Averaging principle for distribution dependent stochastic differential equations driven by fractional Brownian motion and standard Brownian motion

Guangjun Shen, Jie Xiang and Jiang-Lun Wu*

October 29, 2021

Abstract

In this paper, we study distribution dependent stochastic differential equations driven simultaneously by fractional Brownian motion with Hurst index $H > \frac{1}{2}$ and standard Brownian motion. We first establish the existence and uniqueness theorem for solutions of the distribution dependent stochastic differential equations by utilising the Carathéodory approximation. We then show that, under certain averaging condition, the solutions of distribution dependent stochastic differential equations can be approximated by the solutions of the associated averaged distribution dependent stochastic differential equations in the sense of the mean square convergence.

Keywords: Distribution dependent stochastic differential equations; fractional Brownian motion; stochastic averaging principle.

2020 Mathematics Subject Classification: 60G22; 60H10; 34C29; 35Q83

1 Introduction

Recently, stochastic processes possessing self-similarity and long-range dependence property have become an important component of stochastic models in various scientific areas including hydrology, telecommunication, turbulence, image processing and finance (see, for example, Decreusefond and Üstünel [9]). The celebrated and also most widely used stochastic process that exhibits the long-range dependence property and self-similarity is the fractional Brownian motion (fBm in short). Interesting surveys of fBm and related stochastic calculus could be found in Biagini et al. [4], Mishura [27], Nualart[29], Hu [17]. We recall that fBm with Hurst parameter $H \in (0, 1)$ is a zero mean Gaussian process $B^H = \{B^H_t, t \geq 0\}$ with covariance function

$$R_H(s,t) := \frac{1}{2}(s^{2H} + t^{2H} - |t-s|^{2H}),\ \ s,t \geq 0.$$

*Corresponding author. E-mail: j.l.wu@swansea.ac.uk.

This research is supported by the National Natural Science Foundation of China (12071003)

For $H = \frac{1}{2}$, fBm B^H is nothing but standard Brownian motion which, in this paper, is denoted by W. B^H is neither a semimartingale nor a Markov process unless $H = \frac{1}{2}$. Hence, one can not use the classical Itô theory with respect to W to establish stochastic calculus for fBm B^H.

On the other hand, it is now well known that stochastic differential equations (SDEs) play a significant role in modelling evolutions of dynamical systems when taking into account uncertainty features along time in diverse fields ranging from biology, chemistry, and physics, as well as economics and finance, and so on (see, e.g., Sobczyk [38], Mao [23] and references therein). There have been many fundamental studies addressing the existence and uniqueness of solutions of SDEs driven by fBm. Coutin and Qian [8] established the existence of strong solutions for SDEs driven by fBm with Hurst parameter $H > \frac{1}{4}$. Nualart and Răşcanu [30] obtained global existence and uniqueness result of solutions of multidimensional, time-inhomogeneous, SDEs driven by fBm with Hurst parameter $H > \frac{1}{2}$. Furthermore, along systematical studies of fBm, many authors have proposed to study SDEs driven by standard Brownian motion and fBm simultaneously, which led to many interesting theoretical questions about self-similar Gaussian processes and fields in general. Zähle [43] and Kubilius [21] defined the stochastic integral with respect to fBm as an extended Riemann-Stieltjes pathwise integral and used p-variation estimates to obtain the existence and uniqueness of the solutions of the SDEs driven by both fBm and standard Brownian motion, respectively. Using fractional integration and the classical Itô stochastic calculus, Guerra and Nualart [13] proved an existence and uniqueness theorem for solutions of multidimensional, time-inhomogeneous, SDEs driven simultaneously by a multidimensional fBm with Hurst parameter $H > \frac{1}{2}$ and a multidimensional standard Brownian motion. Under mild regularity assumptions on the coefficients, Mishura and Shevchenko [28] proved that SDEs driven by dependent fBm and Brownian motion has a unique solution. Silva et al. [37] established an existence and uniqueness result for solutions of multidimensional, time-dependent, SDEs driven simultaneously by a multidimensional fBm with Hurst parameter $H > \frac{1}{2}$ and a multidimensional standard Brownian motion under a weaker condition than the Lipschitz one. Sönmez [39] considered such SDEs with irregular drift coefficient. Pei et al. [32] obtained averaging principle for fast-slow system driven by mixed fractional Brownian rough path.

Generally, nonlinear Fokker-Planck equations can be characterised by distribution dependent stochastic differential equations (DDSDEs), which are also named as McKean-Vlasov SDEs or mean field SDEs. A distinct feature of such systems is the appearance of probability laws in the coefficients of the resulting equations. There has been an increasing interest to study existence and uniqueness for solutions of DDSDEs. Wang [41] established strong well-posedness of DDSDEs with one-sided Lipschitz continuous drifts and Lipschitz-continuous dispersion coefficients. Under integrability conditions on distribution dependent coefficients, Huang and Wang [18] obtained the existence and uniqueness for DDSDEs with non-degenerate noise. Mehri and Stannat [24] proposed a Lyapunov-type approach to the problem of existence and uniqueness of general law-dependent SDEs. Many interesting studies of DDSDEs have been developed further in Bao et al. [2], Huang et al. [19], Ren and Wang [33], Röckner and Zhang [34], Mishura and Veretennikov [27], Chaudru de Raynal [6], Hammersley et al [16] and references therein. More recently, Fan et al. [11] considered the following DDSDE driven by fBm with Hurst parameter $H > \frac{1}{2}$

$$dX(t) = b(t, X(t), \mathscr{L}(X(t)))dt + \sigma(t, \mathscr{L}(X(t)))dB_t^H \tag{1.1}$$

by showing the well-posedness and by deriving a Bismut type formula for the Lions derivative using Malliavin calculus. Galeati et al. [12] studied DDSDEs with irregular, possibly distributional drift, driven by additive fBm of Hurst parameter $H \in (0, 1)$ and established strong well-posedness under a variety of assumptions on the drifts. Buckdahn and Jing [5] considered mean-field SDEs driven by fBm and related stochastic control problem. Bauer and Meyer-Brandis [3] established existence and uniqueness results of solutions to McKean-Vlasov equations driven by cylindrical fBm in an infinite-dimensional Hilbert space setting with irregular drift.

It is worth to note that the averaging principle, initiated by Khasminskii in the seminal work [20], is a very efficient and important tool in the study of SDEs for modelling problems arising in many practical research areas. Indeed, averaging principle is an effective method for studying dynamical systems with highly oscillating components. Under certain suitable conditions, the highly oscillating components can be "averaged out" to produce an averaged system. The averaged system is easier for analysis which governs the evolution of the original system over long time scales. The fundamental idea of the stochastic averaging principle is to study complex stochastic equations with related averaging stochastic equations, which paves a convenient and easy way to study many important properties (see, e.g., [42], [10], [22], [31], [36], [14], [15]). Although there exist many investigations in the literature devoted to studying stochastic averaging principle for SDEs driven by Brownian motion, fBm, Lévy processes as well as more general stochastic measures inducing semimartingales, etc., as we know, there is not any consideration of averaging principle for DDSDEs driven simultaneously by fBm and Brownian motion. Moreover, due to their distribution dependent nature, they are potentially useful and important for modelling complex systems in diverse areas of applications. Comparing to the classical SDEs driven by Brownian motion, fBm, and Lévy processes, the DDSDEs are much more complex, therefore, a stochastic averaging principle for such SDEs is naturally interesting and would also be very useful. This motivates us to carry out the present paper, aiming to establish a stochastic averaging principle for the DDSDEs driven simultaneously by fBm and Brownian motion

$$dX(t) = b(t, X(t), \mathscr{L}(X(t)))dt + \sigma_W(t, X(t), \mathscr{L}(X(t)))dW_t + \sigma_H(t, \mathscr{L}(X(t)))dB_t^H, \tag{1.2}$$

where B_t^H is an m-dimensional fBm with Hurst parameter $H \in (\frac{1}{2}, 1)$, and W_t is an r-dimensional standard Browinan motion, independent of B^H, and $X_0 = \xi$ is a d-dimensional random variable independent of (B^H, W). Precise assumptions on the coefficients $b : [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$, $\sigma_W : [0, T] \times \mathbb{R}^d \times \mathcal{P}_\theta(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^r$, and $\sigma_H : [0, T] \times \mathcal{P}_\theta(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^m$ will be specified in later sections, here $\mathcal{P}_\theta(\mathbb{R}^d)$ stands for the set of probability measures on $\mathbb{R}^d$ with finite θ-th moment. The motivation to consider such equations comes from some financial applications, where Brownian motion as a model of volatility is inappropriate because of the lack of memory, and fBm with $H > \frac{1}{2}$ might be too "smooth" (i.e., not rough enough). A model equation driven by the both noise processes is then free of such drawbacks (see, for example, [27], [28]).

Throughout this paper, the letter C will denote a positive constant, with or without subscript, its value may change in different occasions. We will write the dependence of the constant on parameters explicitly if it is essential.

The rest of the paper is organised as follows. In Section 2, we present some preliminaries for this paper. In Section 3, we prove the existence and uniqueness of solutions to our DDSDEs driven by fBm and Brownian motion. In Section 4, we establish an approximation theorem as an averaging principle for the solutions of the concerned DDSDEs.

2 Preliminaries

In this section, we briefly give preliminaries for our discussions in the sequel. We will use the following notations. $|\cdot|$ denotes the Euclidean norm and $\langle\cdot,\cdot\rangle$ denotes the scalar product in d-dimensional Euclidean space $\mathbb{R}^d$, and for a matrix, we denote by $\|\cdot\|$ the Euclidean norm. Fix a time interval $[0,T]$ and a complete probability space $(\Omega,\mathscr{F},\mathbb{P})$ associated with fBm B^H and Brownian motion W. We assume that there is a sufficiently rich sub-σ-algebra $\mathscr{F}_0\subset\mathscr{F}$ independent of (B^H,W) such that for any $\mu\in\mathcal{P}_2(\mathbb{R}^d)$ there exists a random variable $X\in L^2(\Omega\to\mathbb{R}^d,\mathscr{F}_0,\mathbb{P})$ with distribution μ. For each $t\in[0,T]$, let $\{\mathscr{F}_t\}_{t\in[0,T]}$ be the σ-field generated by the random variables $\{X_0,B^H_s,W_s,s\in[0,t]\}$ and the P-null sets, completed and augmented by $\mathscr{F}_0$.

For technical reasons, we will work on the following subspace of $\mathcal{P}(\mathbb{R}^d)$ for any fixed $\theta\in[2,\infty)$

$$\mathcal{P}_\theta(\mathbb{R}^d):=\left\{\mu\in\mathcal{P}(\mathbb{R}^d):\mu(|\cdot|^\theta):=\int_{\mathbb{R}^d}|x|^\theta\mu(dx)<\infty\right\}$$

which is a Polish space under the L^θ-Wasserstein distance

$$\mathbb{W}_\theta(\mu_1,\mu_2):=\inf_{\pi\in\mathscr{C}(\mu_1,\mu_2)}\left(\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^\theta\pi(dx,dy)\right)^{\frac{1}{\theta}},\mu_1,\mu_2\in\mathcal{P}_\theta(\mathbb{R}^d),$$

where $\mathscr{C}(\mu_1,\mu_2)$ is the set of probability measures on $\mathbb{R}^d\times\mathbb{R}^d$ with marginals μ_1 and μ_2.

