A Study of SDEs Driven by Brownian Motion and Fractional Brownian Motion

Yongqiang Suo

Student no. [redacted]

Supervisor: Prof. Chenggui Yuan

Submitted to Swansea University in fulfilment of the requirements for the Degree of Doctor of Philosophy

Swansea University

2021

Abstract

In this thesis, we mainly study some properties for certain stochastic differential equations.

The types of stochastic differential equations we are interested in are (i) stochastic differential equations driven by Brownian motion, (ii) stochastic functional differential equations driven by fractional Brownian motion, (iii) McKean-Vlasov stochastic differential equations driven by Brownian motion, (iv) McKean-Vlasov stochastic differential equations driven by fractional Brownian motion.

The properties we investigate include the weak approximation rate of Euler-Maruyama scheme, the central limit theorem and moderate deviation principle for McKean-Vlasov stochastic differential equations. Additionally, we investigate the existence and uniqueness of solution to McKean-Vlasov stochastic differential equations driven by fractional Brownian motion, and then the Bismut formula of Lion's derivatives for this model is also obtained.

The crucial method we utilised to establish the weak approximation rate of Euler-Maruyama scheme for stochastic equations with irregular drift is the Girsanov transformation. More precisely, giving a reference stochastic equations, we construct the equivalent expressions between the aim stochastic equations and associated numerical stochastic equations in another probability spaces in view of the Girsanov theorem.

For the Mckean-Vlasov stochastic differential equation model, we first construct the moderate deviation principle for the law of the approximation stochastic differential equation in view of the weak convergence method. Subsequently, we show that the approximation stochastic equations and the McKean-Vlasov stochastic differential equations are in the same exponentially equivalent family, and then we establish the moderate deviation principle for this model.

Based on the result of Well-posedness for Mckean-Vlasov stochastic differential equation driven by fractional Brownian motion, by using the Malliavin analysis, we first establish a general result of the Bismut type formula for Lions derivative, and then we apply this result to the non-degenerate case of this model.

Keywords: Weak approximation; moderate deviation principle; central limit theorem; Lions derivative; Bismut formula

Declarations and Statements

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed [redacted] .... Date 12/07/2021

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed [redacted] .... Date 12/07/2021

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed [redacted] .... Date 12/07/2021

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans after expiry of a bar on access approved by the Swansea University.

Signed [redacted] .... Date 12/07/2021

Contents

Acknowledgements

First of all, I am very thankful to my supervisor Professor Chenggui Yuan for the patience and effective guidance he provided during my study and the preparation of this doctoral thesis.

I registered as a PhD student in Swansea University in the April of 2018. I appreciate the three-year financial support from the Swansea University. This meant a lot to me, because I could have a good study environment free from stress of financial pressure. I am also grateful for the support and understanding of my parents for my study abroad.

I would like to thank Professors Elian Crooks, Eugene Lytvynov, Andrew Neate, Irina Rodionova, Feng-Yu Wang and Jianglun Wu, who gave wonderful courses and seminars, and the visiting researcher Shao-qin Zhang, who gave me many useful guidance when learning new research materials. I acknowledge the inspiration of the questions from Professors Jianhai Bao, Xiliang Fan and Xing Huang.

It is my pleasure to thank my friends Dimitra Georgopoulou and Danis Kiziridis who gave me great help during my first few months in Swansea. I thank my office colleagues Jane Allwright, Yini Du, Xuan Gao and Chi Zhang, for the communication and help.

Swansea, May, 2021

Notations

$a.s.$:	almost surely, or $\mathbb{P}-$ almost surely, or with probability 1.
$\mathbb{N}$:	set of positive integer numbers $\{1, 2, 3\cdots\}$.
$a \vee b$:	the larger number between a and b in $\mathbb{R}$.
$a \wedge b$:	the smaller number between a and b in $\mathbb{R}$.
$[a]$:	the integer part of a.
$\mathbb{R}^d$:	$d-$ dimensional Euclidean space.
$\langle \cdot, \cdot \rangle$:	the usual inner product on $\mathbb{R}^d$.
$\lvert \cdot \rvert$:	the norm on $\mathbb{R}^d$, corresponding with respect to $\langle \cdot, \cdot \rangle$.
$\mathscr{B}_b(\mathbb{R}^d)$:	the collection of all bounded measurable functions on $\mathbb{R}^d$.
$C^\alpha([a,b];\mathbb{R}^d)$:	the space of $\alpha-$ Hölder continuous functions on $[a,b]$.
$\lVert f \rVert_{a,b,\alpha}$:	the Hölder norm of function f.
$B_x(R)$:	ball with radius R and center x.
$\mathscr{C}$:	the space of continuous functions, i.e. $C([a,b];\mathbb{R}^d)$.
$\lVert f \rVert_p$:	$= \left(\int_{\mathbb{R}^d} \lvert f \rvert^p \mathrm{d}x \right)^{\frac{1}{p}}$ for $p \geq 1$.
$\lVert f \rVert_\infty$:	$= \sup_{x \in \mathbb{R}^d} \lvert f(x) \rvert$.
$\mathscr{P}_p(\mathbb{R}^d)$:	the space of probability measures on $\mathbb{R}^d$ with finite $p-$ th moment.
$Ran(a)$:	the range of matrix a.

Chapter 1

Introduction

In this thesis, we mainly investigate the weak approximation rate of Euler-Maruyama (abbreviated as EM) scheme for certain stochastic differential equations (shorted as SDEs) with irregular coefficients, the moderate deviation principle, and the Bismut formula of Lions derivative to some McKean-Vlasov stochastic differential equations (MV-SDEs for short).

In the following part, we describe the background of the main results of this thesis in more detail.

I Weak approximation rate of EM scheme for stochastic differential equations driven by Brownian motion

SDEs with singular coefficients have been extensively studied recently, see [47, 61, 97, 99, 100, 103] and references therein. Meanwhile, in order to understand the numerical approximation of SDEs with irregular coefficients, numerical schemes have been established. The strong and weak convergence rates of EM scheme for SDEs with singular coefficients have already been obtained, see for instance [4, 5, 38, 45, 46, 48, 49, 50, 76, 79, 80, 89]. The references [23, 51, 71, 72, 73, 77, 84] investigated L^p-approximation of solutions to the SDEs with singular drift, and obtained the corresponding L^p-error rates under the differential assumptions about the coefficients. In specific, reference [72] obtained the L^p-error rate to be at least $1/2$ with $p \in [1, \infty)$ for the scalar SDEs with a piecewise Lipschitz drift, and a Lipschitz diffusion coefficient that is non-zero at the discontinuity points of the drift coefficient. This result has been extended to the case of scalar jump-diffusion SDEs in [84]. Based on the assumptions in [72, 84], [71, 73] showed that the L^p-error rate is at least $3/4$ under additional piecewise smoothness assumptions on the coefficients, where they employed a novel technique by studying equations with coupled noise. They additionally showed that the $3/4$ L^p-error rate cannot in general be improved even when additionally to the assumptions in [73] further piecewise regularity assumptions were imposed on the coefficients of the scalar SDEs. Under the condition of the Sobolev-Slobodeckij-type regularity of order $\kappa \in (0,1)$, [77] obtained the L^2-error rate $\min\{3/4, (1+\kappa)/2\} - \varepsilon$ (for arbitrarily small $\varepsilon > 0$) of the equidistant EM scheme for scalar SDEs with irregular drift and additive noise by using an explicit the Zvonkin-type transformation and the Girsanov transformation.

For simplicity, we generalise the main idea of weak approximation in the following case of stochastic differential equation:

$$X_t = x + \int_0^t b(X_s)\mathrm{d}s + \sigma W_t,\ x \in \mathbb{R}^d,\ t \in [0,T], \tag{1.0.1}$$

where $(W_t)_{t\geq 0}$ is a d-dimensional standard Brownian motion, which is defined in a certain probability measure space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$. Moreover, $b : \mathbb{R}^d \to \mathbb{R}^d$ is a Borel measurable function, σ is a $d \times d$ deterministic, uniformly elliptic matrix and $a := \sigma\sigma^*$.

Let $X_t^{(\delta)}, \delta > 0$ denote the EM approximation of X_t,

$$X_t^{(\delta)} = x + \int_0^t b(X_{s_\delta})\mathrm{d}s + \sigma W_t,\ t \in [0,T],$$

where $s_\delta = [\frac{s}{\delta}]\delta$, $[a]$ denotes the integer part of a.

The weak convergence rate is concerned with the approximation of $\mathbb{E}f(X_t)$ by $\mathbb{E}f(X_t^{(\delta)})$ for a given function f.

To this end, we introduce a reference SDE as follows:

$$Y_t = x + \sigma W_t, t \in [0, T],$$

and

$$\begin{aligned} R_{1,t} &= \exp\Big\{ \int_0^t -\sigma^{-1} b(X_s) \mathrm{d}W_s - \frac{1}{2}\int_0^t |\sigma^{-1} b(X_s)|^2 \mathrm{d}s \Big\}, \\ R_{2,t} &= \exp\Big\{ \int_0^t -\sigma^{-1} b(X_{s_\delta}) \mathrm{d}W_s - \frac{1}{2}\int_0^t |\sigma^{-1} b(X_{s_\delta})|^2 \mathrm{d}s \Big\}. \end{aligned}$$

Under appropriate conditions, the Novikov condition holds, we then can define new measures $\mathbb{Q}_1, \mathbb{Q}_2$ as $\mathrm{d}\mathbb{Q}_1 = R_{1,t}\mathrm{d}\mathbb{P}$ and $\mathrm{d}\mathbb{Q}_2 = R_{2,t}\mathrm{d}\mathbb{P}$. Based on this, one can show an equivalent expression of $\mathbb{E}[f(X_t)] - \mathbb{E}[f(X_{t_\delta})]$ by using the reference equation, that is,

$$\mathbb{E}[f(X_t)] - \mathbb{E}[f(X_{t_\delta})] = \mathbb{E}[f(Y_t)(R_{1,t} - R_{2,t})].$$

The weak error has been obtained for some SDEs with discontinuous drifts in [45, 46, 78]. It is worth noting that the test function f in these references is assumed to be Hölder continuous. When the test function f was relaxed to be just measurable and bounded, the result of weak convergence rate of EM scheme was obtained in [3], where the coefficients of SDEs need to be smooth.

Recently, [7, 89] established the weak convergence rate of EM scheme for SDEs with irregular coefficients by using Girsanov's transformation. Inspired by [5] and [7, 89], we shall give a note on the weak error for (1.0.1) with a possibly discontinuous drift b. Moreover, the given function f is only assumed to be bounded and measurable on $\mathbb{R}^d$.

In Chapter 3, we establish the weak approximation rate of EM scheme for (1.0.1) with a class of low-regular drift by using the Girsanov transformation. In specific, the class of low-regular drift contains cases of non-Lipschitz continuous function, discontinuous function). Moreover, we also give some illustrative examples.

II Weak approximation rate of EM scheme for stochastic functional differential equations driven by fractional Brownian motion

The fractional Brownian motion (abbreviated as fBm) appears naturally in modelling stochastic systems with long-range dependence phenomena. Fractional Brownian motions with Hurst parameter $H \neq 1/2$ are neither Markov processes nor (weak) semi-martingales, which makes the study of stochastic differential equations driven by fractional noise complicated. The existence and uniqueness of solutions to fractional equations have received much attention. [59] obtained existence and uniqueness of solutions to SDEs driven by fractional noise with Hurst parameter $H \in (\frac{1}{2}, 1)$ by using Young integrals (see [101]) and p-variation estimate. [22] derived the existence and uniqueness result for $H \in (\frac{1}{4}, \frac{1}{2})$ through the same rough-type arguments as in [59]. [83] studied SDEs driven by fractional noise by using fractional calculus developed in [102]. For more results on existence and uniqueness of solutions to SDEs driven by fractional noise, we refer to for instance [13, 43, 44, 55, 67, 68, 82, 92]. Stochastic functional differential equations (SFDE for short) are also used to characterise stochastic systems with memory effects. For the existence and uniqueness of solutions for SFDEs with regular coefficients, one can consult to [33, 69, 75]. In recent years, SDEs driven by fBm with irregular coefficients have received much attention, e.g. [31, 44, 65, 66]. However, for fractional SFDEs with irregular coefficients, even the weak existence and uniqueness results are not well studied.

Consider the following SFDE:

$$\mathrm{d}X(t) = \{b(X(t)) + \sigma Z(X_t)\}\mathrm{d}t + \sigma \mathrm{d}B^H(t), t > 0, \tag{1.0.2}$$

with the initial datum $X_0 = \xi \in \mathscr{C} := C([-\tau, 0]; \mathbb{R}^d)$, where $\sigma \in \mathbb{R}^d \otimes \mathbb{R}^m$, $b : \mathbb{R}^d \to \mathbb{R}^d, d \geq m$ and $Z : \mathscr{C} \to \mathbb{R}^m$ are measurable, X_t is the segment process of $X(t)$ defined by $X_t(\theta) = X(t + \theta), \theta \in [-\tau, 0]$. $B^H(t)$ is an m-dimensional fBm on a complete probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$.

In chapter 4, we first study the weak existence and uniqueness for (1.0.2) (see Theorem 4.1.1). Based on the weak existence and uniqueness result, we investigate the weak approximation rate of EM scheme for (1.0.2) by using a measurable bounded test function (see Theorem 4.3.1). The exponential integrability of functionals of the segment process is studied in our work involves fractional calculus, which is nontrivial for the irregular drift with memory, and it is more complicated than those of SFDEs driven by Brownian motion. The main ingredient is giving exact estimates for fractional derivatives of functionals of the segment process truncated by gridpoints (see Lemma 4.4.2).

III Central limit theorem and Moderate deviation principle for McKean-Vlasov stochastic differential equations driven by Brownian motion

As it is well known, the large deviation principle (abbreviated as LDP) is a branch of probability theory that deals with the asymptotic behaviour of rare events. In the case of stochastic process, the idea is to find a deterministic path around which the diffusion is concentrated with high probability, and the stochastic motion can be interpreted as a small perturbation of the deterministic path. Moreover, it has a wide range of applications, such as in mathematical finance, statistic mechanics and biology. Thus, the LDP for stochastic equations has been investigated extensively; e.g. see [8, 9, 40, 70] and references therein.

There are two main methods to investigate the large deviations, one method is based on contraction principle in LDP, that is, it relies on approximation arguments and exponential-type probability estimates; e.g. see [12, 35, 36, 41, 54, 56, 70, 88] and references therein. [35, 56, 88] were concerned about the LDP for SDEs driven by Brownian motion or Poisson measure. In [41], the authors therein investigated how rapid-switching behaviour of solution(X_t^ϵ) affects the small-noise asymptotics of X_t^ϵ-modulated diffusion processes on the certain interval. [36] investigated the LDP for invariant distributions of memory gradient diffusions. Other method is the weak convergence one, which has also been applied in establishing LDP for a various stochastic dynamic systems; e.g. see [8, 9, 15, 16, 17, 18]. According to the compactness argument in this method of the solution space of the corresponding skeleton equation, the weak convergence is done for Borel measurable functions whose existence is based on the Yamada-Watanabe theorem. In [15, 16, 18], the authors study an LDP for SDEs and stochastic partial differential equations (SPDEs for short).

Compared with the theory of large deviations, the central limit theorem (abbreviated as CLT) is interested in the asymptotic behaviour of stochastic motion tends to the corresponding deterministic path in the smallest deviation scale. Similarly, the moderate deviation principle (MDP for short) is concerned with probabilities with a smaller order than in the LDP, which deviation scale fills in the gap between the CLT scale and the LDP scale (see [60]).

To explain these deviations, we introduce the general deviation for MV-SDEs. This is the topic of Chapter 5.

Consider the following MV-SDE on $(\mathbb{R}^d, \langle\cdot,\cdot\rangle, |\cdot|)$:

$$\mathrm{d}X_t^\epsilon = b_t(X_t^\epsilon, \mathscr{L}_{X_t^\epsilon})\mathrm{d}t + \sqrt{\epsilon}\sigma_t(X_t^\epsilon, \mathscr{L}_{X_t^\epsilon})\mathrm{d}W_t, \quad X_0^\epsilon = x, \tag{1.0.3}$$

with $\epsilon > 0$, which is called the scaling parameter. Here W_t is the d-dimensional Brownian motion defined on a complete filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$ and $\mathscr{L}_{X_t^\epsilon}$ is the law of X_t^ϵ.

Intuitively, as the parameter ϵ tends to 0 in (5.2.1), the diffusion term vanishes, and we have the following ordinary differential equation (ODE for short):

$$\mathrm{d}X_t^0 = b_t(X_t^0, \delta_{X_t^0})\mathrm{d}t, \tag{1.0.4}$$

with the same initial datum as (1.0.3), that is, $X_0^0 = x$. Since x is deterministic, we deduce that $\delta_{X_\cdot^0}$ is a Dirac measure centered on the path $X_\cdot^0$.

On the general case, investigating the deviations of solution X_t^ϵ to (1.0.3) from the solution X_t^0 to (1.0.4) is to study the asymptotic behaviour of the trajectory

$$\bar{X}_t^\epsilon = \frac{1}{\sqrt{\epsilon}\lambda(\epsilon)}(X_t^\epsilon - X_t^0), \ t \in [0, T]. \tag{1.0.5}$$

More precisely, as the parameter $\lambda(\varepsilon)$ taking value in different scale, we have the following three cases:

(LDP) The case $\lambda(\epsilon) = 1/\sqrt{\epsilon}$ provides some large deviation estimates. [28] proved that the law of the solution X^ϵ satisfies an LDP by means of the discussion of exponential tightness.

(CLT) If $\lambda(\epsilon) \equiv 1$, we shall show that $\frac{X^\epsilon - X^0}{\sqrt{\epsilon}}$ converges to a stochastic process in a certain sense as $\epsilon \to 0$ (see Theorem 5.3.1).

(MDP) To fill in the gap between the CLT scale and LDP scale, the MDP for X^ϵ is to investigate the LDP of trajectory (1.0.5), where the deviation scale $\lambda(\epsilon)$ satisfies

$$\lambda(\epsilon) \to \infty, \quad \sqrt{\epsilon}\lambda(\epsilon) \to 0, \quad \text{as } \epsilon \to 0. \tag{1.0.6}$$

When coefficients in (1.0.3) are distribution independent, it reduces to the case of classical SDEs. These deviation theories have been intensively investigated in the literature mentioned above and references therein.

Now, we introduce the main contents of Chapter 5, which is devoted to the study of CLT and MDP for MV-SDEs.

The motivation comes from the Freidlin-Wentzell LDP in path space for MV-SDEs in [28]. The authors in [58] investigated large and moderate deviation principles for MV-SDEs with jumps.

In Chapter 5, we first show that the law of solution to a good approximation SDE of the underlying MV-SDE satisfies a MDP via weak convergence method. Then, we show that the solution to an approximation SDE and the solution to the MV-SDE are exponentially equivalent as the deviation scale tends to zero. Then, the exponentially equivalent measures are indistinguishable in the case of LDP yielding our main result.

IV Bismut formula of Lions derivative for McKean-Vlasov stochastic differential equation driven by fractional Brownian motion

McKean-Vlasov (distribution dependent or mean field) stochastic differential equations has been studied intensively in the literature [26, 27, 64, 89, 96] and references therein. The existence of this type of SDEs has been investigated by different methods: [34] used a approximation argument about particle systems to obtain the existence and uniqueness of weak solutions to MV-SDEs; in [96], by iterating in distributions, a strong solution is constructed using SDEs with control; finally, [85] applied the fixed point theorem to establish the existence of strong solutions to MV-SDEs.

The Bismut formula initiated in [11] due to Bismut, is also called Bismut-Elworty-Li formula due to the development by Elworthy and Li in [29]. Since then, it has become a powerful tool to characterise the regularity of distribution for SDEs and SPDEs. The Bismut formula have been derived by different methods: for instance [29] by martingale method; and [98] using the coupling method (or Malliavin calculus) and references therein. It is worthy noting that these results are derived for the classical SDEs.

The Lions derivative (L-derivative for short) was introduced by P.-L. Lions in his lecture [19]. Since then, researchers have applied the L-derivative of solution to this type of SDEs to characterise the properties of partial differential equations, such as [14, 39, 52] and references therein. Recently, [6, 87] used Malliavin calculus to derive the Bismut formula for L-derivative of MV-SDEs and MV-SDEs with memory. It is worthy pointing that the existing literature about the Bismut formula for SDEs and SFDEs driven by fractional noise [2, 30, 32] only works for classical case (distribution independent). Thus, in Chapter 6, we aim to obtain the existence and uniqueness of solutions to MV-SDEs driven by fBm, and establish a general result for the Bismut formula for L-derivative for this type of stochastic equations.

Chapter 2

Preliminaries

In this chapter, we will give some preliminary knowledge, which will be used in the following chapters. In Section 2.1, we recall the fractional integrals and derivatives. In Section 2.2, we introduce the fBm and the Malliavin calculus with respect to fBm. In Section 2.3, we introduce the definition of L-derivative of measure function. Section 2.4 is devoted to the background of the LDP and some of its properties.

2.1 Fractional integrals and derivatives

In this section, we recall some basics of fractional integrals and derivatives, and for more details, see [81, 89].

Let $a, b \in \mathbb{R}$ with $a < b$. For $f \in L^1(a, b)$ and $\alpha > 0$, the left-sided fractional Riemann-Liouville integral of order α of f on $[a, b]$ is given by

$$I_{a+}^{\alpha} f = \frac{1}{\Gamma(\alpha)} \int_a^x \frac{f(y)}{(x-y)^{1-\alpha}} \mathrm{d}y,$$

where $x \in (a, b)$ a.e. $(-1)^{-\alpha} = \mathrm{e}^{-i\alpha\pi}$, and Γ denotes the Euler function. If $\alpha = n \in \mathbb{N}$, this definition coincides with the n-order iterated integrals of f.

By the definition, we have the first composition formula

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

Fractional differentiation may be introduced as an inverse operation. Let $\alpha \in (0,1)$ and $p \geq 1$. If $f \in I^{\alpha}_{a+}(L^p([a,b],\mathbb{R}))$, then the function ϕ satisfying $f = I^{\alpha}_{a+}\phi$ is unique in $L^p([a,b],\mathbb{R})$ and it coincides with the left-sided Riemann-Liouville derivative of f of order α given by

$$D^{\alpha}_{a+}f(x) = \frac{1}{\Gamma(1-\alpha)}\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x \frac{f(y)}{(x-y)^{\alpha}}\mathrm{d}y.$$

The corresponding Weyl representation reads as follows:

$$D^{\alpha}_{a+}f(x) = \frac{1}{\Gamma(1-\alpha)}\Big(\frac{f(x)}{(x-a)^{\alpha}} + \alpha\int_a^x \frac{f(x)-f(y)}{(x-y)^{1+\alpha}}\mathrm{d}y\Big), \tag{2.1.1}$$

where the convergence of the integrals at the singularity $y = x$ holds pointwise for almost all x if $p = 1$ and in the L^p sense if $p > 1$. By the construction, we have

$$I^{\alpha}_{a+}(D^{\alpha}_{a+}f) = f, \qquad f \in I^{\alpha}_{a+}(L^p([a,b],\mathbb{R})),$$

and moreover it holds the second composition formula

$$D^{\alpha}_{a+}(D^{\beta}_{a+}f) = D^{\alpha+\beta}_{a+}f, \qquad f \in I^{\alpha+\beta}_{a+}(L^1([a,b],\mathbb{R})).$$

2.2 Fractional Brownian motion

We first recall some basic facts about the stochastic calculus of variations with respect to the fBm with Hurst parameter $H \in (\frac{1}{2},1)$. We refer the reader to [24] for further details.

For fixed $T > 0$, the d-dimensional fBm $B^H = \{B^H(t), t \in [0,T]\}$ with Hurst parameter H on a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ can be defined as the centered Gaussian process with covariance function

$$\mathbb{E}(B^H(t)B^H(s)) = R_H(t,s) = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H}).$$

In particular, if $H = \frac{1}{2}$, B^H is a Brownian motion. Besides, for $p \geq 1$, we have

$$\mathbb{E}|B^H(t) - B^H(s)|^p = \mathbb{E}|B^H(t-s)|^p = |t-s|^{pH}\mathbb{E}|B^H(1)|^p \leq C(p)|t-s|^{pH}.$$

Then, it follows from the Kolmogorov continuity theorem that B^H has β-Hölder continuous paths, where $\beta \in (0, H)$. For each $t \in [0,T]$, we denote by $\mathscr{F}_t$ the σ-algebra generated by $\{B^H(s) : s \in [0,t]\}$ and the $\mathbb{P}$-null sets.

We denote by $\mathscr{E}$ the set of step functions on $[0,T]$. Let $\mathcal{H}$ be the Hilbert space defined as the closure of $\mathscr{E}$ with respect to the scalar product

$$\langle (I_{[0,t_1]}, \cdots, I_{[0,t_d]}), (I_{[0,s_1]}, \cdots, I_{[0,s_d]}) \rangle_{\mathcal{H}} = \sum_{i=1}^{d} R_H(t_i, s_i).$$

The mapping $(I_{[0,t_1]}, \cdots, I_{[0,t_d]}) \mapsto (B^{H,1}(t_1), \cdots, B^{H,d}(t_d))$ can be extended to an isometry between $\mathcal{H}$ associated with fBm B^H and the Gaussian space $\mathcal{H}_1$. That is, $\mathcal{H}_1$ is a closed subspace whose elements are zero-mean Gaussian random variables. This allows to define the Wiener integrals with respect to B^H, and denote this isometry by $\phi \mapsto B^H(\phi) = \int_0^T \phi(t)\mathrm{d}B^H(t)$.

On the other hand, from [24], we know the covariance kernel $R_H(t,s)$ can be written as

$$R_H(t,s) = \int_0^{t\wedge s} K_H(t,r)K_H(s,r)\mathrm{d}r,$$

where K_H is a square integrable kernel given by

$$K_H(t,s) = \Gamma(H+\frac{1}{2})^{-1}(t-s)^{H-\frac{1}{2}}F(H-\frac{1}{2}, \frac{1}{2}-H, H+\frac{1}{2}, 1-\frac{t}{s}),$$

in which $F(\cdot,\cdot,\cdot,\cdot)$ is Gauss hypergeometric function (see [24]).

Define the linear operator $K_H^*: \mathscr{E} \to L^2([0,T],\mathbb{R}^d)$ as follows:

$$(K_H^*\phi)(s) = K_H(T,s)\phi(s) + \int_s^T (\phi(r)-\phi(s))\frac{\partial K_H}{\partial r}(r,s)\mathrm{d}r.$$

Reformulating the above equality as follows:

$$(K_H^*\phi)(s) = \int_s^T \phi(r)\frac{\partial K_H}{\partial r}(r,s)\mathrm{d}r.$$

It can be shown that for all $\phi, \psi \in \mathscr{E}$,

$$\langle K_H^*\phi, K_H^*\psi\rangle_{L^2([0,T],\mathbb{R}^d)} = \langle \phi, \psi\rangle_{\mathcal{H}},$$

and therefore K_H^* is an isometry between $\mathcal{H}$ and $L^2([0,T],\mathbb{R}^d)$. Consequently, B^H has the following integral representation:

$$B^H(t) = \int_0^t K_H(t,s)\mathrm{d}W(s),$$

where $\{W(t) := B^H((K_H^*)^{-1}I_{[0,t]})\}$ is a standard Brownian motion.

According to [24], the operator $K_H: L^2([0,T],\mathbb{R}^d) \to I_{0+}^{H+\frac{1}{2}}(L^2([0,T],\mathbb{R}^d))$ associated with the kernel $K_H(\cdot,\cdot)$ is defined as follows:

$$(K_H f^i)(t) = \int_0^t K_H(t,s)f^i(s)\mathrm{d}s, \quad i = 1,\cdots,d. \tag{2.2.1}$$

It can be proved that K_H is an isomorphism. Moreover, for each $f \in L^2([0,T],\mathbb{R}^d)$,

$$(K_H f)(s) = I_{0+}^1 s^{H-1/2} I_{0+}^{H-1/2} s^{1/2-H} f, H > \frac{1}{2}. \tag{2.2.2}$$

Consequently, for each $h \in I_{0+}^{H+1/2}(L^2([0,T],\mathbb{R}^d))$, the inverse operator K_H^{-1} is of the form

$$(K_H^{-1}h)(s) = s^{H-1/2} D_{0+}^{H-1/2} s^{1/2-H} h', \ H > \frac{1}{2}. \tag{2.2.3}$$

In what follows, we give a brief account on the Malliavin calculus for fBm. Let Ω be the canonical probability space $C_0([0,T];\mathbb{R}^d)$, the set of continuous functions, null at time 0, equipped with the supremum norm. Let $\mathbb{P}$ be the unique probability measure on Ω such that the canonical process $\{B^H(t); t \in [0,T]\}$ is a d-dimensional fBm with the Hurst parameter H. Subsequently, we will make this assumption on the underlying probability space.

Denote $\mathcal{S}$ by the set of smooth and cylindrical random variables of the form

$$F = f(B^H(\phi_1), \cdots, B^H(\phi_d)),$$

where $d \geq 1, f \in C_b^\infty(\mathbb{R}^d)$, which is the collection of f and all its partial derivatives are bounded, $\phi_i \in \mathcal{H}, 1 \leq i \leq d$. The Malliavin derivative of F, denoted by $\mathbb{D}F$, is defined as the $\mathcal{H}$-valued random variable

$$\mathbb{D}F = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(B^H(\phi_1), \cdots, B^H(\phi_d))\phi_i.$$

For any $p \geq 1$, we define the Sobolev space $\mathbb{D}^{1,p}$ as the completion of $\mathcal{S}$ with respect to the norm

$$\|F\|_{1,p}^p = \mathbb{E}|F|^p + \mathbb{E}\|\mathbb{D}F\|_{\mathcal{H}}^p.$$

For more details and applications on Malliavin calculus with respect to fBm one may refer to [24, 81].

The following lemma is the Fernique-type lemma (see [62, 90]), and some notation for future use.

Lemma 2.2.1. *Let $T > 0, 1/2 < \beta < H < 1$. Then for any $\alpha < \frac{1}{2T}$,*

$$\mathbb{E}\exp\{\alpha\|B^H\|_{0,T,\infty}^2\} < \infty,$$

and for any $\alpha < 1/(128(2T)^{2(H-\beta)})$,

$$\mathbb{E}[\exp(\alpha\|B^H\|_{0,T,\beta}^2)] \leq (1 - 128\alpha(2T)^{2(H-\beta)})^{-1/2}.$$

Moreover, we have the following moment estimate for any $k \geq 1$:

$$\mathbb{E}(\|B^H\|_{0,T,\beta}^{2k}) \leq 32^k (2T)^{2k(H-\beta)} \frac{(2k)!}{k!}.$$

Meanwhile, we will denote by δ and $\mathrm{Dom}\delta$ the divergence operator of $\mathbb{D}$ and its domain, respectively. Let us finish this part by giving a transfer principle that connects the derivative and divergence operators of both processes fBm B^H and Brownian motion W that are needed later on.

Proposition 2.2.2. *[81, Proposition 5.2.1] For any $F \in \mathbb{D}_W^{1,2} = \mathbb{D}^{1,2}$,*

$$K_H^* \mathbb{D} F = \mathbb{D}^W F,$$

where $\mathbb{D}^W$ denotes the derivative operator with respect to the underlying Wiener process W, and $\mathbb{D}_W^{1,2}$ the corresponding Sobolev space.

Proposition 2.2.3. *[81, Proposition 5.2.2] $\mathrm{Dom}\delta = (K_H^*)^{-1}(\mathrm{Dom}\delta_W)$, and for any $\mathcal{H}$-valued random variable u in $\mathrm{Dom}\delta$ we have $\delta(u) = \delta_W(K_H^* u)$, where δ_W denotes the divergence operator with respect to the underlying Wiener process W.*

Remark 2.2.1. The above proposition, together with [81, Proposition 1.3.11], yields that if $K_H^* u \in L_a^2([0,T] \times \Omega, \mathbb{R}^d)$ (the closed subspace of $L^2([0,T] \times \Omega, \mathbb{R}^d)$ formed by the adapted processes), then $u \in \mathrm{Dom}\delta$.

2.3 L-derivative of measure function

We first recall the definition of L-derivative (for more details, see [6, 87]). Let $\mathscr{P}_2(\mathbb{R}^d)$ be the set of all probability measures on $\mathbb{R}^d$ with finite second moment, i.e.

$$\mathscr{P}_2(\mathbb{R}^d) = \Big\{\mu \in \mathscr{P}(\mathbb{R}^d) : \mu(|\cdot|^2) := \int_{\mathbb{R}^d} |x|^2 \mu(\mathrm{d}x) < \infty\Big\},$$

where $\mu(f) := \int f \mathrm{d}\mu$ for a measurable function f. Then $\mathscr{P}_2(\mathbb{R}^d)$ is a Polish space under the Wasserstein distance

$$\mathbb{W}_2(\mu,\nu) := \inf_{\pi \in \mathscr{C}(\mu,\nu)} \Big(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2 \pi(\mathrm{d}x, \mathrm{d}y) \Big)^{\frac{1}{2}}, \quad \mu, \nu \in \mathscr{P}_2(\mathbb{R}^d),$$

where $\mathscr{C}(\mu,\nu)$ is the set of couplings for μ and ν.

For any $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, the tangent space at μ is given by

$$T_{\mu,2} = L^2(\mathbb{R}^d \to \mathbb{R}^d; \mu) := \{\phi : \mathbb{R}^d \to \mathbb{R}^d \text{ is measurable with } \mu(|\phi|^2) < \infty\},$$

which is a Hilbert space under the norm $\|\phi\|_{T_{\mu,2}} := \big(\mu(|\phi|^2)\big)^{\frac{1}{2}}$, (see [87]).

There are many monographs on L-derivative of measure function, see for instance, [19, 86, 87]. We will now introduce the basic case of L-derivative of measure function on the Euclidean space.

Definition 2.3.1. Let $f : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be a continuous function, and let Id be the identity map on $\mathbb{R}^d$.

(1) f is called intrinsically differentiable at a point $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, if

$$T_{\mu,2} \ni \phi \mapsto D^L_\phi f(\mu) := \lim_{\epsilon \downarrow 0} \frac{f(\mu \circ (Id + \epsilon\phi)^{-1}) - f(\mu)}{\epsilon} \in \mathbb{R}$$

is a well-defined bounded linear functional. In this case, by the Riesz representation theorem, the unique element $D^L f(\mu) \in T_{\mu,2}$ satisfying

$$\langle D^L f(\mu), \phi \rangle := \int_{\mathbb{R}^d} \langle D^L f(\mu)(x), \phi(x) \rangle \mu(\mathrm{d}x) = D^L_\phi f(\mu), \quad \phi \in T_{\mu,2},$$

is called the intrinsic derivative of f at μ, and we denote

$$\|D^L f(\mu)\|_{T_{\mu,2}} := \|D^L f(\mu)(\cdot)\|_{T_{\mu,2}}, \quad \mu \in \mathscr{P}_2(\mathbb{R}^d).$$

Moreover, if

$$\lim_{\|\phi\|_{T_{\mu,2}} \downarrow 0} \frac{|f(\mu \circ (Id + \phi)^{-1}) - f(\mu) - D^L_\phi f(\mu)|}{\|\phi\|_{T_{\mu,2}}} = 0.$$

f is called L-differentiable at μ with the L-derivative $D^L f(\mu)$.

(2) We write $f \in C^1(\mathscr{P}_2(\mathbb{R}^d))$ if f is L-differentiable at any point $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, and the L-derivative has a version $D^L f(\mu)(x)$ jointly continuous in $(x,\mu) \in \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d)$. If moreover $D^L f(\mu)(x)$ is bounded, we denote $f \in C_b^1(\mathscr{P}_2(\mathbb{R}^d))$.

For a vector-valued function $f = (f_i)$, or a matrix-valued function $f = (f_{ij})$ with L-differentiable components, we write

$$D_\phi^L f(\mu) = (D_\phi^L f_i(\mu)), \text{ or } D_\phi^L f(\mu) = (D_\phi^L f_{ij}(\mu)), \ \mu \in \mathscr{P}_2(\mathbb{R}^d).$$

The following lemma is the formula of L-derivative (for further details of the proof, refer to [6, 87]).

Lemma 2.3.1. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be an atomless probability space, and let $X, Y \in L^2(\Omega \to \mathbb{R}^d, \mathbb{P})$ with $\mathscr{L}_X = \mu$. If either X and Y are bounded, and f is L-differentiable at μ, or $f \in C_b^1(\mathscr{P}_2(\mathbb{R}^d))$, then*

$$\lim_{\epsilon \to 0} \frac{f(\mathscr{L}_{X+\epsilon Y}) - f(\mu)}{\epsilon} = \mathbb{E}\langle D^L f(\mu)(X), Y\rangle. \tag{2.3.1}$$

Consequently,

$$\left| \lim_{\epsilon \downarrow 0} \frac{f(\mathscr{L}_{X+\epsilon Y}) - f(\mu)}{\epsilon} \right| = |\mathbb{E}\langle D^L f(\mu)(X), Y\rangle| \le \|D^L f(\mu)\|_{T_{\mu,2}} \|Y\|_{T_{\mu,2}}. \tag{2.3.2}$$

2.4 Theory of large deviations

This section is devoted to the preliminaries of the LDP, (see [25, 93]).

Throughout this section, $\mathcal{X}$ is a topological space so that open and closed subsets of $\mathcal{X}$ are well-defined. $\mathcal{B}_\mathcal{X}$ denotes the Borel σ-field on $\mathcal{X}$. The LDP characterises the limiting behaviour, as $\varepsilon \to 0$, of a family of probability measures $\{\mu_\varepsilon\}$ on $(\mathcal{X}, \mathcal{B})$ in terms of a rate function. This characterisation is via asymptotic upper and lower exponential bounds on the values that μ_ε assigns to measurable subsets of $\mathcal{X}$.

Definition 2.4.1. A rate function I is a lower semicontinuous mapping $I : \mathcal{X} \to [0, \infty]$ (such that for all $\alpha \in [0, \infty)$, the level set $\Psi_I(\alpha) = \{x : I(x) \leq \alpha\}$ is a closed subset of $\mathcal{X}$). A good rate function is a rate function for which all the level sets $\Psi_I(\alpha)$ are compact subsets of $\mathcal{X}$.

For any set $\Gamma, \bar{\Gamma}$ denotes the closure of Γ, Γ^0 the interior of Γ, and Γ^c the complement of Γ. The infimum of a function over an empty set is interpreted as ∞.

Definition 2.4.2. $\{\mu_\varepsilon\}$ satisfies the LDP with a rate function I if, for all $\Gamma \in \mathcal{B}$,

$$-\inf_{x \in \Gamma^0} I(x) \leq \liminf_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(\Gamma) \leq \limsup_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(\Gamma) \leq -\inf_{x \in \bar{\Gamma}} I(x). \quad (2.4.1)$$

The right- and left-hand sides of (2.4.1) are referred to as the upper and lower bounds, respectively.

When $\mathcal{B}_\mathcal{X} \subset \mathcal{B}$, the LDP is equivalent to the following bounds:

(a) (Upper bound) For any closed set $F \subset \mathcal{X}$,

$$\limsup_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(F) \leq -\inf_{x \in F} I(x).$$

(b) (Lower bound) For any open set $G \subset \mathcal{X}$,

$$\liminf_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(G) \geq -\inf_{x \in G} I(x).$$

Having defined what is meant by an LDP, the rest of this section is devoted to some properties of the LDP.

Definition 2.4.3. Suppose that all the compact subsets of $\mathcal{X}$ belong to $\mathcal{B}$. A family of probability measures $\{\mu_\varepsilon\}$ on $\mathcal{X}$ is exponentiablly tight if for every $\alpha < \infty$, there exists a compact set $K_\alpha \subset \mathcal{X}$ such that

$$\limsup_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(K_\alpha^c) < -\alpha. \tag{2.4.2}$$

The following definition shows the general result of approximate contractions.

Definition 2.4.4. Let $(\mathcal{Y}, d)$ be a metric space. The probability measures $\{\mu_\varepsilon\}$ and $\{\tilde{\mu}_\varepsilon\}$ on $\mathcal{Y}$ are called exponentially equivalent if there exists probability spaces $\{(\Omega, \mathcal{B}_\varepsilon, \mathbb{P}_\varepsilon)\}$ and two families of $\mathcal{Y}$-valued random variables $\{Z_\varepsilon\}$ and $\{\tilde{Z}_\varepsilon\}$ with joint laws $\{P_\varepsilon\}$ and marginals $\{\mu_\varepsilon\}$ and $\{\tilde{\mu}_\varepsilon\}$, respectively, such that the following condition is satisfied:

For each $\delta > 0$, the set $\{\omega : (\tilde{Z}_\varepsilon, Z_\varepsilon) \in \Gamma_\delta\}$ is $\mathcal{B}_\varepsilon$ measurable, and

$$\limsup_{\varepsilon \to 0} \varepsilon \log P_\varepsilon(\Gamma_\delta) = -\infty,$$

where

$$\Gamma_\delta = \{(\tilde{y}, y) : d(\tilde{y}, y) > \delta\} \subset \mathcal{Y} \times \mathcal{Y}.$$

As far as the LDP is concerned, exponentially equivalent measures are indistinguishable, as the following theorem shows.

Theorem 2.4.1. *If an LDP with a good rate function $I(\cdot)$ holds for the probability measures $\{\mu_\varepsilon\}$, which are exponentially equivalent to $\{\tilde{\mu}_\varepsilon\}$, the the same LDP holds for $\{\tilde{\mu}_\varepsilon\}$.*

The following uniform LDP criteria was presented in [53].

Lemma 2.4.2. *For any $\epsilon > 0$, let Γ^ϵ be a measurable mapping from $C([0,T];\mathbb{R}^d)$ into $C([0,T];\mathbb{R}^d)$. Suppose that $\{\Gamma^\epsilon\}_{\epsilon>0}$ satisfies the following assumptions: there exists a measurable map $\Gamma^0 : C([0,T];\mathbb{R}^d) \to C([0,T];\mathbb{R}^d)$ such that*

(a) *For every $N < +\infty$ and any family $\{h_\epsilon; \epsilon > 0\} \subset \mathscr{A}_N$ satisfying that h_ϵ converges in distribution as S_N-valued random variables to h as $\epsilon \to 0$, then*

$$\Gamma^\epsilon\Big(W_\cdot + \frac{1}{\sqrt{\epsilon}}\int_0^\cdot \dot{h}_\epsilon(s)\mathrm{d}s\Big) \Rightarrow \Gamma^0\Big(\int_0^\cdot \dot{h}(s)\mathrm{d}s\Big) \text{ as } \epsilon \to 0.$$

(b) *For every $N < +\infty$, the set $\{\Gamma^0(\int_0^\cdot \dot{h}(s)\mathrm{d}s); h \in S_N\}$ is a compact subset of $C([0,T];\mathbb{R}^d)$.*

Then the family $\{\Gamma^\epsilon\}_{\epsilon>0}$ satisfies an LDP in $C([0,T];\mathbb{R}^d)$ with the rate function I given by

$$I(g) := \inf_{h\in\mathbb{H}; g=\Gamma^0(\int_0^\cdot \dot{h}(s)\mathrm{d}s)} \Big\{\frac{1}{2}\int_0^T |\dot{h}(s)|^2\mathrm{d}s\Big\}, \quad g \in C([0,T];\mathbb{R}^d), \tag{2.4.3}$$

with $\inf\emptyset = \infty$ *by convention.*

We now state the classical exponential inequality for stochastic integral, which is crucial in proving the exponential approximation, (for more details, refer to [89, lemma 4.7] therein).

Lemma 2.4.3. *Let $\alpha : [0,\infty)\times\Omega \to \mathbb{R}^d\times\mathbb{R}^d$ and $\beta : [0,\infty)\times\Omega \to \mathbb{R}^d$ be $(\mathscr{F}_t)_{t\geq 0}$-progressively measurable processes. Assume that $\|\alpha(\cdot)\|_{HS} \leq A$ and $|\beta| \leq B$. Set $\xi(t) := \int_0^t \alpha(s)\mathrm{d}W(s) + \int_0^t \beta(s)\mathrm{d}s$ for $t \geq 0$. Let $T > 0$ and $R > 0$ satisfy $d^{\frac{1}{2}}BT < R$. Then*

$$P\Big(\sup_{0\leq t\leq T}|\xi(t)| \geq R\Big) \leq 2d\exp\Big(\frac{-(R-d^{\frac{1}{2}}BT)^2}{2A^2dT}\Big). \tag{2.4.4}$$

Chapter 3

Weak approximation for stochastic differential equations driven by Brownian motion

This chapter is devoted to investigate the weak approximation rate of EM scheme for SDEs driven by Brownian motion. The drift in this work contains cases of non-Lipschitz continuous and discontinuous functions. Noting that, the method of Girsanov's transformation for the weak convergence rate of numerical scheme does not work for SDEs with multiplicative noise. We also give the reason.

It is worth noting that [5] obtained strong convergence rates for multidimensional SDEs under an integrability condition with the aid of the Krylov estimate and of the heat kernel estimate of the Gaussian type process established by the parametrix method in [48]. Inspired by this work, we aim to investigate the weak convergence of SDEs with low-regular drift. Note that the weak convergence is concerned with the convergence of the distribution of the solutions of SDEs.

In Section 3.1, we give the low-regular assumption about the drift of the model and obtain the weak approximation rate of EM scheme for this model. Section 3.2 is devoted to the proof of the main result. In Section 3.3, we give some illustrative examples explaining the drift could be some types of low-regular function.

3.1 Weak approximation rate of EM scheme

Let $(\mathbb{R}^d, \langle\cdot,\cdot\rangle, |\cdot|)$ be the d-dimensional Euclidean space. $\|\cdot\|$ denotes the operator norm. Consider the following SDE on $\mathbb{R}^d$:

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma \mathrm{d}W_t,\ X_0 = x \in \mathbb{R}^d, \tag{3.1.1}$$

where $(W_t)_{t\geq 0}$ is a d-dimensional Brownian motion with respect to a complete filtration probability space $(\Omega, (\mathscr{F}_t)_{t\geq 0}, \mathscr{F}, \mathbb{P})$. The associated EM scheme reads as follows: for any $\delta \in (0,1)$,

$$\mathrm{d}X_t^{(\delta)} = b(X_{t_\delta}^{(\delta)})\mathrm{d}t + \sigma \mathrm{d}W_t,\ X_0^{(\delta)} = x, \tag{3.1.2}$$

where $t_\delta = [t/\delta]\delta$ and $[t/\delta]$ denotes the integer part of t/δ.

To obtain the main result, throughout this chapter, we assume that the coefficients of (3.1.1) satisfy the following assumptions:

(H1) $b : \mathbb{R}^d \to \mathbb{R}^d$ is measurable and σ is an invertible $d \times d$-matrix. There exist $\beta \in (0,1)$ and nonnegative constants L_1 and L_2 such that

$$|b(x)| \leq L_1 + L_2|x|^\beta.$$

(H2) There exist $p_0 \geq 2$, $\alpha > 0$ and $\phi \in C((0,+\infty);(0,+\infty))$ with $\int_{0^+} \phi^2(s)\mathrm{d}s < \infty$ such that

$$\sup_{z\in\mathbb{R}^d} \int_{\mathbb{R}^d\times\mathbb{R}^d} |b(y) - b(x)|^{p_0} \frac{\mathrm{e}^{-\frac{|x-z|^2}{s} - \frac{|y-x|^2}{r}}}{s^{\frac{d}{2}} r^{\frac{d}{2}}} \mathrm{d}x\mathrm{d}y \leq (\phi(s)r^\alpha)^{p_0},\ s > 0, r \in [0,1].$$

The index α in (H2) is used to characterise the order of the continuity, and the function ϕ is used to characterise the type of the continuity. From examples of Section 3.3, it is clear that functions sharing the same order of continuity can have different types of continuity.

By [103, Theorem 1.1], (3.1.1) has a unique strong solution under (H1). It is clear that (3.1.2) also has a unique strong solution. We denote $\|f\|_\infty = \sup_{x\in\mathbb{R}^d}|f(x)|$. We now formulate the main result.

Theorem 3.1.1. *Assume (H1)-(H2). Then for any $T > 0$ and any bounded measurable function f on $\mathbb{R}^d$, there exists a constant $C_{T,p_0,\sigma,x} > 0$ such that*

$$|\mathbb{E}f(X_t) - \mathbb{E}f(X_t^{(\delta)})| \leq C_{T,p_0,\sigma,x}\|f\|_\infty \delta^\alpha, \quad t \in [0,T], \tag{3.1.3}$$

where p_0 is defined in (H2). If the growth condition in (H1) is replaced by $|b(x)| \leq L_1 + L_2|x|$, then the conclusion (3.1.3) *also holds for time T satisfying*

$$TL_2\|\sigma^{-1}\|\|\sigma\|\frac{\sqrt{2(p_0+1)(p_0+3)}}{p_0-1} < 1. \tag{3.1.4}$$

Remark 3.1.1. When the drift b is non-regular and b is assumed to be bounded, there are many results, (e.g. see [5, 45, 46, 77] and references therein). In particular, we would like to highlight that authors in [77] have obtained the rate of strong convergence for one-dimensional SDEs if b is in $L^1(\mathbb{R})$ and bounded, and satisfies the Sobolev-Slobodeckij-type regularity. This result is better than the present one in Theorem 3.1.1. However, results in [77] relied on an Zvonkin-type transformation which can be given explicitly in one dimension, and some favourable properties are lost in high dimensions. The weak convergence rate can not be derived from the strong rate directly for a bounded and measurable function f. Here, only the Girsanov transformation is used, while we allow that the SDE is multi-dimensional and that the drift satisfies sub-linear growth condition. Our assumption (H2) also includes the Sobolev-Slobodeckij-type regularity (see Example 3.3.3). To obtain higher convergence rate as in [77], it seems that we need to make a deep investigation on the Zvonkin-type transformation.

In the assumption (H2), if α is a decreasing function of p_0, then we can choose $p_0 = 2$ and obtain the highest rate of convergence in (3.1.3) (see Example 3.3.2). Moreover, we obtain the same convergence rate as long as T satisfies (3.1.4) when b has linear growth.

Remark 3.1.2. In [5], the strong convergence and the convergence rate are investigated with the drift satisfying an integrability condition and boundedness. Here we obtain the weak convergence rate of EM scheme, where the drift does not need to be bounded, and the test function f in (3.1.3) is only bounded and measurable, and the convergence rate is better than the rate obtained in [5, Theorem 1.3].

Remark 3.1.3. In [78], authors considered the weak convergence rate of the EM scheme for (3.1.1) with the drift b being of sub-linear growth and $b = b^H + b^A$, where b^H is α-Hölder for some $\alpha \in (0,1)$ and b^A belongs to a class $\mathcal{A}$, which does not contain any nontrivial Hölder continuous functions. The order of the convergence rate obtained in [78] is $\frac{\alpha}{2} \wedge \frac{1}{4}$, even if $b^A \equiv 0$. However, the order of the convergence rate in Theorem 3.1.1 comes from the continuity order α in (H2), and it can be greater than $\frac{1}{4}$.

The class $\mathcal{A}$ in [78, 79] is given by $\mathcal{A}$-approximation. In contrast to the $\mathcal{A}$-approximation, our condition (H2) is more explicit. For instance, the class $\mathcal{A}$ in [79] is a class of all bounded functions $\zeta : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ such that there exists a sequence of functions $(\zeta_n)_{n\in\mathbb{N}} \subset C^1(\mathbb{R}^d)$ satisfying the following conditions:

$\mathcal{A}(i)$ For any $L > 0, \sup_{t\in[0,T]} \int_{|x|\leq L} |\zeta_n(t,x) - \zeta(t,x)| \mathrm{d}x \to 0$ as $n \to \infty$.

