A STRONG AVERAGING PRINCIPLE RATE FOR TWO-TIME-SCALE COUPLED FORWARD-BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS DRIVEN BY FRACTIONAL BROWNIAN MOTION

JIE XU[†,§], QIQI LIAN[†], JIANG-LUN WU[‡]

[†] *College of Mathematics and Information Science, Henan Normal University,*
Xinxiang, Henan 453007*, P. R. China*
[‡] *Department of Mathematics, Computational Foundry, Swansea University,*
Swansea, SA1 8EN, UK

Abstract. This paper concerns the strong convergence rate of an averaging principle for two-time-scale coupled forward-backward stochastic differential equations (CFBSDEs, for short) driven by fractional Brownian motion (fBm, for short). The fast component is a forward stochastic differential equation (FSDE, for short) driven by Brownian motion, while the slow component is a backward stochastic differential equation (BSDE, for short) driven by fBm with the Hurst index greater than 1/2. Combining Malliavin calculus theory to stochastic integral and Khasminskii's time discretization method, the rate of strong convergence for the slow component towards the solution of the averaging equation in the mean square sense is derived. The strong convergence rate of an averaging principle for fast-slow CFBSDEs driven by fBm is new.

AMS Mathematics Subject Classification (2010): 60H10, 70K65, 70K70.
Keywords: Stochastic averaging principle; Convergence rate; Fractional Brownian motion; Fast-slow forward-backward stochastic differential equations.

1. Introduction and main result

For any $T > 0$, consider the following two-time-scale CFBSDEs :

$$\begin{cases} -\mathrm{d}X_t^\epsilon = a(\eta_t, X_t^\epsilon, Y_t^\epsilon, Z_t^\epsilon)\mathrm{d}t - Z_t^\epsilon \mathrm{d}B_t^H, \\ \mathrm{d}Y_t^\epsilon = \frac{1}{\epsilon} f(X_t^\epsilon, Y_t^\epsilon)\mathrm{d}t + \frac{1}{\sqrt{\epsilon}} g(X_t^\epsilon, Y_t^\epsilon)\mathrm{d}W_t, \end{cases} \tag{1.1}$$

for $t \in [0, T]$ and $\epsilon \in (0, 1)$, with a terminal condition $X_T^\epsilon = \varphi(\eta_T)$ and an initial condition $Y_0^\epsilon = y$, where $X_t^\epsilon, Y_t^\epsilon$ and Z_t^ϵ are n-dimensional, m-dimensional and $n \times d$-dimensional diffusion processes, respectively. The driving process B_t^H is a d-dimensional fBm with the Hurst parameter $H \in (1/2, 1)$, and W_t is a r-dimensional Wiener process. The two driven processes $B^H := \{B_t^H\}_{t\in[0,T]}$ and $W := \{W_t\}_{t\in[0,T]}$ are assumed to be independent, and they are defined on a given complete, filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$, where $\mathscr{F}_t$ is the complete reference family generated by B^H and W (i.e., the usual augmentation of σ-algebra $\sigma(B_s^H, W_s, 0 \leqslant s \leqslant t)$), and $\mathscr{F}_t$ satisfies the usual conditions. Here the integral with respect to B^H is a divergence type integral, and that with respect to W is the usual Itô's integral. The precise conditions on a, f, g, φ and η will be presented in Section 3. Moreover, ϵ is a small positive parameter describing the ratio of time scale between the process X^ϵ and Y^ϵ. With this time scale the variable X^ϵ is referred as the *slow* component and Y^ϵ as the *fast* component.

[§] Corresponding author at: College of Mathematics and Information Science, Henan Normal University, Xinxiang, Henan 453007, P. R. China. Tel.:+8615516552382. E-mail addresses: xujiescu@163.com.

If there is no fast component and $n = d = 1$, Eq.(1.1) will be a one time scale BSDE driven by fBm, namely, Eq.(1.1) becomes

$$\begin{cases} -\mathrm{d}X_t^\epsilon = a(\eta_t, X_t^\epsilon, Z_t^\epsilon)\mathrm{d}t - Z_t^\epsilon \mathrm{d}B_t^H, \\ X_T^\epsilon = \varphi(\eta_T). \end{cases}$$

This equation was firstly studied by Hu and Peng [1], where they obtained the existence and uniqueness of the solution. Later, Maticiuc and Nie developed a rigorous approach for this equation with the help of quasi-conditional expectation and derived fractional backward variational inequalities in [2]. Wen and his coauthors discussed the anticipative and mean-field BSDEs driven by fBm in [3] and [4], respectively. For further investigations, the reader is referred to relevant literatures, which we omit here.

If the slow equation in Eq.(1.1) is replaced by an FSDE driven by Brownian motion, the two-time-scale strong averaging principle was initiated by Khasminskii in the seminal work [5]. Since then, the strong averaging principle of FSDEs has been extensively developed in controls, stability analysis, chemical reaction systems, stochastic approximations, adaptive algorithms and extremum seeking (cf.[6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16], just mention a few). Now the stochastic averaging principle of FSDEs has been extended from various aspects (cf., e.g., [17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 31, 32, 34, 35, 36, 37, 38, 39, 40, 41], and so on.).

If the slow equation in Eq.(1.1) is displaced by an FSDE driven by fBm, there are also a number of papers on the averaging principle. Pei, Inahama and Xu in [31] utilized rough path theory to study the averaging principle for such mixed fast-slow systems, where the slow equation is driven by fBm with the Hurst index $H \in (1/3, 1/2]$. Hairer and Li in [32] discussed the averaging dynamics where the slow system is driven by fBm with the Hurst index $H \in (1/2, 1)$ and proved the convergence in probability via stochastic sewing lemma. Li and Sieber in [33] demonstrated a fractional averaging principle for interacting slow-fast systems in Hölder norm in probability, and established geometric ergodicity for a class of fractional FSDEs.

If fBm is substituted by Brownian motion and $f(x, y)$ and $g(x, y)$ reduce to $f(y)$ and $g(y)$, the weak convergence of the averaging principle of Eq.(1.1) has been studied in the literatures. Let us mention a few here. Pardoux and Veretennikov in [42] was first to establish an averaging of BSDEs and then applied to semi-linear PDE's. Later, Essaky and Ouknine in [43] investigated a homogenization of partial differential equations with periodic coefficients by using averaging of BSDEs. Recently, Bahlalia, Elouaflin and Pardoux in [44] proved an averaging principle for BSDEs with null recurrent fast component and further applied to homogenization in a non periodic media.

There exist some results on the weak convergence of the averaging principle of the two-time-scale BSDEs, but the convergence rate is not given, the driving term is induced by Brownian motion and the system is decoupled (i.e., $f(x, y) = f(y)$ and $g(x, y) = g(y)$)(see [42, 43, 44, 45] and references therein). It seems that it is difficult to get the convergence rate by the method outlined in the above papers [42, 43, 44, 45]. However, the convergence rate is crucial in numerical analysis and engineering linked research fields. On the other hand, most results of the average principle of BSDEs are driven by Brownian motion. Contrary to Brownian motion, the increment of fBm with the Hurst parameter $H \in (0, 1/2) \cup (1/2, 1)$ is not independent and a special case of fBm (with the Hurst parameter $H = 1/2$) is Brownian motion, which indicates that fBm may be applied to describe much more natural or social phenomenon than that aspect of Brownian motion. We note that in the case $H \in (1/2, 1)$, fBm is a process with long memory, and it is widely used in finance, telecommunication networks, physics and statistics, etc.. In addition, the decoupled system is not general. It is well known that coupled system can degenerate into the decoupled system, but not vice versa. To the best of our knowledge, the strong averaging principle of the two-time-scale CFBSDEs driven by fBm has not been established.

We would like to point out that it is not an easy task to study the strong averaging principle of the two-time-scale BSDEs driven by fBm. The main reasons are as follows. First, due to the fact that the solution of BSDEs driven by fBm in general is neither Markovian nor a semimartingale, the classical stochastic analysis theory used in the studies of SDEs is not applicable. Second, since we are considering the coupled system, there is no known consequence of the existence and uniqueness for Eq.(1.1) driven by fbm. Third, by examining the existing research methods on the averaging principle of BSDEs, one can not obtain the convergence rate (e.g.,[42, 43, 44, 45]). Last but not least, on account of the presence of BSDEs and Malliavin calculus theory, the control variable Z is rather hard to manage, which is also a difficulty that can not be ignored.

It is very natural to ask whether the strong averaging principle of two-time-scale CFBSDEs driven by fBm still hold. These motivate us to carry out this paper, aiming to establish the strong averaging principle with an explicit convergence rate for Eq.(1.1). Our main result is the following theorem.

Theorem 1.1. *Suppose that (**A1**)-(**A7**) hold, then for any* $T > 0$ *and* $\beta > 0$*, we have*

$$\sup_{0\leqslant t\leqslant T}\left\{e^{\beta t}\mathbb{E}|X_t^\epsilon - \bar{X}_t|^2 + \mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z_s^\epsilon - \bar{Z}_s|^2 \mathrm{d}s\right\} \leqslant C\epsilon^{\frac{1}{4}}, \tag{1.2}$$

with $\epsilon \in (0,1)$ *and* $H \in (1/2, 1)$*, where* C *is a positive constant which is independent of* ϵ*,* (X^ϵ, Z^ϵ) *is the solution of Eq.(1.1), and* $(\bar{X}, \bar{Z})$ *is the solution of the following effect dynamics equation* :

$$\begin{cases} -\mathrm{d}\bar{X}_t = \bar{a}(\eta_t, \bar{X}_t, \bar{Z}_t)\mathrm{d}t - \bar{Z}_t\mathrm{d}B_t^H, \\ \bar{X}_T = \varphi(\eta_T), \end{cases} \tag{1.3}$$

with

$$\bar{a}(u,x,z) = \int_{\mathbb{R}^m} a(u,x,y,z)\mu^x(\mathrm{d}y), \quad (u,x,z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times d},$$

where μ^x *stands for the unique invariant measure for the following fast equation with the frozen slow component*

$$\begin{cases} \mathrm{d}Y_t = f(x, Y_t)\mathrm{d}t + g(x, Y_t)\mathrm{d}W_t, \\ Y_0 = y, \end{cases} \tag{1.4}$$

for any fixed $x \in \mathbb{R}^n$*, and* W_t *is a r-dimensional standard Wiener process. For convenience, we use* $|\cdot|$ *to denote the norms of vectors and matrices in (1.2).*

Remark 1.2. *Without loss of generality, we will consider only* $n = m = d = r = 1$ *in the sequel assumptions, proofs and discussions. The general multidimensional case can be done in the similar manner.*

It is worthy to point out that the novelty of this paper is to extend the uncoupled results driven by Brownian motion in [42, 43, 44, 45] to the coupled case driven by fBm, and from the perspective of proof techniques, we establish a strong convergence rate by combining Malliavin calculus theory and Khasminskii's time discretization method (cf.[5, 48, 49, 50, 51, 52]). Furthermore, different from [30, 31, 32, 33], our work need to think over Z rigorously for the reason that BSDEs participate in the systems, so the model itself is an innovation. As far as we know, this is the first result on the averaging principle rate for fast-slow CFBSDEs driven by fBm.

This paper is organized as follows. In Section 2, we give some main definitions and results about fBm and Malliavin calculus. In Section 3, we present some conditions on the coefficients of equations throughout this work. In Section 4, the existence and uniqueness theorem of two-time-scale CFBSDEs is established. In Section 5, some a priori estimates are carrieded out and further utilized in the subsequent discussions. In Section 6, we prove the mean-square convergence rate for the averaging principle of two-time-scale CFBSDEs driven by fBm.

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

2. Preliminaries

In this section, we recall some main definitions and results about fBm and Malliavin calculus which are used later. For more details, the readers may refer to, e.g., [48], [49], [50], [51], [52].

Let $B^H=(B^H_t, t\geqslant 0)$ be an fBm defined on a complete probability space $(\Omega,\mathscr{F},\mathbb{P})$, with the Hurst parameter $H\in(1/2,1)$. Define

$$\phi(x)=H(2H-1)|x|^{2H-2},\quad x\in\mathbb{R} \tag{2.1}$$

and

$$\langle\xi,\eta\rangle_T=\int_0^T\int_0^T\phi(r-s)\xi_r\eta_s\mathrm{d}r\mathrm{d}s,\quad \|\xi\|_T=\langle\xi,\xi\rangle_T, \tag{2.2}$$

where ξ and η are continuous functions on $[0,T]$. Then $\langle\xi,\eta\rangle_T$ is a Hilbert scalar product. We denote by $\mathscr{H}$ the completion of continuous functions endowed with this scalar product. Moreover, let $\mathscr{P}_T$ be the set of elementary random variables of the form

$$F=u\Big(\int_0^T\xi_1(t)\mathrm{d}B^H_t,...,\int_0^T\xi_n(t)\mathrm{d}B^H_t\Big),$$

where u is a polynomial function of n variables and $\xi_1,...,\xi_n\in\mathscr{H}$. The Malliavin derivative operator D^H of $F\in\mathscr{P}_T$ is defined by

$$D^H_sF=\sum_{i=1}^n\frac{\partial u}{\partial x_i}\Big(\int_0^T\xi_1(t)\mathrm{d}B^H_t,\dots,\int_0^T\xi_n(t)\mathrm{d}B^H_t\Big)\xi_i(s),\quad s\in[0,T].$$

Due to the fact that the derivative operator $D^H: L^2(\Omega,\mathscr{F},\mathbb{P})\mapsto(\Omega,\mathscr{F},\mathbb{P})$ is closable, we denote by $\mathbb{D}_{1,2}$ the Banach space defined as the completion of $\mathscr{P}_T$ equipped with the following norm

$$\|F\|^2_{1,2}=\mathbb{E}|F|^2+\mathbb{E}\|D^H_sF\|^2_T,\quad F\in\mathscr{P}_T.$$

Now let us introduce another derivative

$$\mathbb{D}^H_tF=\int_0^T\phi(t-v)D^H_vF\mathrm{d}v,\quad t\in[0,T].$$

Moreover, we need the adjoint operator of the Malliavin derivative operator D^H, which is the so-called Skorohod divergence operator. This operator represents the divergence type integral and is denoted by $\delta(\cdot)$.

Definition 2.1. *A process $v\in L^2(\Omega\times[0,T];\mathscr{H})$ is said to belong to the domain Dom(δ), if there exists $\delta(v)\in L^2(\Omega\times[0,T])$ satisfying the following duality relationship* :

$$\mathbb{E}(F\delta(v))=\mathbb{E}(\langle D^H_\cdot F,v\rangle_T),\quad \textit{for all } F\in\mathscr{P}_T.$$

In addition, if $v\in$ Dom(δ), the divergence type integral of v w.r.t.B^H is defined by putting $\int_0^T v_s\mathrm{d}B^H_s=:\delta(v)$.

One has the following result for the divergence-type integrals

Proposition 2.2. *Let $\mathbb{L}^H_{1,2}$ be the space of all stochastic processes $\rho : (\Omega, \mathscr{F}, \mathbb{P}) \mapsto \mathscr{H}$ such that*

$$\mathbb{E}\Big(\|\rho\|_T^2 + \int_0^T \int_0^T |\mathbb{D}_s^H \rho(t)|^2 \mathrm{d}s\mathrm{d}t\Big) < \infty. \tag{2.3}$$

If $\rho \in \mathbb{L}^{1,2}_H$, then the divergence-type integral $\int_0^T \rho(t)\mathrm{d}B_t^H$ exists in $L^2(\Omega, \mathscr{F}, \mathbb{P})$ and

$$\mathbb{E}\Big(\int_0^T \rho(t)\mathrm{d}B_t^H\Big) = 0, \quad \mathbb{E}\Big(\int_0^T \rho(t)\mathrm{d}B_t^H\Big)^2 = \mathbb{E}\Big(\|\rho\|_T^2 + \int_0^T \int_0^T \mathbb{D}_s^H \rho(t)\mathbb{D}_t^H \rho(s)\mathrm{d}s\mathrm{d}t\Big).$$

Next, we shall present the Itô's formula and the integration by parts formula.