Note that for any $x\in\mathbb{R}^d$, the Dirac measure δ_x belongs to $\mathcal{P}_\theta(\mathbb{R}^d)$ for any $\theta\in[2,\infty)$ and if $\mu_1=\mathscr{L}(X)$, $\mu_2=\mathscr{L}(Y)$ are the corresponding distributions of random variables X and Y respectively, then

$$(\mathbb{W}_\theta(\mu_1,\mu_2))^\theta\leq\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^\theta\mathscr{L}((X,Y))(dx,dy)=\mathbb{E}|X-Y|^\theta,$$

in which $\mathscr{L}((X,Y))$ represents the joint distribution of the random pair (X,Y). For arbitrarily fixed $T>0$, let $C([0,T];\mathbb{R}^d)$ be the Banach space of all $\mathbb{R}^d$-valued continuous functions on $[0,T]$, endowing with the supremum norm. Furthermore, we let $L^p(\Omega;C([0,T];\mathbb{R}^d))$ be the totality of $C([0,T];\mathbb{R}^d)$-valued random variables X satisfying $\mathbb{E}[\sup_{0\leq t\leq T}|X(t)|^p]<\infty$. Then, $L^p(\Omega;C([0,T];\mathbb{R}^d))$ is a Banach space under the norm

$$||X||_{L^p}:=(\mathbb{E}[\sup_{0\leq t\leq T}|X(t)|^p])^{\frac{1}{p}}.$$

In the follows, we recall the basic definitions and properties of the fractional calculus. For a detailed presentation of these notions we refer to [35]. Let $a,b\in\mathbb{R}$, $a<b$. Let $f\in L^1(a,b)$

and $\alpha > 0$. The left and right-sided fractional integrals of f of order α are defined for almost all $x \in (a, b)$ by

$$I^{\alpha}_{a+} f(x) = \frac{1}{\Gamma(\alpha)} \int_a^x (x-y)^{\alpha-1} f(y) dy$$

and

$$I^{\alpha}_{b-} f(x) = \frac{1}{\Gamma(\alpha)} \int_x^b (y-x)^{\alpha-1} f(y) dy,$$

respectively. Let $I^{\alpha}_{a+}(L^p)$ (resp.$f \in I^{\alpha}_{b-}(L^p)$) and $0 < \alpha < 1$, then the left and right-sided fractional derivatives are defined by

$$D^{\alpha}_{a+} f(x) = \frac{1}{\Gamma(1-\alpha)} \left(\frac{f(x)}{(x-a)^{\alpha}} + \alpha \int_a^x \frac{f(x)-f(y)}{(x-y)^{\alpha+1}} dy \right),$$

and

$$D^{\alpha}_{b-} f(x) = \frac{1}{\Gamma(1-\alpha)} \left(\frac{f(x)}{(b-x)^{\alpha}} + \alpha \int_x^b \frac{f(x)-f(y)}{(y-x)^{\alpha+1}} dy \right)$$

for almost all $x \in (a, b)$ (the convergence of the integrals at the singularity $y = x$ holds point-wise for almost all $x \in (a, b)$ if $p = 1$ and moreover in L^p-sense if $1 < p < \infty$).

Recall the following properties of these operators:

- If $\alpha < \frac{1}{p}$ and $q = \frac{p}{1-\alpha p}$, then

$$I^{\alpha}_{a+}(L^p) = I^{\alpha}_{b-}(L^p) \subset L^q(a, b).$$

- If $\alpha > \frac{1}{p}$, then

$$I^{\alpha}_{a+}(L^p) \cup I^{\alpha}_{b-}(L^p) \subset C^{\alpha-\frac{1}{p}}(a, b),$$

where $C^{\alpha-\frac{1}{p}}(a, b)$ denotes the space of $(\alpha - \frac{1}{p})$-Hölder continuous functions of order $\alpha - \frac{1}{p}$ in the interval $[a, b]$.

The following inversion formulas hold:

$$I^{\alpha}_{a+}(D^{\alpha}_{a+} f) = f$$

for all $f \in I^{\alpha}_{a+}(L^p)$, and

$$D^{\alpha}_{a+}(I^{\alpha}_{a+} f) = f$$

for all $f \in L^1(a, b)$. Similar inversion formulas hold for the operators I^{α}_{b-} and D^{α}_{b-}.

The following integration by parts formula holds:

$$\int_a^b (D^{\alpha}_{a+} f)(s) g(s) ds = \int_a^b f(s) (D^{\alpha}_{b-} g)(s) ds,$$

for any $f \in I^{\alpha}_{a+}(L^p)$, $g \in I^{\alpha}_{b-}(L^q)$, $\frac{1}{p} + \frac{1}{q} = 1$. To prove our main results, we also present the following Hardy-Littlewood inequality.

Lemma 2.1 *([40]) Let $1 < \tilde{p} < \tilde{q} < \infty$ and $\frac{1}{\tilde{q}} = \frac{1}{\tilde{p}} - \alpha$. If $f : \mathbb{R}^+ \to \mathbb{R}$ belongs to $L^{\tilde{p}}(0, \infty)$, then $I^{\alpha}_{0+} f(x)$ converges absolutely for almost every x, and moreover*

$$\|I^{\alpha}_{0+} f(x)\|_{L^{\tilde{q}}(0,\infty)} \leq C_{\tilde{p},\tilde{q}} \|f\|_{L^{\tilde{p}}(0,\infty)}$$

holds for some positive constant $C_{\tilde{p},\tilde{q}}$.

Lemma 2.2 *(Bihari's inequality,[23]) Let $T > 0$, and $c > 0$. Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a continuous nondecreasing function such that $\psi(t) > 0$ for all $t > 0$. Let $u(\cdot)$ be a Borel measurable bounded nonnegative function on [0,T], and let $v(\cdot)$ be a nonnegative integrable function on [0,T]. If*

$$u(t) \leq c + \int_0^t v(s)\psi(u(s))ds \qquad for \quad all \quad 0 \leq t \leq T,$$

then

$$u(t) \leq G^{-1}(G(c) + \int_0^t v(s)ds)$$

holds for all $t \in [0, T]$ such that

$$G(c) + \int_0^t v(s)ds \in Dom(G^{-1}),$$

where

$$G(r) = \int_1^r \frac{ds}{\psi(s)} \qquad on \quad r > 0,$$

and G^{-1} is the inverse function of G.

3 Existence and Uniqueness

In this section, we will establish the existence and uniqueness theorem for the solution of distribution dependent stochastic differential equations (1.2) driven by fBm and standard Brownian motion under the following Assumption 3.1 using the Carathéodory approximation technique. For this purpose, we define the Carathéodory approximation as follows. For any integer $k \geq 1$, define $X_k(t) = X_0$ for all $-1 \leq t \leq 0$ and

$$\begin{aligned} X_k(t) = X_0 &+ \int_0^t b(s, X_k(s - \frac{1}{k}), \mathscr{L}(X_k(s - \frac{1}{k})))ds \\ &+ \int_0^t \sigma_W(s, X_k(s - \frac{1}{k}), \mathscr{L}(X_k(s - \frac{1}{k})))dW_s \\ &+ \int_0^t \sigma_H(s, \mathscr{L}(X_k(s - \frac{1}{k})))dB_s^H, \quad t \in (0, T]. \end{aligned} \tag{3.1}$$

Note that for $0 \leq t \leq \frac{1}{k}$, $X_k(t)$ can be computed by

$$X_k(t) = X_0 + \int_0^t b(s, X_0, \mathscr{L}(X_0))ds \quad + \int_0^t \sigma_W(s, X_0, \mathscr{L}(X_0))dW_s + \int_0^t \sigma_H(s, \mathscr{L}(X_0))dB_s^H,$$

then for $\frac{1}{k} \leq t \leq \frac{2}{k}$, $X_k(t)$ can be computed by

$$\begin{aligned} X_k(t) = X_k(\frac{1}{k}) &+ \int_{\frac{1}{k}}^t b(s, X_k(s - \frac{1}{k}), \mathscr{L}(X_k(s - \frac{1}{k})))ds \\ &+ \int_{\frac{1}{k}}^t \sigma_W(s, X_k(s - \frac{1}{k}), \mathscr{L}(X_k(s - \frac{1}{k})))dW_s + \int_{\frac{1}{k}}^t \sigma_H(s, \mathscr{L}(X_k(s - \frac{1}{k})))dB_s^H, \end{aligned}$$

and so on. It is well known that comparing with Picard's successive approximation technique, the advantage of using Carathéodory approximation technique is that we do not need to compute $X_1(t), X_2(t), \cdots, X_{k-1}(t)$ to compute $X_k(t)$. In fact, we can compute $X_k(t)$ directly over intervals of length $\frac{1}{k}$. It is noted that our results are new even when the coefficients appeared in (1.2) satisfy Lipschitz condition.

Assumption 3.1 *There exists a non-decreasing function K(t) such that for any $t \in [0,T]$, $x,\ y \in \mathbb{R}^d$, $\mu,\ \nu \in \mathcal{P}_\theta(\mathbb{R}^d)$,*

$$\begin{aligned}|b(t,x,\mu) - b(t,y,\nu)|^p &\le K(t)\psi(|x-y|^p + \mathbb{W}_\theta^p(\mu,\nu)),\\ \|\sigma_W(t,x,\mu) - \sigma_W(t,y,\nu)\|^p &\le K(t)\psi(|x-y|^p + \mathbb{W}_\theta^p(\mu,\nu)),\\ \|\sigma_H(t,\mu) - \sigma_H(t,\nu)\|^p &\le K(t)\psi(\mathbb{W}_\theta^p(\mu,\nu)),\end{aligned}$$

and

$$|b(t,0,\delta_0)|^p + \|\sigma_W(t,0,\delta_0)\|^p + \|\sigma_H(t,\delta_0)\|^p \le K(t),$$

where $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ is continuous and non-decreasing concave function with $\psi(0) = 0$, $\psi(x) > 0$, for every $x > 0$ such that $\int_{0^+} \frac{1}{\psi(x)}dx = +\infty$.

Example 3.2 *We can give a few concrete examples of the function $\psi(\cdot)$. Let $K > 0$, and let $\delta \in (0,1)$ be sufficiently small. Define*
$\psi_1(u) = Ku, u \ge 0.$

$$\psi_2(u) = \begin{cases} u\log(u^{-1}), & 0 \le u \le \delta;\\ \delta\log(\delta^{-1}) + \psi'_2(\delta-)(u-\delta), & u > \delta.\end{cases}$$

$$\psi_3(u) = \begin{cases} u\log(u^{-1})\log\log(u^{-1}), & 0 \le u \le \delta;\\ \delta\log(\delta^{-1})\log\log(\delta^{-1}) + \psi'_3(\delta-)(u-\delta), & u > \delta.\end{cases}$$

where ψ' denotes the derivative of the function ψ. They are all concave nondecreasing functions satisfying $\int_{0^+} \frac{du}{\psi_i(u)} = \infty$. Furthermore, we observed that the Lipschitz condition is a special case of our proposed condition.

Definition 3.3 *An $\mathbb{R}^d$ valued stochastic process $X = (X_t)_{0 \le t \le T}$ is called an unique solution of (1.2), if $X \in L^\theta$ and satisfies the follwing*

(i)

$$\begin{aligned}X(t) = \xi &+ \int_0^t b(s, X(t), \mathscr{L}(X(s)))ds + \int_0^t \sigma_W(s, X(s), \mathscr{L}(X(s)))dW_s\\ &+ \int_0^t \sigma_H(s, \mathscr{L}(X(s)))dB_s^H, t \in [0,T], \mathbb{P}-a.s..\end{aligned} \tag{3.2}$$

(ii) If $Y = (Y_t)_{0 \le t \le T}$ is another solution with $Y(0) = \xi$, then

$$\mathbb{P}(X(t) = Y(t) \text{ for all } 0 \le t \le T) = 1.$$

Note that $\sigma_H(s, \mathscr{L}(X(s)))$ is a deterministic function, then $\int_0^t \sigma_H(s, \mathscr{L}(X(s)))dB_s^H$ can be regarded as a Wiener integral with respect to fBm.

Next, we will prove the uniform boundedness property for the sequence of stochastic processes $\{X_k(t), k \ge 1\}$ given by Equation (3.1).

Lemma 3.4 *Suppose that Assumption 3.1 holds and $\xi \in L^p(\Omega \to \mathbb{R}, \mathscr{F}_0, \mathbb{P})$ with $p \geq \theta$ and $p > 1/H$. Then for all $k \geq 1$,*

$$\mathbb{E}(\sup_{0\leq s\leq T} |X_k(s)|^p) \leq C_2 := (4^{p-1}\mathbb{E}|\xi|^p + C_1)e^{C_1 T},$$

where $C_1 = C_{\lambda,p,H,T}K(T)(2a+1)$.