$\mathcal{A}(ii)$ There exists a positive constant K such that for any

$$\sup_{t\in[0,T]}\sup_{n\in\mathbb{N}}|\zeta_n(x)|\leq K.$$

$\mathcal{A}(iii)$ There exists a positive constant K such that for any $a\in\mathbb{R}^d$ and $u>0$,

$$\sup_{t\in[0,T]}\sup_{n\in\mathbb{N}}\int_{\mathbb{R}^d}\|\nabla\zeta_n(x+a)\|\frac{\mathrm{e}^{-|x|^2/u}}{u^{(d-1)/2}}\mathrm{d}x<K(1+\sqrt{u}).$$

Moreover, for any time independent function ζ in the class $\mathcal{A}$ of [79], ζ satisfies (H2) with $p_0=2,\alpha=\frac{1}{4}$ and $\phi(s)=s^{-\frac{1}{4}}\sqrt{1+\sqrt{s}}$. Indeed, according to definition of $\mathcal{A}$, the boundedness of ζ, and there exists a sequence $\{\zeta_n\}_{n\geq1}$ such that $\zeta_n\in C^1(\mathbb{R}^d)$ is uniformly bounded and converges to ζ locally in $L^1(\mathbb{R}^d)$, and there exists $K>0$ such that

$$\sup_{n\geq1,\ a\in\mathbb{R}^d}\int_{\mathbb{R}^d}\|\nabla\zeta_n(x+a)\|\frac{e^{-\frac{|x|^2}{s}}}{s^{(d-1)/2}}\mathrm{d}x\leq K(1+\sqrt{s}),\tag{3.1.5}$$

noting the fact that

$$\sup_{x\geq0}(x^{\gamma'}\,\mathrm{e}^{-\gamma x^2})=\left(\frac{\gamma'}{2\,\mathrm{e}\,\gamma}\right)^{\gamma'/2},\qquad\gamma',\gamma>0,\tag{3.1.6}$$

we then obtain from (3.1.5) and (3.1.6) that

$$\begin{aligned}
&\int_{\mathbb{R}^d\times\mathbb{R}^d}|\zeta(x)-\zeta(y)|^2\frac{e^{-\frac{|x-z|^2}{s}-\frac{|x-y|^2}{r}}}{(sr)^{\frac{d}{2}}}\mathrm{d}x\mathrm{d}y\\
&\quad\leq\|\zeta\|_\infty\lim_{n\to+\infty}\int_{\mathbb{R}^d\times\mathbb{R}^d}|\zeta_n(x)-\zeta_n(y)|\frac{e^{-\frac{|x-z|^2}{s}-\frac{|x-y|^2}{r}}}{(sr)^{\frac{d}{2}}}\mathrm{d}x\mathrm{d}y\\
&\quad=\|\zeta\|_\infty\lim_{n\to+\infty}\int_0^1\int_{\mathbb{R}^d\times\mathbb{R}^d}\|\nabla_{y-x}\zeta_n(x+\theta(y-x))\|\frac{e^{-\frac{|x-z|^2}{s}-\frac{|x-y|^2}{r}}}{(sr)^{\frac{d}{2}}}\mathrm{d}x\mathrm{d}y\mathrm{d}\theta\\
&\quad\leq\|\zeta\|_\infty\lim_{n\to+\infty}\int_0^1\left(\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}\|\nabla\zeta_n(x+\theta h+z)\|\frac{e^{-\frac{|x|^2}{s}}}{s^{d/2}}\mathrm{d}x\right)\frac{|h|e^{\frac{-|h|^2}{r}}}{r^{\frac{d}{2}}}\mathrm{d}h\right)\mathrm{d}\theta\\
&\quad\leq\|\zeta\|_\infty\int_{\mathbb{R}^d}Ks^{-\frac{1}{2}}(1+\sqrt{s})\frac{|h|e^{\frac{-|h|^2}{r}}}{r^{\frac{d}{2}}}\mathrm{d}h
\end{aligned}$$

$$\leq C\|\zeta\|_\infty s^{-\frac{1}{2}}(1+\sqrt{s})r^{\frac{1}{2}},$$

where the constant C is independent of z. The class $\mathcal{A}$ used in [78] allows functions in $\mathcal{A}$ to be just exponentially bounded (i.e. $|\zeta(x)| \leq K\,\mathrm{e}^{Kx}$, for all $x \in \mathbb{R}^d$). However, they assume that the drift is only sub-linear growth. There is no example showing that the class $\mathcal{A}$ used in [78] can contain functions which are more irregular than functions in $\mathcal{A}$ of [79].

3.2 Proof of the main theorem

The key point for proving the Theorem 3.1.1 is to construct a reference SDE, which can provide new representations of (3.1.1) and its EM approximation SDE (3.1.2) under other probability measures which will be defined in view of the Girsanov theorem.

We denote by $Y_t = x + \sigma W_t$ the reference SDE of (3.1.1). One can see that Y_t is a time homogenous Markov process with heat kernel with respect to the Lebesgue measure as follows:

$$p_t(x,y) = \frac{\exp\left\{-\frac{\langle(\sigma\sigma^*)^{-1}(y-x),(y-x)\rangle}{2t}\right\}}{\sqrt{(2t\pi)^d\det(\sigma\sigma^*)}}, \quad x,y \in \mathbb{R}^d. \tag{3.2.1}$$

To complete the proof of Theorem 3.1.1, we give the following auxiliary lemmas.

The first lemma is on the exponential estimate of $|b(Y_t)|$. More precisely, we give a more general result about the exponential estimate of $|b(Y_t)|$ by using a weaker condition (H1') below in lieu of assumption (H1).

(H1') there exist $\beta \in (0,1)$, nonnegative constants L_1, L_2 and function $F \geq 0$ with $F \in L^{p_1}(\mathbb{R}^d)$ for some $p_1 > d$ such that

$$|b(x)| \leq L_1 + L_2|x|^\beta + F(x). \tag{3.2.2}$$

Lemma 3.2.1. *Assume (H1'). Then, for any $\lambda > 0$, it holds that*

$$\mathbb{E}\exp\left\{\lambda\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\right\} < \infty. \tag{3.2.3}$$

Proof. Noting that for any $\varepsilon > 0$, it holds that

$$L_1 + L_2|x|^\beta \leq L_1 + (1-\beta)L_2^{\frac{1}{1-\beta}}\left(\frac{\beta}{\varepsilon}\right)^{\frac{\beta}{1-\beta}} + \varepsilon|x| =: L(\varepsilon) + \varepsilon|x|, \tag{3.2.4}$$

and the elementary inequality

$$(a+b+c)^2 \leq (2+\frac{1}{\varepsilon_1})a^2 + (1+\varepsilon_1+\varepsilon_2)b^2 + (2+\frac{1}{\varepsilon_2})c^2,$$

where $a, b, c, \varepsilon_1, \varepsilon_2 > 0$.

Combing this with (3.2.4) and the Hölder inequality, we derive from (3.2.2) that

$$\begin{aligned}
&\mathbb{E}\exp\left\{\lambda\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\right\}\\
&\leq \mathbb{E}\exp\left\{\lambda\int_0^T\|\sigma^{-1}\|^2\left(L(\varepsilon)+\varepsilon|Y_s|+F(Y_s)\right)^2\mathrm{d}s\right\}\\
&\leq \mathbb{E}\exp\left\{\lambda\int_0^T\|\sigma^{-1}\|^2\left((L(\varepsilon)+\varepsilon|x|)+\varepsilon|Y_s-x|+F(Y_s)\right)^2\mathrm{d}s\right\}\\
&\leq \mathbb{E}\exp\left\{\lambda\int_0^T\|\sigma^{-1}\|^2\Big(\left(2+\varepsilon_1^{-1}\right)\left(L(\varepsilon)+\varepsilon|x|\right)^2\right.\\
&\qquad\qquad \left. + (1+\varepsilon_1+\varepsilon_2)\varepsilon^2|Y_s-x|^2 + (2+\varepsilon_2^{-1})F^2(Y_s)\Big)\mathrm{d}s\right\}\\
&\leq \exp\{\lambda T\|\sigma^{-1}\|^2(L(\varepsilon)+\varepsilon|x|)^2\left(2+\varepsilon_1^{-1}\right)\}\\
&\quad\times\left(\mathbb{E}\exp\left\{\lambda(1+\varepsilon_1+\varepsilon_2)^2\varepsilon^2\|\sigma^{-1}\|^2\int_0^T|Y_s-x|^2\mathrm{d}s\right\}\right)^{\frac{1}{1+\varepsilon_1+\varepsilon_2}}\\
&\quad\times\left(\mathbb{E}\exp\left\{\frac{\lambda(2+\varepsilon_2^{-1})(1+\varepsilon_1+\varepsilon_2)}{\varepsilon_1+\varepsilon_2}\|\sigma^{-1}\|^2\int_0^T F^2(Y_s)\mathrm{d}s\right\}\right)^{\frac{\varepsilon_1+\varepsilon_2}{1+\varepsilon_1+\varepsilon_2}}\\
&= \exp\{\lambda T\|\sigma^{-1}\|^2(L(\varepsilon)+\varepsilon|x|)^2\left(2+\varepsilon_1^{-1}\right)\}I_{1,T}^{\frac{1}{1+\varepsilon_1+\varepsilon_2}}I_{2,T}^{\frac{\varepsilon_1+\varepsilon_2}{1+\varepsilon_1+\varepsilon_2}}.
\end{aligned} \tag{3.2.5}$$

Noting that $F \in L^{p_1}(\mathbb{R}^d)$, for any $0 \leq S \leq T$ and q satisfying $\frac{d}{p_1} + \frac{1}{q} < 1$, we obtain that (e.g. see [47])

$$\mathbb{E}\left[\int_S^T F^2(Y_s)\mathrm{d}s\Big|\mathscr{F}_S\right] \leq (T-S)^{\frac{1}{q}}\|F\|_{L^{p_1}}, \tag{3.2.6}$$

which yields the following Khasminskii's estimate (e.g. see [100, Lemma 3.5]): for any $C > 0$,

$$\mathbb{E}\exp\left\{C\int_0^T F^2(Y_s)\mathrm{d}s\right\} < \infty. \tag{3.2.7}$$

This implies that for any $\varepsilon_2 > 0$,

$$I_{2,T} < \infty. \tag{3.2.8}$$

For $I_{1,T}$. Noting the arbitrariness of ε, ε_1 and ε_2, we can choose them sufficiently small such that for any $T > 0$,

$$1 - 2T^2(1+\varepsilon_1+\varepsilon_2)^2\lambda\varepsilon^2\|\sigma^{-1}\|^2\|\sigma\|^2 =: \hat{\lambda} > 0.$$

This, together with the Jensen inequality and the heat kernel (3.2.1), yields that

$$\begin{aligned}
I_{1,T} &= \mathbb{E}\exp\left\{\lambda(1+\varepsilon_1+\varepsilon_2)^2\varepsilon^2\|\sigma^{-1}\|^2\int_0^T |Y_s - x|^2\mathrm{d}s\right\}\\
&\le \frac{1}{T}\int_0^T \mathbb{E}\exp\left\{T\lambda(1+\varepsilon_1+\varepsilon_2)^2\varepsilon^2\|\sigma^{-1}\|^2|Y_s-x|^2\right\}\mathrm{d}s\\
&= \int_0^T\int_{\mathbb{R}^d}\frac{\exp\left\{T\lambda(1+\varepsilon_1+\varepsilon_2)^2\varepsilon^2\|\sigma^{-1}\|^2|y|^2 - \frac{|\sigma^{-1}y|^2}{2s}\right\}}{T\sqrt{(2s\pi)^d\det(\sigma\sigma^*)}}\mathrm{d}y\mathrm{d}s\\
&\le \int_0^T\int_{\mathbb{R}^d}\frac{\exp\left\{-(\frac{1-2sT\lambda(1+\varepsilon_1+\varepsilon_2)^2\varepsilon^2\|\sigma^{-1}\|^2\|\sigma\|^2}{2s})|\sigma^{-1}y|^2\right\}}{T\sqrt{(2s\pi)^d\det(\sigma\sigma^*)}}\mathrm{d}y\mathrm{d}s\\
&\le \int_0^T\int_{\mathbb{R}^d}\frac{\exp\left\{-(\frac{\hat{\lambda}}{2s})|\sigma^{-1}y|^2\right\}}{T\sqrt{(2s\pi)^d\det(\sigma\sigma^*)}}\mathrm{d}y\mathrm{d}s\\
&= \hat{\lambda}^{-\frac{d}{2}} < \infty.
\end{aligned} \tag{3.2.9}$$

(3.2.3) follows by plugging (3.2.9) and (3.2.8) into (3.2.5). □

The following lemma deals with the exponential estimate of $|b(Y_{t_\delta})|$, where $\{Y_{t_\delta}\}_{t\in[0,T]}$ denotes solution to the discrete-time EM scheme.

Lemma 3.2.2. *Assume (H1). Then, for any $T > 0$ and $\lambda > 0$, it holds that*

$$\sup_{0<\delta<1\wedge T} \mathbb{E}\exp\left\{\lambda\int_0^T |\sigma^{-1}b(Y_{s_\delta})|^2\mathrm{d}s\right\} < \infty. \tag{3.2.10}$$

Proof. Splitting the time interval and applying (3.2.4), the elementary inequality, it yields that

$$\begin{aligned}
&\mathbb{E}\exp\left\{\lambda\int_0^T |\sigma^{-1}b(Y_{s_\delta})|^2\mathrm{d}s\right\}\\
&= \mathbb{E}\exp\left\{\lambda\int_0^\delta |\sigma^{-1}b(Y_{s_\delta})|^2\mathrm{d}s\right\} + \mathbb{E}\exp\left\{\lambda\int_\delta^T |\sigma^{-1}b(Y_{s_\delta})|^2\mathrm{d}s\right\}\\
&\le \exp\left\{\lambda\delta\|\sigma^{-1}\|^2(L(\varepsilon)+\varepsilon x)^2\right\}\\
&\quad + \mathbb{E}\exp\left\{\lambda\int_\delta^T \|\sigma^{-1}\|^2|L(\varepsilon)+\varepsilon x+\varepsilon(Y_{s_\delta}-x)|^2\mathrm{d}s\right\}\\
&\le \exp\{\lambda\delta\|\sigma^{-1}\|^2(L(\varepsilon)+\varepsilon x)^2\} + \exp\{\lambda(T-\delta)\|\sigma^{-1}\|^2(L(\varepsilon)+\varepsilon|x|)^2\left(1+\varepsilon_1^{-1}\right)\}\\
&\qquad\qquad \times \mathbb{E}\exp\left\{\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2\int_\delta^T |Y_{s_\delta}-x|^2\mathrm{d}s\right\}.
\end{aligned} \tag{3.2.11}$$

Noting the arbitrariness of $\varepsilon, \varepsilon_1$, we can choose them sufficiently small such that for any $T > 0$,

$$1 - 2T^2\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2\|\sigma\|^2 =: \breve{\lambda} > 0.$$

This, together with the Jensen inequality and (3.2.1), we obtain

$$\begin{aligned}
&\mathbb{E}\exp\left\{\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2\int_\delta^T |Y_{s_\delta}-x|^2\mathrm{d}s\right\}\\
&\le \frac{1}{T-\delta}\int_\delta^T \mathbb{E}\exp\{(T-\delta)\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2|Y_{s_\delta}-x|^2\}\mathrm{d}s\\
&\le \int_\delta^T\int_{\mathbb{R}^d} \frac{\exp\{(T-\delta)\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2|y|^2 - \frac{\langle(\sigma\sigma^*)^{-1}y,y\rangle}{2s_\delta}\}}{(T-\delta)\sqrt{(2\pi s_\delta)^d\det(\sigma\sigma^*)}}\mathrm{d}y\mathrm{d}s\\
&\le \int_\delta^T\int_{\mathbb{R}^d} \frac{\exp\{(T-\delta)\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2\|\sigma\|^2|\sigma^{-1}y|^2 - \frac{|\sigma^{-1}y|^2}{2s_\delta}\}}{(T-\delta)\sqrt{(2\pi s_\delta)^d\det(\sigma\sigma^*)}}\mathrm{d}y\mathrm{d}s
\end{aligned}$$

$$\leq \int_\delta^T \int_{\mathbb{R}^d} \frac{\exp\{-\frac{(1-2(T-\delta)^2\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2\|\sigma\|^2)}{2s_\delta}|\sigma^{-1}y|^2\}}{(T-\delta)\sqrt{(2\pi s_\delta)^d \det(\sigma\sigma^*)}} \mathrm{d}y\mathrm{d}s$$
$$\leq \int_\delta^T \int_{\mathbb{R}^d} \frac{\exp\{-\frac{(1-2T^2\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2\|\sigma\|^2)}{2s_\delta}|\sigma^{-1}y|^2\}}{(T-\delta)\sqrt{(2\pi s_\delta)^d \det(\sigma\sigma^*)}} \mathrm{d}y\mathrm{d}s$$
$$= \breve{\lambda}^{-\frac{d}{2}} < \infty. \tag{3.2.12}$$

Combining this with (3.2.11), it implies that (3.2.10) holds. □

Remark 3.2.1. According to the proofs of Lemma 3.2.1 and Lemma 3.2.2 (see (3.2.9), (3.2.12), and the definitions of $\hat{\lambda}$ and $\breve{\lambda}$), we have that $\varepsilon = O(T^{-1})$ as $T \to +\infty$. From (3.2.4), the constant $TL^2(\varepsilon)$ in (3.2.5) and (3.2.11) is of the order $(1-\beta)^2(L_2^{\frac{2}{1+\beta}}T)^{\frac{1+\beta}{1-\beta}}$. Hence, the larger $L_2^{\frac{2}{1+\beta}}T$, the closer β is to 1, the greater the upper bound of (3.2.3) and (3.2.10).

Lemma 3.2.1 and Lemma 3.2.2 serve to use the Novikov condition in the proof of Theorem 3.1.1. For the case of $\beta < 1$, we have that λ in both lemmas is arbitrary. For the case of $\beta = 1$, with $\varepsilon = L_2$ and $L(\varepsilon) = L_1$ in (3.2.4), one can see from $\hat{\lambda}$ and $\breve{\lambda}$ that for any $\lambda > 0$ and $T > 0$ satisfying the following condition

$$2T^2\lambda L_2^2\|\sigma^{-1}\|^2\|\sigma\|^2 < 1. \tag{3.2.13}$$

Since we can choose ε_1 and ε_2 to be sufficiently small, it yields that (3.2.9) and (3.2.12) hold.

Remark 3.2.2. The Krylov estimate (3.2.6) fails for Y_{s_δ} (see [5, Remark 2.5] or [89]). Hence, we use (H1) in Lemma 3.2.2 instead of (H1').

Lemma 3.2.3. *Assume (H2). Then there exists a constant C_σ such that for all $0 < s \leq t \leq T$ we have*

$$\mathbb{E}|b(Y_t) - b(Y_s)|^{p_0} \leq C_\sigma(\phi(2s\|\sigma\|^2)(2(t-s)\|\sigma\|^2)^\alpha)^{p_0}, \tag{3.2.14}$$

where σ is the constant matrix of the reference SDE, p_0, ϕ and α are defined as in assumption (H2).

Proof. By the definition of reference SDE, it is easy to see that

$$\mathbb{E}|b(Y_t) - b(Y_s)|^{p_0} = \mathbb{E}|b(x + \sigma W_t) - b(x + \sigma W_s)|^{p_0}.$$

Noting that $W_t - W_s$ and W_s are mutually independent, we obtain from (3.2.1) and (H2) that

$$\begin{aligned}
&\mathbb{E}|b(x + \sigma W_t) - b(x + \sigma W_s)|^{p_0} \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |b(x+y) - b(x+z)|^{p_0} p_{t-s}(x+z, x+y) p_s(x, x+z) \mathrm{d}y\mathrm{d}z \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |b(x+y) - b(x+z)|^{p_0} \frac{\mathrm{e}^{-\frac{\langle(\sigma\sigma^*)^{-1}(y-z),(y-z)\rangle}{2(t-s)}}}{\sqrt{(2\pi(t-s))^d \det(\sigma\sigma^*)}} \frac{\mathrm{e}^{-\frac{\langle(\sigma\sigma^*)^{-1}z,z\rangle}{2s}}}{\sqrt{(2\pi s)^d \det(\sigma\sigma^*)}} \mathrm{d}y\mathrm{d}z \\
&\le \frac{\|\sigma\|^{2d}}{\pi^d \det(\sigma\sigma^*)} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |b(x+y) - b(x+z)|^{p_0} \frac{\mathrm{e}^{-\frac{|y-z|^2}{2\|\sigma\|^2(t-s)}}\, \mathrm{e}^{-\frac{|z|^2}{2\|\sigma\|^2 s}}}{(2(t-s)\|\sigma\|^2)^{d/2}(2s\|\sigma\|^2)^{d/2}} \mathrm{d}y\mathrm{d}z \\
&= \frac{\|\sigma\|^{2d}}{\pi^d \det(\sigma\sigma^*)} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |b(u) - b(v)|^{p_0} \frac{\mathrm{e}^{-\frac{|u-v|^2}{2\|\sigma\|^2(t-s)}}\, \mathrm{e}^{-\frac{|v-x|^2}{2\|\sigma\|^2 s}}}{(2(t-s)\|\sigma\|^2)^{d/2}(2s\|\sigma\|^2)^{d/2}} \mathrm{d}u\mathrm{d}v \\
&\le \sup_{x\in\mathbb{R}^d} \frac{\|\sigma\|^{2d}}{\pi^d \det(\sigma\sigma^*)} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |b(u) - b(v)|^{p_0} \frac{\mathrm{e}^{-\frac{|u-v|^2}{2\|\sigma\|^2(t-s)}}\, \mathrm{e}^{-\frac{|v-x|^2}{2\|\sigma\|^2 s}}}{(2(t-s)\|\sigma\|^2)^{d/2}(2s\|\sigma\|^2)^{d/2}} \mathrm{d}u\mathrm{d}v \\
&\le \frac{\|\sigma\|^{2d}}{\pi^d \det(\sigma\sigma^*)} (\phi(2s\|\sigma\|^2)(2(t-s)\|\sigma\|^2)^{\alpha})^{p_0},
\end{aligned}$$

which implies that (3.2.14) holds by taking $C_\sigma = \frac{\|\sigma\|^{2d}}{\pi^d \det(\sigma\sigma^*)}$. □

Now, we are in position to finish the Proof of Theorem 3.1.1.

Proof of Theorem 3.1.1. Let

$$\begin{aligned}
\hat{W}_t &= W_t - \int_0^t \sigma^{-1} b(Y_s)\mathrm{d}s, \quad \tilde{W}_t = W_t - \int_0^t \sigma^{-1} b(Y_{s_\delta})\mathrm{d}s, \\
R_{1,T} &= \exp\Big\{ \int_0^T \langle \sigma^{-1} b(Y_s), \mathrm{d}W_s\rangle - \frac{1}{2}\int_0^T |\sigma^{-1} b(Y_s)|^2 \mathrm{d}s \Big\},
\end{aligned}$$

$$R_{2,T} = \exp\Big\{\int_0^T \langle \sigma^{-1} b(Y_{s_\delta}), \mathrm{d}W_s\rangle - \frac{1}{2}\int_0^T \big|\sigma^{-1} b(Y_{s_\delta})\big|^2 \mathrm{d}s\Big\}.$$

The proof is divided into two steps:

Step (i), we shall prove that the assertion holds under (H1) and (H2).

We first show that $\{\hat{W}_t\}_{t\in[0,T]}$ is a Brownian motion under $\mathbb{Q}_1 := R_{1,T}\mathbb{P}$, and $\{\tilde{W}_t\}_{t\in[0,T]}$ is a Brownian motion under $\mathbb{Q}_2 := R_{2,T}\mathbb{P}$. In view of Lemma 3.2.1, the Girsanov theorem implies that $\{R_{1,t}\}_{t\in[0,T]}$ is a martingale and $\{\hat{W}_t\}_{t\in[0,T]}$ is a Brownian motion under $\mathbb{Q}_1$. Similarly, it follows from Lemma 3.2.2 and Novikov's condition that $\{\tilde{W}_t\}_{t\in[0,T]}$ is a Brownian motion under $\mathbb{Q}_2$.

Then, we can reformulate the reference SDE $Y_t = x + \sigma W_t$ as follows:

$$Y_t = x + \int_0^t b(Y_s)\mathrm{d}s + \sigma\hat{W}_t,$$

which means that $(Y_t, \hat{W}_t)$ under $\mathbb{Q}_1$ is a weak solution of (3.1.1). Hence, Y_t under $\mathbb{Q}_1$ has the same law of X_t under $\mathbb{P}$ due to the pathwise uniqueness of the solutions to (3.1.1). Similarly, reformulating $Y_t = x + \sigma W_t$ as follows:

$$Y_t = x + \int_0^t b(Y_{s_\delta})\mathrm{d}s + \sigma\tilde{W}_t, \tag{3.2.15}$$

$(Y_t, \tilde{W}_t)$ under $\mathbb{Q}_2$ is also a weak solution of (3.1.2), which has a pathwise unique solution. Hence Y_t under $\mathbb{Q}_2$ has the same law of $X_t^{(\delta)}$ under $\mathbb{P}$.

From these equivalent relations, we obtain that for any bounded measurable function f on $\mathbb{R}^d$,

$$\begin{aligned}&|\mathbb{E}f(X_t) - \mathbb{E}f(X_t^{(\delta)})| = |\mathbb{E}_{\mathbb{Q}_1} f(Y_t) - \mathbb{E}_{\mathbb{Q}_2} f(Y_t)|\\ &= \mathbb{E}|(R_{1,T} - R_{2,T})f(Y_t)| \le \|f\|_\infty \mathbb{E}|R_{1,T} - R_{2,T}|.\end{aligned}$$

Using the inequality $|\,\mathrm{e}^x - \mathrm{e}^y\,| \le (\mathrm{e}^x \vee \mathrm{e}^y)|x-y|$, Hölder's inequality and Minkowski's inequality, we obtain from the definitions of $R_{1,T}$ and $R_{2,T}$ that

$\mathbb{E}|R_{1,T} - R_{2,T}|$

$$
\begin{aligned}
&\leq \mathbb{E}\Big\{(R_{1,T}\vee R_{2,T})\Big|\int_0^T\langle\sigma^{-1}(b(Y_s)-b(Y_{s_\delta})),\mathrm{d}W_s\rangle \\
&\qquad\quad +\frac{1}{2}\int_0^T\Big(|\sigma^{-1}b(Y_{s_\delta})|^2-|\sigma^{-1}b(Y_s)|^2\Big)\mathrm{d}s\Big|\Big\} \\
&\leq \mathbb{E}\Big[(R_{1,T}\vee R_{2,T})\Big|\int_0^T\langle\sigma^{-1}(b(Y_s)-b(Y_{s_\delta})),\mathrm{d}W_s\rangle\Big|\Big] \\
&\quad +\frac{1}{2}\mathbb{E}\Big[(R_{1,T}\vee R_{2,T})\Big|\int_0^T\Big(|\sigma^{-1}b(Y_{s_\delta})|^2-|\sigma^{-1}b(Y_s)|^2\Big)\mathrm{d}s\Big|\Big] \\
&\leq \Big(\mathbb{E}(R_{1,T}\vee R_{2,T})^{\frac{p_0}{p_0-1}}\Big)^{\frac{p_0-1}{p_0}}\Big(\mathbb{E}\Big|\int_0^T\langle\sigma^{-1}(b(Y_s)-b(Y_{s_\delta})),\mathrm{d}W_s\rangle\Big|^{p_0}\Big)^{\frac{1}{p_0}} \\
&\quad +\frac{1}{2}\Big(\mathbb{E}(R_{1,T}\vee R_{2,T})^{\frac{p_0+1}{p_0-1}}\Big)^{\frac{p_0-1}{p_0+1}}\Big(\mathbb{E}\Big|\int_0^T\big(|\sigma^{-1}b(Y_{s_\delta})|^2-|\sigma^{-1}b(Y_s)|^2\big)\mathrm{d}s\Big|^{\frac{p_0+1}{2}}\Big)^{\frac{2}{p_0+1}} \\
&\leq \Big(\mathbb{E}(R_{1,T}\vee R_{2,T})^{\frac{p_0}{p_0-1}}\Big)^{\frac{p_0-1}{p_0}}\Big(\mathbb{E}\Big|\int_0^T\langle\sigma^{-1}(b(Y_s)-b(Y_{s_\delta})),\mathrm{d}W_s\rangle\Big|^{p_0}\Big)^{\frac{1}{p_0}} \\
&\quad +\frac{1}{2}\Big(\mathbb{E}(R_{1,T}\vee R_{2,T})^{\frac{p_0+1}{p_0-1}}\Big)^{\frac{p_0-1}{p_0+1}}\int_0^T\Big(\mathbb{E}\big||\sigma^{-1}b(Y_{s_\delta})|^2-|\sigma^{-1}b(Y_s)|^2\big|^{\frac{p_0+1}{2}}\Big)^{\frac{2}{p_0+1}}\mathrm{d}s \\
&=: \Big(\mathbb{E}(R_{1,T}\vee R_{2,T})^{\frac{p_0}{p_0-1}}\Big)^{\frac{p_0-1}{p_0}}G_{1,T}+\frac{1}{2}\Big(\mathbb{E}(R_{1,T}\vee R_{2,T})^{\frac{p_0+1}{p_0-1}}\Big)^{\frac{p_0-1}{p_0+1}}G_{2,T}.
\end{aligned}
\tag{3.2.16}
$$

Let

$$
M_{1,t}=\int_0^t\langle\sigma^{-1}b(Y_s),\mathrm{d}W_s\rangle \quad\text{and}\quad M_{2,t}=\int_0^t\langle\sigma^{-1}b(Y_{s_\delta}),\mathrm{d}W_s\rangle.
$$

For any $q>1$, using Hölder's inequality and the fact that $\hat{M}_{i,t}:=\mathrm{e}^{2qM_{i,t}-2q^2\langle M_{i,\cdot}\rangle_t}$, $i=1,2$, is an exponential martingale, we arrive at

$$
\begin{aligned}
\mathbb{E}R_{1,T}^{\frac{p_0}{p_0-1}} &= \mathbb{E}\exp\Big\{\frac{p_0}{p_0-1}\int_0^T\langle\sigma^{-1}b(Y_s),\mathrm{d}W_s\rangle \\
&\qquad\qquad -\frac{p_0}{2(p_0-1)}\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\Big\} \\
&\leq \Big(\mathbb{E}\hat{M}_{1,T}\Big)^{1/2}\Big(\mathbb{E}\exp\Big\{\frac{p_0(p_0+1)}{(p_0-1)^2}\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\Big\}\Big)^{1/2} \\
&= \Big(\mathbb{E}\exp\Big\{\frac{p_0(p_0+1)}{(p_0-1)^2}\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\Big\}\Big)^{1/2}.
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
\mathbb{E}R_{2,T}^{\frac{p_0+1}{p_0-1}} &\le \left(\mathbb{E}\hat{M}_{2,T}\right)^{1/2}\left(\mathbb{E}\exp\left\{\frac{(p_0+3)(p_0+1)}{(p_0-1)^2}\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\right\}\right)^{1/2}\\
&=\left(\mathbb{E}\exp\left\{\frac{(p_0+3)(p_0+1)}{(p_0-1)^2}\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\right\}\right)^{1/2}.
\end{aligned}
$$

In view of Lemma 3.2.1, we then have that

$$
\mathbb{E}\left(R_{1,T}^{\frac{p_0+1}{p_0-1}}+R_{1,T}^{\frac{p_0}{p_0-1}}\right)<\infty. \tag{3.2.17}
$$

Similarly, we can prove by Lemma 3.2.2 that

$$
\mathbb{E}\left(R_{2,T}^{\frac{p_0+1}{p_0-1}}+R_{2,T}^{\frac{p_0}{p_0-1}}\right)<\infty. \tag{3.2.18}
$$

Noting that $\int_{0+}\phi^2(s)\mathrm{d}s<\infty$. Using Riemann sums we obtain that

$$
\begin{aligned}
\lim_{\delta\to0+}\sum_{k=1}^{[T/\delta]}\phi^2(2k\delta\|\sigma\|^2)\delta &= \int_0^T\phi^2(2\|\sigma\|^2r)\mathrm{d}r\\
&=\frac{1}{2\|\sigma\|^2}\int_0^{2\|\sigma\|^2T}\phi^2(s)\mathrm{d}s<\infty.
\end{aligned} \tag{3.2.19}
$$

This, together with the Burkholder-Davis-Gundy (abbreviated as BDG) inequality and Lemma 4.2.15, yields that for $p_0\ge2$,

$$
\begin{aligned}
G_{1,T} &= \left(\mathbb{E}\Big|\int_0^T\langle\sigma^{-1}(b(Y_s)-b(Y_{s_\delta})),\mathrm{d}W_s\rangle\Big|^{p_0}\right)^{1/p_0}\\
&\le\left(\frac{p_0}{p_0-1}\right)^{\frac{p_0}{2}}\left(\frac{p_0(p_0-1)}{2}\right)^{\frac12}\|\sigma^{-1}\|\left(\int_0^T\left(\mathbb{E}|b(Y_s)-b(Y_{s_\delta})|^{p_0}\right)^{\frac{2}{p_0}}\mathrm{d}s\right)^{\frac12}\\
&\le\delta^\alpha\frac{2^{\frac{\alpha-1}{2}}\|\sigma\|^{\frac{2d}{p_0}+\alpha-1}\|\sigma^{-1}\|}{(\pi^d\det(\sigma\sigma^*))^{\frac{1}{p_0}}}\left(\frac{p_0}{p_0-1}\right)^{\frac{p_0}{2}}\left(\frac{p_0(p_0-1)}{2}\right)^{\frac12}\int_0^{2\|\sigma\|^2T}\phi^2(s)\mathrm{d}s\\
&=C_{T,p_0,\sigma,\alpha,\phi}\delta^\alpha.
\end{aligned} \tag{3.2.20}
$$

Noting that for any $p\ge1$

$$
\mathbb{E}|Y_t|^p\le2^{p-1}\left(|x|^p+(\sqrt{t}\|\sigma\|)^p\mathbb{E}|W_1|^p\right), \tag{3.2.21}
$$

we derive from (3.2.4) and (3.2.19) that

$$
\begin{aligned}
&\left(\mathbb{E}|b(Y_s)+b(Y_{s_\delta})|^{\frac{p_0(p_0+1)}{p_0-1}}\right)^{\frac{p_0-1}{p_0(p_0+1)}}\\
&\le \left(\mathbb{E}\left(2L(\varepsilon)+\varepsilon(|Y_s|+|Y_{s_\delta}|)\right)^{\frac{p_0(p_0+1)}{p_0-1}}\right)^{\frac{p_0-1}{p_0(p_0+1)}}\\
&\le 6\left\{L(\varepsilon)+2\varepsilon\left(|x|+\sqrt{T}\|\sigma\|\left(\mathbb{E}|W_1|^{\frac{p_0(p_0+1)}{p_0-1}}\right)^{\frac{p_0-1}{p_0(p_0+1)}}\right)\right\}\\
&=: C_{T,p_0,\sigma,L(\varepsilon),\varepsilon,x}.
\end{aligned}
$$

Combining this with Lemma 4.2.15, (3.2.19) and Hölder's inequality, we obtain

$$
\begin{aligned}
G_{2,T} &= \frac{1}{2}\int_0^T \left(\mathbb{E}\left||\sigma^{-1}b(Y_{s_\delta})|^2-|\sigma^{-1}b(Y_s)|^2\right|^{\frac{p_0+1}{2}}\right)^{\frac{2}{p_0+1}} \mathrm{d}s\\
&\le \frac{\|\sigma^{-1}\|^2}{2}\int_0^T \left(\mathbb{E}|b(Y_s)-b(Y_{s_\delta})|^{\frac{p_0+1}{2}}|b(Y_s)+b(Y_{s_\delta})|^{\frac{p_0+1}{2}}\right)^{\frac{2}{p_0+1}}\mathrm{d}s\\
&\le \frac{\|\sigma^{-1}\|^2}{2}\int_0^T (\mathbb{E}|b(Y_s)-b(Y_{s_\delta})|^{p_0})^{\frac{1}{p_0}}\left(\mathbb{E}|b(Y_s)+b(Y_{s_\delta})|^{\frac{p_0(p_0+1)}{p_0-1}}\right)^{\frac{p_0-1}{p_0(p_0+1)}}\mathrm{d}s\\
&\le \frac{\|\sigma^{-1}\|^2}{2}C_{T,p_0,\sigma,L_1,L_2,x}\int_0^T (\mathbb{E}|b(Y_s)-b(Y_{s_\delta})|^{p_0})^{\frac{1}{p_0}}\,\mathrm{d}s\\
&\le C_{T,p_0,\sigma,L(\varepsilon),\varepsilon,\phi,x}\delta^\alpha,
\end{aligned}
\tag{3.2.22}
$$

where

$$
C_{T,p_0,\sigma,L(\varepsilon),\varepsilon,\phi,x} = \frac{2^{\alpha-2}\|\sigma\|^{\frac{2d}{p_0}+2\alpha-2}\|\sigma^{-1}\|^2 C_{T,p_0,\sigma,L(\varepsilon),\varepsilon,x}}{(\pi^d\det(\sigma\sigma^*))^{\frac{1}{p_0}}}\int_0^{2\|\sigma\|^2T}\phi(s)\mathrm{d}s.
$$

The desired assertion (3.1.3) is proved by substituting (3.2.17), (3.2.18), (3.2.20) and (3.2.22) into (3.2.16). Thus, we verified that the conclusion holds under (H1) and (H2).

Step (ii), we prove that if b satisfies the linear growth condition, then the conclusion (3.1.3) holds for time T satisfying (3.1.4). By Remark 3.2.1, we can arrive at the conclusions of Lemma 3.2.1 and Lemma 3.2.2 for any λ, T

satisfying (3.2.13). Then, by checking step (i), for any λ satisfying (3.2.13), we have that

$$\left\{\mathbb{E}\exp\left\{\lambda\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\right\}\right\}\vee\left\{\sup_{0<\delta\leq 1\wedge T}\mathbb{E}\exp\left\{\lambda\int_0^T|\sigma^{-1}b(Y_{s_\delta})|^2\mathrm{d}s\right\}\right\}<\infty. \tag{3.2.23}$$

Combining this with (3.1.4), we have that

$$\mathbb{E}\exp\left\{\frac{(p_0+3)(p_0+1)}{(p_0-1)^2}\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\right\}<\infty. \tag{3.2.24}$$

It is clear that $\frac{(p_0+3)(p_0+1)}{(p_0-1)^2}>\frac{p_0(p_0+1)}{(p_0-1)^2}>\frac{1}{2}$. Taking the same arguments as in step (i), we can then arrive at the second conclusion. The proof is therefore complete. □

Remark 3.2.3. According to the proof of this theorem, the reason why the test function f in (3.1.3) can only be bounded measurable is that the distributions of $X_t^{(\delta)}$ and X_t come from the same process $Y_t=x+\sigma W_t$ by using Girsanov's transformation. This fails for the multiplicative noise case.

3.3 Illustrative examples

According to the proof of Theorem 3.1.1, the condition (H2) comes from the use of the heat kernel of σW_t, (see (3.2.1) and the proof of Lemma 4.2.15). In this section, we give several examples to illustrate the condition (H2) and the convergence rate α.

Example 3.3.1. *If b is Hölder continuous with exponent β, i.e.*

$$|b(y)-b(x)|\leq L|x-y|^\beta,$$

then (H2) holds with $\alpha=\frac{\beta}{2}$ and a constant function $\phi(s)$. It is clear that b has sublinear growth if $\beta<1$. Then for any $T>0$, (3.1.3) holds with $\alpha=\frac{\beta}{2}$.

Proof. By the Hölder continuity and (3.1.6), the assertion follows from the following inequality

$$\begin{aligned}
&\sup_{z\in\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d}|b(y)-b(x)|^{p_0}\frac{\mathrm{e}^{-\frac{|x-z|^2}{s}-\frac{|y-x|^2}{r}}}{s^{\frac{d}{2}}r^{\frac{d}{2}}}\mathrm{d}x\mathrm{d}y\\
&\le L^{p_0}\sup_{z\in\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d}|y-x|^{\beta p_0}\frac{\mathrm{e}^{-\frac{|x-z|^2}{s}-\frac{|y-x|^2}{r}}}{s^{\frac{d}{2}}r^{\frac{d}{2}}}\mathrm{d}x\mathrm{d}y\\
&\le L^{p_0}\frac{1}{s^{\frac{d}{2}}r^{\frac{d}{2}}}\left(\frac{\beta p_0 r}{\mathrm{e}}\right)^{\frac{\beta p_0}{2}}\sup_{z\in\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d}\mathrm{e}^{-\frac{|x-z|^2}{s}}\mathrm{e}^{-\frac{|y-x|^2}{2r}}\mathrm{d}x\mathrm{d}y\\
&\le CL^{p_0}\left(\frac{\beta p_0 r}{\mathrm{e}}\right)^{\frac{\beta p_0}{2}}.
\end{aligned}$$

□

The following example shows that (H2) can hold even if the drift term b is not piecewise continuous.

Example 3.3.2. *Let A be the Smith-Volterra-Cantor set on $[0,1]$, which is constructed in the following way. As the first step, we let $I_{1,1}=\left(\frac{3}{8},\frac{5}{8}\right)$, $J_{1,1}=\left[0,\frac{3}{8}\right]$, $J_{1,2}=\left[\frac{5}{8},1\right]$ and we remove the open interval $I_{1,1}$. As the second step, we remove the middle $\frac{1}{4^2}$ open intervals, denoting by $I_{2,1}$ and $I_{2,2}$, from $J_{1,1}$ and $J_{1,2}$ respectively, i.e. $I_{2,1}=\left(\frac{5}{32},\frac{7}{32}\right)$, $I_{2,2}=\left(\frac{25}{32},\frac{27}{32}\right)$. The intervals left are denoted by $J_{2,1},J_{2,2},J_{2,3},J_{2,4}$, i.e.*

$$J_{2,1}=\left[0,\frac{5}{32}\right],J_{2,2}=\left[\frac{7}{32},\frac{3}{8}\right],J_{2,3}=\left[\frac{5}{8},\frac{25}{32}\right],J_{2,4}=\left[\frac{27}{32},1\right].$$

For the n-th step, we remove the middle $\frac{1}{4^n}$ open intervals $I_{n,1},\cdots,I_{n,2^{n-1}}$ from $J_{n-1,1},\cdots,J_{n-1,2^{n-1}}$ respectively, and the intervals left are denoted by $J_{n,1},\cdots,J_{n,2^n}$. Let

$$A=\bigcap_{n=1}^{\infty}\left(\bigcup_{k=1}^{2^n}J_{n,k}\right).$$

Then A is a nowhere dense set and the Lebesgue measure of A is $1/2$. Define

$$b(x)=\mathbb{1}_{[0,1]}(x)-\sum_{n=1}^{\infty}\sum_{j=1}^{2^{n-1}}2^{-(n+j)}\mathbb{1}_{I_{n,j}}(x)$$

$$= \mathbb{1}_A(x) + \sum_{n=1}^{\infty}\sum_{j=1}^{2^{n-1}} \left(1 - 2^{-(n+j)}\right) \mathbb{1}_{I_{n,j}}(x).$$

All of the endpoints of the intervals $\bar{I}_{n,j}$ are the discontinuous points of b, which is dense in A. For any interval $I \subset [0,1]$ such that $I \cap A \neq \emptyset$, it always contains the discontinuous points of b. However, any interval $I \subset [0,1]$ such that $I \cap A = \emptyset$, it is a subset of some $I_{n,j}$. Hence, b is not a piecewise continuous function. In the following, we shall show that b satisfies condition (H2) with $p_0 = 2$, $\alpha = \frac{1}{4}$ and $\phi(s) = Cs^{-\frac{1}{4}}$.

Proof. For $u > 0$ and any interval (a_1, a_2) (it is similar for $[a_1, a_2]$), it yields that

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \left|\mathbb{1}_{(a_1,a_2)}(x+u) - \mathbb{1}_{(a_1,a_2)}(x)\right|^2 \mathrm{d}x \\
&= \int_{a_1-u}^{a_2-u} \mathbb{1}_{(a_1,a_2)^c}(x)\mathrm{d}x + \int_{a_1}^{a_2} \mathbb{1}_{(a_1-u,a_2-u)^c}(x)\mathrm{d}x \\
&= \int_{a_1-u}^{(a_2-u)\wedge a_1} \mathrm{d}x + \int_{(a_2-u)\vee a_1}^{a_2} \mathrm{d}x \\
&\le 2\left(|u| \wedge (a_2 - a_1)\right).
\end{aligned}$$

For $u < 0$, we obtain that

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \left|\mathbb{1}_{(a_1,a_2)}(x+u) - \mathbb{1}_{(a_1,a_2)}(x)\right|^2 \mathrm{d}x \\
&= \int_{-\infty}^{+\infty} \left|\mathbb{1}_{(a_1,a_2)}(v) - \mathbb{1}_{(a_1,a_2)}(v-u)\right|^2 \mathrm{d}v \le 2\left(|u| \wedge (a_2 - a_1)\right).
\end{aligned}$$

Hence, by Jensen's inequality, it yields that

$$\begin{aligned}
&\int_{-\infty}^{+\infty} |b(x+u) - b(x)|^2 \mathrm{d}x \\
&\le \int_{-\infty}^{+\infty} \Big(\left|\mathbb{1}_{[0,1]}(x+u) - \mathbb{1}_{[0,1]}(x)\right| \\
&\quad + \sum_{n=1}^{+\infty}\sum_{j=1}^{2^{n-1}} 2^{-(n+j)} \left|\mathbb{1}_{I_{n,j}}(x+z) - \mathbb{1}_{I_{n,j}}(x)\right| \Big)^2 \mathrm{d}x
\end{aligned}$$

$$
\begin{aligned}
&\leq \left(1+\sum_{n=1}^{+\infty}\sum_{j=1}^{2^{n-1}} 2^{-(n+j)}\right)\left\{\int_{-\infty}^{+\infty} \left|\mathbb{1}_{[0,1]}(x+u)-\mathbb{1}_{[0,1]}(x)\right|^2 \mathrm{d}x\right.\\
&\qquad \left.+\sum_{n=1}^{+\infty}\sum_{j=1}^{2^{n-1}} 2^{-(n+j)} \int_{-\infty}^{+\infty} \left|\mathbb{1}_{I_{n,j}}(x+u)-\mathbb{1}_{I_{n,j}}(x)\right|^2 \mathrm{d}x\right\}\\
&\leq 2\left(1+\sum_{n=1}^{\infty}\sum_{j=1}^{2^{n-1}} 2^{-(n+j)}\right)^2 |u| = 4|u|.
\end{aligned}
$$

Combining this with (3.1.6), we obtain that

$$
\begin{aligned}
&\sup_{z\in\mathbb{R}} \int_{\mathbb{R}\times\mathbb{R}} |b(y)-b(x)|^2 \frac{\mathrm{e}^{-\frac{|x-z|^2}{s}}\,\mathrm{e}^{-\frac{|y-x|^2}{r}}}{s^{\frac{1}{2}} r^{\frac{1}{2}}} \mathrm{d}x\mathrm{d}y\\
&\qquad \leq \frac{1}{s^{\frac{1}{2}} r^{\frac{1}{2}}} \int_{\mathbb{R}} \mathrm{e}^{-\frac{|u|^2}{r}} \int_{\mathbb{R}} |b(x+u)-b(x)|^2 \mathrm{d}x\mathrm{d}u\\
&\qquad \leq \frac{4}{s^{\frac{1}{2}} r^{\frac{1}{2}}} \int_{\mathbb{R}} \mathrm{e}^{-\frac{|u|^2}{r}} |u|\mathrm{d}u = \left(C s^{-\frac{1}{4}} r^{\frac{1}{4}}\right)^2.
\end{aligned}
$$

□

A general class of functions that satisfies (H2) is the (fractional) Sobolev space $W^{\beta,p}(\mathbb{R}^d)$, see the following example:

Example 3.3.3. *If there exist $\beta > 0$ and $p \in [2,\infty)\cap(d,+\infty)$ such that the Gagliardo seminorm of b is finite, i.e.*

$$
[b]_{W^{\beta,p}} := \left(\int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|b(x)-b(y)|^p}{|x-y|^{d+\beta p}} \mathrm{d}x\mathrm{d}y\right)^{\frac{1}{p}} < \infty,
$$

then (H2) holds for any $p_0 = p$ with $\alpha = \frac{\beta}{2}$ and $\phi(s) = C_1 s^{-\frac{d}{2}} [b]_{W^{\beta,p}}^p$. Hence, if b satisfies (H1) and $[b]_{W^{\beta,p}} < \infty$ with $p \in [2,\infty)\cap(d,+\infty)$, then (3.1.3) *holds.*

Proof. Indeed, by Hölder's inequality and (3.1.6), it follows that

$$
\frac{1}{(rs)^{\frac{d}{2}}} \int_{\mathbb{R}^d\times\mathbb{R}^d} |b(y)-b(x)|^p\, \mathrm{e}^{-\frac{|x-z|^2}{s}-\frac{|y-x|^2}{r}}\, \mathrm{d}x\mathrm{d}y
$$

$$
\begin{aligned}
&= \frac{1}{(rs)^{\frac{d}{2}}} \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|b(x)-b(y)|^p}{|x-y|^{d+\beta p}}\, \mathrm{e}^{-\frac{|x-z|^2}{s}-\frac{|y-x|^2}{r}}\, |x-y|^{d+\beta p}\mathrm{d}x\mathrm{d}y \\
&\leq C_1 s^{-\frac{d}{2}} r^{\frac{\beta p}{2}} \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|b(x)-b(y)|^p}{|x-y|^{d+\beta p}}\, \mathrm{e}^{-\frac{|x-z|^2}{s}-\frac{|y-x|^2}{2r}}\, \mathrm{d}x\mathrm{d}y \\
&\leq C_1 s^{-\frac{d}{2}} r^{\frac{\beta p}{2}} \, [b]^p_{W^{\beta,p}} .
\end{aligned}
$$

□

From these examples, one can see that the drift could be very irregular. This means that we have extended the results in [3] where the coefficients must be smooth. However, our method is not optimal in the Lipschitz case since the classical weak convergence rate is $\alpha = 1$ for SDEs with smooth coefficients in [3].

Chapter 4

Weak approximation for stochastic functional differential equations driven by fractional Brownian motion

In this chapter, we investigate the weak existence and uniqueness of SFDEs with singular coefficients, and obtain the associated weak approximation of its truncated EM scheme.

In Section 4.1, we first give the associated assumptions about the coefficients of the model, we then obtain the first main result about the existence and uniqueness of solutions to the model. Section 4.2 is devoted to the proof of the first main result (i.e. Theorem 4.1.1). In Section 4.3, we first introduce other assumptions about the coefficients in our framework (that is, weak approximation rate of EM scheme for the model (4.1.1)) and establish the main result (i.e. Theorem 4.3.1). Finally, Section 4.4 is devoted to the proof of Theorem 4.3.1.

4.1 Well-posedness of stochastic functional differential equations

Let $(\mathbb{R}^d, \langle\cdot,\cdot\rangle, |\cdot|)$ be the d-dimensional Euclidean space. Let $\mathbb{R}^d\otimes\mathbb{R}^m$ be the set of all $d\times m$-matrices. Let $\tau>0$ be a fixed number and $\mathscr{C}=C([-\tau,0];\mathbb{R}^d)$, which is endowed with the uniform norm $\|f\|_\infty := \sup_{-\tau\le\theta\le 0}|f(\theta)|$. For $f\in C([-\tau,\infty);\mathbb{R}^d)$ and fixed $t>0$, define the segment $f_t\in\mathscr{C}$ by $f_t(\theta)=f(t+\theta), \theta\in[-\tau,0]$. $\mathscr{B}_b(\mathbb{R}^d)$ be the collection of all bounded measurable functions on $\mathbb{R}^d$. For any $\alpha\in(0,1)$, let $C^\alpha(a,b)$ be the space of α-Hölder continuous functions f on the interval $[a,b]$ and set

$$\|f\|_{a,b,\alpha} := \sup_{a\le s\le t\le b}\frac{|f(t)-f(s)|}{|t-s|^\alpha}.$$

In this chapter, for $H\in(\frac{1}{2},1)$, we consider the following equation:

$$\mathrm{d}X(t) = \{b(X(t)) + \sigma Z(X_t)\}\mathrm{d}t + \sigma\mathrm{d}B^H(t), \quad t>0, \tag{4.1.1}$$

with the initial datum $X_0=\xi\in\mathscr{C}$, where $\sigma\in\mathbb{R}^d\otimes\mathbb{R}^m$, $b:\mathbb{R}^d\to\mathbb{R}^d, d\ge m$ and $Z:\mathscr{C}\to\mathbb{R}^m$ are measurable, X_t is the segment process of $X(t)$, and $B^H(t)$ is an m-dimensional fBm on a complete filtration probability space $(\Omega,\mathscr{F},(\mathscr{F}_t)_{t\ge0},\mathbb{P})$. Consider a reference SDE as follows:

$$\mathrm{d}Y(t) = b(Y(t))\mathrm{d}t + \sigma\mathrm{d}B^H(t), \quad t>0, \quad Y(0)\in\mathbb{R}^d. \tag{4.1.2}$$

Let $\xi\in\mathscr{C}$, and let $Y^{\xi(0)}(\cdot)$ be a solution of (4.1.2) with $Y^{\xi(0)}(0)=\xi(0)$. We extend $Y^{\xi(0)}(\cdot)$ from $[0,\infty)$ to $[-\tau,\infty)$ in the following way:

$$Y^\xi(t) = \xi(t)I_{[-\tau,0)}(t) + Y^{\xi(0)}(t)I_{[0,\infty)}(t), \quad t\in[-\tau,\infty), \quad \xi\in\mathscr{C}. \tag{4.1.3}$$

The weak existence and uniqueness of solutions to (4.1.1) will then be studied by using Girsanov's transform and the extended solutions to the reference equation (4.1.2).

Before move on, we first introduce the following assumptions on b and Z for the weak existence and uniqueness result.

(A1) There exists a constant $K_1 \in \mathbb{R}$ such that

$$\langle b(x)-b(y), x-y\rangle \le K_1|x-y|^2, \quad x, y \in \mathbb{R}^d.$$

(A2) There exist $C_1 > 0$ and $q_0 \ge 0$ such that $|b(x)| \le C_1(1+|x|^{q_0}), \ x \in \mathbb{R}^d$.

(A3) There exist $\alpha \in (H-1/2, 1]$, $p > 0$, $C_2 > 0$, $C_3 \ge 0$ and $q_1 \ge 0$ such that

$$|Z(\eta_1)-Z(\eta_2)| \le C_2\|\eta_1-\eta_2\|_\infty^\alpha \left(1+\|\eta_1\|_\infty^p+\|\eta_2\|_\infty^p\right), \quad (4.1.4)$$

$$\langle \sigma Z(\eta_1+\eta_2), \eta_1(0)\rangle \le C_3\left(1+\|\eta_2\|_\infty^{q_1}+\|\eta_1\|_\infty^2\right), \ \eta_1, \eta_2 \in \mathscr{C}. \quad (4.1.5)$$

Then, we give the first main result of existence and uniqueness of weak solutions to (4.1.1).

