Branching particle systems, stochastic partial differentiable equations and nonlinear rough path analysis

Panqiu Xia

B.A. Mathematics, Wuhan University, 2011

M.A. Mathematics, University of Macau, 2014

Submitted to the graduate degree program in Department of Mathematics and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Committee members

David Nualart, Chairperson

Yaozhong Hu, Chairperson

Zhipeng Liu

Xuemin Tu

Jianbo Zhang, Department of Economics

Date defended: July 20, 2020

The Thesis Committee for Panqiu Xia certifies
that this is the approved version of the following thesis :

Branching particle systems, stochastic partial differentiable equations and nonlinear rough path analysis

____

David Nualart, Chairperson

Date approved: July 20, 2020

Abstract

In this dissertation, we study some problems related to the stochastic partial differential equations, branching particle systems and rough path analysis.

In Chapter 1, we provide a brief introduction and background of the topics considered in this dissertation.

In Charter 2, a branching particle system in a random environment has been studied. Under the Mytnik-Sturm branching mechanism, we prove that the scaling limit of this particle system exists. This limit has a Lebesgue density that is a weak solution to a stochastic partial equation. We also investigate the Hölder continuity of this limit, and prove it is $1/2-\varepsilon$ in time and $1-\varepsilon$ in space.

In Chapter 3, a theory of nonlinear rough paths is developed. Following the idea of Lyons and Gubinelli, we define a nonlinear integral of rough functions. Then we study a rough differential differential equation, and obtain the local and global existence and uniqueness of this solution under suitable sufficient conditions. As an application, we consider the transport equation with rough vector field and observe the classical solution formula does not satisfy the rough equation. Indeed, it is the solution to the transport equation with compensators.

In Chapter 4, we study the parabolic Anderson model of Skorokhod type with very rough noise in time. By using the Feynman-Kac formula for moments, we obtain the upper and lower bounds for moments of the solution.

Acknowledgements

First of all, I would like to take this opportunity to express my sincere gratitude to my advisors Professor Yaozhong Hu and Professor David Nualart for their enthusiastic guidance and support. They provided me a lot of ideas and knowledge in research, and taught me how to think and work as a mathematician. Without their patient guidance, I am not able to move forward even one step on the academic journey.

I also want to thank Professor Zhipeng Liu, Professor Xuemin Tu and Professor Jianbo Zhang for their serving on my dissertation committee.

I gratefully acknowledge Professor Terry Soo, Professor Jin Feng and Professor Weishi Liu for their helpful advices and support. During the days at the University of Kansas, many staffs help me with a lot of efforts. Particularly, I want to thank Kerrie Brecheisen, Gloria Prothe, and Lori Springs for their assistance in different ways.

I thank Professor Jie Xiong at SUSTech for the helpful discussions in branching particle systems.

I would like to thank all my friends who gave me help and encouragement. Especially, I am thankful to Jian Song, Xiaoming Song, Le Chen, Fei Lu, Jingyu Huang, Guannan Hu, Yanghui Liu, Hongjuan Zhou, Arturo Jaramillo, Guangqu Zheng, Junxi Zhang, Yiying Cheng, Yu Yin, Yanru Su, Xiong Wang, Jinjin Zhang, Baobin Wang, Zhi Wang, Yong Chen, Xingchun Peng, Chaozheng Li, Hao Chen, Yi Li, Jinan Wang, Nin Sun and many others, for their numerous discussion in mathematics and other sciences that improve my understanding of the world.

Last by not least, I thank my family, especially, my wife Jiayu Zheng for their unconditional love, encouragement and support. This dissertation is dedicated to them.

Contents

Chapter 1

Introduction

This dissertation concerns with several topics in branching particle systems, stochastic partial differential equations and nonlinear rough paths analysis. It consists of three research articles, which are listed as follows,

1. Hölder continuity of the solutions to a class of SPDE's arising from branching particle systems in a random environment, with Yahozhong Hu and David Nualart, Electron. J. Probab., **24**, no. 105, (2019), 1 - 52. [48]

2. On nonlinear rough paths, with David Nualart, ALEA, **17**, (2020), 545 – 587, [73]

3. Intermittency for the Parabolic Anderson model of Skorohod type driven by a rough noise, with Nicholas Ma and David Nualart, Electron. Commun. Probab., **25**, no. 48, (2020), 1 - 10. [67]

In paper 1, we study a branching particle system in a random environment, observe that after an appropriate scaling, the limit of the empirical measures of the system exists as a measure-valued process. This limit has a Lebesgue density that is a weak solution to a stochastic partial differential equation. Then, by using the techniques of Malliavin calculus, we find that the solution is jointly Hölder continuous with exponents $1/2-\varepsilon$ in time and $1-\varepsilon$ in space.

In paper 2, we develop a theory of nonlinear rough paths following the ideas of Lyons and Gubinelli. By a stability analysis, we obtain suitable sufficient conditions for the local and global existences and uniqueness of the solution to the Kunita-type equation $Y_t = \xi + \int_0^t W(ds, Y_s)$. Finally we apply this theory to the transport equation with rough vector fields.

In paper 3, a Skorohod type parabolic Anderson model with rough noise in time is studied.

By using the Feynman-Kac formula for the moments of the solution, we find the upper and lower bounds for the moments.

In the rest of this chapter, we introduce the background and main results of each topic studied in the present dissertation in details.

1.1 Branching particle systems in a random environment

The study of branching particle systems started from 1960s, since the pioneer work of Jiřina [51] and Watanabe [82]. It has been found that the limits of empirical measures of branching particle systems are measure-valued Markov processes. Then, it was investigated that measure-valued Markov processes are highly connected to stochastic evolution equations by Dawson [20]. The theory of branching particle systems is a combination of spatial motion of particles and a continuous branching phenomenons. Additionally, it is closely related to a class of semilinear partial differential equations. We refer the readers to the monographs of Dawson [21, 22], Dynkin [28, 29], Etheridge [30], Le Gall [59], Li [62] and Perkins [74] for a detailed account of those developments.

In the last century corner, Wang [80, 81] and Skoulakis & Adler [75] introduced a random environment applying to the system, where the motion of each particle is governed by a random environment. On the other hand, Mytnik [70] considered a branching model that a random environment affects the branching rate of the particles with small perturbation, and Sturm [78] modified Mytnik's model such that the branching is a rather rare event and totally depending on the environment. In Chapter 2, we study a branching particle system, in which the particle motion follows Wang's model, while the branching obeys Mytnik-Sturm's mechanism.

Consider a d-dimensional branching particle system in a random environment. For any integer $n \geq 1$, the branching events happen at time $\frac{k}{n}$, $k = 1, 2, \dots$. The dynamics of each particle, labeled by a multi-index α, is described by the stochastic differential equation (SDE):

$$dx_t^{\alpha,n} = dB_t^{\alpha} + \int_{\mathbb{R}^d} h(y - x_t^{\alpha,n}) W(dt, dy), \tag{1.1.1}$$

where h is a $d \times d$ matrix-valued function on $\mathbb{R}^d$, whose entries $h^{ij} \in L^2(\mathbb{R}^d)$, B^α are d-dimensional independent Brownian motions, and W is a d-dimensional space-time white Gaussian random field on $\mathbb{R}_+ \times \mathbb{R}^d$ independent of the family $\{B^\alpha\}$. The random field W can be regarded as the random environment for the particle system. The existence and uniqueness of the Feller process $x^{\alpha,n}$ that solves the SDE (1.1.1) will be proved in Section 2.1.

At any branching time each particle dies and it randomly generates offspring. The new particles are born at the death position of their parents, and inherit the branching-dynamics mechanism. As we already stated before, the branching mechanism follows the one introduced by Mytnik [70], and studied further by Sturm [78]. Let $X^n = \{X_t^n, t \geq 0\}$ denote the empirical measure of the particle system. One of the main results of this work is to prove that the empirical measure-valued processes converge weakly to a process $X = \{X_t, t \geq 0\}$, such that for almost every $t \geq 0$, X_t has a density $u_t(x)$ almost surely. By using the techniques of Malliavin calculus, we also establish the almost surely joint Hölder continuity of u with exponent $\frac{1}{2} - \varepsilon$ in time and $1 - \varepsilon$ in space for any $\varepsilon > 0$.

To compare our results with the classical ones. Let us recall briefly some existing work in the literature. The one-dimensional model was initially introduced and studied by Wang [80, 81]. In these papers, he proved that under the classical Dawson-Watanabe branching mechanism, the empirical measure X^n converges weakly to a process $X = \{X_t, t \geq 0\}$, which is the unique solution to a martingale problem.

For the above one dimensional model, Dawson et al. [25] proved that for almost every $t > 0$, the limit measure-value process X has a density $u_t(x)$ a.s. and u is the weak solution to the following stochastic partial differential equation (SPDE):

$$\begin{aligned} u_t(x) =& \mu(x) + \int_0^t \frac{1}{2}(1 + \|h\|_2^2)\Delta u_s(x)ds - \int_0^t \int_{\mathbb{R}} \nabla_x [h(y-x)u_s(x)]W(ds,dy) \\ &+ \int_0^t \sqrt{u_s(x)} \frac{V(ds,dx)}{dx}, \end{aligned} \tag{1.1.2}$$

where $\|h\|_2$ is the L^2-norm of h, and V is a space-time white Gaussian random field on $\mathbb{R}_+ \times \mathbb{R}$

independent of W.

Suppose further that h is in the Sobolev space $H_2^2(\mathbb{R})$ and the initial measure has a density $\mu \in H_2^1(\mathbb{R})$. Then Li et al. [63] proved that $u_t(x)$ is almost surely jointly Hölder continuous. By using the techniques of Malliavin calculus, Hu et al. [44] improved their result to obtain the sharp Hölder continuity: they improved the Hölder exponents to be $\frac{1}{4} - \varepsilon$ in time and $\frac{1}{2} - \varepsilon$ in space, for any $\varepsilon > 0$.

In Chapter 2, we are interested in higher dimensions ($d > 1$). However in this case, the super Brownian motion (a special case when $h = 0$) does not have a density (c.f. Corollary 2.4 of Dawson & Hochberg [23]). Thus in higher dimensional case we have to abandon the classical Dawson-Watanabe branching mechanism and adopt the Mytnik-Sturm one. As a consequence, the difficult term $\sqrt{u_s(x)}$ in the SPDE (1.1.2) becomes $u_s(x)$ (see equation (2.2.1) in Section 2.2 for the exact form of the equation).

We follow the approach introduced in Hu et al. [44] to study the Hölder continuity of the conditional density of a particle motion using Malliavin calculus. However, because of the multi-dimensional setting considered here, new difficulties arise. On one hand, the integration by parts formulas require higher order Malliavin derivatives which make computations more complex. To lower the order of Malliavin differentiability in our framework, we use the combination of Riesz transform and Malliavin calculus, previously studied in depth by Bally & Caramellino [5] (see Appendix 2.7 for the density formula that we are using). Another difficulty is the fact that in the one-dimensional case considered in Hu et al. [44], the Malliavin derivative can be expressed explicitly and this type of formula for the Malliavin derivative is no longer available here. We have to use another approach to obtain appropriate sharp estimates. More details are given in Appendix 2.7.

Chapter 2 is organized as follows. In Section 2.1 we shall briefly describe the branching mechanism used in Chapter 2. In Section 2.2 we state the main results obtained in this chapter. These include three theorems. The first one (Theorem 2.2.3) is about the existence and uniqueness of a (linear) stochastic partial differential equation (equation (2.2.1)), which is proved (Theorem 2.2.2) to be satisfied by the density of the limiting empirical measure process X^n of the particle system (see (2.1.12)). The core result of this chapter is Theorem 2.2.4 which intends to give sharp Hölder continuity of the solution $u_t(x)$ to (2.2.1).

Section 2.3 presents the proofs for Theorems 2.2.2 and 2.2.3. The proof of Theorem 2.2.4 is the objective of the remaining sections. Firstly, in Section 2.4, we focus on the one-particle motion with no branching. By using the techniques from Malliavin calculus, we obtain a Gaussian type estimates for the transition probability density of the particle motion conditional on W. This estimate plays a crucial role in the proof of Theorem 2.2.4. In Section 2.5, we derive a conditional convolution representation of the weak solution to the SPDE (2.2.1), which is used to establish the Hölder continuity. In Section 2.6, we show that the solution u to (2.2.1) is Hölder continuous.

Lastly, the martingale problem (2.3.4) - (2.3.5) is introduced in Section 2.3 to prove Theorems 2.2.2 and 2.2.3. The well-posedness of the martingale problem can be proved under the assumption that the initial measure has a bounded density. We conjecture that it also holds for an arbitrary finite initial measure. We will not pursue this in this chapter (see Remark 2.3.12 (ii)).

1.2 Nonlinear rough paths

Nonlinear integrals in the sense of Young have been studied in recent years (c.f. Catellier & Gubinelli [13], Chouk & Gubinelli [17, 16] and Hu & Lê [42]). In these papers, the authors consider the following nonlinear integral

$$\mathscr{I}_{s,t} = \int_s^t W(dr, Y_r), \tag{1.2.1}$$

where W is a function on $[0,T] \times \mathbb{R}^d$ with values in $\mathbb{R}^d$, that is τ-Hölder continuous in time and λ-Hölder continuous in space, and $Y : [0,T] \to \mathbb{R}^d$ is γ-Hölder continuous. Under the assumption $\tau + \lambda\gamma > 1$, the nonlinear integral (1.2.1) is well-defined in the sense of Young [86]. That is, $\mathscr{I}_{s,t}$

is the limit of the following linear approximations as $|\pi| \to 0$

$$\sum_{k=1}^{n} W_{t_{k-1},t_k}(Y_{t_{k-1}}) := \sum_{k=1}^{n} \left[W(t_k, Y_{t_{k-1}}) - W(t_{k-1}, Y_{t_{k-1}})\right],$$

where $\pi = (s = t_0 < t_1 < \cdots < t_n = t)$ is a partition of the interval $[s,t]$ and $|\pi| := \max_{1\le k\le n} |t_k - t_{k-1}|$. As an example, one can define a pathwise nonlinear integral of the form (1.2.1), where W is a fractional Brownian sheet with Hurst parameters $H_0 \in (\frac{1}{2}, 1)$ in time and $H_1 = \cdots = H_d = H$ in space, such that $H_0 + \frac{1}{2}H > 1$, and Y is a d-dimensional standard Brownian motion. By applying this theory of nonlinear Young's integrals, Hu & Lê [42] studied the following transport equation with distributional vector field (see also [13, 34]):

$$\frac{\partial}{\partial t}u(t,x) + Du(t,x)\frac{\partial}{\partial t}W(t,x) = 0, \tag{1.2.2}$$

where D denotes the spatial derivative operator. The existence and uniqueness of the solution to (1.2.2) with $\mathscr{C}^{1+\lambda_0}_{loc}(\mathbb{R}^d;\mathbb{R})$-valued initial condition were proved in [42] assuming that $(1+\lambda_0)\tau > 1$. They also provided a formula for the solution:

$$u(t,x) = h(Z_t(x)), \tag{1.2.3}$$

where h is the initial condition, Z_t is the inverse of Y_t, and Y is the solution to the following nonlinear differential equation:

$$Y_t(x) = x + \int_0^t W(ds, Y_s(x)). \tag{1.2.4}$$

On the other hand, applying the theory of nonlinear integrals to the stochastic heat equation, Hu and Lê also gave a pathwise proof of the Feynman-Kac formula, which provides an alternative method to study this topic (c.f. [43, 47] for a probabilistic approach).

The purpose of Chapter 3 is to extend the theory of nonlinear integrals to the case when the functions W and Y are rougher, that is $\tau+\lambda\gamma<1$. In this situation, Young's approach fails. The following example, inspired by the lecture notes from Zanco (see Example 3.6 of Zanco [87]), provides a non-standard nonlinear rough path behavior in $\mathbb{R}$. For any $n\in\mathbb{N}$, $t\in[0,T]$ and $x,y\in\mathbb{R}$, we define

$$F(x,y)=e^{xy},\ X_t^{(n)}=\frac{1}{n}\cos(2\pi n^2 t)\text{ and }Y_t^{(n)}=\frac{1}{n}\sin(2\pi n^2 t).$$

Then $F(X_t^{(n)},y)$ converges to 1 and $Y_t^{(n)}$ converges to 0 uniformly on compact sets as $n\to\infty$. On the other hand, however, the following integral

$$\int_0^1 F(dX_t^{(n)},Y_t^{(n)})=-\frac{1}{4}\int_0^{4\pi}\exp\Big(\frac{1}{2n^2}\sin(n^2 s)\Big)ds\to-\pi,$$

by dominated convergence theorem, as $n\to\infty$.

In the linear situation, a useful tool to deal with the integration of rough functions is the theory of rough paths. This theory has been developed from the pioneering work of Lyons since the early nineties (c.f. Lyons [64, 65]) to study d-dimensional dynamical systems of the form

$$dY_t=f(Y_t)dX_t,\ t\in[0,T],$$

where the driven signal X_t is α-Hölder continuous and $\alpha\in(0,\frac{1}{2}]$. The main idea of the rough path analysis is as follows. Let $p=\lfloor\frac{1}{\alpha}\rfloor$, and let $T^{(p)}$ be a p-step truncated tensor algebra given by the expression

$$T^{(p)}:=\mathbb{R}\oplus(\mathbb{R}^d)\oplus(\mathbb{R}^d)^{\otimes 2}\oplus\cdots\oplus(\mathbb{R}^d)^{\otimes p}.$$

The rough path associated to X is a lifting of X to a $T^{(p)}$-valued function on $[0,T]^2$, denoted by $S^{(p)}(X)$, in such a way that when X is piecewise differentiable, the function

$$S_{s,t}^{(p)}=(1,X_{s,t}^1,X_{s,t}^2,\ldots,X_{s,t}^p),$$

and each component $X^i_{s,t}$ is the ith iterated integral of X on the time interval $[s,t] \subset [0,T]$. Suppose that f is a smooth function, then the integral of $f(X)$ against X on $[s,t]$ can be approximated by

$$\int_s^t f(X_r)dX_r \approx f(X_s)X^1_{s,t} + f'(X_s)X^2_{s,t} + \cdots + f^{(p-1)}(X_s)X^p_{s,t}, \tag{1.2.5}$$

with an error of order $O(|t-s|^{(p+1)\alpha})$. Because $(p+1)\alpha > 1$, the error term vanishes in the limit, which explains the choice $p = \lfloor \frac{1}{\alpha} \rfloor$. This allows us to define the integral by passing the limit as $|\pi| \to 0$ of the following expression

$$\sum_{k=1}^{n} \sum_{i=0}^{p} f^{(i-1)}(X_{t_{k-1}})X^i_{t_{k-1},t_k},$$

where $\pi = (s = t_1 < \cdots < t_n = t)$.

Suppose that $\alpha \in (\frac{1}{3}, \frac{1}{2}]$. Gubinelli (see [37]) generalized the integration of "1-forms", which means the integrand is a function $f(X_t)$ of the driving signal, to a class of rough functions called "controlled rough paths". A controlled rough path (by X), is a function $Y : [0,T] \to \mathbb{R}^d$ whose increment on an interval $[s,t]$ can be written in the following way: $Y_{s,t} = Y'_s X_{s,t} + R^Y_{s,t}$, for some $\mathbb{R}^d \otimes \mathbb{R}^d$-valued α-Hölder continuous function Y' and some $\mathbb{R}^d$-valued 2α-Hölder continuous function R^Y. In this case, the approximation of the integral is the following

$$\int_s^t Y_r dX_r \approx Y_s X^1_{s,t} + Y'_s X^2_{s,t}.$$

For a more detailed account on this topic, we refer the readers to the books of Friz & Hairer [36] and Lyons & Qian [66]. An alternative approach to deal with the integration of "non-1-forms" based on fractional calculus was developed in [6, 45].

In Chapter 3, we will extend the nonlinear Young's integral to the rough case by using Gubinelli's approach, and assuming a Hölder regularity of order $\alpha \in (\frac{1}{3}, \frac{1}{2}]$. Chapter 3 is organized in the following way. In Section 3.1 we give brief review of the preliminaries about (linear) rough paths. In Section 3.2 we introduce a nonlinear variant of rough paths. By definition a nonlinear rough path is a pair $(W,\mathbb{W})$ such that $W(t,x)$ is a function of two variables, $(t,x) \in [0,T] \times V$, where V is Banach space. The component $\mathbb{W}_{s,t}(x,y)$ should be interpreted as the integral

$$\mathbb{W}_{s,t}(x,y) = \int_s^t DW(dr,y)W_{r,s}(x),$$

for any $0 \leq s \leq t \leq T$, where $DW(r,y)$ denotes the Fréchet derivative of W with respect to the spatial argument y. We also assume that $(W,\mathbb{W})$ satisfies certain properties, including α-Hölder continuity and a version of Chen's relation. Then, a nonlinear rough integral can be approximated in the following way:

$$\int_s^t W(dr,Y_r) \approx W_{s,t}(Y) + \mathbb{W}_{s,t}(\dot{Y}_s, Y_s),$$

where $\dot{Y}$ is the Gubinelli derivative of Y in the context of nonlinear rough paths. We prove that the nonlinear rough integral is a nonlinear controlled rough path and we establish some properties of nonlinear rough integrals.

In Section 3.3, we consider the following rough differential equation (RDE):

$$Y_t = \xi + \int_0^t W(dr,Y_r), \tag{1.2.6}$$

where $(W,\mathbb{W})$ is an α-Hölder nonlinear rough path. Local and global existence and uniqueness of the solution to the RDE (1.2.6) are proved in this section. We also obtain some estimates for the solution to this equation. This type of RDEs was previously studied by Bailleul and his collaborators (c.f. [2, 3, 4]) under some boundedness assumption of W. Here, we study this equation via a different approach, and improve their results removing boundedness conditions.

Another approach to equation (1.2.6) was introduced by Brault & Lejay [8, 9, 10]. In these papers, the authors introduced the almost flow $\phi_{s,t}(x)$. In comparison with our setting, $\phi_{s,t}(x)$ is equivalent to $W_{s,t}(x) + \mathbb{W}_{s,t}(x,x) + x$. Then without the analysis of the rough integrals, a solution to equation (1.2.6) can be constructed as the limit of the following iterations over partitions $\pi =$

$(0 = t_0 < t_1 < \cdots < t_n = t)$

$$\phi_{t_{n-1},t_n} \circ \cdots \circ \phi_{t_1,t_0}(\xi), \quad \text{as } |\pi| \to 0.$$

In Section 3.4, we study nonlinear rough paths as function space-valued linear rough paths. Then, we prove that under some assumptions, these two approaches are equivalent. Despite this, we still prefer to keep the analysis in Sections 3.2 and 3.3. Firstly, the approach to define nonlinear rough paths in Section 3.2 is more intuitive than the latter method based on abstract spaces. Additionally, in order to interpret nonlinear rough paths as function space-valued linear rough paths, a stronger assumption is required, namely, the existence of the integral $\mathscr{W}_{s,t} = \int_s^t W_{s,r} \otimes dW_r$, whereas, in Sections 3.2 and 3.3, we only need to define the integral $\mathbb{W}_{s,t}(x,y) = \int_s^t DW(dr,y)W_{r,s}(y)$.

Section 3.5 contains some applications of nonlinear rough paths. In Section 3.5.1, we provide a generalized Itô-type formula for (nonlinear) controlled rough paths. In Section 3.5.2 we analyze the gradient flow of the following equation with spatial parameter,

$$Y_t(x) = x + \int_0^t W(dr, Y_r(x)),$$

where $x \in \mathbb{R}^d$ and $W : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is a nonlinear rough path. We will prove that under some assumptions, $Y_t(x)$ is differentiable in x. In addition for every $(t,x) \in [0,T] \times \mathbb{R}^d$, the gradient $DY_t(x)$ is an invertible matrix. Thus, for any fixed $t \in [0,T]$, by the Inverse Function Theorem in $\mathbb{R}^d$, there exists $Z_t : \mathbb{R}^d \to \mathbb{R}^d$ such that $Z_t(Y_t(x)) = Y_t(Z_t(x)) = x$ for all $x \in \mathbb{R}^d$. Assume that $h \in \mathscr{C}^4_{loc}(\mathbb{R}^d;\mathbb{R})$. Because the structure of W here is rougher than in Young's case, it turns out that $h(Z_t(x))$ does not satisfies the transport equation (1.2.2). In Section 3.5.3 we will prove that $h(Z_t(x))$ is indeed the solution to the following transport equation with compensators

$$\begin{aligned}\frac{\partial}{\partial t}u(t,x) + Du(t,x)\frac{\partial W(t,x)}{\partial t} &= \frac{1}{2}Du(t,x)\frac{\partial \langle\!\langle DW(x), W(x)\rangle\!\rangle_{0,t}}{\partial t} \\ &+ \frac{1}{2}Du(t,x)\frac{\partial \langle\!\langle W(x), DW(x)\rangle\!\rangle_{0,t}}{\partial t} + \frac{1}{2}D^2u(t,x)\frac{\partial \langle W(x)\rangle_{0,t}}{\partial t}. \qquad (1.2.7)\end{aligned}$$

Furthermore, the solution is unique in the space $\mathscr{C}^{\alpha,3}_{loc}([0,T]\times\mathbb{R}^d;\mathbb{R})$. A similar transport equation with rough vector field was also studied by Catellier in [12].

Finally, we finish Chapter 3 by providing some examples for nonlinear rough paths and transport equations with compensators in Section 3.6.

1.3 Parabolic Anderson model of Skorohod type

The parabolic Anderson model is a special case of the random Schrödinger equation. The study of this equation dates back to the late 1950s when Anderson published his ground break paper [1]. In this paper, Anderson predicted a phenomenon, which is now called the Anderson localization, that the absence of diffusion happens in a lattice system in a random medium. Because its connections to other fields in mathematics like branching processes in a random environment, large deviations and stochastic partial differential equations, as well as numerous applications to physical theories (c.f. Carmona & Molchanov [11], Hairer [38], Havlin & ben Avraham [39], Zel'dovich et al. [89, 90]), the parabolic Anderson model became a popular topic among probabilistic and mathematical physicists in recent decades. We refer the readers to the books of Kirsch [54] and König [55] for a detailed survey of this model.

In Chapter 4, we consider the following parabolic Anderson model of Skorohod type

$$\frac{\partial}{\partial t}u(t,x)=\frac{1}{2}\Delta u(t,x)+u(t,x)\diamond\frac{\partial}{\partial t}W(t,x), \tag{1.3.1}$$

where $\diamond$ denotes the Wick product. The noise $W=\{W(t,x),(t,x)\in\mathbb{R}_+\times\mathbb{R}^d\}$ is a Gaussian random field, that is a fractional Brownian motion of Hurst parameter $H\in(0,\frac{1}{2})$ in time, and has a correlation in space given by a function Q, namely,

$$\mathbb{E}[W(t,x)W(s,y)]=\frac{1}{2}\big(t^{2H}+s^{2H}-|t-s|^{2H}\big)Q(x,y),$$

for all $s,t\in\mathbb{R}_+$ and $x,y\in\mathbb{R}^d$.

A similar equation in the Stratonovich sense, where the Wick product in (1.3.1) is replaced by the ordinary product, has been studied by Hu et al. [43] and Chen et al. [14]. In these papers, it has been proved that under suitable sufficient hypotheses on the correlation function Q, the Stratonovich type equation with bounded initial condition has a unique solution, which admits a Feynman-Kac representation. Additionally, by using the Feynman-Kac formula for the moments of the solution, the authors in [14] studied the intermittency phenomenon for the solution. Under certain assumptions on Q, the inequality holds for some $\alpha, \beta \in (0,1)$,

$$\underline{C}_x \exp\left(\underline{C} n^{\frac{2-\beta}{1-\beta}} t^{\frac{2H+\beta}{1-\beta}}\right) \leq \mathbb{E}\left[u(t,x)^n\right] \leq \overline{C}_x \exp\left(\overline{C} n^{\frac{2-\alpha}{1-\alpha}} t^{\frac{2H+\alpha}{1-\alpha}}\right)$$

for all $t \geq 1$, $x \in \mathbb{R}^d$ and $n \geq 1$, where $\underline{C}$ and $\overline{C}$ are positive constants depending on $d, H, \alpha, \|u_0\|_\infty$ and $\underline{C}_x, \overline{C}_x > 0$ depend on $d, H, \alpha, \|u_0\|_\infty$ and x.

In Chapter 4, we will study the intermittency for the Skorohod equation (1.3.1). The upper bounds for the moments of the solution can be easily obtained. This is due to the fact that the solution to the Skorohod equation is bounded by the solution to the equation of Stratonovich type. For the same reason, to get lower bounds is more involved. By using the Feynman-Kac formula for the moments, we see that in comparison with the Stratonovich case, the exponent in our case contains an additional negative term. This increases the difficulty to estimate lower bounds for the moments. To settle this difficulty, we pin the Brownian motion B_t at the middle point $t/2$, and observe that conditional on $B_{t/2} = r$, B_s is a Brownian bridge before time $t/2$, and an independent Brownian motion after $t/2$. Then, we estimate the probability of the event that the supremum and the Hölder norm of the Brownian bridge (motion) are bounded above and below by appropriate constants. This allows us to find a lower bound for the moments of the solution.

Chapter 4 is organized as follows. In Section 4.1, we give a brief introduction on the Malliavin calculus and present the precise definition of the solution to equation (1.3.1). In Section 4.2, following the idea of Hu et al. [43], we prove that equation (1.3.1) has a unique solution and give the Feynman-Kac formula and the chaos expansion of the solution. Then, we provide the upper bounds for the moments. Finally, the lower bounds for the moments are proved in Section 4.3.

Chapter 2

Branching particle systems and related stochastic partial differential equations

In this chpater, we consider a d-dimensional branching particle system in a random environment. Suppose that the initial measures converge weakly to a measure with bounded density. Under the Mytnik-Sturm branching mechanism, we prove that the corresponding empirical measure X_t^n converges weakly in the Skorohod space $D([0,T];M_F(\mathbb{R}^d))$ and the limit has a density $u_t(x)$, where $M_F(\mathbb{R}^d)$ is the space of finite measures on $\mathbb{R}^d$. We also derive a stochastic partial differential equation $u_t(x)$ satisfies. By using the techniques of Malliavin calculus, we prove that $u_t(x)$ is jointly Hölder continuous in time with exponent $\frac{1}{2}-\varepsilon$ and in space with exponent $1-\varepsilon$ for any $\varepsilon > 0$.

2.1 Branching particle systems

We split this section into two parts. In Section 2.1.1, we consider a finite branching-free particle system, and prove the existence and uniqueness of this system. In Section 2.1.2, we give a brief induction to the Mytnik-Sturm branching mechanism.

2.1.1 Finite branching-free particle systems

In this section, we will show the existence and uniqueness of the finite branching-free particle system that is determined by (1.1.1). The one-dimensional analogue is given by Lemma 1.3 of Wang [80].

Fix a time interval $[0,T]$. Let $W=\{W(t,x),(t,x)\in[0,T]\times\mathbb{R}^d\}$ be a d-dimensional space-time white Gaussian random field. For any positive integer n, let $\{B^i\}_{i\in\{1,\dots,n\}}$ be a family of independent d-dimensional Brownian motions that is independent of W. Consider an n-particle system, where the motion of each particle is described by the following stochastic differential equation in a random environment W:

$$dx_t^i = dB_t^i + \int_{\mathbb{R}^d} h(y-x_t^i)W(dt,dy), \tag{2.1.1}$$

with initial condition $x_0^i\in\mathbb{R}^d$ for all $i=1,\dots n$. In the case $n=1$, we omit all upper indexes in the equation (2.1.1) without confusion.

The following hypothesis for h will be used throughout this chapter:

Hypothesis (H0). $h=(h^{ij})_{1\le i,j\le d}\in H_3(\mathbb{R}^d;\mathbb{R}^d\otimes\mathbb{R}^d)$. *That is, the entries h^{ij} of h belongs to the Sobolev space $H_3(\mathbb{R}^d)$.*

For $k=0,1,2,3$, denote by $\|\cdot\|_{k,2}$ the Sobolev norm on $H_k(\mathbb{R}^d;\mathbb{R}^d\otimes\mathbb{R}^d)$, that is

$$\begin{aligned}\|h\|_{k,2}^2 := \sum_{i,j=1}^d \|h^{ij}\|_{k,2}^2 = \sum_{i,j=1}^d \Big(\int_{\mathbb{R}^d}|h^{ij}(x)|^2dx\Big)^{\frac{1}{2}} \\ + \sum_{i,j=1}^d\sum_{l=1}^k\sum_{i_1,\dots,i_l=1}^d\Big(\int_{\mathbb{R}^d}\Big|\frac{\partial^l}{\partial x_1\cdots\partial x_l}h^{ij}(x)\Big|^2dx\Big)^{\frac{1}{2}}.\end{aligned}$$

Let $\rho:\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^d$ be given by

$$\rho(x)=\int_{\mathbb{R}^d} h(z-x)h^*(z)dz, \tag{2.1.2}$$

where $h^*=(h^{ji})_{1\le i,j\le d}$ denotes the transpose of h. Then, for any $1\le i,j\le d$, and $x\in\mathbb{R}^d$, by Cauchy-Schwarz's inequality, we have

$$|\rho^{ij}(x)|\le\sum_{k=1}^d\|h^{ik}\|_2\|h^{kj}\|_2.$$

We denote by $\|\cdot\|_2$ the Hilbert Schmidt norm for matrices. Then, by Cauchy-Schwarz's inequality again, we have

$$\begin{aligned}\|\rho\|_\infty &:= \sup_{x\in\mathbb{R}^d}\|\rho(x)\|_2 = \sup_{x\in\mathbb{R}^d}\Big(\sum_{i,j=1}^d|\rho^{ij}(x)|^2\Big)^{\frac{1}{2}}\\ &\leq\Big(\sum_{i,j=1}^d\Big|\sum_{k=1}^d\|h^{ik}\|_2\|h^{kj}\|_2\Big|^2\Big)^{\frac{1}{2}}\\ &\leq\Big(\sum_{i,k=1}^d\|h^{ik}\|_2^2\sum_{j,k=1}^d\|h^{kj}\|_2^2\Big)^{\frac{1}{2}}\leq\|h\|_2^2\leq\|h\|_{3,2}^2.\end{aligned}$$

Similarly, we can show that the first, second, and third partial derivatives of ρ are bounded in $\mathbb{R}^d$. We make use of the following notations:

$$\|\rho\|_{k,\infty} := \sup_{x\in\mathbb{R}}\Big(\sum_{i,j=1}^d\sum_{i_1,\dots,i_k=1}^d\Big|\frac{\partial^k}{\partial x_{i_1}\cdots\partial x_{i_k}}\rho^{ij}(x)\Big|^2\Big)^{\frac{1}{2}},$$

for $k=1,2,3$. Now let us study the SDE's (2.1.1). These equations are not coupled and we solve them for each i separately. For this reason in the next theorem, which provides the existence and uniqueness of the equation (for each fixed i), we suppress the superscript index i.

Theorem 2.1.1. *Assume the Hypothesis (H0). Then, there exists a d-dimensional stochastic process $x=\{x_t, 0\leq t\leq T\}$ that is the unique strong solution to the SDE (2.1.1) (for each fixed i) with initial condition $x_0=x\in\mathbb{R}^d$.*

Proof. We prove this theorem by Picard iteration. Let

$$x_t^{(0)} = B_t+\int_0^t\int_{\mathbb{R}^d}h(y-x)W(ds,dy),$$

and let

$$x_t^{(m)} = B_t+\int_0^t\int_{\mathbb{R}^d}h(y-x_s^{(m-1)})W(ds,dy)$$

for all $m\geq 1$. Denote by $d_t^{(m)}=x_t^{(m)}-x_t^{(m-1)}$ for all $t\in[0,T]$. Then $d_t^{(m)}$ satisfies the following equation

$$d_t^{(m)} = \int_0^t \int_{\mathbb{R}^d} \left[h(y - x_s^{(m)}) - h(y - x_s^{(m-1)}) \right] W(ds, dy). \tag{2.1.3}$$

An application of the Itô isometry yields that

$$\begin{aligned} \|d_t^{(m)}\|_2^2 =& \left\| \int_0^t \int_{\mathbb{R}^d} \left[h(y - x_s^{(m)}) - h(y - x_s^{(m-1)}) \right] W(ds, dy) \right\|_2^2 \\ =& \mathbb{E} \int_0^t \int_{\mathbb{R}^d} \|h(y - x_s^{(m)}) - h(y - x_s^{(m-1)})\|_2^2 dyds \\ =& 2\|h\|_2^2 t - 2 \sum_{i,j=1}^d \int_0^t \rho^{ij}(d_s^{(m-1)}) ds \\ =& 2 \sum_{i,j=1}^d \int_0^t \left[\rho^{ij}(d_s^{(m-1)}) - \rho^{ij}(0) \right] ds, \end{aligned} \tag{2.1.4}$$

since $\sum_{i,j=1}^d \rho^{ij}(0) = \|h\|_2^2$. Noticing that ρ^{ij} has bounded first partial derivatives, we have

$$\|d_t^{(m)}\|_2^2 \le C \int_0^t \|d_s^{(m-1)}\|_2^2 ds,$$

for some constant C independent of m. On the other hand, we can show that

$$\begin{aligned} \|x_t^{(0)} - x\|_2^2 =& \left\| B_t + \int_0^t \int_{\mathbb{R}^d} h(y - x) W(ds, dy) - x \right\|_2^2 \\ \le& t + t\|h\|_2^2 + \|x\|^2. \end{aligned}$$

By iteration, we can conclude that

$$\|d_t^{(m)}\| \le \frac{1}{(m+1)!}(1 + \|h\|_2^2) t^{m+1} + \frac{1}{m!} \|x\|^2 t^m, \tag{2.1.5}$$

which is summable in m. In other words, for any $t \in [0, T]$, the sequence $x_t^{(m)}$ is convergent in $L^2(\Omega)$. Denote by x_t the limit of this sequence.