Proof. Following the simple inequality

$$|x_1 + x_2 + \cdots + x_k|^p \leq k^{p-1}(|x_1|^p + |x_2|^p + \cdots + |x_k|^p), \tag{3.3}$$

we have

$$\begin{aligned}\mathbb{E}(\sup_{0\leq s\leq t} |X_k(s)|^p) \leq\ & 4^{p-1}\mathbb{E}|\xi|^p + 4^{p-1}\mathbb{E}(\sup_{0\leq s\leq t}\left|\int_0^s b(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))du\right|^p) \\ & + 4^{p-1}\mathbb{E}(\sup_{0\leq s\leq t}\left|\int_0^s \sigma_W(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))dW_u\right|^p) \\ & + 4^{p-1}\mathbb{E}(\sup_{0\leq s\leq t}\left|\int_0^s \sigma_H(u, \mathscr{L}(X_k(u-\frac{1}{k})))dB_u^H\right|^p) \\ =:\ & 4^{p-1}\mathbb{E}|\xi|^p + I_1 + I_2 + I_3.\end{aligned} \tag{3.4}$$

For the term I_1, by Hölder inequality and Assumption 3.1, we have

$$\begin{aligned} I_1 &= 4^{p-1}\mathbb{E}(\sup_{0\leq s\leq t}\left|\int_0^s b(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))du\right|^p) \\ &\leq (4T)^{p-1}\mathbb{E}\int_0^t \left|b(s, X_k(s-\frac{1}{k}), \mathscr{L}(X_k(s-\frac{1}{k})))\right|^p ds \\ &\leq (8T)^{p-1}\mathbb{E}\int_0^t (\left|b(s, X_k(s-\frac{1}{k}), \mathscr{L}(X_k(s-\frac{1}{k}))) - b(s,0,\delta_0)\right|^p + |b(s,0,\delta_0)|^p)ds \\ &\leq (8T)^{p-1}\mathbb{E}\int_0^t K(s)\left[\psi\left(|X_k(s-\frac{1}{k})|^p + \mathbb{W}_\theta^p(\mathscr{L}(X_k(s-\frac{1}{k})), \delta_0)\right) + 1\right]ds \\ &\leq (8T)^{p-1}\int_0^t K(s)\left(\psi\left(2\mathbb{E}|X_k(s-\frac{1}{k})|^p\right) + 1\right)ds. \end{aligned} \tag{3.5}$$

For the term I_2, it comes from Burkholder-Davis-Gundy inequality, Hölder inequality, and Assumption 3.1, we obtain

$$\begin{aligned} I_2 &= 4^{p-1}\mathbb{E}(\sup_{0\leq s\leq t}\left|\int_0^s \sigma_W(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))dW_u\right|^p) \\ &\leq C_p\mathbb{E}[\int_0^t \|\sigma_W(s, X_k(s-\frac{1}{k}), \mathscr{L}(X_k(s-\frac{1}{k})))\|^2 ds]^{\frac{p}{2}} \\ &\leq C_p T^{\frac{p}{2}-1}\mathbb{E}\int_0^t \|\sigma_W(s, X_k(s-\frac{1}{k}), \mathscr{L}(X_k(s-\frac{1}{k})))\|^p ds \\ &\leq C_p T^{\frac{p}{2}-1}\mathbb{E}\int_0^t (\|\sigma_W(s, X_k(s-\frac{1}{k}), \mathscr{L}(X_k(s-\frac{1}{k}))) - \sigma_W(s,0,\delta_0)\|^p + \|\sigma_W(s,0,\delta_0)\|^p)ds \end{aligned} \tag{3.6}$$

$$\le C_p T^{\frac{p}{2}-1}\mathbb{E}\int_0^t K(s)\left[\psi\left(|X_k(s-\frac{1}{k})|^p + \mathbb{W}_\theta^p(\mathscr{L}(X_k(s-\frac{1}{k})),\delta_0)\right)+1\right]ds$$
$$\le C_p T^{\frac{p}{2}-1}\int_0^t K(s)\left(\psi\left(2\mathbb{E}|X_k(s-\frac{1}{k})|^p\right)+1\right)ds.$$

For the term I_3. It follows from Theorem 4 in [1], (3.5) in [11], and Assumption 3.1, we have

$$\begin{aligned}
I_3 &= 4^{p-1}\mathbb{E}(\sup_{0\le s\le t}|\int_0^s \sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))dB_u^H|^p)\\
&\le C_{\lambda,p,H}T^{pH-1}\int_0^t \|\sigma_H(s,\mathscr{L}(X_k(s-\frac{1}{k})))\|^p ds\\
&\le C_{\lambda,p,H}T^{pH-1}\int_0^t (\|\sigma_H(s,\mathscr{L}(X_k(s-\frac{1}{k})))-\sigma_H(s,\delta_0)\|^p + \|\sigma_H(s,\delta_0)\|^p)ds \qquad (3.7)\\
&\le C_{\lambda,p,H}T^{pH-1}\int_0^t K(s)\left[\psi\left(W_\theta^p(\mathscr{L}(X_k(s-\frac{1}{k})),\delta_0)\right)+1\right]ds\\
&\le C_{\lambda,p,H}T^{pH-1}\int_0^t K(s)\left(\psi\left(\mathbb{E}|X_k(s-\frac{1}{k})|^p\right)+1\right)ds,
\end{aligned}$$

where $C_{\lambda,p,H}$ above may depend only on H by choosing proper $1-H<\lambda<1-\frac{1}{p}$.

Given that $\psi(\cdot)$ is concave and increasing, there must exist a positive number a such that

$$\psi(u)\le a(1+u). \qquad (3.8)$$

Hence, we have

$$\begin{aligned}
&\mathbb{E}(\sup_{0\le s\le t}|X_k(s)|^p)\\
&\le 4^{p-1}\mathbb{E}|\xi|^p + ((8T)^{p-1}+C_pT^{\frac{p}{2}-1}+C_{\lambda,p,H}T^{pH-1})\\
&\quad\times\int_0^t K(s)\left(\psi\left(2\mathbb{E}|X_k(s-\frac{1}{k})|^p\right)+1\right)ds \qquad (3.9)\\
&\le 4^{p-1}\mathbb{E}|\xi|^p + C_{\lambda,p,H,T}K(T)(2a+1)\int_0^t\left(\mathbb{E}(\sup_{0\le u\le s}|X_k(u)|^p)+1\right)ds,
\end{aligned}$$

which with the help of Gronwall inequality, gives

$$\mathbb{E}(\sup_{0\le s\le t}|X_k(s)|^p)\le(4^{p-1}\mathbb{E}|\xi|^p + C_1T)e^{C_1T},$$

where $C_1 = C_{\lambda,p,H,T}K(T)(2a+1)$. ■

Lemma 3.5 *Suppose that $\xi\in L^p(\Omega\to\mathbb{R},\mathscr{F}_0,\mathbb{P})$ with $p\ge\theta$ and $p>\frac{1}{H}$. Then*

$$\mathbb{E}|X_k(t)-X_k(s)|^p\le C_3[(t-s)^p+(t-s)^{\frac{p}{2}}+(t-s)^{pH}],\quad t\ge s,$$

where $C_3=(6^{p-1}+C_p+C_{\lambda,p,H})(a+1)K(T)(C_2+1)$.

Proof. It follows from (3.1), we have

$$
\begin{aligned}
X_k(t) - X_k(s) = &\int_s^t b(u, X_k(u - \frac{1}{k}), \mathscr{L}(X_k(u - \frac{1}{k})))du \\
&+ \int_s^t \sigma_W(u, X_k(u - \frac{1}{k}), \mathscr{L}(X_k(u - \frac{1}{k})))dW_u \\
&+ \int_s^t \sigma_H(u, \mathscr{L}(X_k(u - \frac{1}{k})))dB_u^H.
\end{aligned} \tag{3.10}
$$

In view of inequality (3.3), we get

$$
\begin{aligned}
&\mathbb{E}|X_k(t) - X_k(s)|^p \\
&\le 3^{p-1}\mathbb{E}|\int_s^t b(u, X_k(u - \frac{1}{k}), \mathscr{L}(X_k(u - \frac{1}{k})))du|^p \\
&\quad + 3^{p-1}\mathbb{E}|\int_s^t \sigma_W(u, X_k(u - \frac{1}{k}), \mathscr{L}(X_k(u - \frac{1}{k})))dW_u|^p \\
&\quad + 3^{p-1}\mathbb{E}|\int_s^t \sigma_H(u, \mathscr{L}(X_k(u - \frac{1}{k})))dB_u^H|^p \\
&=: J_1 + J_2 + J_3.
\end{aligned} \tag{3.11}
$$

By Hölder inequality and Assumption 3.1, we have

$$
\begin{aligned}
J_1 &= 3^{p-1}\mathbb{E}\left|\int_s^t b(u, X_k(u - \frac{1}{k}), \mathscr{L}(X_k(u - \frac{1}{k})))du\right|^p \\
&\le 3^{p-1}(t-s)^{p-1}\mathbb{E}\int_s^t \left|b(u, X_k(u - \frac{1}{k}), \mathscr{L}(X_k(u - \frac{1}{k})))\right|^p du \\
&\le 6^{p-1}(t-s)^{p-1}\mathbb{E}\int_s^t (\left|b(u, X_k(u - \frac{1}{k}), \mathscr{L}(X_k(u - \frac{1}{k}))) - b(u, 0, \delta_0)\right|^p \\
&\qquad\qquad + |b(u, 0, \delta_0)|^p)du \\
&\le 6^{p-1}(t-s)^{p-1}\mathbb{E}\int_s^t K(u)\left[\psi\left(|X_k(u - \frac{1}{k})|^p + \mathbb{W}_\theta^p(\mathscr{L}(X_k(u - \frac{1}{k})), \delta_0)\right) + 1\right]du \\
&\le 6^{p-1}(t-s)^{p-1}\int_s^t K(u)\left(\psi\left(2\mathbb{E}|X_k(u - \frac{1}{k})|^p\right) + 1\right)du.
\end{aligned} \tag{3.12}
$$

As for J_2, by Burkholder-Davis-Gundy inequality, Hölder inequality and Assumption 3.1, we obtain

$$
\begin{aligned}
J_2 &= 3^{p-1}\mathbb{E}\left|\int_s^t \sigma_W(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))dW_u\right|^p \\
&\le 3^{p-1}\mathbb{E}\left(\sup_{s\le u_1\le t}\left|\int_s^{u_1} \sigma_W(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))dW_u\right|^p\right) \\
&\le C_p\mathbb{E}[\int_s^t \|\sigma_W(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))\|^2 du]^{\frac{p}{2}} \\
&\le C_p(t-s)^{\frac{p}{2}-1}\mathbb{E}\int_s^t \|\sigma_W(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))\|^p du \\
&\le C_p(t-s)^{\frac{p}{2}-1}\mathbb{E}\int_s^t (\|\sigma_W(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k}))) - \sigma_W(u,0,\delta_0)\|^p \\
&\qquad\qquad + \|\sigma_W(u,0,\delta_0)\|^p)du \\
&\le C_p(t-s)^{\frac{p}{2}-1}\mathbb{E}\int_s^t K(u)\left[\psi\left(|X_k(u-\frac{1}{k})|^p + \mathbb{W}_\theta^p(\mathscr{L}(X_k(u-\frac{1}{k})),\delta_0)\right)+1\right]du \\
&\le C_p(t-s)^{\frac{p}{2}-1}\int_s^t K(u)\left(\psi\left(2\mathbb{E}|X_k(u-\frac{1}{k})|^p\right)+1\right)du.
\end{aligned} \tag{3.13}
$$

For the term J_3, the method main comes from [1], [11], for completeness, we give the main proof. Taking λ satisfying $1-H<\lambda<1-\frac{1}{p}$ because $pH>1$. Using the fact that $\int_u^t (t-r)^{-\lambda}(r-u)^{\lambda-1}dr = C(\lambda)$, the stochastic Fubini theorem and the Hölder inequality, we get

$$
\begin{aligned}
&\mathbb{E}\left|\int_s^t \sigma_H(u, \mathscr{L}(X_k(u-\frac{1}{k})))dB_u^H\right|^p \\
&= C(\lambda)^{-p}\mathbb{E}\left(\left|\int_s^t \left(\int_u^t (t-r)^{-\lambda}(r-u)^{\lambda-1}dr\right)\sigma_H(u, \mathscr{L}(X_k(u-\frac{1}{k})))dB_u^H\right|^p\right) \\
&= C(\lambda)^{-p}\mathbb{E}\left(\left|\int_s^t (t-r)^{-\lambda}\left(\int_s^r (r-u)^{\lambda-1}\sigma_H(u, \mathscr{L}(X_k(u-\frac{1}{k})))dB_u^H\right)dr\right|^p\right) \\
&\le \frac{C(\lambda)^{-p}(p-1)^{p-1}}{(p-1-p\lambda)^{p-1}}(t-s)^{p-1-p\lambda} \\
&\quad\times\int_s^t \mathbb{E}\left|\int_s^r (r-u)^{\lambda-1}\sigma_H(u, \mathscr{L}(X_k(u-\frac{1}{k})))dB_u^H\right|^p dr
\end{aligned} \tag{3.14}
$$

Notice that for each $r\in[0,T]$, $\int_s^r (r-u)^{\lambda-1}\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))dB_u^H$ is a centered Gaussian random variable. By Kahane-Khintchine formula, we obtain that there exists a constant C_p such that

$$
\begin{aligned}
&\mathbb{E}\Big|\int_s^r (r-u)^{\lambda-1}\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))dB_u^H\Big|^p \\
&\quad\le C_p\Big(\mathbb{E}\Big|\int_s^r (r-u)^{\lambda-1}\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))dB_u^H\Big|^2\Big)^{\frac{p}{2}} \\
&\quad\le C_p\Big(\int_s^r\int_s^r (r-u)^{\lambda-1}\|\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))\| \\
&\qquad\qquad \times (r-v)^{\lambda-1}\|\sigma_H(v,\mathscr{L}(X_k(v-\frac{1}{k})))\||u-v|^{2H-2}dudv\Big)^{\frac{p}{2}} \\
&\quad\le C_{p,H}\left(\int_s^r (r-u)^{\frac{\lambda-1}{H}}\|\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))\|^{\frac{1}{H}}du\right)^{pH},
\end{aligned}
\tag{3.15}
$$

where the last inequality is due to the argument in Theorem 1.1 in [25].

