱ_ℓ^k and χ_ℓ^k denote the drift and diffusion coefficients for $\bar{Z}_{\ell,t}^1$ in SDE (2.17) and $\bar{Z}_{\ell,t}^2$ in SDE (2.20), respectively.

Note that by definition,

$$
\begin{aligned}
\mathrm{d}Z_t^{1,\varepsilon}&=\frac{1}{\eta_\varepsilon}\Big[F(t,X_t^\varepsilon,Y_t^\varepsilon)-\bar{F}_1(t,Y_t^\varepsilon)\Big]\mathrm{d}t+\frac{1}{\eta_\varepsilon\gamma_\varepsilon}H(t,X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}t\\
&\quad+\left(\frac{1}{\eta_\varepsilon}\Big[\bar{F}_1(t,Y_t^\varepsilon)-\bar{F}_1(t,\bar{Y}_t^1)\Big]\mathrm{d}t+\frac{1}{\eta_\varepsilon}\Big[G(t,Y_t^\varepsilon)-G(t,\bar{Y}_t^1)\Big]\mathrm{d}W_t^2\right),
\end{aligned}
$$

and since $\bar{F}_2(t,y)=\bar{F}_1(t,y)+\overline{c\cdot\nabla_x\Phi}(t,y)$, we have

$$
\begin{aligned}
\mathrm{d}Z_t^{2,\varepsilon}&=\frac{1}{\eta_\varepsilon}\Big[F(t,X_t^\varepsilon,Y_t^\varepsilon)-\bar{F}_1(t,Y_t^\varepsilon)\Big]\mathrm{d}t\\
&\quad+\left(\frac{1}{\eta_\varepsilon\gamma_\varepsilon}H(t,X_t^\varepsilon,Y_t^\varepsilon)\mathrm{d}t-\frac{1}{\eta_\varepsilon}\overline{c\cdot\nabla_x\Phi}(t,Y_t^\varepsilon)\mathrm{d}t\right)
\end{aligned}
$$

$$
+\left(\frac{1}{\eta_\varepsilon}\Big[\bar{F}_2(t,Y_t^\varepsilon)-\bar{F}_2(t,\bar{Y}_t^2)\Big]\mathrm{d}t+\frac{1}{\eta_\varepsilon}\Big[G(t,Y_t^\varepsilon)-G(t,\bar{Y}_t^2)\Big]\mathrm{d}W_t^2\right).
$$

Recall that $\bar{\mathscr{L}}_3^1$, $\mathscr{L}_4^1$ and $\mathscr{L}_5^{k,\varepsilon}$ are defined by (2.23), (5.1) and (5.2), respectively. We define

$$
\mathscr{L}_4^2:=\mathscr{L}_4^2(t,x,y,z):=\sum_{i=1}^{d_2}H^i(t,x,y)\frac{\partial}{\partial z_i}. \tag{6.2}
$$

Given a function $f(t,x,y,z)$ satisfying the centering condition (5.3), let $\Phi^f(t,x,y,z)$ be the solution to the Poisson equation (5.4). We have the following estimate for the fluctuations of $f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{k,\varepsilon})$ over $[0,t]$.

Lemma 6.1. *Let* $(\mathbf{A}_\sigma)$, $(\mathbf{A}_b)$, *(2.1) hold and* $0<\delta,\vartheta\leqslant 2$. *Assume that*
(i) (Regime 1) $b,\sigma\in C_b^{\delta,1\vee\vartheta}$, $F\in C_p^{0,\delta,1}$, $G\in C_b^{0,1}$ *and* $c,H\in L_p^\infty$;
(ii) (Regime 2) $b,\sigma\in C_b^{\delta,1\vee\vartheta}$, $c\in C_p^{0,1}$, $F,H\in C_p^{0,\delta,1}$ *and* $G\in C_b^{0,1}$.
Then for every $f\in C_p^{\vartheta/2,\delta,\vartheta,2}$ *satisfying (5.3), we have for* $k=1,2$,

$$
\begin{aligned}
\mathbb{E}\left(\int_0^t f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{k,\varepsilon})\mathrm{d}s\right)\leqslant C_t\Big(\alpha_\varepsilon^\vartheta+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon}+\frac{\alpha_\varepsilon^{1+(\vartheta\wedge 1)}}{\gamma_\varepsilon}\Big)&\\
+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\mathbb{E}\left(\int_0^T\mathscr{L}_3\Phi^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{k,\varepsilon})\mathrm{d}s\right)&\\
+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon}\mathbb{E}\left(\int_0^T\mathscr{L}_4^2\Phi^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{k,\varepsilon})\mathrm{d}s\right)&.
\end{aligned} \tag{6.3}
$$

where $C_t>0$ *is a constant independent of* δ,ε.

Remark 6.2. *Compared with Lemma 5.1, the term involving* $\mathscr{L}_4^1\Phi^f$ *is replaced by* $\mathscr{L}_4^2\Phi^f$, *since it is of lower order now. As before, we get*

$$
\mathbb{E}\left(\int_0^T\mathscr{L}_3\Phi^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{k,\varepsilon})\mathrm{d}s\right)+\mathbb{E}\left(\int_0^T\mathscr{L}_4^2\Phi^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{k,\varepsilon})\mathrm{d}s\right)<\infty.
$$

Thus we also have

$$
\mathbb{E}\left(\int_0^T f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{k,\varepsilon})\mathrm{d}s\right)\leqslant C_T\Big(\alpha_\varepsilon^\vartheta+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}+\frac{\alpha_\varepsilon^{1+(\vartheta\wedge 1)}}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon}\Big). \tag{6.4}
$$

The homogenization effects of the last two terms on the right hand side of (6.3) will appear in the study of the functional central limit theorems.

Proof. The proof follows by the same arguments as in the proof of Lemma 5.1. We provide some details here in order to make clear which parts should be the leading terms. We only prove (6.3) for $k=2$ (Regime 2), the case $k=1$ (Regime 1) can be proved similarly and is even easier since the operator $\bar{\mathscr{L}}_3^1$ is not involved. Let Φ_n^f be the mollifyer of Φ^f defined as in (3.5). Using Itô's formula for SDE (6.1), we have

$$
\Phi_n^f(t,X_t^\varepsilon,Y_t^\varepsilon,Z_t^{2,\varepsilon})=\Phi_n^f(0,x,y,0)+\frac{1}{\alpha_\varepsilon}\hat{M}_n^1(t)+\hat{M}_n^2(t)+\frac{1}{\eta_\varepsilon}\hat{M}_n^3(t)
$$

$$
\begin{aligned}
&+\int_0^t \big(\partial_s+\mathscr{L}_1+\mathscr{L}_5^{2,\varepsilon}\big)\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\\
&+\frac{1}{\alpha_\varepsilon^2}\int_0^t \mathscr{L}_0\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s+\frac{1}{\beta_\varepsilon}\int_0^t \mathscr{L}_3\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\\
&+\frac{1}{\gamma_\varepsilon}\int_0^t \mathscr{L}_2\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s+\frac{1}{\eta_\varepsilon}\int_0^t \mathscr{L}_4^1\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\\
&+\frac{1}{\eta_\varepsilon\gamma_\varepsilon}\int_0^t \mathscr{L}_4^2\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s-\frac{1}{\eta_\varepsilon}\int_0^t \bar{\mathscr{L}}_3^1\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s,
\end{aligned}
$$

where $\mathscr{L}_4^2$ is given by (6.2), $\hat{M}_n^1(t)$, $\hat{M}_n^2(t)$ and $\hat{M}_n^3(t)$ are martingales defined by

$$
\begin{aligned}
\hat{M}_n^1(t)&:=\int_0^t \sigma(X_s^\varepsilon,Y_s^\varepsilon)\nabla_x\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}W_s^1,\\
\hat{M}_n^2(t)&:=\int_0^t G(s,Y_s^\varepsilon)\nabla_y\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}W_s^2,\\
\hat{M}_n^3(t)&:=\int_0^t \big[G(s,Y_s^\varepsilon)-G(s,\bar{Y}_s^2)\big]\nabla_z\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}W_s^2.
\end{aligned}
$$