Theorem 4.1.1. *Assume (A1)-(A3). For any $\xi \in \mathscr{C}$, there exist $\theta \in (\frac{2H-1}{2\alpha}, 1]$ and $\bar{C}_1 > 0$ such that*

$$|\xi(r)-\xi(s)| \le \bar{C}_1|r-s|^\theta, \quad -\tau \le r \le s \le 0, \quad (4.1.6)$$

then the equation (4.1.1) *has a unique weak solution with $X_0 = \xi$.*

Remark 4.1.1. The condition (4.1.6) is for us to use Girsanov's transformation to remove the drift term $Z(\cdot)$ of equation (4.1.1). For given $T > 0$, $\forall \gamma \in C([-\tau, T], \mathbb{R}^d)$ with $\gamma_0 = \xi_0$, to ensure that $\{\int_0^s Z(\gamma_r)\mathrm{d}r\}_{s\in[0,T]}$ belongs to the Cameron-Martin space of the fBm, it is necessary that the integral $\int_0^\cdot Z(\gamma_s)\mathrm{d}s \in I_{0+}^{H+\frac{1}{2}}(L^2([0,T], \mathbb{R}^d))$. This means that we need $Z(\gamma_\cdot) \in I_{0+}^{H-\frac{1}{2}}(L^2([0,T], \mathbb{R}^d))$. Note that for $t \in [0, T\wedge\tau]$, we have

$$\|\gamma_\cdot\|_{0,t,\theta} = \sup_{0\le r\le u\le t} \frac{\|\gamma_u-\gamma_r\|_\infty}{|u-r|^\theta} = \sup_{0\le r\le u\le t, v\in[-\tau,0]} \frac{|\gamma(u+v)-\gamma(r+v)|}{(u-r)^\theta} \ge \|\xi\|_{-\tau,0,\theta}.$$

Hence, despite imposing the regularity condition (4.1.4) on Z, we also need an additional assumption on the initial value ξ. If Z is α-Hölder continuous and ξ is θ-Hölder continuous, then our conditions on ξ yield that $\theta\alpha > H - \frac{1}{2}$, which ensures that $\{\int_0^s Z(\gamma_r)\mathrm{d}r\}_{s\in[0,T]}$ is in the Cameron-Martin space (see (4.2.13) for more details).

4.2 Proof of the main theorem

Before giving the proof of Theorem 4.1.1, we first introduce the following lemma on the existence and uniqueness of solutions to reference SDE (4.1.2).

Lemma 4.2.1. *Assume (A1). Then* (4.1.2) *has a unique strong solution and*

$$|Y(t)| \leq \mathrm{e}^{\frac{\bar{K}_2 t}{2}}|Y(0)| + \sqrt{\bar{K}_2}\left(\int_0^t e^{\bar{K}_1(t-r)}\left|b(\sigma B^H(r))\right|^2 \mathrm{d}r\right)^{\frac{1}{2}} + |\sigma B^H(t)|,\ t \geq 0. \tag{4.2.1}$$

Furthermore, if (A2) holds, then

$$\mathbb{E}\|Y\|_{0,t,\beta}^q < \infty,\ q > 0, t > 0, 0 < \beta < H.$$

Proof. (1) Let $U(t) = Y(t) - \sigma B^H(t)$. Then $U(t)$ satisfies

$$\mathrm{d}U(t) = b(U(t) + \sigma B^H(t))\mathrm{d}t,\ U(0) = Y(0). \tag{4.2.2}$$

Set $\bar{b}(u,t) = b(u + \sigma B^H(t))$. Then one can see from (A1) that

$$\langle \bar{b}(u_1,t) - \bar{b}(u_2,t), u_1 - u_2\rangle \leq K_1|u_1 - u_2|^2,$$

which implies that (4.2.2) has a unique solution. Moreover, it follows from the chain rule and the Hölder inequality that

$$\mathrm{d}|U(t)|^2 = 2\langle \bar{b}(U(t),t), U(t)\rangle \mathrm{d}t$$

$$
\begin{aligned}
&\leq 2K_1|U(t)|^2 + 2\langle b(\sigma B^H(t)), U(t)\rangle \mathrm{d}t \\
&\leq \bar{K}_1|U(t)|^2\mathrm{d}t + \bar{K}_2\left|b(\sigma B^H(t))\right|^2\mathrm{d}t,
\end{aligned}
$$

which implies that

$$
|U(t)|^2 \leq \left\{\bar{K}_2\int_s^t\left|b(\sigma B^H(r))\right|^2\mathrm{d}r + |U(s)|^2\right\}\mathrm{e}^{\bar{K}_1(t-s)}.
$$

This, together with $U(t) = Y(t) - \sigma B^H(t)$, yields that for any $t \geq s$

$$
\begin{aligned}
|Y(t)| \leq \mathrm{e}^{\frac{1}{2}(t-s)\bar{K}_1}|Y(s)| + \sqrt{\bar{K}_2}\left(\int_s^t e^{\bar{K}_1(t-r)}\left|b(\sigma B^H(r))\right|^2\mathrm{d}r\right)^{\frac{1}{2}} \\
+ \mathrm{e}^{\frac{\bar{K}_1(t-s)}{2}}|\sigma B^H(s)| + |\sigma B^H(t)|,
\end{aligned}
$$

which implies our first claim (4.2.1).

(2) For any $0 < \beta < H$, we derive from (A2) that

$$
\begin{aligned}
&\frac{|Y(t) - Y(s)|}{(t-s)^\beta} \\
&\leq \frac{1}{(t-s)^\beta}\int_s^t|b(Y(r))|\mathrm{d}r + \|\sigma\|\|B^H\|_{0,t,\beta} \\
&\leq \frac{C_1}{(t-s)^\beta}\int_s^t\left(1 + |Y(r)|^{q_0}\right)\mathrm{d}r + \|\sigma\|\|B^H\|_{0,t,\beta} \\
&\leq C_1(t-s)^{1-\beta} + \|\sigma\|\|B^H\|_{0,t,\beta} + C_1 3^{(q_0-1)^+}(t-s)^{1-\beta}\,\mathrm{e}^{\frac{t\bar{K}_1^+ q_0}{2}}|Y(0)|^{q_0} \\
&\quad + C_1 3^{(q_0-1)^+}\left(\sqrt{\bar{K}_2}C_1\left(1 + \|\sigma\|^{q_0}\|B^H\|^{q_0}_{0,t,\infty}\right)\right)^{q_0}\mathrm{e}^{\frac{q_0\bar{K}_1^+ t}{2}}t^{q_0}(t-s)^{1-\beta} \\
&\quad + C_1 3^{(q_0-1)^+}\left(\|\sigma\|\|B^H\|_{0,t,\infty}\right)^{q_0}(t-s)^{1-\beta},
\end{aligned}
$$

which yields

$$
\begin{aligned}
&\|Y\|_{0,t,\beta} \qquad (4.2.3)\\
&\leq C_1 t^{1-\beta} + \|\sigma\|\|B^H\|_{0,t,\beta} + C_1 3^{(q_0-1)^+}t^{1-\beta}\|\sigma\|^{q_0}\|B^H\|^{q_0}_{0,t,\infty} \\
&+ C_1 3^{(q_0-1)^+}t^{1-\beta}\,\mathrm{e}^{\frac{t\bar{K}_1^+ q_0}{2}}\left(\|Y\|_{0,t,\infty} + t^{q_0}\left(\sqrt{\bar{K}_2}C_1\left(1 + \|\sigma\|^{q_0}\|B^H\|^{q_0}_{0,t,\infty}\right)\right)^{q_0}\right). \qquad (4.2.4)
\end{aligned}
$$

Combining this with (4.2.1), it is clear that our second claim holds. □

Fix any $T > 0$. Let

$$\left\{\tilde{B}^H(t)\right\}_{t\in[0,T]} = \left\{B^H(t) - \int_0^t Z(Y_s^\xi)\mathrm{d}s\right\}_{t\in[0,T]},$$

$$R^\xi(t) = \exp\left(\int_0^t \left\langle K_H^{-1}\left(\int_0^\cdot Z(Y_r^\xi)\mathrm{d}r\right)(s), \mathrm{d}B(s)\right\rangle \right.$$
$$\left. -\frac{1}{2}\int_0^t \left|K_H^{-1}\left(\int_0^\cdot Z(Y_r^\xi)\mathrm{d}r\right)\right|^2(s)\mathrm{d}s\right), t \in [0,T], \tag{4.2.5}$$

where $B(t)_{t\geq 0}$ is a standard d-dimensional Brownian motion, and Y^ξ satisfies the following SDE

$$\mathrm{d}Y^\xi(t) = b(Y^\xi(t))\mathrm{d}t + \sigma Z(Y_t^\xi)\mathrm{d}t + \sigma \mathrm{d}\tilde{B}^H(t). \tag{4.2.6}$$

The following lemma is for the investigation of the exponential martingale, which is crucial to prove Theorem 4.1.1.

Lemma 4.2.2. *Let the assumptions of Theorem 4.1.1 hold. Then*

(1) $\left\{\tilde{B}^H(t)\right\}_{t\in[0,T]}$ *is a fBm under* $R^\xi(T)\mathbb{P}$.

(2) *Assume in addition that* $q_0 = 1$ *in (A2). If there exist* $C_4 \geq 0$, $C_5 \geq 0$ *and* $p \in (0,1)$ *such that*

$$|Z(\eta_1) - Z(\eta_2)| \leq C_4\{\|\eta_1 - \eta_2\|_\infty^\alpha \wedge (1 + C_5(\|\eta_1\|_\infty^p + \|\eta_2\|_\infty^p))\}, \tag{4.2.7}$$

then for any $C \geq 0$, *it holds that*

$$\mathbb{E}\exp\left\{C\int_0^T \left|K_H^{-1}\left(\int_0^\cdot Z(Y_r^\xi)\mathrm{d}r\right)\right|^2(s)\mathrm{d}s\right\} < \infty. \tag{4.2.8}$$

(3) *If* $q_0 = 1$, (4.2.7) *holds with* $p = 1$ *and* $T > 0$ *satisfying*

$$\frac{1}{(1-H)\pi}\Big\{24C_4^2C_5^2T^2(H-1/2)^2\|\sigma\|^2 + \Big\{\frac{3\theta^2T^{3-2H}2^{2H-1}}{(\theta - H + \frac{1}{2})^2}\|\sigma\|^2 \tag{4.2.9}$$

$$+ 3C_4^2T^{2(1+\alpha\theta-\theta)}(1+(H-1/2)^2C_0^2)\|\sigma\|^{2\alpha}\Big\}\mathbf{1}_{[\alpha=1]}\Big\}$$
$$\leq 2^{-9}\Big\{1+C_1Te^{\frac{T\bar{K}_1^+}{2}}\left(L_1\Phi(\bar{K}_1,\bar{K}_2,T)+2+T\sqrt{\bar{K}_2}C_1\right)\Big\}^{-2}, \tag{4.2.10}$$

where C_0 is defined in Theorem 4.3.1, then (4.2.8) *holds for some* $C > 1$.

Proof. If (4.1.6) holds for $\theta \geq H$, then one sees that (4.1.6) holds for $\theta \in (H-1/2, H)$. Hence, we shall assume that $\theta \in (H-1/2, H)$ in the following proof.

(1) It follows from (2.2.3) that

$$\begin{aligned}
&K_H^{-1}\left(\int_0^{\cdot} Z(Y_r^\xi)\mathrm{d}r\right)(s)\\
&= s^{H-\frac{1}{2}}D_{0+}^{H-\frac{1}{2}}\left(\cdot^{\frac{1}{2}-H}Z(Y_\cdot^\xi)\right)(s)\\
&= \frac{H-\frac{1}{2}}{\Gamma(\frac{3}{2}-H)}\left[\frac{s^{\frac{1}{2}-H}}{H-\frac{1}{2}}Z(Y_s^\xi)+s^{H-\frac{1}{2}}Z(Y_s^\xi)\int_0^s\frac{s^{\frac{1}{2}-H}-r^{\frac{1}{2}-H}}{(s-r)^{\frac{1}{2}+H}}\mathrm{d}r\right.\\
&\quad\left.+s^{H-\frac{1}{2}}\int_0^s\frac{Z(Y_s^\xi)-Z(Y_r^\xi)}{(s-r)^{\frac{1}{2}+H}}r^{\frac{1}{2}-H}\mathrm{d}r\right]\\
&=: \frac{H-\frac{1}{2}}{\Gamma(\frac{3}{2}-H)}(J_1(s)+J_2(s)+J_3(s)).
\end{aligned} \tag{4.2.11}$$

By (4.1.4), we have

$$\begin{aligned}
&|J_1(s)|^2+|J_2(s)|^2\\
&\leq \frac{s^{1-2H}}{(H-\frac{1}{2})^2}\left(C_2\left(1+\|Y_s^\xi\|_\infty^p\right)\|Y_s^\xi\|_\infty^\alpha+|Z(0)|\right)^2\\
&+s^{2H-1}\left|\int_0^s\frac{s^{\frac{1}{2}-H}-r^{\frac{1}{2}-H}}{(s-r)^{\frac{1}{2}+H}}\mathrm{d}r\right|^2\left(C_2\left(1+\|Y_s^\xi\|_\infty^p\right)\|Y_s^\xi\|_\infty^\alpha+|Z(0)|\right)^2\\
&\leq 3s^{1-2H}\Big(\frac{4}{(2H-1)^2}+C_0^2\Big)\left(|Z(0)|^2+C_2^2(\|Y_s^\xi\|_\infty^{2\alpha}+\|Y_s^\xi\|_\infty^{2(p+\alpha)})\right),
\end{aligned} \tag{4.2.12}$$

and

$$\begin{aligned}
|J_3(s)| &\le C_2 s^{H-\frac{1}{2}} \int_0^s \frac{\|Y_s^\xi - Y_r^\xi\|_\infty^\alpha \left(1 + \|Y_s^\xi\|_\infty^p + \|Y_r^\xi\|_\infty^p\right)}{(s-r)^{H+1/2} r^{H-1/2}} \mathrm{d}r \\
&\le 2C_2 s^{H-\frac{1}{2}} \left(1 + \|Y^\xi\|_{-\tau,s,\infty}^p\right) \int_0^s \frac{\|Y^\xi\|_{-\tau,r,\theta}^\alpha (s-r)^{\theta\alpha}}{(s-r)^{H+1/2} r^{H-1/2}} \mathrm{d}r \\
&= 2C_2 s^{\theta\alpha+\frac{1}{2}-H} \mathcal{B}\left(\frac{3}{2} - H, \theta\alpha + \frac{1}{2} - H\right) \left(1 + \|Y^\xi\|_{-\tau,s,\infty}^p\right) \|Y^\xi\|_{-\tau,s,\theta}^\alpha,
\end{aligned} \tag{4.2.13}$$

where $\mathcal{B}$ is Beta function.

Combining this with (4.2.12), we arrive at

$$\begin{aligned}
&\left|K_H^{-1}\left(\int_0^\cdot Z(Y_r^\xi)\mathrm{d}r\right)\right|^2(s) \\
&\le 9s^{1-2H}\left(\frac{4}{(2H-1)^2} + C_0^2\right)\left(|Z(0)|^2 + C_2^2(\|Y_s^\xi\|_\infty^{2\alpha} + \|Y_s^\xi\|_\infty^{2(p+\alpha)})\right) \\
&+ 12C_2^2 s^{2\theta\alpha+1-2H}\mathcal{B}^2\left(\frac{3}{2} - H, \theta\alpha + \frac{1}{2} - H\right)\left(1 + \|Y^\xi\|_{-\tau,s,\infty}^{2p}\right)\|Y^\xi\|_{-\tau,s,\theta}^{2\alpha},
\end{aligned} \tag{4.2.14}$$

Let

$$\tau_n = \inf\left\{t > 0 \;\Big|\; \int_0^t \left|K_H^{-1}\left(\int_0^\cdot Z(Y_r^\xi)\mathrm{d}r\right)\right|^2(s)\mathrm{d}s \ge n\right\}, \qquad n \in \mathbb{N}.$$

We then know $\{R^\xi(r\wedge\tau_n)\}_{r\in[0,T]}$ which was defined in (4.2.5) is an exponential martingale. The Girsanov theorem (e.g. see [81, Proposition 4.1.2]) implies that

$$\tilde{B}_n(t) := B(t) - \int_0^{t\wedge\tau_n} K_H^{-1}\left(\int_0^\cdot Z(Y_r^\xi)\mathrm{d}r\right)(s)\mathrm{d}s,\ t\in[0,T]$$

is a Brownian motion under $R^\xi(T\wedge\tau_n)\mathbb{P}$. This implies

$$\tilde{B}_n^H(t) := B^H(t) - \int_0^{t\wedge\tau_n} Z(Y_r^\xi)\mathrm{d}r,\ t\ge 0$$

is a fBm under $R^\xi(T\wedge\tau_n)\mathbb{P}$, and Y^ξ satisfies

$$\mathrm{d}Y^\xi(t) = b(Y^\xi(t))\mathrm{d}t + \sigma\mathrm{d}\tilde{B}_n^H(t) + \mathbf{1}_{[0\le t\le\tau_n]}\sigma Z(Y_t^\xi)\mathrm{d}t.$$

Let $u^\xi(t)=Y^\xi(t)-\sigma\tilde{B}^H_n(t)$, Then for $0\le t\le\tau_n$, we derive from (4.1.5) that

$$\begin{aligned}\mathrm{d}u^\xi(t)^2&=2\langle b(u^\xi(t)+\sigma\tilde{B}^H_n(t)),u^\xi(t)\rangle\mathrm{d}t+2\langle\sigma Z(u^\xi_t+\sigma\tilde{B}^H_{n,t}),u^\xi(t)\rangle\mathrm{d}t\\&\le\bar{K}_1(u^\xi(t))^2\mathrm{d}t+C_1^2\bar{K}_2(1+|\sigma\tilde{B}^H_n(t)|^{q_0})^2\mathrm{d}t\\&+2C_3(1+\|u^\xi_t\|^2_\infty+\|\sigma\|^{q_1}\|\tilde{B}^H_{n,t}\|^{q_1}_\infty)\mathrm{d}t.\end{aligned}$$

Then, one gets that

$$\begin{aligned}&|Y^\xi(t)|^2\\&\le|Y^\xi(0)|^2+2(\bar{K}_1+2C_3)\int_0^t\|Y^\xi_s\|^2_\infty\mathrm{d}s+2\bar{K}_2C_1^2\int_0^t(1+|\sigma\tilde{B}^H_n(s)|^{q_0})^2\mathrm{d}s\\&+4C_3\int_0^t(1+\|\sigma\|^{q_1}\|\tilde{B}^H_{n,s}\|^{q_1}_\infty)\mathrm{d}s+2|\sigma\tilde{B}^H_n(t)|^2\\&=:F^2(\tilde{B}^H_n)(t)+2(\bar{K}_1+2C_3)\int_0^t\|Y^\xi_s\|^2_\infty\mathrm{d}s.\end{aligned}$$

Note that $\sup_{0\le s\le t}\|Y^\xi_s\|_\infty\le\|\xi\|_\infty\vee\sup_{0\le s\le t}|Y^\xi(s)|$, we arrive at that for any $q_3\ge 2$,

$$\begin{aligned}&\|Y^\xi\|^{q_3}_{0,t,\infty}\\&\le 3^{q_3/2-1}\Big(\|\xi\|^{q_3}_\infty+F^{q_3}(\tilde{B}^H_n)(t)+2^{q_3/2}(\bar{K}_1+2C_3)^{q_3/2}t^{q_3/2-1}\int_0^t\|Y^\xi\|^{q_3}_{0,s,\infty}\mathrm{d}s\Big),\end{aligned}$$

Combining this with Gronwall's lemma, it yields that

$$\|Y^\xi\|^{q_3}_{0,t,\infty}\le 3^{q_3/2-1}\Big(\|\xi\|^{q_3}_\infty+F^{q_3}(\tilde{B}^H_n)(t)\Big)\,\mathrm{e}^{(3t)^{q_3/2-1}2^{q_3/2}(\bar{K}_1+2C_3)^{q_3/2}t}.\tag{4.2.15}$$

Similarly, following the proof of (4.2.3), we get for $\beta\in(0,H)$ that

$$\begin{aligned}\|Y\|_{0,t,\beta}\le&\Big\{C_1\Big(1+\|Y^\xi\|^{q_0}_{0,t,\infty}\Big)+\|\sigma\||Z(0)|\\&+C_2\|\sigma\|\|Y^\xi\|^\alpha_{-\tau,t,\infty}\Big(1+\|Y^\xi\|^p_{-\tau,t,\infty}\Big)\Big\}t^{1-\beta}+\|\sigma\|\|\tilde{B}^H_n\|_{0,t,\beta},\ t>0.\end{aligned}\tag{4.2.16}$$

Combining this with (4.2.15), and that $\{\tilde{B}_n^H(t)\}_{t\in[0,T]}$ under $R^\xi(T\wedge\tau_n)\mathbb{P}$ has the same distribution as $\{B^H(t)\}_{t\in[0,T]}$ under $\mathbb{P}$, we have

$$\sup_n \mathbb{E}R^\xi(T\wedge\tau_n)\left(\|Y^\xi\|_{-\tau,T,\infty}^{q_3}+\|Y^\xi\|_{-\tau,T,\theta}^{q_3}\right)<\infty,\qquad q_3>0,\,\theta<H.$$

Combining this with (4.2.5), (4.2.14) and (4.2.15), we obtain

$$\begin{aligned}&\sup_{t\in[0,T],n}\ \mathbb{E}R^\xi(t\wedge\tau_n)\log R^\xi(t\wedge\tau_n)\\&\le C\sup_{t\in[0,T],n}\ \mathbb{E}R^\xi(t\wedge\tau_n)\left(1+\|Y^\xi\|_{-\tau,t,\infty}^{2(p+\alpha)}+\|Y^\xi\|_{-\tau,t,\theta}^{2\alpha}+\|Y^\xi\|_{-\tau,t,\theta}^{2\alpha}\|Y\|_{-\tau,t,\infty}^{2p}\right)\\&<\infty.\end{aligned}$$

Hence, it follows from the Fatou lemma and the martingale convergence theorem that $\{R^\xi(t)\}_{t\in[0,T]}$ is a uniformly integrable martingale and

$$\sup_{t\in[0,T]}\ \mathbb{E}R^\xi(t)\log R^\xi(t)<\infty.$$

It follows from Girsanov's theorem that under $R^\xi(T)\mathbb{P}$, the process $\tilde{B}^H$ is a fBm.

(2) By (4.2.7), we derive from (4.2.11) that

$$\begin{aligned}&|J_1(s)|^2+|J_2(s)|^2\\&\le 2C_4^2s^{1-2H}\left(\frac{4}{(2H-1)^2}+C_0^2\right)\left((1+C_5\|Y^\xi\|_{-\tau,s,\infty}^p)^2\wedge\|Y^\xi\|_{-\tau,s,\infty}^{2\alpha}+|Z(0)|^2\right).\end{aligned}\tag{4.2.17}$$

For J_3, one gets from (4.2.7) that

$$\begin{aligned}&|J_3(s)|\\&\le C_4s^{H-\frac{1}{2}}\int_0^s\frac{\|Y_s^\xi-Y_r^\xi\|_\infty^\alpha\wedge\left(1+C_5(\|Y_s^\xi\|_\infty^p+\|Y_r\|_\infty^p)\right)}{(s-r)^{H+1/2}r^{H-1/2}}\mathrm{d}r\\&\le 2C_4s^{H-\frac{1}{2}}\int_0^s\frac{\left(1+C_5\|Y^\xi\|_{-\tau,s,\infty}^p\right)\wedge\left(\|Y^\xi\|_{-\tau,s,\theta}^\alpha(s-r)^{\theta\alpha}\right)}{(s-r)^{H+1/2}r^{H-1/2}}\mathrm{d}r\end{aligned}$$

$$
\leq 2C_4 s^{H-\frac{1}{2}}\left(1+C_5\|Y^\xi\|^p_{-\tau,s,\infty}\right)\int_0^s \frac{\|Y^\xi\|^\alpha_{-\tau,s,\theta}(s-r)^{\theta\alpha-H-1/2}r^{-H+1/2}}{1+C_5\|Y^\xi\|^p_{-\tau,s,\infty}+\|Y^\xi\|^\alpha_{-\tau,s,\theta}(s-r)^{\theta\alpha}}\mathrm{d}r
$$

$$
\leq 2C_4 s^{H-1/2}\left(1+C_5\|Y^\xi\|^p_{-\tau,s,\infty}\right)
$$

$$
\times\left(\frac{2^{2H-1}s^{1-2H}}{3/2-H}+\frac{\theta\alpha 2^{H-1/2}s^{\frac{1}{2}-H}\|Y^\xi\|^{\frac{H-1/2}{\theta}}_{-\tau,s,\theta}\left(1+C_5\|Y^\xi\|^p_{-\tau,s,\infty}\right)^{-\frac{H-1/2}{\theta\alpha}}}{(\alpha\theta-H+1/2)(H-1/2)}\right),
$$

where we used [31, Lemma 3.4] in the last inequality, and it yields that

$$
\begin{aligned}
|J_3(s)|^2 &\leq C_6 s^{1-2H}(1+C_5\|Y^\xi\|^p_{-\tau,s,\infty})^2 \\
&\quad + C_7\left(1+C_5\|Y^\xi\|^p_{-\tau,s,\infty}\right)^{\frac{2\theta\alpha-2H+1}{\theta\alpha}}\|Y^\xi\|^{\frac{2H-1}{\theta}}_{-\tau,s,\theta},
\end{aligned} \tag{4.2.18}
$$

where

$$
C_6=8C_4^2,\qquad C_7=\left(\frac{\theta\alpha 2^H}{(\alpha\theta-H+1/2)(H-1/2)}\right)^2.
$$

Since $\|Y^\xi\|_{-\tau,s,\infty}\leq(s\vee\tau)^\theta\|Y^\xi\|_{-\tau,s,\theta}+|\xi(0)|$, it follows from (4.2.17) and (4.2.18) that

$$
\begin{aligned}
&\mathbb{E}\exp\left\{C\int_0^T\left|K_H^{-1}\left(\int_0^\cdot Z(Y_r^\xi)\mathrm{d}r\right)(s)\right|^2\mathrm{d}s\right\} \\
&\leq \mathbb{E}\exp\left(C_T\left(1+\|Y^\xi\|^{2p\vee\frac{2\theta\alpha p+(2H-1)(\alpha-p)}{\theta\alpha}}_{-\tau,T,\theta}\right)\right) \\
&= \mathbb{E}\exp\left(C_T\left(1+\|Y^\xi\|^{2p+\frac{(2H-1)(\alpha-p)^+}{\theta\alpha}}_{-\tau,T,\theta}\right)\right).
\end{aligned}
$$

For $p<1$, it is clear that

$$
2p+\frac{(2H-1)(\alpha-p)^+}{\theta\alpha}<2.
$$

Then (4.2.8) follows from (4.2.3) with $q_0=1,\beta=\theta$ and the Fernique-type lemma 2.2.1.

(3) For $p=1$, substituting (4.2.17) and (4.2.18) into (4.2.11), we have

$$
\int_0^T\left|K_H^{-1}\left(\int_0^\cdot Z(Y_r^\xi)\mathrm{d}r\right)\right|^2(s)\mathrm{d}s
$$

$$\leq C_8(T) + \frac{3C_4^2 T^{2-2H}(1+(H-1/2)^2 C_0^2)}{(1-H)\Gamma^2(3/2-H)} \|Y^\xi\|_{-\tau,T,\infty}^{2\alpha}$$
$$+ \frac{3C_6 C_5^2 T^{2-2H}(H-1/2)^2}{(1-H)\Gamma^2(3/2-H)} \|Y^\xi\|_{-\tau,T,\infty}^2 + \frac{3C_7 T^{2-2H}(H-1/2)^2}{2(1-H)\Gamma^2(3/2-H)}$$
$$\times (1 + C_5|\xi(0)|) \|Y^\xi\|_{-\tau,T,\theta}^{\frac{2H-1}{\theta}} + \frac{3C_7 T^{2(1-H+\theta)-\frac{2H-1}{\alpha}}(H-1/2)^2}{2(1-H)\Gamma^2(3/2-H)} \|Y^\xi\|_{-\tau,T,\theta}^{2+\frac{2H-1}{\theta}(1-\alpha^{-1})},$$

where

$$C_8(T) = \frac{3T^{2-2H}(1+(H-1/2)^2 C_0^2)}{(1-H)\Gamma^2(3/2-H)} \Big(|Z(0)|^2 + C_6(H-1/2)^2\Big).$$

It follows from (4.2.1) and (A2) with $q_0 = 1$, we have

$$\|Y^\xi\|_{0,T,\infty} \leq \tilde{C}(T) + \big(L_1 \Phi(\bar{K}_1, \bar{K}_2, T) + 1\big) \|\sigma\| \|B^H\|_{0,T,\infty},$$
$$\|Y^\xi\|_{-\tau,T,\theta}$$
$$\leq \bar{C}(T) + \Big\{1 + C_1 T e^{\frac{T\bar{K}_1^+}{2}} \Big(L_1 \Phi(\bar{K}_1, \bar{K}_2, T) + 2 + T\sqrt{\bar{K}_2} C_1\Big)\Big\} \|\sigma\| \|B^H\|_{0,T,\theta}.$$

Therefore, for $T > 0$ such that (4.2.9) holds, it follows from Lemma 2.2.1 that there is some $C > 1$ such that

$$\mathbb{E} \exp\left\{ C \int_0^T \left| K_H^{-1} \left(\int_0^\cdot Z(Y_r^\xi) \mathrm{d}r \right)(s) \right|^2 \mathrm{d}s \right\} < \infty.$$

□

Proof of Theorem 4.1.1

We first show the existence of weak solution to (4.1.1). It follows from (A1)-(A3) and Lemma 4.2.2 that $R^\xi(t)$ is an exponential martingale. Then the Girsanov theorem implies that $\tilde{B}^H(t)$ is a fBm under $\mathbb{Q}^\xi := R^\xi(T)\mathbb{P}$. Reformulating the reference equation (4.1.2) as equation (4.2.6), then under the complete filtration probability $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,T]}, \mathbb{Q}^\xi)$, $(Y^\xi(t), \tilde{B}^H(t))_{t\in[0,T]}$ is a weak solution of (4.1.1).

We shall show the uniqueness of weak solutions to (4.1.1) (see [97, Theorem 2.1] for more proof details). For the rest of this section, we sketch the proof as follows:

For $i=1,2$, let $(Y^{(i),\xi}(t),B_i^H(t))_{t\in[0,T]}$ be two weak solutions to (4.1.1) in the complete filtration probability space $\left(\Omega^{(i)},\{\mathscr{F}_t^{(i)}\}_{t\geq 0},\mathbb{P}^{(i)}\right)$ with $Y_0^{(i),\xi}=\xi$ satisfying (4.1.6). Let $\{B_i(t)\}_{t\in[0,T]}$ be the Brownian motion associated with $\{B_i^H(t)\}_{t\in[0,T]}$. Note that $Y^\xi(\cdot)\in C^\beta([0,T],\mathbb{R}^d)$ for any $\beta\in(0,H)$. Since $\xi\in C^\theta([-\tau,0],\mathbb{R}^d)$, we obtain $Y^\xi_\cdot\in C^{\theta\wedge\beta}([0,T],\mathscr{C})$.

Let $\beta>\frac{2H-1}{2\alpha}$. It yields that $(\theta\wedge\beta)\alpha+\frac{1}{2}-H>0$, which ensures the integrals in (4.2.11) make sense. By (4.2.14), (4.2.15) and (4.2.16), we have

$$\int_0^t\left|K_H^{-1}\left(\int_0^\cdot Z(Y_s^{(i),\xi})\mathrm{d}s\right)\right|^2(r)\mathrm{d}r<\infty,\ t\in[0,T],\ \mathbb{P}\text{-a.s.}$$

Denote by $\mathbb{P}^{(i),\xi}$ the distribution of $Y^{(i),\xi}$. We intend to prove $\mathbb{P}^{(1),\xi}=\mathbb{P}^{(2),\xi}$. To this end, we define for $i=1,2$ that

$$\tau_n^{(i)}=\inf\Big\{t\geq 0:\int_0^t\Big|K_H^{-1}\Big(\int_0^\cdot Z(Y_s^{(i),\xi})\mathrm{d}s\Big)\Big|^2(r)\mathrm{d}r\geq n\Big\}\uparrow\infty,\ \text{as } n\uparrow\infty.$$

For every $i=1,2$ and $n\geq 1$,

$$\begin{aligned}R_n^{(i),\xi}(t):=\exp\Big\{&-\int_0^{t\wedge\tau_n^{(i)}}\left\langle K_H^{-1}\left(\int_0^\cdot Z(Y_s^{(i),\xi})\mathrm{d}s\right)(r),\mathrm{d}B_i(r)\right\rangle\\&-\frac{1}{2}\int_0^{t\wedge\tau_n^{(i)}}\left|K_H^{-1}\left(\int_0^\cdot Z(Y_s^{(i),\xi})\mathrm{d}s\right)\right|^2(r)\mathrm{d}r\Big\}\end{aligned}$$

is a $\mathbb{P}^{(i)}$-martingale. Define the probability measure $\mathbb{Q}_n^{(i),\xi}$ on $\mathscr{F}_\infty^{(i)}$ by letting

$$\mathbb{Q}_n^{(i),\xi}(A)=\mathbb{E}_{\mathbb{P}^{(i)}}[\mathbb{1}_A R_n^{(i),\xi}(T)],\quad T>0,\ A\in\mathscr{F}_T^{(i),\xi}.$$

By the Girsanov theorem and the integral representation of fBm, we can prove that

$$\hat{B}_i^H(t):=B_i^H(t)+\int_0^{t\wedge\tau_n^{(i)}}Z(Y_s^{(i),\xi})\mathrm{d}s,t\in[0,T]$$

is a $\mathbb{Q}_n^{(i)}$-fBm on $\mathbb{R}^m$.

Therefore, under the $\mathbb{Q}_n^{(i)}$, $(Y^{(i),\xi}(t),\hat{B}_i^H(t))_{t\in[0,T\wedge\tau_n^{(i)}]}$ solves (4.1.2) under the $\mathbb{Q}_n^{(i)}$. By the pathwise uniqueness of (4.1.2), the law of $(Y^{(i),\xi}(t),\hat{B}_i^H(t))_{t\in[0,T\wedge\tau_n^{(i)}]}$

under $\mathbb{Q}_n^{(i)}$ coincides with the law of $(Y^\xi(t), B^H(t))_{t\in[0,T\wedge\tau_n^{(i)}]}$ under $\mathbb{P}$. Thus, for any $F\in\mathscr{B}_b(C([0,T];\mathbb{R}^d)\times C([0,T];\mathbb{R}^d))$, we have

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}^{(i)}}\Big[\mathbb{1}_{\{\tau_n^{(i)}\ge T\}}F\Big(Y^{(i),\xi}([0,T]), B_i^H([0,T])\Big)\Big]\\
&=\mathbb{E}_{\mathbb{Q}_n^{(i)}}\Big[\mathbb{1}_{\{\int_0^T|K_H^{-1}(\int_0^\cdot Z(Y_s^{(i),\xi})\mathrm{d}s)|^2(r)\mathrm{d}r\le n\}}\exp\Big\{\int_0^T\Big\langle K_H^{-1}\Big(\int_0^\cdot Z(Y_s^{(i),\xi})\mathrm{d}s\Big)(r),\\
&\qquad\qquad \mathrm{d}B_i(r)\Big\rangle-\frac{1}{2}\int_0^t\Big|K_H^{-1}\Big(\int_0^\cdot Z(Y_s^{(i),\xi})\mathrm{d}s\Big)\Big|^2(r)\mathrm{d}r\Big\}\\
&\quad\times F(Y^{(i),\xi}([0,T]),(\hat{B}_i^H-\int_0^\cdot Z(Y_s^{(i),\xi})\mathrm{d}s)([0,T]))\Big]\\
&=\mathbb{E}_{\mathbb{P}}\Big[\mathbb{1}_{\{\int_0^T|K_H^{-1}(\int_0^\cdot Z(Y_s^{\xi})\mathrm{d}s)|^2(r)\mathrm{d}r\le n\}}\exp\Big\{\int_0^T\Big\langle K_H^{-1}\big(\int_0^\cdot Z(Y_s^{\xi})\mathrm{d}s\big)(r),\mathrm{d}B(r)\Big\rangle\\
&\qquad\qquad -\frac{1}{2}\int_0^t|K_H^{-1}\big(\int_0^\cdot Z(Y_s^{\xi})\mathrm{d}s\big)|^2(r)\mathrm{d}r\Big\}\\
&\quad\times F(Y^\xi([0,T]),(B^H-\int_0^\cdot Z(Y_s^\xi)\mathrm{d}s)([0,T]))\Big],\quad i=1,2.
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}^{(1)}}\Big[\mathbb{1}_{\{\tau_n^{(1)}\ge T\}}F\Big(Y^{(1),\xi}([0,T]),B_1^H([0,T])\Big)\Big]\\
&=\mathbb{E}_{\mathbb{P}^{(2)}}\Big[\mathbb{1}_{\{\tau_n^{(2)}\ge T\}}F\Big(Y^{(2),\xi}([0,T]),B_2^H([0,T])\Big)\Big],
\end{aligned}$$

holds for any $n\ge1$. Letting $n\to\infty$, we obtain

$$\mathbb{E}_{\mathbb{P}^{(1)}}\Big[F\Big(Y^{(1),\xi}([0,T]),B_1^H([0,T])\Big)\Big]=\mathbb{E}_{\mathbb{P}^{(2)}}\Big[F\Big(Y^{(2),\xi}([0,T]),B_2^H([0,T])\Big)\Big].$$

This, together with the arbitrariness of F, yields that $\mathbb{P}^{(1),\xi}=\mathbb{P}^{(2),\xi}$. Thus, the uniqueness of weak solution to (4.1.1) is verified.

4.3 Weak approximation rate of EM scheme

In this section, we shall study the weak convergence of the numerical approximation to (4.1.1). In (4.1.1), σ is a $d\times m$ matrix with $d\ge m$. For $d>m$, this equation is obviously degenerate. In this case, we shall introduce the pseudo-inverse of σ to cover some degenerate models, such as stochastic Hamiltonian systems. Denote by $\mathrm{Ran}(\sigma)$ the range of σ, i.e. $\mathrm{Ran}(\sigma) = \sigma(\mathbb{R}^m)$. If $\mathrm{Ran}(\sigma)$ contains nonzero vectors, then $\sigma\sigma^*$ is a bijective from $\mathrm{Ran}(\sigma)$ onto $\mathrm{Ran}(\sigma)$, whose inverse is denoted by $(\sigma\sigma^*)^{-1}\Big|_{\mathrm{Ran}(\sigma)}$. Let π_* be the orthogonal projection from $\mathbb{R}^d$ to $\mathrm{Ran}(\sigma)$. Then $\mathbb{R}^d$ has the following decomposition:

$$\mathbb{R}^d = \pi_*\mathbb{R}^d \oplus (I_{d\times d} - \pi_*)\mathbb{R}^d \equiv \mathrm{Ran}(\sigma) \oplus (I_{d\times d} - \pi_*)\mathbb{R}^d,$$

where $I_{d\times d}$ is the identity matrix of $\mathbb{R}^d$. We define $\hat{\sigma}^{-1}$, the pseudo-inverse of σ, as follows

$$\hat{\sigma}^{-1}v = \sigma^*\left((\sigma\sigma^*)^{-1}\Big|_{\mathrm{Ran}(\sigma)}\pi_* v\right), \ v \in \mathbb{R}^d.$$

Then $\|\hat{\sigma}^{-1}\| = \left\|(\sigma\sigma^*)^{-1}\Big|_{\mathrm{Ran}(\sigma)}\right\|$. In particular, if σ is of the form $\begin{pmatrix} 0 \\ \sigma_0 \end{pmatrix}$ with σ_0 is an invertible $m\times m$-matrix and 0 is a $(d-m)\times m$ zero matrix, then

$$\hat{\sigma}^{-1} = \left(0^*, \sigma_0^{-1}\right), \qquad \|\hat{\sigma}^{-1}\| = \|\sigma_0^{-1}\|.$$

To obtain the result of numerical approximation, we give stronger assumptions on b and Z as follows:

(H1) (A1) holds and there exists a constant $L_1 > 0$ such that

$$|b(x) - b(y)| \le L_1|x-y|, \ x, y \in \mathbb{R}^d. \tag{4.3.1}$$

Moreover, if $\mathrm{Ran}(\sigma) \neq \mathbb{R}^d$, we also assume that there exist a matrix A on $(I_{d\times d} - \pi_*)(\mathbb{R}^d)$ and a measurable function $b_* : \mathrm{Ran}(\sigma) \to (I_{d\times d} - \pi_*)(\mathbb{R}^d)$ such that

$$(I_{d\times d} - \pi_*)b(x) = A(I_{d\times d} - \pi_*)x + b_*(\pi_* x), \ x \in \mathbb{R}^d.$$

(H2) Z is Hölder continuous with the exponent $\alpha \in (1 - \frac{1}{2H}, 1]$, that is

$$|Z(\xi) - Z(\eta)| \le L_2 \|\xi - \eta\|_\infty^\alpha, \xi, \eta \in \mathscr{C}. \tag{4.3.2}$$

(H3) the initial value $\xi \in \mathscr{C}$ is Hölder continuous with exponent $\theta \in (\frac{2H-1}{2\alpha}, 1]$, that is,

$$|\xi(t) - \xi(s)| \le L_3 |t - s|^\theta, \quad s, t \in [-\tau, 0]. \tag{4.3.3}$$

By these conditions, it follows from Theorem 4.1.1 that (4.1.1) has a unique weak solution with $X_0 = \xi$.

Remark 4.3.1. Since the pseudo-inverse of σ is the inverse of σ if it is invertible, our setting can unify non-degenerate and some degenerate models. A typical example for the equation with $\{0\} \subsetneq \mathrm{Ran}(\sigma) \subsetneq \mathbb{R}^d$ is the following stochastic Hamiltonian system ($d = 2m$):

$$\begin{cases} \mathrm{d}X^{(1)}(t) = X^{(2)}(t)\mathrm{d}t \\ \mathrm{d}X^{(2)}(t) = b_0(X^{(1)}(t), X^{(2)}(t))\mathrm{d}t + Z_0(X_t^{(1)}, X_t^{(2)})\mathrm{d}t + \sigma_0 \mathrm{d}B^H(t), \end{cases}$$

where σ_0 is an invertible $m \times m$-matrix. For any $\eta_1, \eta_2 \in \mathscr{C}, x = (x^{(1)}, x^{(2)}) \in \mathbb{R}^{2m}$, we set

$$b(x^{(1)}, x^{(2)}) = \begin{pmatrix} x^{(2)} \\ b_0(x^{(1)}, x^{(2)}) \end{pmatrix}, Z(\eta_1, \eta_2) = \begin{pmatrix} 0 \\ \sigma_0^{-1} Z_0(\eta_1, \eta_2) \end{pmatrix}, \sigma = \begin{pmatrix} 0 \\ \sigma_0 \end{pmatrix}.$$

Then

$$\mathrm{d}X(t) \equiv \mathrm{d} \begin{pmatrix} X^{(1)}(t) \\ X^{(2)}(t) \end{pmatrix} = (b(X(t)) + \sigma Z(X_t))\, \mathrm{d}t + \sigma \mathrm{d}B^H(t),$$

and in this case, $\pi_*(x^{(1)}, x^{(2)}) = (0, x^{(2)})$, $b_*((0, x^{(2)})) = (x^{(2)}, 0)$ and $A \equiv 0$ in (H1).

We can construct the EM scheme now. Let $\delta \in (0,1)$ be the step-size given by $\delta = \tau/M$ for some $M \in \mathbb{N}$ sufficiently large. The continuous time EM scheme associated with (4.1.1) is defined as below: for $t > 0$,

$$\mathrm{d}X^{(\delta)}(t) = \{(I_{d\times d} - \pi_*)b(X^{(\delta)}(t)) + \pi_* b(X^{(\delta)}(t_\delta)) + \sigma Z(\hat{X}_t^{(\delta)})\}\mathrm{d}t + \sigma \mathrm{d}B^H(t), \tag{4.3.4}$$

with the initial value $X^{(\delta)}(u) = X(u) = \xi(u), u \in [-\tau, 0]$, where $t_\delta := [t/\delta]\delta$, $[t/\delta]$ denotes the integer part of t/δ, and $\hat{X}_t^{(\delta)} \in \mathscr{C}$ is defined as follows

$$\hat{X}_t^{(\delta)}(u) = X^{(\delta)}((t+u) \wedge t_\delta),\ \ u \in [-\tau, 0].$$

For $t \in [0, \delta)$, one gets that

$$\hat{X}_t^{(\delta)}(u) = X^{(\delta)}((t+u) \wedge 0) = \xi((t+u) \wedge 0),$$

and

$$\pi_* X^{(\delta)}(t) = \pi_* X^{(\delta)}(0) + \pi_* b(X^{(\delta)}(0))t + \int_0^t \sigma Z(\hat{X}_s^{(\delta)})\mathrm{d}s + \sigma B^H(t).$$

Then it follows from (H1) that

$$\begin{aligned}
&(I_{d\times d} - \pi_*)X^{(\delta)}(t) \\
&= (I_{d\times d} - \pi_*)X^{(\delta)}(0) + \int_0^t (I_{d\times d} - \pi_*)b(\pi_* X^{(\delta)}(s) + (I_{d\times d} - \pi_*)X^{(\delta)}(s))\mathrm{d}s \\
&= (I_{d\times d} - \pi_*)X^{(\delta)}(0) + \int_0^t A(I_{d\times d} - \pi_*)X^{(\delta)}(s)\mathrm{d}s + \int_0^t b_*(\pi_* X^{(\delta)}(s))\mathrm{d}s,
\end{aligned}$$

which implies that

$$(I_{d\times d} - \pi_*)X^{(\delta)}(t) = \mathrm{e}^{At}(I_{d\times d} - \pi_*)X^{(\delta)}(0) + \int_0^t \mathrm{e}^{A(t-s)}\, b_*(\pi_* X^{(\delta)}(s))\mathrm{d}s.$$

Thus, $X^{(\delta)}(t) = (I_{d\times d} - \pi_*)X^{(\delta)}(t) + \pi_* X^{(\delta)}(t)$ can be obtained explicitly on $[0, \delta]$. By induction, we can get $X^{(\delta)}(t)$ explicitly on $[0, T]$.

Let

$$\bar{K}_1 = 2K_1 + \mathbb{1}_{[K_1\geq 0]} + \frac{|K_1|}{2}\mathbb{1}_{[K_1<0]}, \qquad \bar{K}_2 = \mathbb{1}_{[K_1\geq 0]} + \frac{2}{|K_1|}\mathbb{1}_{[K_1<0]},$$

and

$$\Phi(\bar{K}_1, \bar{K}_2, T) = \sqrt{\frac{\bar{K}_2\left(\mathrm{e}^{\bar{K}_1 T} - 1\right)}{\bar{K}_1}}.$$

Our main result on the weak convergence of EM scheme to (4.1.1) is stated as follows.

Theorem 4.3.1. *Assume (H1)-(H3) and* $\mathrm{Ran}(\sigma) \neq \{0\}$. *For* $\delta \in (0,1)$ *if* T *satisfies*

$$\begin{aligned}
&\frac{2L_1^2 T^{2(H-\beta)+1}\|\hat{\sigma}^{-1}\|^2\|\sigma\|^2\delta^{2\beta}}{\pi(1-H)}\Big\{3\left[1+C_0(H-\frac{1}{2})\right]^2 T^{1-2H} \\
&\quad + 2\delta^{1-2H}\Big[\frac{4(1-H)+4^H}{1-H} + \frac{\pi\delta(1-H)}{6T}\Big]\Big\} \\
&\quad + \frac{2L_2^2}{\pi}\Big\{\frac{3[1+C_0(H-\frac{1}{2})]^2 T^{2(1+\alpha\beta-\beta)}\|\sigma\|^{2\alpha}}{(1-H)} \\
&\quad + T^{2(H-\beta+\alpha\beta-\alpha(\beta\wedge\theta))}\|\sigma\|^{2\alpha}\Big[\frac{\mathcal{B}^2(\frac{1}{2},\frac{1}{2}+\alpha(\beta\wedge\theta)-H)T^{2\alpha(\beta\wedge\theta)+1-2H}}{2\alpha(\beta\wedge\theta)+3-4H} \\
&\quad + \delta^{2\alpha(\beta\wedge\theta)+1-2H}\frac{16^H(1-H)+4^H}{(1-H)^2}\Big]\Big\}\mathbf{1}_{[\alpha=1]} \\
&< 2^{-9}\{1+L_1T(L_1\Phi(\bar{K}_1,\bar{K}_2,T)+1)\}^{-2},
\end{aligned} \tag{4.3.5}$$

where $\beta \in (\frac{2(\alpha+1)H-1}{4\alpha}, H)$ *and* $C_0 = \int_0^1 \frac{u^{\frac{1}{2}-H}-1}{(1-u)^{\frac{1}{2}+H}}\mathrm{d}u$, *then for any bounded measurable function* f *on* $\mathbb{R}^d$, *there exists a constant* C_T *which is independent of* δ *such that for* $t \in [0,T]$

$$|\mathbb{E}f(X(t)) - \mathbb{E}f(X^{(\delta)}(t))| \leq C_T\delta^{\alpha(\beta\wedge\theta)+\frac{1}{2}-H}. \tag{4.3.6}$$

Remark 4.3.2. The convergence result only holds for $t \in [0,T]$ and T satisfies (4.3.5). Letting $\delta \to 0$, (4.3.5) converges to

$$\frac{2L_2^2}{\pi}\Big\{\frac{3[1+C_0(H-\frac{1}{2})]^2 T^{2(1+\alpha\beta-\beta)}\|\sigma\|^{2\alpha}}{(1-H)}$$

$$+\frac{\mathcal{B}^2(\frac{1}{2},\frac{1}{2}+\alpha(\beta\wedge\theta)-H)\|\sigma\|^{2\alpha}T^{1+2\alpha\beta-2\beta}}{2\alpha(\beta\wedge\theta)+3-4H}\Big\}\mathbf{1}_{[\alpha=1]}$$
$$< 2^{-9}\{1+L_1T(L_1\Phi(\bar K_1,\bar K_2,T)+1)\}^{-2}.$$

It is easy to see that for any fixed $\delta\in(0,1)$, there always exists $T>0$ such that (4.3.5) holds.

4.4 Proof of the main theorem

Before giving the proof for Theorem 4.3.1, we prepare two lemmas. The lemma below shows the estimates of $(Y^\xi(t))_{t\in[0,T]}$, the solution to (4.1.2) in the sense of uniform and Hölder norms, respectively.

Lemma 4.4.1. *Assume (H1). Then for any $T>0$*

$$\begin{aligned}
&\|Y^\xi\|_{-\tau,T,\infty} \qquad (4.4.1)\\
&\le \|\xi\|_\infty+|b(0)|\Phi(\bar K_1,\bar K_2,T)+\big(L_1\Phi(\bar K_1,\bar K_2,T)+1\big)\|\sigma\|\|B^H\|_{0,T,\infty}.\\
&\|Y^\xi\|_{-\tau,T,\beta\wedge\theta}\\
&\le T^{1-\beta\wedge\theta}\big(|b(0)|+|b(0)|L_1\Phi(\bar K_1,\bar K_2,T)+|\xi(0)|\big)+\|\sigma\|\|B^H\|_{0,T,\beta\wedge\theta}\\
&\quad +L_1T^{1-\beta\wedge\theta}\big(L_1\Phi(\bar K_1,\bar K_2,T)+1\big)\|\sigma\|\|B^H\|_{0,T,\infty}+\|\xi\|_{-\tau,0,\beta\wedge\theta}.
\end{aligned}$$

Proof. The first inequality follows from (4.2.1) and (H1) directly. Since b is Lipschitz continuous, we have

$$|b(x)|\le|b(0)|+L_1|x|.$$

Taking into account the following inequality

$$\|Y^\xi\|_{-\tau,T,\beta\wedge\theta}\le\|\xi\|_{-\tau,0,\beta\wedge\theta}+\|Y^\xi\|_{0,T,\beta\wedge\theta},$$

the proof of the second inequality is similar to the second part of the proof of Lemma 4.2.1. □

For the sake of simplicity, we denote

$$h^{\xi}(t) = \hat{\sigma}^{-1}\{b(Y^{\xi}(t)) - b(Y^{\xi}(t_{\delta})\} - Z(\hat{Y}_t^{\xi}), \quad t \geq 0,$$

with

$$\hat{Y}_t^{\xi}(u) = Y^{\xi}((t+u) \wedge t_{\delta}), \quad u \in [-\tau, 0].$$

Let

$$\begin{aligned} B_h^H(t) &= B^H(t) + \int_0^t h^{\xi}(s)\mathrm{d}s, \\ R^{\xi,\delta}(t) &= \exp\Big\{ - \int_0^t \Big\langle K_H^{-1}\big(\int_0^{\cdot} h^{\xi}(s)\mathrm{d}s\big)(r), \mathrm{d}B(r)\Big\rangle \\ &\qquad - \frac{1}{2}\int_0^t \Big|K_H^{-1}\big(\int_0^{\cdot} h^{\xi}(s)\mathrm{d}s\big)(r)\Big|^2 \mathrm{d}r\Big\}, t \in [0, T], \end{aligned} \tag{4.4.2}$$

and $\mathrm{d}\mathbb{Q}^{\xi,\delta} = R^{\xi,\delta}(T)\mathrm{d}\mathbb{P}$. Then it follows from Lemma 4.4.2 below and the Girsanov theorem that $\mathbb{Q}^{\xi,\delta}$ is a probability and $(B_h^H(t))_{t\in[0,T]}$ is a fBm under $\mathbb{Q}^{\xi,\delta}$. Since $\sigma\sigma^{-1} = \pi_*$, we can rewrite the reference SDE (4.1.2) into the following form:

$$\mathrm{d}Y^{\xi}(t) = \{(I_{d\times d} - \pi_*)b(Y^{\xi}(t)) + \pi_* b(Y^{\xi}(t_{\delta})) + \sigma Z(\hat{Y}_t^{\xi})\}\mathrm{d}t + \sigma \mathrm{d}B_h^H(t), \tag{4.4.3}$$

which implies that $(Y^{\xi}(t), B_h^H(t))_{t\in[0,T]}$ is a weak solution of (4.3.4). This, together with the pathwise uniqueness of solution to (4.3.4), yields the weak uniqueness. Then, we have

$$\begin{aligned} |\mathbb{E}f(X(t)) - \mathbb{E}f(X^{(\delta)}(t))| &= |\mathbb{E}_{\mathbb{Q}^{\xi}} f(Y^{\xi}(t)) - \mathbb{E}_{\mathbb{Q}^{\xi,\delta}} f(Y^{\xi}(t))| \\ &= |\mathbb{E}(R^{\xi}(t) - R^{\xi,\delta}(t))f(Y^{\xi}(t))|. \end{aligned}$$

Hence, in the following discussion, we shall prove that $\{R^{\xi,\delta}(t)\}_{t\in[0,T]}$ is an exponential martingale, and give estimates of $R^{\xi}(t) - R^{\xi,\delta}(t)$.