Substituting (3.15) into (3.14) and using the condition $1-H<\lambda$ and Lemma 2.1 with $\tilde{q}=pH$ and $\alpha=1-\frac{1-\lambda}{H}$ (imply $\tilde{p}=\frac{pH}{p(\lambda+H-1)+1}$), we have

$$
\begin{aligned}
&\mathbb{E}\Big|\int_s^t \sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))dB_u^H\Big|^p \\
&\le C_{\lambda,p,H}(t-s)^{p-1-\lambda p}\int_s^t\left(\int_s^r (r-u)^{\frac{\lambda-1}{H}}\|\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))\|^{\frac{1}{H}}du\right)^{pH}dr \\
&= C_{\lambda,p,H}(t-s)^{p-1-\lambda p}\left(\|I_{0+}^{\frac{\lambda+H-1}{H}}(\|\sigma_H(r,\mathscr{L}(X_k(r-\frac{1}{k})))\|^{\frac{1}{H}}I_{[s,t]})\|_{L^{pH}[s,t]}\right)^{pH} \\
&\le C_{\lambda,p,H}(t-s)^{p-1-\lambda p}\left(\|\|\sigma_H(r,\mathscr{L}(X_k(r-\frac{1}{k})))\|^{\frac{1}{H}}I_{[s,t]}\|_{L^{\frac{pH}{p(\lambda+H-1)+1}}[s,t]}\right)^{pH} \\
&= C_{\lambda,p,H}(t-s)^{p-1-\lambda p}\left(\int_s^t \|\sigma_H(r,\mathscr{L}(X_k(r-\frac{1}{k})))\|^{\frac{p}{p(\lambda+H-1)+1}}dr\right)^{p(\lambda+H-1)+1} \\
&\le C_{\lambda,p,H}(t-s)^{pH-1}\int_s^t \|\sigma_H(r,\mathscr{L}(X_k(r-\frac{1}{k})))\|^p dr.
\end{aligned}
\tag{3.16}
$$

Hence, by Assumption 3.1, we have

$$
\begin{aligned}
J_3 &= 3^{p-1}\mathbb{E}|\int_s^t \sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))dB_u^H|^p \\
&\le C_{\lambda,p,H}(t-s)^{pH-1}\int_s^t \|\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))\|^p du \\
&\le C_{\lambda,p,H}(t-s)^{pH-1}\int_s^t (\|\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))-\sigma_H(u,\delta_0)\|^p+\|\sigma_H(u,\delta_0)\|^p)du \\
&\le C_{\lambda,p,H}(t-s)^{pH-1}\int_s^t K(u)\left[\psi\left(W_\theta^p(\mathscr{L}(X_k(u-\frac{1}{k})),\delta_0)\right)+1\right]du \\
&\le C_{\lambda,p,H}(t-s)^{pH-1}\int_s^t K(u)\left(\psi\left(\mathbb{E}|X_k(u-\frac{1}{k})|^p\right)+1\right)du.
\end{aligned}
\tag{3.17}
$$

Thus, by (3.8), (3.11) and Lemma 3.4, we can obtain

$$
\begin{aligned}
&\mathbb{E}|X_k(t)-X_k(s)|^p \\
&\le (6^{p-1}+C_p+C_{\lambda,p,H})[(t-s)^{p-1}+(t-s)^{\frac{p}{2}-1}+(t-s)^{pH-1}] \\
&\qquad\qquad \times \int_s^t K(u)\left(\psi\left(2\mathbb{E}|X_k(u-\frac{1}{k})|^p\right)+1\right)du \\
&\le (6^{p-1}+C_p+C_{\lambda,p,H})(2a+1)K(T)(\mathbb{E}(\sup_{0\le u\le T}|X_k(u)|^p)+1) \\
&\qquad\qquad \times [(t-s)^p+(t-s)^{\frac{p}{2}}+(t-s)^{pH}] \\
&\le C_3[(t-s)^p+(t-s)^{\frac{p}{2}}+(t-s)^{pH}],
\end{aligned} \tag{3.18}
$$

where $C_3=(6^{p-1}+C_p+C_{\lambda,p,H})(2a+1)K(T)(C_2+1)$. ■

Theorem 3.6 *Suppose that Assumption 3.1 holds and $\xi\in L^p(\Omega\to\mathbb{R},\ \mathscr{F}_0,\ \mathbb{P})$ with $p\ge\theta$ and $p>1/H$. Then the equation (1.2) has a unique solution $X\in L^p(\Omega;C([0,T];\mathbb{R}^d))$.*

Proof. We split the proof into two step.

Step one: (Existence) We first prove that $(X_k)_{k\ge1}$ is a Cauchy sequence in $L^p(\Omega;C([0,T];\mathbb{R}^d))$ with $p\ge\theta$ and $p>\frac{1}{H}$.

In fact, for $m>k\ge1$, it is routine to obtain

$$
\begin{aligned}
&\sup_{0\le s\le t}|X_m(s)-X_k(s)|^p \\
&\le 3^{p-1}\sup_{0\le s\le t}\Big|\int_0^s\Big(b(u,X_m(u-\frac{1}{m}),\mathscr{L}(X_m(u-\frac{1}{m}))) \\
&\qquad\qquad\qquad -b(u,X_k(u-\frac{1}{k}),\mathscr{L}(X_k(u-\frac{1}{k})))\Big)du\Big|^p \\
&\quad+3^{p-1}\sup_{0\le s\le t}\Big|\int_0^s\Big(\sigma_W(u,X_m(u-\frac{1}{m}),\mathscr{L}(X_m(u-\frac{1}{m}))) \\
&\qquad\qquad\qquad -\sigma_W(u,X_k(u-\frac{1}{k}),\mathscr{L}(X_k(u-\frac{1}{k})))\Big)dW_u\Big|^p \\
&\quad+3^{p-1}\sup_{0\le s\le t}\Big|\int_0^s\Big(\sigma_H(u,\mathscr{L}(X_m(u-\frac{1}{m}))) \\
&\qquad\qquad\qquad -\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))\Big)dB_u^H\Big|^p \\
&=:F_1+F_2+F_3.
\end{aligned} \tag{3.19}
$$

By Hölder inequality, Assumption 3.1 and Lemma 3.5, we have

$$
\begin{aligned}
\mathbb{E}F_1 = 3^{p-1}\mathbb{E}\Big(\sup_{0\le s\le t}\Big|\int_0^s \Big(& b(u, X_m(u-\frac{1}{m}), \mathscr{L}(X_m(u-\frac{1}{m}))) \\
& - b(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))\Big)du\Big|^p\Big) \\
\le (3T)^{p-1}\mathbb{E}\int_0^t & |b(u, X_m(u-\frac{1}{m}), \mathscr{L}(X_m(u-\frac{1}{m}))) \\
& - b(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))|^p du \\
\le (6T)^{p-1}\mathbb{E}\int_0^t & \Big|b(u, X_m(u-\frac{1}{m}), \mathscr{L}(X_m(u-\frac{1}{m}))) \\
& - b(u, X_k(u-\frac{1}{m}), \mathscr{L}(X_k(u-\frac{1}{m})))\Big|^p du \\
+ (6T)^{p-1}\mathbb{E}\int_0^t & \Big|b(u, X_k(u-\frac{1}{m}), \mathscr{L}(X_k(u-\frac{1}{m}))) \\
& - b(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))\Big|^p du \\
\le (6T)^{p-1}\int_0^t & K(u)\psi(2\mathbb{E}|X_m(u-\frac{1}{m}) - X_k(u-\frac{1}{m})|^p)du \\
+ (6T)^{p-1}\int_0^t & K(u)\psi\Big(2C_3\Big((\frac{1}{k}-\frac{1}{m})^p + (\frac{1}{k}-\frac{1}{m})^{\frac{p}{2}} + (\frac{1}{k}-\frac{1}{m})^{pH}\Big)\Big)du.
\end{aligned} \tag{3.20}
$$

By Burkholder-Davis-Gundy inequality, Hölder inequality, Assumption 3.1 and Lemma 3.5, we obtain

$$
\begin{aligned}
\mathbb{E}F_2 = 3^{p-1}\mathbb{E}\Big(\sup_{0\le s\le t}\Big|\int_0^s \Big(& \sigma_W(u, X_m(u-\frac{1}{m}), \mathscr{L}(X_m(u-\frac{1}{m}))) \\
& - \sigma_W(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))\Big)dW_u\Big|^p\Big) \\
\le C_p\mathbb{E}\Big(\int_0^t & \|\sigma_W(u, X_m(u-\frac{1}{m}), \mathscr{L}(X_m(u-\frac{1}{m}))) \\
& - \sigma_W(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))\|^2 du\Big)^{\frac{p}{2}} \\
\le C_p T^{\frac{p}{2}-1}\mathbb{E}\int_0^t & \|\sigma_W(u, X_m(u-\frac{1}{m}), \mathscr{L}(X_m(u-\frac{1}{m}))) \\
& - \sigma_W(u, X_k(u-\frac{1}{k}), \mathscr{L}(X_k(u-\frac{1}{k})))\|^p du
\end{aligned} \tag{3.21}
$$

$$
\begin{aligned}
&\leq C_pT^{\frac{p}{2}-1}\mathbb{E}\int_0^t \|\sigma_W(u,X_m(u-\frac{1}{m}),\mathscr{L}(X_m(u-\frac{1}{m})))\\
&\qquad -\sigma_W(u,X_k(u-\frac{1}{m}),\mathscr{L}(X_k(u-\frac{1}{m})))\|^p du\\
&\quad + C_pT^{\frac{p}{2}-1}\mathbb{E}\int_0^t \|\sigma_W(u,X_k(u-\frac{1}{m}),\mathscr{L}(X_k(u-\frac{1}{m})))\\
&\qquad -\sigma_W(u,X_k(u-\frac{1}{k}),\mathscr{L}(X_k(u-\frac{1}{k})))\|^p du\\
&\leq C_pT^{\frac{p}{2}-1}\int_0^t K(u)\psi\Big(2\mathbb{E}|X_m(u-\frac{1}{m})-X_k(u-\frac{1}{m})|^p\Big)du\\
&\quad + C_pT^{\frac{p}{2}-1}\int_0^t K(u)\psi\Big(2C_3\Big((\frac{1}{k}-\frac{1}{m})^p+(\frac{1}{k}-\frac{1}{m})^{\frac{p}{2}}+(\frac{1}{k}-\frac{1}{m})^{pH}\Big)\Big)du.
\end{aligned}
$$

By Assumption 3.1 and Lemma 3.5, we have

$$
\begin{aligned}
\mathbb{E}F_3 &= 3^{p-1}\mathbb{E}\Big(\sup_{0\leq s\leq t}\Big|\int_0^s \Big(\sigma_H(u,\mathscr{L}(X_m(u-\frac{1}{m})))-\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))\Big)dB_u^H\Big|^p\Big)\\
&\leq C_{\lambda,p,H}T^{pH-1}\int_0^t \|\sigma_H(u,\mathscr{L}(X_m(u-\frac{1}{m})))-\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))\|^p du\\
&\leq C_{\lambda,p,H}T^{pH-1}\int_0^t \|\sigma_H(u,\mathscr{L}(X_m(u-\frac{1}{m})))-\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{m})))\|^p du\\
&\quad + C_{\lambda,p,H}T^{pH-1}\int_0^t \|\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{m})))-\sigma_H(u,\mathscr{L}(X_k(u-\frac{1}{k})))\|^p du\\
&\leq C_{\lambda,p,H}T^{pH-1}\int_0^t K(u)\psi(\mathbb{E}|X_m(u-\frac{1}{m})-X_k(u-\frac{1}{m})|^p)du\\
&\quad + C_{\lambda,p,H}T^{pH-1}\int_0^t K(u)\psi\Big(C_3\Big((\frac{1}{k}-\frac{1}{m})^p+(\frac{1}{k}-\frac{1}{m})^{\frac{p}{2}}+(\frac{1}{k}-\frac{1}{m})^{pH}\Big)\Big)du.
\end{aligned}
\tag{3.22}
$$

Hence, we have

$$
\begin{aligned}
&\mathbb{E}\Big(\sup_{0\leq s\leq t}|X_m(s)-X_k(s)|^p\Big)\\
&\leq ((6T)^{p-1}+C_pT^{\frac{p}{2}-1}+C_{\lambda,p,H}T^{pH-1})\int_0^t K(u)\psi(2\mathbb{E}|X_m(u-\frac{1}{m})-X_k(u-\frac{1}{m})|^p)du\\
&\quad + (6^{p-1}T^p+C_pT^{\frac{p}{2}}+C_{\lambda,p,H}T^{pH})K(T)\psi\Big(2C_3\Big((\frac{1}{k}-\frac{1}{m})^p+(\frac{1}{k}-\frac{1}{m})^{\frac{p}{2}}+(\frac{1}{k}-\frac{1}{m})^{pH}\Big)\Big).
\end{aligned}
\tag{3.23}
$$