Taking expectation and in view of (5.4), we get

$$
\begin{aligned}
\hat{\mathcal{V}}(\varepsilon)&:=\mathbb{E}\bigg(\int_0^t f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\bigg)=\alpha_\varepsilon^2\mathbb{E}\big[\Phi_n^f(0,x,y,0)-\Phi_n^f(t,X_t^\varepsilon,Y_t^\varepsilon,Z_t^{2,\varepsilon})\big]\\
&+\alpha_\varepsilon^2\mathbb{E}\left(\int_0^T \big(\partial_s+\mathscr{L}_1+\mathscr{L}_5^{2,\varepsilon}\big)\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&+\mathbb{E}\left(\int_0^T \big[\mathscr{L}_0\Phi_n^f-\mathscr{L}_0\Phi^f\big](s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&+\frac{\alpha_\varepsilon^2}{\gamma_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_2\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_4^1\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)-\frac{\alpha_\varepsilon^2}{\eta_\varepsilon}\mathbb{E}\left(\int_0^T \bar{\mathscr{L}}_3^1\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&+\bigg[\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\mathbb{E}\left(\int_0^T \big(\mathscr{L}_3\Phi_n^f-\mathscr{L}_3\Phi^f\big)(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&\quad+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon}\mathbb{E}\left(\int_0^T \big(\mathscr{L}_4^2\Phi_n^f-\mathscr{L}_4^2\Phi^f\big)(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\bigg]\\
&+\bigg[\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_3\Phi^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&\quad+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_4^2\Phi^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\bigg]=:\sum_{i=1}^{8}\hat{\mathcal{V}}_i(\varepsilon).
\end{aligned}
$$

Now, due to assumptions that $b,\sigma\in C_b^{\delta,1\vee\vartheta}$, $c\in C_p^{0,1}$, $H\in C_p^{0,\delta,1}$ and by Lemma 3.3, we have $\bar{F}_2\in C_b^{0,1}$. Following the same argument as in (5.7), we can deduce that

$$\mathbb{E}\left(\int_0^T \mathscr{L}_5^{2,\varepsilon}\Phi_n^f(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\leqslant C_T n^{1-(\vartheta\wedge 1)}.$$

As a result, we further have

$$\sum_{i=1}^7 \mathcal{V}_i(\varepsilon)\leqslant C_T\Big(\alpha_\varepsilon^2 n^{2-\vartheta}+n^{-\vartheta}+\frac{\alpha_\varepsilon^2}{\gamma_\varepsilon}n^{1-(\vartheta\wedge 1)}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}n^{-\vartheta}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon}n^{-\vartheta}\Big).$$

Taking $n=\alpha_\varepsilon^{-1}$ and noticing that $\alpha_\varepsilon^2/(\eta_\varepsilon\gamma_\varepsilon)\to 0$ as $\varepsilon\to 0$, we get the desired result. $\square$

6.2. **Proof of Theorem 2.5.** We are now in the position to give:

Proof of Theorem 2.5. We concentrate on the main differences in regard to the proof of Theorem 2.3. For every $\varphi\in C_b^4(\mathbb{R}^{d_2})$ and $\ell=1,2,3$, let u_ℓ^1 be the solution to the Cauchy problem (3.10) on $[0,T]\times\mathbb{R}^{d_2}\times\mathbb{R}^{d_2}$ with $k=1$, and define

$$\tilde{u}_\ell^1(t,y,z):=u_\ell^1(T-t,y,z),\quad t\in[0,T].$$

Then we write for $\ell=1,2,3$,

$$\mathcal{R}_\ell^1(\varepsilon):=\mathbb{E}\varphi(Z_T^{1,\varepsilon})-\mathbb{E}\varphi(\bar{Z}_{\ell,T}^1)=\mathbb{E}\tilde{u}_\ell^1(T,Y_T^\varepsilon,Z_T^{1,\varepsilon})-\tilde{u}_\ell^1(0,y,0).$$

By Theorem 3.4 and using Itô's formula for SDE (6.1), we deduce that

$$\begin{aligned}\tilde{u}_\ell^1(T,Y_T^\varepsilon,Z_T^{1,\varepsilon})&=\tilde{u}_\ell^1(0,y,0)+\int_0^T\Big(\partial_s+\mathscr{L}_1+\mathscr{L}_5^{1,\varepsilon}\Big)\tilde{u}_\ell^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s\\&\quad+\frac{1}{\gamma_\varepsilon}\int_0^T\mathscr{L}_2\tilde{u}_\ell^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s+\frac{1}{\eta_\varepsilon}\int_0^T\mathscr{L}_4^1\tilde{u}_\ell^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s\\&\quad+\frac{1}{\eta_\varepsilon\gamma_\varepsilon}\int_0^T\mathscr{L}_4^2\tilde{u}_\ell^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s+\hat{M}_{\ell,T}^1+\frac{1}{\eta_\varepsilon}\hat{M}_{\ell,T}^2,\end{aligned}$$

where $\hat{M}_{\ell,t}^1$ and $\hat{M}_{\ell,t}^2$ are martingales defined by

$$\begin{aligned}\hat{M}_{\ell,t}^1&:=\int_0^t G(s,Y_s^\varepsilon)\nabla_y\tilde{u}_\ell^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}W_s^2,\\\hat{M}_{\ell,t}^2&:=\int_0^t\big[G(s,Y_s^\varepsilon)-G(s,\bar{Y}_s^1)\big]\nabla_z\tilde{u}_\ell^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}W_s^2.\end{aligned}$$

As a result, we further have

$$\begin{aligned}\mathcal{R}_\ell^1(\varepsilon)&\leqslant\mathbb{E}\left(\int_0^T\Big[\partial_s+\mathscr{L}_1+\mathscr{L}_5^{1,\varepsilon}\Big]\tilde{u}_\ell^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s\right)\\&\quad+\frac{1}{\gamma_\varepsilon}\mathbb{E}\left(\int_0^T\mathscr{L}_2\tilde{u}_\ell^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s\right)+\frac{1}{\eta_\varepsilon}\mathbb{E}\left(\int_0^T\mathscr{L}_4^1\tilde{u}_\ell^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s\right)\\&\quad+\frac{1}{\eta_\varepsilon\gamma_\varepsilon}\mathbb{E}\left(\int_0^T\mathscr{L}_4^2\tilde{u}_\ell^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s\right)=:\sum_{i=1}^4\mathscr{G}_i(\varepsilon).\end{aligned}$$

Note that by definition,

$$\mathscr{G}_2(\varepsilon) = \frac{1}{\gamma_\varepsilon}\mathbb{E}\left(\int_0^T H(s, X_s^\varepsilon, Y_s^\varepsilon)\nabla_y \tilde{u}_\ell^1(s, Y_s^\varepsilon, Z_s^{1,\varepsilon})\mathrm{d}s\right).$$