Lemma 4.4.2. *Under the assumptions of Theorem 4.3.1, we have*

$$\mathbb{E}\exp\left\{C\int_0^t |K_H^{-1}\big(\int_0^\cdot h^\xi(s)\mathrm{d}s\big)|^2(r)\mathrm{d}r\right\} < \infty,$$

for some $C > 1$.

Proof. The definition of inverse operator K_H^{-1} (2.2.3) yields that

$$\begin{aligned}
&K_H^{-1}\Big(\int_0^\cdot h^\xi(s)\mathrm{d}s\Big)(r) = r^{H-\frac{1}{2}} D_{0+}^{H-\frac{1}{2}}[\cdot^{\frac{1}{2}-H} h^\xi(\cdot)](r)\\
&= r^{H-\frac{1}{2}}\frac{1}{\Gamma(\frac{3}{2}-H)}\Big(\frac{r^{\frac{1}{2}-H}h^\xi(r)}{r^{H-\frac{1}{2}}} + (H-\frac{1}{2})\int_0^r \frac{r^{\frac{1}{2}-H}h^\xi(r) - s^{\frac{1}{2}-H}h^\xi(s)}{(r-s)^{H+\frac{1}{2}}}\mathrm{d}s\Big)\\
&= \frac{r^{\frac{1}{2}-H}h^\xi(r)}{\Gamma(\frac{3}{2}-H)} + \frac{r^{H-\frac{1}{2}}}{\Gamma(\frac{3}{2}-H)}(H-\frac{1}{2})\int_0^r \frac{r^{\frac{1}{2}-H}h^\xi(r) - s^{\frac{1}{2}-H}h^\xi(s)}{(r-s)^{H+\frac{1}{2}}}\mathrm{d}s\\
&= \frac{r^{\frac{1}{2}-H}h^\xi(r)}{\Gamma(\frac{3}{2}-H)} + \frac{r^{H-\frac{1}{2}}}{\Gamma(\frac{3}{2}-H)}(H-\frac{1}{2})\int_0^r \frac{(r^{\frac{1}{2}-H} - s^{\frac{1}{2}-H})h^\xi(r)}{(r-s)^{H+\frac{1}{2}}}\mathrm{d}s\\
&\quad + \frac{r^{H-\frac{1}{2}}}{\Gamma(\frac{3}{2}-H)}(H-\frac{1}{2})\int_0^r \frac{s^{\frac{1}{2}-H}(h^\xi(r) - h^\xi(s))}{(r-s)^{H+\frac{1}{2}}}\mathrm{d}s\\
&= [1 + C_0(H-\frac{1}{2})]\frac{r^{\frac{1}{2}-H}h^\xi(r)}{\Gamma(\frac{3}{2}-H)} + \frac{r^{H-\frac{1}{2}}}{\Gamma(\frac{3}{2}-H)}(H-\frac{1}{2})\int_0^r \frac{s^{\frac{1}{2}-H}(h^\xi(r) - h^\xi(s))}{(r-s)^{H+\frac{1}{2}}}\mathrm{d}s\\
&=: \hat{J}_1(r) + \hat{J}_2(r). \qquad (4.4.4)
\end{aligned}$$

For $\hat{J}_1$, it follows from (H1) and (H2) that

$$\begin{aligned}
|h^\xi(r)| &\le \|\hat{\sigma}^{-1}\||b(Y^\xi(r)) - b(Y^\xi(r_\delta))| + |Z(\hat{Y}_r^\xi)|\\
&\le \|\hat{\sigma}^{-1}\|L_1\|Y^\xi\|_{0,r,\beta}\delta^\beta + |Z(0)| + L_2\|Y^\xi\|_{-\tau,r,\infty}^\alpha.
\end{aligned}$$

For $\hat{J}_2$, note that

$$\begin{aligned}
&\int_0^r \frac{s^{\frac{1}{2}-H}|h^\xi(r) - h^\xi(s)|}{(r-s)^{H+\frac{1}{2}}}\mathrm{d}s\\
&= \int_0^r s^{\frac{1}{2}-H}\Big|\frac{\hat{\sigma}^{-1}(b(Y^\xi(r)) - b(Y^\xi(r_\delta))) - Z(\hat{Y}_r^\xi))}{(r-s)^{H+\frac{1}{2}}}\\
&\qquad - \frac{\hat{\sigma}^{-1}(b(Y^\xi(s)) - b(Y^\xi(s_\delta))) + Z(\hat{Y}_s^\xi))}{(r-s)^{H+\frac{1}{2}}}\Big|\mathrm{d}s
\end{aligned}$$

$$\leq \int_0^r \frac{\|\hat{\sigma}^{-1}\| s^{\frac{1}{2}-H} |b(Y^\xi(r)) - b(Y^\xi(r_\delta)) - (b(Y^\xi(s)) - b(Y^\xi(s_\delta)))|}{(r-s)^{H+\frac{1}{2}}} \mathrm{d}s$$
$$+ \int_0^r \frac{s^{\frac{1}{2}-H} |Z(\hat{Y}_r^\xi) - Z(\hat{Y}_s^\xi))|}{(r-s)^{H+\frac{1}{2}}} \mathrm{d}s$$
$$=: I_1(r) + I_2(r).$$

Next, we shall give the estimations of $I_i(r), i = 1, 2$, respectively. For $I_1(r)$, it follows from (H1) that

$$\begin{aligned}
&|b(Y^\xi(r)) - b(Y^\xi(r_\delta)) - (b(Y^\xi(s)) - b(Y^\xi(s_\delta)))| \\
&\leq 2L_1 \|Y^\xi\|_{0,r,\beta} \Big[\delta^\beta \wedge \frac{(r-s)^\beta + (r_\delta - s_\delta)^\beta}{2}\Big] \\
&= L_1 \|Y^\xi\|_{0,r,\beta} \begin{cases} (r-s)^\beta, & r_\delta < s < r, \\ (r-s)^\beta + (r_\delta - s_\delta)^\beta, & r-\delta < s < r_\delta, \\ 2\delta^\beta, & 0 < s < r-\delta. \end{cases}
\end{aligned}$$

Since

$$|r_\delta - s_\delta| = |[\frac{r}{\delta}]\delta - [\frac{s}{\delta}]\delta| \leq |[\frac{r}{\delta}]\delta - [\frac{r-\delta}{\delta}]\delta| \leq \delta, \quad r-\delta < s < r_\delta,$$

and for $r \geq \delta$, one gets that

$$\begin{aligned}
&\int_{r_\delta}^r \frac{s^{1/2-H}}{(r-s)^{H+1/2-\beta}} \mathrm{d}s \leq \frac{2\delta^{1/2+\beta-H}}{1+2\beta-2H} r_\delta^{1/2-H}, \\
&\int_0^{r_\delta} \frac{2\delta^\beta}{(r-s)^{H+1/2} s^{H-1/2}} \mathrm{d}s \\
&= \int_{r_\delta/2}^{r_\delta} \frac{2\delta^\beta}{(r-s)^{1/2+H} s^{H-1/2}} \mathrm{d}s + \int_0^{r_\delta/2} \frac{2\delta^\beta}{(r-s)^{1/2+H} s^{H-1/2}} \mathrm{d}s \\
&\leq \frac{2\delta^\beta (r-r_\delta)^{\frac{1}{2}-H}}{(r_\delta/2)^{H-1/2}(H-1/2)} + \frac{2\delta^\beta \left(\frac{r_\delta}{2}\right)^{\frac{3}{2}-H}}{(r-r_\delta/2)^{H+1/2}\left(\frac{3}{2}-H\right)},
\end{aligned}$$

Taking the above inequalities into account, we obtain

$|I_1(r)|$

$$\leq 2L_1\|\hat{\sigma}^{-1}\|\|Y^\xi\|_{0,r,\beta}\Bigg\{\Bigg[\frac{\delta^{\beta+1/2-H}}{(2\beta+1-2H)r_\delta^{H-1/2}}+\frac{\delta^\beta(r-r_\delta)^{\frac{1}{2}-H}}{(r_\delta/2)^{H-1/2}(H-1/2)}$$
$$+\frac{\delta^\beta\left(\frac{r_\delta}{2}\right)^{\frac{3}{2}-H}}{(r-r_\delta/2)^{H+1/2}\left(\frac{3}{2}-H\right)}\Bigg]\mathbb{1}_{[r\geq\delta]}+\frac{1}{2}\mathcal{B}(\frac{3}{2}-H,\beta+\frac{1}{2}-H)r^{\beta+1-2H}\mathbb{1}_{[0\leq r<\delta]}\Bigg\}.$$

We now calculate $I_2(r)$. One can see that

$$\|\hat{Y}_r^\xi-\hat{Y}_s^\xi\|_\infty^\alpha$$
$$=\sup_{-\tau\leq u\leq 0}\frac{|Y^\xi((r+u)\wedge r_\delta)-Y^\xi((s+u)\wedge s_\delta)|^\alpha}{|(r+u)\wedge r_\delta-(s+u)\wedge s_\delta|^{\alpha(\beta\wedge\theta)}}|(r+u)\wedge r_\delta-(s+u)\wedge s_\delta|^{\alpha(\beta\wedge\theta)}$$
$$\leq\|Y^\xi\|_{-\tau,r,\beta\wedge\theta}^\alpha\sup_{-\tau\leq u\leq 0}|(r+u)\wedge r_\delta-(s+u)\wedge s_\delta|^{\alpha(\beta\wedge\theta)}.$$

Since for $s+u>s_\delta$ and $r+u<r_\delta$, we have

$$(s+u)\wedge s_\delta=s_\delta;\qquad (r+u)\wedge r_\delta=r+u;\qquad s_\delta-s<u<r_\delta-r.$$

Then, it yields that

$$\sup_{s_\delta-s<u<r_\delta-r}|(r+u)\wedge r_\delta-(s+u)\wedge s_\delta|=\sup_{s_\delta-s<u<r_\delta-r}|r+u-s_\delta|=|r_\delta-s_\delta|.$$

Similarly, for $s+u<s_\delta$ and $r+u>r_\delta$, we have

$$\sup_{r_\delta-r<u<s_\delta-s}|(r+u)\wedge r_\delta-(s+u)\wedge s_\delta|=|r_\delta-s_\delta|.$$

Then it is easy to see that

$$\sup_{u\in[-\tau,0]}|(r+u)\wedge r_\delta-(s+u)\wedge s_\delta|=(r-s)\vee(r_\delta-s_\delta).$$

Consequently,

$$\|\hat{Y}_r^\xi-\hat{Y}_s^\xi\|_\infty^\alpha\leq\|Y^\xi\|_{-\tau,r,\beta\wedge\theta}^\alpha((r-s)\vee(r_\delta-s_\delta))^{\alpha(\beta\wedge\theta)},$$

and, it implies that

$$I_2(r)=\int_0^r\frac{s^{\frac{1}{2}-H}|Z(\hat{Y}_r^\xi)-Z(\hat{Y}_s^\xi))|}{(r-s)^{H+\frac{1}{2}}}\mathrm{d}s$$

$$\leq \int_0^r \frac{L_2\|\hat{Y}_r^\xi - \hat{Y}_s^\xi\|_\infty^\alpha}{(r-s)^{1/2+H}s^{H-1/2}}\mathrm{d}s$$
$$\leq L_2\|Y\|_{-\tau,r,\beta\wedge\theta}^\alpha \int_0^r \frac{(r-s)^{\alpha(\beta\wedge\theta)} \vee (r_\delta - s_\delta)^{\alpha(\beta\wedge\theta)}}{(r-s)^{1/2+H}s^{H-1/2}}\mathrm{d}s.$$

Since $r_\delta - s_\delta = 0$ for $s \in [r_\delta, r]$,

$$\int_0^r \frac{(r-s)^{\alpha(\beta\wedge\theta)} \vee (r_\delta - s_\delta)^{\alpha(\beta\wedge\theta)}}{(r-s)^{1/2+H}s^{H-1/2}}\mathrm{d}s$$
$$= \int_0^r \frac{(r-s)^{\alpha(\beta\wedge\theta)}}{(r-s)^{1/2+H}s^{H-1/2}}\mathbf{1}_{[r-s\geq r_\delta - s_\delta]}\mathrm{d}s$$
$$+ \int_0^{r_\delta} \frac{(r_\delta - s_\delta)^{\alpha(\beta\wedge\theta)}}{(r-s)^{1/2+H}s^{H-1/2}}\mathbf{1}_{[r-s< r_\delta - s_\delta]}\mathrm{d}s.$$

For $r - r_\delta + s_\delta < s$, it is clear that $r_\delta - s_\delta - (r-s) \leq \delta$, so

$$(r_\delta - s_\delta)^{\alpha(\beta\wedge\theta)} = (r_\delta - s_\delta - r + s + (r-s))^{\alpha(\beta\wedge\theta)} \leq (r-s)^{\alpha(\beta\wedge\theta)} + \delta^{\alpha(\beta\wedge\theta)},$$

which implies that

$$\int_0^{r_\delta} \frac{(r_\delta - s_\delta)^{\alpha(\beta\wedge\theta)}}{(r-s)^{1/2+H}s^{H-1/2}}\mathbf{1}_{[r-s< r_\delta - s_\delta]}\mathrm{d}s$$
$$\leq \int_0^{r_\delta} \frac{(r-s)^{\alpha(\beta\wedge\theta)} + \delta^{\alpha(\beta\wedge\theta)}}{(r-s)^{1/2+H}s^{H-1/2}}\mathbf{1}_{[r-s< r_\delta - s_\delta]}\mathrm{d}s.$$

Moreover, we have

$$\int_0^T \left(\int_0^{r_\delta} \frac{\delta^{\alpha(\beta\wedge\theta)}\mathbf{1}_{[r-s< r_\delta - s_\delta]}}{(r-s)^{1/2+H}s^{H-1/2}}\mathrm{d}s\right)^2 \mathrm{d}r \tag{4.4.5}$$
$$\leq \int_\delta^T \left(\int_0^{r_\delta} \frac{\delta^{\alpha(\beta\wedge\theta)}}{(r-s)^{1/2+H}s^{H-1/2}}\mathrm{d}s\right)^2 \mathrm{d}r$$
$$\leq \sum_{k=1}^{N-1}\int_{k\delta}^{(k+1)\delta} \left(\frac{\delta^{\alpha(\beta\wedge\theta)}\left(\frac{r_\delta}{2}\right)^{\frac{3}{2}-H}}{(r-r_\delta/2)^{H+1/2}\left(\frac{3}{2}-H\right)} + \frac{\delta^{\alpha(\beta\wedge\theta)}(r-r_\delta)^{\frac{1}{2}-H}}{(r_\delta/2)^{H-1/2}(H-1/2)}\right)^2 \mathrm{d}r$$
$$\leq 2\delta^{2\alpha(\beta\wedge\theta)}\sum_{k=1}^{N-1}\left(\frac{16^H(k\delta)^{2-4H}\delta}{(3-2H)^2} + \frac{2^{2H-1}\delta^{2-2H}}{(H-1/2)^2(2-2H)(k\delta)^{2H-1}}\right)$$
$$\leq 2\delta^{2\alpha(\beta\wedge\theta)+1-2H}\left(\frac{16^H}{(3-2H)^2} + \frac{2^{2H}}{(2H-1)^2(1-H)}\right)\sum_{k=1}^{N-1}(k\delta)^{1-2H}\delta$$

$$\leq \frac{\delta^{2\alpha(\beta\wedge\theta)+1-2H}}{1-H}\left(\frac{16^H}{(3-2H)^2}+\frac{2^{2H}}{(2H-1)^2(1-H)}\right)T^{2-2H}, \tag{4.4.6}$$

and

$$\int_0^T\left(\int_0^r \frac{(r-s)^{\alpha(\beta\wedge\theta)}}{(r-s)^{1/2+H}s^{H-1/2}}\mathrm{d}s\right)^2\mathrm{d}r \tag{4.4.7}$$
$$=\frac{T^{2\alpha(\beta\wedge\theta)+3-4H}\mathcal{B}^2(\frac{3}{2}-H,\alpha(\beta\wedge\theta)+\frac{1}{2}-H)}{2\alpha(\beta\wedge\theta)+3-4H}.$$

Substituting $\hat{J}_1$, $I_1(r)$ and $I_2(r)$ into (4.4.4), and taking into account (4.4.5) and (4.4.7), we arrive at

$$\begin{aligned}
&\int_0^T\left|K_H^{-1}\Big(\int_0^\cdot h^\xi(s)\mathrm{d}s\Big)(r)\right|^2\mathrm{d}r\\
&\leq \frac{2L_1^2\|\hat{\sigma}^{-1}\|^2\|Y^\xi\|^2_{0,T,\beta}}{\Gamma^2(\frac{3}{2}-H)(1-H)}\Big\{3\left[1+C_0(H-\frac{1}{2})\right]^2T^{2-2H}\delta^{2\beta}\\
&+2(2H-1)^2T\delta^{2\beta+1-2H}\Big[\frac{1}{(1+2\beta-2H)^2}+\frac{2^{4H-1}}{(3-2H)^2}+\frac{2^{2H-1}}{(2H-1)^2(1-H)}\\
&+\frac{\mathcal{B}^2(\frac{3}{2}-H,\beta+\frac{1}{2}-H)(1-H)\delta}{8(1+\beta-H)T}\Big]\Big\}\\
&+\frac{6[1+C_0(H-\frac{1}{2})]^2T^{2(1-H)}}{\Gamma^2(\frac{3}{2}-H)(1-H)}\Big(|Z(0)|^2+L_2^2\|Y^\xi\|^{2\alpha}_{-\tau,T,\infty}\Big)\\
&+\frac{2L_2^2(2H-1)^2T^{2H-1}\|Y^\xi\|^{2\alpha}_{-\tau,T,\beta\wedge\theta}}{\Gamma^2(\frac{3}{2}-H)}\Big[\frac{\mathcal{B}^2(\frac{3}{2}-H,\alpha(\beta\wedge\theta)+1/2-H)T^{2\alpha(\beta\wedge\theta)+3-4H}}{2\alpha(\beta\wedge\theta)+3-4H}\\
&+\frac{\delta^{2\alpha(\beta\wedge\theta)+1-2H}T^{2-2H}}{1-H}\Big(\frac{16^H}{(3-2H)^2}+\frac{2^{2H}}{(2H-1)^2(1-H)}\Big)\Big].
\end{aligned} \tag{4.4.8}$$

Since $\Gamma(\frac{3}{2}-H)\geq\Gamma(\frac{1}{2})=\sqrt{\pi}$ and $\beta\in(\frac{2(\alpha+1)H-1}{4\alpha},H)$, it follows from Lemma 4.4.1 and (4.3.5) that there exists $C>1$ such that

$$\mathbb{E}\exp\Big\{C\int_0^T\left|K_H^{-1}\Big(\int_0^\cdot h^\xi(s)\mathrm{d}s\Big)(r)\right|^2\mathrm{d}r\Big\}<\infty.$$

□

We are now in the position to complete the **Proof of Theorem 4.3.1**.

Proof. Let

$$M_1(t) = \int_0^t \Big\langle K_H^{-1}\Big(\int_0^\cdot Z(Y_s^\xi)\mathrm{d}s\Big)(r), \mathrm{d}B(r)\Big\rangle,$$
$$M_2(t) = \int_0^t \Big\langle K_H^{-1}\Big(\int_0^\cdot h^\xi(s)\mathrm{d}s\Big)(r), dB(r)\Big\rangle, \quad t \geq 0.$$

By the weak uniqueness of solution to (4.1.1), the Hölder inequality and the following inequality

$$|\,\mathrm{e}^x - \mathrm{e}^y\,| \leq (\mathrm{e}^x \vee \mathrm{e}^y)|x-y|,$$

we then have that $\forall f \in \mathscr{B}_b(\mathbb{R}^d)$,

$$\begin{aligned}
|\mathbb{E}f(X(t)) - \mathbb{E}f(X^{(\delta)}(t))| &= |\mathbb{E}_{\mathbb{Q}^\xi} f(Y^\xi(t)) - \mathbb{E}_{\mathbb{Q}^{\xi,\delta}} f(Y^\xi(t))| \\
&= |\mathbb{E}(R^\xi(t) - R^{\xi,\delta}(t)) f(Y^\xi(t))| \\
&\leq \|f\|_\infty \mathbb{E}|R^\xi(t) - R^{\xi,\delta}(t)| \\
&\leq \|f\|_\infty \mathbb{E}\left(R^\xi(t) \vee R^{\xi,\delta}(t)\right)\left|\log R^\xi(t) - \log R^{\xi,\delta}(t)\right| \\
&\leq \|f\|_\infty \Theta_1(t)(\Theta_2(t)+\Theta_3(t)), t \in [0,T], \qquad (4.4.9)
\end{aligned}$$

where

$$\Theta_1(t) = \left(\mathbb{E}(R^\xi(t))^q\right)^{\frac{1}{q}} + \left(\mathbb{E}(R^{\xi,\delta}(t))^q\right)^{\frac{1}{q}},$$

$$\Theta_2(t) = \left(\mathbb{E}\left|\int_0^t \langle K_H^{-1}\left(\int_0^\cdot (Z(Y_s^\xi) + h^\xi(s))\mathrm{d}s\right)(r), \mathrm{d}B(r)\rangle\right|^{\frac{q}{q-1}}\right)^{\frac{q-1}{q}},$$

$\Theta_3(t)$

$$= \frac{1}{2}\left(\mathbb{E}\left|\int_0^t \left(\left|K_H^{-1}\left(\int_0^\cdot Z(Y_s^\xi)\mathrm{d}s\right)(r)\right|^2 - \left|K_H^{-1}\left(\int_0^\cdot h^\xi(s)\mathrm{d}s\right)(r)\right|^2\right)\mathrm{d}r\right|^{\frac{q}{q-1}}\right)^{\frac{q-1}{q}},$$

where the parameter $q > 1$.

It follows from Lemma 4.2.2 with $C_5 = 0$ and $C_4 = L_2$ that there is some $C > 1$ such that $\mathbb{E}\exp\{C\langle M_1\rangle(T)\} < \infty$. Thus, for $2q^2 - q \leq C$, we have

$$\mathbb{E}(R^\xi(t))^q = \mathbb{E}\exp\Big(qM_1(t) - q^2\langle M_1\rangle(t) + (q^2 - q/2)\langle M_1\rangle(t)\Big)$$

$$
\begin{aligned}
&\leq (\mathbb{E}\exp(2qM_1(t) - 2q^2\langle M_1\rangle(t)))^{1/2}\big(\mathbb{E}\exp((2q^2 - q)\langle M_1\rangle(t))\big)^{1/2}\\
&\leq \Big(\mathbb{E}\exp\Big((2q^2 - q)\int_0^t \Big|K_H^{-1}\Big(\int_0^\cdot Z(Y_s^\xi)\mathrm{d}s\Big)(r)\Big|^2\mathrm{d}r\Big)\Big)^{1/2}\\
&< \infty.
\end{aligned}
$$

Similarly, following from Lemma 4.4.2, there is $q > 1$ such that

$$
\sup_{t\in[0,T]} \big(\mathbb{E}(R^{\xi,\delta}(t))^q\big)^{\frac{1}{q}} < \infty.
$$

Hence, there is $q > 1$ and some constant C_T such that

$$
\Theta_1(t) \leq C_T. \tag{4.4.10}
$$

In the following proof, we fix some $q > 1$ such that (4.4.10) holds.

It is easy to see that

$$
\begin{aligned}
&K_H^{-1}\Big(\int_0^\cdot (Z(Y_s^\xi) + h^\xi(s))\mathrm{d}s\Big)(r) = r^{H-\frac{1}{2}}D_{0+}^{H-\frac{1}{2}}[\cdot^{\frac{1}{2}-H}(Z(Y_\cdot^\xi) + h^\xi(\cdot))](r)\\
&= \frac{r^{H-\frac{1}{2}}}{\Gamma(\frac{3}{2}-H)}\Big(\frac{r^{1/2-H}(Z(Y_r^\xi) + h^\xi(r))}{r^{H-1/2}}\\
&\qquad + (H-\frac{1}{2})\int_0^r \frac{r^{1/2-H}(Z(Y_r^\xi) + h^\xi(r)) - s^{1/2-H}(Z(Y_s^\xi) + h^\xi(s))}{(r-s)^{H+1/2}}\mathrm{d}s\Big)\\
&\leq [1 + C_0(H-\frac{1}{2})]\frac{r^{\frac{1}{2}-H}(Z(Y_r^\xi) + h^\xi(r))}{\Gamma(\frac{3}{2}-H)}\\
&\qquad + \frac{r^{H-\frac{1}{2}}}{\Gamma(\frac{3}{2}-H)}(H-\frac{1}{2})\int_0^r \frac{s^{\frac{1}{2}-H}(Z(Y_r^\xi) + h^\xi(r) - Z(Y_s^\xi) - h^\xi(s))}{(r-s)^{H+\frac{1}{2}}}\mathrm{d}s\\
&=: I_3(r) + I_4(r).
\end{aligned} \tag{4.4.11}
$$

Next, we give the estimates for $I_i(r), i = 3, 4$, respectively.

For $I_3(r)$, by (H1) and (H2), it yields that

$$
\begin{aligned}
&|I_3(r)| \\
&\leq [1 + C_0(H-\frac{1}{2})]\frac{r^{\frac{1}{2}-H}}{\Gamma(\frac{3}{2}-H)}(\|\hat{\sigma}^{-1}\||b(Y^\xi(r)) - b(Y^\xi(r_\delta))| + |Z(Y_r^\xi) - Z(\hat{Y}_r^\xi)|)
\end{aligned} \tag{4.4.12}
$$

$$
\begin{aligned}
&\leq [1+C_0(H-\frac{1}{2})]\frac{r^{\frac{1}{2}-H}}{\Gamma(\frac{3}{2}-H)}(L_1\|\hat{\sigma}^{-1}\|\|Y^\xi\|_{0,r,\beta}\delta^\beta + L_2\|Y_r^\xi - \hat{Y}_r^\xi\|_\infty^\alpha)\\
&\leq \frac{\left[1+C_0(H-\frac{1}{2})\right]r^{\frac{1}{2}-H}}{\Gamma(\frac{3}{2}-H)}\left(L_1\|\hat{\sigma}^{-1}\|\|Y^\xi\|_{0,r,\beta}\delta^\beta + L_2\|Y^\xi\|_{0,r,\beta}^\alpha\delta^{\alpha\beta}\right). \quad (4.4.13)
\end{aligned}
$$

For $I_4(r)$, it yields from the definition of h^ξ that

$$
\begin{aligned}
|I_4(r)| &\leq \frac{r^{H-\frac{1}{2}}}{\Gamma(\frac{3}{2}-H)}(H-\frac{1}{2})\int_0^r \frac{s^{\frac{1}{2}-H}(Z(Y_r^\xi)+h^\xi(r)-Z(Y_s^\xi)-h^\xi(s))}{(r-s)^{H+\frac{1}{2}}}\mathrm{d}s\\
&\leq \frac{r^{H-\frac{1}{2}}}{\Gamma(\frac{3}{2}-H)}(H-\frac{1}{2})\|\hat{\sigma}^{-1}\|\int_0^r s^{1/2-H}\Big|\frac{b(Y^\xi(r))-b(Y^\xi(r_\delta))}{(r-s)^{H+1/2}}\\
&\qquad\qquad - \frac{b(Y^\xi(s))-b(Y^\xi(s_\delta))}{(r-s)^{H+1/2}}\Big|\mathrm{d}s\\
&\quad + \frac{r^{H-\frac{1}{2}}}{\Gamma(\frac{3}{2}-H)}(H-\frac{1}{2})\int_0^r \frac{s^{1/2-H}|Z(Y_r^\xi)-Z(\hat{Y}_r^\xi)-(Z(Y_s^\xi)-Z(\hat{Y}_s^\xi))|}{(r-s)^{H+1/2}}\mathrm{d}s\\
&= I_{41}(r)+I_{42}(r). \quad (4.4.14)
\end{aligned}
$$

In the same way of estimating I_1 in the proof of Lemma 4.4.2, we have

$$
\begin{aligned}
I_{41}(r) &\leq 2\frac{r^{H-\frac{1}{2}}}{\Gamma(\frac{3}{2}-H)}(H-\frac{1}{2})L_1\|\hat{\sigma}^{-1}\|\|Y^\xi\|_{0,r,\beta}\Big\{\Big[\frac{\delta^{\beta+1/2-H}}{(2\beta+1-2H)r_\delta^{H-1/2}}\\
&\quad + \frac{\delta^\beta(r-r_\delta)^{\frac{1}{2}-H}}{(r_\delta/2)^{H-1/2}(H-1/2)} + \frac{\delta^\beta\left(\frac{r_\delta}{2}\right)^{\frac{3}{2}-H}}{\left(r-r_\delta/2\right)^{H+1/2}\left(\frac{3}{2}-H\right)}\Big]\mathbb{1}_{[r\geq\delta]}\\
&\quad + \frac{1}{2}\mathcal{B}(\beta+\frac{1}{2}-H,\frac{3}{2}-H)r^{\beta+1-2H}\mathbb{1}_{[0\leq r<\delta]}\Big\}. \quad (4.4.15)
\end{aligned}
$$

On the other hand, it follows from (H2) that

$$
\begin{aligned}
|Z(Y_r^\xi)-Z(\hat{Y}_r^\xi)-(Z(Y_s^\xi)-Z(\hat{Y}_s^\xi))| &\leq L_2\|Y_r^\xi-\hat{Y}_r^\xi\|_\infty^\alpha + L_2\|Y_s^\xi-\hat{Y}_s^\xi\|_\infty^\alpha\\
&\leq 2L_2\|Y\|_{-\tau,r,\beta\wedge\theta}^\alpha\delta^{\alpha(\beta\wedge\theta)},
\end{aligned}
$$

and

$$
\begin{aligned}
&|Z(Y_r^\xi)-Z(\hat{Y}_r^\xi)-(Z(Y_s^\xi)-Z(\hat{Y}_s^\xi))|\\
&\leq L_2\|Y_r^\xi-Y_s^\xi\|_\infty^\alpha + L_2\|\hat{Y}_r^\xi-\hat{Y}_s^\xi\|_\infty^\alpha
\end{aligned}
$$

$$\leq L_2\|Y\|^{\alpha}_{-\tau,r,\beta\wedge\theta}|r-s|^{\alpha(\beta\wedge\theta)} + L_2\|Y\|^{\alpha}_{-\tau,r,\beta\wedge\theta}\Big(|r-s|^{\alpha(\beta\wedge\theta)} \vee |r_\delta - s_\delta|^{\alpha(\beta\wedge\theta)}\Big)$$
$$= L_2\|Y\|^{\alpha}_{-\tau,r,\beta\wedge\theta}\Big(|r-s|^{\alpha(\beta\wedge\theta)} + |r-s|^{\alpha(\beta\wedge\theta)} \vee |r_\delta - s_\delta|^{\alpha(\beta\wedge\theta)}\Big).$$

Combining these two upper bounds, we have

$$|Z(Y^\xi_r) - Z(\hat{Y}^\xi_r) - (Z(Y^\xi_s) - Z(\hat{Y}^\xi_s))|$$
$$\leq 2L_2\|Y^\xi\|^{\alpha}_{-\tau,r,\beta\wedge\theta}\Big(\delta^{\alpha(\beta\wedge\theta)} \wedge \frac{|r-s|^{\alpha(\beta\wedge\theta)} + |r-s|^{\alpha(\beta\wedge\theta)} \vee |r_\delta - s_\delta|^{\alpha(\beta\wedge\theta)}}{2}\Big).$$

Since for $r \geq \delta$,

$$\delta^{\alpha(\beta\wedge\theta)} \wedge \frac{|r-s|^{\alpha(\beta\wedge\theta)} + |r-s|^{\alpha(\beta\wedge\theta)} \vee |r_\delta - s_\delta|^{\alpha(\beta\wedge\theta)}}{2} = \delta^{\alpha(\beta\wedge\theta)}, s \in [0, r-\delta],$$

one gets that

$$\int_0^{r_\delta} \frac{\delta^{\alpha(\beta\wedge\theta)} \wedge \frac{|r-s|^{\alpha(\beta\wedge\theta)} + |r-s|^{\alpha(\beta\wedge\theta)} \vee |r_\delta - s_\delta|^{\alpha(\beta\wedge\theta)}}{2}}{(r-s)^{H+1/2}s^{H-1/2}}\mathrm{d}s$$
$$= \int_0^{r_\delta} \frac{\delta^{\alpha(\beta\wedge\theta)}}{(r-s)^{H+1/2}s^{H-1/2}}\mathrm{d}s$$
$$\leq \frac{\delta^{\alpha(\beta\wedge\theta)}(r-r_\delta)^{\frac{1}{2}-H}}{(r_\delta/2)^{H-1/2}(H-1/2)} + \frac{\delta^{\alpha(\beta\wedge\theta)}\left(\frac{r_\delta}{2}\right)^{\frac{3}{2}-H}}{(r-r_\delta/2)^{H+1/2}\left(\frac{3}{2}-H\right)},$$

and

$$\int_{r_\delta}^{r} \frac{\delta^{\alpha(\beta\wedge\theta)} \wedge \frac{|r-s|^{\alpha(\beta\wedge\theta)} + |r-s|^{\alpha(\beta\wedge\theta)} \vee |r_\delta - s_\delta|^{\alpha(\beta\wedge\theta)}}{2}}{(r-s)^{H+1/2}s^{H-1/2}}\mathrm{d}s$$
$$= \int_{r_\delta}^{r} \frac{(r-s)^{\alpha(\beta\wedge\theta)}}{(r-s)^{H+1/2}s^{H-1/2}}\mathrm{d}s$$
$$\leq \frac{r_\delta^{\frac{1}{2}-H}(r-r_\delta)^{\alpha(\beta\wedge\theta)+\frac{1}{2}-H}}{\alpha(\beta\wedge\theta)+\frac{1}{2}-H}$$
$$\leq \frac{2r_\delta^{\frac{1}{2}-H}\delta^{\alpha(\beta\wedge\theta)+\frac{1}{2}-H}}{2\alpha(\beta\wedge\theta)+1-2H}.$$

Thus, we obtain

$$|I_{42}(r)| \leq \frac{2L_2(H-\frac{1}{2})r^{H-\frac{1}{2}}}{\Gamma(\frac{3}{2}-H)}\|Y^\xi\|^{\alpha}_{-\tau,r,\beta\wedge\theta}\Bigg\{\Big(\frac{\delta^{\alpha(\beta\wedge\theta)}(r-r_\delta)^{\frac{1}{2}-H}}{(r_\delta/2)^{H-1/2}(H-1/2)} \quad (4.4.16)$$

$$+\frac{\delta^{\alpha(\beta\wedge\theta)}\left(\frac{r_\delta}{2}\right)^{\frac{3}{2}-H}}{(r-r_\delta/2)^{H+1/2}\left(\frac{3}{2}-H\right)}+\frac{2r_\delta^{\frac{1}{2}-H}\delta^{\alpha(\beta\wedge\theta)+\frac{1}{2}-H}}{2\alpha(\beta\wedge\theta)+1-2H}\Big)\mathbb{1}[r\geq\delta]$$

$$+\mathcal{B}(\frac{3}{2}-H,\alpha(\beta\wedge\theta)+\frac{1}{2}-H)r^{\alpha(\beta\wedge\theta)+\frac{1}{2}-H}\mathbb{1}_{[0\leq r<\delta]}\Big\}.$$

Substituting (4.4.15), (4.4.16), (4.4.12) and (4.4.14) into (4.4.11), we arrive at

$$\begin{aligned}
&\Big|K_H^{-1}\Big(\int_0^{\cdot}(Z(Y_s^\xi)+h^\xi(s))\mathrm{d}s\Big)(r)\Big| \\
&\leq[1+C_0(H-\frac{1}{2})]\frac{r^{\frac{1}{2}-H}}{\Gamma(\frac{3}{2}-H)}(L_1\|\hat{\sigma}^{-1}\|\|Y^\xi\|_{0,r,\beta}\delta^\beta+L_2\|Y^\xi\|_{0,r,\beta}^\alpha\delta^{\alpha\beta}) \\
&+\frac{r^{H-1/2}(2H-1)}{\Gamma(\frac{3}{2}-H)}\Big\{L_1\|\hat{\sigma}^{-1}\|\|Y^\xi\|_{0,r,\beta}\Big[\frac{\delta^{\beta+1/2-H}}{(2\beta+1-2H)r_\delta^{H-1/2}}+\frac{\delta^\beta(r-r_\delta)^{\frac{1}{2}-H}}{(r_\delta/2)^{H-1/2}(H-1/2)} \\
&+\frac{\delta^\beta\left(\frac{r_\delta}{2}\right)^{\frac{3}{2}-H}}{(r-r_\delta/2)^{H+1/2}\left(\frac{3}{2}-H\right)}\Big]+L_2\|Y^\xi\|_{-\tau,r,(\beta\wedge\theta)}^\alpha\Big[\frac{\delta^{\alpha(\beta\wedge\theta)}(r-r_\delta)^{\frac{1}{2}-H}}{(r_\delta/2)^{H-1/2}(H-1/2)} \\
&+\frac{\delta^{\alpha(\beta\wedge\theta)}\left(\frac{r_\delta}{2}\right)^{\frac{3}{2}-H}}{(r-r_\delta/2)^{H+1/2}\left(\frac{3}{2}-H\right)}+\frac{2r_\delta^{\frac{1}{2}-H}\delta^{\alpha(\beta\wedge\theta)+\frac{1}{2}-H}}{2\alpha(\beta\wedge\theta)+1-2H}\Big]\Big\}\mathbb{1}_{[r\geq\delta]} \\
&+\frac{r^{\frac{1}{2}-H}(2H-1)}{\Gamma(\frac{3}{2}-H)}\Big\{\frac{L_1\|\hat{\sigma}^{-1}\|}{2}\mathcal{B}(\frac{3}{2}-H,\frac{1}{2}+\beta-H)\|Y^\xi\|_{0,r,\beta}r^{\beta+1-2H} \\
&+L_2\mathcal{B}(\frac{3}{2}-H,\frac{1}{2}+\alpha(\beta\wedge\theta)-H)\|Y^\xi\|_{-\tau,r,\beta\wedge\theta}^\alpha r^{\alpha(\beta\wedge\theta)+\frac{1}{2}-H}\Big\}\mathbb{1}_{[0\leq r<\delta]}.
\end{aligned}\tag{4.4.17}$$

This, together with the BDG inequality, yields that

$$\begin{aligned}
\Theta_2(t)&\leq C_T\Big(\mathbb{E}\left(\int_0^T\Big|K_H^{-1}\Big(\int_0^{\cdot}(Z(Y_s^\xi)+h^\xi(s))\mathrm{d}s\Big)(r)\Big|^2\mathrm{d}r\right)^{\frac{q}{(q-1)2}}\Big)^{\frac{q-1}{q}} \\
&\leq C_T\delta^{\alpha(\beta\wedge\theta)+\frac{1}{2}-H}.
\end{aligned}$$

For Θ_3, it follows from Hölder's inequality and (4.4.17) that

$$\Theta_3(t)\leq\frac{1}{2}\left(\mathbb{E}\left(\int_0^T\left|K_H^{-1}\left(\int_0^{\cdot}(Z(Y_s^\xi)-h^\xi(s))\mathrm{d}s\right)(r)\right|^2\mathrm{d}r\right)^{\frac{q}{q-1}}\right)^{\frac{q-1}{2q}}$$

$$\times \left(\mathbb{E}\left(\int_0^T \left|K_H^{-1}\Big(\int_0^\cdot (Z(Y_s^\xi)+h^\xi(s))\mathrm{d}s\Big)(r)\right|^2 \mathrm{d}r\right)^{\frac{q}{q-1}}\right)^{\frac{q-1}{2q}}$$

$$\leq C_T\delta^{\alpha(\beta\wedge\theta)+\frac{1}{2}-H}\left(\mathbb{E}\left(\int_0^T \left|K_H^{-1}\left(\int_0^\cdot (Z(Y_s^\xi)-h^\xi(s))\mathrm{d}s\right)(r)\right|^2 \mathrm{d}r\right)^{\frac{q}{q-1}}\right)^{\frac{q-1}{2q}}.$$

Note that

$$\int_0^T \left|K_H^{-1}\left(\int_0^\cdot (Z(Y_s^\xi)-h^\xi(s))\mathrm{d}s\right)\right|^2(r)\mathrm{d}r$$
$$\leq 2\int_0^T \left|K_H^{-1}\left(\int_0^\cdot (Z(Y_s^\xi)+h^\xi(s))\mathrm{d}s\right)\right|^2(r)\mathrm{d}r + 2\int_0^T \left|K_H^{-1}\left(\int_0^\cdot h^\xi(s)\mathrm{d}s\right)\right|^2(r)\mathrm{d}r,$$

it follows from (4.4.17) and (4.4.8) that

$$\left(\mathbb{E}\left(\int_0^T \left|K_H^{-1}\left(\int_0^\cdot (Z(Y_s^\xi)-h^\xi(s))\mathrm{d}s\right)\right|^2(r)\mathrm{d}r\right)^{\frac{q}{q-1}}\right)^{\frac{q-1}{2q}} < \infty.$$

We then obtain that

$$\Theta_3(t) \leq C_T\delta^{\alpha(\beta\wedge\theta)+\frac{1}{2}-H}.$$

Finally, the desired assertion is established from (4.4.9) and the estimates of $\Theta_i(t), i=1,2,3$. □

Chapter 5

Moderate deviations and central limit theorem for McKean-Vlasov stochastic differential equations

Inspired by the Freidlin-Wentzell LDP in path space for MV-SDEs in [28], we will consider the MDP and CLT for MV-SDEs in this chapter.

In Section 5.1, we introduce the MV-SDEs and its particular properties.

In Section 5.2, we recall the general deviations of the solution to the MV-SDEs, and give the associated assumptions about the coefficients of the MV-SDEs (that is, the Lipschitz condition about the coefficients, the gradient of coefficients with respect to the space variable, and the L-derivative of coefficients with respect to the measure variable, respectively).

Section 5.3 describes the CLT for MV-SDEs, and the proof is provided in subsection 5.3.1.

In Section 5.4, we establish the MDP for MV-SDEs.

Section 5.5 is devoted to the proof of Throrem 5.4.1. More precisely, in view of the weak convergence methods and exponential approximation, Subsection 5.5.1 is devoted to the LDP for $\bar{Y}^\varepsilon$. In Subsection 5.5.2, we show $\bar{X}^\varepsilon$ and $\bar{Y}^\varepsilon$ are exponentially equivalent.

Section 5.6 provides an illustrative example to verify that it satisfies the assumptions in this chapter work.

5.1 McKean-Vlasov stochastic differential equation

In recent years, MV-SDEs have been received increasing attention by researchers. They are also called mean-field SDEs or distribution dependent SDEs, which are much more involved than classical SDEs as the drift and diffusion coefficients depend on the solution and the law of solution. In a nutshell, this kind of equations play important role in characterising nonlinear Fokker-Planck equations and environment dependent financial systems (see [26, 27, 32, 37, 64, 94, 95] and references therein). Also, this kind of SDEs have been applied to characterise the PDEs involving the L-derivative, which was introduced by P.-L. Lions in his lecture notes [19], see also [11, 20, 39, 52, 86, 87] for more details. Additionally, the analysis of stochastic particle systems has been developed as a crucial mathematic tool for modelling systems in economics and finance.

Compared with the classical SDEs (the law of solution to this equation satisfies the linear PDE), the law of solution to MV-SDEs satisfies the nonlinear Kolmogorov-Fokker-Planck equation. To explain it, we introduce the following model on $\mathbb{R}^d$:

$$\mathrm{d}X_t^\mu = b(X_t^\mu, \mathscr{L}_{X_t^\mu})\mathrm{d}t + \sigma(X_t^\mu, \mathscr{L}_{X_t^\mu})\mathrm{d}W_t, \quad \mathscr{L}_{X_0} = \mu, \tag{5.1.1}$$

where W_t is a d-dimensional Brownian motion on a complete filtration probability space $(\Omega, \{\mathscr{F}_t\}_{t\geq 0}, \mathscr{F}, \mathbb{P})$. $\mu_t := \mathscr{L}_{X_t^\mu}$ denotes the law of solution to (5.1.1) at time t with the initial distribution μ, and

$$b : \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}^d,\ \sigma : \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^d,$$

are measurable.

Let $L_\mu = \langle \partial_x \cdot, b(x,\mu)\rangle + \frac{1}{2}tr\big(\sigma\sigma^* \partial^2_{x^2} \cdot\big)$. Then, Itô's formula yields that

$$\mathbb{E}f(X_t^\mu) = \mathbb{E}f(X_0) + \int_0^t \mathbb{E}L_\mu f(X_s^\mu)\mathrm{d}s, \quad \forall f \in \mathscr{B}_b(\mathbb{R}^d).$$

This, together with the arbitrariness of f, implies the following non-linear Fokker-Planck equation:

$$\partial_t \mu_t = L^*_{\mu_t}\mu_t,$$

where L^*_μ is the adjoint operator of L_μ. For more properties of MV-SDEs, (see, e.g. [6, 87] and references therein).

5.2 General deviations and assumptions

Let $(\mathbb{R}^d, \langle\cdot,\cdot\rangle, |\cdot|)$ be the d-dimensional Euclidean space. Consider the Cameron-Martin space

$$\mathbb{H} = \Big\{h \in C([0,T];\mathbb{R}^d) : h(0) = \mathbf{0}, \dot{h}(t) \text{ exists for a.e. } t, \|h\|_{\mathbb{H}} := \Big(\int_0^T |\dot{h}(t)|^2 \mathrm{d}t\Big)^{\frac{1}{2}}\Big\},$$

where $\mathbf{0}$ denotes the vector with components 0.

Let $\mathscr{A}$ denote the class of $\mathbb{R}^d$-valued $\{\mathscr{F}_t\}$-predictable processes $h(\omega,\cdot)$ belonging to $\mathbb{H}$ a.s.. For each $N > 0$, let

$$S_N = \Big\{h \in \mathbb{H};\ \int_0^T |\dot{h}(s)|^2 \mathrm{d}s \leq N\Big\}.$$

S_N is endowed with the weak topology induced by $\mathbb{H}$. Define

$$\mathscr{A}_N := \{h \in \mathscr{A}, h(\omega, \cdot) \in S_N,\ \mathbb{P} - a.s.\}.$$

In this chapter, we use the symbol " $\Rightarrow$ " to denote the convergence in distribution.

Consider the following MV-SDE on $(\mathbb{R}^d, \langle \cdot, \cdot \rangle, |\cdot|)$:

$$\mathrm{d}X_t^\epsilon = b_t(X_t^\epsilon, \mathscr{L}_{X_t^\epsilon})\mathrm{d}t + \sqrt{\epsilon}\sigma_t(X_t^\epsilon, \mathscr{L}_{X_t^\epsilon})\mathrm{d}W_t, \quad X_0^\epsilon = x, \tag{5.2.1}$$

with $\epsilon > 0$, which is called as the scaling parameter. Here W_t is the d-dimensional Brownian motion defined on a complete filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$, $\mathscr{L}_{X_t^\epsilon}$ is the law of X_t^ϵ.

Intuitively, as the parameter ϵ tends to 0 in (5.2.1), the diffusion term vanishes and we have the following ordinary differential equation:

$$\mathrm{d}X_t^0 = b_t(X_t^0, \delta_{X_t^0})\mathrm{d}t, \tag{5.2.2}$$

with the same initial datum as (5.2.1), that is, $X_0^0 = x$. Since x is deterministic, we deduce that $\delta_{X_\cdot^0}$ is a Dirac measure centered on the path $X_\cdot^0$.

On the general case, the investigation of the deviations of solution X_t^ϵ to (5.2.1) from the solution X_t^0 to (5.2.2) is to study the asymptotic behaviour of the trajectory

$$\bar{X}_t^\epsilon = \frac{1}{\sqrt{\epsilon}\lambda(\epsilon)}(X_t^\epsilon - X_t^0),\ t \in [0, T], \tag{5.2.3}$$

it yields the following three cases:

(LDP) The case $\lambda(\epsilon) = 1/\sqrt{\epsilon}$ provides an LDP. [28] proved that the law of the solution X^ϵ satisfies an LDP by means of the discussion of exponential tightness.

(CLT) The case $\lambda(\epsilon) \equiv 1$ describes a CLT for solution X^ϵ. That is, $\frac{X^\epsilon - X^0}{\sqrt{\epsilon}}$ converges to a stochastic process in a certain sense as $\epsilon \to 0$, see Theorem 5.3.1.

(MDP) To fill in the gap between the CLT scale and the LDP scale, the MDP for X^ϵ is investigating the LDP of trajectory (5.2.3), where the deviation scale $\lambda(\epsilon)$ satisfies

$$\lambda(\epsilon) \to \infty, \quad \sqrt{\epsilon}\lambda(\epsilon) \to 0, \quad \text{as } \epsilon \to 0. \tag{5.2.4}$$

To obtain the main results of this chapter, we assume that the coefficients b and σ satisfy the following conditions:

(**H1**) For any $t \geq 0$, $b_t \in C^{1,(1,0)}(\mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$. Moreover, there exists an increasing function $K : [0, \infty) \to [0, \infty)$ such that

$$\max\{\|\nabla b_t(\cdot, \mu)(x)\|, \|D^L b_t(x, \cdot)(\mu)\|_{T_{\mu,2}}\} \leq K(t), \tag{5.2.5}$$

$$\|\sigma_t(x, \mu) - \sigma_t(y, \nu)\| \leq K(t)(|x - y| + \mathbb{W}_2(\mu, \nu)), \tag{5.2.6}$$
$$t \geq 0, \quad x, y \in \mathbb{R}^d, \quad \mu, \nu \in \mathscr{P}_2(\mathbb{R}^d),$$

and

$$|b_t(\mathbf{0}, \delta_{\mathbf{0}})| + \|\sigma_t(\mathbf{0}, \delta_{\mathbf{0}})\| \leq K(t), \quad t \geq 0. \tag{5.2.7}$$

(**H2**) $\nabla b_t(\cdot, \mu)(x)$ and $D^L b_t(x, \cdot)(\mu)$ satisfy

$$\|\nabla b_t(\cdot, \mu)(x) - \nabla b_t(\cdot, \nu)(y)\| \leq K(t)(|x - y| + \mathbb{W}_2(\mu, \nu)), \tag{5.2.8}$$

and

$$\begin{aligned}
&|D^L b_t(x, \cdot)(\mu)(z_1) - D^L b_t(y, \cdot)(\nu)(z_2)| \\
&\leq K(t)(|x - y| + \mathbb{W}_2(\mu, \nu) + |z_1 - z_2|),
\end{aligned}$$

for all $t \geq 0, \ x, y, z_1, z_2 \in \mathbb{R}^d$.

Remark 5.2.1. By (**H1**), we have for $t \geq 0, x, y \in \mathbb{R}^d, \mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$ that

$$|b_t(x,\mu) - b_t(y,\mu)| \leq K(t)(|x-y| + \mathbb{W}_2(\mu,\nu)). \tag{5.2.9}$$

5.3 Central limit theorem

The first main result is to investigate the CLT for $(X^\epsilon)_{\epsilon \in (0,1)}$ to (5.2.1), which is stated as follows:

Theorem 5.3.1. *Under assumptions (**H1**) and (**H2**),*

$$\mathbb{E}\Big(\sup_{0\leq t\leq T}\Big|\frac{X_t^\epsilon - X_t^0}{\sqrt{\epsilon}} - Z_t\Big|^p\Big) \lesssim \epsilon^{\frac{p}{2}}, \quad \textit{for any } p \geq 2,$$

where Z_t solves

$$\begin{aligned} \mathrm{d}Z_t &= \nabla_{Z_t} b_t(\cdot, \delta_{X_t^0})(X_t^0)\mathrm{d}t + \mathbb{E}\langle D^L b_t(y,\cdot)(\delta_{X_t^0})(X_t^0), Z_t\rangle|_{y=X_t^0}\mathrm{d}t \\ &\quad + \sigma(X_t^0, \delta_{X_t^0})\mathrm{d}W_t, \ Z_0 = \mathbf{0}. \end{aligned} \tag{5.3.1}$$

Here, and in what follows, for $x, y \in \mathbb{R}^d$ and $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, $\nabla_y f(\cdot,\mu)(x)$ constitutes the directional derivative of function f at x in direction y.