We claim that $x = \{x_t, 0 \le t \le T\}$ is a strong solution to (2.1.1) (recall we suppress the superscript). It suffices to show that as $m \to \infty$,

$$\int_0^t \int_{\mathbb{R}^d} h(y - x_s^{(m)}) W(ds, dy) \to \int_0^t \int_{\mathbb{R}^d} h(y - x_s) W(ds, dy)$$

in $L^2(\Omega)$ for all $t \in [0, T]$. We can easily check this convergence by arguments similar to those in (2.1.3) - (2.1.5).

Suppose that there are two solutions x and $\widetilde{x}$ to the SDE (2.1.1). Let $d = x - \widetilde{x}$. Again, by a similar argument as in (2.1.3) - (2.1.5), we have the following inequality

$$\|d_t\|_2^2 \leq C \int_0^t \|d_s\|_2^2 ds.$$

Notice that

$$\|d_t\|_2^2 \leq 2\|x_t\|_2^2 + 2\|\widetilde{x}_t\|_2^2 \leq 4(t + t\|h\|_2^2) < \infty.$$

An application of Gronwall's inequality yields $\|d_t\|_2^2 \equiv 0$. □

While equations (2.1.1) can be solved separately for each fixed i the solutions $x^1, \ldots, x^n$ are not independent since they all of them depend on the common random environment W. It is easy to see that $(x^1, \ldots, x^n)$ is an nd-dimensional Feller process governed by the generator

$$A^{(n)} f(y_1, \ldots, y_n) = \frac{1}{2}(\Delta^{(n)} + B^{(n)}) f(y_1, \ldots, y_n), \tag{2.1.6}$$

where $\Delta^{(n)}$ is the Laplace operator in $\mathbb{R}^{nd}$,

$$B^{(n)} f(y_1, \ldots, y_n) = \sum_{k_1, k_2 = 1}^{n} \sum_{i,j=1}^{d} \rho^{ij}(y_{k_1} - y_{k_2}) \frac{\partial^2 f}{\partial y_{k_1}^i \partial y_{k_2}^j (y^1, \ldots, y^n)}, \tag{2.1.7}$$

and $y_k = (y_k^1, \ldots, y_k^d) \in \mathbb{R}^d$ for all $k = 1, \ldots, n$. This is similar to (1.19) of Wang [80] for the one-dimensional case.

2.1.2 The Mytnik-Sturm branching mechanism

In this section, we briefly construct the branching particle system. For further study of this branching mechanism, we refer the readers to papers of Mytnik and Sturm [70, 78]).

We start this section by introducing some notation. For any integer $k \geq 0$, we denote by $C_b^k(\mathbb{R}^d)$ the space of k times continuously differentiable functions on $\mathbb{R}^d$ which are bounded together with their derivatives up to the order k. Also, $H_k(\mathbb{R}^d)$ is the Sobolev space of square integrable functions on $\mathbb{R}^d$ which have square integrable derivatives up to the order k. For any differentiable function ϕ on $\mathbb{R}^d$, we make use of the notation $\partial_{i_1\cdots i_m}\phi(x) = \frac{\partial^m}{\partial x_{i_1}\cdots\partial x_{i_m}}\phi(x)$.

We write $M_F(\mathbb{R}^d)$ for the space of finite measures on $\mathbb{R}^d$. We denote by $D([0,T], M_F(\mathbb{R}^d))$ the Skorohod space of càdlàg functions on time interval $[0,T]$, taking values in $M_F(\mathbb{R}^d)$, and equipped with the weak topology. For any $\phi \in C_b(\mathbb{R}^d)$ and $\mu \in M_F(\mathbb{R}^d)$, we write

$$\langle \mu, \phi \rangle = \mu(\phi) := \int_{\mathbb{R}^d} \phi(x)\mu(dx). \tag{2.1.8}$$

Let $\mathscr{I} := \{\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_N), \alpha_0 \in \{1,2,3\ldots\}, \alpha_i \in \{1,2\}, \text{ for } 1 \leq i \leq N\}$ be a set of multi-indexes. In our model $\mathscr{I}$ is the index set of all possible particles. In other words, initially there are a finite number of particles and each particle generates at most 2 offspring. For any particle $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_N) \in \mathscr{I}$, let $\alpha - 1 = (\alpha_0 \ldots, \alpha_{N-1}), \alpha - 2 = (\alpha_0, \ldots, \alpha_{N-2}), \ldots, \alpha - N = (\alpha_0)$ be the ancestors of α. Then, $|\alpha| = N$ is the number of the ancestors of the particle α. It is easy to see that the ancestors of any particle α are uniquely determined.

Fix a time interval $[0,T]$. Let $(\Omega, \mathscr{F}, P)$ be a complete probability space, on which $\{B_t^\alpha, t \in [0,T]\}_{\alpha \in \mathscr{I}}$ are independent d-dimensional standard Brownian motions, and W is a d-dimensional space-time white Gaussian random field on $[0,T] \times \mathbb{R}^d$ independent of the family $\{B^\alpha\}$.

Let $x_t = x(x_0, B^\alpha, r, t)$, where $0 \leq r \leq t \leq T$, be the unique solution to the following SDE:

$$x_t = x_0 + B_t^\alpha - B_r^\alpha + \int_r^t \int_{\mathbb{R}^d} h(y - x_s) W(ds, dy), \tag{2.1.9}$$

where $x_0 \in \mathbb{R}^d$, $r \in [0,t)$ and h is a $d \times d$ matrix-valued function. We assume that h satisfies Hypothesis (H0).

For any $t \in [0,T]$, let $t_n = \frac{\lfloor nt \rfloor}{n}$ be the last branching time before t. For any $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_N)$, if $nt_n = \lfloor nt \rfloor \le N$, let $\alpha_t = (\alpha_0, \dots, \alpha_{\lfloor nt \rfloor})$ be the ancestor of α at time t. Suppose that each particle, which starts from the death place of its parent, moves in $\mathbb{R}^d$ following the motion described by the SDE (2.1.9) during its lifetime. Then, the path of any particle α and all its ancestors, denoted by $x_t^{\alpha,n}$, is given by

$$x_t^{\alpha,n} = x_t^{\alpha_t,n} = \begin{cases} x\left(x_{\alpha_0}^n, B^{(\alpha_0)}, 0, t\right), & 0 \le t < \frac{1}{n}, \\ x\big(x_{t_n^-}^{\alpha_t - 1,n}, B^{\alpha_t}, t_n, t\big), & \frac{1}{n} \le t < \frac{N+1}{n}, \\ \partial, & \text{otherwise.} \end{cases}$$

Here $x_{\alpha_0}^n \in \mathbb{R}^d$ is the initial position of particle (α_0), $x_{t_n^-}^{\alpha_t - 1,n} := \lim_{s \uparrow t_n} x_s^{\alpha_t - 1,n}$, and ∂ denotes the "cemetery"-state, that can be understood as a point at infinity.

Let $\xi = \{\xi(x), x \in \mathbb{R}^d\}$ be a real-valued random field on $\mathbb{R}^d$ with covariance

$$\mathbb{E}\big(\xi(x)\xi(y)\big) = \kappa(x,y), \tag{2.1.10}$$

for all $x, y \in \mathbb{R}^d$. Assume that ξ satisfies the following conditions:

Hypothesis (H1). *(i)* ξ *is a symmetric random field, that is* $\mathbb{P}(\xi(x) > z) = \mathbb{P}(\xi(x) < -z)$ *for all* $x \in \mathbb{R}^d$ *and* $z \in \mathbb{R}$.

(ii) $\sup_{x \in \mathbb{R}^d} \mathbb{E}\big(|\xi(x)|^p\big) < \infty$ *for some* $p > 2$.

(iii) κ *is continuous and bounded on* $\mathbb{R}^d \times \mathbb{R}^d$.

For any $n \ge 1$, the random field ξ is used to define the offspring distribution after a scaling $\frac{1}{\sqrt{n}}$. In order to make the offspring distribution a probability measure, we introduce the truncation of the random field ξ, denoted by ξ^n, as follows:

$$\xi^n(x) = \begin{cases} \sqrt{n}, & \text{if } \xi(x) > \sqrt{n}, \\ -\sqrt{n}, & \text{if } \xi(x) < -\sqrt{n}, \\ \xi(x), & \text{otherwise.} \end{cases} \tag{2.1.11}$$

The correlation function of the truncated random field is then given by

$$\kappa_n(x,y) = \mathbb{E}\big(\xi^n(x)\xi^n(y)\big).$$

Let $(\xi_i^n)_{i\geq 0}$ be independent copies of ξ^n. Denote by ξ_i^{n+} and ξ_i^{n-} the positive and negative part of ξ_i^n respectively. Let $N^{\alpha,n} \in \{0,1,2\}$ be the offspring number of the particle α at the branching time $\frac{|\alpha|+1}{n}$. Assume that $\{N^{\alpha,n}, |\alpha| = i\}$ are conditionally independent given ξ_i^n and the position of α at its branching time, with a distribution given by

$$\begin{aligned} P\Big(N^{\alpha,n} = 2 |\, \xi_i^n, x^{\alpha,n}_{\frac{i+1}{n}-}\Big) &= \frac{1}{\sqrt{n}} \xi_i^{n+}\Big(x^{\alpha,n}_{\frac{i+1}{n}-}\Big), \\ P\Big(N^{\alpha,n} = 0 |\, \xi_i^n, x^{\alpha,n}_{\frac{i+1}{n}-}\Big) &= \frac{1}{\sqrt{n}} \xi_i^{n-}\Big(x^{\alpha,n}_{\frac{i+1}{n}-}\Big), \\ P\Big(N^{\alpha,n} = 1 |\, \xi_i^n, x^{\alpha,n}_{\frac{i+1}{n}-}\Big) &= 1 - \frac{1}{\sqrt{n}} |\xi_i^n|\Big(x^{\alpha,n}_{\frac{i+1}{n}-}\Big). \end{aligned}$$

For any particle $\alpha = (\alpha_0, \ldots, \alpha_N)$, α is called to be alive at time t, denoted by $\alpha \sim_n t$, if the following conditions are satisfied:

(i) There are exactly N branching before or at t: $\lfloor nt \rfloor = N$.

(ii) α has an unbroken ancestors line: $\alpha_{N-i+1} \leq N^{\alpha - i, n}$, for all $i = 1, 2, \ldots, N$.

[Introduction of $N^{\alpha,n}$ allows the particle α produce one more generation, namely, produce new particle $(\alpha, N^{\alpha,n})$. However, $(\alpha, 0)$ is considered no longer alive and will not produce offspring any more.] For any n, denote by $X^n = \{X_t^n, t \in [0,T]\}$ the empirical measure-valued process of the particle system. Then, X^n is a discrete measure-valued process, given by

$$X_t^n = \frac{1}{n} \sum_{\alpha \sim_n t} \delta_{x_t^{\alpha,n}}, \tag{2.1.12}$$

where δ_x is the Dirac measure at $x \in \mathbb{R}^d$, and the sum is among all alive particles at time $t \in [0,T]$. Then, for any $\phi \in C_b^2(\mathbb{R}^d)$, with the notation (2.1.8), we have

$$X_t^n(\phi) = \int_{\mathbb{R}^d} \phi(x) X_t^n(dx) = \frac{1}{n} \sum_{\alpha \sim n} \phi(x_t^{\alpha,n}).$$

2.2 Main results

Let $(\Omega, \mathscr{F}, \{F_t\}_{t\in[0,T]}, P)$ be a complete filtered probability space that satisfies the usual conditions. Suppose that W is a d-dimensional space-time white Gaussian random field on $[0,T] \times \mathbb{R}^d$, and V is a one-dimensional Gaussian random field on $[0,T] \times \mathbb{R}^d$ independent of W, that is time white and spatially colored with correlation κ defined in (2.1.10):

$$\mathbb{E}\big(V(t,x)V(s,y)\big) = (t \wedge s)\kappa(x,y),$$

for all $s,t \in [0,T]$ and $x,y \in \mathbb{R}^d$. Assume that $\{W(t,x), x \in \mathbb{R}^d\}$, $\{V(t,x), x \in \mathbb{R}^d\}$ are $\mathscr{F}_t$-measurable for all $t \in [0,T]$, and $\{W(t,x) - W(s,x), x \in \mathbb{R}^d\}$, $\{V(t,x) - V(s,x), x \in \mathbb{R}^d\}$ are independent of $\mathscr{F}_s$ for all $0 \leq s < t \leq T$.

Denote by A^* the adjoint of A, where $A = A^{(1)}$ is the generator defined in (2.1.6). Consider the following SPDE:

$$\begin{aligned} u_t(x) =& \mu(x) + \int_0^t A^* u_s(x) ds - \sum_{i,j=1}^d \int_0^t \int_{\mathbb{R}^d} \frac{\partial}{\partial x_i} \left[h^{ij}(y-x) u_s(x) \right] W^j(ds,dy) \\ &+ \int_0^t u_s(x) \frac{V(ds,dx)}{dx}. \end{aligned} \tag{2.2.1}$$

Definition 2.2.1. *Let $u = \{u_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ be a random field. Then,* *(i)* *u is said to be a strong solution to the SPDE (2.2.1), if u is jointly measurable on $[0,T]\times\mathbb{R}^d\times\Omega$, adapted to $\{\mathscr{F}_t\}_{t\in[0,T]}$ and for any $\phi\in C_b^2(\mathbb{R}^d)$, the following equation holds for every $t\in[0,T]$:*

$$\begin{aligned}\int_{\mathbb{R}^d}\phi(x)u_t(x)dx=&\int_{\mathbb{R}^d}\phi(x)\mu(x)dx+\int_0^t\int_{\mathbb{R}^d}A\phi(x)u_s(x)dxds\\&+\int_0^t\int_{\mathbb{R}^d}\Big[\int_{\mathbb{R}^d}\nabla\phi(x)^*h(y-x)u_s(x)dx\Big]W(ds,dy)\\&+\int_0^t\int_{\mathbb{R}^d}\phi(x)u_s(x)V(ds,dx),\ a.s.\end{aligned}\tag{2.2.2}$$

where the last two stochastic integrals are Walsh's integral (c.f. Walsh [79]).

The solution to (2.2.1) is said to be pathwise unique, if whenever u and $\widetilde{u}$ are two solutions to (2.2.1), then there exists a set $G\in\mathscr{F}$ with probability one, such that $u_t(\omega)=\widetilde{u}_t(\omega)$ for all $t\in[0,T]$ and $\omega\in G$.

(ii) *u is said to be a weak solution to the SPDE (2.2.1), if there exists a filtered probability space, on which W and V are independent random fields that satisfy the above properties, such that u is a strong solution with this probability space.*

Let $X^n=\{X_t^n,0\le t\le T\}$ be defined by (2.1.12). In order to show the convergence of X^n in $D([0,T];M_F(\mathbb{R}^d))$, we make use of the following hypothesis on the initial measures X_0^n:

Hypothesis (H2). *(i)* $\sup_{n\geq 1}|X_0^{(n)}(1)|<\infty$.

(ii) *$X_0^n\Rightarrow X_0$ in $M_F(\mathbb{R}^d)$ as $n\to\infty$.*

(iii) *X_0 has a bounded density μ.*

In Section 2.3 we prove the following two theorems.

Theorem 2.2.2. *Let X^n be defined in (2.1.12). Then, under hypotheses(H1) and (H2), we have the following results:*

(i) *$X^n\Rightarrow X$ in $D([0,T],M_F(\mathbb{R}^d))$ as $n\to\infty$.*

(ii) The limit $X = \{X_t, t \in [0,T]\}$ is a continuous $M_F(\mathbb{R}^d)$-valued process. In addition, for almost all $\omega \in \Omega$ and every $t \in [0,T]$, as a finite measure on $\mathbb{R}^d$, $X_t(\omega)$ has a density $u_t(x,\omega)$.

(iii) $u = \{u_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ is a weak solution to the SPDE (2.2.1) in the sense of Definition 2.2.1.

Theorem 2.2.3. *Assume the Hypotheses (H1) and (H2) (iii). The SPDE (2.2.1) has a jointly continuous strong solution, which is pathwise unique in the space of jointly continuous solutions in the sense of Definition 2.2.1.*

The last main result in this chapter is the following theorem concerning the Hölder continuity of the solution to the SPDE (2.2.1).

Theorem 2.2.4. *Let $u = \{u_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ be the strong solution to the SPDE (2.2.1) in the sense of Definition 2.2.1. Then, for any $\beta_1, \beta_2 \in (0,1)$ and $p > 1$, there exists a constant C that depends on T, d, h, p, β_1, and β_2, such that for all $x, y \in \mathbb{R}^d$ and $0 < s < t \le T$*

$$\|u_t(x) - u_s(y)\|_{2p} \le C s^{-\frac{1}{2}} \left(|x-y|^{\beta_1} + |t-s|^{\frac{1}{2}\beta_2}\right).$$

Hence by Kolmogorov's criteria, $u_t(x)$ is almost surely jointly Hölder continuous on $(0,T] \times \mathbb{R}^d$, with exponent $\beta_1 \in (0,1)$ in space and $\beta_2 \in (0, \frac{1}{2})$ in time.

2.3 Proof of Theorems 2.2.2 and 2.2.3

We prove Theorems 2.2.2 and 2.2.3 in the following steps:

(i) In Section 2.3.1, we show that $\{X^n\}_{n\ge 1}$ is a tight sequence in $D([0,T]; M_F(\mathbb{R}^d))$, and the limit of any convergent subsequence in law solves a martingale problem.

(ii) In Section 2.3.2, we show that any solution to the martingale problem has a density almost surely.

(iii) In Section 2.3.3, we show the equivalence between martingale problem (c.f. (2.3.4) - (2.3.5) below) and the SPDE (2.2.1). Finally, we prove Theorems 2.2.2 and 2.2.3.

2.3.1 Tightness and martingale problem

Recall the empirical measure-valued process $X^n = \{X^n_t, t \in [0,T]\}$ given by (2.1.12). Let $A = A^{(1)}$ be the generator of one particle motion defined in (2.1.6). For any $\phi \in C^2_b(\mathbb{R}^d)$, similar to the equality (49) of Sturm [78], we can decompose X^n_t as follows:

$$X^n_t(\phi) = X^n_0(\phi) + Z^n_t(\phi) + M^{b,n}_t(\phi) + B^n_t(\phi) + U^n_t(\phi), \tag{2.3.1}$$

where

$$Z^n_t(\phi) = \int_0^t X^n_u(A\phi)du,$$

$$M^{b,n}_t(\phi) = M^{b,n}_{t_n}(\phi) = \frac{1}{n}\sum_{s_n<t_n}\sum_{\alpha\sim_n s_n}\phi\left(x^{\alpha,n}_{s_n+\frac{1}{n}}\right)(N^{\alpha,n}-1),$$

$$B^n_t(\phi) = \frac{1}{n}\Big(\sum_{s_n<t_n}\sum_{\alpha\sim_n s_n}\int_{s_n}^{s_n+\frac{1}{n}}\nabla\phi(x^{\alpha,n}_u)^* dB^\alpha_u + \sum_{\alpha\sim_n t}\int_{t_n}^t \nabla\phi(x^{\alpha,n}_u)^* dB^\alpha_u\Big),$$

and

$$\begin{aligned} U^n_t(\phi) =& \frac{1}{n}\Big(\sum_{s_n<t_n}\sum_{\alpha\sim_n s_n}\int_{s_n}^{s_n+\frac{1}{n}}\int_{\mathbb{R}^d}\nabla\phi(x^{\alpha,n}_u)^* h(y-x^{\alpha,n}_u)W(du,dy) \\ &+ \sum_{\alpha\sim_n t}\int_{t_n}^t\int_{\mathbb{R}^d}\nabla\phi(x^{\alpha,n}_u)^* h(y-x^{\alpha,n}_u)W(du,dy)\Big). \end{aligned}$$

Notice that

$$
\begin{aligned}
&\mathbb{E}\sum_{\alpha\sim_n s_n}\int_{s_n}^{s_n+\frac{1}{n}}\int_{\mathbb{R}^d}\left|\nabla\phi(x_u^{\alpha,n})^*h(y-x_u^{\alpha,n})\right|^2 dydu\\
\leq&\sum_{|\alpha|=\lfloor sn\rfloor}\mathbb{E}\int_{s_n}^{s_n+\frac{1}{n}}\int_{\mathbb{R}^d}\left|\nabla\phi(x_u^{\alpha,n})^*h(y-x_u^{\alpha,n})\right|^d ydu\\
\leq&2^{nT}n^{-1}N_0\|\phi\|_{1,\infty}\|h\|_2<\infty,
\end{aligned}
$$

where N_0 denotes the number of initial particles, that is a finite integer. Therefore, by the stochastic Fubini theorem (c.f. Lemma 4.1 on page 116 of Ikeda & Watanabe [49]), we can write:

$$
U_t^n(\phi)=\int_0^t\int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d}\nabla\phi(x)^*h(y-x)X_u(dx)\Big)W(du,dy).
$$

As in Sturm [78], consider the natural filtration, generated by the process X^n

$$
\mathscr{F}_t^n=\sigma\left(\{x^{\alpha,n},N^{\alpha,n}\big||\alpha|<\lfloor nt\rfloor\}\cup\{x_s^{\alpha,n},s\leq t,|\alpha|=\lfloor nt\rfloor\}\right),
$$

and a discrete filtration at branching times

$$
\widetilde{\mathscr{F}}_{t_n}^n=\sigma\big(\mathscr{F}_{t_n}^n\cup\{x^{\alpha,n}\big||\alpha|=nt_n\}\big)=\mathscr{F}_{(t_n+n^{-1})^-}^n.
$$

Then, $B_t^n(\phi)$ and $U_t^n(\phi)$ are continuous $\mathscr{F}_t^n$-martingales, while $M_t^{b,n}(\phi)$ is a discrete $\widetilde{\mathscr{F}}_{t_n}^n$-martingale.

Lemma 2.3.1. *Assume Hypotheses (H0), (H1), (H2) (i) and (ii). Let $p>2$ be given in Hypothesis (H1). Then, for all $\phi\in C_b^2(\mathbb{R}^d)$,*

(i) $\mathbb{E}\Big(\sup_{0\leq t\leq T}|X_t^n(\phi)|^p\Big)$, $\mathbb{E}\Big(\sup_{0\leq t\leq T}|M_t^{b,n}(\phi)|^p\Big)$ *and* $\mathbb{E}\Big(\sup_{0\leq t\leq T}|U_t^n(\phi)|^p\Big)$ *are bounded uniformly in $n\geq 1$.*

(ii) $\mathbb{E}\Big(\sup_{0\leq t\leq T}|B_t^n(\phi)|^p\Big)\to 0$, *as $n\to\infty$.*

Proof. (i) By the same argument as that for Lemma 3.1 of Sturm [78], we can show that

$$\mathbb{E}\Big(\sup_{0\le t\le T}|M_t^{b,n}(1)|^p\Big)\le C\int_0^T\mathbb{E}\Big(\sup_{0\le s\le t}|X_s^n(1)|^p\Big)dt$$

where the constant $C>0$ does not depend on n. Again similarly as Strum did for (58) of Sturm [78], we can also deduce the following inequality

$$\begin{aligned}\mathbb{E}\Big(\sup_{0\le t\le T}|X_t^n(1)|^p\Big)\le& C\Big(1+\mathbb{E}\Big(\sup_{0\le t\le T}|M_{t_n}^{b,n}(1)|^p\Big)\Big)\\ \le& C_1+C_2\int_0^T\mathbb{E}\Big(\sup_{0\le s\le t}|X_s^n(1)|^p\Big)dt,\end{aligned}$$

where C_1,C_2 are constants independent of n. Notice that

$$\sup_{0\le t\le T}|X_t^n(1)|\le 2^{nT}\frac{N_0^n}{n},$$

that is bounded for fixed n. Then, it follows from Gronwall's inequality that the sequence

$$\Big\{\mathbb{E}\Big(\sup_{0\le t\le T}|X_t^n(1)|^p\Big)\Big\}_{n\ge1}$$

is uniformly bounded in n.

The uniform boundedness of collections

$$\mathbb{E}\Big(\sup_{0\le t\le T}|X_t^n(\phi)|^p\Big)\quad\text{and}\quad\mathbb{E}\Big(\sup_{0\le t\le T}|M_t^{b,n}(\phi)|^p\Big)$$

follows immediately.

We estimate $U_t^n(\phi)$ as follows:

$$\begin{aligned}\langle U^n(\phi)\rangle_t =&\Big\langle \int_0^\cdot \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} \nabla\phi(x)^* h(y-x) X_u^n(dx)\Big) W(du,dy)\Big\rangle_t \\ =&\sum_{j=1}^d \int_0^t \int_{\mathbb{R}^d} \Big(\sum_{i=1}^d \int_{\mathbb{R}^d} \partial_i\phi(x) h^{ij}(y-x) X_u^n(dx)\Big)^2 dydu \\ =&\int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} \nabla\phi(x)^*\rho(x-z)\nabla\phi(z) X_u^n(dx)X_u^n(dz)du \\ \le&\|\rho\|_\infty \|\phi\|_{1,\infty}^2 \int_0^T |X_u^n(1)|^2 du. \end{aligned} \tag{2.3.2}$$

Thus by (2.3.2), Burkholder-Davis-Gundy's and Jensen's inequalities, we have

$$\begin{aligned}\mathbb{E}\Big(\sup_{0\le t\le T} |U_t^n(\phi)|^p\Big) \le& c_p \mathbb{E}\big(\langle U^n(\phi)\rangle_T^{\frac{p}{2}}\big) \le c_p \|\rho\|_\infty^{\frac{p}{2}} \|\phi\|_{1,\infty}^p T^{\frac{p}{2}-1} \mathbb{E}\Big(\int_0^T |X_u^n(1)|^p du\Big) \\ \le& c_p \|\rho\|_\infty^{\frac{p}{2}} \|\phi\|_{1,\infty}^p T^{\frac{p}{2}} \mathbb{E}\Big(\sup_{0\le t\le T} |X_t^n(1)|^p\Big), \end{aligned} \tag{2.3.3}$$

that is also uniformly bounded in n.

(ii) Note that $\{B^\alpha\}$ are independent Brownian motions. Then, by Burkholder-Davis-Gundy's inequality, we have

$$\begin{aligned}\mathbb{E}\Big(\sup_{0\le t\le T} |B_t^n(\phi)|^2\Big) \le& \frac{c_2}{n^2}\Big[\sum_{s_n<T_n}\sum_{\alpha\sim_n s_n} \mathbb{E}\Big(\int_{s_n}^{s_n+\frac{1}{n}} |\nabla\phi(x_u^{\alpha,n})|^2 du\Big) \\ &+ \sum_{\alpha\sim_n t} \mathbb{E}\Big(\int_{T_n}^T |\nabla\phi(x_u^{\alpha,n})|^2 du\Big)\Big] \\ =& \frac{c_2}{n}\mathbb{E}\Big(\int_0^T \int_{\mathbb{R}^d} |\nabla\phi(x)|^2 X_u(dx)du\Big) \le \frac{c_2}{n}\|\phi\|_{1,\infty}^2 T \mathbb{E}\Big(\sup_{0\le t\le T} |X_t^n(1)|^p\Big) \to 0, \end{aligned}$$

because $\mathbb{E}\Big(\sup_{0\le t\le T} |X_t^n(1)|^p\Big)$ is uniformly bounded in n. □

As a consequence of Lemma 2.3.1, the collection

$$\Big\{\sup_{0\le t\le T} |X_t^n(\phi)|^2,\ \sup_{0\le t\le T} |M_t^{b,n}(\phi)|^2,\ \sup_{0\le t\le T} |U_t^n(\phi)|^2\Big\}_{n\ge 1}$$

is uniformly integrable.

Definition 2.3.2. *Let $\{X^{\alpha}\}$ be a collection of real-valued stochastic processes. A family of stochastic processes $\{X^{\alpha}\}$ is said to be C-tight, if it is tight, and the limit of any subsequence is continuous.*

Lemma 2.3.3. *Assume Hypotheses (H0),(H1),(H2) (i) and (ii). For all $\phi \in C_b^2(\mathbb{R}^d)$, $M^{b,n}(\phi)$, $Z^n(\phi)$, and $X^n(\phi)$ and $U^n(\phi)$ are C-tight sequences in $D([0,T],\mathbb{R})$.*

Proof. By an argument to that used by Sturm in the proof of Lemma 3.6 in Sturm [78], we can deduce the C-tightness of $M^{b,n}(\phi)$ and $Z^n(\phi)$.

We prove the tightness of $X_t^n(\phi)$ by checking Aldous's conditions (see Theorem 4.5.4 of Dawson [21]). By Chebyshev's inequality, for any fixed $t \in [0,T]$, and $N > 0$, we have

$$\mathbb{P}\left(|X_t^n(\phi)| > N\right) \le \frac{1}{N^p}\mathbb{E}\left(|X_t^n(\phi)|^p\right) \le \frac{1}{N^p}\mathbb{E}\left(\sup_{0\le t\le T}|X_t^n(\phi)|^p\right) \to 0,$$

uniformly in n as $N \to \infty$ by Lemma 2.3.1 (i).

On the other hand, for any $n \ge 1$, we extend X^n to the time interval $[0, T_n + \frac{1}{n}]$ in such a way that X^n performs the same diffusion/branching mechanism as before on $[T, T_n + \frac{1}{n}]$. Denote by $\widetilde{X}^n = \{\widetilde{X}^n(t), t \in [0, T_n + \frac{1}{n}]\}$ the extended process. Then, by Theorem 10.13 of Dynkin [27], we know that $\widetilde{X}^n$ is a strong Markov process on $[0, T_n + \frac{1}{n}]$.

Let $\{\tau_n\}_{n\ge1}$ be any collection of stopping times bounded by T and let $\{\delta_n\}_{n\ge1}$ be any positive sequence that decreases to 0, such that $\tau_n + \delta_n \le T$. Then, due to the uniform boundedness of $\mathbb{E}\left(\sup_{0\le t\le T}|X_t^n(\phi)|^p\right)$ and the strong Markov property of $\widetilde{X}^n$, we have

$$\begin{aligned}\mathbb{P}\left(\left|X^n_{\tau_n+\delta_n}(\phi) - X^n_{\tau_n}(\phi)\right| > \varepsilon\right) &= \mathbb{P}\left(\left|\widetilde{X}^n_{\tau_n+\delta_n}(\phi) - \widetilde{X}^n_{\tau_n}(\phi)\right| > \varepsilon\right)\\ &= \mathbb{P}\left(\left|\widetilde{X}^n_{\delta_n}(\phi) - \widetilde{X}^n_0(\phi)\right| > \varepsilon\right) \le \frac{1}{\varepsilon^p}\mathbb{E}\left(\left|\widetilde{X}^n_{\delta_n}(\phi) - \widetilde{X}^n_0(\phi)\right|^p\right)\\ &\le \frac{\delta_n^{\frac{p}{2}}}{\varepsilon^p} c_p \|\rho\|_\infty^{\frac{p}{2}} \|\phi\|_\infty^p \left[\mathbb{E}\left(\sup_{0\le t\le T}|X_t^n(1)|^p\right) + \mathbb{E}\left(|X_0^n(1)|^p\right)\right] \to 0,\end{aligned}$$

as $n \to 0$. Thus both of Aldous's conditions are satisfied, and then it follows that $X_t^n(\phi)$ is tight in

$D([0,T],\mathbb{R})$.

Recall the decomposition formula (2.3.1):

$$X_t^n(\phi) = X_0^n(\phi) + Z_t^n(\phi) + M_t^{b,n}(\phi) + B_t^n(\phi) + U_t^n(\phi).$$

Notice that $X^n(\phi)$, $Z^n(\phi)$, $M^{b,n}(\phi)$ are tight sequences as proved just above, $X_0^n(\phi)$ converges weakly by assumption, and $B_t^n(\phi)$ converges 0 in $L^2(\Omega)$ uniformly for all $t \in [0,T]$ by Lemma 2.3.1 (ii). As a consequence, $U^n(\phi)$ is tight in $D([0,T],\mathbb{R})$. In addition, by Proposition VI.3.26 of Jacod & Shiryaev [50], every limit of a tight sequence of continuous process $U^n(\phi)$ is continuous. It follows that $U^n(\phi)$ and $X^n(\phi)$ are C-tight sequences in $D([0,T];\mathbb{R})$. □

Let $\mathscr{S} = \mathscr{S}(\mathbb{R}^d)$ be the Schwartz space on $\mathbb{R}^d$, and let $\mathscr{S}'$ be the Schwartz dual space. Then, we have the following lemma:

Lemma 2.3.4. *Assume Hypotheses (H0), (H1) and (H2) (i), (ii). Then,*

(i) $\{X^n\}_{n\geq 1}$ *is a C-tight sequence in* $D([0,T];M_F(\mathbb{R}^d))$.

(ii) $\{B^n\}_{n\geq 1}$, $\{M^{b,n}\}_{n\geq 1}$, *and* $\{U^n\}_{n\geq 1}$ *are C-tight in* $D([0,T];\mathscr{S}')$.

Proof. Let $\widehat{\mathbb{R}}^d = \mathbb{R}^d \cup \{\partial\}$ be the one point compactification of $\mathbb{R}^d$. Then, by Theorem 4.6.1 of Dawson [21] and Lemma 2.3.3, $\{X^n\}_{n\geq 1}$ is a tight sequence in $D([0,T];M_F(\widehat{\mathbb{R}}^d))$.

On the other hand, by the same argument as in Lemma 3.9 of Sturm [78], we can show that any limit of a weakly convergent subsequence X^{n_k} in $D([0,T];M_F(\widehat{\mathbb{R}}^d))$ belongs to $C([0,T];M_F(\mathbb{R}^d))$, the space of continuous $M_F(\mathbb{R}^d)$-valued functions on $[0,T]$. Therefore, $\{X^n\}_{n\geq 1}$ is a C-tight sequence in $D([0,T];M_F(\mathbb{R}^d))$.

To show the property (ii), notice that $\mathscr{S} \subset C_b^2(\mathbb{R}^d)$. Then, by Theorem 4.1 of Mitoma [69], $\{B^n\}_{n\geq 1}$, $\{M^{b,n}\}_{n\geq 1}$, and $\{U^n\}_{n\geq 1}$ are C-tight in $D([0,T];\mathscr{S}')$. □

Proposition 2.3.5. *Assume Hypotheses (H0), (H1),(H2) (i) and (ii). Let X be the limit of a weakly convergent subsequence* $\{X^{n_k}\}_{k\geq 1}$ *in* $D([0,T];M_F(\mathbb{R}^d))$. *Then, X is a solution to the following* *martingale problem: for any $\phi \in C_b^2(\mathbb{R}^d)$, the process $M(\phi) = \{M_t(\phi) : 0 \le t \le T\}$, given by*

$$M_t(\phi) := X_t(\phi) - X_0(\phi) - \int_0^t X_s(A\phi)ds, \tag{2.3.4}$$

is a continuous and square integrable $\mathscr{F}_t^X$-adapted martingale with quadratic variation:

$$\begin{aligned}\langle M(\phi)\rangle_t =& \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} \nabla\phi(x)^*\rho(x-y)\nabla\phi(y)X_s(dx)X_s(dy)ds \\ &+ \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} \kappa(x,y)\phi(x)\phi(y)X_s(dx)X_s(dy)ds. \end{aligned}\tag{2.3.5}$$

Proof. Let $\{X^{n_k}\}_{k\ge1}$ be a weakly convergent subsequence in $D([0,T];M_F(\mathbb{R}^d))$. By taking further subsequences, we can assume, in view of Lemma 2.3.4 (ii), that $\{B^{n_k}\}_{k\ge1}$, $\{M^{b,n_k}\}_{k\ge1}$, and $\{U^{n_k}\}_{k\ge1}$ are weakly convergent subsequences in $D([0,T];\mathscr{S}')$.

Then, by Skorohod's representation theorem, there exists a probability space $(\widetilde{\Omega},\widetilde{\mathscr{F}},\widetilde{\mathbb{P}})$, on which $(\widetilde{X}^{n_k},\widetilde{M}^{b,n_k},\widetilde{B}^{n_k},\widetilde{U}^{n_k})$ has the same joint distribution as $(X^{n_k},M^{b,n_k},B^{n_k},U^{n_k})$ for all $k\ge1$, and converge a.s. to $(\widetilde{X},\widetilde{M}^b,\widetilde{B},\widetilde{U})$ in the product space $D([0,T],M_F(\widehat{\mathbb{R}}^d))\times D([0,T],\mathscr{S}')^3$.

Then, for any $\phi\in\mathscr{S}'$, $(\widetilde{X}^{n_k}(\phi),\widetilde{M}^{b,n_k}(\phi),\widetilde{B}^{n_k}(\phi),\widetilde{U}^{n_k}(\phi))$ converges a.s. in $D([0,T],\mathbb{R})^4$. Since $\left\{\sup_{0\le t\le T}|X_t^n(\phi)|^2, \sup_{0\le t\le T}|M_t^{b,n}(\phi)|^2, \sup_{0\le t\le T}|U_t^n(\phi)|^2\right\}_{n\ge1}$ is uniformly integrable, the convergence is also in $L^2([0,T]\times\Omega)$.

For any $t\in[0,T]$, let

$$\widetilde{M}_t^{n_k}(\phi) := \widetilde{X}_t^{n_k}(\phi) - \widetilde{X}_0^{n_k}(\phi) - \int_0^t \widetilde{X}_s^{n_k}(A\phi)ds = \widetilde{M}_t^{b,n_k}(\phi) + \widetilde{B}_t^{n_k}(\phi) + \widetilde{U}_t^{n_k}(\phi).$$

Then, it converges to a continuous and square integrable martingale $\widetilde{M}(\phi) = \widetilde{M}^b(\phi) + \widetilde{U}(\phi)$ in $L^2(\widetilde{\Omega})$ with respect to its natural filtration.