Thanks to the assumption $(\mathbf{A}_H)$, it is easily checked that for every $\ell = 1, 2, 3$, the functions $H(t, x, y)\nabla_y \tilde{u}_\ell^1(t, y, z)$ satisfy the centering condition (5.3) and belong to $C_p^{(1+\vartheta)/2,\delta,1+\vartheta,2}$. As a result of estimate (6.4), we get

$$\mathscr{G}_2(\varepsilon) \leqslant C_1\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\gamma_\varepsilon\beta_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon^2}\Big).$$

Similarly, we also have

$$\mathscr{G}_3(\varepsilon) \leqslant C_2\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon}\Big).$$

Concerning the last term, note that

$$\mathscr{G}_4(\varepsilon) = \frac{1}{\eta_\varepsilon\gamma_\varepsilon}\mathbb{E}\left(\int_0^T H(s, X_s^\varepsilon, Y_s^\varepsilon)\nabla_z \tilde{u}_\ell^1(s, Y_s^\varepsilon, Z_s^{1,\varepsilon})\mathrm{d}s\right).$$

For every $\ell = 1, 2, 3$, the functions $H(t, x, y)\nabla_z \tilde{u}_\ell^1(t, y, z)$ satisfy the centering condition (5.3). Recall that $\Phi(t, x, y)$ solves the Poisson equation (2.3), and define

$$\hat{\Phi}_\ell^1(t, x, y, z) := \Phi(t, x, y)\nabla_z \tilde{u}_\ell^1(t, y, z).$$

Then $\hat{\Phi}_\ell^1(t, x, y, z)$ satisfies

$$\mathscr{L}_0(x, y)\hat{\Phi}_\ell^1(t, x, y, z) = -H(t, x, y)\nabla_z \tilde{u}_\ell^1(t, y, z).$$

Consequently, we use Lemma 6.1 to deduce that

$$\begin{aligned}\mathscr{G}_4(\varepsilon) \leqslant C_3\Big(&\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon\gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon^2}\Big)\\ &+ \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon\beta_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_3\hat{\Phi}_\ell^1(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{1,\varepsilon})\mathrm{d}s\right)\\ &+ \frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon^2}\mathbb{E}\left(\int_0^T \mathscr{L}_4^2\hat{\Phi}_\ell^1(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{1,\varepsilon})\mathrm{d}s\right).\end{aligned}$$

Combining the above estimates, we arrive at

$$\begin{aligned}\mathcal{R}_\ell^1(\varepsilon) \leqslant C_4\Big(&\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon\gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\gamma_\varepsilon\beta_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon^2}\Big)\\ &+ \mathbb{E}\left(\int_0^T \Big[\partial_s + \mathscr{L}_1 + \mathscr{L}_5^{1,\varepsilon}\Big]\tilde{u}_\ell^1(s, Y_s^\varepsilon, Z_s^{1,\varepsilon})\mathrm{d}s\right)\\ &+ \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon\beta_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_3\hat{\Phi}_\ell^1(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{1,\varepsilon})\mathrm{d}s\right)\\ &+ \frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon^2}\mathbb{E}\left(\int_0^T \mathscr{L}_4^2\hat{\Phi}_\ell^1(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{1,\varepsilon})\mathrm{d}s\right)\end{aligned}$$

$$=: C_4\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\gamma_\varepsilon\beta_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon^2}\Big)+\sum_{i=1}^{3}\mathscr{V}_i(\varepsilon). \tag{6.5}$$

Below we consider $\ell=1,2,3$ separately, which correspond to Regime 1.1-Regime 1.3 in the conclusion of Theorem 2.5, respectively.

Case $\ell=1$. (Regime 1.1 in (2.14)). Note that we have

$$\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon\beta_\varepsilon}=1 \quad\text{and}\quad \frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon^2}=\frac{\beta_\varepsilon^2}{\alpha_\varepsilon^2}\to 0 \quad\text{as}\quad \varepsilon\to 0.$$

Thus we deduce that

$$\mathscr{V}_3(\varepsilon)\leqslant C_T\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon^2}\mathbb{E}\left(\int_0^T\big(1+|X_s^\varepsilon|^{2m}\big)\mathrm{d}s\right)\leqslant C_T\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon^2}$$

Furthermore, recall that we have $\bar{\mathcal{L}}_1^1=\bar{\mathscr{L}}_5^1+\bar{\mathscr{L}}_3^1$ in this case with $\bar{\mathscr{L}}_3^1$ given by (2.23). As a result,

$$\partial_s\tilde{u}_1^1+\big(\bar{\mathscr{L}}_1+\bar{\mathscr{L}}_5^1+\bar{\mathscr{L}}_3^1\big)\tilde{u}_1^1=0.$$

This in turn yields that

$$\begin{aligned}\mathscr{V}_1(\varepsilon)+\mathscr{V}_2(\varepsilon)\leqslant{}&\mathbb{E}\left(\int_0^T\big(\mathscr{L}_1-\bar{\mathscr{L}}_1\big)\tilde{u}_1^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s\right)\\&+\mathbb{E}\left(\int_0^T\mathscr{L}_3\hat{\Phi}_1^1(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{1,\varepsilon})-\bar{\mathscr{L}}_3^1\tilde{u}_1^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s\right)\\&+\mathbb{E}\left(\int_0^T\big(\mathscr{L}_5^{1,\varepsilon}-\bar{\mathscr{L}}_5^1\big)\tilde{u}_1^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s\right).\end{aligned}$$

By definition,

$$\begin{aligned}&\mathscr{L}_3\hat{\Phi}_1^1(t,x,y,z)-\bar{\mathscr{L}}_3^1\tilde{u}_1^1(t,y,z)\\&=\big[c(x,y)\cdot\nabla_x\Phi(t,x,y)-\overline{c\cdot\nabla_x\Phi}(t,y)\big]\cdot\nabla_z\tilde{u}_1^1(t,y,z),\end{aligned}$$

which satisfies the centering condition (5.3). Using the assumption that $c\in C_p^{\delta,\vartheta}$ and exactly the same arguments as in the proof of Theorem 2.3 (Case $\ell=1$), we get

$$\begin{aligned}\mathscr{V}_1(\varepsilon)+\mathscr{V}_2(\varepsilon)&\leqslant C_T\Big(\alpha_\varepsilon^\vartheta+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}+\frac{\alpha_\varepsilon^{1+\vartheta}}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon}\Big)+C_T\Big(\frac{\alpha_\varepsilon}{\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon\gamma_\varepsilon}\Big)^\vartheta\\&\leqslant C_T\Big(\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon}+\frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta\gamma_\varepsilon^\vartheta}\Big).\end{aligned}$$

Combining the above computations with (6.5), we arrive at

$$\begin{aligned}\mathcal{R}_1^1(\varepsilon)&\leqslant C_T\Big(\frac{\alpha_\varepsilon^2}{\gamma_\varepsilon\beta_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon}+\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon^2}+\frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta\gamma_\varepsilon^\vartheta}\Big)\\&\leqslant C_T\Big(\gamma_\varepsilon+\frac{\beta_\varepsilon^2}{\alpha_\varepsilon^2}+\frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta\gamma_\varepsilon^\vartheta}\Big).\end{aligned}$$