Theorem 2.3. *Assume that $\lambda, \rho : [0, T] \mapsto \mathbb{R}$ are deterministic continuous functions. Let*

$$X_t = X_0 + \int_0^t \lambda(s)\mathrm{d}s + \int_0^t \rho(s)\mathrm{d}B_s^H, t \in [0, T],$$

where the initial X_0 is a constant. Then, for every $F \in C^{1,2}([0, T] \times \mathbb{R})$, the following formula holds:

$$\begin{aligned} F(t, X_t) =& F(0, X_0) + \int_0^t \frac{\partial F}{\partial s}(s, X_s)\mathrm{d}s + \int_0^t \frac{\partial F}{\partial s}(s, X_s)\lambda(s)\mathrm{d}s \\ &+ \int_0^t \frac{\partial F}{\partial s}(s, X_s)\rho(s)\mathrm{d}B_s^H + \frac{1}{2}\int_0^t \frac{\partial^2 F}{\partial s^2}(s, X_s)\Big[\frac{\mathrm{d}}{\mathrm{d}s}\|\rho\|_s^2\Big]\mathrm{d}s, \quad t \in [0, T]. \end{aligned}$$

Theorem 2.4. *Let $T \in (0, \infty)$, $\lambda_2, \rho_2 \in \mathbb{D}^{1,2}$, and for $i = 1, 2$,*

$$\mathbb{E}\Big[\int_0^T |\lambda_i(s)|^2 \mathrm{d}s + \int_0^T |\rho_i(s)|^2 \mathrm{d}s\Big] < \infty.$$

Suppose that $\mathbb{D}_t^H \lambda_2(s)$ and $\mathbb{D}_t^H \rho_2(s)$ are continuously differentiable with respect to $(s, t) \in [0, T]^2$ for $\mathbb{P}$-almost all $\omega \in \Omega$. Furthermore, assume that

$$\mathbb{E}\int_0^T \int_0^T |\mathbb{D}_t^H \lambda_2(s)|^2 \mathrm{d}s\mathrm{d}t < \infty \quad \textit{and} \quad \mathbb{E}\int_0^T \int_0^T |\mathbb{D}_t^H \rho_2(s)|^2 \mathrm{d}s\mathrm{d}t < \infty.$$

Denote

$$F(t) = \int_0^t \lambda_1(s)\mathrm{d}s + \int_0^t \lambda_2(s)\mathrm{d}B_s^H, \quad t \in [0, T]$$

and

$$G(t) = \int_0^t \rho_1(s)\mathrm{d}s + \int_0^t \rho_2(s)\mathrm{d}B_s^H, \quad t \in [0, T].$$

Then

$$\begin{aligned} F(t)G(t) = &\int_0^t G(s)\lambda_1(s)\mathrm{d}s + \int_0^t G(s)\lambda_2(s)\mathrm{d}B_s^H + \int_0^t F(s)\rho_1(s)\mathrm{d}s \\ &+ \int_0^t F(s)\rho_2(s)\mathrm{d}B_s^H + \int_0^t \mathbb{D}_s^H G(s)\lambda_2(s)\mathrm{d}s + \int_0^t \mathbb{D}_s^H F(s)\rho_2(s)\mathrm{d}s. \end{aligned} \tag{2.4}$$

Denote

$$\mathrm{d}G(t) = \rho_1(t)\mathrm{d}t + \rho_2(t)\mathrm{d}B_t^H,$$

which means that

$$\int_0^t \lambda_1(s)\mathrm{d}G(s) = \int_0^t \lambda_1(s)\rho_1(s)\mathrm{d}s + \int_0^t \lambda_1(s)\rho_2(s)\mathrm{d}B_s^H.$$

The same notation will be applied to $\mathrm{d}F(t)$.

Remark 2.5. *With the above notations, the formula* (*2.4*) *can be written formally as*

$$\mathrm{d}(F(t)G(t)) = F(t)\mathrm{d}G(t) + G(t)\mathrm{d}F(t) + [\mathbb{D}_t^H G(t)\lambda_2(t) + \mathbb{D}_t^H F(t)\rho_2(t)]\mathrm{d}t. \tag{2.5}$$

3. OUR ASSUMPTIONS

In the section, we give some assumptions throughout the rest of this work. We assume that the drift coefficients $a(u,x,y,z) : \mathbb{R}^4 \mapsto \mathbb{R}$, $f(x,y) : \mathbb{R}\times\mathbb{R} \mapsto \mathbb{R}$, and the diffusion coefficient $g(x,y) : \mathbb{R}\times\mathbb{R} \mapsto \mathbb{R}$ are Borel measurable and the following conditions hold:

(A1). There exists a constant $\alpha > 0$, which is independent of (x,y), such that

$$g^2(x,y) \geqslant \alpha, \tag{3.1}$$

for all $(x,y) \in \mathbb{R}\times\mathbb{R}$.

(A2). There exists a positive constant K_1 such that for all $(u_i,x_i,y_i,z_i) \in \mathbb{R}^4, i = 1,2$,

$$\begin{aligned}&|a(u_1,x_1,y_1,z_1) - a(u_2,x_2,y_2,z_2)|^2\\ \leqslant& K_1(|u_1-u_2|^2 + |x_1-x_2|^2 + |y_1-y_2|^2 + |\zeta(z_1)-\zeta(z_2)|^2),\end{aligned} \tag{3.2}$$

$$|f(x_1,y_1) - f(x_2,y_2)|^2 + |g(x_1,y_1) - g(x_2,y_2)|^2 \leqslant K_1(|x_1-x_2|^2 + |y_1-y_2|^2), \tag{3.3}$$

where $a(0,0,0,0) = 0$, and $\zeta : \mathbb{R} \mapsto \mathbb{R}$ is a measurable function with $\zeta(0) = 0$, which is bounded by a positive constant $\tilde{K}$ and Lipschitz continuous, that is, there exists some positive constant L such that $|\zeta(z_1) - \zeta(z_2)| \leqslant L|z_1 - z_2|$ for any $z_1, z_2 \in \mathbb{R}$. Then, it is straightforward to verify that

$$|a(u_1,x_1,y_1,z_1) - a(u_2,x_2,y_2,z_2)|^2 \leqslant K_1(|u_1-u_2|^2 + |x_1-x_2|^2 + |y_1-y_2|^2 + L^2|z_1-z_2|^2) \tag{3.4}$$

and

$$|a(u_1,x_1,y_1,z_1)|^2 \leqslant K_1(|u_1|^2 + |x_1|^2 + |y_1|^2 + \tilde{K}^2). \tag{3.5}$$

(A3). There exist constants $\beta_1 > 0$ and $C > 0$, which are both independent of (x,y), such that

$$2yf(x,y) + |g(x,y)|^2 \leqslant -\beta_1|y|^2 + C, \tag{3.6}$$

for all $(x,y) \in \mathbb{R}\times\mathbb{R}$.

(A4). There exist constants $\beta_2 > 0$ and $C > 0$, which are both independent of (x_i,y_i), such that

$$2(y_1-y_2)(f(x_1,y_1) - f(x_2,y_2)) + |g(x_1,y_1) - g(x_2,y_2)|^2 \leqslant -\beta_2|y_1-y_2|^2 + C|x_1-x_2|^2, \tag{3.7}$$

for all $(x_i,y_i) \in \mathbb{R}\times\mathbb{R}$, $i = 1,2$.

(A5). $\varphi : \mathbb{R} \mapsto \mathbb{R}$ is a differentiable function with polynomial growth.

Remark 3.1. *We would like to give some comments on the above assumptions.*

- (***A1***), *(*3.3*)*, (***A3***) *and* (***A4***) *are interpreted as coupled conditions which yield a unique invariant measure possessing exponentially mixing property for a Markov semigroup associated to the fast variable equation* (*see, e.g.,* [25, Proposition 3.9.], [36, Section 4], [54, Theorem 6.6]).
- *The assumption on* ζ *in* (***A2***) *is very important. This point can be seen in* (6.11). *Without the boundedness of* ζ, *the estimate of* (6.11) *can not be derived. The main reason is that* $\mathbb{E}\int_0^T s^{2H-1}e^{\beta s}|Z_s|^2\mathrm{d}s$ *is finite, but we can not get* $\mathbb{E}|Z_s|^2 < +\infty$, *this makes it hard to deal with* (6.11).
- *On the one hand, due to the boundedness of* ζ, *we have* $|\zeta(z)| \leqslant \tilde{K}$. *Moreover, the Lipschitz condition for* ζ, *combined with* $\zeta(0) = 0$, *yields that*

$$|\zeta(z)|^2 = |\zeta(z) - \zeta(0)|^2 \leqslant L^2|z|^2.$$

Therefore, for all $z \in \mathbb{R}$*, we have*

$$|\zeta(z)| \leqslant L|z| \cdot I_{\{L|z| \leqslant \tilde{K}\}} + \tilde{K} \cdot I_{\{L|z| > \tilde{K}\}},$$

which means the boundedness of ζ *and linear growth condition of* ζ *are not contradictory. In fact, one can have examples for such function* ζ*, one example is given as follows:*

Example 3.2. *Let* $\zeta(z) = \arctan z$*. For all* $z \in \mathbb{R}$*, it is obvious that* $\zeta(0) = 0$ *and* $|\zeta(z)| \leqslant \frac{\pi}{2}$*. By utilizing the mean value theorem, one can derive for each* $z \in \mathbb{R}\backslash 0$*,* $|\zeta(z)| = |\zeta(z) - \zeta(0)| = \frac{1}{1+\xi^2}|z| \leqslant |z|$*, where* $\xi \in (0, z)$ *or* $\xi \in (z, 0)$*. Therefore, for all* $z \in \mathbb{R}$*, we have*

$$|\zeta(z)| \leqslant |z| \cdot I_{\{|z| \leqslant \frac{\pi}{2}\}} + \frac{\pi}{2} \cdot I_{\{|z| > \frac{\pi}{2}\}}.$$

The above inequality can also be regarded as a more accurate characterization of ζ*, rather than only the bounded condition involved.*

Next, we present some hypotheses and propositions for the stochastic process η. Let

$$\eta_t = \eta_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}B_s^H, \quad t \in [0, T], \tag{3.8}$$

where the coefficients satisfy the following:
(A6). the initial $\eta_0 \in \mathbb{R}$ is a constant;
(A7). the drift coefficient $b : \mathbb{R} \mapsto \mathbb{R}$ is a deterministic continuous function, and the coefficient $\sigma : \mathbb{R} \mapsto \mathbb{R}$ is a deterministic continuous function such that $\sigma_t \neq 0$, $t \in [0, T]$.
If we define

$$\hat{\sigma}_t = \int_0^t \phi(t - v)\sigma_v \mathrm{d}v, \quad t \in [0, T],$$

then, by the definition of the scalar product (see (2.1) and (2.2)), we have

$$\|\sigma\|_t^2 = H(2H - 1) \int_0^t \int_0^t |u - v|^{2H-2} \sigma_u \sigma_v \mathrm{d}u \mathrm{d}v.$$

Hence, $\|\sigma\|_t^2$ is continuous differentiable w.r.t. t, and

$$\begin{aligned} \frac{\mathrm{d}\|\sigma\|_t^2}{\mathrm{d}t} =& \frac{\mathrm{d}}{\mathrm{d}t}\Big(2H(2H - 1) \int_v^t \int_0^t |u - v|^{2H-2} \sigma_u \sigma_v \mathrm{d}u \mathrm{d}v\Big) \\ =& 2H(2H - 1) \int_0^t |t - v|^{2H-2} \sigma_v \sigma_t \mathrm{d}v \\ =& 2\hat{\sigma}_t \sigma_t > 0, \quad t \in [0, T]. \end{aligned}$$

Furthermore, by Remark 6 in [2], there exists a constant $M > 0$ such that

$$\frac{1}{M} t^{2H-1} \leqslant \frac{\hat{\sigma}_t}{\sigma_t} \leqslant M t^{2H-1}, \quad t \in [0, T]. \tag{3.9}$$

From **(A6)**, **(A7)** and Proposition 2.2, we know that there exists a constant $C_T > 0$ such that

$$\mathbb{E}|\eta_t|^2 \leqslant 3\Big[\eta_0^2 + \Big(\int_0^t b_s \mathrm{d}s\Big)^2 + \|\sigma\|_t\Big] \leqslant C_T, \quad t \in [0, T]. \tag{3.10}$$

Let us finish this section by defining the following spaces:

- $C_{pol}^{1,3}([0, T] \times \mathbb{R}) := \Big\{\varphi \in C^{1,3}([0, T] \times \mathbb{R}),$ and all the derivatives of φ are polynomial growth$\Big\}$.
- $$\mathcal{V}_T := \Big\{X = \phi(\cdot, \eta(\cdot)) : \phi \in C_{pol}^{1,3}([0, T] \times \mathbb{R}) \text{ with } \frac{\partial \phi}{\partial t} \in C_{pol}^{0,1}([0, T] \times \mathbb{R})\Big\}. \tag{3.11}$$

And also, by $\tilde{\mathcal{V}}_T$ and $\tilde{\mathcal{V}}_T^H$ we denote the completion of $\mathcal{V}_T$ under the following norms, respectively

$$\|X\|^2 := \mathbb{E}\int_0^T e^{\beta t}|X_t|^2 \mathrm{d}t, \quad \|Z\|^2 := \mathbb{E}\int_0^T t^{2H-1}e^{\beta t}|Z_t|^2 \mathrm{d}t,$$

where β is a positive constant.

- $\mathcal{S}_T^2 := \Big\{\mathbb{R}$-valued $\mathscr{F}_t$ adapted continuous stochastic processes $Y_t : \mathbb{E}\sup\limits_{0\leqslant t\leqslant T}|Y_t|^2 < \infty\Big\}$.

4. WELL-POSEDNESS

In this section we state and prove the existence and uniqueness of the two-time-scale fractional BSDEs. We proceed to introduce a lemma, which plays an important role in the proof of the well-posedness theorem.

Lemma 4.1. *Suppose that* (***A2***) *and* (***A5***)*-*(***A7***) *hold. Then*

(i) for a pair of fixed adapted stochastic processes $(x,\ z)$*, the following equation admits a unique solution* $(X, Y, Z) \in (\tilde{\mathcal{V}}_T \times \mathcal{S}_T^2 \times \tilde{\mathcal{V}}_T^H)$

$$\begin{cases} X_t = \varphi(\eta_T) + \int_t^T a(\eta_s, X_s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}B_s^H, \\ Y_t = y + \int_0^t f(x_s, Y_s)\mathrm{d}s + \int_0^t g(x_s, Y_s)\mathrm{d}W_s, \end{cases} \tag{4.1}$$

provided that the coefficients are $\mathscr{F}$*-adapted processes and satisfy*

$$\int_0^T \mathbb{E}[|a(\eta_t, 0, 0, 0)|^2 + |f(x_t, 0)|^2 + |g(x_t, 0)|^2]\mathrm{d}t < \infty,$$

where φ*,* y*,* η*,* a*,* f *and* g *are the same as those given in the CFBSDEs (1.1);*

(ii) The following inequality holds

$$\begin{aligned} &\mathbb{E}\sup_{0\leqslant t\leqslant T}|Y_t|^2 + \mathbb{E}\int_0^T e^{\beta t}|X_t|^2\mathrm{d}t + \mathbb{E}\int_0^T e^{\beta t}t^{2H-1}|Z_t|^2\mathrm{d}t \\ \leqslant & C_1\mathbb{E}|e^{\beta T}\varphi^2(\eta_T)| + 3e^{6K_1T(T+4)}\Big(\frac{C_1T^{2H}e^{\beta T}}{LM}+1\Big)|y|^2 + \frac{C_1e^{\beta T}T^{2H-1}}{K_1LM}\int_0^T \mathbb{E}|a(\eta_t, 0, 0, 0)|^2\mathrm{d}t \\ &+ 6e^{6K_1T(T+4)}(T+4)\Big(\frac{C_1T^{2H}e^{\beta T}}{LM}+1\Big)\int_0^T \mathbb{E}[|f(x_t, 0)|^2 + |g(x_t, 0)|^2]\mathrm{d}t, \end{aligned} \tag{4.2}$$

where $C_1 := Te^{\frac{T^{2H}}{2HLM}+\frac{K_1LMT^{2-2H}}{1-H}} + M\Big[1 + e^{\frac{T^{2H}}{2HLM}+\frac{K_1LMT^{2-2H}}{1-H}}\Big(\frac{T^{2H}}{2HLM}+\frac{K_1LMT^{2-2H}}{1-H}\Big)\Big]$.

Proof. This lemma is the combination of Theorem 25 in [2] and Theorem 3.17 in [46], so it is sufficient to check (4.2).