5.3.1 Proof of the central limit theorem

Before giving the proof of Theorem 5.3.1, we prepare the following lemmas.

The existence and uniqueness of solution to (5.2.1) has been proved in [96]. The following Lemma gives the uniformly p-th moment estimates on X_t^ϵ and X_t^0.

Lemma 5.3.2. *Under assumption (**H1**). $\forall p \geq 2$, we have*

$$\mathbb{E}\Big(\sup_{0\leq t\leq T}|X_t^\epsilon|^p\Big) \vee \Big(\sup_{0\leq t\leq T}|X_t^0|^p\Big) < \infty, \tag{5.3.2}$$

with the initial value $X_0^0 = X_0^\epsilon = x \in \mathbb{R}^d$.

Proof. It is easy to get from (**H1**) that

$$|b_t(x,\mu)| \vee \|\sigma_t(x,\mu)\| \leq K(t)(1+|x|+\mathbb{W}_2(\mu,\delta_0)). \tag{5.3.3}$$

Noting that $\mathbb{W}_2(\mathscr{L}_{X_s^\epsilon},\delta_0)^p \leq (\mathbb{E}|X_s^\epsilon|^2)^{p/2}$, by the BDG inequality and (5.3.3), one has

$$\mathbb{E}\Big(\sup_{0\leq t\leq T}|X_t^\epsilon|^p\Big) \leq 3^{p-1}|x|^p + C(T,p)\mathbb{E}\int_0^T (1+|X_s^\epsilon|^p)\mathrm{d}s,$$

and

$$\Big(\sup_{0\leq t\leq T}|X_t^0|^p\Big) \leq C(T,p)\int_0^T (1+|X_s^0|^p)\mathrm{d}s,$$

thus, (5.3.2) follows from Gronwall's inequality. □

Lemma 5.3.3. *Under (***H1***) and (***H2***), we have* $\forall p \geq 2$

$$\mathbb{E}\Big(\sup_{0\leq t\leq T}|Z_t^\epsilon|^p\Big) \vee \mathbb{E}\Big(\sup_{0\leq t\leq T}|Z_t|^p\Big) < \infty, \tag{5.3.4}$$

where $Z_\cdot^\epsilon := \frac{X_\cdot^\epsilon - X_\cdot^0}{\sqrt{\epsilon}}$ *and* Z_t *is defined in* (5.3.1).

Proof. By (5.2.1) and (5.2.2), we know that Z_t^ϵ satisfies

$$\mathrm{d}Z_t^\epsilon = \frac{1}{\sqrt{\epsilon}}(b_t(X_t^\epsilon,\mathscr{L}_{X_t^\epsilon}) - b_t(X_t^0,\delta_{X_t^0}))\mathrm{d}t + \sigma_t(X_t^\epsilon,\mathscr{L}_{X_t^\epsilon})\mathrm{d}W_t. \tag{5.3.5}$$

To prove $\mathbb{E}\Big(\sup_{0\leq t\leq T}|Z_t^\epsilon|^p\Big) < \infty,\ \forall p \geq 2$, it suffices to show

$$\mathbb{E}\Big(\sup_{0\leq t\leq T}|X_t^\epsilon - X_t^0|^p\Big) \leq C(T,p)\epsilon^{\frac{p}{2}}. \tag{5.3.6}$$

Indeed, by (5.2.9), (5.3.3), Hölder's inequality and BDG's inequality, one gets that

$$\begin{aligned}&\mathbb{E}\Big(\sup_{0\leq t\leq T}|X_t^\epsilon - X_t^0|^p\Big)\\ &\leq 2^{p-1}\Big\{\mathbb{E}\Big|\int_0^T |b_s(X_s^\epsilon,\mathscr{L}_{X_s^\epsilon}) - b_s(X_s^0,\delta_{X_s^0})|\mathrm{d}s\Big|^p\end{aligned}$$

$$
\begin{aligned}
&+\epsilon^{p/2}\mathbb{E}\Big(\sup_{0\le t\le T}\Big|\int_0^t \sigma_s(X_s^\epsilon,\mathscr{L}_{X_s^\epsilon})\mathrm{d}W_s\Big|^p\Big)\Big\}\\
&\le C(T,p)\Big\{\int_0^T(|X_s^\epsilon-X_s^0|+\mathbb{W}_2(\mathscr{L}_{X_s^\epsilon},\delta_{X_s^0}))^p\mathrm{d}s\\
&\quad+\epsilon^{p/2}\Big(\int_0^T(\mathbb{E}|X_s^\epsilon|^2+1)\mathrm{d}s\Big)^{p/2}\Big\}\\
&\le C(T,p)\int_0^T\mathbb{E}|X_s^\epsilon-X_s^0|^p\mathrm{d}s+\epsilon^{p/2}C(T,p)\Big(1+\int_0^T\mathbb{E}|X_s^\epsilon|^p\mathrm{d}s\Big),
\end{aligned}
$$

where the last inequality is due to the fact that $\mathbb{W}_2(\mathscr{L}_{X_s^\epsilon},\delta_0)^2\le\mathbb{E}|X_s^\epsilon|^2$. Then, (5.3.6) follows from (5.3.2) and the Gronwall inequality.

Similarly, by (**H2**) and (5.3.2), we derive from (5.3.1) that

$$
\begin{aligned}
\mathbb{E}\Big(\sup_{0\le t\le T}|Z_t|^p\Big)&\le C(T,p)\int_0^T\mathbb{E}|Z_t|^p\mathrm{d}t+C(T,p)\int_0^T(1+|X_t^0|^p)\mathrm{d}t\\
&\le C(T,p)\Big(1+\int_0^T\mathbb{E}|Z_t|^p\mathrm{d}t\Big).
\end{aligned}
$$

This, together with Gronwall's inequality, implies the desired assertion (5.3.4).

□

Now, we are in position to finish the **Proof of Theorem 5.3.1**.

Proof. By the definitions of Z_t^ϵ and Z_t, we derive that

$$
\begin{aligned}
&Z_t^\epsilon-Z_t\\
&=\int_0^t\Big(\frac{1}{\sqrt{\epsilon}}(b_s(X_s^\epsilon,\mathscr{L}_{X_s^\epsilon})-b_s(X_s^0,\mathscr{L}_{X_s^\epsilon}))-\nabla_{Z_s^\epsilon}b_s(\cdot,\mathscr{L}_{X_s^\epsilon})(X_s^0)\Big)\mathrm{d}s\\
&+\int_0^t\Big(\frac{1}{\sqrt{\epsilon}}(b_s(X_s^0,\mathscr{L}_{X_s^\epsilon})-b_s(X_s^0,\delta_{X_s^0}))-\mathbb{E}\langle D^Lb_s(y,\cdot)(\delta_{X_s^0})(X_s^0),Z_s^\epsilon\rangle|_{y=X_s^0}\Big)\mathrm{d}s\\
&+\int_0^t(\nabla_{Z_s^\epsilon}b_s(\cdot,\mathscr{L}_{X_s^\epsilon})(X_s^0)-\nabla_{Z_s}b_s(\cdot,\delta_{X_s^0})(X_s^0))\mathrm{d}s\\
&+\int_0^t(\mathbb{E}\langle D^Lb_s(y,\cdot)(\delta_{X_s^0})(X_s^0),Z_s^\epsilon\rangle|_{y=X_s^0}-\mathbb{E}\langle D^Lb_s(y,\cdot)(\delta_{X_s^0})(X_s^0),Z_s\rangle|_{y=X_s^0})\mathrm{d}s\\
&+\int_0^t(\sigma_s(X_s^\epsilon,\mathscr{L}_{X_s^\epsilon})-\sigma_s(X_s^0,\delta_{X_s^0}))\mathrm{d}W_s.
\end{aligned}
$$

By (**H2**), Lemma 2.3.1, Hölder's inequality and BDG's inequality, one gets that

$$
\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le T}|Z_t^\epsilon - Z_t|^p\Big) \qquad (5.3.7)\\
&\le C(T,p)\int_0^T \mathbb{E}\Big|\int_0^1 \nabla_{Z_s^\epsilon} b_s(\cdot,\mathscr{L}_{X_s^\epsilon})(R_s^\epsilon(r))\mathrm{d}r - \nabla_{Z_s^\epsilon} b_s(\cdot,\mathscr{L}_{X_s^\epsilon})(X_s^0)\Big|^p \mathrm{d}s\\
&\quad + C(T,p)\int_0^T \mathbb{E}\Big|\int_0^1 \mathbb{E}\langle D^L b_s(y,\cdot)(\mathscr{L}_{R_s^\epsilon(r)})(R_s^\epsilon(r)), Z_s^\epsilon\rangle|_{y=X_s^0}\mathrm{d}r\\
&\qquad\qquad - \mathbb{E}\langle D^L b_s(y,\cdot)(\delta_{X_s^0})(X_s^0), Z_s^\epsilon\rangle|_{y=X_s^0}\Big|^p\mathrm{d}s\\
&\quad + C(T,p)\int_0^T \mathbb{E}|\nabla_{Z_s^\epsilon - Z_s} b_s(\cdot,\mathscr{L}_{X_s^\epsilon})(X_s^0)|^p \mathrm{d}s\\
&\quad + C(T,p)\int_0^T \Big(\mathbb{E}|\nabla_{Z_s} b_s(\cdot,\mathscr{L}_{X_s^\epsilon})(X_s^0) - \nabla_{Z_s} b_s(\cdot,\delta_{X_s^0})(X_s^0)|^p\Big)\mathrm{d}s\\
&\quad + C(T,p)\int_0^T \mathbb{E}\Big|\mathbb{E}\langle D^L b_s(y,\cdot)(\delta_{X_s^0})(X_s^0), Z_s^\epsilon\rangle|_{y=X_s^0}\\
&\qquad\qquad - \mathbb{E}\langle D^L b_s(y,\cdot)(\delta_{X_s^0})(X_s^0), Z_s\rangle|_{y=X_s^0}\Big|^p\mathrm{d}s\\
&\quad + C(T,p)\int_0^T \mathbb{E}|\sigma_s(X_s^\epsilon,\mathscr{L}_{X_s^\epsilon}) - \sigma_s(X_s^0,\delta_{X_s^0})|^p\mathrm{d}s\\
&=: \sum_{i=1}^6 J_i(T),\ I=1,2,\cdots,6,
\end{aligned}
$$

where $R_s^\epsilon(r) = X_s^0 + r(X_s^\epsilon - X_s^0), r\in[0,1]$.

By (**H1**), (**H2**), (5.3.4) and Hölder's inequality, we have

$$
\begin{aligned}
&\sum_{i=1}^5 J_i(T) \qquad (5.3.8)\\
&\le C(T,p)\Big\{\int_0^T (\mathbb{E}|Z_s^\epsilon|^2)^{\frac{p}{2}}\mathbb{E}\Big(\int_0^1((\mathbb{E}|R_s^\epsilon(r)-X_s^0|^2)^{1/2} + \mathbb{W}_2(\mathscr{L}_{R_s^\epsilon(r)},\delta_{X_s^0}))\mathrm{d}r\Big)^p\mathrm{d}s\\
&+ \epsilon^{p/2}\int_0^T \mathbb{E}|Z_s^\epsilon|^{2p}\mathrm{d}s + \int_0^T \mathbb{E}|Z_s^\epsilon - Z_s|^p\mathrm{d}s + \int_0^T \mathbb{E}(|Z_s|\mathbb{W}_2(\mathscr{L}_{X_s^\epsilon},\delta_{X_s^0}))^p\mathrm{d}s\Big\}\\
&\le C(T,p)\Big\{\epsilon^{p/2}\int_0^T\big(\mathbb{E}|Z_s^\epsilon|^{2p} + \mathbb{E}|Z_s|^p\mathbb{E}|Z_s^\epsilon|^p\big)\mathrm{d}s + \int_0^T\mathbb{E}|Z_s^\epsilon - Z_s|^p\mathrm{d}s\Big\},
\end{aligned}
$$

where we used $\mathbb{W}_2(\mathscr{L}_{R_s^\epsilon(r)},\delta_{X_s^0}) \le r\sqrt{\epsilon}(\mathbb{E}|Z_s^\epsilon|^2)^{1/2}$ and $\mathbb{W}_2(\mathscr{L}_{X_s^\epsilon},\delta_{X_s^0}) \le \epsilon^{1/2}(\mathbb{E}|Z_s^\epsilon|^2)^{1/2}$ in the last inequality.

Moreover, we obtain from (**H1**), (5.3.4) and Hölder's inequality that

$$\begin{aligned} J_6(T) &\le C(T,p)\int_0^T (\epsilon^{p/2}\mathbb{E}|Z_s^\epsilon|^p + \mathbb{E}\mathbb{W}_2(\mathscr{L}_{X_s^\epsilon},\delta_{X_s^0})^p)\mathrm{d}s \\ &\le C(T,p)\int_0^T \epsilon^{p/2}\mathbb{E}|Z_s^\epsilon|^p\mathrm{d}s. \end{aligned} \tag{5.3.9}$$

Collecting the estimates (5.3.8) and (5.3.9) into (5.3.7), we arrive at

$$\begin{aligned} &\mathbb{E}\Big(\sup_{0\le t\le T}|Z_t^\epsilon - Z_t|^p\Big) \\ &\le C(T,p)\Big\{\epsilon^{p/2}\int_0^T \big(\mathbb{E}|Z_s^\epsilon|^{2p} + \mathbb{E}|Z_s|^p\mathbb{E}|Z_s^\epsilon|^p\big)\mathrm{d}s + \int_0^T \mathbb{E}|Z_s^\epsilon - Z_s|^p\mathrm{d}s\Big\}. \end{aligned}$$

This, together with the Gronwall inequality, yields that

$$\mathbb{E}\Big(\sup_{0\le t\le T}|Z_t^\epsilon - Z_t|^p\Big) \le C_{T,p}\epsilon^{p/2}.$$

The desired assertion is obtained by taking $\epsilon \to 0$. □

5.4 Moderate deviation principle

The main result of this section is the MDP for $(X^\epsilon)_{\epsilon\in(0,1)}$ to (5.2.1), which is stated as follows:

Theorem 5.4.1. *Under assumptions (***H1***) and* (5.2.8) *of (***H2***). Then,* $\bar{X}^\epsilon_\cdot$*, defined in* (5.2.3)*, satisfies an LDP on* $C([0,T];\mathbb{R}^d)$ *with the rate function* I *which is defined by*

$$I(g) := \inf_{\{h\in\mathbb{H};\, g=\Gamma^0(\int_0^\cdot \dot{h}(s)\mathrm{d}s)\}}\Big\{\frac{\|h\|_{\mathbb{H}}^2}{2}\Big\},\quad g\in C([0,T];\mathbb{R}^d), \tag{5.4.1}$$

where, by convention, $I(g) = \infty$ *if* $\{h\in\mathbb{H}; g = \Gamma^0(\int_0^\cdot \dot{h}(s)\mathrm{d}s)\} = \emptyset$ *and* $Y^h_\cdot := \Gamma^0(\int_0^\cdot \dot{h}(s)\mathrm{d}s)$ *satisfies the following equation:*

$$\mathrm{d}Y_t^h = \Big\{\nabla_{Y_t^h} b_t(\cdot,\delta_{X_t^0})(X_t^0) + \sigma_t(X_t^0,\delta_{X_t^0})\dot{h}(t)\Big\}\mathrm{d}t. \tag{5.4.2}$$

Remark 5.4.1. Theorems 5.3.1 and 5.4.1 can be extended to the case of path-distribution dependent SDEs, and the Lipschitz condition imposed on the drift can be relaxed to the monotone condition.

5.5 Proof of the moderate deviation principle

From (5.2.1)-(5.2.3), we can see that $\bar{X}^\epsilon$ satisfies the following equation:

$$\bar{X}_t^\epsilon = \frac{1}{\sqrt{\epsilon}\lambda(\epsilon)} \int_0^t [b_s(X_s^\epsilon, \mathscr{L}_{X_s^\epsilon}) - b_s(X_s^0, \delta_{X_s^0})]\mathrm{d}s + \frac{1}{\lambda(\epsilon)} \int_0^t \sigma_s(X_s^\epsilon, \mathscr{L}_{X_s^\epsilon})\mathrm{d}W_s. \tag{5.5.1}$$

Subsequently, we aim to show that the law of $\bar{X}_t^\epsilon$ satisfies an LDP. To this end, we first recall that the LDP for stochastic processes, the idea is to identify a deterministic path around which the diffusion is concentrated with overwhelming probability, so that the stochastic motion can be seen as a small random perturbation of this deterministic path. In particular, this means that the law of $\bar{X}_t^\epsilon$ is close to some Dirac mass if ϵ is small. We therefore proceed in two steps towards the aim of proving that the law of $\bar{X}^\epsilon$ satisfies an LDP.

Firstly, note that $\mathscr{L}_{X_t^\epsilon}$ will converge to $\delta_{X_t^0}$ in distribution as the deviation scale $\lambda(\epsilon)$ satisfying (5.2.4). We replace $\mathscr{L}_{X_t^\epsilon}$ by $\delta_{X_t^0}$ in (5.5.1) and obtain an approximation equation of (5.5.1) as follows:

$$\bar{Y}_t^\epsilon = \frac{1}{\sqrt{\epsilon}\lambda(\epsilon)} \int_0^t [b_s(\tilde{Y}_s^\epsilon, \delta_{X_s^0}) - b_s(X_s^0, \delta_{X_s^0})]\mathrm{d}s + \frac{1}{\lambda(\epsilon)} \int_0^t \sigma_s(\tilde{Y}_s^\epsilon, \delta_{X_s^0})\mathrm{d}W_s, \tag{5.5.2}$$

where $\mathrm{d}\tilde{Y}_t^\epsilon = b_t(\tilde{Y}_t^\epsilon, \delta_{X_t^0})\mathrm{d}t + \sqrt{\epsilon}\sigma_t(\tilde{Y}_t^\epsilon, \delta_{X_t^0})\mathrm{d}W_t$ and $\bar{Y}_t^\epsilon = \frac{\tilde{Y}_t^\epsilon - X_t^0}{\sqrt{\epsilon}\lambda(\epsilon)}$. Then, we establish the law of $\bar{Y}_t^\epsilon$ satisfying an LDP.

Secondly, we claim that $\bar{X}_t^\epsilon$ and $\bar{Y}_t^\epsilon$ are exponentially equivalent. Thus, we obtain that the law of $\bar{X}_t^\epsilon$ satisfies an LDP with the good rate function $I(g)$ given in (5.4.1) due to the fact the LDP does not distinguish between exponentially equivalent families.

Note that (5.5.2), is indeed a classical SDEs with time dependent variable, the LDP for this type model has been investigated extensively in the existing literatures. To make the contents is self-contained, we sketch the proof of the law of $\bar{Y}^\epsilon$ satisfying an LDP in the following subsection.

5.5.1 Large deviation principle for $\bar{Y}^\epsilon$

Lemma 5.5.1. *Under the assumptions of Theorem 5.4.1, the family of* $(\bar{Y}^\epsilon)_{\epsilon>0}$ *satisfies an LDP in* $C([0,T];\mathbb{R}^d)$ *equipped with the topology of the uniform norm with the good rate function* $I(g)$ *given in* (5.4.1).

According to the Lemma 2.4.2, to complete the proof of Lemma 5.5.1, we only need to verify the conditions (a) and (b) in Lemma 2.4.2.

By the Yamada-Watanabe theorem, there exists a measurable map Γ^ϵ : $C([0,T];\mathbb{R}^d) \to C([0,T];\mathbb{R}^d)$ such that $\bar{Y}^\epsilon_\cdot = \Gamma^\epsilon\Big(\frac{1}{\lambda(\epsilon)}W_\cdot\Big)$.

Since $\mathbb{E}_{\mathbb{P}}\Big(\exp\big\{\frac{1}{2}\int_0^T |\dot{h}_\epsilon(s))|^2\mathrm{d}s\big\}\Big) < \infty$ for $h_\epsilon \in \mathscr{A}_N$, that is, the Novikov condition holds. By the Girsanov theorem, we know that

$$\frac{1}{\lambda(\epsilon)}\tilde{W}_t = \frac{1}{\lambda(\epsilon)}W_t + \int_0^t \dot{h}_\epsilon(s)\mathrm{d}s$$

is a Brownian motion under the probability measure $\mathbb{P}_\epsilon := R_T\mathbb{P}$, where

$$R_T = \exp\Big\{-\int_0^T \dot{h}_\epsilon(s)\mathrm{d}\frac{W_s}{\lambda(\epsilon)} - \frac{1}{2}\int_0^T |\dot{h}_\epsilon(s)|^2\mathrm{d}s\Big\}$$

is an exponential martingale.

Furthermore, we obtain that $\bar{Y}^{\epsilon,h_\epsilon}_\cdot = \Gamma^\epsilon\Big(\frac{1}{\lambda(\epsilon)}W_\cdot + \int_0^\cdot \dot{h}_\epsilon(s)\mathrm{d}s\Big)$, which solves

$$\begin{aligned}\mathrm{d}\bar{Y}^{\epsilon,h_\epsilon}_t = &\frac{1}{\sqrt{\epsilon}\lambda(\epsilon)}[b_t(Y^{\epsilon,h_\epsilon}_t, \delta_{X^0_t}) - b_t(X^0_t, \delta_{X^0_t})]\mathrm{d}t \\ &+ \frac{1}{\lambda(\epsilon)}\sigma_t(Y^{\epsilon,h_\epsilon}_t, \delta_{X^0_t})\mathrm{d}W_t + \sigma_t(Y^{\epsilon,h_\epsilon}_t, \delta_{X^0_t})\dot{h}_\epsilon(t)\mathrm{d}t,\end{aligned} \tag{5.5.3}$$

where $Y^{\epsilon,h_\epsilon}_t := X^0_t + \sqrt{\epsilon}\lambda(\epsilon)\bar{Y}^{\epsilon,h_\epsilon}_t$.

The following Lemmas play the key roles in the proof of Lemma 5.5.1.

Lemma 5.5.2. *Under assumptions* (**H1**) *and* (**H2**). *Then, for any* $h \in \mathbb{H}$, *equation* (5.4.2) *admits a unique solution* $Y_\cdot^h$ *in* $C([0,T];\mathbb{R}^d)$. *Moreover, for any* $N > 0$, *there exists a constant* $C_{N,T}$ *such that*

$$\sup_{h\in S_N}\Big\{\sup_{0\le t\le T}|Y_t^h|\Big\} \le C_{N,T}. \tag{5.5.4}$$

Proof. By (**H1**) and (**H2**), the coefficients of (5.4.2) satisfy the Lipschitz condition, which implies that equation (5.4.2) admits a unique solution. Moreover, note that the coefficient functions satisfy the linear growth condition, and the fact that $\mathbb{W}_2(\mathscr{L}_{Y_t^h},\delta_0)^2 \le \mathbb{E}|Y_t^h|^2$, we can obtain the estimate (5.5.4) by using the Gronwall inequality. Here we omit the details of the proof. $\square$

Firstly, we prove that the condition (b) of Lemma 2.4.2 holds.

Lemma 5.5.3. *Under assumptions* (**H1**) *and* (**H2**). *Then, for any positive number* $N < \infty$, *the family*

$$K_N := \Big\{\Gamma^0\Big(\int_0^\cdot \dot{h}(s)\mathrm{d}s\Big); h \in S_N\Big\},$$

is compact in $C([0,T];\mathbb{R}^d)$, *where the map* Γ^0 *is defined in Theorem 5.4.1.*

Proof. For any $N < \infty$, the set K_N is compact provided that the compactness of S_N and the continuity of the map Γ^0 from S_N to $C([0,T];\mathbb{R}^d)$. To this end, it suffices to claim that Γ^0 is a continuous map from S_N to $C([0,T];\mathbb{R}^d)$. Let $h_n \to h$ in S_N as $n \to \infty$. Then

$$\begin{aligned}Y_t^{h_n} - Y_t^h &= \int_0^t \nabla_{\{Y_s^{h_n}-Y_s^h\}} b_s(\cdot,\delta_{X_s^0})(X_s^0)\mathrm{d}s + \int_0^t \sigma_s(X_s^0,\delta_{X_s^0})(\dot{h}_n(s)-\dot{h}(s))\mathrm{d}s\\ &=: I_1^n(t) + I_2^n(t).\end{aligned}$$

By (**H2**), (5.3.2) and (5.3.3), it is easy to see that

$$|I_1^n(t)| \le \int_0^t K(s)(1+|X_s^0| + \mathbb{W}_2(\delta_{X_s^0},\delta_0))|Y_s^{h_n} - Y_s^h|\mathrm{d}s.$$

Let $g^n(t)=\int_0^t\sigma_s(X_s^0,\delta_{X_s^0})\dot{h}_n(s)\mathrm{d}s$. By **(H1)**, Lemma 5.3.2, and $h_n,h\in S_N$, we derive that

$$\begin{aligned}|g^n(t)|&\leq\Big(\int_0^t\|\sigma_s(X_s^0,\delta_{X_s^0})\|^2\mathrm{d}s\Big)^{1/2}\Big(\int_0^t|\dot{h}_n(s)|^2\mathrm{d}s\Big)^{1/2}\\&\leq\Big(\int_0^t K^2(s)(1+|X_s^0|+W_2(\delta_{X_s^0},\delta_0))^2\mathrm{d}s\Big)^{1/2}\Big(\int_0^t|\dot{h}_n(s)|^2\mathrm{d}s\Big)^{1/2}\\&<\infty.\end{aligned}$$

Similarly, we see that for any $0\leq t_1\leq t_2\leq T$,

$$\begin{aligned}|g^n(t_2)-g^n(t_1)|&\leq\int_{t_1}^{t_2}\|\sigma_s(X_s^0,\delta_{X_s^0})\||\dot{h}_n(s)|\mathrm{d}s\\&\leq\int_{t_1}^{t_2}K(s)(1+|X_s^0|+W_2(\delta_{X_s^0},\delta_0))|\dot{h}_n(s)|\mathrm{d}s\\&\leq C(T)(t_2-t_1)^{1/2}\Big(\int_{t_1}^{t_2}|\dot{h}_n(s)|^2\mathrm{d}s\Big)^{1/2}\\&\leq C(T,N)(t_2-t_1)^{1/2}.\end{aligned}$$

Hence, the family of functions $\{g^n\}_{n\geq1}$ are equicontinuous in $C([0,T];\mathbb{R}^d)$.

According to the Azelà-Ascoli theorem, $\{g^n\}_{n\geq1}$ is relatively compact in $C([0,T];\mathbb{R}^d)$, let g be any limit point of $\{g^n\}_{n\geq1}$. Noting $h_n\to h$ on S_N, we have

$$\lim_{n\to\infty}\int_0^t\sigma_s(X_s^0,\delta_{X_s^0})\dot{h}_n(s)\mathrm{d}s=\int_0^t\sigma_s(X_s^0,\delta_{X_s^0})\dot{h}(s)\mathrm{d}s,\forall t\in[0,T],$$

that is, $\lim_{n\to\infty}\sup_{t\in[0,T]}|I_2^n(t)|=0$. This, together with (5.3.2), yields that

$$\begin{aligned}&\sup_{0\leq t\leq T}|Y_t^{h_n}-Y_t^h|\\&\leq\int_0^T K(t)(1+|X_t^0|+\mathbb{W}_2(\delta_{X_t^0},\delta_0))|Y_t^{h_n}-Y_t^h|\mathrm{d}t+\sup_{0\leq t\leq T}I_2^n(t),\end{aligned}$$

and by the Gronwall inequality, we arrive at

$$\sup_{0\leq t\leq T}|Y_t^{h_n}-Y_t^h|\leq\exp\Big\{\int_0^T K(t)(1+|X_t^0|+\mathbb{W}_2(\delta_{X_t^0},\delta_0))\mathrm{d}t\Big\}\sup_{0\leq t\leq T}I_2^n(t)$$

$$\le C(T,N)\sup_{0\le t\le T} I_2^n(t)\to 0,\text{as}\quad n\to\infty,$$

which yields that Γ^0 is a continuous map, we therefore complete the proof. $\square$

Before verifying condition (a), we give an estimate for the second moment of $\bar{Y}_t^{\epsilon,h_\epsilon}$.

Lemma 5.5.4. *Assume* (**H1**)*. Then, there exists an* $\epsilon_0\in(0,1)$ *such that for some* C_T*,*

$$\mathbb{E}\Big(\sup_{0\le t\le T}|\bar{Y}_t^{\epsilon,h_\epsilon}|^2\Big)\le C_T,\ \epsilon\in(0,\epsilon_0),\ h_\epsilon\in\mathscr{A}_N. \tag{5.5.5}$$

where $\bar{Y}_\cdot^{\epsilon,h_\epsilon}$ *is defined in* (5.5.3).

Proof. Note that $\bar{Y}_\cdot^{\epsilon,h_\epsilon}$ can be decomposed into the following three parts

$$\begin{aligned}\bar{Y}_t^{\epsilon,h_\epsilon}&=\int_0^t\frac{1}{\sqrt{\epsilon}\lambda(\epsilon)}[b_s(Y_s^{\epsilon,h_\epsilon},\delta_{X_s^0})-b_s(X_s^0,\delta_{X_s^0})]\mathrm{d}s\\&\quad+\int_0^t\frac{1}{\lambda(\epsilon)}\sigma_s(Y_s^{\epsilon,h_\epsilon},\delta_{X_s^0})\mathrm{d}W_s+\int_0^t\sigma_s(Y_s^{\epsilon,h_\epsilon},\delta_{X_s^0})\dot{h}_\epsilon(s)\mathrm{d}s\\&=:\sum_{i=1}^3 J_i^{\epsilon,h_\epsilon}(t).\end{aligned}$$

By (**H1**), we have

$$\mathbb{E}\Big(\sup_{0\le t\le T}|J_1^{\epsilon,h_\epsilon}(t)|^2\Big)\le\frac{TK(T)}{\epsilon\lambda^2(\epsilon)}\int_0^T\mathbb{E}|Y_s^{\epsilon,h_\epsilon}-X_s^0|^2\mathrm{d}\mathrm{d}s\le C_T\int_0^T\mathbb{E}|\bar{Y}_s^{\epsilon,h_\epsilon}|^2\mathrm{d}s.$$

By the BDG inequality, (5.3.2) and (5.3.3), one has

$$\begin{aligned}\mathbb{E}\Big(\sup_{0\le t\le T}|J_2^{\epsilon,h_\epsilon}(t)|^2\Big)&\le\frac{C_T}{\lambda^2(\epsilon)}\int_0^T\mathbb{E}[1+|Y_s^{\epsilon,h_\epsilon}|^2+\mathbb{W}_2(\delta_{X_s^0},\delta_{\mathbf{0}})^2]\mathrm{d}s\\&\le\frac{C_T}{\lambda^2(\epsilon)}\int_0^T[1+\mathbb{E}|Y_s^{\epsilon,h_\epsilon}-X_s^0|^2+\mathbb{E}|X_s^0|^2]\mathrm{d}s\\&\le\frac{C_T}{\lambda^2(\epsilon)}\int_0^T[1+\epsilon\lambda^2(\epsilon)\mathbb{E}|\bar{Y}_s^{\epsilon,h_\epsilon}|^2+\mathbb{E}|X_s^0|^2]\mathrm{d}s\\&\le\frac{C_T}{\lambda^2(\epsilon)}+\epsilon C_T\int_0^T\mathbb{E}|\bar{Y}_s^{\epsilon,h_\epsilon}|^2\mathrm{d}s.\end{aligned}$$

Applying the Hölder inequality, and recalling $h_\epsilon \in \mathscr{A}_N$, we obtain from (5.3.2) and (5.3.3) that

$$
\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le T}|J_3^{\epsilon,h_\epsilon}(t)|^2\Big)\\
&\le C_T\mathbb{E}\int_0^T[1+|Y_s^{\epsilon,h_\epsilon}|^2+\mathbb{W}_2(\delta_{X_s^0},\delta_{\mathbf{0}})^2]|\dot h_\epsilon(s)|^2\mathrm{d}s\\
&\le C_T\Big(1+\Big(\sup_{0\le t\le T}|X_t^0|^2\Big)+\epsilon\lambda^2(\epsilon)\mathbb{E}\Big(\sup_{0\le t\le T}|\bar Y_t^{\epsilon,h_\epsilon}|^2\Big)\Big)\int_0^T|\dot h_\epsilon(s)|^2\mathrm{d}s\\
&\le C_T\Big(1+\epsilon\lambda^2(\epsilon)\mathbb{E}\Big(\sup_{0\le t\le T}|\bar Y_t^{\epsilon,h_\epsilon}|^2\Big)\Big).
\end{aligned}
$$

Thus, the above estimates yield that

$$
\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le t\le T}|\bar Y_t^{\epsilon,h_\epsilon}|^2\Big)\\
&\le C_T\Big(1+\frac{1}{\lambda^2(\epsilon)}+\epsilon\lambda^2(\epsilon)\mathbb{E}\Big(\sup_{0\le t\le T}|\bar Y_t^{\epsilon,h_\epsilon}|^2\Big)+(1+\epsilon)\int_0^T\mathbb{E}|\bar Y_t^{\epsilon,h_\epsilon}|^2\mathrm{d}t\Big).
\end{aligned}
$$

Taking $\epsilon>0$ sufficiently small, such that $C_T\epsilon\lambda^2(\epsilon)\le\frac12$, leads to

$$
\mathbb{E}\Big(\sup_{0\le t\le T}|\bar Y_t^{\epsilon,h_\epsilon}|^2\Big)\le C_T\Big(1+\frac{1}{\lambda^2(\epsilon)}+(1+\epsilon)\int_0^T\mathbb{E}\Big(\sup_{0\le s\le t}|\bar Y_s^{\epsilon,h_\epsilon}|^2\Big)\mathrm{d}t\Big).
$$

The desired assertion follows from Gronwall's inequality and due to the fact that $\frac{1}{\lambda^2(\epsilon)}\to 0$ as $\epsilon\to 0$. □

We are now in the position to verify the condition (a) of Lemma 2.4.2.

Lemma 5.5.5. *Under assumptions (***H1***) and (***H2***), for every fixed* $N\in\mathbb{N}$, *let* $h_\epsilon, h\in\mathscr{A}_N$ *be such that* $h_\epsilon\Rightarrow h$ *as* $\epsilon\to 0$. *Then* $\Gamma^\epsilon\Big(\frac{1}{\lambda(\epsilon)}W_\cdot+\int_0^\cdot\dot h_\epsilon(s)\mathrm{d}s\Big)\Rightarrow\Gamma^0\Big(\int_0^\cdot\dot h(s)\mathrm{d}s\Big)$ *in* $C([0,T];\mathbb{R}^d)$.

Proof. By the Skorokhod representation theorem [10, Theorem 6.7, p70], there exists a probability space $(\tilde\Omega,\tilde{\mathscr{F}},\tilde{\mathscr{F}}_t,\tilde{\mathbb{P}})$, and a Brownian motion $\tilde W$ on this basis, a family of $\tilde{\mathscr{F}}_t$-predictable processes $\{\tilde h_\epsilon;\epsilon>0\},\tilde h$ taking values on $\mathscr{A}_N$, $\tilde{\mathbb{P}}$- a.s., such that the joint law of (h_ϵ, h, W) under $\mathbb{P}$ coincides with the law of $(\tilde{h}_\epsilon, \tilde{h}, \tilde{W})$ under $\tilde{\mathbb{P}}$ and

$$\lim_{\epsilon\to 0}\langle \tilde{h}_\epsilon - \tilde{h}, g\rangle = 0, \forall g \in \mathbb{H}, \tilde{\mathbb{P}} - a.s..$$

Let $\tilde{Y}^{\epsilon,\tilde{h}_\epsilon}$ be the solution of (5.5.3) replacing h_ϵ by $\tilde{h}_\epsilon$ and W by $\tilde{W}$, and $\tilde{Y}^{\tilde{h}}$ be the solution of (5.4.2) replacing h by $\tilde{h}$. Thus, to this end, it suffices to verify

$$\lim_{\epsilon\to 0}\|\tilde{Y}^{\epsilon,\tilde{h}_\epsilon} - \tilde{Y}^{\tilde{h}}\| = 0, \quad \text{in probability}.$$

In the following, we drop off the $\tilde{\cdot}$ in the notation for the sake of simplicity.

Note that $\bar{Y}_t^{\epsilon,h_\epsilon} - Y_t^h$ can be decomposed to the following three parts:

$$\begin{aligned}
&\bar{Y}_t^{\epsilon,h_\epsilon} - Y_t^h \\
&= \left[\frac{1}{\sqrt{\epsilon}\lambda(\epsilon)}\int_0^t [b_s(Y_s^{\epsilon,h_\epsilon}, \delta_{X_s^0}) - b_s(X_s^0, \delta_{X_s^0})]\mathrm{d}s - \int_0^t \nabla_{Y_s^h} b_s(\cdot, \delta_{X_s^0})(X_s^0)\mathrm{d}s\right] \\
&\quad + \int_0^t \left[\sigma_s(Y_s^{\epsilon,h_\epsilon}, \delta_{X_s^0})\dot{h}_\epsilon(s) - \sigma_s(X_s^0, \delta_{X_s^0})\dot{h}(s)\right]\mathrm{d}s + \frac{1}{\lambda(\epsilon)}\int_0^t \sigma_s(Y_s^{\epsilon,h_\epsilon}, \delta_{X_s^0})\mathrm{d}W_s \\
&=: \sum_{i=1}^3 I_i^{\epsilon,h_\epsilon}(t).
\end{aligned}$$

By (**H2**), we have

$$\begin{aligned}
&|I_1^{\epsilon,h_\epsilon}(t)| \\
&= \int_0^t \Big|\int_0^1 \nabla_{\bar{Y}_s^{\epsilon,h_\epsilon}} b_s(\cdot, \delta_{X_s^0})(X_s^0 + r(Y_s^{\epsilon,h_\epsilon} - X_s^0))\mathrm{d}r - \nabla_{Y_s^h} b_s(\cdot, \delta_{X_s^0})(X_s^0)\Big|\mathrm{d}s \\
&\le \int_0^t \Big|\int_0^1 \nabla_{\{\bar{Y}_s^{\epsilon,h_\epsilon} - Y_s^h\}} b_s(\cdot, \delta_{X_s^0})(X_s^0 + r(Y_s^{\epsilon,h_\epsilon} - X_s^0))\mathrm{d}r\Big|\mathrm{d}s \\
&+ \int_0^t \Big|\int_0^1 \nabla_{Y_s^h} b_s(\cdot, \delta_{X_s^0})(X_s^0 + r(Y_s^{\epsilon,h_\epsilon} - X_s^0))\mathrm{d}r - \nabla_{Y_s^h} b_s(\cdot, \delta_{X_s^0})(X_s^0)\Big|\mathrm{d}s \\
&\le K(t)\int_0^t \Big(|\bar{Y}_s^{\epsilon,h_\epsilon} - Y_s^h| + \frac{\sqrt{\epsilon}\lambda(\epsilon)}{2}|Y_s^h||\bar{Y}_s^{\epsilon,h_\epsilon}|\Big)\mathrm{d}s.
\end{aligned}$$

By (5.5.4) and (5.5.5), it follows that

$$\mathbb{E}\Big(\sup_{0\le t\le T}|I_1^{\epsilon,h_\epsilon}(t)|^2\Big) \lesssim \epsilon\lambda^2(\epsilon) + \int_0^T \mathbb{E}|\bar{Y}_s^{\epsilon,h_\epsilon} - Y_s^h|^2\mathrm{d}s.$$

By (**H1**) and (5.3.3), it follows that

$$
\begin{aligned}
&|I_2^{\epsilon,h_\epsilon}(t)| \\
&\leq \Big|\int_0^t \Big[\sigma_s(Y_s^{\epsilon,h_\epsilon},\delta_{X_s^0})-\sigma_s(X_s^0,\delta_{X_s^0})\Big]\dot{h}_\epsilon(s)\mathrm{d}s\Big| + \Big|\int_0^t \sigma_s(X_s^0,\delta_{X_s^0})(\dot{h}_\epsilon(s)-\dot{h}(s))\mathrm{d}s\Big| \\
&\leq \int_0^t K(s)|Y_s^{\epsilon,h_\epsilon}-X_s^0||\dot{h}_\epsilon(s)|\mathrm{d}s + \int_0^t |\sigma_s(X_s^0,\delta_{X_s^0})(\dot{h}_\epsilon(s)-\dot{h}(s))|\mathrm{d}s \\
&\leq \sqrt{\epsilon}\lambda(\epsilon)\int_0^t K(s)|\bar{Y}_s^{\epsilon,h_\epsilon}||\dot{h}_\epsilon(s)|\mathrm{d}s + \int_0^t K(s)(1+|X_s^0|)|\dot{h}_\epsilon(s)-\dot{h}(s)|\mathrm{d}s,
\end{aligned}
$$

thus, by Hölder's inequality and (5.3.2), it follows that

$$
\mathbb{E}\Big(\sup_{0\leq t\leq T}|I_2^{\epsilon,h_\epsilon}(t)|^2\Big) \lesssim \epsilon\lambda^2(\epsilon) + \int_0^T \mathbb{E}|\dot{h}_\epsilon(s)-\dot{h}(s)|^2\mathrm{d}s.
$$

By the BDG inequality, (5.3.3) and (5.5.4), we arrive at

$$
\begin{aligned}
&\mathbb{E}\Big(\sup_{0\leq t\leq T}|I_3^{\epsilon,h_\epsilon}(t)|^2\Big) \\
&\leq \frac{1}{\lambda^2(\epsilon)}\int_0^T \mathbb{E}\Big(\|\sigma_s(Y_s^{\epsilon,h_\epsilon},\delta_{X_s^0})-\sigma_s(X_s^0,\delta_{X_s^0})\|^2 + \|\sigma_s(X_s^0,\delta_{X_s^0})\|^2\Big)\mathrm{d}s \\
&\leq \frac{1}{\lambda^2(\epsilon)} + \epsilon\int_0^T \mathbb{E}|\bar{Y}_s^{\epsilon,h_\epsilon}|^2\mathrm{d}s.
\end{aligned}
$$

Taking the above estimates into consideration, it follows that

$$
\begin{aligned}
&\mathbb{E}\Big(\sup_{0\leq t\leq T}|\bar{Y}_t^{\epsilon,h_\epsilon}-Y_t^h|^2\Big) \\
&\leq \frac{1}{\lambda^2(\epsilon)} + \epsilon(\lambda^2(\epsilon)+1) + \int_0^T \mathbb{E}|\dot{h}_\epsilon(s)-\dot{h}(s)|^2\mathrm{d}s + \int_0^T \mathbb{E}|\bar{Y}_s^{\epsilon,h_\epsilon}-Y_s^h|^2\mathrm{d}s,
\end{aligned}
$$

thus, the desired assertion follows from the Gronwall inequality and taking $\epsilon\to 0$. □

Proof of Lemma 5.5.1

The conclusion of Lemma 5.5.1 follows from Lemma 2.4.2, and from Lemmas 5.5.3 and 5.5.5.

5.5.2 $\bar{X}^\epsilon$ and $\bar{Y}^\epsilon$ are exponentially equivalent

The following lemma shows that $\bar{X}^\epsilon$ and $\bar{Y}^\epsilon$ are exponentially equivalent.

Lemma 5.5.6. *For any $\delta > 0$, we have*

$$\limsup_{\epsilon\to 0} \epsilon \log \Big(\mathbb{P}\Big\{ \sup_{0\le t\le T} |\bar{X}_t^\epsilon - \bar{Y}_t^\epsilon| \ge \delta \Big\} \Big) = -\infty. \tag{5.5.6}$$

The proof of Lemma 5.5.6 is based on the following lemma, (for more details, please refer to [25, Lemma 5.6.18]).

Lemma 5.5.7. *Let b_t, σ_t be progressively measurable processes, $(w_t)_{t\ge 0}$ is a d-dimensional Brownian motion, and let*

$$\mathrm{d}z_t = b_t \mathrm{d}t + \sqrt{\epsilon}\sigma_t \mathrm{d}w_t,\ t \ge 0,$$

where z_0 is deterministic. Let $\tau_1 \in [0,1]$ be a stopping time with respect to the filtration of $\{w_t, t \in [0,1]\}$. Suppose that the coefficients of the diffusion matrix σ are uniformly bounded, and for some constants M, B, ρ and any $t \in [0, \tau_1]$,

$$|\sigma_t| \le M(\rho^2 + |z_t|^2)^{1/2}, \quad |b_t| \le B(\rho^2 + |z_t|^2)^{1/2}.$$

Then for any $\delta > 0$ and any $\epsilon \le 1$,

$$\epsilon \log \mathbb{P}\Big(\sup_{t\in[0,\tau_1]} |z_t| \ge \delta \Big) \le K + \log\Big(\frac{\rho^2 + |z_0|^2}{\rho^2 + \delta^2} \Big),$$

where $K = 2B + M^2(2 + d)$.

Now, we are in position to finish the **Proof of Lemma 5.5.6**.

Proof. Without loss of generality, we may choose $R > 0$ such that the initial data x is in the ball $B_{R+1}(0)$ (with center 0 and radius $R+1$). We also assume that X_t^0 does not leave this ball up to time T. We define the stopping time $\tau_R' := \inf \Big\{ t : t \ge 0 \Big| |\bar{X}_t^\epsilon| \vee |\bar{Y}_t^\epsilon| \ge R+1 \Big\}$, then we denote by $\tau_R = \min\{T, \tau_R'\}$.

Subsequently, we consider $\bar{z}_t := \bar{X}_t^\epsilon - \bar{Y}_t^\epsilon$, the new process satisfies the following equation

$$\mathrm{d}\bar{z}_t = \int_0^t b_s \mathrm{d}s + \sqrt{\epsilon}\int_0^t \sigma_s \mathrm{d}W_s,\ \bar{z}_0 = \mathbf{0}, \tag{5.5.7}$$

where

$$b_t := \frac{b_t(X_t^\epsilon, \mathscr{L}_{X_t^\epsilon}) - b_t(\tilde{Y}_t^\epsilon, \delta_{X_t^0})}{\sqrt{\epsilon}\lambda(\epsilon)}, \quad \sigma_t := \frac{\sigma_t(X_t^\epsilon, \mathscr{L}_{X_t^\epsilon}) - \sigma_t(\tilde{Y}_t^\epsilon, \delta_{X_t^0})}{\sqrt{\epsilon}\lambda(\epsilon)}.$$

Note that both b_t and σ_t are progressively measurable processes. Assume $t \leq \tau_R$, we then derive from (5.2.9) that

$$\begin{aligned}|b_t| &= \frac{|b_t(X_t^\epsilon, \mathscr{L}_{X_t^\epsilon}) - b_t(X_t^\epsilon, \delta_{X_t^0}) + b_t(X_t^\epsilon, \delta_{X_t^0}) - b_t(\tilde{Y}_t^\epsilon, \delta_{X_t^0})|}{\sqrt{\epsilon}\lambda(\epsilon)} \\ &\leq \frac{K(t)W_2(\mathscr{L}_{X_t^\epsilon}, \delta_{X_t^0})}{\sqrt{\epsilon}\lambda(\epsilon)} + \frac{K(t)|X_t^\epsilon - \tilde{Y}_t^\epsilon|}{\sqrt{\epsilon}\lambda(\epsilon)} \\ &\leq K(t)(\rho^2(\epsilon) + |\bar{z}_t|^2)^{1/2},\end{aligned}$$

where $\rho^2(\epsilon) = \sup_{0\leq t\leq T}\mathbb{E}|\bar{X}_t^\epsilon|^2$. In the same vein, we have

$$|\sigma_t| \leq K(t)(\rho^2(\epsilon) + |\bar{z}_t|^2)^{1/2}.$$

Note that $\bar{z}_0 = \mathbf{0}$, for any δ, ρ^ϵ and for any small enough ϵ, we derive from Lemma 5.5.7 that

$$\epsilon \log \mathbb{P}\Big(\sup_{t\in[0,\tau_R]} |\bar{z}_t| \geq \delta\Big) \leq KT + \log\Big(\frac{\rho^2(\epsilon)}{\rho^2(\epsilon) + \delta^2}\Big).$$

In the same way as in the proof of (5.3.6), one can show that $\rho^2(\epsilon)$ converges to 0 as $\epsilon \to 0$. Hence, we deduce that

$$\limsup_{\epsilon\to 0} \epsilon \log \mathbb{P}\Big(\sup_{t\in[0,\tau_R]} |\bar{z}_t| \geq \delta\Big) = -\infty. \tag{5.5.8}$$

Now, since

$$\{\|\bar{X}^\epsilon - \bar{Y}^\epsilon\|_\infty \geq \delta\} \subset \{\tau_R \leq T\} \cup \Big\{\sup_{0\leq t\leq \tau_R} |\bar{X}_t^\epsilon - \bar{Y}_t^\epsilon| \geq \delta\Big\},$$

we can conclude as long as we show that

$$\lim_{R\to\infty}\limsup_{\epsilon\to 0}\epsilon\log\Big(\mathbb{P}\{\tau_R<T\}\Big)=-\infty.$$

Define $\eta_R := \{t : t \geq 0, |\bar{Y}^\epsilon_t| \geq R\}$, i.e. the first time of $\bar{Y}^\epsilon$ exits from the ball $B_R(0)$ (with center 0 and radius R).

Letting $\tau_R < T$, we then have $\{|\bar{X}^\epsilon_{\tau_R}| \vee |\bar{Y}^\epsilon_{\tau_R}| = R+1\}$, which yields the following two cases:

(i) If $|\bar{Y}^\epsilon_{\tau_R}| = R+1$, then we have immediately $\eta_R < T$. This implies that $\mathbb{P}\{\tau_R < T\} \leq \mathbb{P}\{\eta_R < T\}$.

(ii) If $|\bar{X}^\epsilon_{\tau_R}| = R+1$, one can derive that

$$\begin{aligned}
&P\{\tau_R<T\}\leq P\{|\bar{X}^\epsilon_{\tau_R}|=R+1\}\\
&=P\Big\{\sup_{t\in[0,\tau_R]}|\bar{z}_t|\geq\frac{1}{2},|\bar{X}^\epsilon_{\tau_R}|=R+1\Big\}+P\Big\{\sup_{t\in[0,\tau_R]}|\bar{z}_t|<\frac{1}{2},|\bar{X}^\epsilon_{\tau_R}|=R+1\Big\}\\
&\leq P\Big\{\sup_{t\in[0,\tau_R]}|\bar{z}_t|\geq\frac{1}{2}\Big\}+P\{\eta_R<T\}.
\end{aligned}$$

To finish the proof, it is sufficient to prove by (5.5.8) that the probability of $\bar{Y}^\epsilon$ exits the ball $B_R(0)$ is very small as ϵ goes to zero, i.e.

$$\lim_{R\to\infty}\limsup_{\epsilon\to 0}\epsilon\log\Big(\mathbb{P}\{\eta_R<T\}\Big)=-\infty.$$

Recall that $\bar{Y}^\epsilon$ satisfies an LDP for the uniform norm with good rate function $I(g)$ given in (5.4.1). Then, for any closed set $F \subset C([0,T];\mathbb{R}^d)$ we have

$$\limsup_{\epsilon\to 0}\epsilon\log\mathbb{P}\{\bar{Y}^\epsilon\in F\}\leq-\inf_{g\in F}I(g).$$

As a consequence,

$$\begin{aligned}
&\limsup_{\epsilon\to 0}\epsilon\log\Big(\mathbb{P}\{\eta_R<T\}\Big)=\limsup_{\epsilon\to 0}\epsilon\log\Big(\mathbb{P}\Big\{\sup_{0\leq t\leq T}|\bar{Y}^\epsilon_t|\geq R\Big\}\Big)\\
&\leq-\inf_{\{h\in\mathbb{H};g=\Gamma^0(\int_0^\cdot\dot{h}(s)\mathrm{d}s),\|g\|_\infty\geq R\}}\frac{1}{2}\int_0^T|\dot{h}(s)|^2\mathrm{d}s.
\end{aligned}$$

We remark that the infimum of $I(g)$ on the set of paths exiting from the ball $B_R(0)$ goes to infinity as R goes to infinity.

By (**H1**) and (5.3.2), we obtain that

$$\begin{aligned}|g(t)| &\leq \int_0^t |\nabla_{g(s)} b_s(\cdot, \delta_{X_s^0})(X_s^0) + \sigma_s(X_s^0, \delta_{X_s^0})\dot{h}(s)|\mathrm{d}s \\ &\leq \int_0^t |K(s)(|g(s)| + (1 + |X_s^0|)|\dot{h}(s)|)\mathrm{d}s \\ &\leq C_t\Big(\int_0^t |g(s)|\mathrm{d}s + \Big(\int_0^t |\dot{h}(s)|^2\mathrm{d}s\Big)^{1/2}\Big),\end{aligned}$$

and by the Gronwall lemma, we have

$$|g(t)| \leq C_t\Big(\int_0^t |\dot{h}(s)|^2\mathrm{d}s\Big)^{1/2} < \infty.$$

By taking $R \to \infty$, it yields that $\{h \in \mathbb{H}; g = \Gamma^0(\int_0^\cdot \dot{h}(s)\mathrm{d}s), \|g\|_\infty \geq R\} = \emptyset$, which implies $I(g) = -\infty$. That is, $\bar{X}^\epsilon$ and $\bar{Y}^\epsilon$ are exponentially equivalent.

□

Proof of Theorem 5.4.1 The conclusion of Theorem 5.4.1 follows from Lemma 5.5.1 and Lemma 5.5.6.

5.6 Illustrative Example

In this section, we give an illustrate example.