Case $\ell = 2$. (Regime 1.2 in (2.14)). Note that we have

$$\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2 \gamma_\varepsilon^2} = 1 \quad \text{and} \quad \frac{\alpha_\varepsilon^2}{\eta_\varepsilon \gamma_\varepsilon \beta_\varepsilon} = \frac{\alpha_\varepsilon}{\beta_\varepsilon} \to 0 \quad \text{as} \quad \varepsilon \to 0.$$

Thus we deduce that

$$\mathscr{V}_2(\varepsilon) \leqslant C_T \frac{\alpha_\varepsilon^2}{\eta_\varepsilon \gamma_\varepsilon \beta_\varepsilon} \mathbb{E}\left(\int_0^T \left(1 + |X_s^\varepsilon|^{2m}\right) \mathrm{d}s \right) \leqslant C_T \frac{\alpha_\varepsilon}{\beta_\varepsilon}.$$

Furthermore, recall that we have $\bar{\mathcal{L}}_2^1 = \bar{\mathscr{L}}_5^1 + \bar{\mathscr{L}}_4^2$ in this case with $\bar{\mathscr{L}}_4^2$ given by (2.26). It then follows that

$$\partial_s \tilde{u}_2^1 + \left(\bar{\mathscr{L}}_1 + \bar{\mathscr{L}}_5^1 + \bar{\mathscr{L}}_4^2\right) \tilde{u}_2^1 = 0.$$

We thus have

$$\begin{aligned}
\mathscr{V}_1(\varepsilon) + \mathscr{V}_3(\varepsilon) \leqslant\ & \mathbb{E}\left(\int_0^T \left(\mathscr{L}_1 - \bar{\mathscr{L}}_1\right) \tilde{u}_2^1(s, Y_s^\varepsilon, Z_s^{1,\varepsilon}) \mathrm{d}s \right) \\
& + \mathbb{E}\left(\int_0^T \mathscr{L}_4^2 \hat{\Phi}_2^1(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{1,\varepsilon}) - \bar{\mathscr{L}}_4^2 \tilde{u}_2^1(s, Y_s^\varepsilon, Z_s^{1,\varepsilon}) \mathrm{d}s \right) \\
& + \mathbb{E}\left(\int_0^T \left(\mathscr{L}_5^{1,\varepsilon} - \bar{\mathscr{L}}_5^1\right) \tilde{u}_2^1(s, Y_s^\varepsilon, Z_s^{1,\varepsilon}) \mathrm{d}s \right).
\end{aligned}$$

By definition,

$$\begin{aligned}
& \mathscr{L}_4^2 \hat{\Phi}_2^1(t, x, y, z) - \bar{\mathscr{L}}_4^2 \tilde{u}_2^1(t, y, z) \\
& = \left[H(t, x, y) \cdot \Phi^*(t, x, y) - \overline{H \cdot \Phi^*}(t, y)\right] \cdot \nabla_z^2 \tilde{u}_2^1(t, y, z),
\end{aligned}$$

which satisfies the centering condition (5.3). We can employ the same argument used before to deduce that

$$\begin{aligned}
\mathscr{V}_1(\varepsilon) + \mathscr{V}_3(\varepsilon) & \leqslant C_T \Big(\alpha_\varepsilon^\vartheta + \frac{\alpha_\varepsilon^2}{\beta_\varepsilon} + \frac{\alpha_\varepsilon^{1+\vartheta}}{\gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon \gamma_\varepsilon} \Big) + C_T \Big(\frac{\alpha_\varepsilon}{\gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\beta_\varepsilon \gamma_\varepsilon} \Big)^\vartheta \\
& \leqslant C_T \Big(\frac{\alpha_\varepsilon^2}{\beta_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon \gamma_\varepsilon} + \frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon^\vartheta} \Big).
\end{aligned}$$

Combining the above computations with (6.5), we arrive at

$$\begin{aligned}
\mathcal{R}_2^1(\varepsilon) & \leqslant C_T \Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon \gamma_\varepsilon} + \frac{\alpha_\varepsilon}{\beta_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2 \gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon \gamma_\varepsilon^2} + \frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon^\vartheta} \Big) \\
& \leqslant C_T \Big(\gamma_\varepsilon + \frac{\alpha_\varepsilon}{\beta_\varepsilon} + \frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon^\vartheta} \Big).
\end{aligned}$$

Case $\ell = 3$. (Regime 1.3 in (2.14)). In this case, we have

$$\partial_s \tilde{u}_2^1 + \left(\bar{\mathscr{L}}_1 + \bar{\mathscr{L}}_5^1 + \bar{\mathscr{L}}_3^1 + \bar{\mathscr{L}}_4^2\right) \tilde{u}_3^1 = 0.$$

Thus we can write

$$\sum_{i=1}^3 \mathscr{V}_i(\varepsilon) \leqslant \mathbb{E}\left(\int_0^T \left(\mathscr{L}_1 - \bar{\mathscr{L}}_1\right) \tilde{u}_3^1(s, Y_s^\varepsilon, Z_s^{1,\varepsilon}) \mathrm{d}s \right)$$

$$
\begin{aligned}
&+\mathbb{E}\left(\int_0^T \mathscr{L}_3\hat{\Phi}_3^1(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{1,\varepsilon})-\bar{\mathscr{L}}_3^1\tilde{u}_3^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s\right)\\
&+\mathbb{E}\left(\int_0^T \mathscr{L}_4^2\hat{\Phi}_3^1(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{1,\varepsilon})-\bar{\mathscr{L}}_4^2\tilde{u}_3^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s\right)\\
&+\mathbb{E}\left(\int_0^T \left(\mathscr{L}_5^{1,\varepsilon}-\bar{\mathscr{L}}_5^1\right)\tilde{u}_3^1(s,Y_s^\varepsilon,Z_s^{1,\varepsilon})\mathrm{d}s\right).
\end{aligned}
$$

By combining the above two cases we have

$$
\mathcal{R}_3^1(\varepsilon)\leqslant C_T\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon\gamma_\varepsilon}+\frac{\alpha_\varepsilon}{\eta_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon^2}+\frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon^\vartheta}\Big)\leqslant C_T\Big(\gamma_\varepsilon+\frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon^\vartheta}\Big).
$$

Hence the whole proof is finished. □

6.3. **Proof of Theorem 2.7.** The central limit theorems in Regime 2 of (1.9) will be the most complicated cases since homogenization already appears in the law of large numbers. Now, we proceed to give:

Proof of Theorem 2.7. For every $\varphi\in C_b^4(\mathbb{R}^{d_2})$ and $\ell=1,2,3$, let u_ℓ^2 be the solution to the Cauchy problem (3.10) on $[0,T]\times\mathbb{R}^{d_2}\times\mathbb{R}^{d_2}$ with $k=2$, and define

$$
\tilde{u}_\ell^2(t,y,z):=u_\ell^2(T-t,y,z),\quad t\in[0,T].
$$

Then following the same arguments as in the proof of Theorem 2.5, we have for $\ell=1,2,3$ that

$$
\mathcal{R}_\ell^2(\varepsilon):=\mathbb{E}\varphi(Z_T^{2,\varepsilon})-\mathbb{E}\varphi(\bar{Z}_{\ell,T}^2)=\mathbb{E}\tilde{u}_\ell^2(T,Y_T^\varepsilon,Z_T^{2,\varepsilon})-\tilde{u}_\ell^2(0,y,0).
$$