Taking limit as $m,k\to\infty$, using the fact that $\psi(0)=0$, we obtain for every $\epsilon>0$,

$$
\begin{aligned}
Z(t) &\leq ((6T)^{p-1}+C_pT^{\frac{p}{2}-1}+C_{\lambda,p,H}T^{pH-1})\int_0^t K(u)\psi(2Z(u))du\\
&\leq \epsilon+((6T)^{p-1}+C_pT^{\frac{p}{2}-1}+C_{\lambda,p,H}T^{pH-1})K(T)\int_0^t \psi(2Z(u))du.
\end{aligned}
\tag{3.24}
$$

where $Z(t) = \lim_{m,k\to\infty} \mathbb{E}(\sup_{0\le s\le t} |X_m(s) - X_k(s)|^p)$. Hence, Bihari inequality yields

$$Z(t) \le \frac{1}{2}G^{-1}[G(2\epsilon) + 2((6T)^{p-1} + C_p T^{\frac{p}{2}-1} + C_{\lambda,p,H}T^{pH-1})K(T)t],$$

where $G(2\epsilon) + 2((6T)^{p-1} + C_p T^{\frac{p}{2}-1} + C_{\lambda,p,H}T^{pH-1})K(T)t \in Dom(G^{-1})$, G^{-1} is the inverse function of $G(\cdot)$ and

$$G(v) = \int_1^v \frac{ds}{\psi(s)}, \quad v > 0$$

By Assumption 3.1, one sees that $\lim_{\epsilon\downarrow 0} G(\epsilon) = -\infty$ and $Dom(G^{-1}) = (-\infty, G(\infty))$. Letting $\epsilon \to 0$ gives $Z(t) = 0$, i.e.,

$$\mathbb{E}(\sup_{0\le s\le t} |X_m(s) - X_k(s)|^p) \to 0, \quad as \quad m, k \to \infty. \tag{3.25}$$

Consequently, $(X_k)_{k\ge 1}$ is a Cauchy sequence in $L^p(\Omega; C([0,T]; \mathbb{R}^d))$ with $p \ge \theta$ and $p > \frac{1}{H}$, and then the limit, denoted by X. Therefore, putting $m \to \infty$ in (3.25), we conclude

$$\lim_{k\to\infty} \mathbb{E}(\sup_{0\le s\le T} |X(s) - X_k(s)|^p) = 0. \tag{3.26}$$

Now, we will prove that the $X(t)$ is a solution to (1.2). For all $0 \le t \le T$, we have

$$\begin{aligned}\mathbb{E}|X(t) - X_k(t - \frac{1}{k})|^p &= \mathbb{E}|X(t) - X_k(t) + X_k(t) - X_k(t - \frac{1}{k})|^p \\ &\le 2^{p-1}\mathbb{E}|X(t) - X_k(t)|^p + 2^{p-1}\mathbb{E}|X_k(t) - X_k(t - \frac{1}{k})|^p.\end{aligned} \tag{3.27}$$

By Lemma 3.5 and (3.26), we obtain

$$\mathbb{E}|X(t) - X_k(t - \frac{1}{k}))|^p \to 0, \quad as \quad k \to \infty.$$

Therefore, as $k \to \infty$ in (3.1), we obtain

$$\begin{aligned}X(t) =&\xi + \int_0^t b(s, X(t), \mathscr{L}(X(s)))ds + \int_0^t \sigma_W(s, X(s), \mathscr{L}(X(s)))dW_s \\ &+ \int_0^t \sigma_H(s, \mathscr{L}(X(s)))dB_s^H, t \in [0,T],\end{aligned} \tag{3.28}$$

that indicates $X(t)$ is a solution to (1.2). Hence, the proof of the existence is completed.

Step Two: (Uniqueness) Let$X(t), Y(t)$ be two solutions for (1.2) on the same probability space with $X(0) = Y(0)$, then, by inequality (3.3),

$$\begin{aligned}&\mathbb{E}\left(\sup_{0\le s\le t} |X(s) - Y(s)|^p\right) \\ &\le 3^{p-1}\mathbb{E}\left(\sup_{0\le s\le t}\left|\int_0^s \Big(b(u, X(u), \mathscr{L}(X(u))) - b(u, Y(u), \mathscr{L}(Y(u)))\Big)du\right|^p\right) \\ &\quad + 3^{p-1}\mathbb{E}\left(\sup_{0\le s\le t}\left|\int_0^s \Big(\sigma_W(u, X(u), \mathscr{L}(X(u))) - \sigma_W(u, Y(u), \mathscr{L}(Y(u)))\Big)dW_u\right|^p\right) \\ &\quad + 3^{p-1}\mathbb{E}\left(\sup_{0\le s\le t}\left|\int_0^s \Big(\sigma_H(u, \mathscr{L}(X(u))) - \sigma_H(u, \mathscr{L}(Y(u)))\Big)dB_u^H\right|^p\right) \\ &=: H_1 + H_2 + H_3.\end{aligned} \tag{3.29}$$

By Hölder inequality and Assumption 3.1, we have

$$\begin{aligned} H_1 &= 3^{p-1}\mathbb{E}\left(\sup_{0\le s\le t}\left|\int_0^s \Big(b(u,X(u),\mathscr{L}(X(u))) - b(u,Y(u),\mathscr{L}(Y(u)))\Big)du\right|^p\right) \\ &\le (3T)^{p-1}\mathbb{E}\int_0^t |b(s,X(s),\mathscr{L}(X(s))) - b(s,Y(s),\mathscr{L}(Y(s)))|^p ds \\ &\le (3T)^{p-1}\mathbb{E}\int_0^t K(s)\psi\Big(|X(s)-Y(s)|^p + \mathbb{W}_\theta^p\Big(\mathscr{L}(X(s)),\mathscr{L}(Y(s))\Big)\Big)ds \\ &\le (3T)^{p-1}K(T)\int_0^t \psi\left(2\mathbb{E}\left(\sup_{0\le u\le s}|X(u)-Y(u)|^p\right)\right)ds. \end{aligned} \tag{3.30}$$

It comes from Burkholder-Davis-Gundy inequality, Hölder inequality and Assumption 3.1, we obtain

$$\begin{aligned} H_2 &= 3^{p-1}\mathbb{E}\Big(\sup_{0\le s\le t}\Big|\int_0^s \Big(\sigma_W(u,X(u),\mathscr{L}(X(u))) - \sigma_W(u,Y(u),\mathscr{L}(Y(u)))\Big)dW_u\Big|^p\Big) \\ &\le C_p\mathbb{E}\left(\int_0^t \|\sigma_W(s,X(s),\mathscr{L}(X(s))) - \sigma_W(s,Y(s),\mathscr{L}(Y(s)))\|^2 ds\right)^{\frac{p}{2}} \\ &\le C_pT^{\frac{p}{2}-1}\mathbb{E}\int_0^t \|\sigma_W(s,X(s),\mathscr{L}(X(s))) - \sigma_W(s,Y(s),\mathscr{L}(Y(s)))\|^p ds \\ &\le C_pT^{\frac{p}{2}-1}\mathbb{E}\int_0^t K(s)\psi\Big(|X(s)-Y(s)|^p + \mathbb{W}_\theta^p\Big(\mathscr{L}(X(s)),\mathscr{L}(Y(s))\Big)\Big)ds \\ &\le C_pT^{\frac{p}{2}-1}K(T)\int_0^t \psi\Big(2\mathbb{E}\Big(\sup_{0\le u\le s}|X(u)-Y(u)|^p\Big)\Big)ds. \end{aligned} \tag{3.31}$$

By Assumption 3.1, we have that

$$\begin{aligned} H_3 &= 3^{p-1}\mathbb{E}\left(\sup_{0\le s\le t}\left|\int_0^s \Big(\sigma_H(u,\mathscr{L}(X(u))) - \sigma_H(u,\mathscr{L}(Y(u)))\Big)dB_u^H\right|^p\right) \\ &\le C_{\lambda,p,H}T^{pH-1}\int_0^t \|\sigma_H(s,\mathscr{L}(X(s))) - \sigma_H(s,\mathscr{L}(Y(s)))\|^p ds \\ &\le C_{\lambda,p,H}T^{pH-1}\int_0^t K(s)\psi\Big(W_\theta^p\Big(\mathscr{L}(X(s)),\mathscr{L}(Y(s))\Big)\Big)ds \\ &\le C_{\lambda,p,H}T^{pH-1}K(T)\int_0^t \psi\Big(\mathbb{E}\Big(\sup_{0\le u\le s}|X(u)-Y(u)|^p\Big)\Big)ds. \end{aligned} \tag{3.32}$$

Therefore,

$$\begin{aligned} \mathbb{E}\Big(\sup_{0\le s\le t}|X(s)-Y(s)|^p\Big) &\le ((3T)^{p-1} + C_pT^{\frac{p}{2}-1} + C_{\lambda,p,H}T^{pH-1})K(T) \\ &\quad\times\int_0^t \psi\Big(2\mathbb{E}\Big(\sup_{0\le u\le s}|X(u)-Y(u)|^p\Big)\Big)ds. \end{aligned} \tag{3.33}$$

Then, the Bihari inequality implies that $X(t) = Y(t)$, $t \in [0,T]$, $\mathbb{P}-a.s.$ This completes the proof. ■

4 Stochastic averaging principle

In this section, our aim is to establish a stochastic averaging principle for the following stochastic integral equations

$$\begin{aligned} X^\epsilon(t) = \xi &+ \int_0^t b(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s)))ds + \int_0^t \sigma_W(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s)))dW_s \\ &+ \int_0^t \sigma_H(\frac{s}{\epsilon}, \mathscr{L}(X^\epsilon(s)))dB_s^H, \quad t \in [0,T], \end{aligned} \tag{4.1}$$

where $\epsilon \in (0, \epsilon_0]$ is a positive parameter with $\epsilon_0 > 0$ being fixed. The coefficients of (4.1) fulfill the same conditions as in (1.2). Thus, Equation (4.1) has a unique solution $X^\epsilon(t), t \in [0,T]$.

Our objective is to show that the solution $X^\epsilon(t)$, $t \in [0,T]$ could be approximated in certain sense by the solution $\bar{X}(t)$,$t \in [0,T]$ of the following averaged equation

$$\begin{aligned} \bar{X}(t) = \xi &+ \int_0^t \bar{b}(\bar{X}(s), \mathscr{L}(\bar{X}(s)))ds + \int_0^t \bar{\sigma}_W(\bar{X}(s), \mathscr{L}(\bar{X}(s)))dW_s \\ &+ \int_0^t \bar{\sigma}_H(\mathscr{L}(\bar{X}(s)))dB_s^H, \quad t \in [0,T], \end{aligned} \tag{4.2}$$

where $\bar{b} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$, $\bar{\sigma}_W : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^r$, and $\bar{\sigma}_H : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^m$ is Borel measurable function.

Remark 4.1 *The averaging principle is also applicable to the following system*

$$\begin{aligned} dX(t) =& \epsilon b(t, X(t), \mathscr{L}(X(t)))dt + \sqrt{\epsilon}\sigma_W(t, X(t), \mathscr{L}(X(t)))dW_t \\ &+ \epsilon^H \sigma_H(t, \mathscr{L}(X(t)))dB_t^H, \end{aligned} \tag{4.3}$$

where $0 < \epsilon \leq 1$. With the time scaling $t \to \frac{t}{\epsilon}$, denote by $\Phi_\epsilon(t) := X(\frac{t}{\epsilon})$, $W_\epsilon(t) := \sqrt{\epsilon} W_{\frac{t}{\epsilon}}$ and $B_\epsilon^H(t) := \epsilon^H B_{\frac{t}{\epsilon}}^H$ for all $t \in \mathbb{R}$, we transform equation (4.3) to

$$\begin{aligned} d\Phi_\epsilon(t) =& b(\frac{t}{\epsilon}, \Phi_\epsilon(t), \mathscr{L}(\Phi_\epsilon(t)))dt + \sigma_W(\frac{t}{\epsilon}, \Phi_\epsilon(t), \mathscr{L}(\Phi_\epsilon(t)))dW_\epsilon(t) \\ &+ \sigma_H(\frac{t}{\epsilon}, \mathscr{L}(\Phi_\epsilon(t)))dB_\epsilon^H(t). \end{aligned}$$

Then we can consider the following equation

$$\begin{aligned} d\tilde{X}_\epsilon(t) =& b(\frac{t}{\epsilon}, \tilde{X}_\epsilon(t), \mathscr{L}(\tilde{X}_\epsilon(t)))dt + \sigma_W(\frac{t}{\epsilon}, \tilde{X}\epsilon(t), \mathscr{L}(\tilde{X}_\epsilon(t)))dW(t) \\ &+ \sigma_H(\frac{t}{\epsilon}, \mathscr{L}(\tilde{X}_\epsilon(t)))dB^H(t). \end{aligned}$$

To ensure the DDSDEs (4.2) also has a unique solution $\bar{X}_t$, $t \in [0,T]$, we will make use of the following assumptions on the coefficients. Moreover, we assume $K(t)$ is bounded.