Example 5.6.1. *For any $g \in C_b^2(\mathbb{R}^d)$, define the function of μ as $\mu \mapsto \mu(g) := \int_{\mathbb{R}^d} g\mathrm{d}\mu$. Consider the following MV-SDE on $\mathbb{R}^d$:*

$$\mathrm{d}X_t^\epsilon = \{X_t^\epsilon + (\mathscr{L}_{X_t^\epsilon}(g))^2\}\mathrm{d}t + \sqrt{\epsilon}\{X_t^\epsilon + \mathscr{L}_{X_t^\epsilon}(g))\}\mathrm{d}W_t, \tag{5.6.1}$$

with the initial value X_0^ϵ. When $\epsilon \to 0$, we obtain the following ordinary differential equation:

$$\mathrm{d}X_t^0 = \{X_t^0 + (\delta_{X_t^0}(g))^2\}\mathrm{d}t. \tag{5.6.2}$$

We now check that the coefficients of (5.6.1) *satisfy (***H1***) and (***H2***).*

Letting $b(x,\mu) = x + (\mu(g))^2$*, we have* $\nabla b(\cdot,\mu)(x) = I$*, where* I *is the* $d \times d$ *identity matrix. It is easy to check that (***H1***) and (***H2***) hold for the spatial component of* b*. Now, we check that (***H1***) and (***H2***) also hold for the measure component of* b*.*

Firstly, we verify the condition (H1). By the Taylor expansion, one gets that

$$\begin{aligned}
&\lim_{\|\phi\|_{T_{\mu,2}}\to 0} \frac{1}{\|\phi\|_{T_{\mu,2}}}\Big|\mu\circ(Id+\phi)^{-1}(g) - \mu(g) - \langle \nabla g, \phi(x)\rangle\Big| \\
&= \lim_{\|\phi\|_{T_{\mu,2}}\to 0} \frac{1}{\|\phi\|_{T_{\mu,2}}}\Big|\int_{\mathbb{R}^d}\{g(x+\phi(x)) - g(x) - \langle\nabla g, \phi(x)\rangle\}\mu(\mathrm{d}x)\Big| \\
&\le \lim_{\|\phi\|_{T_{\mu,2}}\to 0} \frac{\|\nabla^2 g\|_\infty}{2\|\phi\|_{T_{\mu,2}}}\Big|\int_{\mathbb{R}^d}|\phi(x)|^2\mu(\mathrm{d}x) \\
&\le \lim_{\|\phi\|_{T_{\mu,2}}\to 0} \|\nabla^2 g\|_\infty \|\phi\|_{T_{\mu,2}} = 0.
\end{aligned}$$

That is, $D^L\mu(g) = \nabla g$*. Similarly, we can show that* $D^L b(x,\cdot)(\mu) = 2\mu(g)\nabla g$*. This, together with* $g \in C_b^2(\mathbb{R}^d)$*, yields that* $\|D^L b(x,\cdot)(\mathscr{L}_x(g))\|_{T_{\mu,2}} \le K$*, where* $K = 2\max\{\sup_{x\in\mathbb{R}^d}|g(x)|, \sup_{x\in\mathbb{R}^d}\|\nabla g(x)\|\}$*.*

*We now check the condition (***H2***). For* $X, Y, \phi \in L^2(\Omega\to\mathbb{R}^d, \mathbb{P})$*,*

$$\begin{aligned}
&|\mathbb{E}\langle D^L b(x,\cdot)(\mathscr{L}_X(g))(X), \phi\rangle - \mathbb{E}\langle D^L b(x,\cdot)(\mathscr{L}_Y(g))(Y), \phi\rangle| \\
&= \Big|\mathbb{E}\langle 2(\mathscr{L}_X(g)\nabla g)(X), \phi\rangle - \mathbb{E}\langle 2(\mathscr{L}_Y(g)\nabla g)(Y), \phi\rangle\Big| \\
&\le 2(\mathbb{E}|\phi|^2)^{1/2}(\mathbb{E}|(\mathscr{L}_X(g)\nabla g)(X) - (\mathscr{L}_Y(g)\nabla g)(Y)|^2)^{1/2} \\
&\le 4(\mathbb{E}|\phi|^2)^{1/2}\big\{(\mathbb{E}|\mathscr{L}_X(g) - \mathscr{L}_Y(g)\nabla g(X)|^2)^{1/2} + (\mathbb{E}|\mathscr{L}_Y(g)(\nabla g(X) - \nabla g(Y))|^2)^{1/2}\big\} \\
&\le C(\mathbb{E}|\phi|^2)^{1/2}(\mathbb{E}|X-Y|^2)^{1/2}),
\end{aligned}$$

where we have used $\|D^L\mu(g)\|_{T_{\mu,2}} < \infty$ *in the last inequality, .*

Similarly, we can also check that σ *satisfies (***H1***). By theorem 5.3.1, we then obtain that* Z_t *satisfies*

$$\mathrm{d}Z_t = Z_t\mathrm{d}t + \mathbb{E}\langle 2(\delta^0_{X_t}(g)\nabla g)(X_t^0), Z_t\rangle\mathrm{d}t + \{X_t^0 + (\mathscr{L}_{X_t^0}(g))\}\mathrm{d}W_t.$$

Chapter 6

Bismut formula of Lions derivative for Mckean-Vlasov stochastic differential equations driven by fractional Brownian motion

In this chapter, we investigate the Bismut formula of L-derivative for MV-SDEs driven by fBm in view of the Malliavin analysis method.

In Section 6.1, we show the well-posedness of MV-SDEs driven by fBm under the Lipschitz condition of coefficients.

In Section 6.2, we give the results of partial derivative in initial value and Malliavin derivative of MV-SDEs driven by fBm, that is, Propositions 6.2.2 and 6.2.3.

In Section 6.3, we show a general result of Bismut formula of L-derivative for MV-SDEs driven by fBm (Theorem 6.3.1), and Subsection 6.3.1 is devoted to its proof.

In Section 6.4, we apply the general result (Theorem 6.3.1) to the non-degenerate case. More precisely, Subsection 6.4.1 is devoted to the explicit assumptions on the coefficients and the main result. Subsection 6.4.2 is devoted to the proof of Theorem 6.3.1.

6.1 Well-posedness of McKean-Vlasov stochastic differential equations

Given a complete filtration probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ and $(B_t^H)_{t\geq 0}$ is a d-dimensional fBm with Hurst parameter $H \in (\frac{1}{2}, 1)$. Consider the following MV-SDE:

$$\mathrm{d}X_t = b(t, X_t, \mathscr{L}_{X_t})\mathrm{d}t + \sigma(t, \mathscr{L}_{X_t})\mathrm{d}B_t^H, \quad X_0 = \xi, \tag{6.1.1}$$

where $b : \Omega \times [0, T] \times \mathbb{R}^d \times \mathscr{P}_\theta(\mathbb{R}^d) \to \mathbb{R}^d, \sigma : [0, T] \times \mathscr{P}_\theta(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^d$ and $\theta \in [1, \infty)$.

To obtain the existence and uniqueness of (6.1.1), we give the assumption for (b, σ) as follows:

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

$$|b(t, x, \mu) - b(t, y, \nu)| \leq K(t)(|x - y| + \mathbb{W}_\theta(\mu, \nu)),$$
$$\|\sigma(t, \mu) - \sigma(t, \nu)\| \leq K(t)\mathbb{W}_\theta(\mu, \nu),$$

and

$$|b(t, 0, \delta_0)| + \|\sigma(t, \delta_0)\| \leq K(t).$$

For any $p \geq 1$, let $\mathcal{S}^p([0,T])$ be the space of $\mathbb{R}^d$-valued, continuous $\mathscr{F}$-adapted processes ψ on $[0,T]$ satisfying

$$\|\psi\|_{\mathcal{S}^p} := \left(\mathbb{E} \sup_{t\in[0,T]} |\psi(t)|^p\right)^{1/p} < \infty,$$

and let the letter C with or without indices denotes generic constants, whose value may change from line to line.

Definition 6.1.1. A stochastic process $X = (X_t)_{0\leq t\leq T}$ on $\mathbb{R}^d$ is called a solution of (6.1.1), if $X \in \mathcal{S}^p([0,T])$ and $\mathbb{P}$-a.s.,

$$X_t = \xi + \int_0^t b(s, X_s, \mathscr{L}_{X_s})\mathrm{d}s + \int_0^t \sigma(s, \mathscr{L}_{X_s})\mathrm{d}B_s^H, \quad t \in [0,T].$$

Remark 6.1.1. Observing that the diffusion term $\sigma(\cdot, \mathscr{L}_{X_\cdot})$ is a deterministic function, then $\int_0^t \sigma(s, \mathscr{L}_{X_s})\mathrm{d}B_s^H$ is regarded as a Wiener integral with respect to fBm.

Theorem 6.1.1. *Suppose that (H) holds and $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$. Then the equation* (6.1.1) *has a unique solution $X \in \mathcal{S}^p([0,T])$ with any $p \geq \theta$ and $p > 1/H$.*

To finish the proof of this Theorem, we prepare the following Lemmas. The following lemma presents the Hardy-Littlewood inequality (see, e.g., [91, Theroem 1]).

Lemma 6.1.2. *Let $1 < \tilde{p} < \tilde{q} < \infty$ and $\frac{1}{\tilde{q}} = \frac{1}{\tilde{p}} - \alpha$. If $f : \mathbb{R}_+ \to \mathbb{R}$ belongs to $L^{\tilde{p}}(0,\infty)$, then $I_{0+}^\alpha f(x)$ converges absolutely for almost every x, and moreover*

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

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

The below lemma is due to [63, Theorem 1].

Lemma 6.1.3. *Let Z be a fractional Brownian motion with $H > \frac{1}{2}$. We have the inclusion: for every $T < \infty$, there exists a constant $C(H, r)$ such that for every $r > 0$, for every a, b with $0 \leq a < b < \infty$ we have*

$$\mathbb{E}\Big(\Big|\int_a^b f(u)\mathrm{d}Z_u\Big|^r\Big) \leq C(H,r)\|f\|^r_{L^{1/H}(a,b)}, \tag{6.1.2}$$

and

$$\mathbb{E}\Big(\Big|\int_a^b f(u)\mathrm{d}Z_u \int_a^b g(u)\mathrm{d}Z_u\Big|^r\Big) \leq C(H,r)\|f\|^r_{L^{1/H}(a,b)}\|g\|^r_{L^{1/H}(a,b)}. \tag{6.1.3}$$

Now, we are in position to finish the **Proof of Theorem 6.1.1**.

Proof. Define recursively $(X^n)_{n\geq 1}$ as follows: $X^0_t = \xi,\ t \in [0,T]$ and for each $n \geq 1$,

$$X^n_t = \xi + \int_0^t b(s, X^{n-1}_s, \mathscr{L}_{X^{n-1}_s})\mathrm{d}s + \int_0^t \sigma(s, \mathscr{L}_{X^{n-1}_s})\mathrm{d}B^H_s, \quad t \in [0,T].$$

The rest of the proof will be divided into three steps.

Step 1. Claim: For any $p \geq \theta$ and $p > 1/H$, if $\mathbb{E}(\sup_{t\in[0,T]}|X^n_t|^p) < \infty$, then there holds $\mathbb{E}(\sup_{t\in[0,T]}|X^{n+1}_t|^p) < \infty$. Owing to the Hölder inequality and (H), we have for any $p \geq \theta$,

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{t\in[0,T]}|X^{n+1}_t|^p\Big)\\
&\leq 3^{p-1}\mathbb{E}|\xi|^p + 3^{p-1}\mathbb{E}\left|\int_0^T b(s, X^n_s, \mathscr{L}_{X^n_s})\mathrm{d}s\right|^p\\
&\quad + 3^{p-1}\mathbb{E}\left(\sup_{t\in[0,T]}\left|\int_0^t \sigma(s, \mathscr{L}_{X^n_s})\mathrm{d}B^H_s\right|^p\right)\\
&\leq 3^{p-1}\mathbb{E}|\xi|^p + (3T)^{p-1}\mathbb{E}\int_0^T K^p(s)(1 + |X^n_s| + \mathbb{W}_\theta(\mathscr{L}_{X^n_s}, \delta_0))^p\mathrm{d}s\\
&\quad + 3^{p-1}\mathbb{E}\left(\sup_{t\in[0,T]}\left|\int_0^t \sigma(s, \mathscr{L}_{X^n_s})\mathrm{d}B^H_s\right|^p\right)\\
&\leq 3^{p-1}\mathbb{E}|\xi|^p + 3^{2(p-1)}(TK(T))^p\left(1 + 2\mathbb{E}\Big(\sup_{t\in[0,T]}|X^n_t|^p\Big)\right)
\end{aligned}$$

$$+3^{p-1}\mathbb{E}\left(\sup_{t\in[0,T]}\left|\int_0^t \sigma(s,\mathscr{L}_{X_s^n})\mathrm{d}B_s^H\right|^p\right). \tag{6.1.4}$$

Next, we shall provide an estimate for the last term of the right-hand side of (6.1.4), whose argument is partially borrowed from [1, Theorem 4].

We take λ satisfying $1-H<\lambda<1-1/p$ because $pH>1$. Using the fact that $\int_s^t(t-r)^{-\lambda}(r-s)^{\lambda-1}\mathrm{d}r=C_\lambda$, the stochastic Fubini theorem and the Hölder inequality, we get

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\in[0,T]}\left|\int_0^t \sigma(s,\mathscr{L}_{X_s^n})\mathrm{d}B_s^H\right|^p\right)\\
&=C_\lambda^{-p}\mathbb{E}\left(\sup_{t\in[0,T]}\left|\int_0^t\left(\int_s^t(t-r)^{-\lambda}(r-s)^{\lambda-1}\mathrm{d}r\right)\sigma(s,\mathscr{L}_{X_s^n})\mathrm{d}B_s^H\right|^p\right)\\
&=C_\lambda^{-p}\mathbb{E}\left(\sup_{t\in[0,T]}\left|\int_0^t(t-r)^{-\lambda}\left(\int_0^r(r-s)^{\lambda-1}\sigma(s,\mathscr{L}_{X_s^n})\mathrm{d}B_s^H\right)\mathrm{d}r\right|^p\right)\\
&\le\frac{C_\lambda^{-p}}{(p-1-\lambda p)^{p-1}}\mathbb{E}\left(\sup_{t\in[0,T]}t^{p-1-\lambda p}\int_0^t\left|\int_0^r(r-s)^{\lambda-1}\sigma(s,\mathscr{L}_{X_s^n})\mathrm{d}B_s^H\right|^p\mathrm{d}r\right)\\
&\le\frac{C_\lambda^{-p}}{(p-1-\lambda p)^{p-1}}T^{p-1-\lambda p}\int_0^T\mathbb{E}\left|\int_0^r(r-s)^{\lambda-1}\sigma(s,\mathscr{L}_{X_s^n})\mathrm{d}B_s^H\right|^p\mathrm{d}r,
\end{aligned} \tag{6.1.5}$$

where we have used the condition $\lambda<1-1/p$ in the first inequality. Notice that for each $r\in[0,T]$, $\int_0^r(r-s)^{\lambda-1}\sigma(s,\mathscr{L}_{X_s^n})\mathrm{d}B_s^H$ is a centered Gaussian random variable. Then by the Kahane-Khintchine formula, we obtain that there exists a constant $C_p>0$ such that

$$\begin{aligned}
&\mathbb{E}\left|\int_0^r(r-s)^{\lambda-1}\sigma(s,\mathscr{L}_{X_s^n})\mathrm{d}B_s^H\right|^p\\
&\le C_p\left(\mathbb{E}\left|\int_0^r(r-s)^{\lambda-1}\sigma(s,\mathscr{L}_{X_s^n})\mathrm{d}B_s^H\right|^2\right)^{\frac{p}{2}}\\
&\le C_p\left(\int_0^r\int_0^r(r-u)^{\lambda-1}\|\sigma(u,\mathscr{L}_{X_u^n})\|(r-v)^{\lambda-1}\|\sigma(v,\mathscr{L}_{X_v^n})\||u-v|^{2H-2}\mathrm{d}u\mathrm{d}v\right)^{\frac{p}{2}}\\
&\le C_{p,H}\left(\int_0^r(r-s)^{\frac{\lambda-1}{H}}\|\sigma(s,\mathscr{L}_{X_s^n})\|^{\frac{1}{H}}\mathrm{d}s\right)^{pH},
\end{aligned} \tag{6.1.6}$$

where the last inequality is due to the argument of Lemma 6.1.3.

Substituting (6.1.6) into (6.1.5) and using the condition $1-H<\lambda$ and Lemma 6.1.2 with $\tilde{q}=pH$ and $\alpha=1-\frac{1-\lambda}{H}$ (imply $\tilde{p}=\frac{pH}{p(\lambda+H-1)+1}$), we have

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\in[0,T]}\left|\int_0^t \sigma(s,\mathscr{L}_{X_s^n})\mathrm{d}B_s^H\right|^p\right)\\
&\le C_{\lambda,p,H}T^{p-1-\lambda p}\int_0^T\left(\int_0^r (r-s)^{\frac{\lambda-1}{H}}\|\sigma(s,\mathscr{L}_{X_s^n})\|^{\frac{1}{H}}\mathrm{d}s\right)^{pH}\mathrm{d}r\\
&\le C_{\lambda,p,H}T^{p-1-\lambda p}\left(\int_0^T \|\sigma(r,\mathscr{L}_{X_r^n})\|^{\frac{p}{p(\lambda+H-1)+1}}\mathrm{d}r\right)^{p(\lambda+H-1)+1}\\
&\le C_{\lambda,p,H}T^{pH-1}\int_0^T \|\sigma(s,\mathscr{L}_{X_s^n})\|^p\mathrm{d}s, \qquad (6.1.7)
\end{aligned}$$

where we used the Hölder inequality in the last inequality, and remark that $C_{\lambda,p,H}$ above may depend only on p and H by choosing proper λ.

Observe that, by (H) and $p\ge\theta$ we have

$$\begin{aligned}
\int_0^T \|\sigma(s,\mathscr{L}_{X_s^n})\|^p\mathrm{d}s &\le \int_0^T K^p(s)\left(1+\mathbb{W}_\theta(\mathscr{L}_{X_s^n},\delta_0)\right)^p\mathrm{d}s\\
&\le 2^{p-1}K^p(T)T\left(1+\mathbb{E}\Big(\sup_{t\in[0,T]}|X_t^n|^p\Big)\right).
\end{aligned}$$

Then, plugging this into (6.1.7) yields

$$\mathbb{E}\left(\sup_{t\in[0,T]}\left|\int_0^t \sigma(s,\mathscr{L}_{X_s^n})\mathrm{d}B_s^H\right|^p\right)\le C_{p,H}K^p(T)T^{pH}\left(1+\mathbb{E}\Big(\sup_{t\in[0,T]}|X_t^n|^p\Big)\right).$$

Combining this with (6.1.4) and the assumption that $\mathbb{E}\big(\sup_{t\in[0,T]}|X_t^n|^p\big)<\infty$ yields the desired claim.

Step 2. Existence. To this end, we shall prove the convergence of X^n in $\mathcal{S}^p([0,T])$ with any $p>1/H$ and $p\ge\theta$. For any $t\in[0,T]$, we get

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{s\in[0,t]}|X_s^n-X_s^{n-1}|^p\Big) \qquad (6.1.8)\\
&\le 2^{p-1}\mathbb{E}\Big(\sup_{s\in[0,t]}\Big|\int_0^s\Big(b(r,X_r^{n-1},\mathscr{L}_{X_r^{n-1}})-b(r,X_r^{n-2},\mathscr{L}_{X_r^{n-2}})\Big)\mathrm{d}r\Big|^p\Big)
\end{aligned}$$

$$+ 2^{p-1}\mathbb{E}\Big(\sup_{s\in[0,t]}\Big|\int_0^s\Big(\sigma(r,\mathscr{L}_{X_r^{n-1}})-\sigma(r,\mathscr{L}_{X_r^{n-2}})\Big)\mathrm{d}B_r^H\Big|^p\Big)$$

$$=: 2^{p-1}I_1(t)+2^{p-1}I_2(t).$$

For the term $I_1(t)$, from (H) and $p \geq \theta$ we obtain

$$\begin{aligned}
I_1(t) &\leq t^{p-1}\mathbb{E}\int_0^t \big|b(r,X_r^{n-1},\mathscr{L}_{X_r^{n-1}})-b(r,X_r^{n-2},\mathscr{L}_{X_r^{n-2}})\big|^p\,\mathrm{d}r\\
&\leq t^{p-1}\mathbb{E}\int_0^t\big[K(r)\left(|X_r^{n-1}-X_r^{n-2}|+\mathbb{W}_\theta(\mathscr{L}_{X_r^{n-1}},\mathscr{L}_{X_r^{n-2}})\right)\big]^p\,\mathrm{d}r\\
&\leq 2^{p-1}K^p(t)t^{p-1}\mathbb{E}\int_0^t\left(|X_r^{n-1}-X_r^{n-2}|^p+\mathbb{E}|X_r^{n-1}-X_r^{n-2}|^p\right)\mathrm{d}r\\
&\leq 2^pK^p(t)t^{p-1}\int_0^t\mathbb{E}\Big(\sup_{u\in[0,r]}|X_u^{n-1}-X_u^{n-2}|^p\Big)\mathrm{d}r. \qquad (6.1.9)
\end{aligned}$$

As for the term $I_2(t)$, owing to $p > 1/H$ and $p \geq \theta$, (6.1.7) and (H) we have

$$\begin{aligned}
I_2(t) &= \mathbb{E}\Big(\sup_{s\in[0,t]}\Big|\int_0^s\Big(\sigma(r,\mathscr{L}_{X_r^{n-1}})-\sigma(r,\mathscr{L}_{X_r^{n-2}})\Big)\mathrm{d}B_r^H\Big|^p\Big)\\
&\leq C_{p,H}t^{pH-1}\int_0^t\|\sigma(r,\mathscr{L}_{X_r^{n-1}})-\sigma(r,\mathscr{L}_{X_r^{n-2}})\|^p\mathrm{d}r\\
&\leq C_{p,H}t^{pH-1}\int_0^t K^p(r)\mathbb{W}_\theta(\mathscr{L}_{X_r^{n-1}},\mathscr{L}_{X_r^{n-2}})^p\mathrm{d}r\\
&\leq C_{p,H}K^p(t)t^{pH-1}\int_0^t\mathbb{E}|X_r^{n-1}-X_r^{n-2}|^p\mathrm{d}r\\
&\leq C_{p,H}K^p(t)t^{pH-1}\int_0^t\mathbb{E}\Big(\sup_{u\in[0,r]}|X_u^{n-1}-X_u^{n-2}|^p\Big)\mathrm{d}r. \qquad (6.1.10)
\end{aligned}$$

Plugging (6.1.9) and (6.1.10) into (6.1.8) yields

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{s\in[0,t]}|X_s^n-X_s^{n-1}|^p\Big) \qquad (6.1.11)\\
&\leq 2^{p-1}K^p(t)(2^pt^{p-1}+C_{p,H}t^{pH-1})\int_0^t\mathbb{E}\Big(\sup_{u\in[0,r]}|X_u^{n-1}-X_u^{n-2}|^p\Big)\mathrm{d}r\\
&\leq C_{p,H,T}\int_0^t\mathbb{E}\Big(\sup_{u\in[0,r]}|X_u^{n-1}-X_u^{n-2}|^p\Big)\mathrm{d}r
\end{aligned}$$

with $C_{p,H,T} := 2^{p-1}K^p(T)(2^pT^{p-1} + C_{p,H}T^{pH-1})$.

Hence, by the iteration we arrive at

$$\mathbb{E}\Big(\sup_{s\in[0,t]} |X_s^n - X_s^{n-1}|^p\Big) \leq C_1 C_{p,H,T}^n \frac{t^{n-1}}{(n-1)!},$$

where $C_1 := \mathbb{E}\Big(\sup_{t\in[0,T]} |X_t^1 - \xi|^p\Big) < \infty$ due to Step 1.

Consequently, $(X^n)_{n\geq 1}$ is a Cauchy sequence in $\mathcal{S}^p([0,T])$ with any $p \geq \theta$ and $p > \frac{1}{H}$, and then the limit, denoted by X, is a solution of (6.1.1).

Step 3. Uniqueness. Let X and Y be two solutions of (6.1.1). Along the same lines with Step 2, we derive that as in (6.1.11),

$$\mathbb{E}\Big(\sup_{s\in[0,t]} |X_s - Y_s|^p\Big) \leq C_{p,H,T} \int_0^t \mathbb{E}\Big(\sup_{u\in[0,r]} |X_u - Y_u|^p\Big)\mathrm{d}r, \quad t \in [0,T].$$

Then, the Gronwall lemma implies that $X_t = Y_t,\ t \in [0,T],\ \mathbb{P}$-a.s.. The proof is now complete. $\square$

6.2 Partial derivative in initial value and Malliavin derivative

Consider (6.1.1) with distribution independent $\sigma(t)$, i.e.

$$\mathrm{d}X_t = b(t, X_t, \mathscr{L}_{X_t})\mathrm{d}t + \sigma(t)\mathrm{d}B_t^H, \tag{6.2.1}$$

where $X_0 \in L^2(\Omega, \mathscr{F}_0, \mathbb{P})$ with $\mathscr{L}_{X_0} = \mu$.

The drift b satisfies the following assumption:

(A) For every $t \in [0,T]$, $b(t,\cdot,\cdot) \in C^{1,(1,0)}(\mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$. Moreover, there exists a constant $K > 0$ such that

$$\|\nabla b(t,\cdot,\mu)(x)\| + \|D^L b(t,x,\cdot)(\mu)\| \leq K, \quad t \in [0,T],\ x \in \mathbb{R}^d,\ \mu \in \mathscr{P}_2(\mathbb{R}^d).$$

Note that by the fundamental theorem for Bochner integral (see, for instance, [57, Proposition A.2.3]) and the definitions of L-derivative and the Wasserstein distance, (A) implies

$$|b(t,x,\mu) - b(t,y,\nu)| \le K(|x-y| + \mathbb{W}_2(\mu,\nu)), \quad t \in [0,T],$$

with $x, y \in \mathbb{R}^d,\ \mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$.

Then, it follows from Theorem 6.1.1 that (6.2.1) has a unique solution.

6.2.1 Partial derivative in the initial value

To investigate the partial derivative in initial value of (6.2.1), we first introduce a family of auxiliary equations. For any $\varepsilon > 0$ and $\eta \in L^2(\Omega \to \mathbb{R}^d, \mathscr{F}_0, \mathbb{P})$, let $(X_t^\varepsilon)_{t\in[0,T]}$ solve

$$\mathrm{d}X_t^\varepsilon = b(t, X_t^\varepsilon, \mathscr{L}_{X_t^\varepsilon})\mathrm{d}t + \sigma(t)\mathrm{d}B_t^H, \quad X_0^\varepsilon = X_0 + \varepsilon\eta, \tag{6.2.2}$$

and define

$$\Upsilon_t^\varepsilon := \frac{X_t^\varepsilon - X_t}{\varepsilon}, \quad t \in [0,T], \quad \varepsilon > 0.$$

Lemma 6.2.1. *Assume that (A1) holds. Then*

$$\sup_{\varepsilon>0} \mathbb{E}\left(\sup_{t\in[0,T]} |\Upsilon_t^\varepsilon|^2\right) \le 2\,\mathrm{e}^{8K^2T}\,\mathbb{E}|\eta|^2, \tag{6.2.3}$$

and

$$\sup_{\varepsilon\in(0,1],t\in[0,T]} |\Upsilon_t^\varepsilon|^2 \le \left(2|\eta|^2 + 8(KT)^2\,\mathrm{e}^{8(KT)^2}\,\mathbb{E}|\eta|^2\right)\mathrm{e}^{4(KT)^2}. \tag{6.2.4}$$

Proof. By (6.2.1)-(6.2.2) and (A1), we have for any $t \in [0,T]$ and $\varepsilon > 0$,

$$\sup_{s\in[0,t]} |X_s^\varepsilon - X_s|^2$$

$$\leq 2\varepsilon^2|\eta|^2 + 2\sup_{s\in[0,t]}\left|\int_0^s \big(b(r, X_r^\varepsilon, \mathscr{L}_{X_r^\varepsilon}) - b(r, X_r, \mathscr{L}_{X_r})\big)\,\mathrm{d}r\right|^2$$
$$\leq 2\varepsilon^2|\eta|^2 + 4K^2T\sup_{s\in[0,t]}\int_0^s \big(|X_r^\varepsilon - X_r|^2 + \mathbb{W}_2(\mathscr{L}_{X_r^\varepsilon}, \mathscr{L}_{X_r})^2\big)\,\mathrm{d}r.$$

Taking the expectation on both sides of the above inequality, we get

$$\mathbb{E}\bigg(\sup_{s\in[0,t]}|X_s^\varepsilon - X_s|^2\bigg) \leq 2\varepsilon^2\mathbb{E}|\eta|^2 + 8K^2T\int_0^t \mathbb{E}\bigg(\sup_{u\in[0,r]}|X_u^\varepsilon - X_u|^2\bigg)\mathrm{d}r,$$

which implies (6.2.3) and then (6.2.4) due to the Gronwall inequality. □

With Lemma 6.2.1 in hand, we can present the partial derivative in initial value of the equation (6.2.1). Consider now the following linear random ODE on $\mathbb{R}^d$: for any $\eta \in L^2(\Omega \to \mathbb{R}^d, \mathscr{F}_0, \mathbb{P})$ and $t \in [0, T]$,

$$\mathrm{d}\Gamma_t^\eta = \big[\nabla_{\Gamma_t^\eta} b(t, \cdot, \mathscr{L}_{X_t})(X_t) + \big(\mathbb{E}\langle D^L b(t, y, \cdot)(\mathscr{L}_{X_t})(X_t), \Gamma_t^\eta\rangle\big)\,|_{y=X_t}\big]\,\mathrm{d}t, \quad \Gamma_0^\eta = \eta, \tag{6.2.5}$$

where

$$\mathbb{E}\langle D^L b(t, y, \cdot)(\mathscr{L}_{X_t})(X_t), \Gamma_t^\eta\rangle := \big(\mathbb{E}\langle D^L b_i(t, y, \cdot)(\mathscr{L}_{X_t})(X_t), \Gamma_t^\eta\rangle\big)_{1\leq i\leq d} \in \mathbb{R}^d.$$

Obviously, (A1) implies that the ODE has a unique solution $\{\Gamma_t^\eta\}_{t\in[0,T]}$ satisfying

$$\mathbb{E}\left(\sup_{t\in[0,T]}|\Gamma_t^\eta|^2\right) \leq C_{T,K}\mathbb{E}|\eta|^2. \tag{6.2.6}$$

Proposition 6.2.2. *Assume that (A1) holds. Then for any* $\eta \in L^2(\Omega \to \mathbb{R}^d, \mathscr{F}_0, \mathbb{P})$*, the limit* $\nabla_\eta X_t := \lim_{\varepsilon\downarrow 0}\Upsilon_t^\varepsilon$ *exists in* $L^2(\Omega \to C([0, T]; \mathbb{R}^d), \mathbb{P})$ *such that* $\nabla_\eta X_t = \Gamma_t^\eta$ *holds for each* $t \in [0, T]$*, i.e.,* $\nabla_\eta X_t$ *is the unique solution of* (6.2.5).

Proof. To simplify the notation, we denote $X_\theta^\varepsilon(t) = X_t + \theta(X_t^\varepsilon - X_t), \theta \in [0, 1]$. By (6.2.1) and (6.2.2), we obtain that for any $t \in [0, T]$,

$$\mathrm{d}\Upsilon_t^\varepsilon = \frac{b(t, X_t^\varepsilon, \mathscr{L}_{X_t^\varepsilon}) - b(t, X_t, \mathscr{L}_{X_t})}{\varepsilon}\mathrm{d}t$$

$$
\begin{aligned}
&= \left[\frac{1}{\varepsilon}\int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} b(t, X_\theta^\varepsilon(t), \mathscr{L}_{X_t^\varepsilon})\mathrm{d}\theta + \frac{1}{\varepsilon}\int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} b(t, X_t, \mathscr{L}_{X_\theta^\varepsilon(t)})\mathrm{d}\theta\right]\mathrm{d}t \\
&= \left[\int_0^1 \nabla_{\Upsilon_t^\varepsilon} b(t,\cdot,\mathscr{L}_{X_t^\varepsilon})(X_\theta^\varepsilon(t))\mathrm{d}\theta \right.\\
&\qquad \left. + \int_0^1 (\mathbb{E}\langle D^L b(t,y,\cdot)(\mathscr{L}_{X_\theta^\varepsilon(t)})(X_\theta^\varepsilon(t)), \Upsilon_t^\varepsilon\rangle)|_{y=X_t}\mathrm{d}\theta\right]\mathrm{d}t
\end{aligned}
$$

with $\Upsilon_0^\varepsilon = \eta$. Here, we have used Lemma 2.3.1 in the last equality.

Then, combining this with (6.2.5) yields that for each $t \in [0, T]$,

$$
\begin{aligned}
\mathrm{d}(\Upsilon_t^\varepsilon - \Gamma_t^\eta) &= \left[\Phi_1^\varepsilon(t) + \nabla_{\Upsilon_t^\varepsilon - \Gamma_t^\eta} b(t,\cdot,\mathscr{L}_{X_t})(X_t)\right]\mathrm{d}t \\
&\quad + \left[\Phi_2^\varepsilon(t) + \left(\mathbb{E}\langle D^L b(t,y,\cdot)(\mathscr{L}_{X_t})(X_t), \Upsilon_t^\varepsilon - \Gamma_t^\eta\rangle\right)|_{y=X_t}\right]\mathrm{d}t,
\end{aligned}
$$

with $\Upsilon_0^\varepsilon - \Gamma_0^\eta = 0$, where

$$
\begin{aligned}
\Phi_1^\varepsilon(t) &:= \int_0^1 \left[\nabla_{\Upsilon_t^\varepsilon} b(t,\cdot,\mathscr{L}_{X_t^\varepsilon})(X_\theta^\varepsilon(t)) - \nabla_{\Upsilon_t^\varepsilon} b(t,\cdot,\mathscr{L}_{X_t})(X_t)\right]\mathrm{d}\theta, \\
\Phi_2^\varepsilon(t) &:= \int_0^1 \left(\mathbb{E}\langle D^L b(t,y,\cdot)(\mathscr{L}_{X_\theta^\varepsilon(t)})(X_\theta^\varepsilon(t)) - D^L b(t,y,\cdot)(\mathscr{L}_{X_t})(X_t), \Upsilon_t^\varepsilon\rangle\right)|_{y=X_t}\mathrm{d}\theta.
\end{aligned}
$$

Consequently, by (A1) we get

$$
\begin{aligned}
|\Upsilon_t^\varepsilon - \Gamma_t^\eta|^2 &\le 4T\int_0^t \left(|\Phi_1^\varepsilon(s)|^2 + |\Phi_2^\varepsilon(s)|^2\right)\mathrm{d}s \\
&\quad + 4K^2T\int_0^t \left(|\Upsilon_s^\varepsilon - \Gamma_s^\eta|^2 + \mathbb{E}|\Upsilon_s^\varepsilon - \Gamma_s^\eta|^2\right)\mathrm{d}s.
\end{aligned}
$$

Taking into account of (6.2.3) and (6.2.6), the Gronwall inequality leads to

$$
\mathbb{E}\left(\sup_{t\in[0,T]} |\Upsilon_t^\varepsilon - \Gamma_t^\eta|^2\right) \le 4T\,\mathrm{e}^{8(KT)^2}\int_0^T \mathbb{E}\left(|\Phi_1^\varepsilon(s)|^2 + |\Phi_2^\varepsilon(s)|^2\right)\mathrm{d}s. \quad (6.2.7)
$$

By the Hölder inequality and (6.2.3), one sees that

$$
\begin{aligned}
&|\Phi_1^\varepsilon(s)|^2 + |\Phi_2^\varepsilon(s)|^2 \\
&\le \int_0^1 |\nabla b(s,\cdot,\mathscr{L}_{X_s^\varepsilon})(X_\theta^\varepsilon(s)) - \nabla b(s,\cdot,\mathscr{L}_{X_s})(X_s)|^2\mathrm{d}\theta \cdot |\Upsilon_s^\varepsilon|^2
\end{aligned}
$$

$$+\mathbb{E}|\Upsilon_s^\varepsilon|^2\int_0^1\left(\mathbb{E}|D^Lb(s,y,\cdot)(\mathscr{L}_{X_\theta^\varepsilon(s)})(X_\theta^\varepsilon(s))-D^Lb(s,y,\cdot)(\mathscr{L}_{X_s})(X_s)|^2\right)|_{y=X_s}\mathrm{d}\theta \tag{6.2.8}$$

$$\leq 4K^2(|\Upsilon_s^\varepsilon|^2+\mathbb{E}|\Upsilon_s^\varepsilon|^2),$$

and

$$\lim_{\varepsilon\downarrow 0}\mathbb{E}\left(\sup_{\theta\in[0,1]}|X_\theta^\varepsilon(s)-X_s|^2\right)\leq\lim_{\varepsilon\downarrow 0}\mathbb{E}|X_s^\varepsilon-X_s|^2=0.$$

Then using the condition $b(s,\cdot,\cdot)\in C^{1,(1,0)}(\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d))$ of (A1) and (6.2.3) again, we obtain that $|\Phi_1^\varepsilon(s)|^2+|\Phi_2^\varepsilon(s)|^2$ converges to 0 in probability as ε goes to 0.

By (6.2.8), the dominated convergence theorem and the second assertion of Lemma 6.2.1 we conclude that

$$\lim_{\varepsilon\downarrow 0}\mathbb{E}\left[\sup_{s\in[0,T]}\left(|\Phi_1^\varepsilon(s)|^2+|\Phi_2^\varepsilon(s)|^2\right)\right]=0.$$

This, along with (6.2.7), implies

$$\lim_{\varepsilon\downarrow 0}\mathbb{E}\left(\sup_{t\in[0,T]}|\Upsilon_t^\varepsilon-\Gamma_t^\eta|^2\right)=0,$$

which completes the proof. □

6.2.2 Malliavin derivative

For the Malliavin derivative of the equation (6.2.1), consider for each $h\in\mathcal{H}$ and $\varepsilon>0$ the SDE: for $t\in[0,T]$,

$$\mathrm{d}X_t^{\varepsilon,h}=b(t,X_t^{\varepsilon,h},\mathscr{L}_{X_t})\mathrm{d}t+\sigma(t)\mathrm{d}(B_t^H+\varepsilon(R_Hh)(t)),\quad X_0^{\varepsilon,h}=X_0. \tag{6.2.9}$$

It is easy to see that under (A1) there exists a unique solution $X^{\varepsilon,h}$ to (6.2.9). Using the pathwise uniqueness of (6.2.1) and the fact that X_t can be regarded as a functional of B^H and X_0, the Malliavin directional derivative of X_t along $R_H h$ is shown by

$$\mathbb{D}_{R_H h} X_t := \lim_{\varepsilon \downarrow 0} \frac{X_t^{\varepsilon,h} - X_t}{\varepsilon}$$

if the limit exists in $L^2(\Omega \to C([0,T];\mathbb{R}^d), \mathbb{P})$. The above step is partially borrowed from [87, Proposition 3.5]. Noting that $\mathscr{L}_{X_t}$ in (6.2.9) is independent of ε, by the same arguments as in [30, Lemma 3.1 and Proposition 3.1] we have the following result.

Proposition 6.2.3. *Assume that (A1) holds. Then for any* $\eta \in L^2(\Omega \to \mathbb{R}^d, \mathscr{F}_0, \mathbb{P})$ *and* $h \in \mathcal{H}$, *the limit*

$$\lim_{\varepsilon \downarrow 0} \frac{X_t^{\varepsilon,h} - X_t}{\varepsilon}$$

exists in $L^2(\Omega \to C([0,T];\mathbb{R}^d), \mathbb{P})$ *such that* $\mathbb{D}_{R_H h} X_t = (\langle \mathbb{D} X_t^i, h \rangle_{\mathcal{H}})_{1 \le i \le d} \in \mathbb{R}^d$ *holds for every* $t \in [0,T]$ *and satisfies*

$$\mathbb{D}_{R_H h} X_t = \int_0^t \nabla_{\mathbb{D}_{R_H h} X_s} b(s, \cdot, \mathscr{L}_{X_s})(X_s) \mathrm{d}s + \int_0^t \sigma(s) \mathrm{d}(R_H h)(s), \quad t \in [0,T]. \tag{6.2.10}$$

6.3 General result of Bismut formula for Lions derivative

In this section, we aim to establish a general result of Bismut type formula of the L-derivative for (6.2.1). More precisely, for any $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, let $(X_t^\mu)_{t \in [0,T]}$ be the solution to (6.2.1) with $\mathscr{L}_{X_0} = \mu$ and denote $P_t^* \mu = \mathscr{L}_{X_t^\mu}$ for every $t \in [0,T]$. Now, define

$$(P_t f)(\mu) := \int_{\mathbb{R}^d} f \mathrm{d}(P_t^* \mu) = \mathbb{E} f(X_t^\mu), \quad t \in [0,T], \ f \in \mathscr{B}_b(\mathbb{R}^d).$$

For any $t \in (0,T], \mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\phi \in L^2(\mathbb{R}^d \to \mathbb{R}^d, \mu)$, the aim of this section is to find an integrable random variable $M_t(\mu,\phi)$ such that

$$D^L_\phi(P_t f)(\mu) = \mathbb{E}\left(f(X^\mu_t)M_t(\mu,\phi)\right), \quad f \in \mathscr{B}_b(\mathbb{R}^d).$$

To this end, let $X^{\mu_{\varepsilon,\phi}}_t$ denote the solution of (6.2.1) with initial data $X^{\mu_{\varepsilon,\phi}}_0 = (\mathrm{Id}+\varepsilon\phi)(X_0)$, where $\varepsilon \in [0,1]$ and $\phi \in T_{\mu,2}$. According to Proposition 6.2.2 and 6.2.3, $\nabla_{\phi(X_0)}X^{\mu_{\varepsilon,\phi}}_\cdot$ and $\mathbb{D}_{R_H h^{\varepsilon,\phi}_{s_0}}X^{\mu_{\varepsilon,\phi}}_\cdot$ below are both well-defined for any $s_0 \in [0,T)$, and satisfy (6.2.5) with $\eta = \phi(X_0)$ and (6.2.10), respectively. In order to ease notations, we simply write $\mu_{\varepsilon,\phi} = \mathscr{L}_{(\mathrm{Id}+\varepsilon\phi)(X_0)}$, and if $s_0 = 0$ or $\varepsilon = 0$, we often suppress s_0 or ε (e.g., $R_H h^{\varepsilon,\phi}_0 = R_H h^{\varepsilon,\phi}, h^{0,\phi}_0 = h^\phi, X^{\mu_{0,\phi}}_t = X^\mu_t$.).

The theorem below shows the general result of Bismut formula of L-derivative for (6.2.1).

Theorem 6.3.1. *Assume that for any $\varepsilon \in [0,1]$ and $s_0 \in [0,T)$, there exists $h^{\varepsilon,\phi}_{s_0} \in \mathrm{Dom}\delta \cap \mathcal{H}$ such that*

$$\mathbb{D}_{R_H h^{\varepsilon,\phi}_{s_0}} X^{\mu_{\varepsilon,\phi}}_T = \nabla_\phi X^{\mu_{\varepsilon,\phi}}_T, \tag{6.3.1}$$

with $(R_H h^{\varepsilon,\phi}_{s_0})(t) = 0$ for all $t \in [0,s_0]$, where $\nabla_\phi X^{\mu_{\varepsilon,\phi}}_T$ is in (6.2.5) and $\mathbb{D}_{R_H h^{\varepsilon,\phi}_{s_0}} X^{\mu_{\varepsilon,\phi}}_T$ solves (6.2.10) with $h^{\varepsilon,\phi}_{s_0}$ replacing h^ϕ. Moreover, suppose that

$$\int_0^1 \left(\mathbb{E}\delta^2(h^{\theta,\phi}_{s_0})\right)^{\frac{1}{2}} \mathrm{d}\theta < \infty, \tag{6.3.2}$$

and

$$\lim_{\varepsilon\to 0^+} \mathbb{E}|\delta(h^{\varepsilon,\phi}) - \delta(h^\phi)|^2 = 0, \quad \forall \phi \in T_{\mu,2}. \tag{6.3.3}$$

Then, there holds that

(i) For any $f \in \mathscr{B}_b(\mathbb{R}^d)$, then $P_T f$ is intrinsically differentiable at μ, such that

$$D^L_\phi(P_T f)(\mu) = \mathbb{E}(f(X^\mu_T)\delta(h^\phi)), \quad \forall \phi \in T_{\mu,2}.$$

(ii) If moreover

$$\mathbb{E}\delta^2(h^\phi) \le \tilde{L}\|\phi\|^2_{T_{\mu,2}}, \tag{6.3.4}$$

with a constant $\tilde{L} > 0$ *and*

$$\lim_{\|\phi\|_{T_{\mu,2}}\to 0} \sup_{\varepsilon\in(0,1]} \frac{\mathbb{E}|\delta(h^{\varepsilon,\phi}) - \delta(h^\phi)|^2}{\|\phi\|^2_{T_{\mu,2}}} = 0, \tag{6.3.5}$$

then for any $f \in \mathscr{B}_b(\mathbb{R}^d), P_T f$ *is L-differentiable at* μ.

6.3.1 Proof of the general result

Before providing the proof of Theorem 6.3.1, we prepare the following lemma.

Lemma 6.3.2. *Assume that for any* $\varepsilon \in [0,1]$ *and* $s_0 \in [0,T)$*, there exists* $h^{\varepsilon,\phi}_{s_0} \in \mathrm{Dom}\delta \cap \mathcal{H}$ *such that (6.3.1) holds with* $(R_H h^{\varepsilon,\phi}_{s_0})(t) = 0$ *for all* $t \in [0,s_0]$*. Then for any* $\varepsilon \in [0,1], s_0 \in [0,T)$ *and* $f \in \mathscr{B}_b(\mathbb{R}^d)$,

$$\mathbb{E}(f(X_T^{\mu_{\varepsilon,\phi}}) - f(X_T^\mu)|\mathscr{F}_{s_0}) = \int_0^\varepsilon \mathbb{E}\left(f(X_T^{\mu_{\tau,\phi}})\delta(h^{\tau,\phi}_{s_0})\Big|\mathscr{F}_{s_0}\right)\mathrm{d}\tau.$$

In particular, it holds

$$\mathbb{E}(f(X_T^{\mu_{\varepsilon,\phi}}) - f(X_T^\mu)) = \int_0^\varepsilon \mathbb{E}\left(f(X_T^{\mu_{\tau,\phi}})\delta(h^{\tau,\phi})\right)\mathrm{d}\tau.$$

Proof. Since $\mathbb{D}_{R_H h^{\varepsilon,\phi}_{s_0}} X_T^{\mu_{\varepsilon,\phi}} = \nabla_{\phi(X_0)} X_T^{\mu_{\varepsilon,\phi}}$, we deduce that for any $f \in C_b^1(\mathbb{R}^d)$,

$$\begin{aligned}
&\mathbb{E}(f(X_T^{\mu_{\varepsilon,\phi}}) - f(X_T^\mu)|\mathscr{F}_{s_0}) = \mathbb{E}\left(\int_0^\varepsilon \frac{\mathrm{d}}{\mathrm{d}\tau} f(X_T^{\mu_{\tau,\phi}})\mathrm{d}\tau\Big|\mathscr{F}_{s_0}\right)\\
&= \mathbb{E}\left(\int_0^\varepsilon \langle \nabla f(X_T^{\mu_{\tau,\phi}}), \nabla_{\phi(X_0)} X_T^{\mu_{\tau,\phi}}\rangle \mathrm{d}\tau\Big|\mathscr{F}_{s_0}\right)\\
&= \int_0^\varepsilon \mathbb{E}\left(\langle \nabla f(X_T^{\mu_{\tau,\phi}}), \nabla_{\phi(X_0)} X_T^{\mu_{\tau,\phi}}\rangle\Big|\mathscr{F}_{s_0}\right)\mathrm{d}\tau\\
&= \int_0^\varepsilon \mathbb{E}\left(\langle \nabla f(X_T^{\mu_{\tau,\phi}}), \mathbb{D}_{R_H h^{\tau,\phi}_{s_0}} X_T^{\mu_{\tau,\phi}}\rangle\Big|\mathscr{F}_{s_0}\right)\mathrm{d}\tau\\
&= \int_0^\varepsilon \mathbb{E}\left(\mathbb{D}_{R_H h^{\tau,\phi}_{s_0}} f(X_T^{\mu_{\tau,\phi}})\Big|\mathscr{F}_{s_0}\right)\mathrm{d}\tau
\end{aligned}$$

$$= \int_0^\varepsilon \mathbb{E}\left(\langle \mathbb{D} f(X_T^{\mu_\tau,\phi}), h_{s_0}^{\tau,\phi} \rangle_{\mathcal{H}} \Big| \mathscr{F}_{s_0} \right) \mathrm{d}\tau. \tag{6.3.6}$$

Now, let $\zeta \in \mathrm{Dom}\mathbb{D}$ be any bounded and $\mathscr{F}_{s_0}$-measurable smooth random variable, by [81, Proposition 1.2.3] we have for any $\tau \in [0, \varepsilon]$,

$$\begin{aligned}
&\mathbb{E}\left(\zeta \langle \mathbb{D} f(X_T^{\mu_\tau,\phi}), h_{s_0}^{\tau,\phi} \rangle_{\mathcal{H}}\right) \\
&= \mathbb{E}\left[\langle \mathbb{D}(\zeta f(X_T^{\mu_\tau,\phi})), h_{s_0}^{\tau,\phi} \rangle_{\mathcal{H}} - f(X_T^{\mu_\tau,\phi}) \langle \mathbb{D}\zeta, h_{s_0}^{\tau,\phi} \rangle_{\mathcal{H}}\right] \\
&= \mathbb{E}\left[\zeta f(X_T^{\mu_\tau,\phi}) \delta(h_{s_0}^{\tau,\phi}) - f(X_T^{\mu_\tau,\phi}) \langle \mathbb{D}\zeta, h_{s_0}^{\tau,\phi} \rangle_{\mathcal{H}}\right] \\
&= \mathbb{E}\left[\zeta f(X_T^{\mu_\tau,\phi}) \delta(h_{s_0}^{\tau,\phi}) - f(X_T^{\mu_\tau,\phi}) \langle K_H^* \mathbb{D}\zeta, K_H^* h_{s_0}^{\tau,\phi} \rangle_{L^2([0,T];\mathbb{R}^d)}\right],
\end{aligned} \tag{6.3.7}$$

where the last equality is due to the fact that K_H^* is an isometry between $\mathcal{H}$ and a closed subspace of $L^2([0,T];\mathbb{R}^d)$.

Using Proposition 2.2.2 and the fact that $(\mathbb{D}^W \zeta)(t) = 0$ for all $t > s_0$, we get

$$\begin{aligned}
&\left\langle K_H^* \mathbb{D}\zeta, K_H^* h_{s_0}^{\tau,\phi} \right\rangle_{L^2([0,T];\mathbb{R}^d)} \\
&= \left\langle \mathbb{D}^W \zeta, K_H^* h_{s_0}^{\tau,\phi} \right\rangle_{L^2([0,T];\mathbb{R}^d)} \\
&= \int_0^T \left\langle (\mathbb{D}^W \zeta)(t), (K_H^* h_{s_0}^{\tau,\phi})(t) \right\rangle \mathrm{d}t \\
&= \int_0^{s_0} \left\langle (\mathbb{D}^W \zeta)(t), (K_H^* h_{s_0}^{\tau,\phi})(t) \right\rangle \mathrm{d}t = 0.
\end{aligned} \tag{6.3.8}$$

Here we have used $K_H^* h_{s_0}^{\tau,\phi} = K_H^{-1}(R_H h_{s_0}^{\tau,\phi})$ and the fact that $(R_H h_{s_0}^{\tau,\phi})(t) = 0$ for $t \in [0, s_0]$ in the last equality.

Substituting (6.3.8) into (6.3.7) implies

$$\mathbb{E}\left[\zeta \langle \mathbb{D} f(X_T^{\mu_\tau,\phi}), h_{s_0}^{\tau,\phi} \rangle_{\mathcal{H}}\right] = \mathbb{E}\left[\zeta f(X_T^{\mu_\tau,\phi}) \delta(h_{s_0}^{\tau,\phi})\right].$$

Hence, combining this with (6.3.6), we obtain

$$\mathbb{E}(f(X_T^{\mu_\varepsilon,\phi}) - f(X_T^{\mu}) | \mathscr{F}_{s_0}) = \int_0^\varepsilon \mathbb{E}\left(f(X_T^{\mu_\tau,\phi}) \delta(h_{s_0}^{\tau,\phi}) \Big| \mathscr{F}_{s_0} \right) \mathrm{d}\tau, \quad f \in C_b^1(\mathbb{R}^d). \tag{6.3.9}$$

Set

$$\nu_{s_0}^{\varepsilon,\phi}(A) := \int_0^\varepsilon \mathbb{E}\left(\mathrm{I}_A(X_T^{\mu_{\tau,\phi}})|\delta(h_{s_0}^{\tau,\phi})|\right) \mathrm{d}\tau, \quad A \in \mathscr{B}(\mathbb{R}^d),$$

which is a finite measure on $\mathbb{R}^d$. Then $C_b^1(\mathbb{R}^d)$ is dense in $L^1(\mathbb{R}^d, \mathscr{L}_{X_T^{\mu_{\varepsilon,\phi}}} + \mathscr{L}_{X_T^\mu} + \nu_{s_0}^{\varepsilon,\phi})$. Therefore, (6.3.9) holds for any $f \in \mathscr{B}_b(\mathbb{R}^d)$. The proof now is complete. □

Now, we are in position to complete the **Proof of Theorem 6.3.1**

Proof. We divide the proof into two steps.