Using Itô's formula and taking expectation, we further get

$$
\begin{aligned}
\mathcal{R}_\ell^2(\varepsilon)\leqslant\ &\mathbb{E}\left(\int_0^T\Big[\partial_s+\mathscr{L}_1+\mathscr{L}_5^{2,\varepsilon}\Big]\tilde{u}_\ell^2(s,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&+\frac{1}{\gamma_\varepsilon}\mathbb{E}\left(\int_0^T\mathscr{L}_2\tilde{u}_\ell^2(s,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)+\frac{1}{\eta_\varepsilon}\mathbb{E}\left(\int_0^T\mathscr{L}_4^1\tilde{u}_\ell^2(s,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&+\left[\frac{1}{\eta_\varepsilon\gamma_\varepsilon}\mathbb{E}\left(\int_0^T\mathscr{L}_4^2\tilde{u}_\ell^2(s,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)-\frac{1}{\eta_\varepsilon}\mathbb{E}\left(\int_0^T\bar{\mathscr{L}}_3^1\tilde{u}_\ell^2(s,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\right]\\
=:&\sum_{i=1}^4\mathscr{N}_i(\varepsilon),
\end{aligned}
$$

where $\bar{\mathscr{L}}_3^1$, $\mathscr{L}_4^1$ and $\mathscr{L}_4^2$ are given by (2.23), (5.1) and (6.2), respectively. Recall that $\Phi(t,x,y)$ is the solution of the Poisson equation (2.3). Define

$$
\tilde{\Phi}_\ell^2(t,x,y,z):=\Phi(t,x,y)\nabla_y\tilde{u}_\ell^2(t,y,z).
$$

Then $\tilde{\Phi}_\ell^2(t,x,y,z)$ satisfies

$$
\mathscr{L}_0(x,y)\tilde{\Phi}_\ell^2(t,x,y,z)=-H(t,x,y)\nabla_y\tilde{u}_\ell^2(t,y,z).
$$

Note that for every $\ell = 1, 2, 3$, we have $H \cdot \nabla_y \tilde{u}_\ell^2 \in C_p^{(1+\vartheta)/2,\delta,1+\vartheta,2}$. As a result of Lemma 6.1 we obtain

$$\begin{aligned}
\mathscr{N}_2(\varepsilon) &= \frac{1}{\gamma_\varepsilon}\mathbb{E}\left(\int_0^T H(s, X_s^\varepsilon, Y_s^\varepsilon)\nabla_y \tilde{u}_\ell^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right) \\
&\leqslant C_T\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\gamma_\varepsilon\eta_\varepsilon} + \frac{\alpha_\varepsilon^2}{\gamma_\varepsilon^2}\Big) \\
&\quad + \frac{\alpha_\varepsilon^2}{\gamma_\varepsilon\beta_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_3\tilde{\Phi}_\ell^2(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right) \\
&\quad + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon^2}\mathbb{E}\left(\int_0^T \mathscr{L}_4^2\tilde{\Phi}_\ell^2(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right).
\end{aligned}$$

Since $\alpha_\varepsilon^2 = \beta_\varepsilon\gamma_\varepsilon$ in Regime 2, we further get

$$\mathscr{N}_2(\varepsilon) \leqslant C_T\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon^2}\Big) + \mathbb{E}\left(\int_0^T \mathscr{L}_3\tilde{\Phi}_\ell^2(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right).$$

Similarly, recall that $\Upsilon(t, x, y)$ solves the Poisson equation (2.11), and define

$$\tilde{\Upsilon}_\ell^2(t, x, y, z) := \Upsilon(t, x, y)\nabla_z \tilde{u}_\ell^2(t, y, z).$$

Then we have for $\ell = 1, 2, 3$,

$$\mathscr{L}_0(x, y)\tilde{\Upsilon}_\ell^2(t, x, y, z) = -\big[F(t, x, y) - \bar{F}_1(t, y)\big]\nabla_z \tilde{u}_\ell^2(t, y, z).$$

Consequently, we use Lemma 6.1 again to deduce that

$$\begin{aligned}
\mathscr{N}_3(\varepsilon) &= \frac{1}{\eta_\varepsilon}\mathbb{E}\left(\int_0^T \big[F(s, X_s^\varepsilon, Y_s^\varepsilon) - \bar{F}_1(s, Y_s^\varepsilon)\big]\nabla_z \tilde{u}_\ell^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right) \\
&\leqslant C_T\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon}\Big) + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_3\tilde{\Upsilon}_\ell^2(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right) \\
&\quad + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_4^2\tilde{\Phi}_\ell^2(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right) \\
&\leqslant C_T\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon}\Big) + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_3\tilde{\Upsilon}_\ell^2(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right).
\end{aligned}$$

Finally, define

$$\hat{\Phi}_\ell^2(t, x, y, z) := \Phi(t, x, y)\nabla_z \tilde{u}_\ell^2(t, y, z).$$

Then we have

$$\mathscr{L}_0(x, y)\hat{\Phi}_\ell^2(t, x, y, z) = -H(t, x, y)\nabla_z \tilde{u}_\ell^2(t, y, z),$$

which in turn yields that

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{\eta_\varepsilon\gamma_\varepsilon}\int_0^T H(s, X_s^\varepsilon, Y_s^\varepsilon)\nabla_z \tilde{u}_\ell^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right) \\
&\leqslant C_T\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon\gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon^2}\Big) + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon\beta_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_3\hat{\Phi}_\ell^2(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right)
\end{aligned}$$

$$+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon^2}\mathbb{E}\left(\int_0^T \mathscr{L}_4^2\hat{\Phi}_\ell^2(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right).$$

Taken into account the definition of $\mathscr{N}_4(\varepsilon)$ and the fact that $\alpha_\varepsilon^2=\beta_\varepsilon\gamma_\varepsilon$, we obtain

$$\mathscr{N}_4(\varepsilon)\leqslant C_T\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon^2}\Big)+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon^2}\mathbb{E}\left(\int_0^T \mathscr{L}_4^2\hat{\Phi}_\ell^2(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)$$
$$+\frac{1}{\eta_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_3\hat{\Phi}_\ell^2(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})-\bar{\mathscr{L}}_3^1\tilde{u}_\ell^2(s,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right).$$

Note that by definition,

$$\begin{aligned}&\mathscr{L}_3\hat{\Phi}_\ell^2(t,x,y,z)-\bar{\mathscr{L}}_3^1\tilde{u}_\ell^2(t,y,z)\\&=\big[c(x,y)\cdot\nabla_x\Phi(t,x,y)-\overline{c\cdot\nabla_x\Phi}(t,y)\big]\cdot\nabla_z\tilde{u}_\ell^2(t,y,z),\end{aligned}$$

and recall that Ψ solves the Poisson equation (2.19). Define

$$\tilde{\Psi}_\ell^2(t,x,y,z):=\Psi(t,x,y)\nabla_z\tilde{u}_\ell^2(t,y,z).$$

Then we have

$$\mathscr{L}_0(x,y)\tilde{\Psi}_\ell^2(t,x,y,z)=-\big[c(x,y)\cdot\nabla_x\Phi(t,x,y)-\overline{c\cdot\nabla_x\Phi}(t,y)\big]\cdot\nabla_z\tilde{u}_\ell^2(t,y,z).$$