For $|Y_t|^2$, by the basic inequality $(a+b+c)^2 \leqslant 3(a^2+b^2+c^2)$, Cauchy-Schwarz's inequality, BDG's inequality and (3.3), it is not hard to get for any $u \in [0, T]$

$$\begin{aligned} &\mathbb{E}\sup_{0\leqslant t\leqslant u}|Y_t|^2 \\ \leqslant & 3|y|^2 + 3\mathbb{E}\sup_{0\leqslant t\leqslant u}\Big|\int_0^t f(x_s, Y_s)\mathrm{d}s\Big|^2 + 3\mathbb{E}\sup_{0\leqslant t\leqslant u}\Big|\int_0^t g(x_s, Y_s)\mathrm{d}W_s\Big|^2 \\ \leqslant & 3|y|^2 + 3T\int_0^u \mathbb{E}|f(x_s, Y_s)|^2\mathrm{d}s + 12\int_0^u \mathbb{E}|g(x_s, Y_s)|^2\mathrm{d}s \\ \leqslant & 3|y|^2 + 6T\int_0^u \mathbb{E}[|f(x_s, Y_s) - f(x_s, 0)|^2 + |f(x_s, 0)|^2]\mathrm{d}s \end{aligned}$$

$$
\begin{aligned}
&+24\int_0^u \mathbb{E}[|g(x_s,Y_s)-g(x_s,0)|^2+|g(x_s,0)|^2]\mathrm{d}s\\
\leqslant& 3|y|^2+6T\int_0^T \mathbb{E}|f(x_s,0)|^2\mathrm{d}s+24\int_0^T \mathbb{E}|g(x_s,0)|^2\mathrm{d}s+(6TK_1+24K_1)\int_0^u \mathbb{E}|Y_s|^2\mathrm{d}s\\
\leqslant& 3|y|^2+6T\int_0^T \mathbb{E}|f(x_s,0)|^2\mathrm{d}s+24\int_0^T \mathbb{E}|g(x_s,0)|^2\mathrm{d}s+(6TK_1+24K_1)\int_0^u \mathbb{E}\sup_{0\leqslant v\leqslant s}|Y_v|^2\mathrm{d}s,
\end{aligned}
$$

which by Gronwall's inequality implies that

$$
\mathbb{E}\sup_{0\leqslant t\leqslant T}|Y_t|^2\leqslant e^{6K_1T(T+4)}\Big[3|y|^2+6T\int_0^T \mathbb{E}|f(x_s,0)|^2\mathrm{d}s+24\int_0^T \mathbb{E}|g(x_s,0)|^2\mathrm{d}s\Big]. \tag{4.3}
$$

Now we deal with X_t. Due to Theorem 8 of [2], we have

$$
\mathrm{d}|X_t|^2=2X_t\mathrm{d}X_t+2Z_t\mathbb{D}_t^HX_t\mathrm{d}t=-2a(\eta_t,X_t,Y_t,Z_t)X_t\mathrm{d}t+2X_tZ_t\mathrm{d}B_t^H+2Z_t\mathbb{D}_t^HX_t\mathrm{d}t.
$$

By applying the integration by parts formula (2.4) to $e^{\beta t}|X_t|^2$, we get

$$
\begin{aligned}
&e^{\beta t}|X_t|^2+2\int_t^T e^{\beta s}Z_s\mathbb{D}_s^HX_s\mathrm{d}s+\beta\int_t^T e^{\beta s}|X_s|^2\mathrm{d}s\\
=&e^{\beta T}\varphi^2(\eta_T)+2\int_t^T e^{\beta s}a(\eta_s,X_s,Y_s,Z_s)X_s\mathrm{d}s-2\int_t^T e^{\beta s}X_sZ_s\mathrm{d}B_s^H.
\end{aligned} \tag{4.4}
$$

By (3.4) and the inequality $2xy\leqslant\frac{1}{C}x^2+Cy^2$, it is easy to derive that

$$
\begin{aligned}
&2a(\eta_s,X_s,Y_s,Z_s)X_s\\
\leqslant&\frac{s^{2H-1}}{2K_1LM}[a(\eta_s,X_s,Y_s,Z_s)-a(\eta_s,0,0,0)+a(\eta_s,0,0,0)]^2+\frac{2K_1LM}{s^{2H-1}}|X_s|^2\\
\leqslant&\frac{s^{2H-1}}{K_1LM}[a(\eta_s,X_s,Y_s,Z_s)-a(\eta_s,0,0,0)]^2+\frac{s^{2H-1}}{K_1LM}|a(\eta_s,0,0,0)|^2+\frac{2K_1LM}{s^{2H-1}}|X_s|^2\\
\leqslant&\Big(\frac{s^{2H-1}}{LM}+\frac{2K_1LM}{s^{2H-1}}\Big)|X_s|^2+\frac{s^{2H-1}}{LM}|Y_s|^2+\frac{s^{2H-1}}{M}|Z_s|^2+\frac{s^{2H-1}}{K_1LM}|a(\eta_s,0,0,0)|^2.
\end{aligned} \tag{4.5}
$$

Taking expectation on both sides of (4.4) and using (4.5), we have

$$
\begin{aligned}
&\mathbb{E}[e^{\beta t}|X_t|^2]+2\mathbb{E}\int_t^T e^{\beta s}Z_s\mathbb{D}_s^HX_s\mathrm{d}s\\
\leqslant&\mathbb{E}|e^{\beta T}\varphi^2(\eta_T)|+\mathbb{E}\int_t^T e^{\beta s}\Big(\frac{s^{2H-1}}{LM}+\frac{2K_1LM}{s^{2H-1}}\Big)|X_s|^2\mathrm{d}s+\mathbb{E}\int_t^T\frac{e^{\beta s}s^{2H-1}}{LM}|Y_s|^2\mathrm{d}s\\
&+\frac{1}{M}\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z_s|^2\mathrm{d}s+\mathbb{E}\int_t^T\frac{e^{\beta s}s^{2H-1}}{K_1LM}|a(\eta_s,0,0,0)|^2\mathrm{d}s\\
\leqslant&\mathbb{E}|e^{\beta T}\varphi^2(\eta_T)|+\mathbb{E}\int_t^T e^{\beta s}\Big(\frac{s^{2H-1}}{LM}+\frac{2K_1LM}{s^{2H-1}}\Big)|X_s|^2\mathrm{d}s+\frac{T^{2H-1}}{LM}\int_0^T e^{\beta s}\mathbb{E}|Y_s|^2\mathrm{d}s\\
&+\frac{1}{M}\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z_s|^2\mathrm{d}s+\frac{e^{\beta T}T^{2H-1}}{K_1LM}\int_0^T\mathbb{E}|a(\eta_s,0,0,0)|^2\mathrm{d}s.
\end{aligned} \tag{4.6}
$$

According to (4.1) and [2, Proposition 24], we are able to deduce that $\mathbb{D}_s^HX_s=\frac{\hat{\sigma}_s}{\sigma_s}Z_s^\epsilon$. Together with (3.9), (4.3) and (4.6), it follows that for any $t\in[0,T]$

$$
\mathbb{E}[e^{\beta t}|X_t|^2]+\frac{1}{M}\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z_s|^2\mathrm{d}s
$$

$$
\begin{aligned}
&\leqslant \mathbb{E}|e^{\beta T}\varphi^2(\eta_T)| + \frac{e^{\beta T}T^{2H-1}}{K_1LM}\int_0^T \mathbb{E}|a(\eta_s,0,0,0)|^2\mathrm{d}s\\
&\quad + \frac{e^{\beta T}T^{2H-1}}{LM}\int_0^T \mathbb{E}\sup_{0\leqslant s\leqslant T}|Y_s|^2\mathrm{d}s + \mathbb{E}\int_t^T e^{\beta s}\Big(\frac{s^{2H-1}}{LM} + \frac{2K_1LM}{s^{2H-1}}\Big)|X_s|^2\mathrm{d}s\\
&\leqslant \mathbb{E}|e^{\beta T}\varphi^2(\eta_T)| + \frac{T^{2H}}{LM}e^{6K_1T(T+4)+\beta T}\Big[3|y|^2 + 6T\int_0^T \mathbb{E}|f(x_s,0)|^2\mathrm{d}s + 24\int_0^T \mathbb{E}|g(x_s,0)|^2\mathrm{d}s\Big]\\
&\quad + \frac{e^{\beta T}T^{2H-1}}{K_1LM}\int_0^T \mathbb{E}|a(\eta_s,0,0,0)|^2\mathrm{d}s + \mathbb{E}\int_t^T e^{\beta s}\Big(\frac{s^{2H-1}}{LM} + \frac{2K_1LM}{s^{2H-1}}\Big)|X_s|^2\mathrm{d}s. \qquad (4.7)
\end{aligned}
$$

By Gronwall's inequality (cf.[46, Page 581, Corollary 6.62]), (4.7) implies

$$
\begin{aligned}
&\sup_{0\leqslant t\leqslant T}\mathbb{E}[e^{\beta t}|X_t|^2]\\
&\leqslant \Big\{\mathbb{E}|e^{\beta T}\varphi^2(\eta_T)| + \frac{T^{2H}}{LM}e^{6K_1T(T+4)+\beta T}\Big[3|y|^2 + 6T\int_0^T \mathbb{E}|f(x_t,0)|^2\mathrm{d}t + 24\int_0^T \mathbb{E}|g(x_t,0)|^2\mathrm{d}t\Big]\\
&\quad + \frac{e^{\beta T}T^{2H-1}}{K_1LM}\int_0^T \mathbb{E}|a(\eta_t,0,0,0)|^2\mathrm{d}t\Big\}e^{\frac{T^{2H}}{2HLM}+\frac{K_1LMT^{2-2H}}{1-H}},
\end{aligned}
$$

so we obtain

$$
\begin{aligned}
&\mathbb{E}\int_0^T e^{\beta t}|X_t|^2\mathrm{d}t\\
&\leqslant \Big\{\mathbb{E}|e^{\beta T}\varphi^2(\eta_T)| + \frac{T^{2H}}{LM}e^{6K_1T(T+4)+\beta T}\Big[3|y|^2 + 6T\int_0^T \mathbb{E}|f(x_t,0)|^2\mathrm{d}t + 24\int_0^T \mathbb{E}|g(x_t,0)|^2\mathrm{d}t\Big]\\
&\quad + \frac{e^{\beta T}T^{2H-1}}{K_1LM}\int_0^T \mathbb{E}|a(\eta_t,0,0,0)|^2\mathrm{d}t\Big\}Te^{\frac{T^{2H}}{2HLM}+\frac{K_1LMT^{2-2H}}{1-H}}
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}\int_0^T e^{\beta t}t^{2H-1}|Z_t|^2\mathrm{d}t\\
&\leqslant \Big\{\mathbb{E}|e^{\beta T}\varphi^2(\eta_T)| + \frac{T^{2H}}{LM}e^{6K_1T(T+4)+\beta T}\Big[3|y|^2 + 6T\int_0^T \mathbb{E}|f(x_t,0)|^2\mathrm{d}t + 24\int_0^T \mathbb{E}|g(x_t,0)|^2\mathrm{d}t\Big]\\
&\quad + \frac{e^{\beta T}T^{2H-1}}{K_1LM}\int_0^T \mathbb{E}|a(\eta_t,0,0,0)|^2\mathrm{d}t\Big\}M\Big[1 + e^{\frac{T^{2H}}{2HLM}+\frac{K_1LMT^{2-2H}}{1-H}}\Big(\frac{T^{2H}}{2HLM} + \frac{K_1LMT^{2-2H}}{1-H}\Big)\Big].
\end{aligned}
$$

Combining the above two inequalities with (4.3), we finally arrive at (4.2), which then completes the proof. □

Theorem 4.2. *Under* ($\boldsymbol{A2}$) *and* ($\boldsymbol{A5}$)-($\boldsymbol{A7}$), *Eq.*(*1.1*) *admits a unique solution* $(X^\epsilon, Y^\epsilon, Z^\epsilon)$ *satisfying the following*:

(*i*) $(X^\epsilon, Y^\epsilon, Z^\epsilon) \in (\tilde{\mathcal{V}}_T \times \mathcal{S}_T^2 \times \tilde{\mathcal{V}}_T^H)$;

(*ii*) $X_t^\epsilon = \varphi(\eta_T) + \int_t^T a(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon)\mathrm{d}s - \int_t^T Z_s^\epsilon \mathrm{d}B_s^H, \quad 0 \leqslant t \leqslant T$;

(*iii*) $Y_t^\epsilon = y + \frac{1}{\epsilon}\int_0^t f(X_s^\epsilon, Y_s^\epsilon)\mathrm{d}s + \frac{1}{\sqrt{\epsilon}}\int_0^t g(X_s^\epsilon, Y_s^\epsilon)\mathrm{d}W_s, \quad 0 \leqslant t \leqslant T$.

Proof. Without loss of generality, we only prove the case of $\epsilon = 1$. For arbitrarily fixed $x \in \tilde{\mathcal{V}}_T$, we consider the following system

$$
\begin{cases}
X_t = \varphi(\eta_T) + \int_t^T a(\eta_s, X_s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s\mathrm{d}B_s^H,\\
Y_t = y + \int_0^t f(x_s, Y_s)\mathrm{d}s + \int_0^t g(x_s, Y_s)\mathrm{d}W_s.
\end{cases} \qquad (4.8)
$$

Next, we introduce the operator $\Gamma : (\tilde{\mathcal{V}}_T \times \tilde{\mathcal{V}}_T^H) \to (\tilde{\mathcal{V}}_T \times \tilde{\mathcal{V}}_T^H)$, defined by $(x, z) \to \Gamma(x, z) = (X, Z)$. For two elements $(x, z), (x', z') \in (\tilde{\mathcal{V}}_T \times \tilde{\mathcal{V}}_T^H)$, let (X, Y, Z), (X', Y', Z') be the corresponding solution to (4.8). If we set

$$\Delta x = x - x', \Delta z = z - z', \Delta X = X - X', \Delta Y = Y - Y', \Delta Z = Z - Z',$$

we will focus on the system below for any $t \in [0, T]$

$$\begin{cases} \Delta X_t = \int_t^T [\alpha_s^1 \Delta X_s + \alpha_s^2 \Delta Y_s + \alpha_s^3 \Delta Z_s] \mathrm{d}s - \int_t^T Z_s \mathrm{d}B_s^H, \\ \Delta Y_t = \int_0^t [\alpha_s^4 \Delta x_s + \alpha_s^5 \Delta Y_s] \mathrm{d}s + \int_0^t [\alpha_s^6 \Delta x_s + \alpha_s^7 \Delta Y_s] \mathrm{d}W_s, \end{cases} \tag{4.9}$$

where

$$\alpha_s^1 = \begin{cases} \frac{a(\eta_s, X_s, Y_s, Z_s) - a(\eta_s, X_s', Y_s, Z_s)}{\Delta X_s}, & \text{if} \quad \Delta X_s \neq 0, \\ 0, & \text{if} \quad \Delta X_s = 0, \end{cases}$$

and α_s^i are defined similarly, $i = 2, \ldots, 7$. Because of (3.3), we can see $|\alpha_s^4|^2 + |\alpha_s^6|^2 \leqslant K_1$. With the help of Lemma 4.1, it follows that

$$\begin{aligned} &\mathbb{E} \sup_{0 \leqslant t \leqslant T} |\Delta Y_t|^2 + \mathbb{E} \int_0^T e^{\beta t} |\Delta X_t|^2 \mathrm{d}t + \mathbb{E} \int_0^T e^{\beta t} t^{2H-1} |\Delta Z_t|^2 \mathrm{d}t \\ \leqslant & 6 e^{6K_1 T(T+4)} (T+4) \Big(\frac{C_1 T^{2H} e^{\beta T}}{LM} + 1 \Big) \mathbb{E} \int_0^T e^{\beta t} |\alpha_t^4 \Delta x_t|^2 + e^{\beta t} |\alpha_t^6 \Delta x_t|^2 \mathrm{d}t \\ \leqslant & 6 e^{6K_1 T(T+4)} (T+4) \Big(\frac{C_1 T^{2H} e^{\beta T}}{LM} + 1 \Big) K_1 \Big\{ \mathbb{E} \int_0^T e^{\beta t} |\Delta x_t|^2 \mathrm{d}t + \mathbb{E} \int_0^T e^{\beta t} t^{2H-1} |\Delta z_t|^2 \mathrm{d}t \Big\}, \end{aligned}$$

where $C_1 := T e^{\frac{T^{2H}}{2HLM} + \frac{K_1 LMT^{2-2H}}{1-H}} + M\Big[1 + e^{\frac{T^{2H}}{2HLM} + \frac{K_1 LMT^{2-2H}}{1-H}} \Big(\frac{T^{2H}}{2HLM} + \frac{K_1 LMT^{2-2H}}{1-H} \Big)\Big]$. Taking a small enough positive number K_1 so that $6 e^{6K_1 T(T+4)} (T+4) \Big(\frac{C_1 T^{2H} e^{\beta T}}{LM} + 1 \Big) K_1 < 1$, it is easy to derive that the mapping Γ is a contraction in $(\tilde{\mathcal{V}}_T \times \tilde{\mathcal{V}}_T^H)$ and has a unique fixed point (X, Z). When Y is the solution of (1.1) with respect to the fixed point (X, Z), (X, Y, Z) is the unique solution to (1.1) naturally.