The next step is to compute the quadratic variation of $\widetilde{M}(\phi)$. Notice that W and $\{B^\alpha\}$ are independent, then U^n and B^n are orthogonal. As a consequence, $\widetilde{U}^{n_k}$ and $\widetilde{B}^{n_k}$ are also orthogonal. On the other hand, $\widetilde{M}^{b,n}(\phi)$ is a pure jump martingale, while $\widetilde{U}^{n_k}(\phi)$ and $\widetilde{B}^{n_k}(\phi)$ are continuous martingales. Due to Theorem 43 on page 353 of Dellacherie & Meyer [26], $\widetilde{M}^{b,n}(\phi)$, $\widetilde{B}^{n_k}(\phi)$

and $\widetilde{U}^{n_k}(\phi)$ are mutually orthogonal. By the same argument as in Lemma 2.3.1, we can show that $\langle \widetilde{M}^{b,n_k}(\phi)+\widetilde{B}^{b,n_k}(\phi)+\widetilde{U}^{n_k}(\phi)\rangle_t = \langle \widetilde{M}^{b,n_k}(\phi)\rangle_t + \langle \widetilde{B}^{b,n_k}(\phi)\rangle_t + \langle \widetilde{U}^{n_k}(\phi)\rangle_t$ are uniformly integrable. Then, by Theorem II.4.5 of Perkins [74], we have

$$\begin{aligned}\langle \widetilde{M}^{b,n_k}(\phi)+\widetilde{B}^{b,n_k}(\phi)+\widetilde{U}^{n_k}(\phi)\rangle_t &= \langle \widetilde{M}^{b,n_k}(\phi)\rangle_t + \langle \widetilde{B}^{b,n_k}(\phi)\rangle_t + \langle \widetilde{U}^{n_k}(\phi)\rangle_t \\ &\to \langle \widetilde{M}^{b}(\phi)\rangle_t + \langle \widetilde{U}(\phi)\rangle_t = \langle \widetilde{M}(\phi)\rangle_t\end{aligned}$$

as $k\to\infty$ in $D([0,T],\mathbb{R})$ in probability.

On the other hand, by the same argument of Lemma 3.8 of Sturm [78], we have

$$\langle \widetilde{M}^{b}(\phi)\rangle_t = \lim_{k\to\infty}\langle \widetilde{M}^{b,n_k}(\phi)\rangle_t = \int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d}\kappa(x,y)\phi(x)\phi(y)\widetilde{X}_s(dx)\widetilde{X}_s(dy)ds,\ a.s.$$

For $\langle \widetilde{U}(\phi)\rangle_t$, by (2.3.2), since $\widetilde{X}^{n_k}(\phi)\to\widetilde{X}(\phi)$ in $L^2([0,T]\times\Omega)$, it follows that

$$\begin{aligned}\lim_{k\to\infty}\langle \widetilde{U}^{n_k}(\phi)\rangle_t &= \lim_{k\to\infty}\int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d}\nabla\phi(x)^*\rho(x-z)\nabla\phi(z)\widetilde{X}^n_u(dx)\widetilde{X}^n_u(dz)du \\ &= \int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d}\nabla\phi(x)^*\rho(x-z)\nabla\phi(z)\widetilde{X}_u(dx)\widetilde{X}_u(dz)du.\end{aligned}$$

As a consequence, $\widetilde{M}=\{\widetilde{M_t}, t\in[0,T]\}$, where

$$\widetilde{M}_t(\phi)=\widetilde{X}_t(\phi)-\widetilde{X}_0(\phi)-\int_0^t\widetilde{X}_s(A\phi)ds=\widetilde{M}^b_t(\phi)+\widetilde{B}_t(\phi)+\widetilde{U}_t(\phi),$$

is a continuously square integrate martingale with the quadratic variation given by the expression (2.3.5).

Finally, by the same argument as in Theorem II in Section 4.2 of Perkins [74], we can show $\widetilde{M}(\phi)$ is an $\mathscr{F}^{\widetilde{X}}$-adapted martingale. □

2.3.2 Absolute continuity

Assume Hypotheses (H0) and (H1). Let X_t be a solution to the martingale problem (2.3.4) - (2.3.5). In this section, we show that for almost every $t \in [0,T]$, as an $M_F(\mathbb{R}^d)$-valued random variable, X_t has a density almost surely.

For any $n \geq 1$, $f \in C_b^2(\mathbb{R}^{nd})$, and $\mu \in M_F(\mathbb{R}^d)$, we define

$$\mu^{\otimes n}(f) := \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} f(x_1,\ldots,x_n)\mu(dx_1)\cdots\mu(dx_n).$$

We derive the moment formula $\mathbb{E}(X_t^{\otimes n}(f))$ of the process X. In the one-dimensional Dawson-Watanabe branching case, Skoulakis & Adler [75] obtained the formula by computing the limit of particle approximations. An alternative approach by Xiong [83] consists in differentiating a conditional stochastic log-Laplace equation. In current chapter we use the techniques of moment duality to derive the moment formula. It can be also formulated by computing the limit of particle approximations.

For any integers $n \geq 2$ and $k \leq n$, we make use of the notation $x_k = (x_k^1,\ldots,x_k^d) \in \mathbb{R}^d$ and $x = (x_1,\ldots,x_n) \in \mathbb{R}^{nd}$. Let $\Phi_{ij}^{(n)} : C_b^2(\mathbb{R}^{nd}) \to C_b^2(\mathbb{R}^{nd})$, and $F^{(n)}, G^{(n)} : C_b^2(\mathbb{R}^{nd}) \times M_F(\mathbb{R}^d) \to \mathbb{R}$ be given by

$$(\Phi_{ij}^{(n)} f)(x_1,\ldots,x_n) := \kappa(x_i,x_j)f(x_1,\ldots,x_n), \quad i,j \in \{1,2,\ldots,n\},$$

$$F^{(n)}(f,\mu) := \mu^{\otimes n}(f),$$

and

$$G^{(n)}(f,\mu) := \mu^{\otimes n}(A^{(n)}f) + \frac{1}{2}\sum_{\substack{1\leq i,j\leq n \\ i\neq j}} \mu^{\otimes n}(\Phi_{ij}^{(n)} f),$$

where $\kappa \in C_b^2(\mathbb{R}^{2d})$ is the correlation of the random field ξ given by (2.1.10), and $A^{(n)}$ is the generator on n-particle motion defined in (2.1.6).

Lemma 2.3.6. *Let X_t be a solution to the martingale problem (2.3.4) - (2.3.5). Then, for any $n \geq 2$*

and $f \in C_b^2(\mathbb{R}^{nd})$, the following process

$$F^{(n)}(f, X_t) - \int_0^t G^{(n)}(f, X_s) ds$$

is a martingale.

Proof. See Lemma 1.3.2 of Xiong [84]. □

Let $\{T_t^{(n)}\}_{t \geq 0}$ be the semigroup generated by $A^{(n)}$, that is, $T_t^{(n)} : C_b^2(\mathbb{R}^{nd}) \to C_b^2(\mathbb{R}^{nd})$, given by

$$T_t^{(n)} f(x_1, \ldots, x_n) = \int_{\mathbb{R}^{nd}} p(t, (x_1, \ldots, x_n), (y_1, \ldots, y_n)) f(y_1, \ldots, y_n) dy_1 \ldots dy_n,$$

where p is the transition density of n-particle-motion.

Let $\{S_k^{(n)}\}_{k \geq 1}$ be i.i.d. uniformly distributed random variables taking values in the set $\{\Phi_{ij}, 1 \leq i, j \leq n, i \neq j\}$. Let $\{\tau_k\}_{k \geq 1}$ be i.i.d exponential random variables independent of $\{S_k^{(n)}\}_{k \geq 1}$, with rate $\lambda_n = \frac{1}{2} n(n-1)$. Let $\eta_0 \equiv 0$, and $\eta_j = \sum_{i=1}^j \tau_i$ for all $j \geq 1$. For any $f \in C_b^2(\mathbb{R}^{nd})$, we define a $C_b^2(\mathbb{R}^{nd})$-valued random process $Y^{(n)} = \{Y_t^{(n)}, 0 \leq t \leq T\}$ as follows: for any $j \geq 0$ and $t \in [\eta_j, \eta_{j+1})$,

$$Y_t^{(n)} := T_{t-\eta_j}^{(n)} S_j^{(n)} T_{\tau_j}^{(n)} \cdots S_2^{(n)} T_{\tau_2}^{(n)} S_1^{(n)} T_{\tau_1}^{(n)} f. \tag{2.3.6}$$

Then, $Y^{(n)}$ is a Markov process with $Y_0^{(n)} = f$. It involves countable many i.i.d. jumps $S_k^{(n)}$, controlled by i.i.d. exponential clocks τ_k. In between jumps, the process evolves deterministically by the continuous semigroup $T_t^{(n)}$. Notice that the exponential clock is memoryless, and the semigroup $T_t^{(n)}$ is generated by a time homogeneous Markov process. Therefore, $Y^{(n)}$ is also time homogeneous.

Lemma 2.3.7. *For any $n \geq 2$ and $f \in C_b^2(\mathbb{R}^{nd})$, let $Y_t^{(n)}$ be defined in (2.3.6). Then*

$$\mathbb{E}\Big(\sup_{x \in \mathbb{R}^{nd}} \big| Y_t^{(n)}(x) \big| \Big) \leq \|f\|_\infty \exp\left(\|\kappa\|_\infty \lambda_n t \right). \tag{2.3.7}$$

Proof. Since $T_t^{(n)}$ is the semigroup generated by a Markov process, for any $t>0$ and $f \in C_b^2(\mathbb{R}^{nd})$, $\|T_t^{(n)} f\|_\infty \le \|f\|_\infty$. By definition of the jump operators $\{S_j^{(n)}\}_{j\ge 1}$, it is easy to see that $\|S_j^{(n)} f\|_\infty \le \|\kappa\|_\infty \|f\|_\infty$. Thus we have

$$\mathbb{E}\Big(\sup_{x\in\mathbb{R}^{nd}} \big|Y_t^{(n)}(x)\big|\Big) \le \|f\|_\infty \sum_{j=0}^{\infty} \big[\|\kappa\|_\infty^j \mathbb{P}(\eta_j < t)\big]. \tag{2.3.8}$$

Notice that η_j is the sum of i.i.d. exponential random variables. Then, we have

$$\mathbb{P}(\eta_j < t) = 1 - \exp(-\lambda_n t) \sum_{k=0}^{j-1} \frac{(\lambda_n t)^k}{k!} = \exp(\lambda_n (t'-t)) \frac{(\lambda_n t)^j}{j!}, \tag{2.3.9}$$

for some $t' \in (0,t)$. Therefore, (2.3.7) follows from (2.3.8) and (2.3.9). ☐

Let $H^{(n)} : C_b^2(\mathbb{R}^{nd}) \times M_F(R^d) \to \mathbb{R}$ be given by

$$H^{(n)}(f,\mu) := G^{(n)}(f,\mu) - \lambda_n F^{(n)}(f,\mu).$$

Lemma 2.3.8. *Let $n \ge 2$ and $\mu \in M_F(\mathbb{R}^d)$. Then, the process*

$$F^{(n)}(Y_t^{(n)}, \mu) - \int_0^t H^{(n)}(Y_s^{(n)}, \mu) ds \tag{2.3.10}$$

is a martingale.

Proof. Let $\mu^{(n)}$ be any finite measure on $\mathbb{R}^{nd}$. Then, we have

$$\mathbb{E}\big(\mu^{(n)}(Y_t^{(n)})\big) = \mathbb{E}\big(\mu^{(n)}(Y_t^{(n)})\mathbf{1}_{\{\tau_1>t\}}\big) + \mathbb{E}\big(\mu^{(n)}(Y_t^{(n)})\mathbf{1}_{\{\eta_1\le t<\eta_2\}}\big) + o(t). \tag{2.3.11}$$

For the first term, we have

$$\mathbb{E}\big(\mu^{(n)}(Y_t^{(n)})\mathbf{1}_{\{\tau_1>t\}}\big) = \mu^{(n)}(T_t^{(n)} f)\mathbb{P}(\tau_1 > t) = \mu^{(n)}(T_t^{(n)} f)\exp(-\lambda_n t)). \tag{2.3.12}$$

For the second term, since τ_1, τ_2 are independent, then for any $0 \le s \le t$, we have

$$\mathbb{P}(\tau_1 + \tau_2 > t, \tau_1 \le s) = \int_0^s \int_{t-s_1}^{\infty} \lambda_n^2 \exp(-\lambda_n(s_1+s_2)) ds_2 ds_1 = \lambda_n s e^{-\lambda_n t}. \tag{2.3.13}$$

Note that by Lemma 2.3.7, $|Y_\cdot^{(n)}|$ is integrable on $[0,T] \times \mathbb{R}^{nd} \times \Omega$ with respect to the product measure $dt \times \mu^{(n)}(dx) \times P(d\omega)$. Then, by (2.3.13), Fubini's theorem, and the mean value theorem, we have

$$\begin{aligned}
&\mathbb{E}\big(\mu^{(n)}(Y_t^{(n)}) \mathbf{1}_{\{\eta_1 \le t < \eta_2\}}\big) \\
=&\frac{1}{2} \sum_{\substack{1 \le i,j \le n \\ i \ne j}} \int_0^t \int_{\mathbb{R}^{nd}} \big(T_{t-s}^{(n)} \Phi_{ij}^{(n)} T_s^{(n)} f\big)(x) \exp(-\lambda_n t)\, \mu^{(n)}(dx) ds \\
=&\frac{t}{2} \exp(-\lambda_n t) \sum_{\substack{1 \le i,j \le n \\ i \ne j}} \int_{\mathbb{R}^{nd}} \big(T_{t-t'}^{(n)} \Phi_{ij}^{(n)} T_{t'}^{(n)} f\big)(x) \mu^{(n)}(dx),
\end{aligned} \tag{2.3.14}$$

for some $t' \in (0,t)$. Combining (2.3.11), (2.3.12), and (2.3.14), we have

$$\lim_{t \downarrow 0} \frac{\mathbb{E}\big(\mu^{(n)}(Y_t^{(n)})\big) - \mu^{(n)}(f)}{t} = \mu^{(n)}(A^{(n)} f) + \frac{1}{2} \sum_{\substack{1 \le i,j \le n \\ i \ne j}} \mu^{(n)}\big(\Phi_{ij}^{(n)} f - f\big).$$

By Proposition 4.1.7 of Ethier & Kurtz [31], the following process:

$$\mu^{(n)}(Y_t^{(n)}) - \int_0^t \Big[\mu^{(n)}(A^{(n)} Y_s^{(n)}) + \frac{1}{2} \sum_{\substack{1 \le i,j \le n \\ i \ne j}} \mu^{(n)}(\Phi_{ij}^{(n)} Y_s^{(n)} - Y_s^{(n)})\Big] ds, \tag{2.3.15}$$

is a martingale. Then, the lemma follows by choosing $\mu^{(n)} = \mu^{\otimes n}$. □

By Lemma 2.3.6, 2.3.8 and Corollary 3.2 of Dawson & Kurtz [24], we have the following moment equality:

$$\mathbb{E}\big(X_t^{\otimes n}(f)\big) = \mathbb{E}\Big[X_0^{\otimes n}(Y_t^{(n)})\exp\Big(\int_0^t \lambda_n ds\Big)\Big] = \exp\Big(\frac{1}{2}n(n-1)t\Big)\mathbb{E}\big(X_0^{\otimes n}(Y_t^{(n)})\big). \tag{2.3.16}$$

Lemma 2.3.9. *Let $n \geq 2$, and let $f \in C_b^2(\mathbb{R}^{nd})$.*

(i) The following PDE

$$\partial_t v_t(x) = A^{(n)} v_t(x) + \frac{1}{2}\sum_{\substack{1\leq i,j\leq n\\ i\neq j}} \kappa(x_i,x_j)v(t,x), \tag{2.3.17}$$

with initial value $v_0(x) = f(x)$, has a unique solution.

(ii) Let $X = \{X_t, t \in [0,T]\}$ be a solution to the martingale problem (2.3.4) - (2.3.5). Then,

$$\mathbb{E}\big(X_t^{\otimes n}(f)\big) = X_0^{\otimes n}(v_t). \tag{2.3.18}$$

Proof. Firstly, we claim that the operator $A^{(n)} = \frac{1}{2}(\Delta + B^{(n)})$ is uniformly parabolic in the sense of Friedman (see Section 1.1 of Friedman [35]). Because Δ is uniformly parabolic, then it suffices to analyse the properties of $B^{(n)}$. For any $k = 1,\dots,n$, $i = 1,\dots,d$, and $\xi_k^i \in \mathbb{R}$, let $\xi_k = (\xi_k^1,\dots,\xi_k^d)$. Then, we have

$$\sum_{k_1,k_2=1}^n \sum_{i,j=1}^d \rho^{ij}(x_{k_1}-x_{k_2})\xi_{k_1}^i\xi_{k_2}^j = \int_{\mathbb{R}^d}\Big|\sum_{k=1}^n h^*(z-x_k)\xi_k\Big|^2 dz \geq 0.$$

Thus $B^{(n)}$ is parabolic. On the other hand, by Jensen's inequality, we have

$$\sum_{k_1,k_2=1}^n \sum_{i,j=1}^d \rho^{ij}(x_{k_1}-x_{k_2})\xi_{k_1}^i\xi_{k_2}^j = \int_{\mathbb{R}^d}\Big|\sum_{k=1}^n h^*(z-x_k)\xi_k\Big|^2 dz \leq n\|\rho\|_\infty \sum_{k=1}^n |\xi_k|^2.$$

It follows that $A^{(n)} = \frac{1}{2}(\Delta + B^{(n)})$ is uniformly parabolic.

Since $h \in H_3(\mathbb{R}^d;\mathbb{R}^d\otimes\mathbb{R}^d)$, $\rho(x-y) = \int_{\mathbb{R}^d} h(z-x)h^*(z-y)dz$ has bounded derivatives up to order three, then by Theorem 1.12 and 1.16 of Friedman [35], the PDE (2.3.17) has a unique solution.

In order to show (ii), let

$$\widetilde{v}_t(x) = \mathbb{E}\big(Y_t^{(n)}(x)\big),$$

where $Y^{(n)}$ is defined by (2.3.6). By the same argument as we did in the proof of Lemma 2.3.7, we can show that for any $t \in [0,T]$ and $x \in \mathbb{R}^{nd}$

$$\mathbb{E}\Big(\sup_{x\in\mathbb{R}^d} \big|A^{(n)}Y_t^{(n)}(x)\big|\Big) < \infty.$$

Then, by the dominated convergence theorem, we have

$$\mathbb{E}\big(A^{(n)}Y_t^{(n)}(x)\big) = A^{(n)}\mathbb{E}(Y_t^{(n)}(x)).$$

Let $\mu^{(n)}$ be any finite measure on $\mathbb{R}^{nd}$. Recall that the process defined by (2.3.15) is a martingale, then the following equality follows from Fubini's theorem:

$$\begin{aligned}\mu^{(n)}(\widetilde{v}_t) =&\mathbb{E}\big(\mu^{(n)}(Y_t^{(n)})\big) = \mu^{(n)}(f) + \int_0^t \langle \mu^{(n)}, \mathbb{E}\big(A^{(n)}Y_s^{(n)}\big)\rangle ds \\ &+\frac{1}{2}\sum_{\substack{1\le i,j\le n\\ i\neq j}} \int_0^t \langle \mu^{(n)}, [k(\cdot_i,\cdot_j)-1]\mathbb{E}(Y_s^{(n)})\rangle ds \\ =&\mu^{(n)}(f) + \int_0^t \langle \mu^{(n)}, A^{(n)}\widetilde{v}_s\rangle ds + \frac{1}{2}\sum_{\substack{1\le i,j\le n\\ i\neq j}} \int_0^t \langle \mu^{(n)}, [k(\cdot_i,\cdot_j)-1]\widetilde{v}_s\rangle ds.\end{aligned}$$

In other words,

$$\Big\langle \mu^{(n)}, \widetilde{v}_t - f - \int_0^t \Big[A^{(n)}\widetilde{v}_s - \frac{1}{2}\sum_{\substack{1\le i,j\le n\\ i\neq j}} (k(\cdot_i,\cdot_j)-1)\widetilde{v}_s\Big] ds\Big\rangle = 0,$$

for all $\mu^{(n)} \in M_F(\mathbb{R}^{nd})$. It follows that $\widetilde{v} = \{\widetilde{v}_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ solves the following PDE

$$\partial_t \widetilde{v}_t(x) = A^{(n)} \widetilde{v}_t(x) + \frac{1}{2} \sum_{\substack{1 \le i,j \le n \\ i \ne j}} [\kappa(x_i, x_j) - 1] \widetilde{v}_t(x), \tag{2.3.19}$$

with the initial value $\widetilde{v}_0(x) = f(x)$. This solution is unique by the same argument as in part (i). Observe that

$$v_t(x) := \widetilde{v}_t(x) \exp\Big(\frac{1}{2} n(n-1)t\Big), \tag{2.3.20}$$

solves the equation (2.3.17). Therefore, (2.3.18) follows from (2.3.20) and the moment duality (2.3.16). □

In Lemma 2.3.9, we derived the moment formula for $\mathbb{E}\big(X_t^{(n)}(f)\big)$ in the case when $n \ge 2$. If $n = 1$, the dual process only involves a deterministic evolution driven by the semigroup of one particle motion, which makes things much simpler. We write the formula below and skip the proof. Let $p(t,x,y)$ be the transition density of the one particle motion, then for any $\phi \in C_b^2(\mathbb{R}^d)$,

$$\mathbb{E}(X_t(\phi)) = X_0(T_t^{(1)} \phi) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} p(t,x,y) \phi(y) dy X_0(dx).$$

The existence of the density of X_t will be derived following Wang's idea (see Theorem 2.1 of Wang [80]). For any $\varepsilon > 0$, $x \in \mathbb{R}^d$, let p_ε be the heat kernel on $\mathbb{R}^d$, that is

$$p_\varepsilon(x) = (2\pi\varepsilon)^{-\frac{d}{2}} \exp\Big(-\frac{|x|^2}{2\varepsilon}\Big).$$

Lemma 2.3.10. *Let $X = \{X_t, t \in [0,T]\}$ be a solution to the martingale problem (2.3.4) - (2.3.5). Assume that the initial measure $X_0 \in M_F(\mathbb{R}^d)$ has a bounded density μ. Then,*

$$\int_0^T \int_{\mathbb{R}^d} \mathbb{E}\big(\big|X_t(p_\varepsilon(x-\cdot))\big|^2\big) dx dt < \infty, \tag{2.3.21}$$

and

$$\lim_{\varepsilon_1,\varepsilon_2\downarrow 0}\int_0^T\int_{\mathbb{R}^d}\mathbb{E}\big(\big|X_t(p_{\varepsilon_1}(x-\cdot))-X_t(p_{\varepsilon_2}(x-\cdot))\big|^2\big)dxdt=0. \tag{2.3.22}$$

Proof. Let $\Gamma(t,(y_1,y_2);r,(z_1,z_2))$ be the fundamental solution to the PDE (2.3.17) when $n=2$ (see Chapter 1 of Friedman [35] for a detailed account on existence and properties of fundamental solutions to parabolic PDEs). We write $y=(y_1,y_2)$ and $z=(z_1,z_2)\in\mathbb{R}^{2d}$. Then, for $f\in C_b^2(\mathbb{R}^{2d})$,

$$v(t,y)=\int_{\mathbb{R}^{2d}}\Gamma(t,y;0,z)f(z)dz,$$

is the unique solution to the PDE (2.3.17) with initial condition $v_0=f$. Thus by Lemma 2.3.9, we have

$$\begin{aligned}&\mathbb{E}\big[X_t(p_{\varepsilon_1}(x-\cdot))X_t(p_{\varepsilon_2}(x-\cdot))\big]\\&\qquad=\int_{\mathbb{R}^{2d}}\int_{\mathbb{R}^{2d}}\Gamma(t,y;0,z)p_{\varepsilon_1}(x-z_1)p_{\varepsilon_2}(x-z_2)dzX_0^{\otimes 2}(dy).\end{aligned} \tag{2.3.23}$$

By the inequality (6.12) of Friedman [35] on page 24, we know that there exists $C_\Gamma,\lambda>0$, such that for any $0\le r<t\le T$,

$$|\Gamma(t,y;r;z)|\le C_\Gamma p_{\frac{t-r}{\lambda}}(y_1-z_1)p_{\frac{t-r}{\lambda}}(y_2-z_2). \tag{2.3.24}$$

Therefore, by the semigroup property of heat kernels and Fubini's theorem, we have

$$\begin{aligned}&\int_0^T\int_{\mathbb{R}^d}\mathbb{E}\big[X_t(p_{\varepsilon_1}(x-\cdot))X_t(p_{\varepsilon_2}(x-\cdot))\big]dxdt\\&\qquad=\int_0^T\int_{\mathbb{R}^{2d}}\int_{\mathbb{R}^{2d}}\Gamma(t,y;0,z)p_{\varepsilon_1+\varepsilon_2}(z_1-z_2)dzX_0^{\otimes 2}(dy)dt\end{aligned} \tag{2.3.25}$$

From (2.3.24), (2.3.25) and the fact that $X_0\in M_F(\mathbb{R}^d)$ has a bounded density μ, it follows that (2.3.21) is true.

Let $\mathscr{M}$ be the function on $\mathbb{R}^{2d}$, given by

$$\mathscr{M}(z) = \int_0^T \int_{\mathbb{R}^{2d}} \Gamma(t,y;0,z) X_0^{\otimes 2}(dy)dt.$$

Notice that fix $0 \le r < t \le T$, $\Gamma(t,y;r,x)$ is uniformly continuous in the spatial argument (see (6.13) of Friedman [35] on page 24). As a consequence $\mathscr{M}$ is continuous. Therefore, by (2.3.24) and the continuity of $\mathscr{M}$, the function $\mathscr{N}$ on $\mathbb{R}^d$ given by

$$\mathscr{N}(x) := \int_{\mathbb{R}^d} \mathscr{M}(z_1, z_1 - x) dz_1,$$

is integrable and continuous everywhere. It follows that

$$\begin{aligned}
&\lim_{\varepsilon_1,\varepsilon_2\to 0} \int_0^t \int_{\mathbb{R}^d} \mathbb{E}\big[X_t(p_{\varepsilon_1}(x-\cdot))X_t(p_{\varepsilon_2}(x-\cdot))\big] dxdt \\
&\quad = \lim_{\varepsilon_1,\varepsilon_2\to 0} \int_{\mathbb{R}^{2d}} \mathscr{M}(z) p_{\varepsilon_1+\varepsilon_2}(z_1 - z_2) dz \\
&\quad = \lim_{\varepsilon_1,\varepsilon_2\to 0} \int_{\mathbb{R}^d} \mathscr{N}(y) p_{\varepsilon_1+\varepsilon_2}(y) dy \\
&\quad = \mathscr{N}(0) = \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^{2d}} \Gamma(t,y;0,(x,x)) X_0^{\otimes 2}(dy) dxdt \qquad (2.3.26)
\end{aligned}$$

Therefore, (2.3.22) is a consequence of (2.3.26). □

Proposition 2.3.11. *Let $X = \{X_t, t \in [0,T]\}$ be a solution to the martingale problem (2.3.4) - (2.3.5). Assume that the initial measure $X_0 \in M_F(\mathbb{R}^d)$ has a bounded density μ. Then, for almost every $t \in (0,T]$, X_t is absolutely continuous with respect to the Lebesgue measure almost surely.*

Proof. As proved in Lemma 2.3.10, for any $x \in \mathbb{R}^d$ and $\varepsilon_n \downarrow 0$, the sequence $\{X_t(p^x_{\varepsilon_n})\}_{n\ge 1}$ is Cauchy in $L^2(\Omega \times \mathbb{R}^d \times [0,T])$. Then, it converges to some square integrable random field. By the same argument as in Theorem 2.1 of Wang [80], we can show that the limit random field is the density of X_t almost surely. □

Remark 2.3.12. *(i) The assumption in Proposition 2.3.11, that the initial measure has a bounded density, cannot be removed. Actually, if we choose $X_0 = \delta_0$, the Dirac delta mass at 0, then*

$\int_0^T \int_{\mathbb{R}^d} \Gamma(t,0;0,(x,x))dxdt$ behaves like $\int_0^T t^{-\frac{d}{2}}dt$, which is finite only if $d = 1$. This is another difference from the one dimensional situation, in which case $X_0(1) < \infty$ is enough to imply the existence of the density (see Theorem 2.1 Wang [80] for the Dawson-Watanabe branching model).

(ii) *The method of duality is conventionally used to prove the well-posedness of martingale problems arisen from branching mechanisms. In the one-dimensional Dawson-Watanabe model, Wang proved the well-posedness by solving a moment problem (see Section 4 of Wang [81]). This requires a moment bound of the form $\sum_{n=1}^{\infty} r^n \mathbb{E}(|X_t(1)|^n)/n! < \infty$ for some positive r. However, this method does not work in our model and here is the explanation. In the next section, we will prove that the density u is a solution to equation (2.2.1) and when $h \equiv 0$, we have that $\mathbb{E}(\langle u_t, 1\rangle^n)$ behaves like $c_1 e^{c_2 n^{1+\varepsilon}}$ for some $\varepsilon > 0$ (c.f. Theorem 4.4 of Chen et al. [15] and Theorem 4.3 of Hu et al. [41] for some sharp bounds of similar SPDE's). Therefore, the condition $\sum_{n=1}^{\infty} r^n \mathbb{E}(|\langle u_t, 1\rangle|^n)/n! < \infty$ for some positive r cannot be satisfied in our model. In the next section, we prove the well-posedness of the martingale problem (2.3.4) - (2.3.5) by the Yamada-Watanabe argument assuming the existence of the density. Without the existence of the density, we are currently not able to use the moment duality to show the well-posedness of the martingale problem. We are not pursue this in the current chapter.*

2.3.3 Proof of Theorems 2.2.2 and 2.2.3

The proof of Theorems 2.2.2 and 2.2.3 is based on the equivalence of the martingale problem (2.3.4) - (2.3.5) and the SPDE (2.2.1).

The equivalence between martingale problems and SDE's in finite dimensions was observed in the 1970s (c.f. Stroock & Varadhan [77]). An alternative proof given by Kurtz [58] consists of the "Markov mapping theorem". In a recent paper, Biswas et al. [7] generalized this result to the infinite dimensional cases with one noise following Kurtz's idea. Here in the present chapter, we establish a similar result with two noises by using the martingale representation theorem.

Proposition 2.3.13. *Let $\mu \in C_b(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ be a nonnegative function on $\mathbb{R}^d$. Then, $u = \{u_t, t \in [0,T]\}$ is the density of a solution of the martingale problem (2.3.4) - (2.3.5) with initial density μ, if and only if u is a weak solution to the SPDE (2.2.1).*

Proof. If u is a weak solution to (2.2.1), then, as a consequence of Itô's formula, u is the density of a measure-valued process that solves the martingale problem (2.3.4) - (2.3.5). It suffices to show the converse statement.

Let $X = \{X_t, t \in [0,T]\}$ be a solution to the martingale problem (2.3.4) - (2.3.5) with initial density μ. Then, by Proposition 2.3.11, for almost every $t \in [0,T]$, X_t has a density almost surely. We denote by u_t the density of X_t.

Consider $M = \{M_t, t \in [0,T]\}$ defined by (2.3.4) as an $\mathscr{S}'$-martingale (see Definition 2.1.2 of Kallianpur & Xiong [52]). Then, by Theorem 3.1.4 of Kallianpur & Xiong [52], there exists a Hilbert space $\mathscr{H}^* \supset L^2(\mathbb{R}^d)$, such that M is an $\mathscr{H}^*$-valued martingale. Denote by $\mathscr{H}$ the dual space of $\mathscr{H}^*$.

Let $\mathfrak{H}_1 = L^2(\mathbb{R}^d;\mathbb{R}^d)$, and let $\mathfrak{H}_2$ be the completion of $\mathscr{S}$ with the inner product

$$\langle \phi, \varphi \rangle_{\mathfrak{H}_2} := \int_{\mathbb{R}^d \times \mathbb{R}^d} \kappa(x,y)\phi(x)\varphi(y)dxdy.$$

Consider the product space $\mathfrak{H} = \mathfrak{H}_1 \times \mathfrak{H}_2$. Then, $\mathfrak{H}$ is a Hilbert space equipped with the inner product

$$\big\langle (\phi_1,\phi_2),(\varphi_1,\varphi_2) \big\rangle_{\mathfrak{H}} := \langle \phi_1, \varphi_1 \rangle_{\mathfrak{H}_1} + \langle \phi_2, \varphi_2 \rangle_{\mathfrak{H}_2}.$$

For any $t \in [0,T]$, let $\Psi_t : \mathscr{H} \to \mathfrak{H}$ be given by $\Psi_t(\phi)(x,y) = \big(\Psi_t^1(\phi)(x), \Psi_t^2(\phi)(y)\big)$, where

$$\Psi_t^1(\phi)(x) := \int_{\mathbb{R}^d} \nabla\phi(y)^* h(x-y)u_t(y)dy,$$

and

$$\Psi_t^2(\phi)(x) := \phi(x)u_t(x).$$

Then, for any $\phi, \varphi \in \mathscr{H}$, we have

$$\begin{aligned}\langle M(\phi), M(\varphi)\rangle_t &= \int_0^t \nabla\phi(x)^*\rho(x-y)\nabla\varphi(y)X_s(dx)X_s(dy)ds \\ &\quad + \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} \kappa(x,y)\phi(x)\phi(y)X_s(dx)X_s(dy)ds \\ &= \int_0^t \langle \Phi_s(\phi), \Phi_s(\varphi)\rangle_{\mathfrak{H}} ds,\end{aligned}$$

Therefore, by the martingale representation theorem (c.f. Theorem 3.3.5 of Kallianpur & Xiong [52]), there exists a $\mathfrak{H}$-cylindrical Brownian motion $\mathfrak{B} = \{\mathfrak{B}_t, 0 \le t \le T\}$, such that

$$M_t(\phi) = \int_0^t \left\langle \Psi_s(\phi), d\mathfrak{B}_s \right\rangle.$$

Let $\mathfrak{B}^1 = \{\mathfrak{B}^1_t(\phi), t \in [0,T], \phi \in \mathfrak{H}_1\}$ and $\mathfrak{B}^2 = \{\mathfrak{B}^2_t(\varphi), t \in [0,T], \varphi \in \mathfrak{H}_2\}$ be given by

$$\mathfrak{B}^1_t(\phi) = \mathfrak{B}_t(\phi, 0) \text{ and } \mathfrak{B}^2_t(\varphi) = \mathfrak{B}_t(0, \varphi).$$

Then, $\mathfrak{B}^1$ and $\mathfrak{B}^2$ are $\mathfrak{H}^1$- and $\mathfrak{H}^2$-cylindrical Brownian motion respectively, and they are independent. As a consequence, we have

$$M_t(\phi) = \int_0^t \left\langle \int_{\mathbb{R}^d} \nabla\phi(z)^* h(\cdot - z)X_s(dz), d\mathfrak{B}^1_s \right\rangle + \int_0^t \left\langle \phi u_s, d\mathfrak{B}^2_s \right\rangle. \tag{2.3.27}$$

Let $\{e_j\}_{j\ge 1}$ be a complete orthonormal basis of $\mathfrak{H}_2$. Then, by Theorem 3.2.5 of Kallianpur & Xiong [52], $V = \{V_t, t \in [0,T]\}$, defined by

$$V_t := \sum_{j=1}^{\infty} \mathfrak{B}^2_t(e_j)e_j,$$

is a $\mathscr{S}'$-valued Wiener process with covariance

$$\mathbb{E}\left[V_s(\phi)V_t(\varphi)\right] = s \wedge t \int_{\mathbb{R}^d\times\mathbb{R}^d} \kappa(x,y)\phi(x)\varphi(y)dxdy,$$

for any $\phi, \varphi \in \mathscr{S}$. Therefore, by (2.3.27) and the equivalence of stochastic integrals with respect to Hilbert space valued Brownian motion and Walsh's integrals (c.f. Proposition 2.6 of Dalang & Sardanyons [18] for spatial homogeneous noises), u is a weak solution to the SPDE (2.2.1). □

Proof of Theorems 2.2.2 and 2.2.3. By Propositions 2.3.5 and 2.3.13, the SPDE (2.2.1) has a weak solution, that can be obtained by the branching particle approximation. We do not prove the continuity here, because later in Section 2.6, we will show that the solution is not only continuous, but also Hölder continuous. The continuity of u yields an improved version of Proposition 2.3.11. Namely, if X_t is a continuous measure-valued process (e.g. the limit of the particle approximation), then it has a density for all $t \in [0,T]$ almost surely.