By the assumption that $c\in C_p^{\delta,1+\vartheta}$ and Lemma 6.1, we further have

$$\begin{aligned}&\frac{1}{\eta_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_3\hat{\Phi}_\ell^2(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})-\bar{\mathscr{L}}_3^1\tilde{u}_\ell^2(s,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\&\leqslant C_T\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon}\Big)+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_3\tilde{\Psi}_\ell^2(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right).\end{aligned}$$

Combining the above computations, we arrive at

$$\begin{aligned}\mathcal{R}_\ell^2(\varepsilon)&\leqslant C_T\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon^2}\Big)+\mathbb{E}\left(\int_0^T\Big[\partial_s+\mathscr{L}_1+\mathscr{L}_5^{2,\varepsilon}\Big]\tilde{u}_\ell^2(s,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\&\quad+\mathbb{E}\left(\int_0^T \mathscr{L}_3\tilde{\Phi}_\ell^2(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon^2}\mathbb{E}\left(\int_0^T \mathscr{L}_4^2\hat{\Phi}_\ell^2(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\&\quad+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_3\tilde{\Upsilon}_\ell^2(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\&\quad+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon}\mathbb{E}\left(\int_0^T \mathscr{L}_3\tilde{\Psi}_\ell^2(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\&=:C_T\Big(\frac{\alpha_\varepsilon^{1+\vartheta}}{\eta_\varepsilon\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon^2}\Big)+\sum_{i=1}^5\mathscr{S}_i(\varepsilon).\end{aligned}$$

Below, we consider $\ell=1,2,3$ separately, which correspond to Regime 2.1-Regime 2.3 in the conclusion of Theorem 2.7, respectively.

Case $\ell=1$**.** (Regime 2.1 in (2.18)). Note that we have

$$\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon}=1\quad\text{and}\quad\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon^2}=\frac{\beta_\varepsilon^2}{\alpha_\varepsilon^2\gamma_\varepsilon^2}\to0\quad\text{as}\quad\varepsilon\to0.$$

Thus we deduce that

$$\mathscr{S}_3(\varepsilon) \leqslant C_T \frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2 \gamma_\varepsilon^2} \mathbb{E}\left(\int_0^T \left(1+|X_s^\varepsilon|^{2m}\right)\mathrm{d}s\right) \leqslant C_T \frac{\beta_\varepsilon^2}{\alpha_\varepsilon^2 \gamma_\varepsilon^2}.$$

Furthermore, recall that we have $\bar{\mathscr{L}_2} = \bar{\mathscr{L}_1} + \overline{c \cdot \nabla_x \Phi}(t,y) \cdot \nabla_y$ and $\bar{\mathcal{L}}_1^2 = \bar{\mathscr{L}_5^2} + \bar{\mathscr{L}_3^0} + \bar{\mathscr{L}_3^2}$ in this case with $\bar{\mathscr{L}_3^2}$ given by (2.24). As a result,

$$\partial_s \tilde{u}_1^2 + (\bar{\mathscr{L}_2} + \bar{\mathscr{L}_5^2} + \bar{\mathscr{L}_3^0} + \bar{\mathscr{L}_3^2})\tilde{u}_1^2 = 0.$$

This in turn yields that

$$\begin{aligned}
&\mathscr{S}_1(\varepsilon) + \mathscr{S}_2(\varepsilon) + \mathscr{S}_4(\varepsilon) + \mathscr{S}_5(\varepsilon) \\
&\leqslant \mathbb{E}\left(\int_0^T \left(\mathscr{L}_1 - \bar{\mathscr{L}_1}\right)\tilde{u}_1^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right) \\
&\quad + \mathbb{E}\left(\int_0^T \left(\mathscr{L}_5^{2,\varepsilon} - \bar{\mathscr{L}_5^2}\right)\tilde{u}_1^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right) \\
&\quad + \mathbb{E}\left(\int_0^T \mathscr{L}_3 \tilde{\Phi}_1^2(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{2,\varepsilon}) - \overline{c \cdot \nabla_x \Phi}(s, Y_s^\varepsilon) \cdot \nabla_y \tilde{u}_1^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right) \\
&\quad + \mathbb{E}\left(\int_0^T \mathscr{L}_3 \tilde{\Upsilon}_1^2(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{2,\varepsilon}) - \bar{\mathscr{L}_3^0}\tilde{u}_1^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right) \\
&\quad + \mathbb{E}\left(\int_0^T \mathscr{L}_3 \tilde{\Psi}_1^2(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{2,\varepsilon}) - \bar{\mathscr{L}_3^2}\tilde{u}_1^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right).
\end{aligned}$$

By definition, we have

$$\begin{aligned}
&\mathscr{L}_3 \tilde{\Phi}_1^2(t,x,y,z) - \overline{c \cdot \nabla_x \Phi}(t,y) \cdot \nabla_y \tilde{u}_1^2(t,y,z) \\
&= \left[c(x,y) \cdot \nabla_x \Phi(t,x,y) - \overline{c \cdot \nabla_x \Phi}(t,y)\right] \cdot \nabla_y \tilde{u}_1^2(t,y,z), \\
&\mathscr{L}_3 \tilde{\Upsilon}_1^2(t,x,y,z) - \bar{\mathscr{L}_3^0}\tilde{u}_1^2(t,y,z) \\
&= \left[c(x,y) \cdot \nabla_x \Upsilon(t,x,y) - \overline{c \cdot \nabla_x \Upsilon}(t,y)\right] \cdot \nabla_z \tilde{u}_1^2(t,y,z), \\
&\mathscr{L}_3 \tilde{\Psi}_1^2(t,x,y,z) - \bar{\mathscr{L}_3^2}\tilde{u}_1^2(t,y,z) \\
&= \left[c(x,y) \cdot \nabla_x \Psi(t,x,y) - \overline{c \cdot \nabla_x \Psi}(t,y)\right] \cdot \nabla_z \tilde{u}_1^2(t,y,z).
\end{aligned}$$

Using (2.7) and the same arguments as before, we deduce that

$$\begin{aligned}
\mathscr{S}_1(\varepsilon) + \mathscr{S}_2(\varepsilon) + \mathscr{S}_4(\varepsilon) + \mathscr{S}_5(\varepsilon) &\leqslant C_T\Big(\alpha_\varepsilon^\vartheta + \frac{\alpha_\varepsilon^2}{\beta_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon \gamma_\varepsilon} + \frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon^\vartheta} + \frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta}\Big) \\
&\leqslant C_T\Big(\alpha_\varepsilon^\vartheta + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon \gamma_\varepsilon} + \frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta}\Big).
\end{aligned}$$

Combining the above computations, we arrive at

$$\mathcal{R}_1^2(\varepsilon) \leqslant C_T\Big(\alpha_\varepsilon^\vartheta + \frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta} + \frac{\beta_\varepsilon^2}{\alpha_\varepsilon^2 \gamma_\varepsilon^2}\Big) \leqslant C_T\Big(\frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta} + \frac{\beta_\varepsilon^2}{\alpha_\varepsilon^2 \gamma_\varepsilon^2}\Big).$$

Case $\ell = 2$. (Regime 2.1 in (2.18)). Note that in this case, we have

$$\frac{\alpha_\varepsilon^2}{\eta_\varepsilon^2\gamma_\varepsilon^2} = 1 \quad \text{and} \quad \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon} = \frac{\alpha_\varepsilon\gamma_\varepsilon}{\beta_\varepsilon} \to 0 \quad \text{as} \quad \varepsilon \to 0..$$