Assumption 4.1 *(Averaging condition) There is a bounded positive function $\varphi : (0, \infty) \to (0, \infty)$ with $\lim_{T\to\infty} \varphi(T) = 0$, such that for any $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$*

$$\sup_{t\geq 0} \left| \frac{1}{T} \int_t^{t+T} [b(s, x, \mu) - \bar{b}(x, \mu)] ds \right|^2 \leq \varphi(T)(1 + |x|^2 + \mathbb{W}_2^2(\mu, \delta_0)),$$

$$\sup_{t\geq 0} \frac{1}{T} \int_t^{t+T} \|\sigma_W(s, x, \mu) - \bar{\sigma}_W(x, \mu)\|^2 ds \leq \varphi(T)(1 + |x|^2 + \mathbb{W}_2^2(\mu, \delta_0)),$$

$$\sup_{t\geq 0} \frac{1}{T} \int_t^{t+T} \|\sigma_H(s, \mu) - \bar{\sigma}_H(\mu)\|^2 ds \leq \varphi(T)(1 + \mathbb{W}_2^2(\mu, \delta_0)).$$

Remark 4.2 *(i) Noting that*

$$\sup_{t\geq 0} \left| \frac{1}{T} \int_t^{t+T} [b(s, x, \mu) - \bar{b}(x, \mu)] ds \right|^2 \leq \sup_{t\geq 0} \frac{1}{T} \int_t^{t+T} |b(s, x, \mu) - \bar{b}(x, \mu)|^2 ds,$$

this shows that Assumption 4.1 is weaker than the following averaging condition:

$$\sup_{t\geq 0} \frac{1}{T} \int_t^{t+T} |b(s, x, \mu) - \bar{b}(x, \mu)|^2 ds \leq \varphi(T)(1 + |x|^2 + \mathbb{W}_2^2(\mu, \delta_0)).$$

Hence, we need to overcome the difficulties with the weaker condition to obtain the averaging principle for the concerned DDSDEs.

(ii) For any $x, y \in \mathbb{R}^d$, and any $T > 0$, we have

$$\begin{aligned}
&\left|\bar{b}(x, \mu) - \bar{b}(y, \nu)\right|^2 \\
&\leq 3 \left| \frac{1}{T} \int_0^T [b(s, x, \mu) - \bar{b}(x, \mu)] ds \right|^2 + 3 \left| \frac{1}{T} \int_0^T [b(s, y, \nu) - \bar{b}(y, \nu)] ds \right|^2 \\
&\quad + 3 \left| \frac{1}{T} \int_0^T [b(s, x, \mu) - b(s, y, \nu)] ds \right|^2 \\
&\leq 3\varphi(T)(2 + |x|^2 + |y|^2 + \mathbb{W}_2^2(\mu, \delta_0) + \mathbb{W}_2^2(\nu, \delta_0)) + 3K(T)\psi(|x - y|^2 + \mathbb{W}_2^2(\mu, \nu)),
\end{aligned} \tag{4.4}$$

$$\begin{aligned}
|\bar{b}(0, \delta_0)|^2 &\leq 2 \left| \frac{1}{T} \int_0^T [b(s, 0, \delta_0) - \bar{b}(0, \delta_0)] ds \right|^2 + 2 \left| \frac{1}{T} \int_0^T b(s, 0, \delta_0) ds \right|^2 \\
&\leq 2\varphi(T) + 2K(T).
\end{aligned}$$

Taking $T \to \infty$, because K(t) is bounded, there exists a constant L, such that

$$|\bar{b}(x, \mu) - \bar{b}(y, \nu)|^2 \leq L\psi(|x - y|^2 + \mathbb{W}_2^2(\mu, \nu)), \quad |\bar{b}(0, \delta_0)|^2 \leq L.$$

Similarly, for any x, $y \in \mathbb{R}^d$, and any $T > 0$, we have

$$\begin{aligned}
&\|\bar{\sigma}_W(x, \mu) - \bar{\sigma}_W(y, \nu)\|^2 \\
&\leq 3\varphi(T)(2 + |x|^2 + |y|^2 + \mathbb{W}_2^2(\mu, \delta_0) + \mathbb{W}_2^2(\nu, \delta_0)) + 3K(T)\psi(|x - y|^2 + \mathbb{W}_2^2(\mu, \nu)),
\end{aligned}$$

$$\|\bar{\sigma}_H(\mu)-\bar{\sigma}_H(\nu)\|^2 \le 3\varphi(T)(2+\mathbb{W}_2^2(\mu,\delta_0)+\mathbb{W}_2^2(\nu,\delta_0))+3K(T)\psi(\mathbb{W}_2^2(\mu,\nu)),$$
$$\|\bar{\sigma}_W(0,\delta_0)\|^2 \le 2\varphi(T)+2K(T), \quad \|\bar{\sigma}_H(\delta_0)\|^2 \le 2\varphi(T)+2K(T).$$

Taking $T\to\infty$, we have

$$\|\bar{\sigma}_W(x,\mu)-\bar{\sigma}_W(y,\nu)\|^2 \le L\psi(|x-y|^2+\mathbb{W}_2^2(\mu,\nu)), \quad \|\bar{\sigma}_H(\mu)-\bar{\sigma}_H(\nu)\|^2 \le L\psi(\mathbb{W}_2^2(\mu,\nu)),$$
$$\|\bar{\sigma}_W(0,\delta_0)\|^2 \le L, \quad \|\bar{\sigma}_H(\delta_0)\|^2 \le L.$$

Thus, the coefficients $\bar{b},\bar{\sigma}_w,\bar{\sigma}_H$ satisfy the Assumptions 3.1, Therefore, there is a unique solution $\bar{X}_t$ to the averaged equationthe (4.2).

Remark 4.3 *Using the similar proof methods as Lemma 3.4, Lemma 3.5 and Theorem 3.6, we have that for any initial value $X^\epsilon(0)=\xi$ satisfying $\mathbb{E}|\xi|^2<\infty$, under Assumption 3.1, there exists an unique solution $X^\epsilon(t)\in L^2(\Omega;\mathbb{R}^d)$ for equation (4.1). Moreover, this solution satisfies*

$$\mathbb{E}\bigg[\sup_{0\le t\le T}|X^\epsilon(t)|^2\bigg] \le C_{4,\epsilon},$$

and

$$\mathbb{E}|X^\epsilon(t)-X^\epsilon(s)|^2 \le C_{5,\epsilon}K(\frac{T}{\epsilon})(|t-s|+|t-s|^2+|t-s|^{2H}),$$

where $C_{4,\epsilon}$, $C_{5,\epsilon}$ are two positive constants depend on ϵ.

Lemma 4.2 *Suppose that Assumptions 3.1 and 4.1 hold and $\mathbb{E}|\xi|^2<+\infty$. Then, we have*

$$\lim_{\epsilon\to 0}\mathbb{E}\left(\sup_{0\le t\le T}\left|\int_0^t\left(b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))\right)ds\right|^2\right)=0. \tag{4.5}$$

Proof. Let $\{t_1,t_2,\cdots,t_N\}$ be a partition of $[0,T]$:

$$t_i=i\sqrt{\epsilon},\ 0\le i\le N-1,\ 0<T-t_{N-1}\le\sqrt{\epsilon},\ t_N=T.$$

Then, it is not difficult to obtain that $T\le N\sqrt{\epsilon}<T+\sqrt{\epsilon}$. We have

$$\begin{aligned}&\left|\int_0^t[b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))]ds\right|^2\\ &\le N\left|\int_{[\frac{t}{\sqrt{\epsilon}}]\sqrt{\epsilon}}^t[b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))]ds\right|^2+N\sum_{i=0}^{N-2}|X_i|^2,\end{aligned} \tag{4.6}$$

where $X_i:=\int_{t_i}^{t_{i+1}}[b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))]ds$. By Hölder inequality, Assumptions 3.1 and Remark 4.2, we have

$$\begin{aligned}&\left|\int_{[\frac{t}{\sqrt{\epsilon}}]\sqrt{\epsilon}}^t[b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))]ds\right|^2\\ &\le 2(t-[\frac{t}{\sqrt{\epsilon}}]\sqrt{\epsilon})\int_{[\frac{t}{\sqrt{\epsilon}}]\sqrt{\epsilon}}^t\left(|b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))|^2+|\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))|^2\right)ds\\ &\le 4(t-[\frac{t}{\sqrt{\epsilon}}]\sqrt{\epsilon})\int_{[\frac{t}{\sqrt{\epsilon}}]\sqrt{\epsilon}}^t(K(\frac{s}{\epsilon})+L)\left[\psi\left(|X^\epsilon(s)|^2+\mathbb{E}|X^\epsilon(s)|^2\right)+1\right]ds\\ &\le 4(K(\frac{T}{\epsilon})+L)\epsilon(a+1+a\sup_{0\le t\le T}|X^\epsilon(t)|^2+a\mathbb{E}(\sup_{0\le t\le T}|X^\epsilon(t)|^2)).\end{aligned} \tag{4.7}$$

By (4.7) and Remark 4.3, we get

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le t\le T}\left|\int_0^t [b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))]ds\right|^2\right)\\
&\le C\epsilon N(K(\frac{T}{\epsilon})+L)+N\mathbb{E}\sum_{i=0}^{N-2}|X_i|^2 \\
&\le C(K(\frac{T}{\epsilon})+L)\sqrt{\epsilon}(T+\sqrt{\epsilon})+N\sum_{i=0}^{N-2}\mathbb{E}|X_i|^2.
\end{aligned}
\tag{4.8}
$$

By Assumptions 3.1, 4.1 and Remark 4.2, we have

$$
\begin{aligned}
|X_i|^2 &= \left|\int_{t_i}^{t_{i+1}}[b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))]ds\right|^2\\
&\le 3\left|\int_{t_i}^{t_{i+1}}[b(\frac{s}{\epsilon},X^\epsilon(t_i),\mathscr{L}(X^\epsilon(t_i)))-\bar{b}(X^\epsilon(t_i),\mathscr{L}(X^\epsilon(t_i)))]ds\right|^2\\
&\quad+3\left|\int_{t_i}^{t_{i+1}}[b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-b(\frac{s}{\epsilon},X^\epsilon(t_i),\mathscr{L}(X^\epsilon(t_i)))]ds\right|^2\\
&\quad+3\left|\int_{t_i}^{t_{i+1}}[\bar{b}(X^\epsilon(t_i),\mathscr{L}(X^\epsilon(t_i)))-\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))]ds\right|^2\\
&\le 3\left|\epsilon\int_{\frac{t_i}{\epsilon}}^{\frac{t_{i+1}}{\epsilon}}[b(s,X^\epsilon(t_i),\mathscr{L}(X^\epsilon(t_i)))-\bar{b}(X^\epsilon(t_i),\mathscr{L}(X^\epsilon(t_i)))]ds\right|^2\\
&\quad+3\sqrt{\epsilon}(K(\frac{T}{\epsilon})+L)\int_{t_i}^{t_{i+1}}\psi(|X^\epsilon(s)-X^\epsilon(t_i)|^2+\mathbb{E}|X^\epsilon(s)-X^\epsilon(t_i)|^2)ds\\
&\le 3\epsilon\varphi(\frac{1}{\sqrt{\epsilon}})\left(1+\sup_{t\in[0,T]}|X^\epsilon(t)|+\mathbb{E}(\sup_{t\in[0,T]}|X^\epsilon(t)|^2)\right)\\
&\quad+3\sqrt{\epsilon}(K(\frac{T}{\epsilon})+L)\int_{t_i}^{t_{i+1}}\psi(|X^\epsilon(s)-X^\epsilon(t_i)|^2+\mathbb{E}|X^\epsilon(s)-X^\epsilon(t_i)|^2)ds.
\end{aligned}
\tag{4.9}
$$