In the next step, we prove the pathwise uniqueness of equation (2.2.1). Assume that u and $\widetilde{u}$ are two continuous strong solutions to (2.2.1) with initial condition μ. Let $d = u - \widetilde{u}$. Then, $d = \{d_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ is a solution to (2.2.1), with initial condition $\mu \equiv 0$, that is continuous in two parameters. Thus d is also the density of a solution to the martingale problem (2.3.4) - (2.3.5), with initial measure $X_0 \equiv 0$.[1] By the moment duality (2.3.16), for any $\phi \in C_b^2(\mathbb{R}^d)$, we have

$$\mathbb{E}\langle d_t, \phi \rangle^2 = \exp(t)\mathbb{E}\big(X_0(Y_t^{(2)})\big) \equiv 0,$$

where $Y^{(2)}$ is the dual process defined by (2.3.6) in the case $n = 2$. Since d is continuous in t, it follows that $u = \widetilde{u}$ almost surely. Therefore, by the Yamada-Watanabe argument (c.f. Yamada & Watanabe [85] and Kurtz [57]), we obtain the strong existence and weak uniqueness of equation (2.2.1). This proves Theorem 2.2.3. Recall Propositions 2.3.5 and 2.3.13. The weak uniqueness of equation (2.2.1) also implies that every limit of the convergent subsequence of $\{X^n\}_{n\geq 1}$ constructed in Section 2.3.1 is continuous (see Lemma 2.3.4) and unique in law. In other words, $\{X^n\}_{n\geq 1}$ is convergent in $D([0,T]; M_F)$ to a continuous $M_F(\mathbb{R}^d)$-valued process in law. The limit has a density almost surely, that is a weak solution the SPDE (2.2.1). □

[1] $d_t(x)$ may be negative for some $(t,x) \in [0,T] \times \mathbb{R}^d$. In this case d_t is considered as the density of a signed measure ν, where $|\nu|(1) \leq |u_t(1)| + |\widetilde{u}_t(1)| < \infty$ a.s.. The moment duality still holds.

The following corollary is a direct result of Theorem 2.2.3 and Proposition 2.3.13.

Corollary 2.3.14. *Assume Hypotheses (H0),(H1), and assume that $X_0 \in M_F(\mathbb{R}^d)$ has a bounded density. Then, the martingale problem (2.3.4) - (2.3.5) is well-posed.*

2.4 Moment estimates for one-particle motion

In this section, we focus on the one-particle motion without branching. By using the techniques of Malliavin calculus, we will obtain moment estimates for the transition probability density of the particle motion conditional on the environment W. A brief introduction and several theorems on Malliavin calculus are stated in Appendix 2.7. For a detailed account on this topic, we refer the readers to the book of Nualart [71].

Fix a time interval $[0,T]$. Let $B = \{B_t, 0 \le t \le T\}$ be a standard d-dimensional Brownian motion and let W be a d-dimensional space-time white Gaussian random field on $[0,T] \times \mathbb{R}^d$ that is independent of B. Assume that $h \in H_3(\mathbb{R}^d;\mathbb{R}^d \otimes \mathbb{R}^d)$. For any $0 \le r < t \le T$, we denote by $\xi_t = \xi_t^{r,x}$, the path of one-particle motion, with initial position $\xi_r = x$. It satisfies the SDE:

$$\xi_t = x + B_t - B_r + \int_r^t \int_{\mathbb{R}^d} h(y - \xi_u) W(du, dy). \tag{2.4.1}$$

We will apply the Malliavin calculus on ξ_t with respect to the Brownian motion B. Let $H = L^2([0,T];\mathbb{R}^d)$ be the associated Hilbert space. By the Picard iteration scheme (c.f. Theorem 2.2.1 of Nualart [71]), we can prove that for any $t \in (r,T]$, $\xi_t \in \cap_{p\ge 1} \mathbb{D}^{3,p}(\mathbb{R}^d)$. Particularly, $D\xi_t$ satisfies the following system of SDE's:

$$D_\theta^{(k)} \xi_t^i = \delta_{ik} - \sum_{j_1,j_2=1}^d \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_1} h^{ij_2}(y - \xi_s) D_\theta^{(k)} \xi_s^{j_1} W^{j_2}(ds, dy), \quad 1 \le i,k \le d, \tag{2.4.2}$$

for any $\theta \in [r,t]$, and $D_\theta^{(k)} \xi_t^i = 0$ for all $\theta > t$.

In order to simplify the expressions, we rewrite the stochastic integrals in (2.4.2) as integrals with respect to martingales. To this end, let $M = \{M_t, r \le t \le T\}$ be the $d \times d$ matrix-valued process given by

$$M_t = \sum_{k=1}^{d} \int_r^t \int_{\mathbb{R}^d} g_k(s,y) W^k(ds,dy),$$

where $g_k : \Omega \times [r,T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is given by

$$g_k^{ij}(t,y) = \partial_i h^{jk}(y - \xi_t), \quad 1 \le i,j,k \le d.$$

Notice that M_t is the sum of stochastic integrals, so it is a matrix-valued martingale. The quadratic covariations of $\{M^{ij}\}_{i,j=1}^d$ are bounded and deterministic:

$$\begin{aligned} \left\langle M^{i_1 j_1}, M^{i_2 j_2} \right\rangle_t &= \sum_{k=1}^{d} \int_r^t \int_{\mathbb{R}^d} \partial_{i_1} h^{j_1 k}(y - \xi_s) \partial_{i_2} h^{j_2 k}(y - \xi_s) dy ds \\ &= (t-r) \sum_{k=1}^{d} \int_{\mathbb{R}^d} \partial_{i_1} h^{j_1 k}(y) \partial_{i_2} h^{j_2 k}(y) dy := Q_{i_2, j_2}^{i_1, j_1}(t-r) \le \|h\|_{3,2}(t-r). \end{aligned} \tag{2.4.3}$$

Now the equation (2.4.2) can be rewritten as follows:

$$D_\theta^{(k)} \xi_t^i = \delta_{ik} - \sum_{j=1}^{d} \int_\theta^t \int_{\mathbb{R}^d} D_\theta^{(k)} \xi_s^j dM_s^{ji}, \quad 1 \le i,k \le d. \tag{2.4.4}$$

Lemma 2.4.1. *For any $0 \le r < t \le T$, $x \in \mathbb{R}^d$, let $\gamma_t = \gamma_{\xi_t}$ be the Malliavin matrix of $\xi_t = \xi_t^{r,x}$, then γ_t is nondegenerate almost surely.*

Proof. We prove the lemma following Stroock's idea (see Chapter 8 of Stroock [76]). Let $\lambda_\theta(t)$ be the $d \times d$ symmetric random matrix given by

$$\lambda_\theta^{ij}(t) = \sum_{k=1}^{d} D_\theta^{(k)} \xi_t^i D_\theta^{(k)} \xi_t^j.$$

Then, the Malliavin matrix of ξ_t is the integral of $\lambda_\theta(t)$:

$$\gamma_t = \int_r^t \lambda_\theta(t) d\theta.$$

By (2.4.2), (2.4.3) and Itô's formula, we have

$$\begin{aligned}D_\theta^{(k)}\xi_t^i D_\theta^{(k)}\xi_t^j =&\delta_{ik}\delta_{kj}-\sum_{k_1=1}^d\int_\theta^t D_\theta^{(k)}\xi_s^i D_\theta^{(k)}\xi_s^{k_1}dM_s^{k_1j}-\sum_{k_2=1}^d\int_\theta^t D_\theta^{(k)}\xi_s^j D_\theta^{(k)}\xi_s^{k_2}dM_s^{k_2i}\\&+\sum_{k_1,k_2=1}^d Q_{k_2,i}^{k_1,j}\int_\theta^t D_\theta^{(k)}\xi_s^{k_1}D_\theta^{(k)}\xi_s^{k_2}ds.\end{aligned}$$

Therefore,

$$\begin{aligned}\lambda_\theta(t)=&I-\int_\theta^t\lambda_\theta(s)dM_s-\int_\theta^t dM_s^*\cdot\lambda_\theta(s)\\&+\sum_{k=1}^d\int_\theta^t\int_{\mathbb{R}^d}g_k^*(s,y)\lambda_\theta(s)g_k(s,y)dyds.\end{aligned}\tag{2.4.5}$$

For any $\theta\in[r,t]$, we claim that $\lambda_\theta(t)$ is invertible almost surely, and its inverse $\beta_\theta(t)$ satisfies the following SDE:

$$\begin{aligned}\beta_\theta(t)=&I+\int_\theta^t\beta_\theta(s)dM_s^*+\int_\theta^t dM_s\cdot\beta_\theta(s)\\&+\sum_{k=1}^d\int_\theta^t\int_{\mathbb{R}^n}\left(g_k(s,y)^2\beta_\theta(s)+g_k(s,y)\beta_\theta(s)g_k^*(s,y)+\beta_\theta(s)g_k^*(s,y)^2\right)dyds.\end{aligned}\tag{2.4.6}$$

Indeed, by Itô's formula, we have

$$\begin{aligned}d[\lambda_\theta(t)\beta_\theta(t)]=&-dM_t^*\cdot[\lambda_\theta(t)\beta_\theta(t)]+[\lambda_\theta(t)\beta_\theta(t)]dM_t^*\\&+\sum_{k=1}^d\Big(\int_{\mathbb{R}^d}\left([\lambda_\theta(t)\beta_\theta(t)]g_k^*(t,y)^2-g_k^*(t,y)[\lambda_\theta(t)\beta_\theta(t)]g_k^*(t,y)\right)dy\Big)dt.\end{aligned}\tag{2.4.7}$$

Note that $\lambda_\theta(t)\beta_\theta(t)\equiv I$ solves the SDE (2.4.7) with initial value $\lambda_\theta(\theta)\beta_\theta(\theta)=I$. Therefore, the strong uniqueness of the linear SDE (2.4.7) implies that $\lambda_\theta^{-1}(t)=\beta_\theta(t)$ almost surely.

Denote by $\|\cdot\|_2$ the Hilbert-Schmidt norm of matrices. By Jensen's inequality (see Lemma

8.14 of Stroock [76]), the following inequality holds almost surely

$$\|\gamma_t^{-1}\|_2 = \Big\|\Big(\int_r^t \lambda_\theta(t)d\theta\Big)^{-1}\Big\|_2 \le \frac{1}{(t-r)^2}\Big\|\int_r^t \beta_\theta(t)d\theta\Big\|_2. \tag{2.4.8}$$

It is easy to show that $\sup_{\theta\in[r,t]} \big\|\|\beta_\theta(t)\|_2\big\|_{2p} < \infty$ for all $p \ge 1$. Therefore, the right-hand side of (2.4.8) is finite a.s., and thus γ_t is nondegenerate almost surely. □

We denote by $\sigma_t = \gamma_t^{-1}$ the inverse of the Malliavin matrix of ξ_t. In the following lemma, we obtain some moment estimates for the derivatives of ξ_t and σ_t. Before estimates, we introduce the following generalized Cauchy-Schwarz's inequality.

Lemma 2.4.2. *Let n_1, n_2 be nonnegative integers, $u_1 \in L^{2p}(\Omega;(H^{\otimes n_1}))$, and $u_2 \in L^{2p}(\Omega,(H^{\otimes n_2}))$, then $u_1 \otimes u_2 \in L^p(\Omega;(H^{\otimes(n_1+n_2)}))$, and*

$$\big\|\|u_1 \otimes u_2\|_{H^{\otimes(n_1+n_2)}}\big\|_p \le \big\|\|u_1\|_{H^{\otimes n_1}}\big\|_{2p}\big\|\|u_2\|_{H^{\otimes n_2}}\big\|_{2p}. \tag{2.4.9}$$

Proof. The lemma can be obtained by the classical Cauchy-Schwarz inequality. □

Lemma 2.4.3. *For any $p \ge 1$ and $0 \le r < t \le T$, there exists a constant $C > 0$, that depends on T, d, $\|h\|_{3,2}$, p, such that*

$$\max_{1\le i\le d}\big\|\|D\xi_t^i\|_H\big\|_{2p} \le C(t-r)^{\frac{1}{2}}. \tag{2.4.10}$$

$$\max_{1\le i,j\le d}\Big\|\sigma_t^{ij}\Big\|_{2p} \le C(t-r)^{-1}, \tag{2.4.11}$$

$$\max_{1\le i,j\le d}\Big\|\|D\sigma_t^{ij}\|_H\Big\|_{2p} \le C, \tag{2.4.12}$$

$$\max_{1\le i\le d}\big\|\|D^2\xi_t^i\|_{H^{\otimes 2}}\big\|_{2p} \le C(t-r)^{\frac{3}{2}}. \tag{2.4.13}$$

$$\max_{1\le i,j\le d}\Big\|\|D^2\sigma_t^{ij}\|_{H^{\otimes 2}}\Big\|_{2p} \le C(t-r)^{\frac{1}{2}}, \tag{2.4.14}$$

$$\max_{1\le i\le d}\big\|\|D^3\xi_t^i\|_{H^{\otimes 3}}\big\|_{2p} \le C(t-r)^2. \tag{2.4.15}$$

Proof of (2.4.10). By (2.4.3), (2.4.4), Jensen's, Burkholder-Davis-Gundy's, and Minkowski's inequalities, we have

$$\begin{aligned}\sum_{i,k=1}^{d}\|D_\theta^{(k)}\xi_t^i\|_{2p}^2 \leq &\sum_{i,k=1}^{d}\left(\delta_{ik}+\sum_{j=1}^{d}\left\|\int_\theta^t\int_{\mathbb{R}^d}D_\theta^{(k)}\xi_s^j dM_s^{ji}\right\|_{2p}\right)^2\\ \leq&(d+1)\sum_{i,k=1}^{d}\left(\delta_{ik}+\sum_{j=1}^{d}\left\|\int_\theta^t\int_{\mathbb{R}^d}D_\theta^{(k)}\xi_s^j dM_s^{ji}\right\|_{2p}^2\right)\\ \leq&d(d+1)+(d+1)c_p\sum_{i,j,k=1}^{d}Q_{ji}^{ji}\left\|\int_\theta^t|D_\theta^{(k)}\xi_s^j|^2ds\right\|_p\\ \leq&d(d+1)+2c_pd(d+1)\|h\|_{3,2}^2\sum_{j,k=1}^{d}\int_\theta^t\|D_\theta^{(k)}\xi_s^j\|_{2p}^2ds. \qquad (2.4.16)\end{aligned}$$

Thus by Gronwall's lemma, we have

$$\sum_{i,j=1}^{d}\|D_\theta^{(k)}\xi_t^j\|_{2p}^2\leq d(d+1)\exp\left(2c_pd(d+1)\|h\|_{3,2}^2T\right):=C. \qquad (2.4.17)$$

Therefore, by (2.4.17) and Minkowski's inequality, we have

$$\big\|\|D\xi_t^i\|_H\big\|_{2p}^2=\left\|\sum_{k=1}^{d}\int_r^t|D_\theta^{(k)}\xi_t^i|^2d\theta\right\|_p\leq\sum_{k=1}^{d}\int_r^t\|D_\theta^{(k)}\xi_t^i\|_{2p}^2d\theta\leq C(t-r).$$

□

Proof of (2.4.11). In order to prove (2.4.11), we rewrite the SDE (2.4.6) in the following way:

$$\begin{aligned}\beta_\theta^{ij}(t)=&\delta_{ij}+\sum_{k_1=1}^{d}\int_\theta^t\beta_\theta^{ik_1}(s)dM_s^{jk_1}+\sum_{k_2=1}^{d}\int_\theta^t\beta_\theta^{k_2j}(s)dM_s^{ik_2}\\ &+\sum_{k_1,k_2=1}^{d}\left(Q_{k_1,k_2}^{i,k_1}\int_\theta^t\beta_\theta^{k_2j}(s)ds\right)+\sum_{k_1,k_2=1}^{d}\left(Q_{j,k_2}^{i,k_1}\int_\theta^t\beta_\theta^{k_1k_2}(s)ds\right)\\ &+\sum_{k_1,k_2=1}^{d}\left(Q_{j,k_2}^{k_2,k_1}\int_\theta^t\beta_\theta^{ik_1}(s)ds\right). \qquad (2.4.18)\end{aligned}$$

Similarly as we did in step **(i)**, by Burkholder-Davis-Gundy's, and Minkowski's inequalities, we can show that the martingale terms satisfies the following inequality

$$\Big\| \int_\theta^t \beta_\theta^{ik_1}(s) dM_s^{jk_1} \Big\|_{2p}^2 \leq 2c_p \|h\|_{3,2}^2 \int_\theta^t \|\beta_\theta^{ik_1}(s)\|_{2p}^2 ds. \tag{2.4.19}$$

For the drift terms, by Minkowski's and Jensen's inequality, we have

$$\Big\| \int_\theta^t \beta_\theta^{k_1 k_2}(s) ds \Big\|_{2p}^2 \leq (t-\theta) \int_\theta^t \Big\| \beta_\theta^{k_1 k_2}(s) \Big\|_{2p}^2 ds. \tag{2.4.20}$$

Then, by (2.4.18) - (2.4.20), and Gronwall's lemma, we have

$$\sum_{i,j=1}^d \Big\| \beta_\theta^{ij}(t) \Big\|_{2p}^2 \leq C.$$

Thus by Minkowski's and Jensen's inequalities, we have

$$\Big\| \Big\| \int_r^t \beta_\theta(t) d\theta \Big\|_2 \Big\|_{2p} \leq c_d \sum_{i,j=1}^d \int_r^t \Big\| \beta_\theta^{ij}(t) \Big\|_{2p} d\theta \leq C(t-r). \tag{2.4.21}$$

Therefore, (2.4.11) follows from (2.4.8), (2.4.21), Minkowski's and Jensen's inequalities. □

Proof of (2.4.12). By integrating equation (2.4.5) on both sides with respect to θ, and applying the stochastic Fubini theorem (c.f. Lemma 4.1 on page 116 of Ikeda & Watanabe [49]), we have

$$\begin{aligned}\gamma_t = \int_r^t \lambda_\theta(t) d\theta =& I(t-r) - \int_r^t \gamma_s dM_s - \int_r^t dM_s^* \cdot \gamma_s \\ &+ \sum_{m=1}^d \int_r^t \int_{\mathbb{R}^d} g_m^*(y,s) \gamma_s g_m(y,s) dy ds.\end{aligned} \tag{2.4.22}$$

Taking the Malliavin derivative on both sides of (2.4.22), we have the following SDE:

$$
\begin{aligned}
D_\theta^{(k)}\gamma_t^{ij} = & -\sum_{k_1=1}^d \int_\theta^t D_\theta^{(k)}\gamma_s^{ik_1} dM_s^{k_1 j} - \sum_{k_1=1}^d \int_\theta^t \gamma_s^{ik_1} d\left(D_\theta^{(k)} M_s^{k_1 j}\right) \\
& - \sum_{k_2=1}^d \int_\theta^t D_\theta^{(k)}\gamma_s^{k_2 j} dM_s^{k_2 i} - \sum_{k_2=1}^d \int_\theta^t \gamma_s^{k_2 j} d\left(D_\theta^{(k)} M_s^{k_2 i}\right) \\
& + \sum_{k_1,k_2=1}^d \left(Q_{k_2,j}^{k_1,i} \int_\theta^t D_\theta^{(k)}\gamma_s^{k_1 k_2} ds \right),
\end{aligned}
\tag{2.4.23}
$$

where

$$
D_\theta^{(k)} M_s^{ij} = -\sum_{i_1,i_2=1}^d \int_\theta^s \int_{\mathbb{R}^d} \partial_{i,i_2} h^{ji_1}(y-\xi_r) D_\theta^{(k)} \xi_r^{i_2} W^{i_1}(dr,dy). \tag{2.4.24}
$$

For the first and the third term, by similar arguments as in (2.4.16), we can show that

$$
\left\| \int_\theta^t D_\theta^{(k)}\gamma_s^{ik_1} dM_s^{k_1 j} \right\|_{2p}^2 \leq c_{d,p} \|h\|_{3,2}^2 \int_\theta^t \|D_\theta^{(k)}\gamma_s^{ik_1}\|_{2p}^2 ds. \tag{2.4.25}
$$

To estimate the second and the fourth term, notice that by (2.4.10), we have

$$
\begin{aligned}
\max_{1\leq i,j\leq d} \left\| \gamma_t^{ij} \right\|_{2p} = & \max_{1\leq i,j\leq d} \left\| \langle D\xi_t^i, D\xi_t^j \rangle_H \right\|_{2p} \\
& \leq \max_{1\leq i\leq d} \left\| \|D\xi_t^i\|_H \right\|_{4p} \max_{1\leq j\leq d} \left\| \|D\xi_t^j\|_H \right\|_{4p} \leq C(t-r).
\end{aligned}
\tag{2.4.26}
$$

Therefore, by (2.4.17), (2.4.24), (2.4.26), Jensen's, Burkholder-Davis-Gundy's, Minkowski's, and Cauchy-Schwarz's inequalities, we have

$$
\begin{aligned}
\left\| \int_\theta^t \gamma_s^{ik_1} d\left(D_\theta^{(k)} M_s^{k_1 j}\right) \right\|_{2p}^2 \leq & c_{d,p} \|h\|_{3,2}^2 \sum_{k_2=1}^d \int_\theta^t \|\gamma_s^{ik_1}\|_{4p}^2 \|D_\theta^{(k)}\xi_s^{k_2}\|_{4p}^2 ds \\
\leq & C(t-r)^3.
\end{aligned}
\tag{2.4.27}
$$

For the last term, by Minkowski's and Jensen's inequalities, we have

$$\Big\| \int_\theta^t D_\theta^{(k)} \gamma_s^{k_1k_2} ds \Big\|_{2p}^2 \leq (t-\theta) \int_\theta^t \|D_\theta^{(k)} \gamma_s^{k_1k_2}\|_{2p}^2 ds \leq T \int_\theta^t \|D_\theta^{(k)} \gamma_s^{k_1k_2}\|_{2p}^2 ds. \tag{2.4.28}$$

Combining (2.4.23) - (2.4.28), we obtain the following inequality

$$\sum_{i,j=1}^d \|D_\theta^{(k)} \gamma_t^{ij}\|_{2p}^2 \leq c_1 \int_\theta^t \sum_{i,j=1}^d \|D_\theta^{(k)} \gamma_s^{ij}\|_{2p}^2 ds + c_2(t-r)^3, \tag{2.4.29}$$

where c_1, c_2 depends on T, d, $\|h\|_{3,2}^2$, and p. Thus by Gronwall's lemma, we have

$$\sum_{i,j=1}^d \|D_\theta^{(k)} \gamma_t^{ij}\|_{2p}^2 \leq C(t-r)^3. \tag{2.4.30}$$

It follows that

$$\Big\| \|D\gamma_t^{ij}\|_H \Big\|_{2p} \leq C(t-r)^2 \tag{2.4.31}$$

Notice that $\gamma_t \sigma_t = I$, a.s., as a consequence, $D(\gamma_t \sigma_t) = DI \equiv 0$. That implies

$$D\sigma_t^{ij} = -\sum_{i_1,i_2=1}^d \sigma_t^{ii_1} D\gamma_t^{i_1i_2} \sigma_t^{i_2j}. \tag{2.4.32}$$

Then, (2.4.12) follows from (2.4.9), (2.4.11), (2.4.31) and (2.4.32). □

Proof of (2.4.13). Fix $0 \leq r < t \leq T$. For any $\theta_1, \theta_2 \in [r,t]$, let $\theta = \theta_1 \vee \theta_2$. Taking the Malliavin derivative on both sides of (2.4.4), we have the following SDE:

$$\begin{aligned} D_{\theta_1,\theta_2}^{(k_1,k_2)} \xi_t^i = &- \sum_{j_1=1}^d \int_\theta^t D_{\theta_1,\theta_2}^{(k_1,k_2)} \xi_s^{j_1} dM_s^{j_1 i} \\ &+ \sum_{j_1,j_2,j_3=1}^d \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2,j_3} h^{ij_1}(y-\xi_s) D_{\theta_1}^{(k_1)} \xi_s^{j_2} D_{\theta_2}^{(k_2)} \xi_s^{j_3} W^{j_1}(ds,dy). \end{aligned} \tag{2.4.33}$$

Similarly as in (2.4.16), we can show the following inequalities

$$\Big\| \int_\theta^t D^{(k_1,k_2)}_{\theta_1,\theta_2} \xi_s^{j_1} dM_s^{j_1 i} \Big\|_{2p}^2 \le c_{d,p} \|h\|_{3,2}^2 \int_\theta^t \|D^{(k_1,k_2)}_{\theta_1,\theta_2} \xi_s^{j_1}\|_{2p}^2 ds, \tag{2.4.34}$$

and

$$\begin{aligned} &\Big\| \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2,j_3} h^{ij_1}(y-\xi_s) D^{(k_1)}_{\theta_1} \xi_s^{j_2} D^{(k_2)}_{\theta_2} \xi_s^{j_3} W^{j_1}(ds,dy) \Big\|_{2p}^2 \\ &\qquad \le c_p \|h\|_{3,2}^2 \int_\theta^t \|D^{(k_1)}_{\theta_1} \xi_s^{j_2}\|_{4p}^2 \|D^{(k_2)}_{\theta_2} \xi_s^{j_3}\|_{4p}^2 ds \le C(t-r). \end{aligned} \tag{2.4.35}$$

Thus combining (2.4.33) - (2.4.35), we have

$$\sum_{i=1}^d \|D^{(k_1,k_2)}_{\theta_1,\theta_2} \xi_t^i\|_{2p}^2 \le c_1 \sum_{i=1}^d \int_\theta^t \|D^{(k_1,k_2)}_{\theta_1,\theta_2} \xi_s^i\|_{2p}^2 ds + c_2(t-r).$$

Then, it follows from Gronwall's lemma that

$$\sum_{i=1}^d \|D^{(k_1,k_2)}_{\theta_1,\theta_2} \xi_t^i\|_{2p}^2 \le C(t-r). \tag{2.4.36}$$

The inequality (2.4.13) is a consequence of (2.4.36), Jensen's and Minkowski's inequalities. □

Proof of (2.4.14). For any $\theta_1, \theta_2 \in [r,t]$ and $\theta = \theta_1 \vee \theta_2$, by taking the Malliavin derivative on both sides of (2.4.23), we have

$$\begin{aligned} D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma^{ij}_t = &- \sum_{i_1=1}^{d}\Big(\int_\theta^t D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma^{ii_1}_s dM^{i_1 j}_s + \int_\theta^t D^{(k_1)}_{\theta_2}\gamma^{ii_1}_s d\left(D^{(k_2)}_{\theta_1}M^{i_1 j}_s\right)\Big) \\ &- \sum_{i_1=1}^{d}\Big(\int_\theta^t D^{(k_2)}_{\theta_2}\gamma^{ii_1}_s d\left(D^{(k_1)}_{\theta_1}M^{i_1 j}_s\right) + \int_\theta^t \gamma^{ii_1}_s d\left(D^{(k_1,k_2)}_{\theta_1,\theta_2}M^{i_1 j}_s\right)\Big) \\ &- \sum_{i_2=1}^{d}\Big(\int_\theta^t D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma^{i_2 j}_s dM^{i_2 i}_s + \int_\theta^t D^{(k_1)}_{\theta}\gamma^{i_2 j}_s d\left(D^{(k_2)}M^{i_2 i}_s\right)\Big) \\ &- \sum_{i_2=1}^{d}\Big(\int_\theta^t D^{(k_2)}_{\theta_1}\gamma^{i_2 j}_s d\left(D^{(k_1)}_{\theta_2}M^{i_2 i}_s\right) + \int_\theta^t \gamma^{i_2 j}_s d\left(D^{(k_1,k_2)}_{\theta_1,\theta_2}M^{i_2 i}_s\right)\Big) \\ &+ \sum_{i_1,i_2=1}^{d}\Big(Q^{i_1,i}_{i_2,j}\int_\theta^t D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma^{i_1 i_2}_s ds\Big), \end{aligned} \tag{2.4.37}$$

where

$$\begin{aligned} D^{(k_1,k_2)}_{\theta_1,\theta_2}M^{ij}_s = &- \sum_{j_1,j_2,j_3=1}^{d}\int_\theta^s\int_{\mathbb{R}^d}\partial_{i,j_2,j_3}h^{jj_1}(y-\xi_r)\,D^{(k_1)}_{\theta_1}\xi^{j_2}_r D^{(k_2)}_{\theta_2}\xi^{j_3}_r W^{j_1}(dr,dy) \\ &+ \sum_{j_1,j_2=1}^{d}\int_\theta^s\int_{\mathbb{R}^d}\partial_{i,j_2}h^{jj_1}(y-\xi_r)\,D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^{j_2}_r W^{j_1}(dr,dy). \end{aligned}$$

By (2.4.17), (2.4.26), (2.4.30), (2.4.36), Burkholder-Davis-Gundy's, Minkowski's and Hölder's inequalities, we have the following inequalities

$$\Big\|\int_\theta^t D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma^{ii_1}_s dM^{i_1 j}_s\Big\|^2_{2p} \le c_{d,p}\|h\|^2_{3,2}\int_\theta^t \|D^{(k_1,k_2)}_{\theta_1,\theta_2}\gamma^{ii_1}_s\|^2_{2p}ds, \tag{2.4.38}$$

$$\begin{aligned} &\Big\|\int_\theta^t D^{(k_1)}_{\theta_2}\gamma^{ii_1}_s d\left(D^{(k_2)}_{\theta_1}M^{i_1 j}_s\right)\Big\|^2_{2p} \le c_{d,p}\|h\|^2_{3,2}\sum_{i_2=1}^{d}\int_\theta^t\Big\|D^{(k_1)}_{\theta_2}\gamma^{ii_1}_s D^{(k_2)}_{\theta_2}\xi^{i_2}_s\Big\|^2_{2p}ds \\ \le& c_{d,p}\|h\|^2_{3,2}\sum_{i_2=1}^{d}\int_\theta^t\Big\|D^{(k_1)}_{\theta_2}\gamma^{ii_1}_s\Big\|^2_{4p}\Big\|D^{(k_2)}_{\theta_2}\xi^{i_2}_s\Big\|^2_{4p}ds \le C(t-r)^4, \end{aligned} \tag{2.4.39}$$

and

$$
\begin{aligned}
&\Big\| \int_\theta^t \gamma_t^{ii_1} d\left(D_{\theta_1,\theta_2}^{(k_1,k_2)} M_s^{i_1 j}\right) \Big\|_{2p}^2 \\
\le & c_d \Bigg(\sum_{j_1,j_2,j_3=1}^d \Big\| \int_\theta^t \int_{\mathbb{R}^d} \gamma_s^{ii_1} \partial_{i_1,j_2,j_3} h^{jj_1}(y-\xi_r) D_{\theta_1}^{(k_1)} \xi_s^{j_2} D_{\theta_2}^{(k_2)} \xi_s^{j_3} W^{j_1}(ds,dy) \Big\|_{2p}^2 \\
&+ \sum_{j_1,j_2=1}^d \Big\| \int_\theta^t \int_{\mathbb{R}^d} \gamma_s^{ii_1} \partial_{i_1,j_2} h^{jj_1}(y-\xi_s) D_{\theta_1,\theta_2}^{(k_1,k_2)} \xi_s^{j_2} W^{j_1}(ds,dy) \Big\|_{2p}^2 \Bigg) := c_d (I_1 + I_2).
\end{aligned}
$$

We estimate I_1, I_2 as follows:

$$
I_1 \le d \|h\|_{3,2}^2 \sum_{j_2,j_3=1}^d \int_\theta^t \|\gamma_s^{ii_1}\|_{6p}^2 \|D_{\theta_1}^{(k_1)} \xi_s^{j_2}\|_{6p}^2 \|D_{\theta_2}^{(k_2)} \xi_s^{j_3}\|_{6p}^2 ds \le C(t-r)^3,
$$

and

$$
I_2 \le d \|h\|_{3,2}^2 \sum_{j_2=1}^d \int_\theta^t \|\gamma_s^{ii_1}\|_{4p}^2 \|D_{\theta_1,\theta_2}^{(k_1,k_2)} \xi_s^{j_2}\|_{4p}^2 ds \le C(t-r)^4 \le CT(t-r)^3.
$$

Thus we have

$$
\Big\| \int_\theta^t \gamma_t^{ii_1} d\left(D_{\theta_1,\theta_2}^{(k_1,k_2)} M_s^{i_1 j}\right) \Big\|_{2p}^2 \le C(t-r)^3. \tag{2.4.40}
$$

Therefore, combine (2.4.37) - (2.4.40), we have

$$
\sum_{i,j=1}^d \Big\| D_{\theta_1,\theta_2}^{(k_1,k_2)} \gamma_t^{ij} \Big\|_{2p}^2 \le c_1 (t-r)^3 + c_2 \sum_{i,j=1}^d \int_\theta^t \Big\| D_{\theta_1,\theta_2}^{(k_1,k_2)} \gamma_s^{ij} \Big\|_{2p}^2 ds,
$$

By Gronwall's lemma, we have

$$
\sum_{i,j=1}^d \|D_{\theta_1,\theta_2}^{(k_1,k_2)} \gamma_t^{ij}\|_{2p}^2 \le C(t-r)^3, \tag{2.4.41}
$$

which implies

$$\Big\| \|D^2\gamma_t^{jj}\|_{H^{\otimes 2}} \Big\|_{2p} \le C(t-r)^{\frac{5}{2}}.$$

By taking the second Malliavin derivative of $\gamma_t\sigma_t \equiv I$, we have

$$D^2\sigma_t^{ij} = -\sum_{i_1,i_2=1}^{d} \sigma_t^{ii_1}\big(D^2\gamma_t^{i_1i_2}\sigma_t^{i_2j} + D\gamma_t^{i_1i_2}\otimes D\sigma_t^{i_2j} + D\sigma_t^{i_2j}\otimes D\gamma_t^{i_1i_2}\big). \tag{2.4.42}$$

Then, (2.4.14) can be deduced by (2.4.9), (2.4.11), (2.4.12), (2.4.31) and (2.4.42). □

Proof of (2.4.15). For any $\theta_1,\theta_2,\theta_3 \in [r,t]$, let $\theta = \theta_1\vee\theta_2\vee\theta_3$. Taking the Malliavin derivative on both sides of (2.4.33), we have

$$\begin{aligned}
D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3}\xi_t^i = &\sum_{j_1,j_2,j_3=1}^{d}\int_\theta^t\int_{\mathbb{R}^d}\partial_{j_2,j_3}h^{ij_1}(y-\xi_s)D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi_s^{j_2}D^{(k_3)}_{\theta_3}\xi_s^{j_3}W^{j_1}(ds,dy)\\
&-\sum_{j_1,j_2=1}^{d}\int_\theta^t\int_{\mathbb{R}^d}\partial_{j_2}h^{ij_1}(y-\xi_s)D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3}\xi_s^{j_2}W^{j_1}(ds,dy)\\
&-\sum_{j_1,j_2,j_3,j_4=1}^{d}\int_\theta^t\int_{\mathbb{R}^d}\partial_{j_2,j_3,j_4}h^{ij_1}(y-\xi_s)D^{(k_1)}_{\theta_1}\xi_s^{j_2}D^{(k_2)}_{\theta_2}\xi_s^{j_3}D^{(k_3)}_{\theta_3}\xi_s^{j_4}W^{j_1}(ds,dy)\\
&+\sum_{j_1,j_2,j_3=1}^{d}\int_\theta^t\int_{\mathbb{R}^d}\partial_{j_2,j_3}h^{ij_1}(y-\xi_s)D^{(k_1,k_3)}_{\theta_1,\theta_3}\xi_s^{j_2}D^{(k_2)}_{\theta_2}\xi_s^{j_3}W^{j_1}(ds,dy)\\
&+\sum_{j_1,j_2,j_3=1}^{d}\int_\theta^t\int_{\mathbb{R}^d}\partial_{j_2,j_3}h^{ij_1}(y-\xi_s)D^{(k_1)}_{\theta_1}\xi_s^{j_2}D^{(k_2,k_3)}_{\theta_2,\theta_3}\xi_s^{j_3}W^{j_1}(ds,dy).
\end{aligned}\tag{2.4.43}$$

By (2.4.17), (2.4.36), Burkholder-Davis-Gundy's, Minkowski's, and Hölder's inequalities, we have the following inequalities:

$$\begin{aligned}
&\Big\|\int_\theta^t\int_{\mathbb{R}^d}\partial_{j_2,j_3}h^{ij_1}(y-\xi_s)D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi_s^{j_2}D^{(k_3)}_{\theta_3}\xi_s^{j_3}W^{k_1}(ds,dy)\Big\|_{2p}^2\\
\le& c_p\|h\|_{3,2}^2\int_\theta^t\|D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi_s^{j_2}\|_{4p}^2\|D^{(k_3)}_{\theta_3}\xi_s^{j_3}\|_{4p}^2 ds \le C(t-r)^2,
\end{aligned}\tag{2.4.44}$$

$$\Big\| \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2} h^{ij_1}(y-\xi_s) D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3} \xi_s^{j_2} W^{j_1}(ds,dy) \Big\|_{2p}^2$$
$$\leq c_p \|h\|_{3,2}^2 \int_\theta^t \|D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3} \xi_s^{j_2}\|_{2p}^2 ds, \tag{2.4.45}$$

and

$$\Big\| \int_\theta^t \int_{\mathbb{R}^d} \partial_{j_2,j_3,j_4} h^{ij_1}(y-\xi_s) D^{(k_1)}_{\theta_1} \xi_s^{j_2} D^{(k_2)}_{\theta_2} \xi_s^{j_3} D^{(k_3)}_{\theta_3} \xi_s^{j_3} W^{j_1}(ds,dy) \Big\|_{2p}^2$$
$$\leq c_p \|h\|_{3,2}^2 \int_\theta^t \|D^{(k_1)}_{\theta_1} \xi_s^{j_2}\|_{6p}^2 \|D^{(k_2)}_{\theta_2} \xi_s^{j_3}\|_{6p}^2 \|D^{(k_3)}_{\theta_3} \xi_s^{j_3}\|_{6p}^2 ds \leq C(t-r). \tag{2.4.46}$$

Thus combining (2.4.43) - (2.4.46), by Jensen's inequality, we have

$$\sum_{i=1}^d \|D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3} \xi_t^i\|_{2p}^2 \leq c_1 \sum_{i=1}^d \int_\theta^t \|D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3} \xi_t^i\|_{2p}^2 ds + c_2(t-r).$$

Then, the following inequality follows from Gronwall's lemma

$$\sum_{i=1}^d \|D^{(k_1,k_2,k_3)}_{\theta_1,\theta_2,\theta_3} \xi_t^i\|_{2p}^2 \leq C(t-r). \tag{2.4.47}$$

Therefore, (2.4.15) is a consequence of (2.4.47). □

In the next lemma, we derive estimates for the moments of increments of the derivatives of ξ_t and σ_t.