Thus we deduce that

$$\mathscr{S}_4(\varepsilon) + \mathscr{S}_5(\varepsilon) \leqslant C_T \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\beta_\varepsilon}\mathbb{E}\left(\int_0^T \left(1 + |X_s^\varepsilon|^{2m}\right)\mathrm{d}s\right) \leqslant C_T \frac{\alpha_\varepsilon\gamma_\varepsilon}{\beta_\varepsilon}.$$

Furthermore, recall that we have $\bar{\mathcal{L}}_2^2 = \bar{\mathscr{L}}_5^2 + \bar{\mathscr{L}}_4^2$ in this case. As a result,

$$\partial_s \tilde{u}_2^2 + (\bar{\mathscr{L}}_2 + \bar{\mathscr{L}}_5^2 + \bar{\mathscr{L}}_4^2)\tilde{u}_2^2 = 0.$$

This in turn yields that

$$\begin{aligned}
&\mathscr{S}_1(\varepsilon) + \mathscr{S}_2(\varepsilon) + \mathscr{S}_3(\varepsilon)\\
&\leqslant \mathbb{E}\left(\int_0^T \left(\mathscr{L}_1 - \bar{\mathscr{L}}_1\right)\tilde{u}_2^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&\quad + \mathbb{E}\left(\int_0^T \left(\mathscr{L}_5^{2,\varepsilon} - \bar{\mathscr{L}}_5^2\right)\tilde{u}_2^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&\quad + \mathbb{E}\left(\int_0^T \mathscr{L}_3\tilde{\Phi}_2^2(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{2,\varepsilon}) - \overline{c\cdot\nabla_x\Phi}(s, Y_s^\varepsilon)\cdot\nabla_y\tilde{u}_2^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&\quad + \mathbb{E}\left(\int_0^T \mathscr{L}_4^2\hat{\Phi}_2^2(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^{2,\varepsilon}) - \bar{\mathscr{L}}_4^2\tilde{u}_2^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right).
\end{aligned}$$

By definition, we have

$$\begin{aligned}
&\mathscr{L}_4^2\hat{\Phi}_2^2(t, x, y, z) - \bar{\mathscr{L}}_4^2\tilde{u}_2^2(t, y, z)\\
&= \left[H(t, x, y)\cdot\Phi^*(t, x, y) - \overline{H\cdot\Phi^*}(t, y)\right]\cdot\nabla_z^2\tilde{u}_2^2(t, y, z).
\end{aligned}$$

Then we can get

$$\mathscr{S}_1(\varepsilon) + \mathscr{S}_2(\varepsilon) + \mathscr{S}_3(\varepsilon) \leqslant C_T\Big(\alpha_\varepsilon^\vartheta + \frac{\alpha_\varepsilon^2}{\beta_\varepsilon} + \frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon} + \frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon^\vartheta} + \frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta}\Big) \leqslant C_T\Big(\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon} + \frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon^\vartheta}\Big).$$

Combining the above computations, we arrive at

$$\mathcal{R}_2^2(\varepsilon) \leqslant C_T\Big(\frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon^\vartheta} + \frac{\alpha_\varepsilon\gamma_\varepsilon}{\beta_\varepsilon}\Big).$$

Case $\ell = 3$. (Regime 2.1 in (2.18)). In this case, we have

$$\partial_s \tilde{u}_3^2 + (\bar{\mathscr{L}}_2 + \bar{\mathscr{L}}_5^2 + \bar{\mathscr{L}}_3^0 + \bar{\mathscr{L}}_3^2 + \bar{\mathscr{L}}_4^2)\tilde{u}_3^2 = 0.$$

Thus we write

$$\begin{aligned}
\sum_{i=1}^5 \mathscr{S}_i(\varepsilon) \leqslant{}& \mathbb{E}\left(\int_0^T \left(\mathscr{L}_1 - \bar{\mathscr{L}}_1\right)\tilde{u}_3^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&+ \mathbb{E}\left(\int_0^T \left(\mathscr{L}_5^{2,\varepsilon} - \bar{\mathscr{L}}_5^2\right)\tilde{u}_3^2(s, Y_s^\varepsilon, Z_s^{2,\varepsilon})\mathrm{d}s\right)
\end{aligned}$$

$$
\begin{aligned}
&+\mathbb{E}\left(\int_0^T \mathscr{L}_3\tilde{\Phi}_3^2(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})-\overline{c\cdot\nabla_x\Phi}(s,Y_s^\varepsilon)\cdot\nabla_y\tilde{u}_3^2(s,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&+\mathbb{E}\left(\int_0^T \mathscr{L}_3\tilde{\Upsilon}_3^2(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})-\bar{\mathscr{L}}_3^0\tilde{u}_3^2(s,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&+\mathbb{E}\left(\int_0^T \mathscr{L}_3\tilde{\Psi}_3^2(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})-\bar{\mathscr{L}}_3^2\tilde{u}_3^2(s,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right)\\
&+\mathbb{E}\left(\int_0^T \mathscr{L}_4^2\hat{\Phi}_3^2(s,X_s^\varepsilon,Y_s^\varepsilon,Z_s^{2,\varepsilon})-\bar{\mathscr{L}}_4^2\tilde{u}_3^2(s,Y_s^\varepsilon,Z_s^{2,\varepsilon})\mathrm{d}s\right).
\end{aligned}
$$

By combining the above two cases we deduce that

$$
\mathcal{R}_3^2(\varepsilon)\leqslant C_T\Big(\alpha_\varepsilon^\vartheta+\frac{\alpha_\varepsilon^2}{\beta_\varepsilon}+\frac{\alpha_\varepsilon^2}{\eta_\varepsilon\gamma_\varepsilon}+\frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon^\vartheta}+\frac{\alpha_\varepsilon^{2\vartheta}}{\beta_\varepsilon^\vartheta}\Big)\leqslant C_T\frac{\alpha_\varepsilon^\vartheta}{\gamma_\varepsilon^\vartheta},
$$

and the whole proof is finished. □

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** (2004), 157–181.

[3] Bakhvalov N. S.: Averaging of partial differential equations with rapidly oscillating coefficients. *Sov. Math. Dokl.*, **16** (1975), 351–355.

[4] 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.

[5] Bouchet F., Grafke T., Tangarife T. and Vanden-Eijnden E.: Large deviations in fast-slow systems. *J. Statistical Physics*, **162** (2016), 793–812.

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

[7] Bréhier C. E.: Orders of convergence in the averaging principle for SPDEs: the case of a stochastically forced slow component. *Stoch. Proc. Appl.*, https://doi.org/10.1016/j.spa.2019.09.015.

[8] Budhiraja A., Dupuis P. and Ganguly A.: Large deviations for small noise diffusions in a fast Markovian environment. *Electron. J. Probab.*, **23** (2018), 1–33.

[9] Cattiaux P., Chafaï D. and Guillin A.: Central limit theorems for additive functionals of ergodic Markov diffusions processes. *ALEA, Lat. Am. J. Probab. Math. Stat.*, **9** (2012), 337–382.

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

[13] 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.

[14] Chevyrev I., Friz P., Korepanov A. and Melbourne I.: Superdiffusive limits for deterministic fast-slow dynamical systems. arXiv:1907.04825v1.