Hence,

$$
\begin{aligned}
&N\sum_{i=0}^{N-2}\mathbb{E}|X_i|^2\\
&\le 3N\epsilon\sum_{i=0}^{N-2}\varphi(\frac{1}{\sqrt{\epsilon}})(1+2\mathbb{E}(\sup_{t\in[0,T]}|X^\epsilon(t)|^2))\\
&\quad+3N\sqrt{\epsilon}(K(\frac{T}{\epsilon})+L)\sum_{i=0}^{N-2}\int_{t_i}^{t_{i+1}}\psi(2\mathbb{E}|X^\epsilon(s)-X^\epsilon(t_i)|^2)ds
\end{aligned}
\tag{4.10}
$$

$$
\begin{aligned}
&\le 3N\epsilon \sum_{i=0}^{N-2} \varphi(\frac{1}{\sqrt{\epsilon}})(1 + 2\mathbb{E}(\sup_{t\in[0,T]} |X^\epsilon(t)|^2)) \\
&\quad + 3N^2\sqrt{\epsilon}(K(\frac{T}{\epsilon}) + L) \int_{t_i}^{t_{i+1}} \psi\left(2C_{5,\epsilon}K(\frac{s}{\epsilon})((s-t_i) + (s-t_i)^2 + (s-t_i)^{2H})\right) ds \\
&\le CN^2\epsilon \left[\varphi(\frac{1}{\sqrt{\epsilon}}) + (K(\frac{T}{\epsilon}) + L)\psi\left(2C_{5,\epsilon}K(\frac{T}{\epsilon})(\sqrt{\epsilon} + \epsilon + \epsilon^H)\right)\right] \\
&\le C(T + \sqrt{\epsilon})^2 \left[\varphi(\frac{1}{\sqrt{\epsilon}}) + (K(\frac{T}{\epsilon}) + L)\psi\left(2C_{5,\epsilon}K(\frac{T}{\epsilon})(\sqrt{\epsilon} + \epsilon + \epsilon^H)\right)\right].
\end{aligned}
$$

Combining (4.10) with (4.8), we get

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{0\le t\le T}\left|\int_0^t [b(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \bar{b}(X^\epsilon(s), \mathscr{L}(X^\epsilon(s)))]ds\right|^2\right) \\
&\le C(K(\frac{T}{\epsilon}) + L)\sqrt{\epsilon}(T + \sqrt{\epsilon}) \\
&\quad + C(T + \sqrt{\epsilon})^2 \left[\varphi(\frac{1}{\sqrt{\epsilon}}) + (K(\frac{T}{\epsilon}) + L)\psi\left(2C_{5,\epsilon}K(\frac{T}{\epsilon})(\sqrt{\epsilon} + \epsilon + \epsilon^H)\right)\right] \to 0,
\end{aligned} \tag{4.11}
$$

as ϵ tends to zero.

■

Lemma 4.3 *Suppose that Assumptions 3.1 and 4.1 hold and $\mathbb{E}|\xi|^2 < +\infty$. Then, we have*

$$
\lim_{\epsilon\to 0} \mathbb{E}\int_0^T \|\sigma_W(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_W(X^\epsilon(s), \mathscr{L}(X^\epsilon(s)))\|^2 ds = 0. \tag{4.12}
$$

Proof. It's similar to Lemma 4.2, let $\{t_1, t_2, \cdots, t_N\}$ satisfy

$$
t_i = i\sqrt{\epsilon},\ 0 \le i \le N-1,\ 0 < T - t_{N-1} \le \sqrt{\epsilon},\ t_N = T.
$$

Denote

$$
Y_i := \int_{t_i}^{t_{i+1}} \|\sigma_W(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_W(X^\epsilon(s), \mathscr{L}(X^\epsilon(s)))\|^2 ds.
$$

Thus, we have

$$
\sum_{i=0}^{N-1} \mathbb{E}Y_i = \mathbb{E}\int_0^T \|\sigma_W(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_W(X^\epsilon(s), \mathscr{L}(X^\epsilon(s)))\|^2 ds.
$$

By Assumptions 4.1, 3.1, and Remark 4.2, we obtain

$$
\begin{aligned}
Y_i &= \int_{t_i}^{t_{i+1}} \|\sigma_W(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_W(X^\epsilon(s), \mathscr{L}(X^\epsilon(s)))\|^2 ds \\
&\le 3\int_{t_i}^{t_{i+1}} \|\sigma_W(\frac{s}{\epsilon}, X^\epsilon(t_i), \mathscr{L}(X^\epsilon(t_i))) - \bar{\sigma}_W(X^\epsilon(t_i), \mathscr{L}(X^\epsilon(t_i)))\|^2 ds \\
&\quad + 3\int_{t_i}^{t_{i+1}} \|\sigma_W(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \sigma_W(\frac{s}{\epsilon}, X^\epsilon(t_i), \mathscr{L}(X^\epsilon(t_i)))\|^2 ds \\
&\quad + 3\int_{t_i}^{t_{i+1}} \|\bar{\sigma}_W(X^\epsilon(t_i), \mathscr{L}(X^\epsilon(t_i))) - \bar{\sigma}_W(X^\epsilon(s), \mathscr{L}(X^\epsilon(s)))\|^2 ds \\
&\le 3\epsilon\int_{\frac{t_i}{\epsilon}}^{\frac{t_{i+1}}{\epsilon}} \|\sigma_W(s, X^\epsilon(t_i), \mathscr{L}(X^\epsilon(t_i))) - \bar{\sigma}_W(X^\epsilon(t_i), \mathscr{L}(X^\epsilon(t_i)))\|^2 ds \\
&\quad + 3(K(\frac{T}{\epsilon}) + L)\int_{t_i}^{t_{i+1}} \psi\left(|X^\epsilon(s) - X^\epsilon(t_i)|^2 + \mathbb{E}|X^\epsilon(s) - X^\epsilon(t_i)|^2\right) ds \\
&\le 3\sqrt{\epsilon}\varphi(\frac{1}{\sqrt{\epsilon}})(1 + \sup_{t\in[0,T]}|X^\epsilon(t)|^2 + \mathbb{E}(\sup_{t\in[0,T]}|X^\epsilon(t)|^2)) \\
&\quad + 3(K(\frac{T}{\epsilon}) + L)\int_{t_i}^{t_{i+1}} \psi\left(|X^\epsilon(s) - X^\epsilon(t_i)|^2 + \mathbb{E}|X^\epsilon(s) - X^\epsilon(t_i)|^2\right) ds.
\end{aligned}
\tag{4.13}
$$

By Remark 4.3, we have

$$
\begin{aligned}
\sum_{i=0}^{N-1}\mathbb{E}Y_i &\le 3\sqrt{\epsilon}\sum_{i=0}^{N-1}\varphi(\frac{1}{\sqrt{\epsilon}})(1 + 2\mathbb{E}(\sup_{t\in[0,T]}|X^\epsilon(t)|^2)) \\
&\quad + 3(K(\frac{T}{\epsilon}) + L)\sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}} \psi\left(2\mathbb{E}|X^\epsilon(s) - X^\epsilon(t_i)|^2\right) ds \\
&\le CN\sqrt{\epsilon}\left[\varphi(\frac{1}{\sqrt{\epsilon}}) + (K(\frac{T}{\epsilon}) + L)\psi\left(2C_{5,\epsilon}K(\frac{T}{\epsilon})(\sqrt{\epsilon} + \epsilon + \epsilon^H)\right)\right] \\
&\le C(T + \sqrt{\epsilon})\left[\varphi(\frac{1}{\sqrt{\epsilon}}) + (K(\frac{T}{\epsilon}) + L)\psi\left(2C_{5,\epsilon}K(\frac{T}{\epsilon})(\sqrt{\epsilon} + \epsilon + \epsilon^H)\right)\right].
\end{aligned}
\tag{4.14}
$$

Hence,

$$
\begin{aligned}
&\mathbb{E}\int_0^T |\sigma_W(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_W(X^\epsilon(s), \mathscr{L}(X^\epsilon(s)))|^2 ds \\
&\le C(T + \sqrt{\epsilon})\left[\varphi(\frac{1}{\sqrt{\epsilon}}) + (K(\frac{T}{\epsilon}) + L)\psi\left(2C_{5,\epsilon}K(\frac{T}{\epsilon})(\sqrt{\epsilon} + \epsilon + \epsilon^H)\right)\right] \to 0,
\end{aligned}
\tag{4.15}
$$

as ϵ tends to zero.

■

Lemma 4.4 *Suppose that Assumptions 3.1 and 4.1 hold and $\mathbb{E}|\xi|^2 < +\infty$. Then, we have*

$$
\lim_{\epsilon\to 0}\int_0^T \|\sigma_H(\frac{s}{\epsilon}, \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_H(\mathscr{L}(X^\epsilon(s)))\|^2 ds = 0. \tag{4.16}
$$

The proof is the same as Lemma 4.3, here we omit the proof.

Theorem 4.5 *Assume that* $\mathbb{E}|\xi|^2 < +\infty$. *Then, under Assumptions 3.1 and 4.1, the following averaging principle holds*

$$\lim_{\epsilon\to 0}\mathbb{E}\left(\sup_{t\in[0,T]}|X^\epsilon(t)-\bar{X}(t)|^2\right)=0.$$

Proof. For any $r\in[0,T]$, we have

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\in[0,r]}|X^\epsilon(t)-\bar{X}(t)|^2\right)\\
&\leq 3\mathbb{E}\left(\sup_{t\in[0,r]}\left|\int_0^t\left(b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(\bar{X}(s),\mathscr{L}(\bar{X}(s)))\right)ds\right|^2\right)\\
&\quad+3\mathbb{E}\left(\sup_{t\in[0,r]}\left|\int_0^t\left(\sigma_W(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{\sigma}_W(\bar{X}(s),\mathscr{L}(\bar{X}(s)))\right)dW_s\right|^2\right)\\
&\quad+3\mathbb{E}\left(\sup_{t\in[0,r]}\left|\int_0^t\left(\sigma_H(\frac{s}{\epsilon},\mathscr{L}(X^\epsilon(s)))-\bar{\sigma}_H(\mathscr{L}(\bar{X}(s)))\right)dB^H_s\right|^2\right)\\
&:=3V_1+3V_2+3V_3.
\end{aligned}\tag{4.17}$$

By Hölder inequality and Remark 4.2, we have

$$\begin{aligned}
V_1&=\mathbb{E}\left(\sup_{t\in[0,r]}\left|\int_0^t\left(b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(\bar{X}(s),\mathscr{L}(\bar{X}(s)))\right)ds\right|^2\right)\\
&\leq 2\mathbb{E}\left(\sup_{t\in[0,r]}\left|\int_0^t\left(b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))\right)ds\right|^2\right)\\
&\qquad+2\mathbb{E}\left(\sup_{t\in[0,r]}\left|\int_0^t\left(\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(\bar{X}(s),\mathscr{L}(\bar{X}(s)))\right)ds\right|^2\right)\\
&\leq 2\mathbb{E}\left(\sup_{t\in[0,r]}\left|\int_0^t\left(b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))\right)ds\right|^2\right)\\
&\qquad+2r\mathbb{E}\int_0^r|\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(\bar{X}(s),\mathscr{L}(\bar{X}(s)))|^2ds\\
&\leq 2\mathbb{E}\left(\sup_{t\in[0,r]}\left|\int_0^t\left(b(\frac{s}{\epsilon},X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))-\bar{b}(X^\epsilon(s),\mathscr{L}(X^\epsilon(s)))\right)ds\right|^2\right)\\
&\qquad+2TL\int_0^r\psi(2\mathbb{E}|X^\epsilon(s)-\bar{X}(s)|^2)ds.
\end{aligned}\tag{4.18}$$

By Doob's martingale inequality and Remark 4.2, we have

$$\begin{aligned}
V_2 &= \mathbb{E}\left(\sup_{t\in[0,r]}\left|\int_0^t \left(\sigma_W(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_W(\bar{X}(s), \mathscr{L}(\bar{X}(s)))\right) dW(s)\right|^2\right)\\
&\leq 4\mathbb{E}\int_0^r \|\sigma_W(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_W(\bar{X}(s), \mathscr{L}(\bar{X}(s)))\|^2 ds\\
&\leq 8\mathbb{E}\int_0^r \|\sigma_W(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_W(X^\epsilon(s), \mathscr{L}(X^\epsilon(s)))\|^2 ds\\
&\quad + 8\mathbb{E}\int_0^r \|\bar{\sigma}_W(X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_W(\bar{X}(s), \mathscr{L}(\bar{X}(s)))\|^2 ds\\
&\leq 8\mathbb{E}\int_0^r \|\sigma_W(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_W(X^\epsilon(s), \mathscr{L}(X^\epsilon(s)))\|^2 ds\\
&\quad + 8L\int_0^r \psi(2\mathbb{E}|X^\epsilon(s) - \bar{X}(s)|^2) ds.
\end{aligned} \tag{4.19}$$