Step 1. Claim: For any $f \in \mathscr{B}_b(\mathbb{R}^d), P_T f$ is intrinsically differentiable at $\mu = \mathscr{L}_{X_0}$ (namely $(P_T f)(\mu \circ (\mathrm{Id} + \cdot)^{-1}) : L^2(\mathbb{R}^d \to \mathbb{R}^d, \mu) \to \mathbb{R}$ is Gâteaux differentiable at 0), and moreover $D_\phi^L(P_T f)(\mu) = \mathbb{E}(f(X_T^\mu)\delta(h^\phi))$ holds for each $\phi \in L^2(\mathbb{R}^d \to \mathbb{R}^d, \mu)$.

Due to Lemma 6.3.2, we deduce that for any $f \in \mathscr{B}_b(\mathbb{R}^d)$ and $\phi \in L^2(\mathbb{R}^d \to \mathbb{R}^d, \mu)$,

$$\begin{aligned}
&\frac{(P_T f)(\mu \circ (\mathrm{Id} + \varepsilon\phi)^{-1}) - (P_T f)(\mu)}{\varepsilon} - \mathbb{E}(f(X_T^\mu)\delta(h^\phi)) \\
&= \frac{(P_T f)(\mathscr{L}_{(\mathrm{Id}+\varepsilon\phi)(X_0)}) - (P_T f)(\mathscr{L}_{X_0})}{\varepsilon} - \mathbb{E}(f(X_T^\mu)\delta(h^\phi)) \\
&= \frac{\mathbb{E} f(X_T^{\mu_{\varepsilon,\phi}}) - \mathbb{E} f(X_T^\mu)}{\varepsilon} - \mathbb{E}(f(X_T^\mu)\delta(h^\phi)) \\
&= \frac{1}{\varepsilon}\int_0^\varepsilon \mathbb{E}\left(f(X_T^{\mu_{\tau,\phi}})\delta(h^{\tau,\phi})\right) \mathrm{d}\tau - \mathbb{E}(f(X_T^\mu)\delta(h^\phi)) \\
&= \frac{1}{\varepsilon}\int_0^\varepsilon \mathbb{E}\left[f(X_T^{\mu_{\tau,\phi}})(\delta(h^{\tau,\phi}) - \delta(h^\phi))\right] \mathrm{d}\tau + \frac{1}{\varepsilon}\int_0^\varepsilon \mathbb{E}\left[(f(X_T^{\mu_{\tau,\phi}}) - f(X_T^\mu))\delta(h^\phi)\right] \mathrm{d}\tau \\
&=: I_1(\phi) + I_2(\phi).
\end{aligned} \tag{6.3.10}$$

By (6.3.3) and $f \in \mathscr{B}_b(\mathbb{R}^d)$, we obtain

$$\begin{aligned}
\limsup_{\varepsilon\to0^+} |I_1(\phi)| &\le \|f\|_\infty \lim_{\varepsilon\to0^+} \frac{1}{\varepsilon}\int_0^\varepsilon \left(\mathbb{E}|\delta(h^{\tau,\phi}) - \delta(h^\phi)|^2\right)^{\frac{1}{2}} \mathrm{d}\tau \\
&\le \|f\|_\infty \lim_{\varepsilon\to0^+} \left(\mathbb{E}|\delta(h^{\varepsilon,\phi}) - \delta(h^\phi)|^2\right)^{\frac{1}{2}} = 0.
\end{aligned} \tag{6.3.11}$$

For $I_2(\phi)$, we get for any $s_0 \in (0, T)$,

$$\begin{aligned}
|I_2(\phi)| &\le \frac{1}{\varepsilon} \int_0^\varepsilon \left| \mathbb{E}\left[\big(f(X_T^{\mu_\tau,\phi}) - f(X_T^\mu) \big) \big(\delta(h^\phi) - \mathbb{E}(\delta(h^\phi)|\mathscr{F}_{s_0}) \big) \right] \right| \mathrm{d}\tau \\
&\quad + \frac{1}{\varepsilon} \int_0^\varepsilon \left| \mathbb{E}\left[(f(X_T^{\mu_\tau,\phi}) - f(X_T^\mu)) \mathbb{E}\big(\delta(h^\phi)|\mathscr{F}_{s_0} \big) \right] \right| \mathrm{d}\tau \\
&\le 2\|f\|_\infty \left| \mathbb{E}\delta(h^\phi) - \mathbb{E}(\delta(h^\phi)|\mathscr{F}_{s_0}) \right| \\
&\quad + \frac{1}{\varepsilon} \int_0^\varepsilon \left| \mathbb{E}\left[(f(X_T^{\mu_\tau,\phi}) - f(X_T^\mu)) \mathbb{E}\big(\delta(h^\phi)|\mathscr{F}_{s_0} \big) \right] \right| \mathrm{d}\tau.
\end{aligned} \tag{6.3.12}$$

On the one hand, it is easy to see that

$$\lim_{s_0 \to T^-} \mathbb{E} \left| \delta(h^\phi) - \mathbb{E}(\delta(h^\phi)|\mathscr{F}_{s_0}) \right| = 0. \tag{6.3.13}$$

On the other hand, note that by Lemma 6.3.2 again, we have

$$\begin{aligned}
&\left| \mathbb{E}[(f(X_T^{\mu_\tau,\phi}) - f(X_T^\mu)) \mathbb{E}(\delta(h^\phi)|\mathscr{F}_{s_0})] \right| \\
&= \left| \mathbb{E}\Big[\mathbb{E}(\delta(h^\phi)|\mathscr{F}_{s_0}) \mathbb{E}\big(f(X_T^{\mu_\tau,\phi}) - f(X_T^\mu) |\mathscr{F}_{s_0} \big) \Big] \right| \\
&= \left| \mathbb{E}\left[\mathbb{E}\big(\delta(h^\phi)|\mathscr{F}_{s_0} \big) \int_0^\tau \mathbb{E}\left(f(X_T^{\mu_\theta,\phi}) \delta(h_{s_0}^{\theta,\phi}) \middle| \mathscr{F}_{s_0} \right) \mathrm{d}\theta \right] \right| \\
&= \left| \int_0^\tau \mathbb{E}\left[\mathbb{E}\big(\delta(h^\phi)|\mathscr{F}_{s_0} \big) f(X_T^{\mu_\theta,\phi}) \delta(h_{s_0}^{\theta,\phi}) \right] \mathrm{d}\theta \right| \\
&\le \|f\|_\infty (\mathbb{E}\delta^2(h^\phi))^{\frac{1}{2}} \int_0^\tau \big(\mathbb{E}\delta^2(h_{s_0}^{\theta,\phi}) \big)^{\frac{1}{2}} \mathrm{d}\theta,
\end{aligned}$$

which goes to zero as $\tau \to 0$ due to (6.3.2). This means that the function $\tau \mapsto \mathbb{E}[(f(X_T^{\mu_\tau,\phi}) - f(X_T^\mu)) \mathbb{E}(\delta(h^\phi)|\mathscr{F}_{s_0})]$ is continuous at 0. We then derive that for each $s_0 \in (0, T)$,

$$\lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \int_0^\varepsilon \left| \mathbb{E}\left[(f(X_T^{\mu_\tau,\phi}) - f(X_T^\mu)) \mathbb{E}(\delta(h^\phi)|\mathscr{F}_{s_0}) \right] \right| \mathrm{d}\tau = 0. \tag{6.3.14}$$

Hence, plugging (6.3.13) and (6.3.14) into (6.3.12) implies that $\lim_{\varepsilon \to 0^+} |I_2(\phi)| = 0$. Combining this with (6.3.10) and (6.3.11), it yields the desired assertion.

Step 2. Claim: For any $f \in \mathscr{B}_b(\mathbb{R}^d)$, $P_T f$ is L-differentiable at $\mu = \mathscr{L}_{X_0}$ (namely $(P_T f)(\mu \circ (\mathrm{Id} + \cdot)^{-1}) : L^2(\mathbb{R}^d \to \mathbb{R}^d, \mu) \to \mathbb{R}$ is Fréchet differentiable at 0).

According to the definition of L-derivative, it is enough to show that for any $f \in \mathscr{B}_b(\mathbb{R}^d)$,

$$\lim_{\|\phi\|_{T_{\mu,2}} \to 0} \frac{|(P_T f)(\mu \circ (\mathrm{Id} + \phi)^{-1}) - (P_T f)(\mu) - \mathbb{E}(f(X_T^\mu)\delta(h^\phi))|}{\|\phi\|_{T_{\mu,2}}} = 0.$$

Applying Lemma 6.3.2 with $\varepsilon = 1$, we deduce from (6.3.4) that for any $f \in \mathscr{B}_b(\mathbb{R}^d)$,

$$\begin{aligned}
&\frac{|(P_T f)(\mu \circ (\mathrm{Id} + \phi)^{-1}) - (P_T f)(\mu) - \mathbb{E}(f(X_T^\mu)\delta(h^\phi))|}{\|\phi\|_{T_{\mu,2}}} \\
&= \frac{|\mathbb{E} f(X_T^{\mu_{1,\phi}}) - \mathbb{E} f(X_T^\mu) - \mathbb{E}(f(X_T^\mu)\delta(h^\phi))|}{\|\phi\|_{T_{\mu,2}}} \\
&= \frac{\left|\int_0^1 [\mathbb{E}(f(X_T^{\mu_{\tau,\phi}})\delta(h^{\tau,\phi})) - \mathbb{E}(f(X_T^\mu)\delta(h^\phi))] \mathrm{d}\tau\right|}{\|\phi\|_{T_{\mu,2}}} \\
&\le \frac{\left|\int_0^1 \mathbb{E}[f(X_T^{\mu_{\tau,\phi}})(\delta(h^{\tau,\phi}) - \delta(h^\phi))] \mathrm{d}\tau\right|}{\|\phi\|_{T_{\mu,2}}} + \frac{\left|\int_0^1 \mathbb{E}[(f(X_T^{\mu_{\tau,\phi}}) - f(X_T^\mu))\delta(h^\phi)] \mathrm{d}\tau\right|}{\|\phi\|_{T_{\mu,2}}} \\
&\le \frac{\|f\|_\infty \int_0^1 \left(\mathbb{E}\left|\delta(h^{\tau,\phi}) - \delta(h^\phi)\right|^2\right)^{\frac{1}{2}} \mathrm{d}\tau}{\|\phi\|_{T_{\mu,2}}} + \tilde{L} \int_0^1 (\mathbb{E}|f(X_T^{\mu_{\tau,\phi}}) - f(X_T^\mu)|^2)^{\frac{1}{2}} \mathrm{d}\tau \\
&=: J_1(\phi) + J_2(\phi).
\end{aligned}$$

Obviously, it follows from (6.3.5) that $\lim_{\|\phi\|_{T_{\mu,2}} \to 0} J_1(\phi) = 0$.
For $J_2(\phi)$, note first that by the Lusin theorem (see, e.g. [21, Theorem 7.4.4]), there exist $\{f_n\}_{n \ge 1} \subset C_b(\mathbb{R}^d)$ and compact sets $\{K_n\}_{n \ge 1}$ such that

$$f_n|_{K_n} = f|_{K_n}, \quad \|f_n\|_\infty \le \|f\|_\infty, \quad \mathscr{L}_{X_T^{\mu_{\tau,\phi}}}(K_n^c) + \mathscr{L}_{X_T^\mu}(K_n^c) \le \frac{1}{n^2}.$$

Then, we obtain

$$(\mathbb{E}|f(X_T^{\mu_{\tau,\phi}}) - f(X_T^\mu)|^2)^{\frac{1}{2}}$$

$$\leq \Big(\mathbb{E}|f(X_T^{\mu_\tau,\phi}) - f_n(X_T^{\mu_\tau,\phi}|^2 I\{f \neq f_n\}\Big)^{\frac{1}{2}} + (\mathbb{E}|f_n(X_T^{\mu_\tau,\phi}) - f_n(X_T^{\mu})|^2)^{\frac{1}{2}}$$
$$+ \Big(\mathbb{E}|f_n(X_T^{\mu}) - f(X_T^{\mu})|^2 I\{f \neq f_n\}\Big)^{\frac{1}{2}}$$
$$\leq \frac{4\|f\|_\infty}{n} + (\mathbb{E}|f_n(X_T^{\mu_\tau,\phi}) - f_n(X_T^{\mu})|^2)^{\frac{1}{2}}. \tag{6.3.15}$$

Note that for any $\tau \in [0,1]$, we have

$$\limsup_{\|\phi\|_{T_{\mu,2}}\to 0} \mathbb{E}|X_T^{\mu_\tau,\phi} - X_T^{\mu}|^2 \leq C \lim_{\|\phi\|_{T_{\mu,2}}\to 0} \|\phi\|_{T_{\mu,2}}^2 = 0,$$

where C is a positive constant. Consequently, the dominated convergence theorem yields that for every $n \geq 1$,

$$\lim_{\|\phi\|_{T_{\mu,2}}\to 0} \mathbb{E}|f_n(X_T^{\mu_\tau,\phi}) - f_n(X_T^{\mu})|^2 = 0.$$

Combining this with (6.3.15), we obtain that $\lim_{\|\phi\|_{T_{\mu,2}}\to 0} J_2(\phi) = 0$, which completes the proof. □

6.4 Bismut type formula: the non-degenerate case

This part is devoted to applying the general Theorem 6.3.1 to the non-degenerate case of (6.2.1).

6.4.1 Assumptions and main result

To applying our general result to the non-degenerate case, in additional to (A1), we also need the following assumptions:

(A2) There exists a constant $\tilde{K} > 0$ such that

(i) for any $t, s \in [0,T],\ x, y, z_1, z_2 \in \mathbb{R}^d,\ \mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$,

$$\begin{aligned}&\|\nabla b(t,\cdot,\mu)(x) - \nabla b(s,\cdot,\nu)(y)\| + |D^L b(t,x,\cdot)(\mu)(z_1) - D^L b(s,y,\cdot)(\nu)(z_2)|\\ &\le \tilde{K}(|t-s|^{\alpha_0} + |x-y|^{\beta_0} + |z_1 - z_2|^{\gamma_0} + \mathbb{W}_2(\mu,\nu)),\end{aligned}$$

where $\alpha_0 \in (H - 1/2, 1]$ and $\beta_0, \gamma_0 \in (1 - 1/(2H), 1]$.

(ii) σ is invertible and σ^{-1} is Hölder continuous of order $\delta_0 \in (H - 1/2, 1]$:

$$\|\sigma^{-1}(t) - \sigma^{-1}(s)\| \le \tilde{K}|t-s|^{\delta_0}, \quad \forall t, s \in [0,T].$$

(A3) The derivatives

$$\partial_t(D^L b(\cdot,x,\cdot)(\mu)(y))(t),\ \nabla(D^L b(t,\cdot,\cdot)(\mu)(y))(x),$$

$$D^L(D^L b(t,x,\cdot)(\cdot)(y))(\mu)(z),\ \nabla(D^L b(t,x,\cdot)(\mu)(\cdot))(y)$$

exist and are bounded continuous in the corresponding elements (t,x,μ,y) or (t,x,μ,y,z). We denote the bounded constants by a common one $\bar{K} > 0$.

Now, we state the main result of this part as follows:

Theorem 6.4.1. *Assume (A1), (A2) and (A3). Then for any $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $f \in \mathscr{B}_b(\mathbb{R}^d), P_T f$ is L-differentiable at μ such that*

$$D^L_\phi(P_T f)(\mu) = \mathbb{E}\left(f(X_T^\mu) \int_0^T \langle K_H^{-1}(R_H h^\phi)(t), \mathrm{d}W_t\rangle \right), \quad \forall \phi \in T_{\mu,2}, \quad (6.4.1)$$

where $h^\phi \in \mathrm{Dom}\delta \cap \mathcal{H}$ and satisfies for every $t \in [0,T]$,

$$\begin{aligned}(R_H h^\phi)(t) = \int_0^t \sigma^{-1}(s)\Big[&\frac{1}{T}\nabla_{\phi(X_0)} X_s^\mu + \frac{s}{T}\\ &\times \left(\mathbb{E}\langle D^L b(s,y,\cdot)(\mathscr{L}_{X_s^\mu})(X_s^\mu), \nabla_{\phi(X_0)} X_s^\mu\rangle\right)|_{y=X_s^\mu}\Big]\mathrm{d}s. \quad (6.4.2)\end{aligned}$$

Remark 6.4.1. By (2.2.3), one can recast the term $K_H^{-1}(R_H h^\phi)(t)$ in the theorem as

$$
\begin{aligned}
&K_H^{-1}(R_H h^\phi)(t)\\
&= \frac{(H-\frac{1}{2})t^{H-\frac{1}{2}}}{\Gamma(\frac{3}{2}-H)}\Bigg[\frac{t^{1-2H}\sigma^{-1}(t)\varrho(t)}{H-\frac{1}{2}} + \sigma^{-1}(t)\varrho(t)\int_0^t \frac{t^{\frac{1}{2}-H}-s^{\frac{1}{2}-H}}{(t-s)^{\frac{1}{2}+H}}\mathrm{d}s\\
&\qquad + \varrho(t)\int_0^t \frac{\sigma^{-1}(t)-\sigma^{-1}(s)}{(t-s)^{\frac{1}{2}+H}}s^{\frac{1}{2}-H}\mathrm{d}s + \int_0^t \frac{\varrho(t)-\varrho(s)}{(t-s)^{\frac{1}{2}+H}}\sigma^{-1}(s)s^{\frac{1}{2}-H}\mathrm{d}s\Bigg],
\end{aligned}
$$

where for any $s \in [0,T]$,

$$
\varrho(s) = \frac{1}{T}\nabla_{\phi(X_0)}X_s^\mu + \frac{s}{T}\left(\mathbb{E}\langle D^L b(s,y,\cdot)(\mathscr{L}_{X_s^\mu})(X_s^\mu), \nabla_{\phi(X_0)}X_s^\mu\rangle\right)|_{y=X_s^\mu}.
$$

6.4.2 Proof for the result of non-degenrate case

In order to prove the theorem, we prepare the following lemmas. The lemma below provides the estimates of the process $\{\nabla_{\phi(X_0)}X_t^{\mu_\varepsilon,\phi} - \nabla_{\phi(X_0)}X_t^\mu\}_{t\in[0,T]}$ in the sense of L^1-norm and in the sense of L^2-norm conditionally to $\mathscr{F}_0$, respectively.

Lemma 6.4.2. *Assume that (A1) and (A2) are satisfied. Then for any $t \in [0,T]$,*

$$
\mathbb{E}|\nabla_{\phi(X_0)}X_t^{\mu_\varepsilon,\phi} - \nabla_{\phi(X_0)}X_t^\mu| \le C_{T,K,\tilde{K}}\ell(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}, \tag{6.4.3}
$$

and

$$
\begin{aligned}
&\mathbb{E}(|\nabla_{\phi(X_0)}X_t^{\mu_\varepsilon,\phi} - \nabla_{\phi(X_0)}X_t^\mu|^2|\mathscr{F}_0)\\
&\le C_{T,K,\tilde{K}}\left(\tilde{\ell}_1^2(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}^2 + \tilde{\ell}_2^2(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}^2 + \tilde{\ell}_3^2(\varepsilon,\phi)|\phi(X_0)|^2\right),
\end{aligned} \tag{6.4.4}
$$

where

$$
\ell(\varepsilon,\phi) = \varepsilon^{\beta_0}\|\phi\|_{T_{\mu,2}}^{\beta_0} + \varepsilon^{\gamma_0}\|\phi\|_{T_{\mu,2}}^{\gamma_0} + \varepsilon\|\phi\|_{T_{\mu,2}},
$$

$$\tilde{\ell}_1(\varepsilon,\phi)=\varepsilon^{\beta_0}\|\phi\|_{T_{\mu,2}}^{\beta_0}+\varepsilon^{\gamma_0}\|\phi\|_{T_{\mu,2}}^{\gamma_0}+\varepsilon\|\phi\|_{T_{\mu,2}}+\varepsilon^{\frac{\beta_0}{2}}\|\phi\|_{T_{\mu,2}}^{\frac{\beta_0}{2}}+\varepsilon^{\frac{1}{2}}\|\phi\|_{T_{\mu,2}}^{\frac{1}{2}},$$
$$\tilde{\ell}_2(\varepsilon,\phi)=\varepsilon^{\frac{\beta_0}{2}}|\phi(X_0)|^{\frac{\beta_0}{2}}+\varepsilon^{\beta_0}|\phi(X_0)|^{\beta_0},$$
$$\tilde{\ell}_3(\varepsilon,\phi)=\varepsilon^{\frac{\beta_0}{2}}\|\phi\|_{T_{\mu,2}}^{\frac{\beta_0}{2}}+\varepsilon^{\frac{1}{2}}\|\phi\|_{T_{\mu,2}}^{\frac{1}{2}}+\varepsilon^{\frac{\beta_0}{2}}|\phi(X_0)|^{\frac{\beta_0}{2}}. \tag{6.4.5}$$

Remark 6.4.2. By a straightforward calculation, one can see that

$$\lim_{\varepsilon\to 0}\left[\ell(\varepsilon,\phi)+\tilde{\ell}_1(\varepsilon,\phi)+\mathbb{E}\left(\tilde{\ell}_2(\varepsilon,\phi)+\tilde{\ell}_3^2(\varepsilon,\phi)\right)\right]=0,$$
$$\lim_{\|\phi\|_{T_{\mu,2}}\to 0}\sup_{\varepsilon\in(0,1]}\left[\ell(\varepsilon,\phi)+\tilde{\ell}_1(\varepsilon,\phi))+\mathbb{E}\left(\tilde{\ell}_2(\varepsilon,\phi)+\tilde{\ell}_3^2(\varepsilon,\phi)\right)\right]=0.$$

Proof. By Proposition 6.2.2 with $\eta=\phi(X_0)$, we get for every $t\in[0,T]$,

$$\begin{aligned}
&\nabla_{\phi(X_0)}X_t^{\mu_\varepsilon,\phi}-\nabla_{\phi(X_0)}X_t^{\mu}\\
&=\int_0^t\Big[\nabla_{\nabla_{\phi(X_0)}X_s^{\mu_\varepsilon,\phi}}b(s,\cdot,\mathscr{L}_{X_s^{\mu_\varepsilon,\phi}})(X_s^{\mu_\varepsilon,\phi})-\nabla_{\nabla_{\phi(X_0)}X_s^{\mu}}b(s,\cdot,\mathscr{L}_{X_s^{\mu}})(X_s^{\mu})\\
&\qquad+\left(\mathbb{E}\langle D^Lb(s,y,\cdot)(\mathscr{L}_{X_s^{\mu_\varepsilon,\phi}})(X_s^{\mu_\varepsilon,\phi}),\nabla_{\phi(X_0)}X_s^{\mu_\varepsilon,\phi}\rangle\right)\Big|_{y=X_s^{\mu_\varepsilon,\phi}}\\
&\qquad-\left(\mathbb{E}\langle D^Lb(s,y,\cdot)(\mathscr{L}_{X_s^{\mu}})(X_s^{\mu}),\nabla_{\phi(X_0)}X_s^{\mu}\rangle\right)|_{y=X_s^{\mu}}\Big]\mathrm{d}s.
\end{aligned}$$

Let $\zeta_t=|\nabla_{\phi(X_0)}X_t^{\mu_\varepsilon,\phi}-\nabla_{\phi(X_0)}X_t^{\mu}|$. Then, by (A1) and (A2) we have for any $t\in[0,T]$,

$$\begin{aligned}
&\mathbb{E}(\zeta_t|\mathscr{F}_0)\le K\int_0^t(\mathbb{E}(\zeta_s|\mathscr{F}_0)+\mathbb{E}\zeta_s)\mathrm{d}s\\
&+\tilde{K}\int_0^t\mathbb{E}\left((|X_s^{\mu_\varepsilon,\phi}-X_s^{\mu}|^{\beta_0}+\mathbb{W}_2(\mathscr{L}_{X_s^{\mu_\varepsilon,\phi}},\mathscr{L}_{X_s^{\mu}}))|\nabla_{\phi(X_0)}X_s^{\mu}|\big|\mathscr{F}_0\right)\mathrm{d}s\\
&+\tilde{K}\int_0^t\left(\mathbb{E}(|X_s^{\mu_\varepsilon,\phi}-X_s^{\mu}|^{\beta_0}|\mathscr{F}_0)+\mathbb{W}_2(\mathscr{L}_{X_s^{\mu_\varepsilon,\phi}},\mathscr{L}_{X_s^{\mu}})\right)\mathbb{E}|\nabla_{\phi(X_0)}X_s^{\mu}|\mathrm{d}s\\
&+\tilde{K}\int_0^t\mathbb{E}\left(|X_s^{\mu_\varepsilon,\phi}-X_s^{\mu}|^{\gamma_0}|\nabla_{\phi(X_0)}X_s^{\mu}|\right)\mathrm{d}s.
\end{aligned}$$

Notice that by (A1), we derive for any $p>0$,

$$\sup_{s\in[0,T]}\mathbb{E}(|X_s^{\mu_\varepsilon,\phi}-X_s^{\mu}|^p|\mathscr{F}_0)\le C_{p,T,K}\varepsilon^p\left[\|\phi\|_{T_{\mu,2}}^p+|\phi(X_0)|^p\right], \tag{6.4.6}$$

and

$$\sup_{s\in[0,T],\varepsilon\in[0,1]} \mathbb{E}(|\nabla_{\phi(X_0)} X_s^{\mu_\varepsilon,\phi}|^p|\mathscr{F}_0) \le C_{p,T,K}\left[\|\phi\|_{T_\mu,2}^p + |\phi(X_0)|^p\right]. \quad (6.4.7)$$

Consequently, by (6.4.6) and (6.4.7) we obtain that for any $t\in[0,T]$,

$$\mathbb{E}(\zeta_t|\mathscr{F}_0) \le K\int_0^t (\mathbb{E}(\zeta_s|\mathscr{F}_0) + \mathbb{E}\zeta_s)\mathrm{d}s + C_{T,K,\tilde{K}}\chi(\varepsilon,\phi),$$

where

$$\begin{aligned}\chi(\varepsilon,\phi) &= \varepsilon^{\beta_0}\|\phi\|_{T_\mu,2}^{\beta_0+1} + \varepsilon^{\gamma_0}\|\phi\|_{T_\mu,2}^{\gamma_0+1} + \varepsilon\|\phi\|_{T_\mu,2}^2 + \varepsilon^{\beta_0}\|\phi\|_{T_\mu,2}|\phi(X_0)|^{\beta_0}\\ &\quad + \left[\varepsilon^{\beta_0}\left(\|\phi\|_{T_\mu,2}^{\beta_0} + |\phi(X_0)|^{\beta_0}\right) + \varepsilon\|\phi\|_{T_\mu,2}\right]|\phi(X_0)|.\end{aligned}$$

Taking the expectation on both sides and applying the Gronwall lemma, we obtain

$$\mathbb{E}\zeta_t \le C_{T,K,\tilde{K}}\mathbb{E}\chi(\varepsilon,\phi) \le C_{T,K,\tilde{K}}\ell(\varepsilon,\phi)\|\phi\|_{T_\mu,2}, \quad (6.4.8)$$

where $\ell(\varepsilon,\phi)$ is defined in (6.4.5). Hence, this leads to our first claim (6.4.3).

Next, we focus on proving (6.4.4). Applying the chain rule to ζ_t^2 and using (A2) yield that for any $t\in[0,T]$,

$$\begin{aligned}\mathrm{d}\zeta_t^2 \le \Big\{&2K\zeta_t^2 + 2K\zeta_t\mathbb{E}\zeta_t\\ &+ 3\tilde{K}(|\nabla_{\phi(X_0)}X_t^{\mu_\varepsilon,\phi}|^2 + |\nabla_{\phi(X_0)}X_t^\mu|^2)\left(|X_t^{\mu_\varepsilon,\phi} - X_t^\mu|^{\beta_0} + \mathbb{W}_2(\mathscr{L}_{X_t^{\mu_\varepsilon,\phi}}, \mathscr{L}_{X_t^\mu})\right)\\ &+ 2\tilde{K}\zeta_t(\mathbb{E}|\nabla_{\phi(X_0)}X_t^\mu|^2)^{\frac{1}{2}}\Big(|X_t^{\mu_\varepsilon,\phi} - X_t^\mu|^{\beta_0} + \mathbb{W}_2(\mathscr{L}_{X_t^{\mu_\varepsilon,\phi}}, \mathscr{L}_{X_t^\mu})\\ &\quad + (\mathbb{E}|X_t^{\mu_\varepsilon,\phi} - X_t^\mu|^{2\gamma_0})^{\frac{1}{2}}\Big)\Big\}\mathrm{d}t.\end{aligned}$$

Then, by the Hölder inequality we deduce that for any $t\in[0,T]$,

$$\begin{aligned}\mathbb{E}(\zeta_t^2|\mathscr{F}_0) \le{}& 2K\int_0^t \left[\mathbb{E}(\zeta_s^2|\mathscr{F}_0) + \mathbb{E}(\zeta_s|\mathscr{F}_0)\mathbb{E}\zeta_s\right]\mathrm{d}s\\ &+ 3\tilde{K}\int_0^t \left[\left(\mathbb{E}(|\nabla_{\phi(X_0)}X_s^{\mu_\varepsilon,\phi}|^4|\mathscr{F}_0)\right)^{\frac{1}{2}} + \left(\mathbb{E}\left(|\nabla_{\phi(X_0)}X_s^\mu|^4|\mathscr{F}_0\right)\right)^{\frac{1}{2}}\right]\end{aligned}$$

$$\times\left[\left(\mathbb{E}(|X_s^{\mu_\varepsilon,\phi}-X_s^\mu|^{2\beta_0}|\mathscr{F}_0)\right)^{\frac12}+\mathbb{W}_2(\mathscr{L}_{X_s^{\mu_\varepsilon,\phi}},\mathscr{L}_{X_s^\mu})\right]\mathrm{d}s$$

$$+2\tilde K\int_0^t\mathbb{E}(\zeta_s^2|\mathscr{F}_0)^{\frac12}(\mathbb{E}|\nabla_{\phi(X_0)}X_s^\mu|^2)^{\frac12}\times\Big[(\mathbb{E}(|X_s^{\mu_\varepsilon,\phi}-X_s^\mu|^{2\beta_0}|\mathscr{F}_0))^{\frac12}$$

$$+\mathbb{W}_2(\mathscr{L}_{X_s^{\mu_\varepsilon,\phi}},\mathscr{L}_{X_s^\mu})+(\mathbb{E}|X_s^{\mu_\varepsilon,\phi}-X_s^\mu|^{2\gamma_0})^{\frac12}\Big]\mathrm{d}s$$

$$\le\int_0^t\left[(2K+K^2+\tilde K^2)\mathbb{E}(\zeta_s^2|\mathscr{F}_0)+(\mathbb{E}\zeta_s)^2\right]\mathrm{d}s$$

$$+3\tilde K\int_0^t\left[\left(\mathbb{E}(|\nabla_{\phi(X_0)}X_s^{\mu_\varepsilon,\phi}|^4|\mathscr{F}_0)\right)^{\frac12}+\left(\mathbb{E}\left(|\nabla_{\phi(X_0)}X_s^\mu|^4|\mathscr{F}_0\right)\right)^{\frac12}\right]$$

$$\times\left[\left(\mathbb{E}(|X_s^{\mu_\varepsilon,\phi}-X_s^\mu|^{2\beta_0}|\mathscr{F}_0)\right)^{\frac12}+(\mathbb{E}|X_s^{\mu_\varepsilon,\phi}-X_s^\mu|^2)^{\frac12}\right]\mathrm{d}s$$

$$+3\int_0^t\mathbb{E}|\nabla_{\phi(X_0)}X_s^\mu|^2\Big[\mathbb{E}(|X_s^{\mu_\varepsilon,\phi}-X_s^\mu|^{2\beta_0}|\mathscr{F}_0)+\mathbb{E}|X_s^{\mu_\varepsilon,\phi}-X_s^\mu|^2$$

$$+\mathbb{E}|X_s^{\mu_\varepsilon,\phi}-X_s^\mu|^{2\gamma_0}\Big]\mathrm{d}s.$$

Combining this with (6.4.6), (6.4.7) and (6.4.8) and applying the Gronwall lemma, we conclude that that for any $t\in[0,T]$,

$$\mathbb{E}(\zeta_t^2|\mathscr{F}_0)\le C_{T,K,\tilde K}\left(\ell^2(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}^2+\tilde\chi(\varepsilon,\phi)\right),$$

where

$$\tilde\chi(\varepsilon,\phi)=\left(\|\phi\|_{T_{\mu,2}}^2+|\phi(X_0)|^2\right)\left[\varepsilon^{\beta_0}\left(\|\phi\|_{T_{\mu,2}}^{\beta_0}+|\phi(X_0)|^{\beta_0}\right)+\varepsilon\|\phi\|_{T_{\mu,2}}\right]$$
$$+\|\phi\|_{T_{\mu,2}}^2\left[\varepsilon^{2\beta_0}\left(\|\phi\|_{T_{\mu,2}}^{2\beta_0}+|\phi(X_0)|^{2\beta_0}\right)+\varepsilon^{2\gamma_0}\|\phi\|_{T_{\mu,2}}^{2\gamma_0}+\varepsilon^2\|\phi\|_{T_{\mu,2}}^2\right].$$

This, together with (6.4.5), yields that (6.4.4) holds. The proof is therefore complete. □

The following lemma describes the time continuity for the processes $X_t^{\mu_\varepsilon,\phi}$ in the sense of L^p-norm and $\nabla_{\phi(X_0)}X_t^{\mu_\varepsilon,\phi}$ in the sense of L^p-norm conditionally to $\mathscr{F}_0$, respectively.

Lemma 6.4.3. *Under assumptions (A1) and (A2). Then, we have for any $p\in(1/H,2]$,*

$$\sup_{\varepsilon\in[0,1]}\mathbb{E}|X_t^{\mu_\varepsilon,\phi}-X_s^{\mu_\varepsilon,\phi}|^p\le C_{p,T,K,H}|t-s|^{pH},\quad\forall t,s\in[0,T],\qquad(6.4.9)$$

and

$$\mathbb{E}(|\nabla_{\phi(X_0)}X_t^{\mu_\varepsilon,\phi} - \nabla_{\phi(X_0)}X_s^{\mu_\varepsilon,\phi}|^p|\mathscr{F}_0) \leq C_{p,T,K}\left(\|\phi\|_{T_\mu,2}^p + |\phi(X_0)|^p\right)|t-s|^p. \tag{6.4.10}$$

Proof. (6.4.9) derives from the same lines as in Step 1 of Theorem 6.1.1. By (6.4.7), it is easy to see that (6.4.10) holds. Thus, we omit the proof here. □

In the sequel, we are going to finish the proof of Theorem 6.4.1. To apply the general type Bismut formula (Theorem 6.3.1), it suffices to check the conditions therein in the non-degenerate setting. We will verify conditions in Theorem 6.3.1 one by one.

Firstly, for any $\varepsilon \in [0,1]$ and $s_0 \in [0,T)$, let

$$\begin{aligned}\tilde{h}_{s_0}^{\varepsilon,\phi}(t) = &\int_{t\wedge s_0}^t \sigma^{-1}(s)\left[\frac{1}{T-s_0}\nabla_{\phi(X_0)}X_s^{\mu_\varepsilon,\phi}\right.\\ &\left.+\frac{s-s_0}{T-s_0}\left(\mathbb{E}\langle D^L b(s,y,\cdot)(\mathscr{L}_{X_s^{\mu_\varepsilon,\phi}})(X_s^{\mu_\varepsilon,\phi}), \nabla_{\phi(X_0)}X_s^{\mu_\varepsilon,\phi}\rangle\right)|_{y=X_s^{\mu_\varepsilon,\phi}}\right]\mathrm{d}s\\ =: &\int_0^t \sigma^{-1}(s)\varrho_{\varepsilon,s_0}(s)\mathrm{I}_{\{s>s_0\}}\mathrm{d}s, \quad t\in[0,T].\end{aligned}$$

Owing to (A1) and (A2), one can verify that $\tilde{h}_{s_0}^{\varepsilon,\phi}(t) \in I_{0+}^{H+\frac{1}{2}}(L^2([0,T],\mathbb{R}^d))$, which means that $h_{s_0}^{\varepsilon,\phi} \in \mathcal{H}$ such that $R_H h_{s_0}^{\varepsilon,\phi}$ is well-defined.

Next, we intend to show $h_{s_0}^{\varepsilon,\phi} \in \mathrm{Dom}\delta$.

It is easy to check that $(R_H h_{s_0}^{\varepsilon,\phi})(t) = 0$ for all $t \in [0,s_0]$. Moreover, applying the chain rule to $\frac{t-s_0}{T-s_0}\nabla_{\phi(X_0)}X_t^{\mu_\varepsilon,\phi}$ yields $\mathbb{D}_{R_H h_{s_0}^{\varepsilon,\phi}}X_T^{\mu_\varepsilon,\phi} = \nabla_{\phi(X_0)}X_T^{\mu_\varepsilon,\phi}$. By Remark 2.2.1 and Proposition 2.2.3, we then obtain that $h_{s_0}^{\varepsilon,\phi} \in \mathrm{Dom}\delta$ and $\delta(h_{s_0}^{\varepsilon,\phi}) = \delta_W(K_H^* h_{s_0}^{\varepsilon,\phi}) = \int_0^T \langle K_H^{-1}(R_H h_{s_0}^{\varepsilon,\phi})(t), \mathrm{d}W_t\rangle$, provided that $K_H^{-1}(R_H h_{s_0}^{\varepsilon,\phi}) = K_H^* h_{s_0}^{\varepsilon,\phi} \in L^2([0,T]\times\Omega,\mathbb{R}^d)$.

Subsequently, we prove that $K_H^* h_{s_0}^{\varepsilon,\phi} \in L^2([0,T]\times\Omega,\mathbb{R}^d)$.

It is clear that the operator K_H^{-1} preserves the adaptability property.

With the help of (2.1.1) and (2.2.3), we have

$$
\begin{aligned}
&K_H^{-1}\left(\int_0^{\cdot} \sigma^{-1}(s)\varrho_{\varepsilon,s_0}(s)\mathrm{I}_{\{s>s_0\}}\mathrm{d}s\right)(t)\\
&= t^{H-\frac{1}{2}} D_{0+}^{H-\frac{1}{2}}\left[\cdot^{\frac{1}{2}-H}\sigma^{-1}(\cdot)\varrho_{\varepsilon,s_0}(\cdot)\mathrm{I}_{\{\cdot>s_0\}}\right](t)\\
&= \frac{H-\frac{1}{2}}{\Gamma(\frac{3}{2}-H)}\left[\frac{t^{\frac{1}{2}-H}\sigma^{-1}(t)\varrho_{\varepsilon,s_0}(t)\mathrm{I}_{\{t>s_0\}}}{H-\frac{1}{2}} + \sigma^{-1}(t)\varrho_{\varepsilon,s_0}(t)\int_0^t \frac{\mathrm{I}_{\{t>s_0\}}-\mathrm{I}_{\{s>s_0\}}}{(t-s)^{\frac{1}{2}+H}}\mathrm{d}s\right.\\
&\qquad + t^{H-\frac{1}{2}}\sigma^{-1}(t)\varrho_{\varepsilon,s_0}(t)\int_0^t \frac{t^{\frac{1}{2}-H}-s^{\frac{1}{2}-H}}{(t-s)^{\frac{1}{2}+H}}\mathrm{I}_{\{s>s_0\}}\mathrm{d}s\\
&\qquad + t^{H-\frac{1}{2}}\varrho_{\varepsilon,s_0}(t)\int_0^t \frac{\sigma^{-1}(t)-\sigma^{-1}(s)}{(t-s)^{\frac{1}{2}+H}}s^{\frac{1}{2}-H}\mathrm{I}_{\{s>s_0\}}\mathrm{d}s\\
&\qquad \left. + t^{H-\frac{1}{2}}\int_0^t \frac{\varrho_{\varepsilon,s_0}(t)-\varrho_{\varepsilon,s_0}(s)}{(t-s)^{\frac{1}{2}+H}}\sigma^{-1}(s)s^{\frac{1}{2}-H}\mathrm{I}_{\{s>s_0\}}\mathrm{d}s\right]\\
&=: \frac{H-\frac{1}{2}}{\Gamma(\frac{3}{2}-H)}[I_1(t)+I_2(t)+I_3(t)+I_4(t)+I_5(t)]. \qquad (6.4.11)
\end{aligned}
$$

From (6.4.7), it follows that

$$\sup_{s\in[0,T],\varepsilon\in[0,1]} \mathbb{E}|\varrho_{\varepsilon,s_0}(s)|^2 \le C_{s_0,T,K}\|\phi\|^2_{T_{\mu},2}.$$

Additionally, one sees that

$$\int_0^t \frac{\mathrm{I}_{\{t>s_0\}}-\mathrm{I}_{\{s>s_0\}}}{(t-s)^{\frac{1}{2}+H}}\mathrm{d}s = \frac{1}{H-\frac{1}{2}}\left((t-s_0)^{\frac{1}{2}-H}-t^{\frac{1}{2}-H}\right)\mathrm{I}_{\{t>s_0\}}$$

and

$$\int_0^t \frac{s^{\frac{1}{2}-H}-t^{\frac{1}{2}-H}}{(t-s)^{\frac{1}{2}+H}}\mathrm{d}s = t^{1-2H}\int_0^1 \frac{r^{\frac{1}{2}-H}-1}{(1-r)^{\frac{1}{2}+H}}\mathrm{d}r < \infty. \qquad (6.4.12)$$

Combining the above estimates, it yields from (A2)(ii) that

$$\mathbb{E}|I_1(t)|^2 + \mathbb{E}|I_3(t)|^2 \le C_{s_0,T,K,H}\|\phi\|^2_{T_{\mu},2}t^{1-2H}, \qquad (6.4.13)$$

$$\mathbb{E}|I_2(t)|^2 \le C_{s_0,T,K,H}\|\phi\|^2_{T_{\mu},2}(t-s_0)^{1-2H}, \qquad (6.4.14)$$

$$\mathbb{E}|I_4(t)|^2 \le C_{s_0,T,K,\tilde{K},H}\|\phi\|^2_{T_{\mu},2}t^{2\delta_0-2H+1}, \qquad (6.4.15)$$

which means that $I_i \in L^2([0,T]\times\Omega,\mathbb{R}^d), i=1,\cdots,4$.

Before handing I_5, we set for any $\varepsilon\in(0,1], t\in[0,T]$ and $y\in\mathbb{R}^d$,

$$\bar{b}^\varepsilon(t,y) := D^L b(t,y,\cdot)(\mathscr{L}_{X_t^{\mu_\varepsilon,\phi}})(X_t^{\mu_\varepsilon,\phi}), \quad \bar{b}(t,y) := D^L b(t,y,\cdot)(\mathscr{L}_{X_t^\mu})(X_t^\mu).$$

By a direct calculation, we can reduce the integrability of I_5 to that of the following three terms in $L^2([0,T]\times\Omega,\mathbb{R}^d)$:

$$t^{H-\frac{1}{2}}\int_0^t \frac{|\nabla_{\phi(X_0)}X_t^{\mu_\varepsilon,\phi} - \nabla_{\phi(X_0)}X_s^{\mu_\varepsilon,\phi}|}{(t-s)^{\frac{1}{2}+H}} s^{\frac{1}{2}-H}\mathrm{d}s,$$
$$t^{H-\frac{1}{2}}\int_0^t \frac{\left|\left(\mathbb{E}\langle\bar{b}^\varepsilon(t,y)-\bar{b}^\varepsilon(s,z),\nabla_{\phi(X_0)}X_t^{\mu_\varepsilon,\phi}\rangle\right)|_{y=X_t^{\mu_\varepsilon,\phi},z=X_s^{\mu_\varepsilon,\phi}}\right|}{(t-s)^{\frac{1}{2}+H}} s^{\frac{1}{2}-H}\mathrm{d}s,$$
$$t^{H-\frac{1}{2}}\int_0^t \frac{\left|\left(\mathbb{E}\langle\bar{b}^\varepsilon(s,z),\nabla_{\phi(X_0)}X_t^{\mu_\varepsilon,\phi}-\nabla_{\phi(X_0)}X_s^{\mu_\varepsilon,\phi}\rangle\right)|_{z=X_s^{\mu_\varepsilon,\phi}}\right|}{(t-s)^{\frac{1}{2}+H}} s^{\frac{1}{2}-H}\mathrm{d}s.$$

Combining this with (A2), we then obtain that

$$\begin{aligned}\mathbb{E}|I_5(t)|^2 \le C_{s_0,T,K,\tilde{K},H}\|\phi\|_{T_{\mu,2}}^2\Big[&t^{3-2H}+t^{2\alpha_0-2H+1}+t^{2\gamma_0 H-2H+1}+t\\ &+t^{2H-1}\mathbb{E}\Big(\int_0^t \frac{|X_t^{\mu_\varepsilon,\phi}-X_s^{\mu_\varepsilon,\phi}|^{\beta_0}}{(t-s)^{\frac{1}{2}+H}}s^{\frac{1}{2}-H}\mathrm{d}s\Big)^2\Big]. \quad (6.4.16)\end{aligned}$$

Note that there hold

$$\sup_{r\in[0,T]}|X_r^{\mu_\varepsilon,\phi}| \le C_{T,K,H}\Big(1+\|\mathrm{Id}+\varepsilon\phi\|_{T_{\mu,2}}+|X_0+\varepsilon\phi(X_0)|+\Big\|\int_0^\cdot \sigma(r)\mathrm{d}B_r^H\Big\|_\infty\Big),$$

and

$$\mathbb{E}\Big(\sup_{r\in[0,T]}|X_r^{\mu_\varepsilon,\phi}|^2\Big) \le C_{T,K,H}\left(1+\|\mathrm{Id}+\varepsilon\phi\|_{T_{\mu,2}}^2\right),$$

where $\|\int_0^\cdot\sigma(r)\mathrm{d}B_r^H\|_\infty := \sup_{t\in[0,T]}|\int_0^t\sigma(r)\mathrm{d}B_r^H|$.

Then, it follows from (A1) that

$$\begin{aligned}&\left|\int_s^t b(r,X_r^{\mu_\varepsilon,\phi},\mathscr{L}_{X_r^{\mu_\varepsilon,\phi}})\mathrm{d}r\right|\\ &\le K\Big(1+\sup_{r\in[0,T]}|X_r^{\mu_\varepsilon,\phi}|+\big(\mathbb{E}\sup_{r\in[0,T]}|X_r^{\mu_\varepsilon,\phi}|^2\big)^{\frac{1}{2}}\Big)(t-s)\end{aligned}$$

$$
\leq C_{T,K,H}\Big(1+\|\mathrm{Id}+\varepsilon\phi\|_{T_\mu,2}+|X_0+\varepsilon\phi(X_0)|
+\Big\|\int_0^\cdot\sigma(r)\mathrm{d}B_r^H\Big\|_\infty\Big)(t-s). \tag{6.4.17}
$$

Consequently, this implies

$$
\begin{aligned}
&t^{2H-1}\mathbb{E}\bigg(\int_0^t\frac{|X_t^{\mu_{\varepsilon,\phi}}-X_s^{\mu_{\varepsilon,\phi}}|^{\beta_0}}{(t-s)^{\frac12+H}}s^{\frac12-H}\mathrm{d}s\bigg)^2\\
&\leq 2t^{2H-1}\mathbb{E}\bigg(\int_0^t\frac{|\int_s^t b(r,X_r^{\mu_{\varepsilon,\phi}},\mathscr{L}_{X_r^{\mu_{\varepsilon,\phi}}})\mathrm{d}r|^{\beta_0}}{(t-s)^{\frac12+H}}s^{\frac12-H}\mathrm{d}s\bigg)^2\\
&\quad+2t^{2H-1}\mathbb{E}\bigg(\int_0^t\frac{|\int_s^t\sigma(r)\mathrm{d}B_r^H|^{\beta_0}}{(t-s)^{\frac12+H}}s^{\frac12-H}\mathrm{d}s\bigg)^2\\
&\leq C_{T,K,H}\left(1+\|\mathrm{Id}+\varepsilon\phi\|_{L_\mu^2}^{2\beta_0}+\mathbb{E}\left\|\int_0^\cdot\sigma(r)\mathrm{d}B_r^H\right\|_\infty^{2\beta_0}\right)t^{1+2(\beta_0-H)}\\
&\quad+C_H\mathbb{E}\left(\left\|\int_0^\cdot\sigma(r)\mathrm{d}B_r^H\mathrm{d}r\right\|_{H-\varsigma_0}^{2\beta_0}\right)t^{1+2(H-\varsigma_0)\beta_0-2H},
\end{aligned}\tag{6.4.18}
$$

where we use the Hölder continuity of $\int_0^\cdot\sigma(r)\mathrm{d}B_r^H$ of order $H-\varsigma_0$ with $\varsigma_0\in(0,1/2)$ and

$$
\Big\|\int_0^\cdot\sigma(r)\mathrm{d}B_r^H\Big\|_{H-\varsigma_0}:=\sup_{0\leq s<t\leq T}\frac{\Big|\int_0^t\sigma(r)\mathrm{d}B_r^H-\int_0^s\sigma(r)\mathrm{d}B_r^H\Big|}{|t-s|^{H-\varsigma_0}}.
$$

Plugging (6.4.18) into (6.4.16) yields that $I_5\in L^2([0,T]\times\Omega,\mathbb{R}^d)$. Then we get the desired claim.