Remark 4.3. *The system we considered here is not fully coupled, more precisely, the coefficients f and g in the forward equation are independent of Z. If not, the existence and uniqueness of such fully CFBSDEs can still be done, but unfortunately, the dependence will bring new and almost insurmountable difficulties during the course of demonstrating the averaging principle. This will be considered in our future work.*

□

5. SOME A PRIORI ESTIMATES

In this section, we first derive some a priori estimates for solution processes X^ϵ, Z^ϵ and Y^ϵ.

Lemma 5.1. *Suppose that (**A1**)-(**A7**) hold, then for any $T > 0$ there exists a positive constant C such that for all $\epsilon \in (0, 1)$,*

$$\sup_{0 \leqslant t \leqslant T} \mathbb{E} |Y_t^\epsilon|^2 \leqslant |y|^2 + C \tag{5.1}$$

and

$$\sup_{0 \leqslant t \leqslant T} \Big\{ \mathbb{E}[e^{\beta t} |X_t^\epsilon|^2] + \mathbb{E} \int_t^T e^{\beta s} s^{2H-1} |Z_s^\epsilon|^2 \mathrm{d}s \Big\} \leqslant C(1 + |y|^2 + \mathbb{E}|e^{\beta T} \varphi^2(\eta_T)|), \tag{5.2}$$

where C is independent of ϵ.

Proof. For $|Y_t^\epsilon|^2$, we have by the classical Itô's formula

$$d\mathbb{E}|Y_t^\epsilon|^2 = \frac{2}{\epsilon}\mathbb{E}(Y_t^\epsilon f(X_t^\epsilon, Y_t^\epsilon))dt + \frac{1}{\epsilon}\mathbb{E}|g(X_t^\epsilon, Y_t^\epsilon)|^2 dt. \tag{5.3}$$

By **(A3)**, we have

$$2Y_t^\epsilon f(X_t^\epsilon, Y_t^\epsilon) + |g(X_t^\epsilon, Y_t^\epsilon)|^2 \leqslant -\beta_1 |Y_t^\epsilon|^2 + C. \tag{5.4}$$

In terms of (5.3) and (5.4) we have

$$d\mathbb{E}|Y_t^\epsilon|^2 \leqslant -\frac{\beta_1}{\epsilon}\mathbb{E}|Y_t^\epsilon|^2 dt + \frac{C}{\epsilon}dt.$$

Furthermore, we have by Gronwall's inequality (cf.[47, Page 13, 2.4])

$$\mathbb{E}|Y_t^\epsilon|^2 \leqslant |y|^2 e^{-\frac{\beta_1}{\epsilon}t} + \frac{C}{\beta_1}(1 - e^{-\frac{\beta_1}{\epsilon}t}) \leqslant |y|^2 + C, \tag{5.5}$$

which means (5.1) holds.

For $|X_t^\epsilon|^2$, thanks to Theorem 8 of [2], it is easy to get

$$d|X_t^\epsilon|^2 = 2X_t^\epsilon dX_t^\epsilon + 2Z_t^\epsilon \mathbb{D}_t^H X_t^\epsilon dt = -2a(\eta_t, X_t^\epsilon, Y_t^\epsilon, Z_t^\epsilon)X_t^\epsilon dt + 2X_t^\epsilon Z_t^\epsilon dB_t^H + 2Z_t^\epsilon \mathbb{D}_t^H X_t^\epsilon dt.$$

By applying the integration by parts formula (2.4) to $e^{\beta t}|X_t^\epsilon|^2$, we have

$$\begin{aligned} &e^{\beta t}|X_t^\epsilon|^2 + 2\int_t^T e^{\beta s} Z_s^\epsilon \mathbb{D}_s^H X_s^\epsilon ds + \beta \int_t^T e^{\beta s}|X_s^\epsilon|^2 ds \\ =&e^{\beta T}\varphi^2(\eta_T) + 2\int_t^T e^{\beta s} a(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon)X_s^\epsilon ds - 2\int_t^T e^{\beta s} X_s^\epsilon Z_s^\epsilon dB_s^H. \end{aligned} \tag{5.6}$$

By (3.5) and the inequality $2xy \leqslant \frac{1}{K_1}x^2 + K_1 y^2$, it is easy to derive that there is a positive constant C such that

$$2a(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon)X_s^\epsilon \leqslant \frac{C}{K_1}a^2(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon) + \frac{K_1}{C}|X_s^\epsilon|^2 \leqslant C(\tilde{K}^2 + |\eta_s|^2 + |Y_s^\epsilon|^2) + \beta|X_s^\epsilon|^2, \tag{5.7}$$

where $C + \frac{K_1}{C} = \beta$.

Taking expectation on both sides of (5.6) and using (5.7), we have

$$\mathbb{E}[e^{\beta t}|X_t^\epsilon|^2] + 2\mathbb{E}\int_t^T e^{\beta s} Z_s^\epsilon \mathbb{D}_s^H X_s^\epsilon ds \leqslant C(1 + \mathbb{E}|e^{\beta T}\varphi^2(\eta_T)|) + C\mathbb{E}\int_t^T e^{\beta s}(|\eta_s|^2 + |Y_s^\epsilon|^2)ds. \tag{5.8}$$

According to (1.1) and [2, Proposition 24], we are able to deduce that $\mathbb{D}_s^H X_s^\epsilon = \frac{\hat{\sigma}_s}{\sigma_s}Z_s^\epsilon$. Together with (3.9), (3.10), (5.5) and (5.8), it follows that

$$\begin{aligned} \mathbb{E}[e^{\beta t}|X_t^\epsilon|^2] + \frac{2}{M}\mathbb{E}\int_t^T e^{\beta s} s^{2H-1}|Z_s^\epsilon|^2 ds &\leqslant C(1 + \mathbb{E}|e^{\beta T}\varphi^2(\eta_T)|) + C\mathbb{E}\int_t^T e^{\beta s}|Y_s^\epsilon|^2 ds \\ &\leqslant C(1 + |y|^2 + \mathbb{E}|e^{\beta T}\varphi^2(\eta_T)|). \end{aligned} \tag{5.9}$$

By choosing $M \geqslant 2$, the inequality (5.2) can be derived from (5.9). The proof is complete. □

Lemma 5.2. *Suppose that (**A1**)-(**A7**) hold, then for any $T > 0$ there exists a positive constant C such that*

$$\mathbb{E}[e^{\beta t}|X_t^\epsilon - X_{t+h}^\epsilon|^2] + \mathbb{E}\int_t^{t+h} e^{\beta s} s^{2H-1}|Z_s^\epsilon|^2 ds \leqslant Ch, \tag{5.10}$$

for all $t \in [0, T]$, $h \in (0, 1)$ and $t + h \leqslant T$, where C is independent of (ϵ, h).

Proof. It is clear that for all $t \in [0, T]$, $h \in (0, 1)$ and $t + h \leqslant T$,

$$X_t^\epsilon - X_{t+h}^\epsilon = \int_t^{t+h} a(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon)\mathrm{d}s - \int_t^{t+h} Z_s^\epsilon \mathrm{d}B_s^H.$$

Applying Itô's formula to $e^{\beta t}|X_t^\epsilon - X_{t+h}^\epsilon|^2$ (see Theorem 2.4), we have

$$\begin{aligned} &e^{\beta t}|X_t^\epsilon - X_{t+h}^\epsilon|^2 + 2\int_t^{t+h} e^{\beta s} Z_s^\epsilon \mathbb{D}_s^H (X_s^\epsilon - X_{s+h}^\epsilon)\mathrm{d}s + \beta \int_t^{t+h} e^{\beta s}|X_s^\epsilon - X_{s+h}^\epsilon|^2 \mathrm{d}s \\ =&2\int_t^{t+h} e^{\beta s} a(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon)(X_s^\epsilon - X_{s+h}^\epsilon)\mathrm{d}s - 2\int_t^{t+h} e^{\beta s}(X_s^\epsilon - X_{s+h}^\epsilon) Z_s^\epsilon \mathrm{d}B_s^H. \end{aligned} \tag{5.11}$$

Combined with the fact that $\mathbb{D}_s^H(X_s^\epsilon - X_{s+h}^\epsilon) = \frac{\hat{\sigma}_s}{\sigma_s} Z_s^\epsilon$ (cf. [2, Proposition 24]) and (3.9), if we take expectation on both sides of (5.11), this yields

$$\begin{aligned} &\mathbb{E}[e^{\beta t}|X_t^\epsilon - X_{t+h}^\epsilon|^2] + \frac{2}{M}\mathbb{E}\int_t^{t+h} e^{\beta s} s^{2H-1}|Z_s^\epsilon|^2 \mathrm{d}s + \beta \mathbb{E}\int_t^{t+h} e^{\beta s}|X_s^\epsilon - X_{s+h}^\epsilon|^2 \mathrm{d}s \\ \leqslant&2\mathbb{E}\int_t^{t+h} e^{\beta s} a(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon)(X_s^\epsilon - X_{s+h}^\epsilon)\mathrm{d}s. \end{aligned} \tag{5.12}$$

With the help of Young's inequality and (3.5), we get

$$\begin{aligned} &2a(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon)(X_s^\epsilon - X_{s+h}^\epsilon) \\ \leqslant&\frac{1}{\beta}|a(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon)|^2 + \beta|X_s^\epsilon - X_{s+h}^\epsilon|^2 \\ \leqslant&C(1 + |\eta_s|^2 + |X_s^\epsilon|^2 + |Y_s^\epsilon|^2) + \beta|X_s^\epsilon - X_{s+h}^\epsilon|^2. \end{aligned} \tag{5.13}$$

By (5.12) and (5.13), we have

$$\mathbb{E}[e^{\beta t}|X_t^\epsilon - X_{t+h}^\epsilon|^2] + \frac{2}{M}\mathbb{E}\int_t^{t+h} e^{\beta s} s^{2H-1}|Z_s^\epsilon|^2 \mathrm{d}s \leqslant C\mathbb{E}\int_t^{t+h} e^{\beta s}(1 + |\eta_s|^2 + |X_s^\epsilon|^2 + |Y_s^\epsilon|^2)\mathrm{d}s.$$

Thus, the above inequality, (3.10), (5.1) and (5.2) allow to conclude (5.10) by choosing $M \geqslant 2$. □

Next, we introduce two auxiliary processes $(\hat{X}_t^\epsilon, \hat{Y}_t^\epsilon) \in \mathbb{R} \times \mathbb{R}$. Fix a positive number $\delta < 1$ and do a partition of time interval $[0, T]$ of size δ. We construct a process $\hat{Y}_t^\epsilon$, with initial datum $\hat{Y}_0^\epsilon = y$, by means of the equations

$$\mathrm{d}\hat{Y}_t^\epsilon = \frac{1}{\epsilon} f(X_{k\delta}^\epsilon, \hat{Y}_t^\epsilon)\mathrm{d}t + \frac{1}{\sqrt{\epsilon}} g(X_{k\delta}^\epsilon, \hat{Y}_t^\epsilon)\mathrm{d}W_t,\ \hat{Y}_{k\delta}^\epsilon = Y_{k\delta}^\epsilon,$$

for $t \in [k\delta, \min\{(k+1)\delta, T\})$, $k \geqslant 0$, where $X_{k\delta}^\epsilon$ is slow solution process at time $k\delta$, respectively. Denote $\lfloor \cdot \rfloor$ to be the integer function and define the process $\hat{X}_t^\epsilon$ by integral

$$\hat{X}_t^\epsilon = \varphi(\eta_T) + \int_t^T a(\eta_{s(\delta)}, X_{s(\delta)}^\epsilon, \hat{Y}_s^\epsilon, Z_{s(\delta)}^\epsilon)\mathrm{d}s - \int_t^T Z_{s(\delta)}^\epsilon \mathrm{d}B_s^H, \tag{5.14}$$

for $t \in [0, T]$, where $s(\delta) = \lfloor s/\delta \rfloor \delta$ is the nearest breakpoint preceding s. We will establish convergence of the auxiliary processes $\hat{Y}_t^\epsilon$ to the fast solution process Y_t^ϵ and $\hat{X}_t^\epsilon$ to the slow solution process X_t^ϵ, respectively.

Lemma 5.3. *Suppose that* (*A*1)-(*A*7) *hold, then for any* $T > 0$ *there is a positive constant* C *such that*

$$\mathbb{E}|Y_t^\epsilon - \hat{Y}_t^\epsilon|^2 \leqslant C\delta, \tag{5.15}$$

where $t \in [0, T]$.

Proof. Because the proof of this lemma can be concluded from [36, Page 853, (48)] by taking $\gamma(x) = x$, we omit the details. □

Lemma 5.4. *Suppose that* (*A*1)-(*A*7) *hold, then for any* $T > 0$ *there is a positive constant* C *such that*

$$\sup_{0\leqslant t\leqslant T}\Big\{\mathbb{E}[e^{\beta t}|X_t^\epsilon - \hat{X}_t^\epsilon|^2] + \mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z_s^\epsilon - Z_{s(\delta)}^\epsilon|^2\mathrm{d}s\Big\} \leqslant C\delta, \tag{5.16}$$

where C *is independent of* (ϵ, δ).

Proof. Note that

$$X_t^\epsilon - \hat{X}_t^\epsilon = \int_t^T [a(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon) - a(\eta_{s(\delta)}, X_{s(\delta)}^\epsilon, \hat{Y}_s^\epsilon, Z_{s(\delta)}^\epsilon)]\mathrm{d}s - \int_t^T (Z_s^\epsilon - Z_{s(\delta)}^\epsilon)\mathrm{d}B_s^H.$$

By Itô's formula (Theorem 2.4) we have for any $t \in [0, T]$

$$\begin{aligned}
&e^{\beta t}|X_t^\epsilon - \hat{X}_t^\epsilon|^2 + 2\int_t^T e^{\beta s}(Z_s^\epsilon - Z_{s(\delta)}^\epsilon)\mathbb{D}_s^H(X_s^\epsilon - \hat{X}_s^\epsilon)\mathrm{d}s + \beta\int_t^T e^{\beta s}|X_s^\epsilon - \hat{X}_s^\epsilon|^2\mathrm{d}s\\
=&2\int_t^T e^{\beta s}[a(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon) - a(\eta_{s(\delta)}, X_{s(\delta)}^\epsilon, \hat{Y}_s^\epsilon, Z_{s(\delta)}^\epsilon)](X_s^\epsilon - \hat{X}_s^\epsilon)\mathrm{d}s\\
&-2\int_t^T e^{\beta s}(X_s^\epsilon - \hat{X}_s^\epsilon)(Z_s^\epsilon - Z_{s(\delta)}^\epsilon)\mathrm{d}B_s^H.
\end{aligned} \tag{5.17}$$

Combined with the fact that $\mathbb{D}_s^H(X_s^\epsilon - \hat{X}_s^\epsilon) = \frac{\hat{\sigma}_s}{\sigma_s}(Z_s^\epsilon - Z_{s(\delta)}^\epsilon)$ (see [2, Proposition 24]) and (3.9), if we take expectation on both sides of (5.17), this yields

$$\begin{aligned}
&\mathbb{E}[e^{\beta t}|X_t^\epsilon - \hat{X}_t^\epsilon|^2] + \frac{2}{M}\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z_s^\epsilon - Z_{s(\delta)}^\epsilon|^2\mathrm{d}s\\
\leqslant&2\mathbb{E}\int_t^T e^{\beta s}[a(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon) - a(\eta_{s(\delta)}, X_{s(\delta)}^\epsilon, \hat{Y}_s^\epsilon, Z_{s(\delta)}^\epsilon)](X_s^\epsilon - \hat{X}_s^\epsilon)\mathrm{d}s.
\end{aligned} \tag{5.18}$$