For the term V_3, by Remark 4.2, we have

$$\begin{aligned}
V_3 &= \mathbb{E}\left(\sup_{0\leq t\leq r}|\int_0^t [\sigma_H(\frac{s}{\epsilon}, \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_H(\mathscr{L}(\bar{X}(s)))] dB^H(s)|^2\right)\\
&\leq C_{\lambda,H} r^{2H-1}\int_0^r \|\sigma_H(\frac{s}{\epsilon}, \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_H(\mathscr{L}(\bar{X}(s)))\|^2 ds\\
&\leq C_{\lambda,H} r^{2H-1}\int_0^r \|\sigma_H(\frac{s}{\epsilon}, \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_H(\mathscr{L}(X^\epsilon(s)))\|^2 ds\\
&\quad + C_{\lambda,H} r^{2H-1}\int_0^r \|\bar{\sigma}_H(\mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_H(\mathscr{L}(\bar{X}(s)))\|^2 ds\\
&\leq C_{\lambda,H} r^{2H-1}\int_0^r \|\sigma_H(\frac{s}{\epsilon}, \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_H(\mathscr{L}(X^\epsilon(s)))\|^2 ds\\
&\quad + C_{\lambda,H} r^{2H-1} L\int_0^r \psi(\mathbb{E}|X^\epsilon(s) - \bar{X}(s)|^2) ds.
\end{aligned} \tag{4.20}$$

Therefore,

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\in[0,r]}|X^\epsilon(t) - \bar{X}(t)|^2\right)\\
&\leq \Big[6\mathbb{E}\Big(\sup_{t\in[0,r]}\left|\int_0^t \left(b(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \bar{b}(X^\epsilon(s), \mathscr{L}(X^\epsilon(s)))\right) ds\right|^2\Big)\\
&\quad + 24\mathbb{E}\int_0^r \|\sigma_W(\frac{s}{\epsilon}, X^\epsilon(s), \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_W(X^\epsilon(s), \mathscr{L}(X^\epsilon(s)))\|^2 ds\\
&\quad + C_{\lambda,H} T^{2H-1}\int_0^r \|\sigma_H(\frac{s}{\epsilon}, \mathscr{L}(X^\epsilon(s))) - \bar{\sigma}_H(\mathscr{L}(X^\epsilon(s)))\|^2 ds\Big]\\
&\quad + (6T + 24 + C_{\lambda,H} T^{2H-1}) L\int_0^r \psi\left(2\mathbb{E}(\sup_{l\in[0,s]}|X^\epsilon(l) - \bar{X}(l)|^2)\right) ds.
\end{aligned} \tag{4.21}$$

By Lemmas 4.2-4.4, we get

$$
\begin{aligned}
&\lim_{\epsilon\to 0}\mathbb{E}\left(\sup_{t\in[0,r]}|X^{\epsilon}(t)-\bar{X}(t)|^2\right)\\
&\le (6T+24+C_{\lambda,H}T^{2H-1})L\int_0^r\psi\left(2\lim_{\epsilon\to 0}\mathbb{E}(\sup_{l\in[0,s]}|X^{\epsilon}(l)-\bar{X}(l)|^2)\right)ds \qquad (4.22)\\
&\le \epsilon_1+(6T+24+C_{\lambda,H}T^{2H-1})L\int_0^r\psi\left(2\lim_{\epsilon\to 0}\mathbb{E}(\sup_{l\in[0,s]}|X^{\epsilon}(l)-\bar{X}(l)|^2)\right)ds.
\end{aligned}
$$

Hence, Bihari inequality reads

$$
\lim_{\epsilon\to 0}\mathbb{E}\left(\sup_{t\in[0,r]}|X^{\epsilon}(t)-\bar{X}(t)|^2\right)\le\frac{1}{2}G^{-1}[G(2\epsilon_1)+12(T+4+C_{\lambda,H}T^{2H-1})LT],
$$

where $G(2\epsilon_1)+12(T+4+C_{\lambda,H}T^{2H-1})LT\in Dom(G^{-1})$, G^{-1} is the inverse function of $G(\cdot)$ and

$$
G(v)=\int_1^v\frac{ds}{\psi(s)},v>0
$$

By Assumption 3.1, one sees that $\lim_{\epsilon_1\downarrow 0^+}G(2\epsilon_1)=-\infty$ and $Dom(G^{-1})=(-\infty,G(\infty))$.

Letting $\epsilon_1\to 0$ gives

$$
\lim_{\epsilon\to 0}\left(\mathbb{E}\sup_{t\in[0,T]}|X^{\epsilon}(t)-\bar{X}(t)|^2\right)=0.
$$

This completes the proof. ■

Remark 4.4 *By the Chebyshev-Markov inequality and Theorem 4.5, for any given number $\delta>0$, we have*

$$
\lim_{\epsilon\to 0}\mathbb{P}\Big(\sup_{0\le t\le T}|X^{\epsilon}(t)-\bar{X}(t)|>\delta\Big)\le\frac{1}{\delta^2}\lim_{\epsilon\to 0}\mathbb{E}\Big(\sup_{0\le t\le T}|X^{\epsilon}(t)-\bar{X}(t)|^2\Big)=0.
$$

This implies the convergence in probability of the solutions $X^{\epsilon}(t)$ to the averaged solution $\bar{X}(t)$.

References

[1] E. Alòs and D. Nualart. Stochastic integration with respect to the fractional Brownian motion. *Stochastics.* **75**, 129-152 (2003).

[2] J. Bao, P. Ren and F-Y. Wang. Bismut formula for Lions derivative of distribution-path dependent SDEs. *J. Differential Equations.* **282**, 285-329 (2021).

[3] M. Bauer and T. Meyer-Brandis. McKean-Vlasov equations on infinite-dimensional Hilbert spaces with irregular drift and additive fractional noise. ArXiv: 1912. 07427v1.

[4] F. Biagini, Y. Hu, B. Øksendal, and T. Zhang. *Stochastic Calculus for Fractional Brownian Motions and Applications.* London: Springer-Verlag, 2008.

[5] R. Buckdahn and S. Jing. Mean-field SDE driven by a fractional Brownian motion and related stochastic control problem. *SIAM J. Control Optim.* **55**, 1500-1533 (2017).

[6] P. E. Chaudru and de Raynal. Strong well-posedness of McKean-Vlasov stochastic differential equations with Hölder drift. *Stochastic Process. Appl.* **130**, 79-107 (2020).

[7] P. Cheridito. Mixed fractional Brownian motion. *Bernoulli.* **7**, 913-934 (2001).

[8] L. Coutin and Z. Qian. Stochastic analysis, rough path analysis and fractional Brownian motions. *Probab. Theory Related Fields.* **122**, 108-140 (2002).

[9] L. Decreusefond and A. S. Üstünel. Fractional brownian motion: theory and applications. *ESAIM Proceedings.* **5**, 75-86 (1998).

[10] Z. Dong, X. Sun, H. Xiao and J. Zhai. Averaging principle for one dimensional stochastic Burgers equation. *J. Differential Equations.* **265**, 4749-4797 (2018).

[11] X. Fan, X. Huang, Y. Suo and C. Yuan. Distribution dependent SDEs driven by fractional Brownian motions. ArXiv: 2105. 14341v1.

[12] L. Galeati, F. A. Harang and A. Mayorcas. Distribution dependent SDEs driven by additive fractional Brownian motions. ArXiv: 2105. 14063v1.

[13] J. Guerra and D. Nualart. Stochastic differential equations driven by fractional Brownian motion and standard Brownian motion. *Stoch. Anal. Appl.* **26**, 1053-1075 (2008).

[14] Z. Guo, G. Lv and J. Wei. Averaging principle for stochastic differential equations under a weak condition. *Chaos.* **30**, 123139 (2020).

[15] M. Hairer and X.-M. Li. Averaging dynamics driven by fractional Brownian motion. *Ann. Probab.* **48**, 1826-1860 (2020).

[16] W. Hammersley, D. Šiška and L. Szpruch. McKean-Vlasov SDEs under measure dependent Lyapunov conditions. ArXiv: 1802. 03974v1.

[17] Y. Hu. *Analysis on Gaussian spaces.* Hackensack: World Scientific Publishing Co. Pte. Ltd., 2017.

[18] X. Huang, F.-Y. Wang. Distribution dependent SDEs with singular coefficients. *Stochastic Process. Appl.* **129**, 4747–4770 (2019).

[19] X. Huang, P. Ren and F-Y Wang. Distribution dependent stochastic differential equations. *Front. Math. China.* **16**, 257-301 (2021).

[20] R. Khasminskii. On the principle of averaging the Itô stochastic differential equations. *Kibernetika.* **4**, 260-279 (1968).

[21] K. Kubilius. The existence and uniqueness of the solution of the integral equation driven by a p-semimartingale of special type. *Stochastic Process. Appl.* **98**, 289-315 (2002).

[22] D. Luo, Q. Zhu and Z. Luo. An averaging principle for stochastic fractional differential equations with time-delays. *Appl. Math. Lett.* **105**, 106290 (2020).

[23] X. Mao. *Stochastic differential equations and applications.* Chichester: Horwood Publishing Limited, first edition (1997), second edition (2008).

[24] S. Mehri and W. Stannat. Weak solutions to Vlasov-McKean equations under Lyapunov-type conditions. *Stoch. Dyn.* **19**, 1950042 (2019).

[25] J. Mémin, Y. Mishura and E. Valkeila. Inequalities for the moments of Wiener integrals with respect to a fractional Brownian motion. *Statist. Probab. Lett.* **51**, 197-206 (2001).

[26] Y. Mishura and A. Yu. Veretennikov. Existence and uniqueness theorems for solutions of McKean-Vlasov stochstic equations. ArXiv: 1603. 02212.

[27] Y. Mishura. *Stochastic calculus for fractional Brownian motions and related processes.* Lecture Notes in Math, 1929. Berlin-Heidelberg: Springer-Verlag, 2008.

[28] Y. Mishura and G. M. Shevchenko. Existence and uniqueness of the solution of stochastic differential equation involving Wiener process and fractional Brownian motion with Hurst index $H > \frac{1}{2}$. *Commun. Stat. Theory Methods.* **40**, 3492-3508 (2011).

[29] D. Nualart. *The Malliavin Calculus and Related Topics.* Berlin: Springer-Verlag, 2006.

[30] D. Nualart and A. Răşcanu. Differential equations driven by fractional Brownian motion. *Collect. Math.* **53**, 55-81 (2002).

[31] B. Pei, Y. Xu and J.-L. Wu. Stochastic averaging for stochastic differential equations driven by fractional Brownian motion and standard Brownian motion. *Appl. Math. Lett.* **100**, 106006 (2020).

[32] B. Pei, Y. Inahama and Y. Xu. Averaging principle for fast-slow system driven by mixed fractional Brownian rough path. *J. Differential Equations.* **301**, 202-235 (2021).

[33] P. Ren and F-Y Wang. Bismut formula for Lions derivative of distribution dependent SDEs and applications. *J. Differential Equations.* **267**, 4745-4777 (2019).

[34] M. Röckner and X. Zhang. Well-posedness of distribution dependent SDEs with singular drifts. *Bernoulli.* **27**, 1131-1158 (2021).

[35] S. G. Samko, A. A. Kilbas and O. I. Marichev. Fractional Integrals and Derivatives: Theory and Applications. Gordon and Breach Science Publishers, 1993.

[36] G. Shen, J.-L. Wu and X. Yin. Averaging principle for fractional heat equations driven by stochastic measures. *Appl. Math. Lett.* **106**, 106404 (2020).

[37] J. L. D. Silva, M. Erraoui and El H. Essaky. Mixed stochastic differential equations: existence and uniqueness result. *J. Theor. Probab.* **31**, 1119-1141 (2018).

[38] K. Sobczyk. *Stochastic differential equations: with applications to physics and engineering. Mathematics and its Applications (East European Series).* Dordrecht: Kluwer Academic Publishers, 1990.

[39] E. Sönmez. On mixed fractional SDEs with discontinuous drift coefficient. ArXiv: 2010. 14176 (2020).

[40] E. M. Stein. *Singular integrals and differentiability properties of functions.* Princeton: Princeton University Press, 1970.

[41] F.-Y. Wang. Distribution dependent SDEs for Landau type equations. *Stochastic Process. Appl.* **128**, 595–621 (2018).

[42] Y. Xu, J. Duan and W. Xu. An averaging principle for stochastic dynamical systems with Lévy noise. *Phys. D.* **240**, 1395-1401 (2011).

[43] M. Zähle. Integration with respect to fractal functions and stochastic calculus I, *Probab. Theory Related Fields.* **111**, 333-374 (1998).

Guangjun Shen
Department of Mathematics, Anhui Normal University, Wuhu 241002, China
gjshen@163.com

Jie Xiang
Department of Mathematics, Anhui Normal University, Wuhu 241002, China
jiexiangahnu@163.com

Jiang-Lun Wu
Department of Mathematics, Computational Foundry, Swansea University
Swansea, SA1 8EN, UK
j.l.wu@swansea.ac.uk