Lemma 2.4.4. *For any $p \geq 1$, $0 \leq r < s < t \leq T$, and $1 \leq i,j \leq d$, there exists a constant $C > 0$ depends on T, d, p, and $\|h\|_{3,2}$, such that*

$$\max_{1\leq i\leq d} \big\| \|D\xi_t^i - D\xi_s^i\|_H \big\|_{2p} \leq C(t-s)^{\frac{1}{2}}, \tag{2.4.48}$$
$$\max_{1\leq i,j\leq d} \Big\| \sigma_t^{ij} - \sigma_s^{ij} \Big\|_{2p} \leq C(t-r)^{-\frac{1}{2}}(s-r)^{-1}(t-s)^{\frac{1}{2}}, \tag{2.4.49}$$
$$\max_{1\leq i,j\leq d} \Big\| \|D\sigma_t^{ij} - D\sigma_s^{ij}\|_H \Big\|_{2p} \leq C(t-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}, \tag{2.4.50}$$
$$\max_{1\leq i\leq d} \big\| \|D\xi_t^i - D^2\xi_s^i\|_{H^{\otimes 2}} \big\|_{2p} \leq C(t-r)(t-s)^{\frac{1}{2}}. \tag{2.4.51}$$

Proof of (2.4.48). **(i)** By (2.4.4), we have

$$D_\theta^{(k)}\xi_t^i - D_\theta^{(k)}\xi_s^i = \delta_{ik}\mathbf{1}_{[s,t]}(\theta) - \sum_{j=1}^d \int_{\theta\vee s}^t D_\theta^{(k)}\xi_u^j dM_u^{ji}.$$

Thus by (2.4.17), Burkholder-Davis-Gundy's, Jensen's, and Minkowski's inequalities, we have

$$\left\|D_\theta^{(k)}\xi_t^i - D_\theta^{(k)}\xi_s^i\right\|_{2p}^2 \le C\left[\delta_{ik}\mathbf{1}_{[s,t]}(\theta) + (t-s)\right].$$

Thus we can show (2.4.48) by Minkowski's inequality:

$$\begin{aligned}\left\|\|D\xi_t^i - D\xi_s^i\|_H\right\|_{2p}^2 &\le \sum_{k=1}^d \int_r^t \left\|D_\theta^{(k)}\xi_t^i - D_\theta^{(k)}\xi_s^i\right\|_{2p}^2 d\theta \\ &\le \sum_{k=1}^d C\Big(\int_s^t \delta_{ik} d\theta + \int_r^t (t-s)d\theta\Big) \le C(t-s).\end{aligned}$$

□

Proof of (2.4.49). Note that $\sigma_t - \sigma_s = \sigma_t(\gamma_s - \gamma_t)\sigma_s$. Then, by (2.4.11) and Hölder's inequality, it suffices to estimate the moment of $\gamma_t - \gamma_s$. By (2.4.22), we have

$$\begin{aligned}\gamma_t^{ij} - \gamma_s^{ij} =& \delta_{ij}(t-s) - \sum_{k_1=1}^d \int_s^t \gamma_u^{jk_1} dM_u^{k_1 j} - \sum_{k_2=1}^d \int_s^t \gamma_u^{jk_2} dM_u^{k_2 i} \\ &+ \sum_{k_1,k_2=1}^d Q_{k_2,j}^{i,k_1} \int_s^t \gamma_u^{k_1 k_2} du.\end{aligned}$$

Then, by (2.4.26), Minkowski's, Jensen's, and Burkholder-Davis-Gundy's inequalities, for all $1 \le i, j \le d$, we have

$$\begin{aligned}\left\|\gamma_t^{ij} - \gamma_s^{ij}\right\|_{2p}^2 \le& C\left((t-s)^2 + (t-r)^2(t-s) + (t-r)^2(t-s)^2\right) \\ \le& C(1+T)^2(t-r)(t-s). \qquad (2.4.52)\end{aligned}$$

Then, (2.4.49) is a consequence of (2.4.11) and (2.4.52). □

Proof of (2.4.50). By (2.4.23), we have the following equation:

$$
\begin{aligned}
D_\theta^{(k)}\gamma_t^{ij} - D_\theta^{(k)}\gamma_s^{ij} = &-\sum_{k_1=1}^{d}\int_{\theta\vee s}^{t} D_\theta^{(k)}\gamma_u^{ik_1} dM_u^{k_1 j} - \sum_{k_1=1}^{d}\int_{\theta\vee s}^{t}\gamma_u^{ik_1} d\left(D_\theta^{(k)}M_u^{k_2 j}\right) \\
&-\sum_{k_2=1}^{d}\int_{\theta\vee s}^{t} D_\theta^{(k)}\gamma_u^{k_2 j} dM_u^{k_2 i} - \sum_{k_2=1}^{d}\int_{\theta\vee s}^{t}\gamma_u^{k_2 j} d\left(D_\theta^{(k)}M_u^{k_2 i}\right) \\
&+\sum_{k_1,k_2=1}^{d}\left(Q_{k_2,j}^{k_1,i}\int_{\theta\vee s}^{t} D_\theta^{(k)}\gamma_u^{k_1 k_2} du\right).
\end{aligned}
$$

Then, by (2.4.17), (2.4.26), and (2.4.30), Burkholder-Davis-Gundy's, Jensen's, Minkowski's, and Cauchy-Schwarz's inequalities, we have

$$
\begin{aligned}
\left\|D_\theta^{(k)}\gamma_t^{ij} - D_\theta^{(k)}\gamma_s^{ij}\right\|_{2p}^2 \le c_{d,p}\|h\|_{3,2}^2 \Bigg[&\sum_{k_1=1}^{d}\int_{\theta\vee s}^{t}\left\|D_\theta^{(k)}\gamma_u^{ik_1}\right\|_{2p}^2 du \\
&+\sum_{k_2=1}^{d}\int_{\theta\vee s}^{t}\left\|\gamma_u^{ik_1}\right\|_{4p}^2\left\|D_\theta^{(k)}\xi_u^{k_2}\right\|_{4p}^2 du + (t-s)\int_{\theta\vee s}^{t}\left\|D_\theta^{(k)}\gamma_u^{k_1 k_2}\right\|_{2p}^2 du\Bigg] \\
\le\, & C(t-r)^2(t-s).
\end{aligned}
$$

This implies

$$
\left\|\left\|D\gamma_t^{ij} - D\gamma_s^{ij}\right\|_H\right\|_{2p} \le C(t-r)^{\frac{3}{2}}(t-s)^{\frac{1}{2}}. \tag{2.4.53}
$$

By (2.4.32), we have

$$
\begin{aligned}
D\sigma_t^{ij} - D\sigma_s^{ij} &= \sum_{i_1,i_2=1}^{d}\left(\sigma_t^{ii_1}D\gamma_t^{i_1 i_2}\sigma_t^{i_2 j} - \sigma_s^{ii_1}D\gamma_s^{i_1 i_2}\sigma_s^{i_2 j}\right) \\
&= \sum_{i_1,i_2=1}^{d}\sigma_t^{ii_1}\left(D\gamma_t^{i_1 i_2} - D\gamma_s^{i_1 i_2}\right)\sigma_t^{i_2 j} + \sum_{i_1,i_2=1}^{d}\left(\sigma_t^{ii_1} - \sigma_s^{ii_1}\right)D\gamma_s^{i_1 i_2}\sigma_t^{i_2 j} \\
&\quad + \sum_{i_1,i_2=1}^{d}\sigma_s^{ii_1}D\gamma_s^{i_1 i_2}\left(\sigma_t^{i_2 j} - \sigma_s^{i_2 j}\right).
\end{aligned}
$$

Thus (2.4.50) follows from (2.4.9), (2.4.11), (2.4.31), (2.4.49), and (2.4.53). □

Proof of (2.4.51). Let $\theta = \theta_1 \vee \theta_2$, by (2.4.33), we have the following equation:

$$D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^i_t - D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^i_s = -\sum_{j_1,j_2=1}^d \int_{\theta\vee s}^t \int_{\mathbb{R}^d} \partial_{j_2} h^{ij_1}(y-\xi_u) D^{(k_1,k_2)}_{\theta_1,\theta_2}\xi^{j_2}_u W^{j_1}(du,dy)$$
$$+ \sum_{j_1,j_2,j_3=1}^d \int_{\theta\vee s}^t \int_{\mathbb{R}^d} \partial_{j_2,j_3} h^{ij_1}(y-\xi_u) D^{(k_1)}_{\theta_1}\xi^{j_2}_u D^{(k_2)}_{\theta_2}\xi^{j_3}_u W^{j_1}(du,dy).$$

As a consequence, by (2.4.17), (2.4.36), Burkholder-Davis-Gundy's, Minkowski's, and Cauchy-Schwarz's inequalities, we have

$$\begin{aligned}\left\|D^{(i,j)}_{\theta_1,\theta_2}\xi^k_t - D^{(i,j)}_{\theta_1,\theta_2}\xi^k_s\right\|^2_{2p} \le & c_p\Big[\sum_{j_1=1}^d \|h\|^2_{3,2}\int_{\theta\vee s}^t \left\|D^{(i,j)}_{\theta_1,\theta_2}\xi^{j_1}_u\right\|^2_{2p} du \\ & + \sum_{j_1,j_2}^d \|h\|^2_{3,2}\int_{\theta\vee s}^t \left\|D^{(i)}_{\theta_1}\xi^{j_1}_u\right\|^2_{4p}\left\|D^{(j)}_{\theta_2}\xi^{j_2}_u\right\|^2_{4p} du\Big] \\ \le & C(t-s) \end{aligned} \tag{2.4.54}$$

Therefore, we obtain (2.4.51) by integrating (2.4.54) and Minkowski's inequality. □

We define the following functionals of ξ_t

$$H_{(i)}(\xi_t,1) = -\sum_{j=1}^d \delta\left(\sigma^{ji}_t D\xi^j_t\right), \quad 1\le i\le d, \tag{2.4.55}$$

and

$$H_{(i,j)}(\xi_t,1) = -\sum_{k=1}^d \sigma\left(H_{(i)}(\xi_t,1)\sigma^{kj}_t D\xi^k_t\right), \quad 1\le i,j\le d. \tag{2.4.56}$$

A more detailed description of these functionals can be seen in Appendix 2.7. In the next lemma, we establish moment estimates for the functionals $H_{(i)}(\xi_t,1)$ and $H_{(i,j)}(\xi_t,1)$.

Lemma 2.4.5. *Suppose that $h \in H_3(\mathbb{R}^d;\mathbb{R}^d \otimes \mathbb{R}^d)$, then the following inequalities are satisfied:*

$$\max_{1\le i\le d} \left\| H_{(i)}(\xi_t,1) \right\|_{2p} \le C(t-r)^{-\frac{1}{2}}, \tag{2.4.57}$$

$$\max_{1\le i,j\le d} \left\| H_{(i,j)}(\xi_t,1) \right\|_{2p} \le C(t-r)^{-1}. \tag{2.4.58}$$

Proof. Due to Meyer's inequality (c.f. Proposition 1.5.4 and 2.1.4 of Nualart [71]), it suffices to estimate

$$\left\| \|\sigma_t^{ji} D\xi_t^j\|_H \right\|_{2p}, \quad \left\| \|D\left(\sigma_t^{ji} D\xi_t^j\right)\|_{H^{\otimes 2}} \right\|_{2p}, \text{ and } \left\| \|D^2\left(\sigma_t^{ji} D\xi_t^j\right)\|_{H^{\otimes 3}} \right\|_{2p}.$$

By (2.4.10) and Lemma 2.4.2 - 2.4.3, we have

$$\left\| \|\sigma_t^{ji} D\xi_t^j\|_H \right\|_{2p} \le \left\| \sigma_t^{ji} \right\|_{4p} \left\| \|D\xi_t^j\|_H \right\|_{4p} \le C(t-r)^{-\frac{1}{2}},$$

$$\begin{aligned} \left\| \|D\left(\sigma_t^{ji} D\xi_t^j\right)\|_{H^{\otimes 2}} \right\|_{2p} \le & \left\| \|D\sigma_t^{ji} \otimes D\xi_t^j\|_{H^{\otimes 2}} \right\|_{2p} + \left\| \|\sigma_t^{ji} D^2\xi_t^j\|_{H^{\otimes 2}} \right\|_{2p} \\ \le & \left\| \|D\sigma_t^{ji}\|_H \right\|_{4p} \left\| \|D\xi_t^j\|_H \right\|_{4p} + \left\| \sigma_t^{ji} \right\|_{4p} \left\| \|D^2\xi_t^j\|_{H^{\otimes 2}} \right\|_{4p} \\ \le & C(t-r)^{\frac{1}{2}}, \end{aligned}$$

and

$$\begin{aligned} \left\| \|D^2\left(\sigma_t^{ji} D\xi_t^j\right)\|_{H^{\otimes 3}} \right\|_{2p} \le & \left\| \|D^2\sigma_t^{ji} \otimes D\xi_t^j\|_{H^{\otimes 2}} \right\|_{2p} \\ & + \left\| \|D\sigma_t^{ji} \otimes D^2\xi_t^j\|_{H^{\otimes 2}} \right\|_{2p} + \left\| \|\sigma_t^{ji} D^3\xi_t^j\|_{H^{\otimes 2}} \right\|_{2p} \\ \le & C(t-r) \end{aligned}$$

The above inequalities hold for all $1 \le i,j \le d$. Then, (2.4.57) and (2.4.58) follows. $\square$

The next lemma provides the moment estimate for the increment of $H_{(i)}(\xi_t,1)$.

Lemma 2.4.6. *Suppose that $h \in H_3(\mathbb{R}^d;\mathbb{R}^d \otimes \mathbb{R}^d)$. Then,*

$$\max_{1\le i\le d} \left\| H_{(i)}(\xi_t,1) - H_{(i)}(\xi_s,1) \right\|_{2p} \le C(s-r)^{-\frac{1}{2}}(t-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}. \tag{2.4.59}$$

Proof. Notice that, by definition, we have

$$\begin{aligned} H_{(i)}(\xi_t,1) - H_{(i)}(\xi_s,1) &= -\sum_{j=1}^{d} \delta\left(\sigma_t^{ji} D\xi_t^j\right) + \sum_{j=1}^{d} \delta\left(\sigma_s^{ji} D\xi_s^j\right) \\ &= -\sum_{j=1}^{d} \delta\left(\sigma_t^{ji} D\xi_t^j - \sigma_s^{ji} D\xi_s^j\right). \end{aligned}$$

Thus by Meyer's inequality again, it suffices to estimate

$$I_1 := \left\| \|\sigma_t^{ji} D\xi_t^j - \sigma_s^{ji} D\xi_s^j\|_H \right\|_{2p} \text{ and } I_2 := \left\| \|D\left(\sigma_t^{ji} D\xi_t^j - \sigma_s^{ji} D\xi_s^j\right)\|_{H^{\otimes 2}} \right\|_{2p}.$$

For I_1, we have

$$I_1 \le \left\| \|\left(\sigma_t^{ji} - \sigma_s^{ji}\right) D\xi_s^k\|_H \right\|_{2p} + \left\| \|\sigma_t^{ji}\left(D\xi_t^j - D\xi_s^j\right)\|_H \right\|_{2p}.$$

Notice that by Lemmas 2.4.2 - 2.4.4, we can write

$$\begin{aligned} \left\| \|\left(\sigma_t^{ji} - \sigma_s^{ji}\right) D\xi_s^j\|_H \right\|_{2p} &\le \left\| \sigma_t^{ji} - \sigma_s^{ji} \right\|_{4p} \left\| \|D\xi_s^j\|_H \right\|_{4p} \\ &\le C(t-r)^{-\frac{1}{2}}(s-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}} \end{aligned}$$

and

$$\begin{aligned} \left\| \|\sigma_t^{ji}\left(D\xi_t^j - D\xi_s^j\right)\|_H \right\|_{2p} &\le \left\| \sigma_t^{ji} \right\|_{4p} \left\| \|D\xi_t^j - D\xi_s^j\|_H \right\|_{4p} \\ &\le C(t-r)^{-1}(t-s)^{\frac{1}{2}} \le C(t-r)^{-\frac{1}{2}}(s-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}. \end{aligned}$$

Thus combining the above inequalities, we have the following estimate for I_1:

$$I_1 \leq C(t-r)^{-\frac{1}{2}}(s-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}. \tag{2.4.60}$$

By Lemmas 2.4.2 - 2.4.4, we have the following estimate for I_2:

$$\begin{aligned}
I_2 \leq & \left\| D\sigma_t^{ji} \otimes D\xi_t^j - D\sigma_t^{ji} \otimes D\xi_s^j \right\|_{2p,H^{\otimes 2}} + \left\| \sigma_t^{ji} D^2 \xi_s^j - \sigma_s^{ji} D^2 \xi_s^j \right\|_{2p,H^{\otimes 2}} \\
\leq & \left\| \|D\sigma_t^{ji}\|_H \right\|_{4p} \left\| \| \left(D\xi_t^j - D\xi_s^j \right) \|_H \right\|_{4p} + \left\| \| \left(D\sigma_t^{ji} - D\sigma_s^{ji} \right) \|_H \right\|_{4p} \left\| \|D\xi_s^j\|_H \right\|_{4p} \\
& + \left\| \sigma_t^{ji} \right\|_{4p} \left\| \|D^2\xi_t^j - D^2\xi_s^j\|_{H^{\otimes 2}} \right\|_{4p} + \left\| \sigma_t^{ji} - \sigma_s^{ji} \right\|_{4p} \left\| \|D^2\xi_s^j\|_{H^{\otimes 2}} \right\|_{2p} \\
\leq & C(t-s)^{\frac{1}{2}}.
\end{aligned} \tag{2.4.61}$$

Therefore, (2.4.59) follows from (2.4.60), (2.4.61) and Meyer's inequality. □

The next lemma shows that ξ is a d-dimensional Gaussian process in the whole probability space. Notice that, however, conditionally on W, the process ξ is no longer Gaussian, because it is the solution to a nonlinear SDE.

Lemma 2.4.7. *The process ξ given by equation (2.4.1) is a d-dimensional Gaussian process, with mean x and covariance matrix*

$$\Sigma_{s,t} = (t \wedge s - r)(I + \rho(0)), \tag{2.4.62}$$

where $\rho(0)$ is defined in (2.1.2). Moreover, the probability density of ξ_t, denoted by $p_{\xi_t}(y)$, is bounded by a Gaussian density:

$$p_{\xi_t}(y) \leq (2\pi(t-r))^{-\frac{d}{2}} \exp\left(-\frac{k|x-y|^2}{t-r}\right), \tag{2.4.63}$$

where

$$k=[2(d\|h\|_{2,3}^2+1)]^{-1}. \tag{2.4.64}$$

Proof. Since B is a d-dimensional Brownian motion and W is a d-dimensional space-time white Gaussian random field independent of B, then $\xi=\{\xi_t, r\le t\le T\}$ is a square integrable d-dimensional martingale. The quadratic covariation of ξ is given by

$$\begin{aligned}\langle \xi^i,\xi^j\rangle_t =&\delta_{ij}(t-r)+\sum_{k=1}^d\int_r^t\int_{\mathbb{R}^d}h^{ik}(\xi_s-y)h^{jk}(\xi_s-y)dyds\\ =&\left(\delta_{ij}+\rho^{ij}(0)\right)(t-r).\end{aligned} \tag{2.4.65}$$

Note that $\rho(0)$ is a symmetric nonnegative definite matrix. As a consequence, $I+\rho(0)$ is strictly positive definite, and thus nondegenerate. Therefore, we can find a nondegenerate matrix M, such that $M^*(I+\rho(0))M=I$. Let $\eta=M\xi$, then $\eta=\{\eta_t,t\in[0,T]\}$ is a martingale with quadratic covariation

$$\langle \eta^i,\eta^j\rangle_t=(t-r)\sum_{k_1,k_2=1}^d M^{ik_1}M^{jk_2}\langle\xi^{k_1},\xi^{k_2}\rangle_t=\delta_{ij}(t-r).$$

By Levy's martingale characterization, η is a d-dimensional Brownian motion. Then, $\xi=M^{-1}\eta$ is a Gaussian process, with covariance matrix (2.4.62).

Since for any $t>r$, $\Sigma_t:=\Sigma_{t,t}=(t-r)(I+\rho(0))$ is symmetric and positive definite, the probability density of the Gaussian random vector ξ_t is given by

$$p_{\xi_t}(y)=\frac{1}{\sqrt{(2\pi)^d|\Sigma_t|}}\exp\Big(-\frac{1}{2}(y-x)^*\Sigma_t^{-1}(y-x)\Big). \tag{2.4.66}$$

Recall that $\rho(0)$ is symmetric and nonnegative definite. Then it has eigenvalues $\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_d\ge 0$. Let λ be the diagonal matrix with diagonal elements $\lambda_1,\dots,\lambda_d$. There is an orthogonal matrix U, such that $\rho(0) = U^*\lambda U$. Let k be defined in (2.4.64). It follows that

$$\lambda_1 + 1 \leq \sum_{i,j=1}^{d} |\rho^{ij}(0)| + 1 \leq \|\rho\|_\infty + 1 \leq d\|h\|_{3,2}^2 + 1 = \frac{1}{2k}.$$

Thus for any nonzero $x \in \mathbb{R}^d$, we have

$$\begin{aligned}\frac{1}{2}x^*\Sigma_t^{-1}x - \frac{k}{t-r}x^*x =& \frac{1}{2}x^*\Big(\Sigma_t^{-1} - \frac{2k}{t-r}I\Big)x \\ =& \frac{1}{2(t-r)}x^*U^*\left((I+\lambda)^{-1} - 2kI\right)Ux \geq 0,\end{aligned}$$

because $(I+\lambda)^{-1} - 2kI$ is a nonnegative diagonal matrix. Thus for any $x, y \in \mathbb{R}^d$, $t > r$, we have

$$\exp\Big(-\frac{1}{2}(y-x)^*\Sigma_t^{-1}(y-x)\Big) \leq \exp\Big(-\frac{k|x-y|^2}{t-r}\Big), \tag{2.4.67}$$

On the other hand, we have

$$|\Sigma_t| = |U^*(I+\lambda)U(t-r)| \geq (t-r)^d. \tag{2.4.68}$$

Therefore, we obtain (2.4.63) by plugging (2.4.67) - (2.4.68) into (2.4.66). □

Denote by $\mathbb{P}^W$, $\mathbb{E}^W$, and $\|\cdot\|_p^W$ the probability, expectation and L^p-norm conditional on W. The following two propositions are estimates for the conditional distribution of ξ.

Proposition 2.4.8. *Fix $0 \leq r < t \leq T$ and recall that $\xi_r = \xi_r^{r,x} = x$. Let $c > 0$, choose $\rho \in (0, c\sqrt{t-r}]$. Then, for any $p_1, p_2 \geq 1$ and $y \in \mathbb{R}^d$, there exist $C > 0$, depending on p_1, p_2, c, $\|h\|_2$, and d, such that*

$$\Big\|\mathbb{P}^W(|\xi_t - y| \leq \rho)^{\frac{1}{p_1}}\Big\|_{p_2} \leq C\exp\Big(-\frac{k|x-y|^2}{p(t-r)}\Big), \tag{2.4.69}$$

where k is defined in (2.4.64) and $p = p_1 \vee p_2$.

Proof. Let $p = p_1 \vee p_2$. Then, by Jensen's inequality, we have

$$\left\| \mathbb{P}^W(|\xi_t - y| \le \rho)^{\frac{1}{p_1}} \right\|_{p_2} = \left\| \|\mathbf{1}_{\{\xi_t - y| \le \rho\}}\|_{p_1}^W \right\|_{p_2} \le \left\| \mathbf{1}_{\{\xi_t - y| \le \rho\}} \right\|_p,$$

We consider two different cases.

(i) Suppose that $2\rho \le |x-y|$. If $|\xi_t - y| \le \rho \le c\sqrt{t-r}$, then

$$|\xi_t - x| \ge |x-y| - |\xi_t - y| \ge |x-y| - \rho \ge \frac{|x-y|}{2},$$

and equivalently $\{|\xi_t - y| < \rho\} \subset \{|\xi_t - x| \ge \frac{|x-y|}{2}\}$. Then, by Lemma 2.4.7, we have

$$\begin{aligned}\left\| \mathbb{P}^W(|\xi_t - y| \le \rho)^{\frac{1}{p_1}} \right\|_{p_2} &= \left\| \mathbf{1}_{\{|\xi_t - x| \ge \frac{|x-y|}{2}\} \cap \{|\xi_t - y| < \rho\}} \right\|_p \le C\Big[V_d \rho^d \sup_{|z-x| \ge \frac{|x-y|}{2}} p_{\xi_t}(z)\Big]^{\frac{1}{p}} \\ &\le C\left[V_d c^d (2\pi)^{-\frac{d}{2}} \exp\Big(-\frac{k|x-y|^2}{t-r}\Big)\right]^{\frac{1}{p}} \end{aligned} \tag{2.4.70}$$

where $V_d = \frac{\pi^{\frac{d}{2}}}{\Gamma(1+\frac{d}{2})}$ is the volume of the unit sphere in $\mathbb{R}^d$.

(ii) On the other hand, suppose that $2\rho > |x-y|$. Then $|x-y| \le 2\rho \le 2c\sqrt{t-r}$. Thus by Lemma 2.4.7 again, we have

$$\begin{aligned}\left\| \mathbb{P}^W(|\xi_t - y| \le \rho)^{\frac{1}{p_1}} \right\|_{p_2} &\le C\big(V_d \rho^d (2\pi(t-r))^{-\frac{d}{2}}\big)^{\frac{1}{p}} \\ &\le C\big(V_d c^d (2\pi)^{-\frac{d}{2}}\big)^{\frac{1}{p}} \exp\Big(\frac{4kc^2}{p} - \frac{4kc^2}{p}\Big) \\ &\le C\big(V_d c^d (2\pi)^{-\frac{d}{2}}\big)^{\frac{1}{p}} e^{\frac{4kc^2}{p}} \exp\Big(-\frac{k|x-y|^2}{p(t-r)}\Big). \end{aligned} \tag{2.4.71}$$

Therefore, (2.4.69) follows from (2.4.70) - (2.4.71). □

Denote by $p^W(r,x;t,y)$ the transition probability density of ξ conditional on W. In other words, $p^W(r,x;t,y)$ is the conditional probability density of $\xi_t = \xi_t^{r,x}$. The existence of $p^W(r,x;t,y)$ is guaranteed by Theorem 2.7.3. By applying Theorem 2.7.4, we can further obtain the following estimate:

Proposition 2.4.9. *For any $0 \le r < t \le T$, $p \ge 1$, and $y \in \mathbb{R}^d$, there exist $C > 0$, depending on T, d, $\|h\|_{3,2}$, p, and q, such that*

$$\left\|p^W(r,x;t,y)\right\|_{2p} \le C\exp\left(-\frac{k|x-y|^2}{6pd(t-r)}\right)(t-r)^{-\frac{d}{2}}, \tag{2.4.72}$$

where k is defined in (2.4.64).

Proof. Choose $p_1 \in (d, 3pd]$, let $p_2 = 2p_1$, and $p_3 = \frac{p_1 p_2}{p_2 - p_1} = p_2$. Then, by (2.7.12) and Hölder's inequality, we have

$$\begin{aligned}\left\|p^W_{\xi_t}(y)\right\|_{2p} \le C \max_{1\le i\le d}\Big\{&\left\|\mathbb{P}^W(|\xi_t - y| < 2\rho)^{\frac{1}{p_2}}\right\|_{6p}\left\|\left(\|H_{(i)}(\xi_t,1)\|^W_{p_1}\right)^{d-1}\right\|_{6p} \\ &\times\left[\frac{1}{\rho} + \left\|\|H_{(i)}(\xi_t,1)\|^W_{p_2}\right\|_{6p}\right]\Big\},\end{aligned} \tag{2.4.73}$$

By Jensen's inequality, we have for any $1 \le i \le d$

$$\left\|\left(\left\|H_{(i)}(\xi_t,1)\right\|^W_{p_1}\right)^{d-1}\right\|_{6p} \le \left\|H_{(i)}(\xi_t,1)\right\|^{d-1}_{6p\vee p_1} \le \left\|H_{(i)}(\xi_t,1)\right\|^{d-1}_{6pd}, \tag{2.4.74}$$

and

$$\left\|\left\|H_{(i)}(\xi_t,1)\right\|^W_{p_2}\right\|_{6p} \le \left\|H_{(i)}(\xi_t,1)\right\|_{6pd}. \tag{2.4.75}$$

Let $\rho = \frac{\sqrt{t-r}}{4}$. (2.4.72) is a consequence of (2.4.73) - (2.4.75), Lemma 2.4.5, and Proposition 2.4.8. $\square$

2.5 A conditional convolution representation

In this section, we follow the idea of Li et al. (see Section 3 of [63]) to obtain a conditional convolution formulation of the SPDE (2.2.1). Consider the following SPDE:

$$u_t(x) = \int_{\mathbb{R}^d} \mu(z) p^W(0,z;t,x)dz + \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x) u_r(z) V(dr,dz), \tag{2.5.1}$$

where W and V are the same random fields as in (2.2.1), p^W is the transition density of ξ_t given by (2.4.1) conditional on W.

In order to define the stochastic integral on the right-hand side of (2.5.1), we introduce the following filtrations. First, for any $t \in [0,T]$, we set

$$\mathscr{F}_t := \sigma\{W(s,x),(s,x) \in [0,T] \times \mathbb{R}^d\} \vee \sigma\{V(s,x),(s,x) \in [0,t] \times \mathbb{R}^d\}. \tag{2.5.2}$$

The stochastic integral in (2.5.1) is defined for all $\mathscr{F}_t$-adapted processes. But later we will see that the solution u, as a limit of Picard iteration, is in fact adapted to a smaller filtration defined as follows: for any $t \in [0,T]$,

$$\mathscr{G}_t := \sigma\{W(s,x),(s,x) \in [0,t] \times \mathbb{R}^d\} \vee \sigma\{V(s,x),(s,x) \in [0,t] \times \mathbb{R}^d\}. \tag{2.5.3}$$

Definition 2.5.1. *A random field $u = \{u_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ is said to be a strong solution to the SPDE (2.5.1), if the following properties are satisfied:*

(i) u is $\mathscr{G}_t$-adapted.

(ii) u is square integrable in the following sense:

$$\mathbb{E}\Big(\int_0^T \int_{\mathbb{R}^d} |u_t(x)|^2 dxdt\Big) < \infty. \tag{2.5.4}$$

(iii) The stochastic integral in (2.5.1) is defined as Walsh's integral and the equality holds almost

surely for all $t \in [0,T]$ and almost every $x \in \mathbb{R}^d$.

Lemma 2.5.2. *Assume that κ and μ are bounded. Then the SPDE (2.5.1) has a unique strong solution (in the sense of Definition 2.5.1). Denote the solution by $u = \{u_t(x), 0 \le t \le T, x \in \mathbb{R}^d\}$. Then, for any $p \ge 1$, the following inequality holds:*

$$\sup_{0 \le t \le T} \sup_{x \in \mathbb{R}^d} \|u_t(x)\|_{2p} < \infty. \tag{2.5.5}$$

Proof. We prove the lemma by the Picard iteration. Let $u_0(t,x) \equiv \mu(x)$ and let

$$u_n(t,x) = \int_{\mathbb{R}^d} \mu(z) p^W(0,z;t,x) dz + \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x) u_{n-1}(r,z) V(dr,dz), \tag{2.5.6}$$

for all $n \ge 1$ and $0 \le t \le T$. Since W and V are independent, then V is a martingale with respect to the filtration $\{\mathscr{F}_t\}_{t \in [0,T]}$. Notice that for any $r \in [0,T]$, $\mathscr{F}_t$ includes all the information of W, and p^W depends only on W. Then, $p^W(r,z;t,x)$ is $\mathscr{F}_r$-measurable, and by induction $u_{n-1}(r,z)$ is $\mathscr{F}_r$-measurable for all $[r,t] \subset [0,T]$ and $x,z \in \mathbb{R}^d$. Thus the stochastic integral is well-defined, and u_n is an $\mathscr{F}_t$-adapted random field. In addition, we know that $p^W(r,z;t,x)$ is $\mathscr{G}_t$-measurable, and by induction we can assume that $u_{n-1}(t)$ is $\mathscr{G}_t$-measurable as well. Thus the stochastic integral in (2.5.6) is $\mathscr{G}_t$-measurable. Therefore, the limit of $u_n(t,x)$ in $L^2(\Omega)$, if exists, is also $\mathscr{G}_t$-measurable.

Let $d_n(t,x) := u_{n+1}(t,x) - u_n(t,x)$. Then

$$d_n(t,x) := \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x) d_{n-1}(r,z) V(dr,dz).$$

For any $p \ge 1$, let

$$d_n^*(t) := \int_{\mathbb{R}^d} \|d_n(t,x)\|_{2p}^2 dx. \tag{2.5.7}$$

We aim to prove the existence and convergence of $\{u^n\}_{n \ge 1}$ in $L^{2p}(\Omega; L^2(\mathbb{R}^d))$ by showing that $\sqrt{d_n^*(t)}$ is summable in n. Then, we will show that the limit is a solution to (2.5.1).

By the definition of $u_n(t)$, Burkholder-Davis-Gundy, Minkowski's and Cauchy-Schwarz's inequalities, we have

$$d_n^*(t) \leq c_p \|\kappa\|_\infty \int_{\mathbb{R}^d} \int_0^t \Big(\int_{\mathbb{R}^d} \big\| p^W(r,z;t,x) d_{n-1}(r,z) \big\|_{2p} dz \Big)^2 dr dx. \tag{2.5.8}$$

By the Markov property, $p^W(r,z;t,x)$ depends only on $\{W(s,z) - W(r,z), s \in (r,t], z \in \mathbb{R}^d\}$. On the other hand, $d_{n-1}(r,z)$ depends on V and $\{W(s,z), s \in [0,r], z \in \mathbb{R}^d\}$. Thus, $p^W(r,z;t,x)$ and $d_{n-1}(r,z)$ are independent. That implies

$$\mathbb{E}\big(|p^W(r,z;t,x) d_{n-1}(r,z)|^{2p}\big) = \mathbb{E}\big(|p^W(r,z;t,x)|^{2p}\big) \mathbb{E}\big(|d_{n-1}(r,z)|^{2p}\big). \tag{2.5.9}$$

Then, by (2.5.8), (2.5.9), Young's convolution inequality, Fubini's theorem and Proposition 2.4.9, we have

$$\begin{aligned} d_n^*(t) \leq & c_p \|\kappa\|_\infty \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} \int_{\mathbb{R}^d} \|p^W(r,z_1;t,x)\|_{2p} \|p^W(r,z_2;t,x)\|_{2p} dx \\ & \qquad \times \|d_{n-1}(r,z_1)\|_{2p} \|d_{n-1}(r,z_2)\|_{2p} dz_1 dz_2 dr \\ \leq & C \int_0^t (t-r)^{-\frac{d}{2}} \exp\Big(-\frac{k|z_1 - z_2|^2}{12pd(t-r)}\Big) \|d_{n-1}(r,z_1)\|_{2p} \|d_{n-1}(r,z_2)\|_{2p} dz_1 dz_2 dr \\ \leq & C \int_0^t d_{n-1}^*(r) dr. \end{aligned} \tag{2.5.10}$$

where $C > 0$ depends on p, T, d, h, and $\|\kappa\|_\infty$.

Thus by iteration, we have

$$d_n^*(t) \leq C^n \int_0^t \int_0^{r_n} \cdots \int_0^{r_2} d_0^*(r_1) dr_1 \cdots dr_n, \tag{2.5.11}$$

To estimate d_0^*, we observe that

$$\begin{aligned}d_0^*(t) =& \int_{\mathbb{R}^d} \Big\| \int_{\mathbb{R}^d} (\mu(z)-\mu(x))\, p^W(0,z;t,x)dz \\ &\qquad + \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x)\mu(z)V(dr,dz)\Big\|_{2p}^2 dx \\ \leq& 3\int_{\mathbb{R}^d} \Big\| \int_{\mathbb{R}^d} \mu(z)p^W(0,z;t,x)dz\Big\|_{2p}^2 dx + 3\int_{\mathbb{R}^d} \Big\| \int_{\mathbb{R}^d} \mu(x)p^W(0,z;t,x)dz\Big\|_{2p}^2 dx \\ &+3\int_{\mathbb{R}^d} \Big\| \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x)\mu(z)V(dr,dz)\Big\|_{2p}^2 dx. \end{aligned} \tag{2.5.12}$$

By an argument similar to the proof of (2.5.10), we can show that $d_0^*(t) < C$. Therefore, we have

$$d_n^*(t) \leq C \int_0^t \int_0^{r_n} \cdots \int_0^{r_2} 1 dr_1 \ldots dr_n = C\frac{t^n}{n!}. \tag{2.5.13}$$

Notice that $\sqrt{d^n(t)}$ is summable in n and the corresponding series is bounded on [0,T]. Therefore, for any fixed $t \in [0,T]$, $\{u_n(t,\cdot)\}_{n\geq 0}$ is convergent in $L^{2p}(\Omega;L^2(\mathbb{R}^d))$. Denote by $u_t(x)$ the limit of this sequence. We claim that $u = \{u_t(x), t\in[0,T], x\in\mathbb{R}^d\}$ is a strong solution to (2.5.1). Clearly u satisfies (2.5.4) and is $\mathscr{G}_t$-adapted. Therefore, it suffices to show that as $n\to\infty$,

$$\int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,\cdot)u_n(r,z)V(dr,dz) \to \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,\cdot)u(r,z)V(dr,dz) \tag{2.5.14}$$

in $L^{2p}(\Omega)$ for all $t\in[0,T]$. Actually, by Burkholder-Davis-Gundy's, Minkowski's, Young's convolution inequalities, and the fact that $\{p^W(r,z;t,x), x,z\in\mathbb{R}^d\}$ and $\{u_n(r,z)-u(r,z), z\in\mathbb{R}^d\}$ are independent, we can write

$$\begin{aligned}&\Big\| \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x)\,(u_n(r,z)-u(r,z))\,V(dr,dz)\Big\|_{2p}^2 \\ \leq& C\int_0^t \int_{\mathbb{R}^d} \|u_n(r,z)-u(r,z)\|_{2p}^2 dzdr,\end{aligned}$$

This implies that (2.5.14) is true. As we discussed before, the limit $u(t,x)$ is $\mathscr{G}_t$-measurable, it follows that $u(t,x)$ is a strong solution to (2.5.1).