By (3.4) and the inequality $2xy \leqslant \frac{1}{C}x^2 + Cy^2$, we obtain

$$\begin{aligned}
&2[a(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon) - a(\eta_{s(\delta)}, X_{s(\delta)}^\epsilon, \hat{Y}_s^\epsilon, Z_{s(\delta)}^\epsilon)](X_s^\epsilon - \hat{X}_s^\epsilon)\\
\leqslant&\frac{s^{2H-1}}{MK_1L}[a(\eta_s, X_s^\epsilon, Y_s^\epsilon, Z_s^\epsilon) - a(\eta_{s(\delta)}, X_{s(\delta)}^\epsilon, \hat{Y}_s^\epsilon, Z_{s(\delta)}^\epsilon)]^2 + \frac{MK_1L}{s^{2H-1}}|X_s^\epsilon - \hat{X}_s^\epsilon|^2\\
\leqslant&\frac{s^{2H-1}}{ML}[|\eta_s - \eta_{s(\delta)}|^2 + |X_s^\epsilon - X_{s(\delta)}^\epsilon|^2 + |Y_s^\epsilon - \hat{Y}_s^\epsilon|^2 + L|Z_s^\epsilon - Z_{s(\delta)}^\epsilon|^2] + \frac{MK_1L}{s^{2H-1}}|X_s^\epsilon - \hat{X}_s^\epsilon|^2.
\end{aligned} \tag{5.19}$$

Now recall that

$$\eta_t = \eta_0 + \int_0^t b_s\mathrm{d}s + \int_0^t \sigma_s\mathrm{d}B_s^H, \quad t \in [0, T].$$

So, with the aid of the inequality $(x + y)^2 \leqslant 2(x^2 + y^2)$, Hölder's inequality, (**A6**), (**A7**) and Proposition 2.2, we can derive that

$$\begin{aligned}
&\mathbb{E}|\eta_s - \eta_{s(\delta)}|^2\\
\leqslant&2\mathbb{E}\Big(\int_{s(\delta)}^s b_u\mathrm{d}u\Big)^2 + 2\mathbb{E}\Big(\int_{s(\delta)}^s \sigma_u\mathrm{d}B_u^H\Big)^2\\
\leqslant&2\delta\mathbb{E}\int_{s(\delta)}^s (b_u)^2\mathrm{d}u + 2\mathbb{E}\Big(\int_0^{s-s(\delta)} \sigma_{u+s(\delta)}\mathrm{d}B_{u+s(\delta)}^H\Big)^2
\end{aligned}$$

$$\leqslant C\delta + 2H(2H-1)\int_0^{s-s(\delta)}\int_0^{s-s(\delta)} |u-v|^{2H-2}\sigma_u\sigma_v \mathrm{d}u\mathrm{d}v$$
$$\leqslant C\delta + C\delta^2\cdot\delta^{2H-2} \leqslant C\delta. \tag{5.20}$$

From Lemma 5.2, Lemma 5.3, (5.18), (5.19) and (5.20), we have

$$\begin{aligned}
&\mathbb{E}[e^{\beta t}|X_t^\epsilon - \hat{X}_t^\epsilon|^2] + \frac{1}{M}\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z_s^\epsilon - Z_{s(\delta)}^\epsilon|^2\mathrm{d}s\\
\leqslant&\mathbb{E}\int_t^T \frac{e^{\beta s}s^{2H-1}}{ML}[|\eta_s-\eta_{s(\delta)}|^2 + |X_s^\epsilon - X_{s(\delta)}^\epsilon|^2 + |Y_s^\epsilon - \hat{Y}_s^\epsilon|^2]\mathrm{d}s + \mathbb{E}\int_t^T \frac{MK_1Le^{\beta s}}{s^{2H-1}}|X_s^\epsilon - \hat{X}_s^\epsilon|^2\mathrm{d}s\\
\leqslant&\frac{C\delta}{ML}\int_t^T e^{\beta s}s^{2H-1}\mathrm{d}s + \mathbb{E}\int_t^T \frac{MK_1Le^{\beta s}}{s^{2H-1}}|X_s^\epsilon - \hat{X}_s^\epsilon|^2\mathrm{d}s\\
\leqslant&C\delta + \mathbb{E}\int_t^T \frac{MK_1Le^{\beta s}}{s^{2H-1}}|X_s^\epsilon - \hat{X}_s^\epsilon|^2\mathrm{d}s.
\end{aligned}$$

Moreover, we have by Gronwall's inequality (cf.[46, Page 581, Corollary 6.62])

$$\sup_{0\leqslant t\leqslant T}\Big\{\mathbb{E}[e^{\beta t}|X_t^\epsilon - \hat{X}_t^\epsilon|^2] + \frac{1}{M}\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z_s^\epsilon - Z_{s(\delta)}^\epsilon|^2\mathrm{d}s\Big\} \leqslant C\delta\exp\Big\{\frac{T^{2-2H}-t^{2-2H}}{2-2H}\Big\} \leqslant C\delta.$$

Taking $M \geqslant 1$, the estimate (5.16) is obtained. □

Secondly, we give some estimates for solution processes $\hat{X}^\epsilon$, Z^ϵ and $\hat{Y}^\epsilon$.

Lemma 5.5. *Let* (*A*1)-(*A*7) *hold. Then for any* $T > 0$ *there is a positive constant C such that*

$$\sup_{0\leqslant t\leqslant T}\Big\{\mathbb{E}[e^{\beta t}|\hat{X}_t^\epsilon|^2] + \mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z_{s(\delta)}^\epsilon|^2\mathrm{d}s\Big\} \leqslant C(1+|y|^2+\mathbb{E}|e^{\beta T}\varphi^2(\eta_T)|) \tag{5.21}$$

and

$$\sup_{0\leqslant t\leqslant T}\mathbb{E}|\hat{Y}_t^\epsilon|^2 \leqslant |y|^2 + C, \tag{5.22}$$

where C is independent of (ϵ,δ).

Proof. Because the proof can follow the same as Lemma 5.1, we omit it. □

Lemma 5.6. *Suppose that* (*A*1)-(*A*7) *hold. Then for any* $T > 0$ *and* $h \in (0,1)$ *there exists a positive constants C such that*

$$\mathbb{E}[e^{\beta t}|\hat{X}_t^\epsilon - \hat{X}_{t+h}^\epsilon|^2] + \mathbb{E}\int_t^{t+h} e^{\beta s}s^{2H-1}|Z_{s(\delta)}^\epsilon|^2\mathrm{d}s \leqslant Ch, \tag{5.23}$$

where $t\in[0,T]$, $t+h\leqslant T$ *and C is independent of* (ϵ,δ).

Proof. Because the proof can follow the same as Lemma 5.2, we omit the details. □

6. AVERAGING PRINCIPLE

In this section, our aim is to derive a strong convergence rate of the averaging principle for Eq.(1.1). Namely, we are going to verify that the sequences $\{X_t^\epsilon : t\geqslant 0\}_{\epsilon>0}$ and $\{Z_t^\epsilon : t\geqslant 0\}_{\epsilon>0}$ strongly converge to the solution processes $\{\bar{X}_t : t\geqslant 0\}$ and $\{\bar{Z}_t : t\geqslant 0\}$ of the averaged system (1.3) as ϵ goes to zero in the corresponding spaces.

By the definition of $\bar{a}$, **(A2)** and **(A4)**, we can get that the mapping $\bar{a} : \mathbb{R}\times\mathbb{R}\times\mathbb{R} \mapsto \mathbb{R}$ is Lipschitz continuous (cf. [25, Lemma 3.12], [36, Lemme 6.1]). By (3.5) and the definition of $\bar{a}$ it is easy to derive that $|\bar{a}(u,x,z)|^2 \leqslant C(1+|u|^2+|x|^2)$. Thus, we can conclude that the well-posedness of the solutions for backward stochastic averaged equations (1.3). In what follows, we shall study the regularity of $\bar{X}$.

Lemma 6.1. *Assume that* $(A1)$-$(A7)$ *hold. For all* $t \in [0, T]$ *and* $h \in (0, 1)$*, there exists a positive constant* C *such that for any* $T > 0$

$$\sup_{0 \leqslant t \leqslant T} \left\{ \mathbb{E}[e^{\beta t} |\bar{X}_t|^2] + \mathbb{E} \int_t^T e^{\beta s} s^{2H-1} |\bar{Z}_s|^2 \mathrm{d}s \right\} \leqslant C(1 + |y|^2 + \mathbb{E}|e^{\beta T} \varphi^2(\eta_T)|), \tag{6.1}$$

$$\mathbb{E}[e^{\beta t} |\bar{X}_t - \bar{X}_{t+h}|^2] + \mathbb{E} \int_t^{t+h} e^{\beta s} s^{2H-1} |\bar{Z}_s|^2 \mathrm{d}s \leqslant Ch, \tag{6.2}$$

where $t + h \leqslant T$ *and* C *is independent of* (ϵ, h).

Proof. The Lemma can be proved in the same way we did for Lemma 5.1 and Lemma 5.2. So we omit the details. □

Next, we shall explore the differences between the solution of backward stochastic averaged equation and backward stochastic auxiliary process $\hat{X}_t^\epsilon$. By the construction of $\hat{Y}_t^\epsilon$ and a time shift transformation, we have for any fixed k and $s \in [0, \delta)$

$$\begin{aligned} \hat{Y}_{s+k\delta}^\epsilon =& \hat{Y}_{k\delta}^\epsilon + \frac{1}{\epsilon} \int_{k\delta}^{k\delta+s} f(X_{k\delta}^\epsilon, \hat{Y}_r^\epsilon) \mathrm{d}r + \frac{1}{\sqrt{\epsilon}} \int_{k\delta}^{k\delta+s} g(X_{k\delta}^\epsilon, \hat{Y}_r^\epsilon) \mathrm{d}W_r \\ =& \hat{Y}_{k\delta}^\epsilon + \frac{1}{\epsilon} \int_0^s f(X_{k\delta}^\epsilon, \hat{Y}_{r+k\delta}^\epsilon) \mathrm{d}r + \frac{1}{\sqrt{\epsilon}} \int_0^s g(X_{k\delta}^\epsilon, \hat{Y}_{r+k\delta}^\epsilon) \mathrm{d}W_r^*, \end{aligned}$$

where $W_t^* = W_{t+k\delta} - W_{k\delta}$ is the shift version of W_t, and hence they have the same distribution. Let $\bar{W}_t$ be a Wiener process and independent of B_t^H and W_t. Construct a process $Y^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon}$ by means of

$$\begin{aligned} Y_{s/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon} =& \hat{Y}_{k\delta}^\epsilon + \int_0^{s/\epsilon} f(X_{k\delta}^\epsilon, Y_r^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon}) \mathrm{d}r + \int_0^{s/\epsilon} g(X_{k\delta}^\epsilon, Y_r^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon}) \mathrm{d}\bar{W}_r \\ =& \hat{Y}_{k\delta}^\epsilon + \frac{1}{\epsilon} \int_0^s f(X_{k\delta}^\epsilon, Y_{r/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon}) \mathrm{d}r + \frac{1}{\sqrt{\epsilon}} \int_0^s g(X_{k\delta}^\epsilon, Y_{r/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon}) \mathrm{d}\bar{\bar{W}}_r^\epsilon, \end{aligned}$$

where $\bar{\bar{W}}_t^\epsilon = \sqrt{\epsilon} \bar{W}_{t/\epsilon}$ is the scaled version of $\bar{W}_t$. By comparing the above two equations, it is easy to derive that

$$(X_{k\delta}^\epsilon, \hat{Y}_{s+k\delta}) \sim (X_{k\delta}^\epsilon, Y_{s/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon}), \quad s \in [0, \delta), \tag{6.3}$$

where $\sim$ denotes coincidence in distribution sense. Set

$$\mathscr{L}_k^\epsilon := \mathbb{E} \left| \int_0^{\delta/\epsilon} e^{\beta k\delta} [a(\eta_{k\delta}, X_{k\delta}^\epsilon, \hat{Y}_{s\epsilon+k\delta}^\epsilon, Z_{k\delta}^\epsilon) - \bar{a}(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon)] \mathrm{d}s \right|^2, \ 0 \leqslant k \leqslant \lfloor T/\delta \rfloor - 1,$$

then we state the critical lemma, which will be used later.

Lemma 6.2. *Suppose that* $(A1)$-$(A7)$ *hold, then for any* $T > 0$ *there exists a positive constant* C *such that*

$$\mathscr{L}_k^\epsilon \leqslant C \frac{\delta}{\epsilon}, \ 0 \leqslant k \leqslant \lfloor T/\delta \rfloor - 1, \tag{6.4}$$

where C *is independent of* (ϵ, δ).

Proof. Let $\mathbb{Q}^y$ denote the probability law of the diffusion process $\{Y_t^x : t \geqslant 0\}$ which is governed by the differential equation

$$\mathrm{d}Y_t^x = f(x, Y_t^x) \mathrm{d}t + g(x, Y_t^x) \mathrm{d}\bar{W}_t.$$

When its initial value is $Y_0^x = y$ and we denote the solution by $Y_t^{x,y}$. The expectation with respect to $\mathbb{Q}^y$ is denoted by $\mathbb{E}^y$. Hence we have $\mathbb{E}^y(\psi(Y_t^x)) = \mathbb{E}(\psi(Y_t^{x,y}))$ for all bounded function ψ. For more details on $\mathbb{Q}^y$ the readers are referred to [53]. First we note that it is easy to show that $\mathscr{L}_k^\epsilon < \infty,\ k = 0, 1, \cdots \lfloor T/\delta \rfloor - 1$. Then, for $k = 0, 1, \cdots \lfloor T/\delta \rfloor - 1$, by Fubini's theorem, we have

$$\begin{aligned}
\mathscr{L}_k^\epsilon \leqslant & \frac{e^{2\beta T}}{\epsilon^2}\mathbb{E}\Big|\int_0^\delta [a(\eta_{k\delta}, X_{k\delta}^\epsilon, \hat{Y}_{s+k\delta}^\epsilon, Z_{k\delta}^\epsilon) - \bar{a}(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon)]\mathrm{d}s\Big|^2 \\
= & \frac{e^{2\beta T}}{\epsilon^2}\mathbb{E}\Big|\int_0^\delta [a(\eta_{k\delta}, X_{k\delta}^\epsilon, Y_{s/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon}, Z_{k\delta}^\epsilon) - \bar{a}(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon)]\mathrm{d}s\Big|^2 \\
= & \frac{e^{2\beta T}}{\epsilon^2}\mathbb{E}\int_0^\delta\int_0^\delta [a(\eta_{k\delta}, X_{k\delta}^\epsilon, Y_{s/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon}, Z_{k\delta}^\epsilon) - \bar{a}(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon)] \\
& \qquad \times [a(\eta_{k\delta}, X_{k\delta}^\epsilon, Y_{\tau/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon}, Z_{k\delta}^\epsilon) - \bar{a}(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon)]\mathrm{d}s\mathrm{d}\tau \\
= & \frac{2e^{\beta T}}{\epsilon^2}\mathbb{E}\int_0^\delta\int_\tau^\delta [a(\eta_{k\delta}, X_{k\delta}^\epsilon, Y_{s/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon}, Z_{k\delta}^\epsilon) - \bar{a}(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon)] \\
& \qquad \times [a(\eta_{k\delta}, X_{k\delta}^\epsilon, Y_{\tau/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon}, Z_{k\delta}^\epsilon) - \bar{a}(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon)]\mathrm{d}s\mathrm{d}\tau,
\end{aligned}$$

where the first equality we used

$$\mathbb{P}\{(\eta_{k\delta}, X_{k\delta}^\epsilon, \hat{Y}_{s+k\delta}, Z_{k\delta}^\epsilon) \in (\cdot)\} = \mathbb{P}\{(\eta_{k\delta}, X_{k\delta}^\epsilon, Y_{s/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon}, Z_{k\delta}^\epsilon) \in (\cdot)\}. \tag{6.5}$$

Indeed, if $\mathbb{P}\{(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon) \in (\cdot)\} = 0$, (6.5) is obvious; on the other hand, if $\mathbb{P}\{(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon) \in (\cdot)\} > 0$, we get

$$\mathbb{P}\{(\eta_{k\delta}, X_{k\delta}^\epsilon, \hat{Y}_{s+k\delta}, Z_{k\delta}^\epsilon) \in (\cdot)\} = \mathbb{P}\{(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon) \in (\cdot)\}\mathbb{P}\{\hat{Y}_{s+k\delta} \in (\cdot)|(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon) \in (\cdot)\} \tag{6.6}$$

and

$$\mathbb{P}\{(\eta_{k\delta}, X_{k\delta}^\epsilon, Y_{s/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon}, Z_{k\delta}^\epsilon) \in (\cdot)\} = \mathbb{P}\{(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon) \in (\cdot)\}\mathbb{P}\{Y_{s/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon} \in (\cdot)|(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon) \in (\cdot)\}. \tag{6.7}$$

In light of $\mathbb{P}\{(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon) \in (\cdot)\} > 0$, we have $\mathbb{P}\{X_{k\delta}^\epsilon \in (\cdot)\} > 0$ by monotonicity of probability. Moreover, by (6.3) and $\mathbb{P}\{X_{k\delta}^\epsilon \in (\cdot)\} > 0$, we obtain

$$\mathbb{P}\{\hat{Y}_{s+k\delta} \in (\cdot)|X_{k\delta}^\epsilon \in (\cdot)\} = \mathbb{P}\{Y_{s/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon} \in (\cdot)|X_{k\delta}^\epsilon \in (\cdot)\}. \tag{6.8}$$

By the tower rule of conditional probability and (6.8), we have

$$\mathbb{P}\{\hat{Y}_{s+k\delta} \in (\cdot)|(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon) \in (\cdot)\} = \mathbb{P}\{Y_{s/\epsilon}^{X_{k\delta}^\epsilon, \hat{Y}_{k\delta}^\epsilon} \in (\cdot)|(\eta_{k\delta}, X_{k\delta}^\epsilon, Z_{k\delta}^\epsilon) \in (\cdot)\}. \tag{6.9}$$

By (6.6), (6.7) and (6.9), it is easy to derive that (6.5) holds.