Since

$$
\mathbb{E}\delta^2(h_{s_0}^{\varepsilon,\phi})=\mathbb{E}\delta_W^2(K_H^*h_{s_0}^{\varepsilon,\phi})=\int_0^T\mathbb{E}|K_H^{-1}(R_Hh_{s_0}^{\varepsilon,\phi})(t)|^2\mathrm{d}t,
$$

By (6.4.13)-(6.4.16), we then have that

$$
\int_0^1\left(\mathbb{E}\delta^2(h_{s_0}^{\tau,\phi})\right)^{\frac12}\mathrm{d}\tau<\infty\ \text{and}\ \mathbb{E}\delta^2(h^\phi)\leq C_{T,K,\tilde K,H}\|\phi\|_{T_\mu,2}^2. \tag{6.4.19}
$$

Finally, we shall estimate $\mathbb{E}|\delta(h^{\varepsilon,\phi})-\delta(h^\phi)|$. As before, we write $\varrho_\varepsilon=\varrho_{\varepsilon,0}$ and $\varrho=\varrho_{0,0}$ for simplicity. Using the linearity of the operator K_H^{-1} and applying the BDG inequality and the Hölder inequality, we have

$$\begin{aligned}
&\mathbb{E}|\delta(h^{\varepsilon,\phi}) - \delta(h^{\phi})| \qquad (6.4.20)\\
&= \mathbb{E}\left|\int_0^T \langle K_H^{-1}(R_H h^{\varepsilon,\phi})(t), \mathrm{d}W_t\rangle - \int_0^T \langle K_H^{-1}(R_H h^{\phi})(t), \mathrm{d}W_t\rangle\right|\\
&\le \left(\int_0^T \mathbb{E}|K_H^{-1}(R_H h^{\varepsilon,\phi} - R_H h^{\phi})(t)|^2 \mathrm{d}t\right)^{\frac{1}{2}}\\
&\le \mathbb{E}\left(\int_0^T \mathbb{E}\Big(|K_H^{-1}(R_H h^{\varepsilon,\phi} - R_H h^{\phi})(t)|^2 \big|\mathscr{F}_0\Big)\mathrm{d}t\right)^{\frac{1}{2}}\\
&= \mathbb{E}\left(\int_0^T \mathbb{E}\left(\left|K_H^{-1}\left(\int_0^{\cdot} \sigma^{-1}(s)(\varrho_\varepsilon - \varrho)(s)\mathrm{d}s\right)(t)\right|^2 \big|\mathscr{F}_0\right)\mathrm{d}t\right)^{\frac{1}{2}}.
\end{aligned}$$

(6.4.21)

By (2.1.1) and (2.2.3) again, we have

$$\begin{aligned}
&K_H^{-1}\left(\int_0^{\cdot} \sigma^{-1}(s)(\varrho_\varepsilon - \varrho)(s)\mathrm{d}s\right)(t) = t^{H-\frac{1}{2}} D_{0+}^{H-\frac{1}{2}}\left[\cdot^{\frac{1}{2}-H}\sigma^{-1}(\cdot)(\varrho_\varepsilon - \varrho)(\cdot)\right](t)\\
&= \frac{H-\frac{1}{2}}{\Gamma(\frac{3}{2}-H)}\Bigg[\frac{t^{\frac{1}{2}-H}\sigma^{-1}(t)(\varrho_\varepsilon - \varrho)(t)}{H-\frac{1}{2}} + t^{H-\frac{1}{2}}\sigma^{-1}(t)(\varrho_\varepsilon - \varrho)(t)\int_0^t \frac{t^{\frac{1}{2}-H} - s^{\frac{1}{2}-H}}{(t-s)^{\frac{1}{2}+H}}\mathrm{d}s\\
&\qquad + t^{H-\frac{1}{2}}(\varrho_\varepsilon - \varrho)(t)\int_0^t \frac{\sigma^{-1}(t) - \sigma^{-1}(s)}{(t-s)^{\frac{1}{2}+H}} s^{\frac{1}{2}-H}\mathrm{d}s\\
&\qquad + t^{H-\frac{1}{2}}\int_0^t \frac{(\varrho_\varepsilon - \varrho)(t) - (\varrho_\varepsilon - \varrho)(s)}{(t-s)^{\frac{1}{2}+H}}\sigma^{-1}(s) s^{\frac{1}{2}-H}\mathrm{d}s\Bigg]\\
&=: \frac{H-\frac{1}{2}}{\Gamma(\frac{3}{2}-H)}[J_1(t) + J_2(t) + J_3(t) + J_4(t)].
\end{aligned}$$

(6.4.22)

Owing to (A2)(ii) and (6.4.12), we get

$$\sum_{i=1}^{3} \mathbb{E}(|J_i(t)|^2|\mathscr{F}_0) \le C_{T,\tilde{K},H}\left(t^{1-2H} + t^{2\delta_0-2H+1}\right)\mathbb{E}\left(|(\varrho_\varepsilon - \varrho)(t)|^2|\mathscr{F}_0\right),$$

which leads to

$$\mathbb{E}\left(\int_0^T\sum_{i=1}^3\mathbb{E}(|J_i(t)|^2|\mathscr{F}_0)\mathrm{d}t\right)^{\frac{1}{2}}\le C_{T,\tilde{K},H}\mathbb{E}\left(\sup_{t\in[0,T]}\mathbb{E}\left(|(\varrho_\varepsilon-\varrho)(t)|^2|\mathscr{F}_0\right)\right)^{\frac{1}{2}}. \tag{6.4.23}$$

For notation simplicity, we set $\Gamma_t^{\mu_\varepsilon,\phi}:=\nabla_{\phi(X_0)}X_t^{\mu_\varepsilon,\phi}$ and $\Gamma_t^{\mu,\phi}:=\nabla_{\phi(X_0)}X_t^{\mu}$. Note that by (A1) and (A2), we obtain for any $t\in[0,T]$,

$$\begin{aligned}
&\mathbb{E}\left(|(\varrho_\varepsilon-\varrho)(t)|^2|\mathscr{F}_0\right)\\
&\le\frac{3}{T^2}\mathbb{E}(|\Gamma_t^{\mu_\varepsilon,\phi}-\Gamma_t^{\mu,\phi}|^2|\mathscr{F}_0)\\
&\quad+6\tilde{K}^2\left((\mathbb{E}(|X_t^{\mu_\varepsilon,\phi}-X_t^\mu|^{\beta_0}|\mathscr{F}_0)+\mathbb{W}_2(\mathscr{L}_{X_t^{\mu_\varepsilon,\phi}},\mathscr{L}_{X_t^\mu})\right)^2(\mathbb{E}|\Gamma_t^{\mu,\phi}|)^2\\
&\quad+6\tilde{K}^2\left(\mathbb{E}(|X_t^{\mu_\varepsilon,\phi}-X_t^\mu|^{\gamma_0}|\nabla_{\phi(X_0)}X_t^\mu|)\right)^2+3K^2(\mathbb{E}|\Gamma_t^{\mu_\varepsilon,\phi}-\Gamma_t^{\mu,\phi}|)^2\\
&\le C_{T,K,\tilde{K}}\Big[\left(\ell^2(\varepsilon,\phi)+\tilde{\ell}_1^2(\varepsilon,\phi)+\varepsilon^{2\beta_0}\|\phi\|_{T_{\mu,2}}^{2\beta_0}+\varepsilon^2\|\phi\|_{T_{\mu,2}}^2+\varepsilon^{2\gamma_0}\|\phi\|_{T_{\mu,2}}^{2\gamma_0}\right)\|\phi\|_{T_{\mu,2}}^2\\
&\qquad\qquad+(\tilde{\ell}_2^2(\varepsilon,\phi)+\varepsilon^{2\beta_0}|\phi(X_0)|^{2\beta_0})\|\phi\|_{T_{\mu,2}}^2+\tilde{\ell}_3^2(\varepsilon,\phi)|\phi(X_0)|^2\Big],
\end{aligned}$$

where the last inequality is due to (6.4.6), (6.4.7) and Lemma 6.4.2. Then, combining this with Remark 6.4.2 and (6.4.23) yields

$$\lim_{\varepsilon\to0}\mathbb{E}\left(\int_0^T\sum_{i=1}^3\mathbb{E}(|J_i(t)|^2|\mathscr{F}_0)\mathrm{d}t\right)^{\frac{1}{2}}=0, \tag{6.4.24}$$

and

$$\lim_{\|\phi\|_{T_{\mu,2}}\to0}\sup_{\varepsilon\in(0,1]}\frac{\mathbb{E}\left(\int_0^T\sum_{i=1}^3\mathbb{E}(|J_i(t)|^2|\mathscr{F}_0)\mathrm{d}t\right)^{\frac{1}{2}}}{\|\phi\|_{T_{\mu,2}}}=0. \tag{6.4.25}$$

For the term associated with $J_4(t)$, observe first that for any $0\le s<t\le T$,

$$(\varrho_\varepsilon-\varrho)(t)-(\varrho_\varepsilon-\varrho)(s)=\sum_{i=1}^6\Theta_i(t,s), \tag{6.4.26}$$

where

$$\Theta_1(t,s)=\frac{1}{T}\left[\left(\Gamma_t^{\mu_\varepsilon,\phi}-\Gamma_s^{\mu_\varepsilon,\phi}\right)-\left(\Gamma_t^{\mu,\phi}-\Gamma_s^{\mu,\phi}\right)\right],$$

$$\Theta_2(t,s) = \frac{t-s}{T}\left[\left(\mathbb{E}\langle \bar{b}^{\varepsilon}(t,y), \Gamma_t^{\mu_\varepsilon,\phi}\rangle\right) - \left(\mathbb{E}\langle \bar{b}(t,\tilde{y}), \Gamma_t^{\mu,\phi}\rangle\right)\right],$$

$$\Theta_3(t,s) = \frac{s}{T}\left(\mathbb{E}\left\langle \bar{b}^{\varepsilon}(t,y), \left(\Gamma_t^{\mu_\varepsilon,\phi} - \Gamma_s^{\mu_\varepsilon,\phi}\right) - \left(\Gamma_t^{\mu,\phi} - \Gamma_s^{\mu,\phi}\right)\right\rangle\right),$$

$$\Theta_4(t,s) = \frac{s}{T}\left(\mathbb{E}\langle \bar{b}^{\varepsilon}(t,y) - \bar{b}^{\varepsilon}(s,z), \Gamma_s^{\mu_\varepsilon,\phi} - \Gamma_s^{\mu,\phi}\rangle\right),$$

$$\Theta_5(t,s) = \frac{s}{T}\left(\mathbb{E}\langle \bar{b}^{\varepsilon}(s,z) - \bar{b}(s,\tilde{z}), \Gamma_t^{\mu,\phi} - \Gamma_s^{\mu,\phi}\rangle\right),$$

$$\Theta_6(t,s) = \frac{s}{T}\left(\mathbb{E}\langle (\bar{b}^{\varepsilon}(t,y) - \bar{b}^{\varepsilon}(s,z)) - (\bar{b}(t,\tilde{y}) - \bar{b}(s,\tilde{z})), \Gamma_t^{\mu,\phi}\rangle\right),$$

with $y = X_t^{\mu_\varepsilon,\phi}, \tilde{y} = X_t^{\mu}, z = X_s^{\mu_\varepsilon,\phi}$ and $\tilde{z} = X_s^{\mu}$.

Owing to (A1), (A2), (6.4.6), (6.4.7), (6.4.10) and Lemma 6.4.2, one has that

$$\mathbb{E}\left(t^{2H-1}\left(\int_0^t \frac{|\Theta_1(t,s)|s^{\frac{1}{2}-H}}{(t-s)^{\frac{1}{2}+H}}\mathrm{d}s\right)^2 |\mathscr{F}_0\right)$$
$$\leq C_{T,K,\tilde{K},H} t^{3-2H}\left[\tilde{\ell}_1^2(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}^2 + \tilde{\ell}_2^2(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}^2 + (\tilde{\ell}_3^2(\varepsilon,\phi) + \tilde{\ell}_3^4(\varepsilon,\phi))|\phi(X_0)|^2\right], \tag{6.4.27}$$

$$\mathbb{E}\left(t^{2H-1}\left(\int_0^t \frac{|\Theta_2(t,s)|s^{\frac{1}{2}-H}}{(t-s)^{\frac{1}{2}+H}}\mathrm{d}s\right)^2 |\mathscr{F}_0\right)$$
$$\leq C_{T,K,\tilde{K},H} t^{3-2H}\left(\ell^2(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}^2 + \varepsilon^{2\beta_0}|\phi(X_0)|^{2\beta_0}\|\phi\|_{T_{\mu,2}}^2\right), \tag{6.4.28}$$

$$\mathbb{E}\left(t^{2H-1}\left(\int_0^t \frac{|\Theta_3(t,s)|s^{\frac{1}{2}-H}}{(t-s)^{\frac{1}{2}+H}}\mathrm{d}s\right)^2 |\mathscr{F}_0\right) \tag{6.4.29}$$
$$\leq C_{T,K,\tilde{K},H} t^{3-2H}\ell^2(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}^2,$$

$$\mathbb{E}\left(t^{2H-1}\left(\int_0^t \frac{|\Theta_5(t,s)|s^{\frac{1}{2}-H}}{(t-s)^{\frac{1}{2}+H}}\mathrm{d}s\right)^2 |\mathscr{F}_0\right)$$
$$\leq C_{T,K,\tilde{K},H} t^{3-2H}\left[\varepsilon^2\|\phi\|_{T_{\mu,2}}^2 + \varepsilon^{2\gamma_0}\|\phi\|_{T_{\mu,2}}^{2\gamma_0} + \varepsilon^{2\beta_0}\left(\|\phi\|_{T_{\mu,2}}^{2\beta_0} + |\phi(X_0)|^{2\beta_0}\right)\right]\|\phi\|_{T_{\mu,2}}^2. \tag{6.4.30}$$

For $\Theta_4(t,s)$, by (A2)(i), (6.4.3) and (6.4.9) we first have

$$|\Theta_4(t,s)| \leq C_{T,K,\tilde{K},H}\left[(t-s)^{\alpha_0} + (t-s)^H + |X_t^{\mu_\varepsilon,\phi} - X_s^{\mu_\varepsilon,\phi}|^{\beta_0}\right]\ell(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}$$

$$
\begin{aligned}
&+\tilde{K}\mathbb{E}\left(|X_t^{\mu_\varepsilon,\phi}-X_s^{\mu_\varepsilon,\phi}|^{\gamma_0}\cdot|\Gamma_s^{\mu_\varepsilon,\phi}-\Gamma_s^{\mu,\phi}|\right)\\
&\le C_{T,K,\tilde{K},H}\left[(t-s)^{\alpha_0}+(t-s)^H+|X_t^{\mu_\varepsilon,\phi}-X_s^{\mu_\varepsilon,\phi}|^{\beta_0}\right]\ell(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}\\
&+\tilde{K}\mathbb{E}\left(\left|\int_s^t b(r,X_r^{\mu_\varepsilon,\phi},\mathscr{L}_{X_r^{\mu_\varepsilon,\phi}})\mathrm{d}r\right|^{\gamma_0}\cdot|\Gamma_s^{\mu_\varepsilon,\phi}-\Gamma_s^{\mu,\phi}|\right)\\
&+\tilde{K}\mathbb{E}\left(\left|\int_s^t \sigma(r)\mathrm{d}B_r^H\right|^{\gamma_0}\cdot|\Gamma_s^{\mu_\varepsilon,\phi}-\Gamma_s^{\mu,\phi}|\right).
\end{aligned}
\tag{6.4.31}
$$

Next, we focus on dealing with the last two terms of the right-hand side of (6.4.31). Using (6.4.17), (6.4.3), (6.4.4) and the fact that B^H is independent of $\mathscr{F}_0$, we obtain

$$
\begin{aligned}
&\mathbb{E}\Big(\Big|\int_s^t b(r,X_r^{\mu_\varepsilon,\phi},\mathscr{L}_{X_r^{\mu_\varepsilon,\phi}})\mathrm{d}r\Big|^{\gamma_0}\cdot|\nabla_{\phi(X_0)}X_s^{\mu_\varepsilon,\phi}-\nabla_{\phi(X_0)}X_s^{\mu}|\Big)\\
&\le C_{T,K,\tilde{K},H}\Big\{\Big(1+\|\mathrm{Id}+\varepsilon\phi\|_{T_{\mu,2}}^{\gamma_0}\Big)\ell(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}\\
&\qquad+\Big(\mathbb{E}\Big\|\int_0^\cdot \sigma(r)\mathrm{d}B_r^H\Big\|_\infty^{2\gamma_0}\Big)^{\frac{1}{2}}\Big[\tilde{\ell}_1(\varepsilon,\phi)+\mathbb{E}\tilde{\ell}_2(\varepsilon,\phi)+\Big(\mathbb{E}\tilde{\ell}_3^2(\varepsilon,\phi)\Big)^{\frac{1}{2}}\Big]\|\phi\|_{T_{\mu,2}}\\
&\qquad+\mathbb{E}\Big(|X_0+\varepsilon\phi(X_0)|^{\gamma_0}\cdot|\nabla_{\phi(X_0)}X_s^{\mu_\varepsilon,\phi}-\nabla_{\phi(X_0)}X_s^{\mu}|\Big)\Big\}(t-s)^{\gamma_0}.
\end{aligned}
\tag{6.4.32}
$$

Observe that by (6.4.4), we derive

$$
\begin{aligned}
&\mathbb{E}\left(|X_0+\varepsilon\phi(X_0)|^{\gamma_0}\cdot|\nabla_{\phi(X_0)}X_s^{\mu_\varepsilon,\phi}-\nabla_{\phi(X_0)}X_s^{\mu}|\right)\\
&\le C_{T,K,\tilde{K}}\mathbb{E}\left[|X_0+\varepsilon\phi(X_0)|^{\gamma_0}\left(\tilde{\ell}_1(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}+\tilde{\ell}_2(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}+\tilde{\ell}_3(\varepsilon,\phi)|\phi(X_0)|\right)\right]\\
&\le C_{T,K,\tilde{K}}\Bigg[\|\mathrm{Id}+\varepsilon\phi\|_{T_{\mu,2}}^{\gamma_0}\left(\tilde{\ell}_1(\varepsilon,\phi)+\left(\mathbb{E}\tilde{\ell}_2^2(\varepsilon,\phi)\right)^{\frac{1}{2}}+\varepsilon^{\frac{\beta_0}{2}}\|\phi\|_{T_{\mu,2}}^{\frac{\beta_0}{2}}+\varepsilon^{\frac{1}{2}}\|\phi\|_{T_{\mu,2}}^{\frac{1}{2}}\right)\|\phi\|_{T_{\mu,2}}\\
&\qquad+\varepsilon^{\frac{\beta_0}{2}}\mathbb{E}\left(|X_0+\varepsilon\phi(X_0)|^{\gamma_0}\cdot|\phi(X_0)|^{1+\frac{\beta_0}{2}}\right)\Bigg]\\
&\le C_{T,K,\tilde{K}}\|\mathrm{Id}+\varepsilon\phi\|_{T_{\mu,2}}^{\gamma_0}\left(\tilde{\ell}_1(\varepsilon,\phi)+\left(\mathbb{E}\tilde{\ell}_2^2(\varepsilon,\phi)\right)^{\frac{1}{2}}+\varepsilon^{\frac{\beta_0}{2}}\|\phi\|_{T_{\mu,2}}^{\frac{\beta_0}{2}}+\varepsilon^{\frac{1}{2}}\|\phi\|_{T_{\mu,2}}^{\frac{1}{2}}\right)\|\phi\|_{T_{\mu,2}}\\
&=C_{T,K,\tilde{K}}\|\mathrm{Id}+\varepsilon\phi\|_{T_{\mu,2}}^{\gamma_0}\left(\tilde{\ell}_1(\varepsilon,\phi)+\left(\mathbb{E}\tilde{\ell}_2^2(\varepsilon,\phi)\right)^{\frac{1}{2}}\right)\|\phi\|_{T_{\mu,2}},
\end{aligned}
\tag{6.4.33}
$$

where we use the Hölder inequality with $\frac{2+\beta_0}{4}+\frac{2-\beta_0}{4}=1$ and the relation $(1-\frac{1}{2H}\le)\gamma_0\le 1-\frac{\beta_0}{2}$ in the last inequality. Note that if $\gamma_0\in(1-\frac{\beta_0}{2},1]$, we may choose $\tilde{\gamma}_0 \in [1-\frac{1}{2H}, 1-\frac{\beta_0}{2}]$ to replace such γ_0 in the first inequality of (6.4.31) due to the boundedness of $D^L b$. In this case, (6.4.35) below holds with γ_0 replaced by $\tilde{\gamma}_0$, which also implies the desired convergence of the term involved Θ_4.

Substituting (6.4.33) into (6.4.32) and recalling that $\varepsilon \in [0,1]$ imply

$$\begin{aligned}
&\mathbb{E}\Big(\Big|\int_s^t b(r, X_r^{\mu_\varepsilon,\phi}, \mathscr{L}_{X_r^{\mu_\varepsilon,\phi}})\mathrm{d}r\Big|^{\gamma_0} \cdot |\nabla_{\phi(X_0)} X_s^{\mu_\varepsilon,\phi} - \nabla_{\phi(X_0)} X_s^{\mu}|\Big) \\
&\le C_{T,K,\tilde{K},H}(t-s)^{\gamma_0}\Big(\ell(\varepsilon,\phi) + \tilde{\ell}_1(\varepsilon,\phi) + \big(\mathbb{E}\tilde{\ell}_2^2(\varepsilon,\phi)\big)^{\frac{1}{2}} \\
&\qquad + \big(\mathbb{E}\tilde{\ell}_3^2(\varepsilon,\phi)\big)^{\frac{1}{2}}\Big)\|\phi\|_{T_{\mu,2}},
\end{aligned} \tag{6.4.34}$$

For the other term, applying the fact that B^H is independent of $\mathscr{F}_0$ again and (6.4.4), one sees that

$$\begin{aligned}
&\mathbb{E}\Big(\Big|\int_s^t \sigma(r)\mathrm{d}B_r^H\Big|^{\gamma_0} \cdot |\nabla_{\phi(X_0)} X_s^{\mu_\varepsilon,\phi} - \nabla_{\phi(X_0)} X_s^{\mu}|\Big) \\
&\le \mathbb{E}\Big\{\Big[\mathbb{E}\Big(\Big|\int_s^t \sigma(r)\mathrm{d}B_r^H\Big|^{2\gamma_0}|\mathscr{F}_0\Big)\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\Big(|\nabla_{\phi(X_0)} X_s^{\mu_\varepsilon,\phi} - \nabla_{\phi(X_0)} X_s^{\mu}|^2|\mathscr{F}_0\Big)\Big]^{\frac{1}{2}}\Big\} \\
&\le C_{T,K,\tilde{K},H}(t-s)^{\gamma_0 H}\Big(\tilde{\ell}_1(\varepsilon,\phi) + \mathbb{E}\tilde{\ell}_2(\varepsilon,\phi) + \Big(\mathbb{E}\tilde{\ell}_3^2(\varepsilon,\phi)\Big)^{\frac{1}{2}}\Big)\|\phi\|_{T_{\mu,2}}.
\end{aligned}$$

Plugging this and (6.4.34) into (6.4.31), we arrive at

$$\begin{aligned}
|\Theta_4(t,s)| \le C_{T,K,\tilde{K},H}\|\phi\|_{T_{\mu,2}}\Big[&|X_t^{\mu_\varepsilon,\phi} - X_s^{\mu_\varepsilon,\phi}|^{\beta_0}\ell(\varepsilon,\phi) \\
&+ (t-s)^{\alpha_0\wedge(\gamma_0 H)}\Big(\ell(\varepsilon,\phi) + \tilde{\ell}_1(\varepsilon,\phi) + \Big(\mathbb{E}\tilde{\ell}_2^2(\varepsilon,\phi)\Big)^{\frac{1}{2}} + \Big(\mathbb{E}\tilde{\ell}_3^2(\varepsilon,\phi)\Big)^{\frac{1}{2}}\Big)\Big].
\end{aligned}$$

Hence, combining this with (6.4.18) and the fact that B^H is independent of $\mathscr{F}_0$ again leads to

$$\begin{aligned}
&\mathbb{E}\Big(t^{2H-1}\Big(\int_0^t \frac{|\Theta_4(t,s)|s^{\frac{1}{2}-H}}{(t-s)^{\frac{1}{2}+H}}\mathrm{d}s\Big)^2|\mathscr{F}_0\Big) \\
&\le C_{T,K,\tilde{K},H}\Big(1 + \|\mathrm{Id}+\varepsilon\phi\|_{T_{\mu,2}}^{2\beta_0} + \mathbb{E}\Big\|\int_0^\cdot \sigma(r)\mathrm{d}B_r^H\Big\|_\infty^{2\beta_0}\Big)t^{1+2(\beta_0-H)}\ell^2(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}^2 \\
&\quad + C_{T,K,\tilde{K},H}\mathbb{E}\Big(\Big\|\int_0^\cdot \sigma(r)\mathrm{d}B_r^H\mathrm{d}r\Big\|_{H-\varsigma_0}^{2\beta_0}\Big)t^{1+2(H-\varsigma_0)\beta_0-2H}\ell^2(\varepsilon,\phi)\|\phi\|_{T_{\mu,2}}^2
\end{aligned}$$

$$
\begin{aligned}
&+ C_{T,K,\tilde{K},H} t^{1+2(\alpha_0\wedge(\gamma_0 H)-H)} \Big(\ell^2(\varepsilon,\phi) + \tilde{\ell}_1^2(\varepsilon,\phi) + \mathbb{E}\tilde{\ell}_2^2(\varepsilon,\phi) + \mathbb{E}\tilde{\ell}_3^2(\varepsilon,\phi)\Big) \|\phi\|_{T_\mu,2}^2 \\
&\le C_{T,K,\tilde{K},H} \|\phi\|_{T_\mu,2}^2 \Big((t^{1+2(\beta_0-H)} + t^{1+2(H-\varsigma_0)\beta_0-2H}) \ell^2(\varepsilon,\phi) + t^{1+2(\alpha_0\wedge(\gamma_0 H)-H)} \\
&\qquad \big(\ell^2(\varepsilon,\phi) + \tilde{\ell}_1^2(\varepsilon,\phi) + \mathbb{E}\tilde{\ell}_2^2(\varepsilon,\phi) + \mathbb{E}\tilde{\ell}_3^2(\varepsilon,\phi)\big)\Big).
\end{aligned} \tag{6.4.35}
$$

As far as $\Theta_6(t,s)$ is concerned, using (A3) and Lemma 2.3.1, we derive that for any $\varepsilon \in [0,1], s,t \in [0,T]$ and $y,z \in \mathbb{R}^d$,

$$
\begin{aligned}
&\bar{b}^\varepsilon(t,y) - \bar{b}^\varepsilon(s,z) = D^L b(t,y,\cdot)(\mathscr{L}_{X_t^{\mu_\varepsilon,\phi}})(X_t^{\mu_\varepsilon,\phi}) - D^L b(s,z,\cdot)(\mathscr{L}_{X_s^{\mu_\varepsilon,\phi}})(X_s^{\mu_\varepsilon,\phi}) \\
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} D^L b(\theta_{s,t},y,\cdot)(\mathscr{L}_{X_t^{\mu_\varepsilon,\phi}})(X_t^{\mu_\varepsilon,\phi}) \mathrm{d}\theta \\
&\quad + \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} D^L b(s,z+\theta(y-z),\cdot)(\mathscr{L}_{X_t^{\mu_\varepsilon,\phi}})(X_t^{\mu_\varepsilon,\phi}) \mathrm{d}\theta \\
&\quad + \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} D^L b(s,z,\cdot)(\mathscr{L}_{X_{s,t}^{\varepsilon,\phi}(\theta)})(X_t^{\mu_\varepsilon,\phi}) \mathrm{d}\theta \\
&\quad + \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} D^L b(s,z,\cdot)(\mathscr{L}_{X_s^{\mu_\varepsilon,\phi}})(X_{s,t}^{\varepsilon,\phi}(\theta)) \mathrm{d}\theta \\
&= \int_0^1 \partial_{\theta_{s,t}} (D^L b(\cdot,y,\cdot)(\mathscr{L}_{X_t^{\mu_\varepsilon,\phi}})(X_t^{\mu_\varepsilon,\phi}))(\theta_{s,t})(t-s) \mathrm{d}\theta \\
&\quad + \int_0^1 \nabla (D^L b(s,\cdot,\cdot)(\mathscr{L}_{X_t^{\mu_\varepsilon,\phi}})(X_t^{\mu_\varepsilon,\phi}))(z+\theta(y-z))(y-z) \mathrm{d}\theta \\
&\quad + \int_0^1 \Big(\mathbb{E}\langle D^L(D^L b(s,z,\cdot)(\cdot)(u))(\mathscr{L}_{X_{s,t}^{\varepsilon,\phi}(\theta)})(X_{s,t}^{\varepsilon,\phi}(\theta)), X_t^{\mu_\varepsilon,\phi} - X_s^{\mu_\varepsilon,\phi}\rangle\Big)|_{u=X_t^{\mu_\varepsilon,\phi}} \mathrm{d}\theta \\
&\quad + \int_0^1 \nabla (D^L b(s,z,\cdot)(\mathscr{L}_{X_s^{\mu_\varepsilon,\phi}})(\cdot))(X_{s,t}^{\varepsilon,\phi}(\theta))(X_t^{\mu_\varepsilon,\phi} - X_s^{\mu_\varepsilon,\phi}) \mathrm{d}\theta,
\end{aligned}
$$

where for any $\theta \in [0,1]$, $\theta_{s,t} := s + \theta(t-s)$ and $X_{s,t}^{\varepsilon,\phi}(\theta) := X_s^{\mu_\varepsilon,\phi} + \theta(X_t^{\mu_\varepsilon,\phi} - X_s^{\mu_\varepsilon,\phi})$.

Then by (A1), (A3) and (6.4.9), we have

$$
\begin{aligned}
|\Theta_6(t,s)| \le C_{T,K,\bar{K},H} \Big[\sum_{i=1}^4 \Lambda_i + |(X_t^{\mu_\varepsilon,\phi} - X_s^{\mu_\varepsilon,\phi}) - (X_t^\mu - X_s^\mu)| \\
+ \varepsilon(t-s)\|\phi\|_{T_\mu,2}\Big] \|\phi\|_{T_\mu,2},
\end{aligned}
$$

where

$$\Lambda_1 := (t-s)\Big(\mathbb{E}\int_0^1 \big|\partial_{\theta_{s,t}}(D^L b(\cdot,y,\cdot)(\mathscr{L}_{X_t^{\mu_\varepsilon,\phi}})(X_t^{\mu_\varepsilon,\phi}))(\theta_{s,t})$$

$$- \partial_{\theta_{s,t}}(D^L b(\cdot,\tilde{y},\cdot)(\mathscr{L}_{X_t^\mu})(X_t^\mu))(\theta_{s,t})\big|^2 \mathrm{d}\theta\Big)^{\frac{1}{2}}\big|_{y=X_t^{\mu_\varepsilon,\phi},\tilde{y}=X_t^\mu},$$

$$\Lambda_2 := \Big(\mathbb{E}\int_0^1 |\nabla(D^L b(s,\cdot,\cdot)(\mathscr{L}_{X_t^{\mu_\varepsilon,\phi}})(X_t^{\mu_\varepsilon,\phi}))(z+\theta(y-z))$$

$$- \nabla(D^L b(s,\cdot,\cdot)(\mathscr{L}_{X_t^\mu})(X_t^\mu))(\tilde{z}+\theta(\tilde{y}-\tilde{z}))|^2 \mathrm{d}\theta\Big)^{\frac{1}{2}}\big|_{y=X_t^{\mu_\varepsilon,\phi},z=X_s^{\mu_\varepsilon,\phi},\tilde{y}=X_t^\mu,\tilde{z}=X_s^\mu}$$

$$\times |X_t^{\mu_\varepsilon,\phi} - X_s^{\mu_\varepsilon,\phi}|,$$

$$\Lambda_3 := (t-s)^H\Big(\mathbb{E}\int_0^1 \Big(\mathbb{E}|D^L(D^L b(s,z,\cdot)(\cdot)(u))(\mathscr{L}_{X_{s,t}^{\varepsilon,\phi}(\theta)})(X_{s,t}^{\varepsilon,\phi}(\theta))$$

$$- D^L(D^L b(s,\tilde{z},\cdot)(\cdot)(v))(\mathscr{L}_{X_{s,t}(\theta)})(X_{s,t}(\theta))|^2\Big)\big|_{u=X_t^{\mu_\varepsilon,\phi},v=X_t^\mu}\mathrm{d}\theta\Big)^{\frac{1}{2}}\big|_{z=X_s^{\mu_\varepsilon,\phi},\tilde{z}=X_s^\mu},$$

$$\Lambda_4 := \Big(\mathbb{E}\Big(\int_0^1 |\nabla(D^L b(s,z,\cdot)(\mathscr{L}_{X_s^{\mu_\varepsilon,\phi}})(\cdot))(X_{s,t}^{\varepsilon,\phi}(\theta))$$

$$- \nabla(D^L b(s,\tilde{z},\cdot)(\mathscr{L}_{X_s^\mu})(\cdot))(X_{s,t}(\theta))|^2 \mathrm{d}\theta\cdot|X_t^{\mu_\varepsilon,\phi} - X_s^{\mu_\varepsilon,\phi}|^2\Big)\Big)^{\frac{1}{2}}\big|_{z=X_s^{\mu_\varepsilon,\phi},\tilde{z}=X_s^\mu},$$

and recall that for any $\theta\in[0,1]$, $X_{s,t}(\theta)=X_s^\mu+\theta(X_t^\mu-X_s^\mu)$.

Note that due to (6.4.6), it follows that as ε or $\|\phi\|_{T_{\mu,2}}$ goes to zero, $X_s^{\mu_\varepsilon,\phi}$ and $X_{s,t}^{\varepsilon,\phi}(\theta)$ converge respectively to X_s^μ and $X_{s,t}(\theta)$ in probability for any $s,t\in[0,T]$ and $\theta\in[0,1]$. Then, using (A3) again and applying the dominated convergence theorem, we deduce that

$$\lim_{\varepsilon\to 0}\mathbb{E}\left(\int_0^T \mathbb{E}\left(t^{2H-1}\left(\int_0^t \frac{|\Theta_6(t,s)|s^{\frac{1}{2}-H}}{(t-s)^{\frac{1}{2}+H}}\mathrm{d}s\right)^2\Big|\mathscr{F}_0\right)\mathrm{d}t\right)^{\frac{1}{2}} = 0 \tag{6.4.36}$$

and

$$\lim_{\|\phi\|_{T_{\mu,2}}\to 0}\sup_{\varepsilon\in(0,1]}\frac{\mathbb{E}\left(\int_0^T \mathbb{E}\left(t^{2H-1}\left(\int_0^t \frac{|\Theta_6(t,s)|s^{\frac{1}{2}-H}}{(t-s)^{\frac{1}{2}+H}}\mathrm{d}s\right)^2\Big|\mathscr{F}_0\right)\mathrm{d}t\right)^{\frac{1}{2}}}{\|\phi\|_{T_{\mu,2}}} = 0. \tag{6.4.37}$$

Hence, combining (6.4.27)-(6.4.30), (6.4.35)-(6.4.36) with (6.4.26) and Remark 6.4.2, we conclude that

$$\lim_{\varepsilon\to 0}\mathbb{E}\left(\int_0^T \mathbb{E}(|J_4(t)|^2|\mathscr{F}_0)\mathrm{d}t\right)^{\frac{1}{2}} = 0$$

and

$$\lim_{\|\phi\|_{T_{\mu,2}}\to 0}\sup_{\varepsilon\in(0,1]}\frac{\mathbb{E}\left(\int_0^T \mathbb{E}(|J_4(t)|^2|\mathscr{F}_0)\mathrm{d}t\right)^{\frac{1}{2}}}{\|\phi\|_{T_{\mu,2}}} = 0.$$

In conjunction with (6.4.20), (6.4.24) and (6.4.25), the above inequalities imply

$$\lim_{\varepsilon\to 0^+}\mathbb{E}|\delta(h^{\varepsilon,\phi}) - \delta(h^{\phi})| = 0, \tag{6.4.38}$$

and

$$\lim_{\|\phi\|_{T_{\mu,2}}\to 0}\sup_{\varepsilon\in(0,1]}\frac{\mathbb{E}|\delta(h^{\varepsilon,\phi}) - \delta(h^{\phi})|}{\|\phi\|_{T_{\mu,2}}} = 0. \tag{6.4.39}$$

Now, we are going to finish the **Proof of Theorem 6.4.1.**

Proof. (a) By (6.4.19), (6.4.38), we verify the conditions (6.3.1)-(6.3.3) of Theorem 6.3.1. We then obtain that $P_T f$ is intrinsically differentiable at μ, and $D^L_\phi(P_T f)(\mu)$ satisfies (6.4.1).

(b) By (6.4.19) and (6.4.39), it is easy to see that (6.3.4) and (6.3.5) of Theorem 6.3.1 hold in this case. Combining this with the result of (a), we obtain that $P_T f$ is L-differentiable at μ. The proof is therefore complete. □

We conclude this section with a remark.

Remark 6.4.3. (i) Compared with the relevant result on MV-SDE driven by the standard Brownian motion ($H = \frac{1}{2}$) shown in [87, Theorem 2.1], one can see that our result Theorem 6.4.1 applies to more general SDEs since we replace $B^{\frac{1}{2}}$ with fractional Brownian motion B^H with arbitrary $H \in (\frac{1}{2}, 1)$ as driving process. Furthermore, due to the appearance of $J_4(t)$ in (6.4.22), essential difficulties are overcome in the analysis of Bismut formula for the L-derivative.

(ii) Combining the above proof with Remark 6.4.1, we can derive the estimate of the L-derivative as the following:

$$\|D^L(P_T f)(\mu)\| = \sup_{\|\phi\|_{T_{\mu,2}} \le 1} |D^L_\phi (P_T f)(\mu)| \le C \left[(P_T f^2)(\mu) - (P_T f(\mu))^2 \right]^{\frac{1}{2}} a(T), \tag{6.4.40}$$

where C is a positive constant depending only on $K, \tilde{K}, H$, and

$$a(T) = C(1+T)\Big(\frac{1}{T^H} + T^{H(\beta_0 - 1)} + T^{H(\gamma_0 - 1)} + T^{\alpha_0 - H} + T^{\delta_0 - H} + 1 + T^{1-H} \Big).$$

Indeed, according to Theorem 6.4.1 and the Hölder inequality, we have

$$\begin{aligned} |D^L_\phi (P_T f)(\mu)|^2 &= \left[\mathbb{E}\Big(f(X^\mu_T) \int_0^T \langle K_H^{-1}(R_H h^\phi)(t), \mathrm{d}W_t \rangle \Big) \right]^2 \\ &= \left[\mathbb{E}\Big((f(X^\mu_T) - P_T f(\mu)) \int_0^T \langle K_H^{-1}(R_H h^\phi)(t), \mathrm{d}W_t \rangle \Big) \right]^2 \\ &\le \left[(P_T f^2)(\mu) - (P_T f(\mu))^2 \right] \int_0^T \mathbb{E}|K_H^{-1}(R_H h^\phi)(t)|^2 \mathrm{d}t. \end{aligned}$$

Taking the same argument as in (6.4.13), (6.4.15) and (6.4.16), applying Remark 6.4.1 and taking into account of the relation $\sup_{s\in[0,T]} \mathbb{E}|\varrho(s)|^2 \le C(\frac{1}{T}+1)^2 \|\phi\|^2_{T_{\mu,2}}$, we obtain the estimate (6.4.40).
In addition, following the same argument as in the proof of [87, Corollary 2.2 (2)] and using (6.4.40), we give the total variation distance estimate for the difference between $\mathscr{L}_{X^\mu_T}$ and $\mathscr{L}_{X^\nu_T}$ with different initial distributions μ and ν:

$$\|\mathscr{L}_{X^\mu_T} - \mathscr{L}_{X^\nu_T}\|_{\mathrm{var}} := \sup_{A \in \mathscr{B}(\mathbb{R}^d)} |\mathscr{L}_{X^\mu_T}(A) - \mathscr{L}_{X^\nu_T}(A)| \le C \mathbb{W}_2(\mu, \nu) a(T).$$

Remark 6.4.4. The general result of the Bismut formula of L-derivative for MV-SDEs driven by fBm can also be applied to the degenerate case by imposing the similar conditions as (A2) and (A3), on the coefficients of the degenerate model. Moreover, as the byproduct, the associated gradient estimate can also be established.

Bibliography

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

[2] O. Amine, E. Coffie, F. Harang, F. Proke, A Bismut-Elworthy-Li formula for singular SDE's driven by a fractional Brownian motion and applications to rough volatility modeling, arXiv:1805.11435v1.

[3] V. Bally, D. Talay, The law of the Euler scheme for stochastic differential equations, I. Convergence rate of the distribution function, *Probab. Theory Related Fields.* **104** (1) (1996), 43–60.

[4] J. Bao, X. Huang, C.Yuan, Convergence rate of Euler-Maruyama scheme for SDEs with Hölder-Dini continuous drifts. *J. Theoret. Probab.* **32** (2019), 848–871.

[5] J. Bao, X. Huang, S.-Q. Zhang, Convergence rate of EM algorithm for SDEs under integrability condition, Arxiv:2009.04781v1

[6] J. Bao, P. Ren, F.Y. Wang, Bismut Formula for Lions Derivative of McKean-Vlasov SDEs with Memory, *J. Differential Equations.* **282** (2021), 285–329.

[7] J. Bao, J. Shao, Weak convergence of path-dependent SDEs with irregular coefficients, arXiv:1809.03088

[8] J. Bao, G. Yin, C. Yuan, Asymptotic analysis for functional stochastic differential equations, Springer, Cham, (2016).

[9] J. Bao, C. Yuan, Large deviations for neutral functional SDEs with jumps, *Stochastics.* **87**(1) (2015), 48–70.

[10] P. Billingsley, Convergence of probability measures. Second edition. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. New York, 1999.

[11] J. M. Bismut, Large deviation and Malliavin Calculus, Birkhäuser, 1984.

[12] L. Bo, T. Zhang, Large deviations for perturbed reflected diffusion processes, *Stochastics: An International Journal of Probability and Stochastics Processes.* **81** (2009), 531–543.

[13] B. Boufoussi, S. Hajji, Functional differential equations driven by a fractional Brownian motion, *Comput. Math. Appl.* **62**(2) (2011), 746–754.

[14]] R. Buckdahn, J. Li, S. Peng, C. Rainer, Mean-field stochastic differential equations and associated PDEs, *Annal. Probab.* **2** (2017), 824–878.

[15] A. Budhiraja, J. Chen, P. Dupuis, Large deviations for stochastic partial differential equations driven by Poisson random measure, *Stochastic Process. Appl.* **123**(2) (2013), 523–560.

[16] A. Budhiraja, P. Dupuis, M. Fischer, Large deviation properties of weakly interacting processes via weak convergence methods, *Ann. Probab.* **40**(1) (2012), 74–102.

[17] A. Budhiraja, P. Dupuis, A. Ganguly, Large deviations for small noise diffusions in a fast Markovian environment, arxiv.org/abs/1705.02948.

[18] A. Budhiraja, P. Nyquist, Large deviations for multidimensional state-dependent shot-noise processes, *J. Appl. Probab.* **52**(4) (2015), 1097–1114.

[19] P. Cardaliaguet, Notes on mean field games, P.-L. Lions lectures at College de France. https://www.researchgate.net/publication/228702832 Notes on Mean Field Games.

[20] R. Carmona, F. Delarue, Probability theory of mean field games with applications I, 1st ed. Probability theory and Stochastic Modelling 84. (2017a). Springer International Publishing.

[21] D. L. Cohn, Measure Theory (Second Edition), Birkhäuser, Boston, Mass., 2013.

[22] L. Coutin, Z. Qian, Stochastic analysis, rough path analysis and fractional Brownian motions. *Probab. Theory Related Fields.* **122**(1) (2002), 108–140.

[23] K. Dareiotis, M. Gerencsér, On the regularisation of the noise for the Euler-Maruyama scheme with irregular drift, *Electron. J. Probab.* **25** (2020), 1–18.

[24]] L. Decreusefond, A.S. Ustunel, S tochstic analysis of the fractional Brownian motion, *Potential Analysis.* **10** (1999), 177–214.

[25] A. Dembo, A. Zeitouni, Large deviations techniques and applications, Springer-Verlag, Berlin Heidelberg, 1998.

[26]] L. Desvillettes, C. Villani, On the spatially homogeneous Landau equation for hard potentials, Part I : existence, uniqueness and smothness, *Comm. Part. Diff. Equat.* **25** (2000), 179–259.

[27] L. Desvillettes, C. Villani, On the spatially homogeneous Landau equation for hard potentials, Part II: H-Theorem and Applications, *Comm. Part. Diff. Equat.* **25** (2000), 261–298.

[28] G. Dos Reis, W. Salkekd, J. Tugaut, Freidlin-Wentzell LDP in path space for McKean-Vlasov equations and the functional iterated logarithm law, *Ann. Appl. Probab.* **29** (2019), 1487–1540.

[29] K.D. Elworthy, X.-M. Li, Formulae for the derivatives of heat semigroups, *J. Funct. Anal.* **125** (1994), 252–286.

[30] X. Fan, Y. Ren, Bismut formulas and applications for stochastic (functional) differential equations driven by fractional Brownian motions, *Stoch. Dyn.* **17** (2017), 19pp.

[31] X. Fan, S. Zhang, On SDEs driven by fractional Brownian motions with irregular drifts. arxiv:1810.01669

[32] X. Fan, Derivative formulas and applications for degenerate stochastic differential equations with fractional noises, *J. Theor. Probab.* **32** (2019), 1360–1381.

[33] M. Ferrante, C. Rovira, Stochastic delay differential equations driven by fractional Brownian motion with Hurst parameter $H > \frac{1}{2}$. *Bernoulli.* **12** (1) (2006), 85–100.

[34] N. Fournier, A. Guillin, From a Kac-like particle system to the Landau equation for hard potentials and Maxwell molecules, *Annales Scientifiques de l'École Normale Supérieure.* **50** (2017), 157–199.

[35] M. Freidlin, Random perturbations of reaction-diffusion equations: the quasideterministic approximation, *Trans. Amer. Math. Soc.* **305** (2) (1988), 665–697.

[36] S. Gadat, F. Panloup, C. Pellegrini, Large deviation principle for invariant distributions of memory gradient diffusions, *Electron. J. Probab.* **18** (81) (2013), 34 pp.

[37] H. Guérin, Existence and regularity of a weak function-solution for some Landau equations with a stochastic approach, *Stoch. Proc. Appl.* **101** (2002), 303–325.

[38] N. Halidias, P.E. Kloeden, A note on the Euler-Maruyama scheme for stochastic differential equations with a discontinuous monotone drift coefficient, *BIT Numer Math.* **48** (2008), 51–59.

[39] W. Hammersley, D. Šiška, L. Szpruch, McKean-Vlasov SDE under measure dependent Lyapunov conditions, arXiv:1802.03974v1

[40] S. Herrmann, P. Imkeller, D. Peithmann, Large deviations and Kramers' type law for self-stabilizing diffusions, *Ann. Appl. Probab.***18** (2008), 1379–1423.

[41] G. Huang, M. Mandjes, P. Spreij, Large deviations for Markov-modulated diffusion processes with rapid switching, *Stochastic Process. Appl.* **126** (6) (2016), 1785–1818.

[42] Y. Hu, Y. Liu, D. Nualart, Rate of convergence and asymptotic error distribution of Euler approximation schemes for fractional diffusions, *Ann. Appl. Probab.* **26**(2) (2016), 1147–1207.

[43] Y. Hu, D. Nualart, Differential equations driven by Hölder continuous functions of order greater than $\frac{1}{2}$, *Stochastic analysis and applications.* 399–413, Abel Symp., 2, Springer, Berlin, 2007.

[44] Y. Hu, D. Nualart, X. Song, A singular stochastic differential equation driven by fractional Brownian motion, *Statist. Probab. Lett.* **78**(14) (2008), 2075–2085.

[45] A. Kohatsu-Higaa, A. Lejay, K. Yasuda, Weak rate of convergence of the Euler-Maruyama scheme for stochastic differential equations with non-regular drift, *Journal of Computational and Applied Mathematics.***326** (2017), 138–158.

[46] V. Konakov, S. Menozzi, Weak error for the Euler scheme approximation of diffusions with non-smooth coefficients, *Electron. J. Probab.***22**(46) (2017), 1–47.

[47] N.V. Krylov, M. Röckner, Strong solutions of stochastic equations with singular time dependent drift, *Probab. Theory Related Fields.* **131** (2005), 154–196.

[48] V. Lemaire and S. Menozzi, On Some non Asymptotic Bounds for the Euler Scheme, *Electron. J. Probab.* **15** (2010), 1645–1681.

[49] G. Leobacher, M. Szölgyenyi, A numerical method for SDEs with discontinuous drift, *BIT Numer Math.* **56** (2016), 151–162.

[50] G. Leobacher, M. Szölgyenyi, A strong order 1/2 method for multidimensional SDEs with discontinuous drift, *Ann. Appl. Probab.* **27** (2017), 2383–2418.

[51] G. Leobacher, M. Szölgyenyi, Convergence of the Euler–Maruyama method for multidimensional SDEs with discontinuous drift and degenerate diffusion coefficient, *Numerische Mathematik.***138** (2018), 219–239.

[52] J. Li, Mean-field forward and backward SDEs with jumps and associated nonlocal quasi- linear integral-PDEs, *Stoch. Proc. Appl.* **128** (2018), 3118–3180.

[53] Y. Li, R. Wang, N. Yao, S. Zhang, A moderate deviation principle for stochastic Volterra equation. *Statistics and Probability Letters.* **122** (2017), 79–85.

[54] R.S. Liptser, A.A. Pukhalskii, Limit theorems on large deviations for semimartingales, *Stochastics and Stochastics Reports.* **38** (4) (1992), 201–249.

[55] J. Liu, L. Yan, On a nonlinear stochastic pseudo-differential equation driven by fractional noise, *Stoch. Dyn.* **18** (1) (2018), 1850002.

[56] K. Liu, T. Zhang, A large deviation principle of retarded Ornstein-Uhlenbeck processes driven by Lévy noise, *Stoch. Anal. Appl.* **32** (5) (2014), 889–910.

[57] W. Liu and M. Röckner, Stochastic Partial Differential Equations: An Introduction, Universitext, Springer, 2015.

[58] W. Liu, Y. Song, J. Zhai, T. Zhang, Large and Moderate deviation principles for McKean-Vlasov SDEs with jumps. arXiv:2011.08403

[59] T. Lyons, Differential equations driven by rough signals, *Revista Matemáática Iberoamericana.* **14** (1998), 215–310.

[60] T.M. Manil, Moderate deviation priciple for the $2D$ stochastic convective Brinkman-Forchheimer equations. *Stochastics.* (2020), 1–55.

[61] X. Mao, Stochastic differential equations and applications, Second Ed, Horwood Publishing Limited, Chichester, 2008.

[62] M. M. Meerschaert, W. Wang, Y. Xiao, Fernique-type inequalities and moduli of continuity for anisotropic Gaussian random fields, *Trans Am Math Soc.* **365** (2013), 1081–1107.

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

[64] Y. S. Mishura, A. Yu. Veretennikov, Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations, arXiv:1603.02212v4

[65] Y. Mishura, G. Shevchenko, The rate of convergence for Euler approximations of solutions of stochastic differential equations driven by fractional Brownian motion, *Stochastics.* **80** (2008), 489–511.

[66] Y. Mishura, G. Shevchenko, Rate of convergence of Euler approximations of solution to mixed stochastic differential equation involving Brownian motion and fractional Brownian motion, *Random Oper. Stoch. Equ.* **19** (4) (2011), 387–406.

[67] Y. Mishura, G. Shevchenko, Existence and uniqueness of the solution of stochastic differential equation involving Wiener process and fractional

Brownian motion with Hurst index $H > 1/2$, *Comm. Statist. Theory Methods.* **40** (2011), 3492–3508.

[68] Y. Mishura, G. Shevchenko, Mixed stochastic differential equations with long-range dependence: Existence, uniqueness and convergence of solutions. *Comput. Math. Appl.* **64** (10) (2012), 3217–3227.

[69] Y. Mishura, T. Shalaiko, Shevchenko, G., Convergence of solutions of mixed stochastic delay differential equations with applications, *Applied Mathematics and Computation.* **257** (2015), 487–497.

[70] S. A. Mohammed, T. Zhang, Large deviations for stochastic systems with memory, *Discrete Contin. Dyn. Syst. Ser.* **6** (4) (2006), 881–893.

[71] T. Müller-Gronbach, L. Yaroslavtseva, Sharp lower error bounds for strong approximation of SDEs with discontinuous drift coefficient by coupling of noise, arXiv:2010.00915

[72] T. Müller-Gronbach, L. Yaroslavtseva, On the performance of the Euler–Maruyama scheme for SDEs with discontinuous drift coefficient, *Ann. Inst. H. Poincaré Probab. Statist.* **56** (2020), 1162–1178.

[73] T. Müller-Gronbach, L. Yaroslavtseva, A strong order 3/4 method for SDEs with discontinuous drift coefficient, *IMA Journal of Numerical Analysis.* **00** (2020), 1–31.

[74] A. Neuenkirch, Optimal approximation of SDE's with additive fractional noise, *J. Complexity.* **22** (4) (2006), 459–474.

[75] A. Neuenkirch, I. Nourdin, S. Tindel, Delay equations driven by rough paths, *Electron. J. Probab.* **13** (2008), 2031–2068.

[76] A. Neuenkirch, M. Szölgyenyi, L. Szpruch, An adaptive Euler-Maruyama scheme for stochastic differential equations with discontinuous drift and its convergence analysis, *SIAM J. Numer. Anal.* **57** (2019), 378–403.

[77] A. Neuenkirch, M. Szölgyenyi, The Euler-Maruyama Scheme for SDEs with Irregular Drift: Convergence Rates via Reduction to a Quadrature Problem, arXiv:1904.07784

[78] H.-L. Ngo, D. Taguchi, Approximation for non-smooth functionals of stochastic differential equations with irregular drift, *J. Math. Anal. Appl.* **457** (2018), 361–388.

[79] H. L. Ngo, D. Taguchi, Strong rate of convergence for Euler-Maruyama approximation of stochastic differential equations with irregular coefficients, *Math. Comp.* **85** (2016), 1793–1819.

[80] H. L. Ngo, D. Taguchi, On the Euler-Maruyama approximation for one-dimensional stochastic differential equations with irregular coefficients, *IMA J. Numer. Anal.* **37** (2017), 1864–1883.

[81] D. Nualart, The malliavin calculus and related topics, Second edition, Springer-Verlag, Berlin, 2006.

[82] D. Nualart, Y. Ouknine, Regularization of differential equations by fractional noise,*Stochastic Process. Appl.* **102** (1) (2002), 103–116.

[83] D. Nualart, A. Răşcanu, Differential equations driven by fractional Brownian motion, *Collect. Math.* **53** (1) (2002), 55–81.

[84] P. Pawel, S. Michaela, Existence, uniqueness, and approximation of solutions of jump-diffusion SDEs with discontinuous drift, arXiv:1912.04215

[85] G. Reis, W.Salkeld, J.Tugaut, Freidlin-Wentzell LDP in path space for McKean-Vlasov equations and the Functional Iterated Logarithm Law, *Ann. Appl. Probab.* **29** (2019), 1487–1540.

[86] P. Ren, F.-Y. Wang, Space-distribution PDEs for path independent additive functionals of McKean-Vlasov SDEs, arXiv:1805.10841

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

[88] M. Röckner, T. Zhang, Stochastic evolution equations of jump type: existence, uniqueness and large deviation principles, *Potential Anal.* **26** (3) (2007), 255–279.

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

[90] B. Saussereau, Transportation inequalities for stochastic differential equations driven by a fractional Brownian motion, *Bernoulli.* **18** (1) (2012), 1–23.

[91] E. M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton University Press, Princeton, 1970.

[92] S. Shklyar, G. Shevchenko, Y. Mishura, V. Doroshenko, O. Banna, Approximation of fractional Brownian motion by martingales, *Methodol. Comput. Appl. Probab.* **16** (3) (2014), 539–560.

[93] D.W. Stroock, An introduction to the theory of large deviations, Springer-Verlag, Berlin, 1984.

[94] A.-S. Sznitman, Topics in propagation of chaos, In Ecole d'Etd́e Probabilites de Sain-Flour XIX-1989, Lecture Notes in Mathematics 1464, 165-251, Springer, Berlin, 1991.

[95] C. Villani, On the spatially homogeneous Landau equation for Maxwellian Mocecules, *Math. Mod. Meth. Appl. Sci.* **8** (1998), 957–983.

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

[97] F.-Y. Wang, Estimates for invariant probability measures of degenerate SPDEs with singular and path-dependent drifts, *Probab. Theory Relat. Fields.* **172** (2018), 1181–1214.

[98] F.-Y. Wang, Harnack inequalities and applications for stochastic partial differential equations, Springer, 2013, Berlin.

[99] L. Xie, X. Zhang, Sobolev differentiable flows of SDEs with local Sobolev and super-linear growth coefficients, *Ann. Probab.* **44** (2016), 3661–3687.

[100] L. Xie, X. Zhang, Ergodicity of stochastic differential equations with jumps and singular coefficients, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques.* **56** (2020), 175–229.

[101] L. C. Young, An inequality of the Hölder type, connected with Stieltjes integration, *Acta Math.* **67** (1) (1936), 251–282.

[102] M. Zähle, Integration with respect to fractal functions and stochastic calculus, *I. Probab. Theory Related Fields.* **111** (3)(1998), 333–374.

[103] X. Zhang, Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients, *Stochastic Process. Appl.* **115** (2015), 1805–1818.