[15] De Simoi J. and Liverani C.: Limit theorems for fast-slow partially hyperbolic systems. *Invent. Math.*, **213** (2018), 811–1016.

[16] Dolgopyat D and Liverani C.: Energy transfer in a fast-slow Hamiltonian system. *Commu. Math. Phys.*, 308 (2011), 201–225.
[17] Dupuis P. and Spiliopoulos K.: Large deviations for multiscale problems via weak convergence methods. *Stoch. Proc. Appl.*, **122** (2012), 1947–1987.
[18] E W., Liu D. and Vanden-Eijnden E.: Analysis of multiscale methods for stochastic differential equations. *Comm. Pure Appl. Math.*, **58** (2005), 1544–1585.
[19] Fedrizzi E. and Flandoli F.: Pathwise uniqueness and continuous dependence for SDEs with non-regular drift. *Stochastics*, **83** (2011), 241–257.
[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] Flandoli F., Gubinelli M. and Priola E.: Remarks on the stochastic transport equation with Hölder drift. *Rend. Semin. Mat.*, **70** (2013), 53–73.
[22] Friedlin M.: Functional integration and partial differential equations. Princeton Univ. Press, Princeton, N.J., 1985.
[23] 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.
[24] Gonzales-Gargate I. I. and Ruffino P. R.: An averaging principle for diffusions in foliated spaces. *Ann. Probab.*, **44** (2016) 567–588.
[25] Guillin A.: Moderate deviations of inhomogeneous functionals of Markov processes and application to averaging. *Stoch. Proc. Appl.*, **92** (2001), 287–313.
[26] Guillin A.: Averaging principle of SDE with small diffusion: moderate deviations. *Ann. Probab.*, **31** (1) (2003), 413–443.
[27] Hairer M. and Li X. M.: Averaging dynamics driven by fractional Brownian motion. https://arxiv.org/pdf/1902.11251.pdf.
[28] Hairer M. and Pardoux E.: Homogenization of periodic linear degenerate PDEs. *J. Func. Anal.*, **255** (2008), 2462–2487.
[29] Hairer M. and Pardoux E.: Fluctuations around a homogenised semilinear random PDE. https://arxiv.org/abs/1911.02865.
[30] Hu W. and Li C. J.: A convergence analysis of the perturbed compositional gradient flow: averaging principle and normal deviations. *Discrete Cont. Dynam. Syst.-A*, **38** (10) (2018), 4951–4977.
[31] 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.
[32] Kelly D. and Vanden-Eijnden E.: Fluctuations in the heterogeneous multiscale methods for fast-slow systems. *Res. Math. Sci.*, **4** (2017), DOI: 10.1186/s40687-017-0112-2.
[33] Khasminskii R. Z.: On stochastic processes defined by differential equations with a small parameter. *Theory Probab. Appl.*, **11** (1966), 211–228.
[34] Khasminskii R. Z. and Yin G.: On averaging principles: an asymptotic expansion approach. *SIAM J. Math. Anal.*, **35** (2004), 1534–1560.
[35] Khasminskii R. Z and Yin G.: Limit behavior of two-time-scale diffusions revised. *J. Differential Equ.*, **212** (2005), 85–113.
[36] Kifer Y.: Averaging in dynamical systems and large deviations. *Invent. Math.*, **110** (1992), 337–370.
[37] Kifer Y.: Averaging and climate models. *Stochastic Climate Models*, Progress in Probability book series, Springer, **49** (2001), 171–188.
[38] Kifer Y.: L^2 diffusion approximation for slow motion in averaging. *Stoch. Dyna.*, **3** (2003), 213–246.
[39] Kifer Y. and Varadhan S. R. S.: Nonconventional limit theorems in discrete and continuous time via martingales. *Ann. Probab.*, **42** (2) (2014), 649–688.
[40] Konda V. R. and Tsitsiklis J. N.: Convergence rate of linear two-time-scale stochastic approximation. *Ann. Appl. Probab.*, **14** (2004), 796–819.
[41] Kosygina E., Rezakhanlou F. and Varadhan S. R. S.: Stochastic homogenization of Hamilton-Jacobi-Bellman equations. *Comm. Pure Appl. Math.*, **59** (2006), 1489–1521.
[42] Kuehn C.: Multiple time scale dynamics, volume 191 of Applied Mathematical Sciences. Springer, Cham, 2015.

[43] Ladyženskaja O. A., Solonnikov V. A. and Ural'ceva N. N.: Linear and Quasi-linear Equations of Parabolic Type. Translated from Russian by S. Smmith. Amercian Mathematical Society, 1968.

[44] Li X. M.: An averaging principle for a completely integrable stochastic Hamiltonian system. *Nonlinearity*, **21** (2008), 803–822.

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

[46] Morse M. R. and Spiliopoulos K.: Moderate deviations principle for systems of slow-fast diffusions. *Asymptotic Anal.*, **105** (2017), 97–135.

[47] Pakdaman K., Thieullen M. and Wainrib G.: Asymptotic expansion and central limit theorem for multiscale piecewise-deterministic Markov processes. *Stoch. Proc. Appl.*, **122** (2012), 2292–2318.

[48] Papanicolaou G. C., Stroock D. and Varadhan S. R. S.: Martingale approach to some limit theorems. Proceedings of the 1976 Duke University Conference on Turbulence, Durham, NC, 1976.

[49] Pardoux E.: Homogenization of linear and semilinear second order parabolic PDEs with periodic coefficients: a probabilistic approach. *J. Func. Anal.*, **167** (1999), 498–520.

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

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

[52] Pavliotis G. A. and Stuart A. M.: Multiscale methods: averaging and homogenization. Texts Appl. Math., 53, Springer, New York, 2008.

[53] Puhalskii A. A.: On large deviations of coupled diffusions with time scale separation. *Ann. Probab.*, **44** (2016), 3111–3186.

[54] Röckner M., Sun X. and Xie L.: Strong and weak convergence in averaging principle for SDEs with Hölder coefficients. https://arxiv.org/pdf/1907.09256.pdf.

[55] Röckner M. and Xie L.: Diffusion approximation for fully coupled stochastic differential equations. To appear in *Ann. Probab.*, https://arxiv.org/pdf/2008.04817.pdf.

[56] H. A. Simon and Ando A.: Aggregation of variables in dynamical systems. *Econometrica*, **29** (1961), 111–138.

[57] Spiliopoulos K.: Fluctuation analysis and short time asymptotics for multiple scales diffusion processes. *Stoch. Dyna.*, **14** (2014), pp. 1350026.

[58] Spiliopoulos K.: Quenched large deviations for multiscale diffusion processes in random environments. *Electron. J. Probab.*, **20 (15)** (2015), 1–29.

[59] Veretennikov A. Yu.: On the strong solutions of stochastic differential equations. *Theory Probab. Appl.*, **24** (1979), 354–366.

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

[61] Veretennikov A. Yu.: On large deviations in the averaging principle for SDEs with a "full dependence". *Ann. Probab.*, **27** (1999), 284–296.

[62] Veretennikov A. Yu.: On polynomial mixing bounds for stochastic differential equations. *Stoch. Proc. Appl.*, **70** (1997), 115–127.

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

[64] Zhang X.: Stochastic homemomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.*, **16** (2011), 1096–1116.

Michael Röckner: Fakultät für Mathematik, Universität Bielefeld, 33615, Bielefeld, Germany, Email: roeckner@math.uni-bielefeld.de

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