Set

$$J_k(\tau, s, u, x, y, z) := \mathbb{E}\big[\big(a(u, x, Y_s^{x,y}, z) - \bar{a}(u, x, z)\big) \times \big(a(u, x, Y_\tau^{x,y}) - \bar{a}(u, x, z)\big)\big].$$

In view of Markov property of $Y_t^{x,y}$, we have

$$\begin{aligned}
J_k(\tau, s, x, y, z) = & \mathbb{E}^y\Big\{\mathbb{E}^y\Big[\big(a(u, x, Y_s^x, z) - \bar{a}(u, x, z)\big) \times \big(a(u, x, Y_\tau^x, z) - \bar{a}(u, x, z)\big)\Big|\mathscr{M}_\tau^x\Big]\Big\} \\
= & \mathbb{E}^y\Big\{\big[a(u, x, Y_\tau^x, z) - \bar{a}(u, x, z)\big] \times \mathbb{E}^{Y_\tau^{x,y}}\big[a(u, x, Y_{s-\tau}^x, z) - \bar{a}(u, x, z)\big]\Big\},
\end{aligned}$$

where $\mathscr{M}_t^x$ denotes the σ-field generated by $\{Y_r^x; r \leqslant t\}$, $\mathbb{E}^{Y_\tau^{x,y}}(a(u, x, Y_{s-\tau}^x, z) - \bar{a}(u, x, z))$ means the function $\mathbb{E}^{\tilde{y}}(a(u, x, Y_{s-\tau}^x, z) - \bar{a}(u, x, z))$ evaluated at $\tilde{y} = Y_\tau^{x,y}$. Therefore the Cauchy- Schwarz's inequality yields

$$J_k(\tau, s, u, x, y, z) \leqslant \{\mathbb{E}^y|a(u, x, Y_\tau^x, z) - \bar{a}(u, x, z)|^2\}^{\frac{1}{2}}\{\mathbb{E}^y(\mathbb{E}^{\tilde{y}}(a(u, x, Y_{s-\tau}^x, z) - \bar{a}(u, x, z))^2|_{\tilde{y}=Y_\tau^{x,y}})\}^{\frac{1}{2}}, \tag{6.10}$$

which, with the help of nonlinear growth bound of $(a, \bar{a})$, implies that

$$\begin{aligned}\mathbb{E}^y|a(u, x, Y^x_\tau, z) - \bar{a}(u, x, z)|^2 \leqslant & 2\mathbb{E}^y(|a(u, x, Y^x_\tau, z)|^2 + |\bar{a}(u, x, z)|^2)\\ \leqslant & C(1 + u^2 + x^2 + \mathbb{E}^y|Y^x_\tau|^2)\\ \leqslant & C(1 + u^2 + x^2 + y^2). \end{aligned} \tag{6.11}$$

By (6.10), (6.11) and [36, Page 850, (38)], we find

$$J_k(\tau, s, x, y, z) \leqslant C(1 + u^2 + x^2 + y^2)e^{\frac{-\beta_2(s-\tau)}{2}}. \tag{6.12}$$

Let $\mathcal{M}^\epsilon_{k\delta}$ be the $\sigma-$field generated by $X^\epsilon_{k\delta}$ and $Y^\epsilon_{k\delta}$, which is independent of $\{Y^{x,y}_r : r \geqslant 0\}$. By adopting the approach in [53, Theorem 7.1.2], we can deduce from (6.12) and Lemma 5.1 that

$$\begin{aligned}\mathscr{L}^\epsilon_k \leqslant & \frac{2e^{\beta T}}{\epsilon^2}\int_0^\delta\int_\tau^\delta \mathbb{E}\Big\{\mathbb{E}\Big[\Big(a(\eta_{k\delta}, X^\epsilon_{k\delta}, Y^{X^\epsilon_{k\delta}, Y^\epsilon_{k\delta}}_{s/\epsilon}, Z^\epsilon_{k\delta}) - \bar{a}(\eta_{k\delta}, X^\epsilon_{k\delta}, Z^\epsilon_{k\delta})\Big)\\ & \times\Big(a(\eta_{k\delta}, X^\epsilon_{k\delta}, Y^{X^\epsilon_{k\delta}, Y^\epsilon_{k\delta}}_{\tau/\epsilon}, Z^\epsilon_{k\delta}) - \bar{a}(\eta_{k\delta}, X^\epsilon_{k\delta}, Z^\epsilon_{k\delta})\Big)\Big|\mathcal{M}^\epsilon_{k\delta}\Big]\Big\}\mathrm{d}s\mathrm{d}\tau\\ = & \frac{2e^{\beta T}}{\epsilon^2}\int_0^\delta\int_\tau^\delta \mathbb{E}\Big((J_k(\tau/\epsilon, s/\epsilon, u, x, y, z))\Big|_{(u,x,y,z)=(\eta_{k\delta}, X^\epsilon_{k\delta}, Y^\epsilon_{k\delta}, Z^\epsilon_{k\delta})}\Big)\mathrm{d}s\mathrm{d}\tau\\ \leqslant & \frac{C}{\epsilon^2}\int_0^\delta\int_\tau^\delta e^{-\beta_2(s-\tau)/2\epsilon}\mathrm{d}s\mathrm{d}\tau \leqslant \frac{C}{\epsilon}[\delta - \frac{2\epsilon}{\beta_2^2}(1 - e^{-\frac{\delta\beta_2}{2\epsilon}})], \end{aligned}$$

which completes the proof. □

Lemma 6.3. *Suppose that* (A1)-(A7) *hold. Then we have for any* $T > 0$

$$\sup_{0\leqslant t\leqslant T}\mathbb{E}\Big|\int_0^t e^{\beta s(\delta)}(a(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)}))(\hat{X}^\epsilon_{s(\delta)} - \bar{X}_{s(\delta)})\mathrm{d}s\Big| \leqslant C(\sqrt{\delta} + \sqrt{\frac{\epsilon}{\delta}}), \tag{6.13}$$

where C is a positive constant and independent of (ϵ, δ).

Proof. For any $t \in [0, T]$, there exists an $n_t = \lfloor t/\delta \rfloor$ such that $t \in [n_t\delta, (n_t + 1)\delta \wedge T]$. Therefore, the integral term can be rewritten as

$$\int_0^t e^{\beta s(\delta)}(a(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)}))(\hat{X}^\epsilon_{s(\delta)} - \bar{X}_{s(\delta)})\mathrm{d}s := \Theta_1(t, \epsilon) + \Theta_2(t, \epsilon), \tag{6.14}$$

where

$$\begin{aligned}\Theta_1(t, \epsilon) = & \sum_{k=0}^{n_t-1}\int_{k\delta}^{(k+1)\delta} e^{\beta k\delta}(a(\eta_{k\delta}, X^\epsilon_{k\delta}, \hat{Y}^\epsilon_s, Z^\epsilon_{k\delta}) - \bar{a}(\eta_{k\delta}, X^\epsilon_{k\delta}, Z^\epsilon_{k\delta}))(\hat{X}^\epsilon_{k\delta} - \bar{X}_{k\delta})\mathrm{d}s,\\ \Theta_2(t, \epsilon) = & \int_{n_t\delta}^t e^{\beta n_t\delta}(a(\eta_{n_t\delta}, X^\epsilon_{n_t\delta}, \hat{Y}^\epsilon_s, Z^\epsilon_{n_t\delta}) - \bar{a}(\eta_{n_t\delta}, X^\epsilon_{n_t\delta}, Z^\epsilon_{n_t\delta}))(\hat{X}^\epsilon_{n_t\delta} - \bar{X}_{n_t\delta})\mathrm{d}s. \end{aligned}$$

For all $T > 0$, by the nonlinear growth conditions of the functions $(a, \bar{a})$, (3.10), (5.2), (5.21), (5.22), (6.1), Cauchy-Schwarz's inequality and Fubini's Theorem, we have

$$\begin{aligned}& \mathbb{E}|\Theta_2(t, \epsilon)|\\ = & \mathbb{E}\Big|(\hat{X}^\epsilon_{n_t\delta} - \bar{X}_{n_t\delta})\int_{n_t\delta}^t e^{\beta n_t\delta}(a(\eta_{n_t\delta}, X^\epsilon_{n_t\delta}, \hat{Y}^\epsilon_s, Z^\epsilon_{n_t\delta}) - \bar{a}(\eta_{n_t\delta}, X^\epsilon_{n_t\delta}, Z^\epsilon_{n_t\delta}))\mathrm{d}s\Big|\\ \leqslant & \Big(e^{\beta n_t\delta}\mathbb{E}|\hat{X}^\epsilon_{n_t\delta} - \bar{X}_{n_t\delta}|^2\Big)^{\frac{1}{2}}\Big(\mathbb{E}\Big|\int_{n_t\delta}^t e^{\beta n_t\delta}(a(\eta_{n_t\delta}, X^\epsilon_{n_t\delta}, \hat{Y}^\epsilon_s, Z^\epsilon_{n_t\delta}) - \bar{a}(\eta_{n_t\delta}, X^\epsilon_{n_t\delta}, Z^\epsilon_{n_t\delta}))\mathrm{d}s\Big|^2\Big)^{\frac{1}{2}} \end{aligned}$$

$$
\begin{aligned}
&\leqslant \sqrt{\delta}\Big(\sup_{0\leqslant t\leqslant T} e^{\beta t}\mathbb{E}|\hat{X}_t^{\epsilon}-\bar{X}_t|^2\Big)^{\frac{1}{2}}\Big(e^{\beta t}\mathbb{E}\int_{n_t\delta}^{t} e^{\beta n_t\delta}(a(\eta_{n_t\delta},X^{\epsilon}_{n_t\delta},\hat{Y}^{\epsilon}_s,Z^{\epsilon}_{n_t\delta})-\bar{a}(\eta_{n_t\delta},X^{\epsilon}_{n_t\delta},Z^{\epsilon}_{n_t\delta}))^2\mathrm{d}s\Big)^{\frac{1}{2}}\\
&\leqslant \sqrt{2\delta}\Big[\sup_{0\leqslant t\leqslant T}\mathbb{E}e^{\beta t}(|\hat{X}_t^{\epsilon}|^2+|\bar{X}_t|^2)\Big]^{\frac{1}{2}}\Big(\mathbb{E}\int_{n_t\delta}^{t} e^{\beta n_t\delta}(a^2(\eta_{n_t\delta},X^{\epsilon}_{n_t\delta},\hat{Y}^{\epsilon}_s,Z^{\epsilon}_{n_t\delta})+\bar{a}^2(\eta_{n_t\delta},X^{\epsilon}_{n_t\delta},Z^{\epsilon}_{n_t\delta}))\mathrm{d}s\Big)^{\frac{1}{2}}\\
&\leqslant C\sqrt{\delta}\Big(\mathbb{E}\int_0^T e^{\beta n_t\delta}(a^2(\eta_{n_t\delta},X^{\epsilon}_{n_t\delta},\hat{Y}^{\epsilon}_s,Z^{\epsilon}_{n_t\delta})+\bar{a}^2(\eta_{n_t\delta},X^{\epsilon}_{n_t\delta},Z^{\epsilon}_{n_t\delta}))\mathrm{d}s\Big)^{\frac{1}{2}}\\
&\leqslant C\sqrt{\delta}\Big(\mathbb{E}\int_0^T e^{\beta n_t\delta}(1+|\eta_{n_t\delta}|^2+|X^{\epsilon}_{n_t\delta}|^2+|\hat{Y}^{\epsilon}_s|^2)\mathrm{d}s\Big)^{\frac{1}{2}}\\
&\leqslant \sqrt{\delta}\int_0^T C\mathrm{d}s \leqslant C\sqrt{\delta}. \qquad (6.15)
\end{aligned}
$$

For $\Theta_1(t,\epsilon)$, by (5.21), (6.1), Cauchy-Schwarz's inequality and Fubini's Theorem, it can be deduced that

$$
\begin{aligned}
&\mathbb{E}|\Theta_1(t,\epsilon)|\\
\leqslant&\mathbb{E}\sup_{0\leqslant t\leqslant T}\Big|\sum_{k=0}^{\lfloor t/\delta\rfloor-1}\int_{k\delta}^{(k+1)\delta} e^{\beta k\delta}(a(\eta_{k\delta},X^{\epsilon}_{k\delta},\hat{Y}^{\epsilon}_s,Z^{\epsilon}_{k\delta})-\bar{a}(\eta_{k\delta},X^{\epsilon}_{k\delta},Z^{\epsilon}_{k\delta}))(\hat{X}^{\epsilon}_{k\delta}-\bar{X}_{k\delta})\mathrm{d}s\Big|\\
\leqslant&\sum_{k=0}^{\lfloor T/\delta\rfloor-1}\mathbb{E}\Big|\int_{k\delta}^{(k+1)\delta} e^{\beta k\delta}(a(\eta_{k\delta},X^{\epsilon}_{k\delta},\hat{Y}^{\epsilon}_s,Z^{\epsilon}_{k\delta})-\bar{a}(\eta_{k\delta},X^{\epsilon}_{k\delta},Z^{\epsilon}_{k\delta}))(\hat{X}^{\epsilon}_{k\delta}-\bar{X}_{k\delta})\mathrm{d}s\Big|\\
\leqslant&\frac{T}{\delta}\max_{0\leqslant k\leqslant\lfloor T/\delta\rfloor-1}\mathbb{E}\Big|\int_{k\delta}^{(k+1)\delta} e^{\beta k\delta}(a(\eta_{k\delta},X^{\epsilon}_{k\delta},\hat{Y}^{\epsilon}_s,Z^{\epsilon}_{k\delta})-\bar{a}(\eta_{k\delta},X^{\epsilon}_{k\delta},Z^{\epsilon}_{k\delta}))(\hat{X}^{\epsilon}_{k\delta}-\bar{X}_{k\delta})\mathrm{d}s\Big|\\
\leqslant&\frac{C}{\delta}\max_{0\leqslant k\leqslant\lfloor T/\delta\rfloor-1}(\mathbb{E}|\hat{X}^{\epsilon}_{k\delta}|^2+\mathbb{E}|\bar{X}^{\epsilon}_{k\delta}|^2)^{\frac{1}{2}}\Big[\mathbb{E}\Big|\int_{k\delta}^{(k+1)\delta} e^{\beta k\delta}(a(\eta_{k\delta},X^{\epsilon}_{k\delta},\hat{Y}^{\epsilon}_s,Z^{\epsilon}_{k\delta})-\bar{a}(\eta_{k\delta},X^{\epsilon}_{k\delta},Z^{\epsilon}_{k\delta}))\mathrm{d}s\Big|^2\Big]^{\frac{1}{2}}\\
=&C\frac{\epsilon}{\delta}\max_{0\leqslant k\leqslant\lfloor T/\delta\rfloor-1}\Big[\mathbb{E}\Big|\int_0^{\delta/\epsilon} e^{\beta k\delta}(a(\eta_{k\delta},X^{\epsilon}_{k\delta},\hat{Y}^{\epsilon}_{s\epsilon+k\delta},Z^{\epsilon}_{k\delta})-\bar{a}(\eta_{k\delta},X^{\epsilon}_{k\delta},Z^{\epsilon}_{k\delta}))\mathrm{d}s\Big|^2\Big]^{\frac{1}{2}}\\
=&C\frac{\epsilon}{\delta}\max_{0\leqslant k\leqslant\lfloor T/\delta\rfloor-1}\sqrt{\mathscr{L}^{\epsilon}_k}. \qquad (6.16)
\end{aligned}
$$