In order to show the uniqueness, we assume that $v = \{v_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ is another strong solution to (2.5.1). Let $d_t(x) = u_t(x) - v_t(x)$ for any $t \in [0,T]$ and $x \in \mathbb{R}^d$. Then,

$$d_t(x) = \int_0^t \int_{\mathbb{R}^d} p^W(r,z;t,x) d_r(z) V(dr,dz).$$

By the Ito isometry, Minkowski's and Young's convolution inequalities and the fact that the families $\{d_r(x), x \in \mathbb{R}^d\}$ and $\{p^W(r,z;t,x), x,z \in \mathbb{R}^d\}$ are independent, we have

$$\begin{aligned}\int_{\mathbb{R}^d} \|d_t(x)\|_2^2 dx \leq & \int_0^t \sup_{x\in\mathbb{R}^d} \|d_r(x)\|_2^2 \Big(\int_{\mathbb{R}^d} \big\|p^W(r,z;t,x)\big\|_2 dz \Big)^2 dr \\ \leq & C \int_0^t \int_{\mathbb{R}^d} \|d_r(x)\|_2^2 dx dr. \end{aligned} \tag{2.5.15}$$

Notice that by definition,

$$\int_{\mathbb{R}^d} \|d_t(x)\|_2^2 dx \leq \int_{\mathbb{R}^d} \mathbb{E}|u_t(x)|^2 dx + \int_{\mathbb{R}^d} \mathbb{E}|v_t(x)|^2 dx < \infty$$

for almost every $t \in [0,T]$. As a consequence of Gronwall's lemma and the fact that $d_0 \equiv 0$, the inequality (2.5.15) implies $d(t,x) \equiv 0$, a.s for almost every $(t,x) \in [0,T] \times \mathbb{R}^d$. It follows that the solution to (2.5.1) in the sense of Definition 2.5.1 is unique.

In order to obtain the uniform boundedness (2.5.5), we need to estimate the following expression when applying the Picard iteration:

$$\widetilde{d_n^*}(t) := \sup_{x\in\mathbb{R}^d} \|d_n(t,x)\|_{2p}^2,$$

instead of $d_n^*(t)$ defined in (2.5.7). By a similar argument as we did before, the following inequality can be proved:

$$\widetilde{d_n^*}(t) \leq C \frac{T^n}{n!},$$

where $C > 0$ is independent of n. Then, the inequality (2.5.5) follows immediately. $\square$

Proposition 2.5.3. *Assume that κ and μ are bounded. Let $u = \{u_t(x), 0 < t \le T, x \in \mathbb{R}^d\}$ be the unique strong solution to (2.5.1) in the sense of Definition 2.5.1. Then, u is the strong solution to (2.2.1) in the sense of Definition 2.2.1.*

Proof. Let $u = \{u_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ be the unique solution to the SPDE (2.5.1), and write $Z(dt,dx) = u_t(x)V(dt,dx)$ for all $t \in [0,T]$ and $x \in \mathbb{R}^d$. Then, it suffices to show that u satisfies the following equation:

$$\begin{aligned}\langle u_t, \phi\rangle =& \langle \mu, \phi\rangle + \int_0^t \langle u_s, A\phi\rangle ds + \int_0^t \int_{\mathbb{R}^d} \langle u_s, \nabla\phi^* h(y-\cdot)\rangle W(ds,dy) \\ &+ \int_0^t \int_{\mathbb{R}^d} \phi(x) Z(ds,dx),\end{aligned} \tag{2.5.16}$$

for any $\phi \in C_b^2\left(\mathbb{R}^d\right)$.

Denote by

$$\mathbb{E}^W_{s,x}(\phi(\xi_t)) := \mathbb{E}\big(\phi(\xi_t)|W, \xi_s = x\big) = \int_{\mathbb{R}^d} \phi(z) p^W(s,x;t,z)dz.$$

As u is the strong solution to (2.5.1), the following equations are satisfied

$$\langle u_t, \phi\rangle = \big\langle \mu, \mathbb{E}^W_{0,\cdot}(\phi(\xi_t))\big\rangle + \int_0^t \int_{\mathbb{R}^d} \mathbb{E}^W_{s,z}(\phi(\xi_t)) Z(ds,dz),$$

$$\int_0^t \langle u_s, A\phi\rangle ds = \int_0^t \big\langle \mu, \mathbb{E}^W_{0,\cdot}(A\phi(\xi_s))\big\rangle ds + \int_0^t \int_0^s \int_{\mathbb{R}^d} \mathbb{E}^W_{r,z}(A\phi(\xi_s)) Z(dr,dz)ds,$$

and

$$\begin{aligned}\int_0^t \int_{\mathbb{R}^d} \langle u_s, \nabla\phi^* h(y-\cdot)\rangle W(ds,dy) =& \int_0^t \int_{\mathbb{R}^d} \big\langle \mu, \mathbb{E}^W_{0,\cdot}\big(\nabla\phi(\xi_s)^* h(y-\xi_s)\big)\big\rangle W(ds,dy) \\ &+ \int_0^t \int_{\mathbb{R}^d} \int_0^s \int_{\mathbb{R}^d} \mathbb{E}^W_{r,z}\big((\nabla\phi(\xi_s)^* h(y-\xi_s)\big) Z(dr,dz) W(ds,dy).\end{aligned}$$

Notice that $\phi \in C_b^2(\mathbb{R}^d)$, $h \in H_3(\mathbb{R}^d; \mathbb{R}^d \otimes \mathbb{R}^d)$, and $\|u_t(x)\|_2^2$ is integrable on $[0,T] \times \mathbb{R}^d$. These properties allow us to write

$$\mathbb{E}\Big(\int_0^T\int_{\mathbb{R}^d}|\nabla\phi(\xi_s)^*h(y-\xi_s)|^2dyds\Big)\le T\|\phi\|_{1,\infty}\|h\|_2^2<\infty,$$

$$\begin{aligned}\mathbb{E}&\int_0^T\int_0^T\int_{\mathbb{R}^d\times\mathbb{R}^d}|A\phi(\xi_s)||\kappa(z_1,z_2)u_r(z_1)u_r(z_2)|dz_1dz_2dsdr\\&\le\|\phi\|_{2,\infty}\|\kappa\|_\infty\int_0^T\int_{\mathbb{R}^d}\|u_r(x)\|_2^2dxdr<\infty,\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}&\Big(\int_0^T\int_0^T\int_{\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d}|\nabla\phi(\xi_s)^*h(y-\xi_s)|^2|\kappa(z_1,z_2)u_r(z_1)u_r(z_2)|dydz_1dz_2dsdr\Big)\\&\le\|\phi\|_{1,\infty}\|h\|_2\|\kappa\|_\infty\int_0^T\int_{\mathbb{R}^d}\|u_r(x)\|_2^2dxdr<\infty.\end{aligned}$$

Thus by the stochastic Fubini theorem (c.f. Lemma 4.1 on page 116 of Ikeda & Watanabe [49]), we have

$$\begin{aligned}&\langle u_t,\phi\rangle-\langle\mu,\phi\rangle-\int_0^t\langle u_s,A\phi\rangle ds-\int_0^t\int_{\mathbb{R}^d}\langle u_s,\nabla\phi^*h(y-\cdot)\rangle W(ds,dy) \qquad (2.5.17)\\=&\Big\langle\mu,\mathbb{E}^W_{0,\cdot}\Big(\phi(\xi_t)-\phi(\xi_0)-\int_0^tA\phi(\xi_s)ds-\int_0^t\int_{\mathbb{R}^d}\nabla\phi(\xi_s)^*h(y-\xi_s)W(ds,dy)\Big)\Big\rangle\\&+\int_0^t\int_{\mathbb{R}^d}\mathbb{E}^W_{s,z}\Big(\phi(\xi_t)-\int_s^tA\phi(\xi_r)dr-\int_s^t\int_{\mathbb{R}^d}\nabla\phi(\xi_r)^*h(y-\xi_r)W(dr,dy)\Big)Z(ds,dz).\end{aligned}$$

The last stochastic integral in (2.5.17) is well-defined, because the integrand is $\mathscr{F}_s$-adapted where $\mathscr{F}_s$ is defined in (2.5.2). Notice that by Itô's formula, we have

$$\begin{aligned}\phi(\xi_t^{s,x})=&\phi(x)+\int_s^tA\phi(\xi_r^{s,x})dr+\int_s^r\nabla\phi(\xi_r^{s,x})^*dB_r\\&+\int_s^t\int_{\mathbb{R}^d}\nabla\phi(\xi_r^{s,x})^*h(y-\xi_r^{s,x})W(dr,dy). \qquad (2.5.18)\end{aligned}$$

Then, (2.5.16) follows from (2.5.17) and (2.5.18). □

2.6 Proof of Theorem 2.2.4

In this section, we prove Theorem 2.2.4 by showing the the Hölder continuity of $u_t(x)$ in spatial and time variables separately:

Proposition 2.6.1. *Suppose that $h \in H_3\left(\mathbb{R}^d\right)$, $\|\kappa\|_\infty < \infty$, and $\mu \in L^1\left(\mathbb{R}^d\right)$ is bounded. Then, for any $0 < s < t \leq T$, $x, y \in \mathbb{R}^d$ $\beta \in (0,1)$ and $p > 1$, there exists a constant C depending on T, d, $\|h\|_{3,2}$, $\|\mu\|_\infty$, $\|\kappa\|_\infty$, p, and β, such that the following inequalities are satisfied:*

$$\|u_t(y) - u_t(x)\|_{2p} \leq Ct^{-\frac{1}{2}}(y-x)^\beta, \tag{2.6.1}$$

$$\|u_t(x) - u_s(x)\|_{2p} \leq Cs^{-\frac{1}{2}}(t-s)^{\frac{1}{2}\beta}. \tag{2.6.2}$$

Then, Theorem 2.2.4 is simply a corollary of Proposition 2.6.1. In order to prove Proposition 2.6.1, we need the following Hölder continuity results for the conditional transition density $p^W(r,z;t,x)$:

Lemma 2.6.2. *Suppose that $h \in H_3(\mathbb{R}^d)$, $0 \leq r < s < t \leq T$, $x, y \in \mathbb{R}^d$, and $\beta \in (0,1)$. Then, there exists $C > 0$, depending on T, d, $\|h\|_{3,2}$, p and β, such that the following inequalities are satisfied:*

$$\int_{\mathbb{R}^d} \left\|p^W(r,z;t,y) - p^W(r,z;t,x)\right\|_{2p} dz \leq C(t-r)^{-\frac{1}{2}\beta} |y-x|^\beta, \tag{2.6.3}$$

$$\int_{\mathbb{R}^d} \left\|p^W(r,z;t,x) - p^W(r,z;s,x)\right\|_{2p} dz \leq C(s-r)^{-\frac{1}{2}\beta} (t-s)^{\frac{1}{2}\beta}. \tag{2.6.4}$$

Before showing the proof, let us firstly derive a variant of the density formula (2.7.11). It will be used in the proof of (2.6.4). Choose $\phi \in C_b^2(\mathbb{R}^n)$, such that $\mathbf{1}_{B(0,1)} \leq \phi \leq \mathbf{1}_{B(0,4)}$, and its first and second partial derivatives are all bounded by 1. For any $x \in \mathbb{R}^d$ and $\rho > 0$, we set $\phi_\rho^x := \phi(\frac{\cdot - x}{\rho})$. Assume that F satisfies all the properties in Theorem 2.7.3. Let Q_n be the n-dimensional Poisson kernel (see (2.7.10)). Then, the density of F can be represented as follows:

$$
\begin{aligned}
p_F(x) &= \sum_{i,j_1,j_2=1}^{n} \mathbb{E}\left[\partial_{j_1} Q_n(F-x)\left\langle DF^{j_1}, DF^{j_2}\right\rangle_H \sigma^{j_2 i} H_{(i)}(F, \phi_\rho^x(F))\right] \\
&= \mathbb{E}\left[\left\langle DQ_n(F-x), \sum_{i,j_2=1}^{m} H_{(i)}(F, \phi_\rho^x(F))\sigma^{j_2 i} DF^{j_2}\right\rangle_H\right] \\
&= \sum_{i=1}^{m} \mathbb{E}\left[Q_n(F-x) \sum_{j_2=1}^{m} \delta\left[H_{(i)}(F, \phi_\rho^x(F))\sigma^{j_2 i} DF^{j_2}\right]\right] \\
&= -\sum_{i=1}^{m} \mathbb{E}\left[Q_n(F-x) H_{(i,i)}(F, \phi_\rho^x(F))\right]. \qquad (2.6.5)
\end{aligned}
$$

Let $\xi_t = \xi_t^{r,z}$ be defined in (2.4.1).

Proof of (2.6.3). Choose $p_1 \in (d, 3pd]$, let $p_2 = 2p_1$, and $p_3 = \frac{p_1 p_2}{p_2 - p_1} = p_2$. Then, by (2.7.13) and Hölder's inequality, for any fixed $z, x, y \in \mathbb{R}^d$ and $\rho > 0$, we can show that

$$
\begin{aligned}
I(z) :=& \| p^W(r,z;t,x) - p^W(r,z;t,y) \|_{2p} \\
\le& C|y-x| \left\| \mathbb{P}^W(\xi_t - \tau \le 4\rho)^{\frac{1}{p_2}} \right\|_{6p} \max_{1\le i\le d} \left\{ \left\| \left(\|H_{(i)}(\xi_t;1)\|_{p_2}^W \right)^{d-1} \right\|_{6p} \right. \\
&\times \left. \left(\frac{1}{\rho^2} + \frac{2}{\rho} \left\| \|H_{(i)(\xi_t;1)}\|_{p_2}^W \right\|_{6p} + \left\| \|H_{(i,j)}(\xi_t;1)\|_{p_2}^W \right\|_{6p} \right) \right\},
\end{aligned}
$$

where $\tau = cx + (1-c)y$, for some $c \in (0,1)$ that depends on z, x, y.

Let $\rho = \frac{\sqrt{t-r}}{8}$. Similarly as proved in Proposition 2.4.9, we can show that

$$
\begin{aligned}
I(z) \le& C|y-x|(t-r)^{-\frac{d+1}{2}} \exp\left(-\frac{k|\tau - z|^2}{(6p \vee p_2)(t-r)} \right) \\
\le& C|y-x|(t-r)^{-\frac{d+1}{2}} \exp\left(-\frac{k|\tau - z|^2}{6pd(t-r)} \right), \qquad (2.6.6)
\end{aligned}
$$

where k is defined in (2.4.64) and $C > 0$ depends on T, d, p, and $\|h\|_{3,2}$.

Notice that even if we fix $x, y \in \mathbb{R}^d$, τ is still a function of z that does not have an explicit formulation. Thus it is not easy to calculate the integral of I directly. Without losing generality, assume that $x = 0$, and $y = (y_1, 0, \ldots, 0)$, where $y_1 \ge 0$. Then $\tau = ((1-c)y_1, 0, \ldots, 0)$, where

$c = c(z) \in (0,1)$. Let $\widehat{k} = \frac{k}{6pd}$. For any $z = (z_1, \ldots, z_d) \in \mathbb{R}^d$, we consider the following cases.

(a) If $z_1 \leq 0$, then

$$\exp\Big(-\frac{k|\tau - z|^2}{6pd(t-r)}\Big) \leq \exp\Big(-\frac{\widehat{k}|z|^2}{t-r}\Big). \tag{2.6.7}$$

(b) If $z_1 \geq y_1$, then

$$\exp\Big(-\frac{k|\tau - z|^2}{6pd(t-r)}\Big) \leq \exp\Big(-\frac{\widehat{k}|y-z|^2}{t-r}\Big). \tag{2.6.8}$$

(c) If $0 < z_1 < y_1$, then

$$\exp\Big(-\frac{k|\tau - z|^2}{6pd(t-r)}\Big) \leq \exp\Big(-\frac{\widehat{k}|\tau_0 - z|^2}{t-r}\Big), \tag{2.6.9}$$

where $\tau_0 = (z_1, 0, \ldots, 0)$.

Therefore, combining (2.6.6) - (2.6.9), we have

$$\int_{\mathbb{R}^d} I(z)dz \leq C|y-x|(t-r)^{-\frac{d+1}{2}}\left(I_1 + I_2 + I_3\right), \tag{2.6.10}$$

where

$$I_1 = \int_{-\infty}^{0} dz_1 \int_{\mathbb{R}^{d-1}} \exp\Big(-\frac{\widehat{k}|z|^2}{t-r}\Big) dz_d \ldots dz_2,$$
$$I_2 = \int_{|y|}^{\infty} dz_1 \int_{\mathbb{R}^{d-1}} \exp\Big(-\frac{\widehat{k}|y-z|^2}{t-r}\Big) dz_d \ldots dz_2,$$
$$I_3 = \int_{0}^{|y|} dz_1 \int_{\mathbb{R}^{d-1}} \exp\Big(-\frac{\widehat{k}|\tau_0 - z|^2}{t-r}\Big) dz_d \ldots dz_2.$$

By a changing of variables, it is easy to show that

$$I_1 + I_2 = \int_{\mathbb{R}^d} \exp\Big(-\frac{\widehat{k}|z|^2}{t-r}\Big) dz = \widehat{k}^{-\frac{d}{2}}(t-r)^{\frac{d}{2}}. \tag{2.6.11}$$

For I_3, we compute the integral as follows:

$$\begin{aligned} I_3 =& \int_0^{|y|} dz_1 \int_{\mathbb{R}^{d-1}} \exp\Big(-\frac{\widehat{k}\,(z_2^2+\dots z_d^2)}{t-r}\Big) dz_d \dots dz_2 \\ =& \big(2\pi\widehat{k}^{-1}\big)^{\frac{d-1}{2}} (t-r)^{\frac{d-1}{2}} |y|. \end{aligned} \tag{2.6.12}$$

Thus combining (2.6.10) - (2.6.12), we have

$$\begin{aligned} \int_{\mathbb{R}^d} I(z)dz \leq& C\big[(t-r)^{-\frac{1}{2}}|y| + (t-r)^{-1}|y|^2\big] \\ =& C\big[(t-r)^{-\frac{1}{2}}|y-x| + (t-r)^{-1}|y-x|^2\big]. \end{aligned} \tag{2.6.13}$$

It is easy to see that the inequality (2.6.13) holds for all $x, y \in \mathbb{R}^d$.

On the other hand, by Proposition 2.4.9, we have

$$\int_{\mathbb{R}^d} I(z)dz \leq \int_{\mathbb{R}^d} \|p^W(r,z;t,y)\|_{2p} + \|p^W(r,z;t,x)\|_{2p} dz \leq C. \tag{2.6.14}$$

Therefore by (2.6.13) and (2.6.14), for any $\beta_1, \beta_2 \in (0,1)$, we have

$$\int_{\mathbb{R}^d} I(z)dz \leq C\big[(t-r)^{-\frac{1}{2}\beta_1} |y-x|^{\beta_1} + (t-r)^{-\beta_2} |y-x|^{2\beta_2}\big]$$

Then, (2.6.3) follows by choosing $\beta = \beta_1 = 2\beta_2$. □

Proof of (2.6.4). Let $\rho_1 = \sqrt{t-r}$ and $\rho_2 = \sqrt{s-r}$. By density formula (2.6.5), we have

$$\begin{aligned} &\big|p^W(r,z;t,x) - p^W(r,z;s,x)\big| \\ \leq& \sum_{i=1}^d \Big|\mathbb{E}^W\Big\{[Q_d(\xi_t - x) - Q_d(\xi_s - x)] H_{(i,i)}(\xi_s, \phi^x_{\rho_2}(\xi_s))\Big\}\Big| \\ &+ \sum_{i=1}^d \Big|\mathbb{E}^W\Big\{Q_d(\xi_t - x)\Big[H_{(i,i)}(\xi_t, \phi^x_{\rho_1}(\xi_t)) - H_{(i,i)}(\xi_s, \phi^x_{\rho_2}(\xi_s))\Big]\Big\}\Big| \\ =& I_1 + I_2. \end{aligned} \tag{2.6.15}$$

Estimation for I_1: Note that by the local property of δ (c.f. Proposition 1.3.15 of Nualart [71]), $H_{(i,i)}(\xi_s, \phi^x_{\rho_2}(\xi_s))$ vanishes except if $\xi_s \in B(x, 4\rho_2)$. Choose $p_1 \in (d, 2pd]$. Let $p_2 = 3p_1$ and $p_3 = \frac{3p_1}{3p_1-2}$. Then, $\frac{2}{p_2} + \frac{1}{p_3} = 1$. Thus, by Hölder's inequality, we have

$$\begin{aligned}\|I_1\|_{2p} \leq & d\big\|\|\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|^W_{p_2}\big\|_{6p}\big\|\|Q_d(\xi_t - x) - Q_d(\xi_s - x)\|^W_{p_3}\big\|_{6p} \\ & \times \max_{1\leq i\leq d}\big\|\|H_{(i,i)}(\xi_s, \phi^x_{\rho_2}(\xi_s))\|^W_{p_2}\big\|_{6p}. \end{aligned} \tag{2.6.16}$$

By Proposition 2.4.8, and the fact that $p_2 = 3p_1 \leq 6pd$, the first factor satisfies the following inequality

$$\big\|\|\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|^W_{p_2}\big\|_{6p} = \big\|\mathbb{P}^W(|\xi_s - x| < 4\rho_2)^{\frac{1}{p_2}}\big\|_{6p} \leq C\exp\Big(-\frac{k|z-x|}{6pd(s-r)}\Big). \tag{2.6.17}$$

By Lemmas 2.4.5 and 2.7.2, for all $1 \leq i \leq d$, the last factor can be estimated as follows:

$$\begin{aligned}\big\|\|H_{(i,i)}(\xi_s, \phi^x_{\rho_2}(\xi_s))\|^W_{p_2}\big\|_{6p} \leq & \frac{1}{\rho_2^2} + \frac{2}{\rho_2}\big\|\|H_{(i)}(\xi_s, 1)\|^W_{p_2}\big\|_{6p} + \big\|\|H_{(i,i)}(\xi_s, 1)\|^W_{p_2}\big\|_{6p} \\ \leq & C(s-r)^{-1}. \end{aligned} \tag{2.6.18}$$

We estimate the second factor by the mean value theorem. Let $\eta_1 = |\xi_t - x|$ and $\eta_2 = |\xi_s - x|$. Then, we can write

$$Q_d(\xi_t - x) - Q_d(\xi_s - x) = \begin{cases} A_2^{-1}(\log\eta_1 - \log\eta_2), & \text{if } d = 2, \\ -A_d^{-1}\big[\eta_1^{-(d-2)} - \eta_2^{-(d-2)}\big], & \text{if } d \geq 3. \end{cases}$$

Thus, by the mean value theorem, it follows that

$$|Q_d(\xi_t - x) - Q_d(\xi_s - x)| = \frac{c_d|\eta_1 - \eta_2|}{|\zeta\eta_1 + (1-\zeta)\eta_2|^{d-1}},$$

where c_d is a constant coming from the Poisson kernel, and $\zeta \in (0,1)$ is a random number that depends on η_1 and η_2. Notice that $f(x) = x^{-(d-1)}$ is a convex function on $(0,\infty)$, and $\mathbb{P}(\eta_1 > 0) = \mathbb{P}(\eta_2 > 0) = 1$, then we have

$$|\zeta\eta_1 + (1-\zeta)\eta_2|^{-(d-1)} \leq |\zeta\eta_1|^{-(d-1)} + |(1-\zeta)\eta_2|^{-(d-1)},\ a.s.$$

Let $q = \frac{p_1}{p_1-1}$, then $\frac{1}{q} + \frac{1}{p_2} = \frac{1}{p_3}$. As a consequence of Hölder's inequality, we have

$$\begin{aligned}\big\|\|Q_d(\xi_t - x) - Q_d(\xi_s - x)\|_{p_3}^W\big\|_{6p} &\leq c_d\Big\|\Big\|\frac{|\eta_1-\eta_2|}{|\zeta\eta_1 + (1-\zeta)\eta_2|^{d-1}}\Big\|_{p_3}^W\Big\|_{6p} \\ &\leq C\Big\|\|\eta_1-\eta_2\|_{p_2}^W\|_{12p}\Big\|\ \Big\||\zeta\eta_1 + (1-\zeta)\eta_2|^{-(d-1)}\Big\|_q^W\ \Big\|_{12p} \\ &\leq C\|\eta_1-\eta_2\|_{12pd}\Big[\Big\|\|\zeta\eta_1^{-(d-1)}\|_q^W\Big\|_{12p} + \Big\|\|(1-\zeta)\eta_2^{-(d-1)}\|_q^W\Big\|_{12p}\Big] \\ &\leq C\|\xi_t - \xi_s\|_{12pd}\Big[\Big\|\||\xi_t - y|^{-(d-1)}\|_q^W\Big\|_{12p} + \Big\|\||\xi_s - y|^{-(d-1)}\|_q^W\Big\|_{12p}\Big].\end{aligned} \tag{2.6.19}$$

The negative moments of $\xi_t - y$ can be estimated by (2.4.57), Jensen's inequality, and Lemma 2.7.6:

$$\begin{aligned}\Big\|\||\xi_t - x|^{-(d-1)}\|_q^W\Big\|_{12p} \leq& C\max_{1\leq i\leq d}\big\|\big(\|H_i(\xi_t,1)\|_{p_1}^W\big)^{d-1}\big\|_{12p} \\ \leq& C\max_{1\leq i\leq d}\big\|H_{(i)}(\xi_t,1)\big\|_{12pd}^{d-1} \leq C(t-r)^{-\frac{d-1}{2}}.\end{aligned} \tag{2.6.20}$$

Then, by (2.6.19) - (2.6.20), we have

$$\big\|\|Q_d(\xi_t - x) - Q_d(\xi_s - x)\|_{p_3}^W\big\|_{6p} \leq C(t-s)^{\frac{1}{2}}(s-r)^{-\frac{d-1}{2}}. \tag{2.6.21}$$

Thus combining (2.6.16), (2.6.17), (2.6.18) and (2.6.21), we have

$$\|I_1\|_{2p} \leq C\exp\Big(-\frac{k|z-x|}{6pd(s-r)}\Big)(s-r)^{-\frac{d+1}{2}}(t-s)^{\frac{1}{2}}.$$

This implies

$$\int_{\mathbb{R}^d} \|I_1\|_{2p} dz \le C(s-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}. \tag{2.6.22}$$

Estimates for I_2: Recall that $\gamma_t = (\langle D\xi^i, D\xi^j\rangle_H)_{i,j=1}^d = \sigma_t^{-1}$. By computation analogue to (2.6.5) going backward, we can show that

$$\begin{aligned}
&\mathbb{E}^W\left[Q_d(\xi_t - x)\left(H_{(i,i)}(\xi_t, \phi^x_{\rho_1}(\xi_t)) - H_{(i,i)}(\xi_s, \phi^x_{\rho_2}(\xi_s))\right)\right] \\
=&-\sum_{j_1,j_2=1}^d \mathbb{E}^W\left[\partial_{j_2} Q_d(\xi_t - x)\langle D\xi_t^{j_2}, D\xi_t^{j_1}\rangle_H H_{(i)}\left(\xi_t, \phi^x_{\rho_1}(\xi_t)\right)\sigma_t^{j_1 i}\right] \\
&+\sum_{j_1,j_2=1}^d \mathbb{E}^W\left[\partial_{j_2} Q_d(\xi_t - x)\langle D\xi_t^{j_2}, D\xi_s^{j_1}\rangle_H H_{(i)}\left(\xi_s, \phi^x_{\rho_2}(\xi_s^{r,z})\right)\sigma_s^{j_1 i}\right] \\
=&-\mathbb{E}^W\left[\partial_i Q_d(\xi_t - x)\left(H_{(i)}\left(\xi_t, \phi^x_{\rho_1}(\xi_t)\right) - H_{(i)}\left(\xi_s, \phi^x_{\rho_2}(\xi_s)\right)\right)\right] \\
&+\sum_{j_1,j_2=1}^d \mathbb{E}^W\left[\partial_{j_2} Q_d(\xi_t - x)\langle D\xi_t^{j_2} - D\xi_s^{j_2}, D\xi_s^{j_1}\rangle_H H_{(i)}\left(\xi_s, \phi^x_{\rho_2}(\xi_s)\right)\sigma_s^{j_1 i}\right] \\
:=&J_1 + J_2.
\end{aligned} \tag{2.6.23}$$

By Lemma 2.7.2, we have

$$\begin{aligned}
\left|H_{(i)}\left(\xi_t, \phi^x_{\rho_1}(\xi_t)\right) - H_{(i)}\left(\xi_s, \phi^x_{\rho_2}(\xi_s)\right)\right| &\le \left|\partial_i \phi^x_{\rho_1}(\xi_t) - \partial_i \phi^x_{\rho_2}(\xi_s)\right| \\
&+ |\phi^x_{\rho_2}(\xi_s)|\left|H_{(i)}(\xi_t, 1) - H_{(i)}(\xi_s, 1)\right| + \left|H_{(i)}(\xi_t, 1)\right|\left|\phi^x_{\rho_1}(\xi_t) - \phi^x_{\rho_2}(\xi_s)\right|.
\end{aligned} \tag{2.6.24}$$

By the mean value theorem, for some random numbers $c_1, c_2 \in (0,1)$, we have

$$\begin{aligned}\left|\phi^x_{\rho_1}(\xi_t)-\phi^x_{\rho_2}(\xi_s)\right| =&\left|\mathbf{1}_{B(x,4\rho_1)}(\xi_t)\vee\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\right|\left|\phi\left(\frac{\xi_t-x}{\rho_1}\right)-\phi\left(\frac{\xi_s-x}{\rho_2}\right)\right| \\ =&\left|\mathbf{1}_{B(x,4\rho_1)}(\xi_t)\vee\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\right| \\ &\times\left|\nabla\phi\left(c_1\frac{\xi_t-x}{\rho_1}+(1-c_1)\frac{\xi_s-x}{\rho_2}\right)^*\cdot\left(\frac{\xi_t-x}{\rho_1}-\frac{\xi_s-x}{\rho_2}\right)\right| \\ \leq&\left|\mathbf{1}_{B(x,4\rho_1)}(\xi_t)\vee\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\right|\left|\frac{\xi_t-x}{\rho_1}-\frac{\xi_s-x}{\rho_2}\right|, \end{aligned} \tag{2.6.25}$$

and

$$\left|\partial_i\phi^x_{\rho_1}(\xi_t)-\partial_i\phi^x_{\rho_2}(\xi_s)\right| = \left|\rho_1^{-1}\partial_i\phi\left(\frac{\xi_t-x}{\rho_1}\right)-\rho_2^{-1}\partial_i\phi\left(\frac{\xi_s-x}{\rho_2}\right)\right| \tag{2.6.26}$$

$$\leq \frac{1}{\rho_1}\left|\nabla\partial_i\phi\left(c_2\frac{\xi_t-x}{\rho_1}+(1-c_2)\frac{\xi_s-x}{\rho_2}\right)^*\cdot\left(\frac{\xi_t-x}{\rho_1}-\frac{\xi_s-x}{\rho_2}\right)\right| \tag{2.6.27}$$

$$+\left|\partial_i\phi^x_{\rho_2}(\xi_s)\right|\left|\frac{1}{\rho_1}-\frac{1}{\rho_2}\right|$$

$$\leq \frac{1}{\rho_1}\left(\mathbf{1}_{B(x,4\rho_1)}(\xi_t)\vee\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\right)\left|\frac{\xi_t-x}{\rho_1}-\frac{\xi_s-x}{\rho_2}\right|+\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\left|\frac{1}{\rho_1}-\frac{1}{\rho_2}\right|.$$

Choose $q \in (d, 3pd]$, let $p_1 = \frac{q}{q-1}$, $p_2 = 2q$, $p_3 = 4q$. Then,

$$\frac{1}{p_1}+\frac{2}{p_2}=\frac{1}{p_1}+\frac{1}{p_2}+\frac{2}{p_3}=1.$$

Then, by (2.6.24) - (2.6.26), and Hölder's inequality, we have

$$
\begin{aligned}
\|J_1\|_{2p} \le& \rho_1^{-1} \big\| \|\partial_i Q_d(\xi_t - x)\|_{p_1}^W \big\|_{6p} \Big\| \|\mathbf{1}_{B(x,4\rho_1)}(\xi_t) \vee \mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|_{p_2}^W \Big\|_{6p} \\
&\times \Big\| \Big\| \frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2} \Big\|_{p_2}^W \Big\|_{6p} \\
&+ \big\| \|\partial_i Q_d(\xi_t - x)\|_{p_1}^W \big\|_{6p} \Big\| \|\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|_{p_2}^W \Big\|_{6p} \big\| \|\rho_1^{-1} - \rho_2^{-1}\|_{p_2} \big\|_{6p} \\
&+ \big\| \|\partial_i Q_d(\xi_t - x)\|_{p_1}^W \big\|_{6p} \Big\| \|\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|_{p_2}^W \Big\|_{6p} \big\| \|H_{(i)}(\xi_t, 1) - H_{(i)}(\xi_s, 1)\|_{p_2}^W \big\|_{6p} \\
&+ \big\| \|\partial_i Q_d(\xi_t - x)\|_{p_1}^W \big\|_{6p} \Big\| \|\mathbf{1}_{B(x,4\rho_1)}(\xi_t) \vee \mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|_{p_2}^W \Big\|_{6p} \\
&\times \Big\| \Big\| \frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2} \Big\|_{p_3}^W \Big\|_{12p} \quad \big\| \|H_{(i)}(\xi_t, 1)\|_{p_3}^W \big\|_{12p} \\
:=& L_1 + L_2 + L_3 + L_4. \qquad (2.6.28)
\end{aligned}
$$

In order to estimate the moments of $\frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2}$, we rewrite this random vector in the following way:

$$
\frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2} = \frac{\xi_t - \xi_s}{\rho_1} + (\xi_s - z)\left(\frac{1}{\rho_1} - \frac{1}{\rho_2}\right) + (z - x)\left(\frac{1}{\rho_1} - \frac{1}{\rho_2}\right).
$$

It follows that

$$
\begin{aligned}
\Big\| \frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2} \Big\|_{12p \vee p_3} \le& (t - r)^{-\frac{1}{2}} \|\xi_t - \xi_s\|_{12pd} \\
&+ \frac{(t-r)^{\frac{1}{2}} - (s-r)^{\frac{1}{2}}}{(t-r)^{\frac{1}{2}}(s-r)^{\frac{1}{2}}} \|\xi_s - z\|_{12pd} + |z - x| \frac{(t-r)^{\frac{1}{2}} - (s-r)^{\frac{1}{2}}}{(t-r)^{\frac{1}{2}}(s-r)^{\frac{1}{2}}}.
\end{aligned}
$$

According to Lemma 2.4.7, $\xi_t - \xi_s$ and $\xi_s - z$ are Gaussian random vectors with mean 0, and covariance matrix $(t-s)(I + \rho(0))$ and $(s-r)(I + \rho(0))$ respectively. Therefore, we have

$$
\begin{aligned}
\Big\| \frac{\xi_t - x}{\rho_1} - \frac{\xi_s - x}{\rho_2} \Big\|_{12pd} \le& c_{p,d}(t-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}} + c_{p,d} \frac{(t-r)^{\frac{1}{2}} - (s-r)^{\frac{1}{2}}}{(t-r)^{\frac{1}{2}}(s-r)^{\frac{1}{2}}} (s-r)^{\frac{1}{2}} \\
&+ |z - x| \frac{(t-r)^{\frac{1}{2}} - (s-r)^{\frac{1}{2}}}{(t-r)^{\frac{1}{2}}(s-r)^{\frac{1}{2}}} \\
\le& C\big(|z - x|(s-r)^{-\frac{1}{2}} + 1\big)(t-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}} \qquad (2.6.29)
\end{aligned}
$$

Therefore, by (2.6.29), Proposition 2.4.8 and Lemma 2.7.6, we have

$$L_1 + L_4 \leq C(t-r)^{-\frac{d}{2}} \left[\exp\left(-\frac{k|z-x|^2}{6pd(t-r)}\right) + \exp\left(-\frac{k|z-x|^2}{6pd(s-r)}\right) \right]$$
$$\times \left(1 + |z-x|(s-r)^{-\frac{1}{2}}\right)(t-s)^{\frac{1}{2}}, \tag{2.6.30}$$

and

$$L_2 + L_3 \leq C(t-r)^{-\frac{d}{2}} \exp\left(-\frac{k|z-x|^2}{6pd(s-r)}\right)(s-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}. \tag{2.6.31}$$

Plugging (2.6.30) and (2.6.31) into (2.6.28), we have

$$\int_{\mathbb{R}^d} \|J_1\|_{2p}\, dz \leq C(s-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}. \tag{2.6.32}$$

For J_2, notice that, by definition,

$$\langle D\xi_t^{j_2} - D\xi_s^{j_2}, D\xi_s^{j_1}\rangle_H = \sum_{k=1}^{d} \int_r^s \left(D_\theta^{(k)}\xi_t^{j_2} - D_\theta^{(k)}\xi_s^{j_2}\right) D_\theta^{(k)}\xi_s^{j_1}\, d\theta.$$