Moreover, by (6.4) and (6.16), it can be concluded that

$$
\mathbb{E}|\Theta_1(t,\epsilon)|\leqslant C\sqrt{\frac{\epsilon}{\delta}},
$$

which, taking into account (6.14) and (6.15), provides (6.13). This completes the proof. □

Lemma 6.4. *Suppose that* $(A1)$-$(A7)$ *hold, then*

$$
\sup_{0\leqslant t\leqslant T}\Big\{e^{\beta t}\mathbb{E}|\hat{X}^{\epsilon}_t-\bar{X}_t|^2+\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z^{\epsilon}_{s(\delta)}-\bar{Z}_s|^2\mathrm{d}s\Big\}\leqslant C(\sqrt{\delta}+\sqrt{\frac{\epsilon}{\delta}}),
$$

for any $T>0$*, where C is a positive constant and independent of* (ϵ,δ)*.*

Proof. For any $t\in[0,T]$, by (1.3) and (5.14), we have

$$
\hat{X}^{\epsilon}_t-\bar{X}_t=\int_t^T[a(\eta_{s(\delta)},X^{\epsilon}_{s(\delta)},\hat{Y}^{\epsilon}_s,Z^{\epsilon}_{s(\delta)})-\bar{a}(\eta_s,\bar{X}_s,\bar{Z}_s)]\mathrm{d}s-\int_t^T(Z^{\epsilon}_{s(\delta)}-\bar{Z}_s)\mathrm{d}B^H_s. \qquad (6.17)
$$

Then, applying Itô's formula (Theorem 2.4) to $e^{\beta t}|\hat{X}^{\epsilon}_t - \bar{X}_t|^2$ on $[0, T]$ and using the fact that $\mathbb{D}^H_s(\hat{X}^{\epsilon}_s - \bar{X}_s) = \frac{\hat{\sigma}_s}{\sigma_s}(Z^{\epsilon}_{s(\delta)} - \bar{Z}_s)$ (see [2, Proposition 24]), together with (3.9), we have

$$
\begin{aligned}
&e^{\beta t}|\hat{X}^{\epsilon}_t - \bar{X}_t|^2 + \frac{2}{M}\int_t^T e^{\beta s} s^{2H-1}|Z^{\epsilon}_{s(\delta)} - \bar{Z}_s|^2 \mathrm{d}s + \beta \int_t^T e^{\beta s}|\hat{X}^{\epsilon}_s - \bar{X}_s|^2 \mathrm{d}s \\
\leqslant& 2\int_t^T e^{\beta s}(a(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, \hat{Y}^{\epsilon}_s, Z^{\epsilon}_{s(\delta)}) - \bar{a}(\eta_s, \bar{X}_s, \bar{Z}_s))(\hat{X}^{\epsilon}_s - \bar{X}_s)\mathrm{d}s \\
&- 2\int_t^T e^{\beta s}(\hat{X}^{\epsilon}_s - \bar{X}_s)(Z^{\epsilon}_{s(\delta)} - \bar{Z}_s)\mathrm{d}B^H_s := \sum_{i=1}^{8}\mathcal{U}_i(t),
\end{aligned}
\tag{6.18}
$$

where

$$
\begin{aligned}
\mathcal{U}_1(t) &= 2\int_t^T (e^{\beta s} - e^{\beta s(\delta)})(a(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, \hat{Y}^{\epsilon}_s, Z^{\epsilon}_{s(\delta)}) - \bar{a}(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, Z^{\epsilon}_{s(\delta)}))(\hat{X}^{\epsilon}_s - \bar{X}_s)\mathrm{d}s, \\
\mathcal{U}_2(t) &= 2\int_t^T e^{\beta s(\delta)}(a(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, \hat{Y}^{\epsilon}_s, Z^{\epsilon}_{s(\delta)}) - \bar{a}(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, Z^{\epsilon}_{s(\delta)}))(\hat{X}^{\epsilon}_{s(\delta)} - \bar{X}_{s(\delta)})\mathrm{d}s, \\
\mathcal{U}_3(t) &= 2\int_t^T e^{\beta s(\delta)}(a(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, \hat{Y}^{\epsilon}_s, Z^{\epsilon}_{s(\delta)}) - \bar{a}(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, Z^{\epsilon}_{s(\delta)}))(\hat{X}^{\epsilon}_s - \hat{X}^{\epsilon}_{s(\delta)})\mathrm{d}s, \\
\mathcal{U}_4(t) &= 2\int_t^T e^{\beta s(\delta)}(a(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, \hat{Y}^{\epsilon}_s, Z^{\epsilon}_{s(\delta)}) - \bar{a}(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, Z^{\epsilon}_{s(\delta)}))(\bar{X}_{s(\delta)} - \bar{X}_s)\mathrm{d}s, \\
\mathcal{U}_5(t) &= 2\int_t^T e^{\beta s}[\bar{a}(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, Z^{\epsilon}_{s(\delta)}) - \bar{a}(\eta_s, X^{\epsilon}_s, Z^{\epsilon}_{s(\delta)})](\hat{X}^{\epsilon}_s - \bar{X}_s)\mathrm{d}s, \\
\mathcal{U}_6(t) &= 2\int_t^T e^{\beta s}[\bar{a}(\eta_s, X^{\epsilon}_s, Z^{\epsilon}_{s(\delta)}) - \bar{a}(\eta_s, \hat{X}^{\epsilon}_s, Z^{\epsilon}_{s(\delta)})](\hat{X}^{\epsilon}_s - \bar{X}_s)\mathrm{d}s, \\
\mathcal{U}_7(t) &= 2\int_t^T e^{\beta s}[\bar{a}(\eta_s, \hat{X}^{\epsilon}_s, Z^{\epsilon}_{s(\delta)}) - \bar{a}(\eta_s, \bar{X}_s, \bar{Z}_s)](\hat{X}^{\epsilon}_s - \bar{X}_s)\mathrm{d}s, \\
\mathcal{U}_8(t) &= -2\int_t^T e^{\beta s}(\hat{X}^{\epsilon}_s - \bar{X}_s)(Z^{\epsilon}_{s(\delta)} - \bar{Z}_s)\mathrm{d}B^H_s.
\end{aligned}
$$

For $\mathcal{U}_1(t)$, by the nonlinear growth conditions of the functions $(a, \bar{a})$, Cauchy-Schwarz's inequality, Young's inequality, Fubini's Theorem, mean valve Theorem, (3.10), (5.2) and (5.22), we have

$$
\begin{aligned}
&\mathbb{E}|\mathcal{U}_1(t)| \\
\leqslant& \mathbb{E}\int_t^T 2(e^{\beta s} - e^{\beta s(\delta)})|a(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, \hat{Y}^{\epsilon}_s, Z^{\epsilon}_{s(\delta)}) - \bar{a}(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, Z^{\epsilon}_{s(\delta)})||\hat{X}^{\epsilon}_s - \bar{X}_s|\mathrm{d}s \\
\leqslant& \mathbb{E}\int_t^T 2\beta e^{\beta s}\delta(|a(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, \hat{Y}^{\epsilon}_s, Z^{\epsilon}_{s(\delta)})| + |\bar{a}(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, Z^{\epsilon}_{s(\delta)})|)|\hat{X}^{\epsilon}_s - \bar{X}_s|\mathrm{d}s \\
\leqslant& C\delta\int_t^T (e^{\beta s}\mathbb{E}|a(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, \hat{Y}^{\epsilon}_s, Z^{\epsilon}_{s(\delta)})|^2 + e^{\beta s}\mathbb{E}|\bar{a}(\eta_{s(\delta)}, X^{\epsilon}_{s(\delta)}, Z^{\epsilon}_{s(\delta)})|^2)^{\frac{1}{2}}(e^{\beta s}\mathbb{E}|\hat{X}^{\epsilon}_s - \bar{X}_s|^2)^{\frac{1}{2}}\mathrm{d}s \\
\leqslant& C\delta\int_t^T e^{\beta s}(\mathbb{E}(1 + |\eta_{s(\delta)}|^2 + |X^{\epsilon}_{s(\delta)}|^2 + |\hat{Y}^{\epsilon}_s|^2))^{\frac{1}{2}}(e^{\beta s}\mathbb{E}|\hat{X}^{\epsilon}_s - \bar{X}_s|^2)^{\frac{1}{2}}\mathrm{d}s \\
\leqslant& C\delta\int_t^T e^{\beta s}\mathbb{E}(1 + |\eta_{s(\delta)}|^2 + |X^{\epsilon}_{s(\delta)}|^2 + |\hat{Y}^{\epsilon}_s|^2) + C\int_t^T (e^{\beta s}\mathbb{E}|\hat{X}^{\epsilon}_s - \bar{X}_s|^2)\mathrm{d}s
\end{aligned}
$$

$$\leqslant C\delta + C\int_t^T (e^{\beta s}\mathbb{E}|\hat{X}_s^\epsilon - \bar{X}_s|^2)\mathrm{d}s. \tag{6.19}$$

For $\mathcal{U}_2(t)$, by the elementary inequality $|x-y| \leqslant |x| + |y|$ and (6.13), we have

$$\begin{aligned}
&\mathbb{E}|\mathcal{U}_2(t)|\\
=&\mathbb{E}\Big|2\int_0^T e^{\beta s(\delta)}[a(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)})](\hat{X}^\epsilon_{s(\delta)} - \bar{X}_{s(\delta)})\mathrm{d}s\\
&-2\int_0^t e^{\beta s(\delta)}[a(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)})](\hat{X}^\epsilon_{s(\delta)} - \bar{X}_{s(\delta)})\mathrm{d}s\Big|\\
\leqslant&\mathbb{E}\Big|2\int_0^T e^{\beta s(\delta)}[a(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)})](\hat{X}^\epsilon_{s(\delta)} - \bar{X}_{s(\delta)})\mathrm{d}s\Big|\\
&+\mathbb{E}\Big|2\int_0^t e^{\beta s(\delta)}[a(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)})](\hat{X}^\epsilon_{s(\delta)} - \bar{X}_{s(\delta)})\mathrm{d}s\Big|\\
\leqslant&C(\sqrt{\delta} + \sqrt{\frac{\epsilon}{\delta}}).
\end{aligned}\tag{6.20}$$

For $\mathcal{U}_3(t)$, by the nonlinear growth conditions of the functions $(a, \bar{a})$, Cauchy-Schwarz's inequality, Fubini's Theorem, (3.10), (5.2), (5.22) and (5.23), we have

$$\begin{aligned}
&\mathbb{E}|\mathcal{U}_3(t)|\\
\leqslant&\mathbb{E}\int_t^T 2e^{\beta s}|a(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)})||\hat{X}^\epsilon_s - \hat{X}^\epsilon_{s(\delta)}|\mathrm{d}s\\
\leqslant&\mathbb{E}\int_t^T 2e^{\beta s}(|a(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s, Z^\epsilon_{s(\delta)})| + |\bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)})|)|\hat{X}^\epsilon_s - \hat{X}^\epsilon_{s(\delta)}|\mathrm{d}s\\
\leqslant&C\int_t^T (e^{\beta s}\mathbb{E}|a(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s, Z^\epsilon_{s(\delta)})|^2 + e^{\beta s}\mathbb{E}|\bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)})|^2)^{\frac{1}{2}}(e^{\beta s}\mathbb{E}|\hat{X}^\epsilon_s - \hat{X}^\epsilon_{s(\delta)}|^2)^{\frac{1}{2}}\mathrm{d}s\\
\leqslant&C\int_t^T e^{\beta s}\mathbb{E}(1 + |\eta_{s(\delta)}|^2 + |X^\epsilon_{s(\delta)}|^2 + |\hat{Y}^\epsilon_s|^2)(e^{\beta s}\mathbb{E}|\hat{X}^\epsilon_s - \hat{X}^\epsilon_{s(\delta)}|^2)^{\frac{1}{2}}\mathrm{d}s\\
\leqslant&C\int_t^T (e^{\beta s}\mathbb{E}|\hat{X}^\epsilon_s - \hat{X}^\epsilon_{s(\delta)}|^2)^{\frac{1}{2}}\mathrm{d}s \leqslant C\delta^{\frac{1}{2}}.
\end{aligned}\tag{6.21}$$

For $\mathcal{U}_4(t)$, by the nonlinear growth conditions of the functions $(a, \bar{a})$, Cauchy-Schwarz's inequality, Fubini's Theorem, (3.10), (5.2), (5.22) and (6.2), we have

$$\begin{aligned}
&\mathbb{E}|\mathcal{U}_4(t)|\\
\leqslant&\mathbb{E}\int_t^T 2e^{\beta s}|a(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)})||\bar{X}^\epsilon_s - \bar{X}^\epsilon_{s(\delta)}|\mathrm{d}s\\
\leqslant&\mathbb{E}\int_t^T 2e^{\beta s}(|a(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s, Z^\epsilon_{s(\delta)})| + |\bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)})|)|\bar{X}^\epsilon_s - \bar{X}^\epsilon_{s(\delta)}|\mathrm{d}s\\
\leqslant&C\int_t^T (e^{\beta s}\mathbb{E}|a(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s, Z^\epsilon_{s(\delta)})|^2 + e^{\beta s}\mathbb{E}|\bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)})|^2)^{\frac{1}{2}}(e^{\beta s}\mathbb{E}|\bar{X}^\epsilon_s - \bar{X}^\epsilon_{s(\delta)}|^2)^{\frac{1}{2}}\mathrm{d}s\\
\leqslant&C\int_t^T e^{\beta s}\mathbb{E}(1 + |\eta_{s(\delta)}|^2 + |X^\epsilon_{s(\delta)}|^2 + |\hat{Y}^\epsilon_s|^2)(e^{\beta s}\mathbb{E}|\bar{X}^\epsilon_s - \bar{X}^\epsilon_{s(\delta)}|^2)^{\frac{1}{2}}\mathrm{d}s\\
\leqslant&C\int_t^T (e^{\beta s}\mathbb{E}|\bar{X}^\epsilon_s - \bar{X}^\epsilon_{s(\delta)}|^2)^{\frac{1}{2}}\mathrm{d}s \leqslant C\delta^{\frac{1}{2}}.
\end{aligned}\tag{6.22}$$

For $\mathcal{U}_5(t)$, by Young's inequality, Fubini's Theorem, (5.10), (5.20) and the Lipschitz property of $\bar{a}$, we have

$$\begin{aligned}
&\mathbb{E}|\mathcal{U}_5(t)|\\
\leqslant&\mathbb{E}\int_t^T 2|e^{\beta s(\delta)}(\bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_s, X^\epsilon_s, Z^\epsilon_{s(\delta)}))(\hat{X}^\epsilon_s - \bar{X}_s)|\mathrm{d}s\\
\leqslant&\mathbb{E}\int_t^T e^{\beta s}|\bar{a}(\eta_{s(\delta)}, X^\epsilon_{s(\delta)}, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_s, X^\epsilon_s, Z^\epsilon_{s(\delta)})|^2\mathrm{d}s + \int_t^T e^{\beta s}\mathbb{E}|\hat{X}^\epsilon_s - \bar{X}_s|^2\mathrm{d}s\\
\leqslant&C\int_t^T e^{\beta s}(\mathbb{E}|\eta^\epsilon_s - \eta^\epsilon_{s(\delta)}|^2 + \mathbb{E}|X^\epsilon_s - X^\epsilon_{s(\delta)}|^2 + \mathbb{E}|\hat{X}^\epsilon_s - \bar{X}_s|^2)\mathrm{d}s.\\
\leqslant&C\delta + C\int_t^T e^{\beta s}\mathbb{E}|\hat{X}^\epsilon_s - \bar{X}_s|^2\mathrm{d}s. \qquad (6.23)
\end{aligned}$$