By (2.4.2), we have

$$D_\theta^{(k)}\xi_t^{j_2} - D_\theta^{(k)}\xi_s^{j_2} = \mathbf{1}_{[s,t]}(\theta)\delta_{j_2 k} - \sum_{i=1}^{d} \mathbf{1}_{[r,t]}(\theta) \int_s^t D_\theta^{(k)}\xi_r^i\, dM_r^{ij_2}.$$

By a argument similar to the one used in the proof of Lemma 2.4.3, we can show that

$$\left\|\mathbf{1}_{[r,s]}(\theta)\left(D_\theta^{(k)}\xi_t^{j_2} - D_\theta^{(k)}\xi_s^{j_2}\right)\right\|_{2p}^2 \leq C\mathbf{1}_{[r,s]}(\theta)(t-s).$$

Therefore, by Hölder's and Minkowski's inequalities, we have

$$\begin{aligned}\left\|\langle D\xi_t^{j_2}-D\xi_s^{j_2},D\xi_s^{j_1}\rangle_H\right\|_{2p}\le&\sum_{k=1}^{d}\int_r^s\|\mathbf{1}_{[r,s]}(\theta)\big(D_\theta^{(k)}\xi_t^{j_2}-D_\theta^{(k)}\xi_s^{j_2}\big)\|_{4p}\|D_\theta^{(k)}\xi_s^{j_1}\|_{4p}d\theta\\ \le&C(s-r)(t-s)^{\frac{1}{2}}.\end{aligned}\tag{2.6.33}$$

Choose $q\in(d,3pd]$. Let $p_1=\frac{q}{q-1}$, $p_2=2q$ and $p_3=6q$. Then $\frac{1}{p_1}+\frac{1}{p_2}+\frac{3}{p_3}=1$. Thus, by (2.6.33), Hölder's inequality, Lemmas 2.4.3, 2.4.5, 2.7.6, and Proposition 2.4.8, we have

$$\begin{aligned}\|J_2\|_{2p}\le&\sum_{j_1,j_2=1}^{d}\big\|\|\mathbf{1}_{B(x,4\rho_2)}(\xi_s)\|_{p_2}^W\big\|_{6p}\big\|\|\partial_{j_2}Q_d(\xi_t-x)\|_{p_1}^W\big\|_{6p}\\ &\times\big\|\|\langle D\xi_t^{j_2}-D\xi_s^{j_2},D\xi_s^{j_1}\rangle_H\|_{p_3}^W\big\|_{18p}\big\|\|H_{(i)}(\xi_s,\phi_{\rho_2}^y(\xi_s))\|_{p_3}^W\big\|_{18p}\big\|\|\sigma_s^{j_1i}\|_{p_3}^W\big\|_{18p}\\ \le&C\exp\Big(-\frac{k|z-x|^2}{6pd(s-r)}\Big)(t-r)^{-\frac{d-1}{2}}(t-s)^{\frac{1}{2}}(s-r)^{-\frac{1}{2}}.\end{aligned}$$

As a consequence, we have

$$\int_{\mathbb{R}^d}\|J_2\|_{2p}dz\le C(t-s)^{\frac{1}{2}}.\tag{2.6.34}$$

Finally, combining (2.6.22), (2.6.32) and (2.6.34), we have

$$\int_{\mathbb{R}^d}\left\|p^W(r,z;t,x)-p^W(r,z;s,x)\right\|_{2p}dz\le C(s-r)^{-\frac{1}{2}}(t-s)^{\frac{1}{2}}.\tag{2.6.35}$$

On the other hand, by (2.4.72), we have

$$\int_{\mathbb{R}^d}\|I_2\|_{2p}dz\le\int_{\mathbb{R}^d}\|p^W(r,z;t,y)\|_{2p}+\|p^W(r,z;s,y)\|_{2p}\le C.\tag{2.6.36}$$

Thus (2.6.4) follows from (2.6.35) and (2.6.36). □

Proof of Proposition 2.6.1. By the convolution representation (2.5.1), Burkholder-Davis-Gundy's, and Minkowski's inequalities, we have

$$
\begin{aligned}
\|u_t(y)-u_t(x)\|_{2p} \le & \Big\| \int_{\mathbb{R}^d} \mu(z)\left(p^W(0,z;t,y)-p^W(0,z;t,x)\right)dz\Big\|_{2p} \\
& + \Big\| \int_0^t \int_{\mathbb{R}^d} u_r(z)\left(p^W(r,z;t,y)-p^W(r,z;t,x)\right)V(dz,dr)\Big\|_{2p} \\
\le & \|\mu\|_\infty \int_{\mathbb{R}^d} \left\|p^W(0,z;t,y)-p^W(0,z;t,x)\right\|_{2p} dz \\
& + \|\kappa\|_\infty^{\frac{1}{2}} \left(\int_0^t \left(\int_{\mathbb{R}^d} \left\|u_r(z)\left(p^W(r,z;t,y)-p^W(r,z;t,x)\right)\right\|_{2p} dz\right)^2 dr\right)^{\frac{1}{2}} \\
:= & I_1 + \|\kappa\|_\infty^{\frac{1}{2}} I_2. \qquad (2.6.37)
\end{aligned}
$$

Note that I_1 can be estimated by Lemma 2.6.2. For I_2, recall that $u(r,z)$ is independent of $p^W(r,z;t,y)$[2]. Then, by Lemma 2.5.2 and 2.6.2, we have

$$
\begin{aligned}
I_2 \le & \left(\int_0^t \sup_{z\in\mathbb{R}^d} \|u_r(z)\|_{2p}^2 \left(\int_{\mathbb{R}^d} \left\|p^W(r,z;t,y)-p^W(r,z;t,x)\right\|_{2p} dz\right)^2 dr\right)^{\frac{1}{2}} \\
\le & C|y-x|^\beta \left(\int_0^t (t-r)^{-\beta} dr\right)^{\frac{1}{2}} \le \frac{Ct^{\frac{1-\beta}{2}}}{\sqrt{1-\beta}} |y-x|^\beta. \qquad (2.6.38)
\end{aligned}
$$

Therefore (2.6.1) follows from (2.6.3), (2.6.37) and (2.6.38).

The proof of (2.6.2) is quite similar. As in (2.6.37), we can show that

$$
\begin{aligned}
\|u_t(x)-u_s(x)\|_{2p} \le & \|\mu\|_\infty \int_{\mathbb{R}^d} \left\|p^W(0,z;t,x)-p^W(0,z;s,x)\right\|_{2p} dz \\
& + C\|\kappa\|_\infty^{\frac{1}{2}} \left[\int_s^t \sup_{z\in\mathbb{R}^d} \|u_r(z)\|_{2p}^2 \left(\int_{\mathbb{R}^d} \left\|p^W(r,z;t,x)\right\|_{2p} dz\right)^2 dr\right]^{\frac{1}{2}} \\
& + C\|\kappa\|_\infty^{\frac{1}{2}} \left[\int_0^s \sup_{z\in\mathbb{R}^d} \|u_r(z)\|_{2p}^2 \left(\int_{\mathbb{R}^d} \left\|\left(p^W(r,z;t,x)-p^W(r,z;s,x)\right)\right\|_{2p} dz\right)^2 dr\right]^{\frac{1}{2}}.
\end{aligned}
$$

Then, the estimate (2.6.2) follows from (2.6.4), Proposition 2.4.9 and Lemma 2.5.2. □

[2] The same idea has been used in the proof of Lemma 2.5.2.

2.7 Appendix: a brief introduction on Malliavin calculus

In this section, we present some preliminaries on the Malliavin calculus. We refer the readers to book of Nualart [71] for a detailed account on this topic.

Fix a time interval $[0,T]$. Let $B=\{B_t^1,\dots,B_t^d, 0\le t\le T\}$ be a standard d-dimensional Brownian motion on $[0,T]$. Denote by $\mathscr{S}$ the class of smooth random variables of the form

$$G=g\left(B_{t_1},\dots,B_{t_m}\right)=g\left(B_{t_1}^1,\dots,B_{t_1}^d,\dots,B_{t_m}^1,\dots,B_{t_m}^d\right), \tag{2.7.1}$$

where m is any positive integer, $0\le t_1<\cdots<t_m\le T$, and $g:\mathbb{R}^{md}\to\mathbb{R}$ is a smooth function that has all partial derivatives with at most polynomial growth. We make use of the notation $x=\left(x_i^k\right)_{1\le i\le m, 1\le k\le d}$ for any element $x\in\mathbb{R}^{md}$. The basic Hilbert space associated with B is $H=L^2\left([0,T];\mathbb{R}^d\right)$.

Definition 2.7.1. *For any $G\in\mathscr{S}$ given by (2.7.1), the Malliavin derivative, is the H-valued random variable DG given by*

$$D_\theta^{(k)}G=\sum_{i=1}^m\frac{\partial g}{\partial x_i^k}\left(B_{t_1},\dots,B_{t_m}\right)\mathbf{1}_{[0,t_i]}(\theta),\quad 1\le k\le d,\ \theta\in[0,T].$$

In the same way, for any $n\ge1$, the iterated derivative D^nG of a random variable of the form (2.7.1) is a random variable with values in $H^{\otimes n}=L^2\left([0,T]^n;\mathbb{R}^{d^n}\right)$. For each $p\ge1$, the iterated derivative D^n is a closable and unbounded operator on $L^p(\Omega)$ taking values in $L^p(\Omega;H^{\otimes n})$. For any $n\ge1$, $p\ge1$ and any Hilbert space V, we can introduce the Sobolev space $\mathbb{D}^{n,p}(V)$ of V-valued random variables as the closure of $\mathscr{S}$ with respect to the norm

$$\begin{aligned}\|G\|_{n,p,V}^2 &=\|G\|_{L^p(\Omega;V)}^2+\sum_{k=1}^n\|D^kG\|_{L^p(\Omega;H^{\otimes k}\otimes V)}^2\\ &=\left[\mathbb{E}\left(\|G\|_V^p\right)\right]^{\frac{2}{p}}+\sum_{k=1}^n\left[\mathbb{E}\left(\|D^kG\|_{H^{\otimes k}\otimes V}^p\right)\right]^{\frac{2}{p}}.\end{aligned}$$

By definition, the divergence operator δ is the adjoint operator of D in $L^2(\Omega)$. More precisely, δ

is an unbounded operator on $L^2(\Omega;H)$, taking values in $L^2(\Omega)$. We denote by $\text{Dom}(\delta)$ the domain of δ. Then, for any $u = (u^1, \dots, u^d) \in \text{Dom}(\delta)$, $\delta(u)$ is characterized by the duality relationship: for all for all $G \in \mathbb{D}^{1,2} = \mathbb{D}^{1,2}(\mathbb{R})$.

$$\mathbb{E}(\delta(u)G) = \mathbb{E}(\langle DG, u\rangle_H). \tag{2.7.2}$$

Let F be an n-dimensional random vector, with components $F^i \in \mathbb{D}^{1,1}, 1 \le i \le n$. We associate to F an $n \times n$ random symmetric nonnegative definite matrix, called the Malliavin matrix of F, denoted by γ_F. The entries of γ_F are defined by

$$\gamma_F^{ij} = \left\langle DF^i, DF^j \right\rangle_H = \sum_{k=1}^{d} \int_0^T D_\theta^{(k)} F^i D_\theta^{(k)} F^j d\theta. \tag{2.7.3}$$

Suppose that $F \in \cap_{p\ge1}\mathbb{D}^{2,p}(\mathbb{R}^n)$, and its Malliavin matrix γ_F is invertible. Denote by σ_F the inverse of γ_F. Assume that $\sigma_F^{ij} \in \cap_{p\ge1}\mathbb{D}^{1,p}$ for all $1 \le i, j \le n$. Let $G \in \cap_{p\ge1}\mathbb{D}^{1,2}$. Then $G\sigma_F^{ij}DF^k \in \text{Dom}(\delta)$ for all $1 \le i, j, k \le n$. Under the hypotheses, we define

$$H_{(i)}(F,G) = -\sum_{j=1}^{n} \delta\left(G\sigma_F^{ji}DF^j\right), \quad 1 \le i \le n. \tag{2.7.4}$$

If furthermore $H_{(i)}(F,G) \in \cap_{p\ge1}\mathbb{D}^{1,p}$ for all $1 \le i \le n$, then we define

$$H_{(i,j)}(F,G) = H_{(j)}\left(F, H_{(i)}(F,G)\right), \quad 1 \le i, j \le n. \tag{2.7.5}$$

The following lemma is a Wiener functional version of Lemma 9 of Bally & Caramellino [5].

Lemma 2.7.2. *Suppose that $F \in \cap_{p\ge1}\mathbb{D}^{2,p}(\mathbb{R}^n)$, $(\gamma_F^{-1})^{ij} = \sigma_F^{ij} \in \cap_{p\ge1}\mathbb{D}^{2,p}$ for all $1 \le i, j \le n$, and $\phi \in C_b^1(\mathbb{R}^n)$. Then, for any $1 \le i \le n$, we have*

$$H_{(i)}(F, \phi(F)) = \partial_i\phi(F) + \phi(F)H_{(i)}(F,1). \tag{2.7.6}$$

Suppose that $F \in \cap_{p\geq 1}\mathbb{D}^{3,p}(\mathbb{R}^n)$ and $\phi \in C_b^2(\mathbb{R}^n)$. Then, for any $1 \leq i,j \leq n$, we have

$$\begin{aligned} H_{(i,j)}(F,\phi(F)) = \partial_{ij}\phi(F) + \partial_i\phi(F)H_{(j)}(F,1) \\ + \partial_j\phi(F)H_{(i)}(F,1) + \phi(F)H_{(i,j)}(F,1). \end{aligned} \tag{2.7.7}$$

Proof. For any $F \in \cap_{p\geq 1}\mathbb{D}^{2,p}(\mathbb{R}^n)$ and $\phi \in C_b^1(\mathbb{R}^n)$, it is easy to check that $\phi(F) \in \cap_{p\geq 1}\mathbb{D}^{1,p}$. Then, $H_{(i)}(F,\phi(F))$ is well defined. For any $G \in \mathbb{D}^{1,2}$, by the duality of D and δ, we have

$$\begin{aligned} \mathbb{E}\left(H_{(i)}(F,\phi(F))G\right) &= -\sum_{j=1}^{n}\mathbb{E}\left(\delta\left(\phi(F)\sigma_F^{ji}DF^j\right)G\right) \\ &= -\sum_{j=1}^{n}\mathbb{E}\left(\phi(F)\sigma_F^{ji}\left\langle DF^j, DG\right\rangle_H\right). \end{aligned} \tag{2.7.8}$$

On the other hand, by the product rule for the operator D, we have

$$\begin{aligned} \mathbb{E}\left(\phi(F)H_{(i)}(F,1)G\right) &= -\sum_{j=1}^{m}\mathbb{E}\left(\left\langle \sigma_F^{ji}DF^j, D(\phi(F)G)\right\rangle_H\right) \\ &= -\sum_{j=1}^{m}\mathbb{E}\left(\phi(F)\sigma_F^{ji}\left\langle DF^j, DG\right\rangle_H\right) - \sum_{j_1,j_2=1}^{m}\mathbb{E}\left(G\partial_{j_2}\phi(F)\sigma_F^{j_1 i}\left\langle DF^{j_1}, DF^{j_2}\right\rangle_H\right). \end{aligned}$$

Note that σ_F is the inverse of $\gamma_F = \left(\langle DF^i, DF^j\rangle_H\right)_{i,j=1}^n$, then

$$\sum_{j_1,j_2=1}^{m}\mathbb{E}\left(G\partial_{j_2}\phi(F)\sigma_F^{j_1 i}\left\langle DF^{j_1}, DF^{j_2}\right\rangle_H\right) = \mathbb{E}\left(G\partial_i\phi(F)\right). \tag{2.7.9}$$

Then, (2.7.6) follows from (2.7.8) - (2.7.9). The equality (2.7.7) can be proved similarly. □

The next theorem is a density formula using the Riesz transformation. The formula was first introduced by Malliavin and Thalmaier (see Theorem Section 2.3.23 of Malliavin & Thalmaier [68]), then further studied by Bally & Caramellino [5].

For any integer $n \geq 2$, let Q_n be the n-dimensional Poisson kernel. That is,

$$Q_n(x) = \begin{cases} A_2^{-1} \log |x|, & n = 1, \\ -A_n^{-1} |x|^{2-n}, & n > 2, \end{cases} \tag{2.7.10}$$

where A_n is the area of the unit sphere in $\mathbb{R}^n$. Then, $\partial_i Q_n(x) = c_n x_i |x|^{-n}$, where $c_2 = A_2^{-1}$ and $c_n = (\frac{n}{2} - 1) A_n^{-1}$ for $n > 2$.

The theorem below is the density formula for a class of differentiable random variables.

Theorem 2.7.3. *(Proposition 10 of Bally & Caramellino [5]) Let $F \in \cap_{p \geq 1} \mathbb{D}^{2,p}(\mathbb{R}^n)$. Assume that $(\gamma_F^{-1})^{ij} = \sigma_F^{ij} \in \cap_{p \geq 1} \mathbb{D}^{1,p}$ for all $1 \leq i, j \leq n$. Then, the law of F has a density p_F.*

More precisely, for any $x \in \mathbb{R}^n$ and $r > 0$, let $B(x,r)$ be the sphere on $\mathbb{R}^n$ centered at x with radius r. Suppose that $\phi \in C_b^1(\mathbb{R}^d)$, such that $\mathbf{1}_{B(0,1)} \leq \phi \leq \mathbf{1}_{B(0,2)}$, and $|\nabla \phi| \leq 1$. Define $\phi_\rho^x := \phi(\frac{\cdot - x}{\rho})$ for any $\rho > 0$ and $x \in \mathbb{R}^n$. Then,

$$\begin{aligned} p_F(x) &= \sum_{i=1}^n \mathbb{E}\big(\partial_i Q_n(F - x) H_{(i)}(F, 1)\big) \\ &= \sum_{i=1}^n \mathbb{E}\big(\partial_i Q_n(F - x) H_{(i)}(F, \phi_\rho^x(F))\big) \\ &= \sum_{i=1}^n \mathbb{E}\big(\mathbf{1}_{B_{(x,2\rho)}}(F) \partial_i Q_n(F - x) H_{(i)}(F, \phi_\rho^x(F))\big). \end{aligned} \tag{2.7.11}$$

The next theorem provides the estimates for the density and its increment.

Theorem 2.7.4. *Suppose that F satisfies the conditions in Theorem 2.7.3. Then, for any $p_2 > p_1 > n$, let $p_3 = \frac{p_1 p_2}{p_2 - p_1}$, there exists a constant C that depends on p_1, p_2 and n, such that*

$$p_F(x) \leq C \mathbb{P}(|F - x| < 2\rho)^{\frac{1}{p_3}} \max_{1 \leq i \leq n} \left[\left\| H_{(i)}(F, 1) \right\|_{p_1}^{n-1} \left(\frac{1}{\rho} + \left\| H_{(i)}(F, 1) \right\|_{p_2} \right) \right]. \tag{2.7.12}$$

If furthermore, $F \in \cap_{p \geq 1} \mathbb{D}^{3,p}(\mathbb{R}^n)$, for any $x_1, x_2 \in \mathbb{R}^n$, we can find $y = c x_1 + (1 - c) x_2$ for some $c \in (0,1)$ that depends on x_1, x_2. Then, there exist a constant F the constant C that depends on p_1,

p_2, and m, such that

$$
\begin{aligned}
|p_F(x_1) - p_F(x_2)| \le & C|x_1 - x_2| \mathbb{P}(|F-y| < 4\rho)^{\frac{1}{p_3}} \\
& \times \max_{1 \le i,j \le n} \left[\left\| H_{(i)}(F,1) \right\|_{p_1}^{n-1} \left(\frac{1}{\rho^2} + \frac{2}{\rho} \left\| H_{(i)}(F,1) \right\|_{p_2} + \left\| H_{(i,j)}(F,1) \right\|_{q_2} \right) \right].
\end{aligned} \tag{2.7.13}
$$

Remark 2.7.5. *The inequalities stated in Theorem 2.7.4 are an improved version of those estimates by Bally and Caramillino (see Theorem 8 of Bally & Caramellino [5]). We refer to (see Lemma 7.3.2 of Nualart & Nualart [72]) for a related result. For the sake of completeness, we present below a proof of Theorem 2.7.4. The proof follows the same idea as in Theorem 8 of Bally & Caramellino [5]. The only difference occurs when choosing the radius of the ball in the estimate for the Poisson kernel. If we optimize the radius, then the exponent of $\|H_{(i)}(F,1)\|_p$ is $n-1$, instead of $\frac{q_1(n-1)}{q_1-n} > n-1$ in Bally & Caramellino [5].*

In order to prove Theorem 2.7.4, we first give the estimate for the Poisson kernel:

Lemma 2.7.6. *Suppose that F satisfy the conditions in Theorem 2.7.3. For any $p > n$, let $q = \frac{p}{p-1}$. Then, there exists a constant $C > 0$ depends on m and p, such that*

$$
\sup_{x \in \mathbb{R}^n} \|\partial_i Q_n(F-x)\|_q \le \sup_{x \in \mathbb{R}^n} \left\| |F-x|^{-(n-1)} \right\|_q \le C \max_{1 \le i \le n} \left\| H_{(i)}(F,1) \right\|_p^{n-1}. \tag{2.7.14}
$$

Proof. Assume that

$$
\|p_F\|_\infty := \sup_{x \in \mathbb{R}^d} p_F(x) < \infty.
$$

Denote by $M = \sup_{1 \le i \le n} \|H_{(i)}(F,1)\|_p$. Then by Hölder's inequality, for all $x \in \mathbb{R}^d$, we have

$$
\begin{aligned}
p_F(x) = & \sum_{i=1}^{n} \mathbb{E}\big(\partial_i Q_n(F-x) H_{(i)}(F,1)\big) \le \sum_{i=1}^{m} \|\partial_i Q_n(F-x)\|_q \|H_{(i)}(F,1)\|_p \\
\le & n \sup_{x \in \mathbb{R}^n} \left\| |F-x|^{-(n-1)} \right\|_q M,
\end{aligned}
$$

which implies

$$\|p_F\|_\infty \leq n \sup_{x\in\mathbb{R}^n} \left\| |F-x|^{-(n-1)} \right\|_q M. \tag{2.7.15}$$

In order to estimate $\||F-x|^{-(n-1)}\|_q$, choose any $\rho > 0$. Then for all $x \in \mathbb{R}^n$,

$$\begin{aligned}
\mathbb{E}(|F-x|^{-(n-1)q}) =& \int_{\mathbb{R}^d} |y-x|^{-(n-1)q} p_F(y)dy \\
=& \int_{|y-x|\leq\rho} |y-x|^{-(n-1)q} p_F(y)dy + \int_{|y-x|>\rho} |y-x|^{-(n-1)q} p_F(y)dy \\
\leq& \|p_F\|_\infty \int_0^\rho r^{-(n-1)q} r^{n-1} dr + \rho^{-(n-1)q} \\
=& k_{n,q} \|p_F\|_\infty \rho^{1-(n-1)(q-1)} + \rho^{-(n-1)q},
\end{aligned} \tag{2.7.16}$$

where $k_{n,q} = [1-(n-1)(q-1)]^{-1}$. The last equality is due to the fact that $1-(n-1)(q-1) > 0$.

Combining (2.7.15) and (2.7.16), we have

$$\|p_F\|_\infty \leq \left[n k_{n,q}^{\frac{1}{q}} \|p_F\|_\infty^{\frac{1}{q}} \rho^{\frac{1-(n-1)(q-1)}{q}} + \rho^{-(n-1)} \right] M. \tag{2.7.17}$$

By optimizing the right-hand side of (2.7.17), we choose

$$\rho = \rho^* := \left[\frac{(n-1)q}{n} \right]^{\frac{q}{n}} \|p_F\|_\infty^{-\frac{1}{n}}.$$

Plugging ρ^* into (2.7.17), we obtain

$$\|p_F\|_\infty \leq \left(n k_{n,q}^{\frac{1}{q}} \left[\frac{(n-1)q}{n} \right]^{\frac{1-(n-1)(q-1)}{n}} + \left[\frac{(n-1)qM}{n} \right]^{-\frac{q(n-1)}{n}} \right) M |p_F\|_\infty^{\frac{n-1}{n}}.$$

Then, it follows that

$$\|p_F\|_\infty \leq CM^n = C \max_{1\leq i\leq n} \|H_{(i)}(F,1)\|_p^n \tag{2.7.18}$$

where C is a constant that depends on p and n. Thus (2.7.14) follows from (2.7.17) and (2.7.18).

The result can be generalized to the case without the assumption $\|p_F\|_\infty < \infty$ by the same argument as in Theorem 5 of Bally & Caramellino [5]. □

Proof of Theorem 2.7.4. Choose $p_2 > p_1 > n$, let $p_3 = \frac{p_1 p_2}{p_2 - p_1}$ and $q = \frac{p_1}{p_1 - 1}$. Then $\frac{1}{q} + \frac{1}{p_2} + \frac{1}{p_3} = 1$. Thus by density formula (2.7.11) and Hölder's inequality, we have

$$p_F(x) \leq \sum_{i=1}^{n} \|\mathbf{1}_{B_{(x,2\rho)}}(F)\|_{p_3} \|\partial_i Q_n(F - x)\|_q \|H_{(i)}(F, \phi_\rho^x(F))\|_{p_2}. \tag{2.7.19}$$

Then, (2.7.12) is a consequence of (2.7.19), Lemma 2.7.2 and 2.7.6. The inequality (2.7.13) can be proved similarly. □

Chapter 3

Nonlinear rough paths

In this chapter, we develop the theory of nonlinear rough paths. Following the ideas of Lyons and Gubinelli, we define the nonlinear rough integral $\int_s^t W(dr, Y_r)$, where W and Y are only α-Hölder continuous in time with $\alpha \in (\frac{1}{3}, \frac{1}{2}]$. Also, we study the Kunita-type equation $Y_t = \xi + \int_0^t W(dr, Y_s)$, obtaining the local and global existence and uniqueness of the solution under suitable sufficient conditions. As an application, we study transport equations with rough vector fields and observe that the classical solution formula for smooth and Young's cases does not provide a solution to the rough equation. Indeed this formula satisfies a transport equation with additional compensator terms (see (1.2.7)).

3.1 Preliminaries

Fix a time interval $[0,T]$. Assume that $\alpha \in (\frac{1}{3}, \frac{1}{2}]$. Let V and K be Banach spaces. We follow the construction of Friz & Hairer [36, Chapters 2, 4] to introduce the basic framework of the theory of (linear) rough paths.

Definition 3.1.1. *(i) $\mathscr{C}^\alpha([0,T];V)$ is the space of functions on $[0,T]$ taking values in V such that the following α-Hölder seminorm is finite*

$$\|\Phi\|_\alpha := \sup_{s \neq t \in [0,T]} \frac{\|\Phi_{s,t}\|_V}{|t-s|^\alpha}, \tag{3.1.1}$$

where $\Phi_{s,t} := \Phi_t - \Phi_s$.

(ii) $\mathscr{C}_2^\alpha([0,T]^2;K)$ is the space of functions on $[0,T]^2$ taking values in K and such that the fol-

lowing α-Hölder seminorm is finite

$$\|\Psi\|_\alpha := \sup_{s\neq t\in[0,T]} \frac{\|\Psi_{s,t}\|_K}{|t-s|^\alpha}. \tag{3.1.2}$$

A V-valued rough path, introduced below, is defined as a pair of a rough function and a double integral term.

Definition 3.1.2. *The space of rough paths $\mathscr{C}^\alpha([0,T];V)$ is the collection of pairs $\mathbf{X}=(X,\mathbb{X})$ satisfying the following properties:*

(i) $X\in\mathscr{C}^\alpha([0,T];V)$.

(ii) $\mathbb{X}\in\mathscr{C}_2^{2\alpha}([0,T]^2;V\otimes V)$*, where $V\otimes V$ is the tensor product space equipped with the projective norm.*

(iii) $(X,\mathbb{X})$ *satisfies Chen's relation: for all $(s,u,t)\in[0,T]^3$,*

$$\mathbb{X}_{s,t}-\mathbb{X}_{s,u}-\mathbb{X}_{u,t}=X_{s,u}\otimes X_{u,t}. \tag{3.1.3}$$

Here $\mathbb{X}$ has to be interpreted as a version of the following double integral:

$$\int_s^t X_{s,r}\otimes dX_r=\int_s^t\int_s^r dX_u\otimes dX_r:=\mathbb{X}_{s,t}.$$

Let $X\in\mathscr{C}^\alpha([0,T];V)$. We define rough paths controlled by X as follows:

Definition 3.1.3 (Definition 4.6 of Friz & Hairer [36])**.** *Let $X\in\mathscr{C}^\alpha([0,T];V)$. An element $Y\in\mathscr{C}^\alpha([0,T];K)$ is said to be controlled by X, if there exist functions $Y'\in\mathscr{C}^\alpha([0,T];\mathscr{L}(V;K))$ and $R^Y\in\mathscr{C}_2^{2\alpha}([0,T]^2;K)$, such that*

$$Y_{s,t}=Y'_s(X_{s,t})+R^Y_{s,t}$$

for any $s,t \in [0,T]$. Here $\mathscr{L}(V;K)$ denotes the space of continuous linear operators from V to K equipped with the operator norm. The function Y' is called the Gubinelli derivative of Y.

Denote by $\mathscr{D}_X^{2\alpha}(K)$ the space of such pairs (Y,Y'). With an abuse of notations, we sometimes write $Y \in \mathscr{D}_X^{2\alpha}(K)$ instead of $(Y,Y') \in \mathscr{D}_X^{2\alpha}(K)$.

Suppose that $X \in \mathscr{C}^\alpha([0,T];V)$ and $(Y,Y') \in \mathscr{D}_X^{2\alpha}(\mathscr{L}(V;K))$. Then, the Gubinelli derivative Y' takes values in $\mathscr{L}(V;\mathscr{L}(V;K))$, which can be identified with $\mathscr{L}(V \otimes V;K)$. The next theorem defines the (linear) rough integral.

Theorem 3.1.4 (Theorem 4.10 (a) of Friz & Hairer [36]). *Let $\mathbf{X} = (X,\mathbb{X}) \in \mathscr{C}^\alpha([0,T];V)$. Suppose that $(Y,Y') \in \mathscr{D}_X^{2\alpha}(\mathscr{L}(V;K))$. Then the following "compensated Riemann-Stieltjes sum"*

$$\sum_{k=1}^{n} \Xi_{t_k,t_{k-1}} := \sum_{k=1}^{n} \left[Y_{t_{k-1}}(X_{t_{k-1},t_k}) + Y'_{t_{k-1}}(\mathbb{X}_{t_{k-1},t_k})\right], \tag{3.1.4}$$

converges as $|\pi| \to 0$, where $\pi = (s = t_1 < t_2 < \cdots < t_n = t)$ and $|\pi| = \max_{1 \le k \le n} |t_k - t_{k-1}|$. Denote by $\mathscr{I}_{s,t}(\Xi)$ the limit of (3.1.4). Then, $\mathscr{I}_{s,t}(\Xi)$ is additive, that is $\mathscr{I}_{s,t}(\Xi) = \mathscr{I}_{s,u}(\Xi) + \mathscr{I}_{u,t}(\Xi)$ for any $0 \le s < u < t \le T$. Moreover, the following estimate is satisfied for all $0 \le s < t \le T$,

$$\left\| \mathscr{I}_{s,t}(\Xi) - \Xi_{s,t} \right\|_K \le k_\alpha(\|X\|_\alpha \|R^Y\|_{2\alpha} + \|\mathbb{X}\|_{2\alpha}\|Y'\|_\alpha)|t-s|^{3\alpha}, \tag{3.1.5}$$

where $k_\alpha = (1-2^{1-3\alpha})^{-1}$. By definition, the rough integral of Y against $\mathbf{X} = (X,\mathbb{X})$ is defined as follows, for all $0 \le s < t \le T$,

$$\int_s^t Y_r d\mathbf{X}_r := \mathscr{I}_{s,t}(\Xi). \tag{3.1.6}$$

Similarly we can define the rough integral $\int_s^t Y_r \otimes d\mathbf{X}_r \in V_1 \otimes V_2$, for any $\mathbf{X} = (X,\mathbb{X}) \in \mathscr{C}^\alpha([0,T];V_1)$ and $(Y,Y') \in \mathscr{D}_X^{2\alpha}(V_2)$. Theorem 3.1.4 can be proved by using the following Sewing Lemma. In this case, $\gamma = 3\alpha > 1$ and k_α comes from inequality (3.1.7) below. The Sewing Lemma is cited from Lemma 2.1 of Feyel & De la Pradelle [33] (see also Gubinelli [37]). It will also be used later in the theory of nonlinear rough paths.

Lemma 3.1.5 (Sewing Lemma)**.** *Let $\beta \in (0,1]$, and let $\Xi \in \mathscr{C}_2^\beta([0,T]^2;K)$. Suppose there exist $C>0$ and $\gamma>1$ such that the following inequality holds:*

$$\|\delta\Xi(s,u,t)\|_K := \|\Xi_{s,t}-\Xi_{s,u}-\Xi_{u,t}\|_K \le C|t-s|^\gamma,$$

for any $0\le s\le u\le t\le T$. Then there exists a unique (up to an additive constant) function $\mathscr{I}(\Xi)\in\mathscr{C}^\beta([0,T];V)$, such that the following inequality holds

$$\|\mathscr{I}_{s,t}(\Xi)-\Xi_{s,t}\|_K = \|\mathscr{I}_t(\Xi)-\mathscr{I}_s(\Xi)-\Xi_{s,t}\|_K \le (1-2^{1-\gamma})^{-1}C|t-s|^\gamma. \tag{3.1.7}$$

Moreover, $\mathscr{I}_{s,t}(\Xi)$ can be represented as follows,

$$\mathscr{I}_{s,t}(\Xi) = \lim_{|\pi|\to 0}\sum_{k=1}^n \Xi_{t_{k-1},t_k}, \tag{3.1.8}$$

where $\pi=(s=t_0<t_1<\cdots<t_n=t)$ and the limit is independent of the choice of π.

The next proposition shows that the rough integral is controlled by X.

Proposition 3.1.6 (Theorem 4.10 (b) of Friz & Hairer [36])**.** *Suppose that $(X,\mathbb{X})\in\mathscr{C}^\alpha([0,T];V)$ and $(Y,Y')\in\mathscr{D}_X^{2\alpha}(\mathscr{L}(V;K))$. Let*

$$Z_t=\int_0^t Y_r d\mathbf{X}_r.$$

Then, Z is an α-Hölder continuous function taking values in K. Moreover Z is controlled by X with Y as a Gubinelli derivative.

In the next proposition, we define the integration of two controlled rough paths.

Proposition 3.1.7. *Let V, K_1 and K_2 be Banach spaces. Suppose that $\mathbf{X}=(X,\mathbb{X})\in\mathscr{C}^\alpha([0,T];V)$ and $(Y,Y')\in\mathscr{D}_X^{2\alpha}(K_1)$.*

(i) [Remark 4.11 of Friz & Hairer [36]] Suppose that $(Z,Z')\in\mathscr{D}_X^{2\alpha}(K_2)$. The following limit

exists

$$\lim_{|\pi|\to 0}\sum_{k=1}^{n}\big[Z_{t_{k-1}}\otimes Y_{t_{k-1},t_k}+(Z'_{t_{k-1}}\otimes Y'_{t_{k-1}})(\mathbb{X}_{t_{k-1},t_k})\big], \tag{3.1.9}$$

where $\pi=(s=t_0<t_1<\cdots<t_n=t)$ *and defines the integral* $\int_s^t Z_r\otimes dY_r$.

(ii) [Proposition 7.1 of Friz & Hairer [36]] Let $\mathbb{Y}:[0,T]^2\to K_1\otimes K_1$ *be given by*

$$\mathbb{Y}_{s,t}=\int_s^t Y_r\otimes dY_r-Y_s\otimes Y_{s,t}, \tag{3.1.10}$$

and the integral in (3.1.10) *is defined by* (3.1.9). *Then,* $\mathbf{Y}:=(Y,\mathbb{Y})$ *is a rough path. Suppose that* $(Z,\widetilde{Z}')\in\mathscr{D}_Y^{2\alpha}(K_2)$. *Let* $Z'_t=\widetilde{Z}'_tY'_t$ *for all* $t\in[0,T]$. *Then,* $(Z,Z')\in\mathscr{D}_X^{2\alpha}(K_2)$. *In addition, the following equality holds*

$$\int_s^t Z_r\otimes d\mathbf{Y}_r=\int_s^t Z_r\otimes dY_r, \tag{3.1.11}$$

where the integral on the left-hand side is in the sense of Theorem 3.1.4, and the integral on the right -hand side is in the sense of (3.1.9).

Remark 3.1.8. *Assume the conditions of Proposition 3.1.7 (i) where* $K_2=\mathscr{L}(K_1;K)$. *Then,*

$$\int_s^t Z_r dY_r:=\lim_{|\pi|\to 0}\sum_{k=1}^{n}\big[Z_{t_{k-1}}(Y_{t_{k-1},t_k})+(Z'_{t_{k-1}}Y'_{t_{k-1}})\mathbb{X}_{t_{k-1},t_k}\big], \tag{3.1.12}$$

and

$$\int_s^t dZ_r(Y_r):=\lim_{|\pi|\to 0}\sum_{k=1}^{n}\big[Z_{t_{k-1},t_k}(Y_{t_{k-1}})+(Z'_{t_{k-1}}Y'_{t_{k-1}})\mathbb{X}^*_{t_{k-1},t_k}\big], \tag{3.1.13}$$

are well-defined, where $\pi=(s=t_0<t_1<\cdots<t_n=t)$, $(Z'_tY'_t):V\otimes V\to K$ *is given by*

$$(Z'_tY'_t)(x,y)=Z'_t(x)\big[Y'_t(y)\big].$$

and $*$ *denotes the transpose operator on the tensor product space* $V \otimes V$*, namely,* $(x \otimes y)^* = y \otimes x$*.*

Next, we define the "quadratic compensator" as follows (c.f. (2.7) of Keller & Zhang [53] for an equivalent definition in finite dimensions). It will be used in Section 3.5.