For $\mathcal{U}_6(t)$, by Young's inequality, Fubini's Theorem, (5.16) and the Lipschitz property of $\bar{a}$, we have

$$\begin{aligned}
&\mathbb{E}|\mathcal{U}_6(t)|\\
\leqslant&\mathbb{E}\int_t^T 2|e^{\beta s}(\bar{a}(\eta_s, X^\epsilon_s, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_s, \hat{X}^\epsilon_s, Z^\epsilon_{s(\delta)}))(\hat{X}^\epsilon_s - \bar{X}_s)|\mathrm{d}s\\
\leqslant&\mathbb{E}\int_t^T e^{\beta s}[\bar{a}(\eta_s, X^\epsilon_s, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_s, \hat{X}^\epsilon_s, Z^\epsilon_{s(\delta)})]^2\mathrm{d}s + \int_t^T e^{\beta s}\mathbb{E}|\hat{X}^\epsilon_s - \bar{X}_s|^2\mathrm{d}s\\
\leqslant&C\int_t^T e^{\beta s}(\mathbb{E}|X^\epsilon_s - \hat{X}^\epsilon_s|^2 + \mathbb{E}|\hat{X}^\epsilon_s - \bar{X}_s|^2)\mathrm{d}s.\\
\leqslant&C\delta + C\int_t^T e^{\beta s}\mathbb{E}|\hat{X}^\epsilon_s - \bar{X}_s|^2\mathrm{d}s. \qquad (6.24)
\end{aligned}$$

For $\mathcal{U}_7(t)$, by Young's inequality, Fubini's Theorem, (5.16) and the Lipschitz property of $\bar{a}$, we have

$$\begin{aligned}
&\mathbb{E}|\mathcal{U}_7(t)|\\
\leqslant&\mathbb{E}\int_t^T 2|e^{\beta s}(\bar{a}(\eta_s, \hat{X}^\epsilon_s, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_s, \bar{X}_s, \bar{Z}_s))(\hat{X}^\epsilon_s - \bar{X}_s)|\mathrm{d}s\\
\leqslant&\frac{1}{MC}\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}[\bar{a}(\eta_s, \hat{X}^\epsilon_s, Z^\epsilon_{s(\delta)}) - \bar{a}(\eta_s, \bar{X}_s, \bar{Z}_s)]^2\mathrm{d}s + MC\int_t^T \frac{e^{\beta s}}{s^{2H-1}}\mathbb{E}|\hat{X}^\epsilon_s - \bar{X}_s|^2\mathrm{d}s\\
\leqslant&\frac{1}{M}\int_t^T e^{\beta s}s^{2H-1}\mathbb{E}(C|\hat{X}^\epsilon_s - \bar{X}_s|^2 + |Z^\epsilon_{s(\delta)} - \bar{Z}_s|^2)\mathrm{d}s + MC\int_t^T \frac{e^{\beta s}}{s^{2H-1}}\mathbb{E}|\hat{X}^\epsilon_s - \bar{X}_s|^2\mathrm{d}s\\
\leqslant&C\int_t^T (1 + \frac{1}{s^{2H-1}})e^{\beta s}\mathbb{E}|\hat{X}^\epsilon_s - \bar{X}_s|^2\mathrm{d}s + \frac{1}{M}\int_t^T e^{\beta s}s^{2H-1}\mathbb{E}|Z^\epsilon_{s(\delta)} - \bar{Z}_s|^2\mathrm{d}s. \qquad (6.25)
\end{aligned}$$

Here, we shall prove that $\mathbb{E}\mathcal{U}_8(t)$ is equal to zero. From Proposition 2.2, it suffices to check $e^{\beta s}(\hat{X}^\epsilon_s - \bar{X}_s)(Z^\epsilon_{s(\delta)} - \bar{Z}_s) \in \mathbb{L}^{1,2}_H$. Indeed, we have $\mathcal{V}_T \subset \mathbb{L}^{1,2}_H$ (see Lemma 7 in [2]) and $e^{\beta s}(\hat{X}^\epsilon_s - \bar{X}_s)(Z^\epsilon_{s(\delta)} - \bar{Z}_s) \in \mathcal{V}_T$ from (3.11). So

$$\mathbb{E}\mathcal{U}_8(t) = -2\mathbb{E}\int_t^T e^{\beta s}(\hat{X}^\epsilon_s - \bar{X}_s)(Z^\epsilon_{s(\delta)} - \bar{Z}_s)\mathrm{d}B^H_s = 0. \qquad (6.26)$$

Now taking expectation on both sides of (6.18) and employing (6.20)-(6.26), we find that

$$\begin{aligned}&e^{\beta t}\mathbb{E}|\hat{X}^\epsilon_t-\bar{X}_t|^2+\frac{1}{M}\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z^\epsilon_{s(\delta)}-\bar{Z}_s|^2\mathrm{d}s\\ \leqslant& C(\sqrt{\delta}+\sqrt{\frac{\epsilon}{\delta}})+C\int_t^T(1+\frac{1}{s^{2H-1}})e^{\beta s}\mathbb{E}|\hat{X}^\epsilon_s-\bar{X}_s|^2\mathrm{d}s,\end{aligned}$$

which, with the aid of Gronwall's inequality (see [46, Page 581, Corollary 6.62]), yields

$$\begin{aligned}&\sup_{0\leqslant t\leqslant T}\left\{e^{\beta t}\mathbb{E}|\hat{X}^\epsilon_t-\bar{X}_t|^2+\frac{1}{M}\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z^\epsilon_{s(\delta)}-\bar{Z}_s|^2\mathrm{d}s\right\}\\ \leqslant& C(\sqrt{\delta}+\sqrt{\frac{\epsilon}{\delta}})\exp\left\{CT+C\frac{T^{2-2H}}{2-2H}\right\}.\end{aligned}$$

This completes the proof. □

We are now in a position to give the proof of Theorem 1.1.

Proof. By Lemma 5.4 and Lemma 6.4 and taking $\delta=\sqrt{\epsilon}$ with $\epsilon\in(0,1)$, we have

$$\begin{aligned}&\sup_{0\leqslant t\leqslant T}\left\{e^{\beta t}\mathbb{E}|X^\epsilon_t-\bar{X}_t|^2+\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z^\epsilon_s-\bar{Z}_s|^2\mathrm{d}s\right\}\\ =&\sup_{0\leqslant t\leqslant T}\left\{e^{\beta t}\mathbb{E}|X^\epsilon_t-\hat{X}^\epsilon_t+\hat{X}^\epsilon_t-\bar{X}_t|^2+\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|(Z^\epsilon_s-Z^\epsilon_{s(\delta)})+(Z^\epsilon_{s(\delta)}-\bar{Z}_s)|^2\mathrm{d}s\right\}\\ \leqslant&2\sup_{0\leqslant t\leqslant T}\left\{e^{\beta t}\mathbb{E}|X^\epsilon_t-\hat{X}^\epsilon_t|^2+\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z^\epsilon_s-Z^\epsilon_{s(\delta)}|^2\mathrm{d}s\right\}\\ &+2\sup_{0\leqslant t\leqslant T}\left\{e^{\beta t}\mathbb{E}|\hat{X}^\epsilon_t-\bar{X}_t|^2+\mathbb{E}\int_t^T e^{\beta s}s^{2H-1}|Z^\epsilon_{s(\delta)}-\bar{Z}_s|^2\mathrm{d}s\right\}\\ \leqslant& C(\sqrt{\delta}+\sqrt{\frac{\epsilon}{\delta}})\leqslant C\epsilon^{\frac{1}{4}},\end{aligned}$$

which finishes the proof. □

Remark 6.5. *It should be pointed out that the convergence rate is $\epsilon^{\frac{1}{8}}$, which is independent of the Hurst parameter H of fBm. It seems strange at the first sight, but indeed it is reasonable due to the fact that $\delta+\delta^{2H}\leqslant 2\delta$ for $\delta\in(0,1)$ and $H\in(1/2,1)$(see (5.20) for details).*

ACKNOWLEDGEMENTS

The first author acknowledges Professor Jinqiao Duan careful reading of manuscript, correcting errors, detailed comments and valuable suggestions, which improve the quality of this paper.

DATA AVAILABILITY STATEMENT

All data, models, and code generated or used during the study appear in the submitted article.

DECLARATION OF INTERESTS

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

REFERENCES

[1] Y. Hu, S. Peng, Backward stochastic differential equation driven by fractional Brownian motion, SIAM Journal on Control and Optimization, 48(3) (2009) 1675-1700.

[2] L. Maticiuc, T. Nie, Fractional backward stochastic differential equations and fractional backward variational inequalities, Journal of Theoretical Probability, 28(1)(2015) 337-395.

[3] J. Wen, Y. Shi, Anticipative backward stochastic differential equations driven by fractional Brownian motion, Statistics Probability Letters, 122(2017) 118-127.

[4] S. Douissi, J. Wen, Y. Shi, Mean-field anticipated BSDEs driven by fractional Brownian motion and related stochastic control problem, Applied Mathematics and Computation, 355(2019) 282-298.

[5] R. Z. Khasminskii, On the principle of averaging the Itô's stochastic differential equations, kibernetika. 4 (1968) 260-279 (in Russian).

[6] A. Benveniste, M. Métivier, P. Priouret, Adaptive Algorithms and Stochastic Approximations, New York: Springer-Verlag, 1990.

[7] H. J. Kushner, G. Yin, Stochastic Approximation and Recursive Algorithms and Applications, 2nd ed. New York: Spinger-Verlag, 2003.

[8] L. Luo, E. Schuster, Mixing enhancement in 2D magnetohydrodynamic channel flow by extremum seeking boundary control, in Proc. Amer. Control Conf., St. Louis, MO, Jun. 10-12, 2009.

[9] V. Solo, X. Kong, Adaptive Signal Processing Algorithms: Stability and Performance, Englewood Cliffs, NJ: Prentice Hall, 1994.

[10] J. C. Spall, Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control, New York: Wiley-Interscience, 2003.

[11] F. Wu, T. Tian, J.B. Rawlings, G. Yin, Approximate method for stochastic chemical kinetics with two-time scales by chemical Langevin equations, J. Chem. Phys., 144(17)(2016) 174112.

[12] Y. Kifer, Stochastic versions of Anosov and Neistadt theorems on averaging, Stoch. Dyn., 1(1) (2001) 1-21.

[13] D. Givon, I.G. Kevrekidis, Multiscale integration schemes for jump-diffusion systems, SIAM J. Multi. Model. Simul., 7 (2008) 495-516.

[14] S. Liu, M. Krstic, Stochastic averaging in continuous time and its applications to extremum Seeking, IEEE Transactions on Automatic Control, 55(10) (2010) 2235-2250.

[15] S. Liu, M. Krstic, Stochastic averaging in discrete time and its applications to extremum Seeking, IEEE Transactions on Automatic Control, 61(10) (2016) 90-102.

[16] L. Wang, X. Han, Y. Cao, HN. Najm, Computational singular perturbation analysis of stochastic chemical systems with stiffness, Journal of Computational Physics, 335 (2017) 404-425.

[17] R. Z. Khasminskii, G. Yin, On averaging principles: an asymptotic expansion approach, SIAM J. Math. Anal., 35 (2004) 1534-1560.

[18] W. E, D. Liu and E. Vanden-Eijnden, Analysis of multiscale methods for stochastic differential equations, Commun, Pure Appl. Math., 58(1) (2005) 1544-1585.

[19] D. Givon, Strong convergence rate for two-time-scale jump-diffusion stochastic differential systems, SIAM J. Multi. Model. Simul., 6 (2) (2007) 577-594.

[20] D. Liu, Strong convergence of principle of averaging for multiscale stochastic dynamical systems, Commun. Math. Sci., 8 (4) (2010) 999-1020.

[21] D. Liu, Strong convergence rate of principle of averaging for jump-diffusion processes, Front. Math. China., 7 (2) (2012) 305-320.

[22] X. Li, An averaging principle for a completely integrable stochastic Hamiltonian system, Nonlinearity, 21 (2008) 803-822.

[23] G. Wainrib, Double averaging principle of periodically forced slow-fast stochastic systems, Electron. Commun. Probab., 18 (51) (2013) 1-12.

[24] S. Liu, M. Krstic, Continuous-time stochastic averaging on the infinite interval for locally Lipschitz systems, SIAM Journal on Control and Optimization, 48 (5)(2010) 3589-3622.

[25] W. Liu, M. Röckner, X. Sun, Y. Xie, Averaging principle for slow-fast stochastic differential equations with time dependent locally Lipschitz coefficients, Journal of Differential Equations, 268(6) (2020) 2910-2948.

[26] F. Wu, G. Yin, An averaging principle for two-time-scale stochastic functional differential equations, Journal of Differential Equations, 269(2020) 1037-1077.

[27] M. Röckner, X. Sun. Y. Xie, Strong convergence order for slow-fast McKean-Vlasov stochastic differential equations, Ann. Inst. H. Poincaré Probab. Statist., 57 (1)(2021) 547-576.

[28] X. Sun, L. Xie, Y. Xie, Averaging principle for slow-fast stochastic partial differential equations with Hölder continuous coefficients, Journal of Differential Equations, 270 (2021) 476-504.

[29] F. Feo, The averaging principle for non-autonomous slow-fast stochastic differential equations and an application to a local stochastic volatility model, Journal of Differential Equations, 302(2021) 406-443.

[30] B. Pei, Y. Inahama, Y. Xu, Averaging principles for mixed fast-slow systems driven by fractional brownian motion, arXiv:2001.06945v3, accepted by Kyoto Journal of Mathematics, 2021.

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

[32] M. Hairer, X. Li, Averaging dynamics driven by fractional Brownian motion, Ann. Probab., 48 (4) (2020) 1826-1860.

[33] X. Li, J. Sieber, Slow-fast systems with fractional environment and dynamics, arXiv preprint arXiv:2012.01910, 2020.

[34] M. Röckner, L. Xie, Averaging principle and normal deviations for multiscale stochastic systems, Communications in Mathematical Physics, 383 (2021) 1889-1937.

[35] M. Röckner, L. Xie, Diffusion approximation for fully coupled stochastic differential equations, Ann. Probab. 49 (3) (2021) 1205-1236.

[36] J. Xu, J. Liu, J. Liu, Y. Miao, Strong averaging principle for two-time-scale stochastic McKean-Vlasov equations, Applied Mathematics and Optimization, 84(2021) 837-867.

[37] J. Bao, G. Yin, C. Yuan, Two-time-scale stochastic partial differential equations driven by α-stable noises: Averaging principles, Bernoulli, 23(1) (2017) 645-669.

[38] F. Wu, G. Yin, Fast-slow-coupled stochastic functional differential equations, Journal of Differential Equations, 323(2022) 1-37.

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

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

[41] C. E. Bréhier, Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component, Stochastic Process. Appl., 130 (2020) 3325-3368.

[42] E. Pardoux, A.Yu. Veretennikov, Averaging of backward stochastic differential equations, with application to semi-linear PDE's, Stochastics and Stochastic Reports, 60 (1997) 255-270.

[43] E. H. Essaky, Y. Ouknine, Averaging of backward stochastic differential equations and homogenization of partial differential equations with periodic coefficients, Stoch. Anal. Appl., 24 (2) (2006) 277-301.

[44] K. Bahlalia, A. Elouaflin, E. Pardoux, Averaging for BSDEs with null recurrent fast component. Application to homogenization in a non periodic media, Stochastic Processes and their Applications, 127(4)(2017) 1321-1353.

[45] K. Bahlali, A. Elouaflin, E. Pardoux, Homogenization of semilinear PDEs with discontinuous averaged coefficients, Electron. J. Probab., 14 (2009) 477-499.

[46] E. Pardoux, A. Rascanu, Stochastic differential equations, Backward SDEs, Partial differential equations, Springer, (2016).

[47] J. Duan, W. Wang, Effective dynamics of stochastic partial differential equations, Elsevier, 2014.

[48] Y. Hu, Integral transformations and anticipative calculus for fractional Brownian motions, Memoirs of the American Mathematical Society, 175(825)(2005) viii+127.

[49] F. Biagini, Y. Hu, B. Øksendal, T. Zhang, Stochastic calculus for fractional Brownian motion and applications , Springer, 2008.

[50] Y. Mishura, Stochastic calculus for fractional Brownian motion and related processes, Springer, 2008.

[51] D. Nualart, The Malliavin calculus and related topics, Springer, 2006.

[52] M. Sanz-Solé, Malliavin calculus with applications to stochastic partial differential equations, EPFL press, 2005.

[53] B. Øksendal, Stochastic Differential Equations, sixth ed., Springer, Berlin, 2003.

[54] F. Xi, C. Zhu, Jump type stochastic differential equations with non-lipschitz coefficients : Non confluence, feller and strong feller properties, and exponential ergodicity, Journal of Differential Equations, 266(8)(2019) 4668-4711.