Definition 3.1.9. *Let* $\mathbf{X} = (X, \mathbb{X}) \in \mathscr{C}^\alpha([0,T];V)$*. Suppose that* $(Y, Y') \in \mathscr{D}_X^{2\alpha}(K_1)$ *and* $(Z, Z') \in \mathscr{D}_X^{2\alpha}(K_2)$*.*

(i) The quadratic compensator $\langle X \rangle$ *is a function on* $[0,T]^2$ *with values in* $V \otimes V$ *given by*

$$\langle X \rangle_{s,t} := X_{s,t} \otimes X_{s,t} - 2\mathbb{X}_{s,t}. \tag{3.1.14}$$

(ii) The quadratic compensator $\langle Z, Y \rangle : [0,T]^2 \to K_2 \otimes K_1$ *is given by*

$$\langle Z, Y \rangle_{s,t} := Z_{s,t} \otimes Y_{s,t} - 2\int_s^t Z_{s,r} \otimes dY_r. \tag{3.1.15}$$

Remark 3.1.10. *(i) Similar as the quadratic variation of Itô processes, the following equality holds:*

$$\langle Y, Z \rangle_{s,t} = \int_s^t Y'_r \otimes Z'_r d\langle X \rangle_r. \tag{3.1.16}$$

(ii) Particularly, if $K_2 = \mathscr{L}(K_1; K)$*, we write*

$$\langle\langle Z, Y \rangle\rangle_{s,t} := Z_{s,t} Y_{s,t} - 2\int_s^t Z_{s,r} dY_r \tag{3.1.17}$$

and

$$\langle\langle Y, Z \rangle\rangle_{s,t} := Z_{s,t} Y_{s,t} - 2\int_s^t dZ_r (Y_{s,r}). \tag{3.1.18}$$

(iii) It is easy to verify that $\langle X \rangle \in \mathscr{C}_2^{2\alpha}([0,T]; V \otimes V)$*. Similarly,* $\langle Y, Z \rangle$*,* $\langle Z, Y \rangle$*,* $\langle\langle Y, Z \rangle\rangle$ *and* $\langle\langle Z, Y \rangle\rangle$ *are also* 2α*-Hölder continuous in corresponding spaces.*

Finally, we finish this section by introducing the following Taylor's theorem (c.f. Theorem 4.C of Zeidler [88]) for Banach space valued functions. It will be used frequently in estimating residuals.

Theorem 3.1.11 (Taylor's Theorem)**.** *Let V and K be Banach spaces. Assume that the map $\phi : V \to K$ is $\mathscr{C}^n$ in the sense of Fréchet differentiability. Then for any $v, h \in V$, the following generalized Taylor formula holds*

$$\phi(v+h) = \phi(v) + \sum_{k=1}^{n-1} \frac{1}{k!} D^k \phi(v) h^{\otimes k} + R_n,$$

where the residual R_n satisfies the following inequality

$$\|R_n\|_K \leq \frac{1}{n!} \sup_{0 \leq \tau \leq 1} \|D^k \phi(u + \tau h) h^{\otimes n}\|_K.$$

3.2 Nonlinear rough integrals

3.2.1 Definitions

Fix a time interval $[0,T]$. Suppose that $\alpha \in (\frac{1}{3}, \frac{1}{2}]$. In this section, we aim to define the following nonlinear integral:

$$\int_s^t W(dr, Y_r).$$

Here W is α-Hölder continuous in time, and differentiable in space, and Y is α-Hölder continuous. The idea is as follows. Assume that Y is controlled by W, that is $Y_{s,t} = W_{s,t}(\dot{Y}_s) + O(|t-s|^{2\alpha})$. Then, we approximate the nonlinear integral by the following expression:

$$\begin{aligned}\int_s^t W(dr, Y_r) \approx & \int_s^t W(dr, Y_s) + \int_s^t DW(dr, Y_s) Y_{r,s} \\ \approx & \int_s^t W(dr, Y_s) + \int_s^t DW(dr, Y_s) W_{s,r}(\dot{Y}_s) \\ = & W_{s,t}(Y_s) + \int_s^t DW(dr, y) W_{s,r}(x) \Big|_{(x,y)=(\dot{Y}_s, Y_s)},\end{aligned}$$

with the error of order $O(|t-s|^{3\alpha})$. This allows us to pass to the limit as $|\pi| \to 0$ in the following expression

$$\sum_{k=1}^{n}\Big[W_{t_{k-1},t_k}(Y_{t_{k-1}}) + \int_{t_{k-1}}^{t_k} DW(dr,y)W_{t_{k-1},r}(x)\Big|_{(x,y)=(\dot{Y}_{t_{k-1}},Y_{t_{k-1}})}\Big],$$

where $\pi = (s = t_0 < t_1 < \cdots < t_n = t)$. The limit is the desired nonlinear integral.

To this end, we need to introduce the following definitions. Let n be any nonnegative integer. We denote by $\mathscr{I}_n$ the set of all multi-indexes $\boldsymbol{\beta}_n$ of length $n+1$. That is, $\boldsymbol{\beta}_n = (\beta_0,\dots,\beta_n)$, where $\beta_0,\dots,\beta_n$ are nonnegative real numbers. These multi-indexes will be used to characterize the growth of a function and its spatial derivatives.

Definition 3.2.1. *(i) $\mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;K)$ is the space of functions such that the following seminorm is finite:*

$$\|\Phi\|_{\alpha,\boldsymbol{\beta}_n} := \sum_{k=0}^{n} \sup_{\substack{s\neq t\in[0,T]\\ x\in V}} \frac{\|D^k\Phi_{s,t}(x)\|_{\mathfrak{L}_k(V;K)}}{|t-s|^{\alpha}(1+\|x\|_V)^{\beta_k}}, \tag{3.2.1}$$

where D^k is the k-th Fréchet derivative operator, and $\mathfrak{L}_k(V;K)$ is the corresponding space of linear operators. That is, $\mathfrak{L}_0(V;K) = K$ and $\mathfrak{L}_k(V;K) = \mathscr{L}(V;\mathfrak{L}_{k-1}(V;K))$ for all $k = 1,2,\dots,n$.

(ii) $\mathscr{C}_2^{\alpha,\boldsymbol{\beta}_n^1,\boldsymbol{\beta}_n^2}([0,T]^2\times V^2;K)$ is the space of functions such that the following seminorm is finite:

$$\|\Psi\|_{\alpha,\boldsymbol{\beta}_n^1,\boldsymbol{\beta}_n^2} := \sum_{k=0}^{n} \sup_{\substack{s\neq t\in[0,T]\\ \mathbf{x}=(x_1,x_2)\in V^2}} \frac{\|D^k\Psi_{s,t}(\mathbf{x})\|_{\mathfrak{L}_k(V^2;K)}}{|t-s|^{\alpha}(1+\|x_1\|_V)^{\beta_k^1}(1+\|x_2\|_V)^{\beta_k^2}}, \tag{3.2.2}$$

where $\mathfrak{L}_k(V^2;K)$ are the corresponding linear spaces of derivatives and the product space V^2 is treated as a Banach space equipped with the norm $\|\mathbf{x}\|_{V^2} = \|x_1\|_V + \|x_2\|_V$.

For any positive integer $m \le n$, we write $\boldsymbol{\beta}_n - m = (\beta_0,\dots,\beta_{n-m})$. Then, by definition, it is easy to verify that $\mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;K) \subset \mathscr{C}^{\alpha,\boldsymbol{\beta}_n-m}([0,T]\times V;K)$. Let $\boldsymbol{\beta}_n, \widetilde{\boldsymbol{\beta}}_n \in \mathscr{I}_n$, we write $\boldsymbol{\beta}_n \le \widetilde{\boldsymbol{\beta}}_n$

if $\beta_k \le \widetilde{\beta}_k$ for all $k = 0, \dots, n$. Then, $\mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;K) \subset \mathscr{C}^{\alpha,\widetilde{\boldsymbol{\beta}}_n}([0,T]\times V;K)$ if $\boldsymbol{\beta}_n \le \widetilde{\boldsymbol{\beta}}_n$. The space $\mathscr{C}_2^{\alpha,\boldsymbol{\beta}_n^1,\boldsymbol{\beta}_n^2}([0,T]^2\times V^2;K)$ also has a similar property. Given a multi-index $\boldsymbol{\beta}_n$ where $n \ge 1$, we make use of the following notations:

$$\boldsymbol{\beta}_{n-1}^* = (\beta_0^*, \dots, \beta_{n-1}^*) \text{ and } \boldsymbol{\beta}_{n-1}^{**} = (\beta_0^{**}, \dots, \beta_{n-1}^{**}), \tag{3.2.3}$$

where $\beta_k^* := \max\{\beta_0, \dots, \beta_k\}$ and $\beta_k^{**} := \max\{\beta_1, \dots, \beta_{k+1}\}$ for all $0 \le k \le n-1$.

Given multi-indexes $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_1^*$ and $\boldsymbol{\beta}_1^{**}$, let $\Phi \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_2}([0,T]\times V;K)$ and let $\Psi \in \mathscr{C}_2^{\alpha,\boldsymbol{\beta}_1^*,\boldsymbol{\beta}_1^{**}}([0,T]^2\times V^2;K)$. We make use of the following notations: $\mathscr{R}^\Phi : [0,T]\times V^2 \to K$ and $\mathscr{D}^\Psi : [0,T]^2 \times V^4 \to K$ are given by

$$\mathscr{R}_t^\Phi(x,y) := \Phi_t(y) - \Phi_t(x) - D\Phi_t(x)(y-x),\ x,y \in V \tag{3.2.4}$$

and

$$\mathscr{D}_{s,t}^\Psi(\mathbf{x},\mathbf{y}) = \Psi_{s,t}(\mathbf{y}) - \Psi_{s,t}(\mathbf{x}),\ \mathbf{x},\mathbf{y} \in V^2. \tag{3.2.5}$$

The following lemma provides the estimates for $\mathscr{R}^\Phi$, $\mathscr{D}^\Psi$ and their derivatives. It will be used in the proof of the stability of nonlinear rough integrals.

Lemma 3.2.2. *Suppose that $\mathscr{R}^\Phi$ and $\mathscr{D}^\Psi$ are given as in (3.2.4) and (3.2.5), respectively. Then, for any $x, y \in V$, and $\mathbf{x} = (x_1,x_2), \mathbf{y} = (y_1,y_2) \in V^2$, the following inequalities are satisfied:*

$$\begin{aligned}
\|\mathscr{R}_{s,t}^\Phi(x,y)\|_K \le& \frac{1}{2}\|\Phi\|_{\alpha,\boldsymbol{\beta}_2}(1+\|x\|_V+\|y\|_V)^{\beta_2}\|y-x\|_V^2|t-s|^\alpha, && (3.2.6)\\
\|\mathscr{D}_{s,t}^\Psi(\mathbf{x},\mathbf{y})\|_K \le& \|\Psi\|_{\alpha,\boldsymbol{\beta}_1^1,\boldsymbol{\beta}_1^2}(1+\|x_1\|_V+\|y_1\|_V)^{\beta_1^1}(1+\|x_2\|_V+\|y_2\|_V)^{\beta_1^2}\\
&\times \|\mathbf{y}-\mathbf{x}\|_{V^2}|t-s|^\alpha. && (3.2.7)
\end{aligned}$$

If furthermore $\Phi \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_3}([0,T]\times V;K)$ and $\Psi \in \mathscr{C}_2^{\alpha,\boldsymbol{\beta}_2^1,\boldsymbol{\beta}_2^2}([0,T]^2\times V^2;K)$, then, for all $\mathbf{z}^1, \mathbf{z}^2 \in V^2$, *the following inequalities are satisfied:*

$$\begin{aligned}\|D\mathscr{R}^{\Phi}_{s,t}(x,y)(z_1,z_2)\|_K \leq & \|\Phi\|_{\alpha,\boldsymbol{\beta}_3}(1+\|x\|_V+\|y\|_V)^{\beta_2\vee\beta_3} \\ & \times\left[\|y-x\|_V^2\|z_2\|_V+\|y-x\|_V\|z_1-z_2\|_V\right]|t-s|^{\alpha} \end{aligned} \tag{3.2.8}$$

and

$$\begin{aligned}\|D\mathscr{D}^{\Psi}_{s,t}(\mathbf{x},\mathbf{y})(\mathbf{z}^1,\mathbf{z}^2)\|_K \leq & \|\Psi\|_{\alpha,\boldsymbol{\beta}^1_2,\boldsymbol{\beta}^2_2}(1+\|x_1\|_V+\|y_1\|_V)^{\beta^1_1\vee\beta^1_2}(1+\|x_2\|_V+\|y_2\|_V)^{\beta^2_1\vee\beta^2_2} \\ & \times\left[\|\mathbf{y}-\mathbf{x}\|_{V^2}\|\mathbf{z}^2\|_{V^2}+\|\mathbf{z}^1-\mathbf{z}^2\|_{V^2}\right]|t-s|^{\alpha}. \end{aligned} \tag{3.2.9}$$

Proof. Inequality (3.2.6) is a consequence of Taylor's Theorem 3.1.11 and the linearity of D:

$$\begin{aligned}\|\mathscr{R}^{\Phi}_{s,t}(x,y)\|_K \leq & \frac{1}{2}\sup_{0\leq\tau\leq 1}\|D^2\Phi_{s,t}(\tau x+(1-\tau)y)(y-x,y-x)\|_K \\ \leq & \frac{1}{2}\|\Phi\|_{\alpha,\boldsymbol{\beta}_2}(1+\|x\|_V+\|y\|_V)^{\beta_2}\|y-x\|_V^2|t-s|^{\alpha}. \end{aligned}$$

For inequality (3.2.8), we assume that $\Phi\in\mathscr{C}^{\alpha,\boldsymbol{\beta}_3}([0,T]\times V;K)$. Then, by differentiating $\mathscr{R}^{\Phi}_{s,t}$ on the spatial argument, for any $(x,y),(z_1,z_2)\in V^2$, we have

$$\begin{aligned}D\mathscr{R}^{\Phi}_{s,t}(x,y)(z_1,z_2)= & -D\Phi_{s,t}(x)(z_1)-D^2\Phi_{s,t}(x)(z_1,y-x)+D\Phi_{s,t}(x)(z_1) \\ & +D\Phi_{s,t}(y)(z_2)-D\Phi_{s,t}(x)(z_2) \\ = & D\Phi_{s,t}(y)(z_2)-D\Phi_{s,t}(x)(z_2)-D^2\Phi_{s,t}(x)(z_1,y-x) \\ = & D\Phi_{s,t}(y)(z_2)-D\Phi_{s,t}(x)(z_2)-D^2\Phi_{s,t}(x)(z_2,y-x) \\ & +D^2\Phi_{s,t}(x)(z_2-z_1,y-x). \end{aligned}$$

By Taylor's Theorem 3.1.11 again, we can deduce that

$$\begin{aligned}&\|D\Phi_{s,t}(y)(z_2)-D\Phi_{s,t}(x)(z_2)-D^2\Phi_{s,t}(x)(z_2,y-x)\|_K\\ \leq&\frac{1}{2}\sup_{0\leq\tau\leq 1}\|D^3\Phi_{s,t}(\tau x+(1-\tau)y)(z_2,y-x,y-x)\|_K.\end{aligned}$$

Thus inequality (3.2.8) is a consequence of above two inequalities. Inequalities (3.2.7) and (3.2.9) can be proved similarly. □

In the rest of this chapter, we focus on the case when $K=V$. A nonlinear rough path is defined as follows.

Definition 3.2.3. *Assume that $n\geq 1$. An α-Hölder continuous nonlinear rough path* $\mathbf{W}$ *on the space $\mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$ is defined as a pair $(W,\mathbb{W})$ that satisfies the following properties:*

(i) $W\in\mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$.

(ii) $\mathbb{W}\in\mathscr{C}_2^{2\alpha,\boldsymbol{\beta}^*_{n-1},\boldsymbol{\beta}^{**}_{n-1}}([0,T]^2\times V^2;V)$, *where $\boldsymbol{\beta}^*_{n-1}$ and $\boldsymbol{\beta}^{**}_{n-1}$ are defined in (3.2.3).*

(iii) $(W,\mathbb{W})$ *satisfies Chen's relation:*

$$\mathbb{W}_{s,t}(x,y)-\mathbb{W}_{s,u}(x,y)-\mathbb{W}_{u,t}(x,y)=DW_{u,t}(y)(W_{s,u}(x)),\tag{3.2.10}$$

for all $(x,y)\in V^2$ and $s,u,t\in[0,T]$.

The collection of such rough paths is denoted by $\mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$.

Remark 3.2.4. *(i) In the smooth case, $\mathbb{W}$ can be interpreted as the following integral*

$$\int_s^t DW(dr,y)(W_{s,r}(x))=\int_s^t\frac{\partial}{\partial r}DW(r,y)(W_{s,r}(x))dr=\mathbb{W}_{s,t}(x,y).$$

This explains the choice of the multi-indexes $\boldsymbol{\beta}^*_{n-1}$ and $\boldsymbol{\beta}^{**}_{n-1}$ in point (ii) of Definition 3.2.3. For example, assume that W is twice differentiable with growth index $\boldsymbol{\beta}_2$. Then, one can *bound the growth of* $\mathbb{W}_{s,t}$ *as follows*

$$\frac{\|\mathbb{W}_{s,t}(x,y)\|_{V\otimes V}}{(1+\|x\|_V)^{\beta_0}(1+\|y\|_V)^{\beta_1}} \leq \limsup_{|\pi|\to 0} \frac{\sum_{k=1}^n \left|\frac{\partial}{\partial r}DW(t_{k-1},y)(W_{s,t_{k-1}}(x))\right||t_k-t_{k-1}|}{(1+\|x\|_V)^{\beta_0}(1+\|y\|_V)^{\beta_1}}$$
$$\leq \|W\|_{1,\boldsymbol{\beta}_2}^2|t-s|.$$

By taking the derivative of $\mathbb{W}_{s,t}$*, it can be deduced that the growth of* $D\mathbb{W}_{s,t}(x,y)$ *is bounded by* $\beta_0\vee\beta_1$ *and* $\beta_1\vee\beta_2$ *in x and y,respectively.*

(ii) By definition, we can deduce that $\mathscr{C}^{\alpha,\boldsymbol{\beta}_n^1}([0,T]\times V;V)\subset\mathscr{C}^{\alpha,\boldsymbol{\beta}_n^2-m}([0,T]\times V;V)$ *for all* $m\in\{0,\ldots,n\}$ *and* $\boldsymbol{\beta}_n^1\leq\boldsymbol{\beta}_n^2$*.*

(iii) Assume that $W(t,x)=W_t(x)$ *where* $W_t\in\mathscr{L}(V;V)$*. Then the nonlinear rough path degenerates to the linear rough path. In this case,* $DW_t(x)=W_t$ *and thus*

$$\mathbb{W}_{s,t}(x,y)=\int_s^t W(dr)(W_{s,r}(x)).$$

Let $\mathbf{W}=(W,\mathbb{W})\in\mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$. We make use of the notation

$$\|\mathbf{W}\|_{\mathscr{C}_n}:=\|W\|_{\alpha,\boldsymbol{\beta}_n}+\|\mathbb{W}\|_{\alpha,\boldsymbol{\beta}_{n-1}^*,\boldsymbol{\beta}_{n-1}^{**}}. \tag{3.2.11}$$

Notice that $\mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$ is not a linear space with the usual addition and scalar product. Thus $\|\cdot\|_{\mathscr{C}_n}$ is not a seminorm in the usual sense. We introduce the pseudometric on $\mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$ given by

$$\rho_{\alpha,\boldsymbol{\beta}_n}(\mathbf{W},\widetilde{\mathbf{W}})=\|W-\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_n}+\|\mathbb{W}-\widetilde{\mathbb{W}}\|_{2\alpha,\boldsymbol{\beta}_{n-1}^*,\boldsymbol{\beta}_{n-1}^{**}}. \tag{3.2.12}$$

Consider the following equivalent relation: $\mathbf{W}\sim\widetilde{\mathbf{W}}$ if and only if there exists $f\in\mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$ such that $W(t,x)-\widetilde{W}(t,x)=f(x)$ for all $(t,x)\in[0,T]\times V$. Then, $\rho_{\alpha,\boldsymbol{\beta}_n}$ is really a metric on the quotient space $\mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)/\sim$.

Let $W \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$. Like in the linear case, we also define the space of nonlinear rough paths controlled by W.

Definition 3.2.5. *The space of basic nonlinear rough paths controlled by W, denoted by $\mathscr{E}_W^{2\alpha}$, is the collection of pairs $(Y,\dot{Y}) \in \mathscr{C}^\alpha([0,T];V)\times\mathscr{C}^\alpha([0,T];V)$ (see (3.1.1)) such that, for all $s,t\in[0,T]$,*

$$Y_{s,t} = W_{s,t}(\dot{Y}_s) + R^Y_{s,t}, \tag{3.2.13}$$

where $R^Y \in \mathscr{C}_2^{2\alpha}([0,T]^2;V)$ (see (3.1.2)). The function $\dot{Y}$ above is called the Gubinelli derivative of Y with respect to W.

Remark 3.2.6. *(i) In the linear case, the set of controlled rough paths is a linear space. However, in the nonlinear case, the set $\mathscr{E}_W^{2\alpha}$ does not need to be a linear space with the usual addition and scalar product, because it may be not closed under these operations.*

(ii) Assume that $V=\mathbb{R}$ and $W(t,x)=xW_t$, then the controlled rough path satisfies the following equality

$$Y_{s,t} = \dot{Y}_s W_{s,t} + R^Y_{s,t},$$

which coincides with the classical definition (see Definition 3.1.3) in the linear case.

(iii) With an abuse of notations, we sometimes write $Y\in\mathscr{E}_W^{2\alpha}$ instead of $(Y,\dot{Y})\in\mathscr{E}_W^{2\alpha}$.

Suppose that $W,\widetilde{W}\in\mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$. Let $(Y,\dot{Y})\in\mathscr{E}_W^{2\alpha}$ and $(\widetilde{Y},\dot{\widetilde{Y}})\in\mathscr{E}_{\widetilde{W}}^{2\alpha}$. A "distance" between $(Y,\dot{Y})$ and $(\widetilde{Y},\dot{\widetilde{Y}})$ is defined as follows (c.f. Friz & Hairer [36, Section 4.4] for the linear case):

$$d_{\alpha,W,\widetilde{W}}\big((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})\big) = \|\dot{Y}-\dot{\widetilde{Y}}\|_\alpha + \|R^Y - R^{\widetilde{Y}}\|_{2\alpha}. \tag{3.2.14}$$

Notice that the definition of $d_{\alpha,W,\widetilde{W}}$ does not include the term $\|Y-\widetilde{Y}\|_\alpha$. Indeed, this term can be estimated in terms of $d_{\alpha,W,\widetilde{W}}\big((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})\big)$ as it is shown in the next lemma. On the other hand, one will see in the next lemma, that $\|Y-\widetilde{Y}\|_\alpha$ depends also on $\|W-\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_1}$ without a factor

T^{α}. If we include $\|W-\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_1}$ in $d_{\alpha,W,\widetilde{W}}\big((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})\big)$, the absence of this factor T^{α} will cause difficulties in the proof of the existence of solutions to RDEs in Section 3.3. For this reason, the term $\|W-\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_1}$ is not included in $d_{\alpha,W,\widetilde{W}}\big((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})\big)$, and it is treated independently.

Lemma 3.2.7. *Let $W,\widetilde{W}\in\mathscr{C}^{\alpha,\boldsymbol{\beta}_1}([0,T]\times V;V)$. Suppose that $(Y,\dot{Y})\in\mathscr{E}_W^{2\alpha}$ and $(\widetilde{Y},\dot{\widetilde{Y}})\in\mathscr{E}_{\widetilde{W}}^{2\alpha}$. Then the following estimate holds:*

$$\begin{aligned}\|Y-\widetilde{Y}\|_{\alpha}\leq&(1+\|\dot{Y}\|_{\infty})^{\beta_0}\|W-\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_1}\\&+\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_1}(1+\|\dot{Y}\|_{\infty}+\|\dot{\widetilde{Y}}\|_{\infty})^{\beta_1}\|\dot{Y}_0-\dot{\widetilde{Y}}_0\|_V\\&+T^{\alpha}(1+\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_1})(1+\|\dot{Y}\|_{\infty}+\|\dot{\widetilde{Y}}\|_{\infty})^{\beta_1}d_{\alpha,W,\widetilde{W}}\big((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})\big).\end{aligned}\tag{3.2.15}$$

Proof. Since Y and $\widetilde{Y}$ are controlled by W and $\widetilde{W}$ respectively, then we have

$$\|Y_{s,t}-\widetilde{Y}_{s,t}\|_V\leq\big\|W_{s,t}(\dot{Y}_s)-\widetilde{W}_{s,t}(\dot{Y}_s)\big\|_V+\big\|\widetilde{W}_{s,t}(\dot{Y}_s)-\widetilde{W}_{s,t}(\dot{\widetilde{Y}}_s)\big\|_V+\big\|R^Y_{s,t}-R^{\widetilde{Y}}_{s,t}\big\|_V.$$

Notice that by Taylor's Theorem 3.1.11,

$$\begin{aligned}\big\|\widetilde{W}_{s,t}(\dot{Y}_s)-\widetilde{W}_{s,t}(\dot{\widetilde{Y}}_s)\big\|_V\leq&\sup_{1\leq\tau\leq1}\big\|D\widetilde{W}_{s,t}(\tau\dot{Y}_s+(1-\tau)\dot{\widetilde{Y}}_s)(\dot{Y}_s-\dot{\widetilde{Y}}_s)\big\|_V\\\leq&\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_1}(1+\|\dot{Y}\|_{\infty}+\|\dot{\widetilde{Y}}\|_{\infty})^{\beta_1}\|\dot{Y}_s-\dot{\widetilde{Y}}_s\|_V.\end{aligned}$$

On the other hand, for any $Y\in\mathscr{C}^{\alpha}([0,T];V)$ we have

$$\|Y_s\|_V\leq\|Y_0\|_V+\|Y_s-Y_0\|_V\leq\|Y_0\|_V+\|Y\|_{\alpha}s^{\alpha}.$$

As a consequence, we can write

$$
\begin{aligned}
\|Y_{s,t}-\widetilde{Y}_{s,t}\|_V \leq & \|W-\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_1}(1+\|\dot{Y}\|_\infty)^{\beta_0}|t-s|^\alpha \\
& +\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_1}(1+\|\dot{Y}\|_\infty+\|\dot{\widetilde{Y}}\|_\infty)^{\beta_1}\left(\|\dot{Y}_0-\dot{\widetilde{Y}}_0\|_V+s^\alpha\|\dot{Y}-\dot{\widetilde{Y}}\|_\alpha\right)|t-s|^\alpha \\
& +\left\|R^Y-R^{\widetilde{Y}}\right\|_{2\alpha}|t-s|^{2\alpha}.
\end{aligned}
$$

This proves inequality (3.2.15). □

Applying Lemma 3.2.7, the supremum norm of $Y-\widetilde{Y}$ can be estimated as follows:

$$
\begin{aligned}
\|Y-\widetilde{Y}\|_\infty \leq & \|Y_0-\widetilde{Y}_0\|_V+T^\alpha\|Y-\widetilde{Y}\|_\alpha \\
\leq & T^\alpha(1+\|\dot{Y}\|_\infty)^{\beta_0}\|W-\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3} \\
& +(1+T^\alpha)(1+\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3})(1+\|\dot{Y}\|_\infty+\|\dot{\widetilde{Y}}\|_\infty)^{\beta_1}\left(\|Y_0-\widetilde{Y}_0\|_V+\|\dot{Y}_0-\dot{\widetilde{Y}}_0\|_V\right) \\
& +T^{2\alpha}(1+\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3})(1+\|\dot{Y}\|_\infty+\|\dot{\widetilde{Y}}\|_\infty)^{\beta_1}d_{\alpha,W,\widetilde{W}}\left((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})\right).
\end{aligned}
\tag{3.2.16}
$$

Both inequalities (3.2.15) and (3.2.16) represent how the difference between Y and $\widetilde{Y}$ depends on $\|W-\widetilde{W}\|_{\alpha,\beta_1}$, $\|\dot{Y}_0-\dot{\widetilde{Y}}_0\|_V$ and $d_{\alpha,W,\widetilde{W}}\left((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})\right)$. As we stated before, the factors T^α and $T^{2\alpha}$ in each inequality are critical for the existence of the solution to equation (3.3.1) in Section 3.3.

Remark 3.2.8. *$d_{\alpha,W,\widetilde{W}}$ defined in a subspace of $\mathscr{C}^\alpha([0,T];V)\times\mathscr{C}^\alpha([0,T];V)$ is not a metric, because the values of $Y_0,\widetilde{Y}_0$ or $\dot{Y}_0,\dot{\widetilde{Y}}_0$ may differ even if $d_{\alpha,W,\widetilde{W}}((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}}))=0$. For any $\mathbf{y}=(y_1,y_2)\in V^2$, let*

$$
\mathscr{E}^{2\alpha}_{W,\mathbf{y}}=\left\{(Y,\dot{Y})\in\mathscr{E}^{2\alpha}_W,(Y_0,\dot{Y}_0)=(y_1,y_2)\right\}.
$$

Then $d_{\alpha,W}=d_{\alpha,W,W}$ is really a metric on $\mathscr{E}^{2\alpha}_{W,\mathbf{y}}$.

The next lemma shows that $\mathscr{E}^{2\alpha}_{W,\mathbf{y}}$ is complete under the metric $d_{\alpha,W}$.

Lemma 3.2.9. *Suppose that $W\in\mathscr{C}^{\alpha,\boldsymbol{\beta}_1}([0,T]\times V;V)$. Let $\mathbf{y}=(y_1,y_2)\in V^2$. Then $(\mathscr{E}^{2\alpha}_{W,\mathbf{y}},d_{\alpha,W})$ is a complete metric space.*

Proof. Suppose that $\{(Y^n, \dot{Y}^n)\}_{n\geq 1} \subset \mathscr{E}^{2\alpha}_{W,\mathbf{y}}$ is a Cauchy sequence under the metric $d_{\alpha,W}$. We first show that $\{(Y^n, \dot{Y}^n, R^{Y^n})\}_{n\geq 1}$ converges to $(Y, \dot{Y}, R^Y)$ in the product space $\mathscr{C}^\alpha([0,T];V) \times \mathscr{C}^\alpha([0,T];V) \times \mathscr{C}^{2\alpha}_2([0,T]^2;V)$ equipped with the Hölder seminorms. Notice that $\mathscr{C}^\alpha([0,T];V)$ is complete with respect to the norm

$$\|Y\|_{\mathscr{C}^\alpha([0,T];V)} := \|Y_0\|_V + \|Y\|_\alpha.$$

Thus there exists $\dot{Y} \in \mathscr{C}^\alpha([0,T];V)$, such that $\dot{Y}^n \to \dot{Y}$ as $n \to \infty$ pointwise and in $\mathscr{C}^\alpha([0,T];V)$. Next, we will show the convergence of $\{R^{Y^n}\}_{n\geq 1}$. Fix $(s,t) \in [0,T]^2$. Then, for and $n,m \geq 1$, we have

$$\|R^{Y^n}_{s,t} - R^{Y^m}_{s,t}\|_V \leq |t-s|^{2\alpha} \|R^{Y^n} - R^{Y^m}\|_{2\alpha}.$$

Therefore, $\{R^{Y^n}_{s,t}\}_{n\geq 1}$ is a Cauchy sequence in V, and thus has a limit denoted by $R^Y_{s,t}$. On the other hand, we can show that

$$\limsup_{n\to\infty} \sup_{s\neq t\in[0,T]} \frac{\|R^Y_{s,t} - R^{Y^n}_{s,t}\|_V}{|t-s|^{2\alpha}} \leq \lim_{n\to\infty} \lim_{m\to\infty} \sup_{s\neq t\in[0,T]} \frac{\|R^{Y^m}_{s,t} - R^{Y^n}_{s,t}\|_V}{|t-s|^{2\alpha}} = 0.$$

This implies that, as a sequence of functions, $\{R^{Y^n}\}_{n\geq 1}$ is also convergent in $\mathscr{C}^{2\alpha}_2([0,T]^2;V)$. To prove the convergence of $\{Y^n\}_{n\geq 1}$, it suffices to show that $\{Y^n\}_{n\geq 1}$ is Cauchy in $\mathscr{C}^\alpha([0,T];V)$ with the α-Hölder seminorm. Notice that for any $n,m \geq 1$, Y^n and Y^m are both controlled by W, then, as a consequence of Lemma 3.2.7, we have

$$\begin{aligned} \|Y^n - Y^m\|_\alpha \leq & T^\alpha (1 + \|W\|_{\alpha,\boldsymbol{\beta}_1})(1 + \|\dot{Y}^n\|_\infty + \|\dot{Y}^m\|_\infty)^{\beta_1} \\ & \times d_{\alpha,W,\widetilde{W}}\big((Y^n, \dot{Y}^n), (Y^m, \dot{Y}^m)\big). \end{aligned} \tag{3.2.17}$$

Observe that

$$\sup_{n\geq 1} \|\dot{Y}^n\|_\infty \leq y_2 + T^\alpha \sup_{n\geq 1} \|\dot{Y}^n\|_\alpha = C < \infty.$$

Therefore, $\{Y^n\}_{n\geq 1}$ converges to a function Y in $\mathscr{C}^\alpha([0,T];V)$.

Finally, notice that for any $s,t \in [0,T]$,

$$Y_{s,t} = \lim_{n\to\infty} Y^n_{s,t} = \lim_{n\to\infty} \left[W_{s,t}(\dot{Y}^n_s) + R^{Y^n}_{s,t}\right] = W_{s,t}(\dot{Y}_s) + R^Y_{s,t}. \tag{3.2.18}$$

Thus $(Y,\dot{Y}) \in \mathscr{E}^{2\alpha}_{W,\mathbf{y}}$ with the remainder R^Y. □

In the next theorem, we define the nonlinear rough integral of a basic controlled rough path against a nonlinear rough path.

Theorem 3.2.10. *Suppose that* $\mathbf{W} = (W,\mathbb{W}) \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_2}([0,T]\times V;V)$. *Let* $(Y,\dot{Y}) \in \mathscr{E}^{2\alpha}_W$. *We define* $\Xi \in \mathscr{C}^{\alpha}_2([0,T]^2;V)$ *as follows:*

$$\Xi_{s,t} = W_{s,t}(Y_s) + \mathbb{W}_{s,t}(\dot{Y}_s, Y_s).$$

Then the following limit exists

$$\mathscr{J}_{s,t}(\Xi) := \lim_{|\pi|\to 0} \sum_{k=1}^{n} \Xi_{t_{k-1},t_k} \tag{3.2.19}$$

where $\pi = (s = t_0 < t_1 < \cdots < t_n = t)$. *Moreover,*

$$\left\| \mathscr{J}_{s,t}(\Xi) - \Xi_{s,t} \right\|_V \le C_1 |t-s|^{3\alpha}, \tag{3.2.20}$$

where

$$C_1 = k_\alpha \|\mathbf{W}\|_{\mathscr{C}_2} (1+2\|\dot{Y}\|_\infty)^{\beta_0\vee\beta_1} (1+2\|Y\|_\infty)^{\beta_1\vee\beta_2} \left(\|Y\|_\alpha + \|Y\|^2_\alpha + \|\dot{Y}\|_\alpha + \|R^Y\|_{2\alpha}\right), \tag{3.2.21}$$ [1]

and k_α *is defined in (3.1.5).*

Proof. For any $0 \le s \le u \le t \le T$, we write

$$\begin{aligned}\delta\Xi_{s,u,t} =& \Xi_{s,t} - \Xi_{s,u} - \Xi_{u,t} \\ =& -[W_{u,t}(Y_u) - W_{u,t}(Y_s)] + [\mathbb{W}_{s,t}(\dot{Y}_s, Y_s) - \mathbb{W}_{s,u}(\dot{Y}_s, Y_s) - \mathbb{W}_{u,t}(\dot{Y}_u, Y_u)].\end{aligned} \tag{3.2.22}$$

According to Lemma 3.1.5, to prove (3.2.19) and (3.2.20), it suffices to show that $\|\delta\Xi_{s,u,t}\|_V$ is of order $O(|t-s|^{3\alpha})$. Recall notations (3.2.4) and (3.2.5). Since Y is controlled by W, we can write

$$\begin{aligned}W_{u,t}(Y_u) - W_{u,t}(Y_s) =& DW_{u,t}(Y_s)(Y_{s,u}) + \mathscr{R}^W_{u,t}(Y_s, Y_u) \\ =& DW_{u,t}(Y_s)(W_{s,u}(\dot{Y}_s)) + DW_{u,t}(Y_s)(R^Y_{s,u}) + \mathscr{R}^W_{u,t}(Y_s, Y_u).\end{aligned} \tag{3.2.23}$$

On the other hand, by Chen's relation (3.2.10), we have

$$\begin{aligned}&\mathbb{W}_{s,t}(\dot{Y}_s, Y_s) - \mathbb{W}_{s,u}(\dot{Y}_s, Y_s) - \mathbb{W}_{u,t}(\dot{Y}_u, Y_u) \\ &\qquad = DW_{u,t}(Y_s)(W_{s,u}(\dot{Y}_s)) - \mathscr{D}^{\mathbb{W}}_{u,t}\big((\dot{Y}_s, Y_s), (\dot{Y}_u, Y_u)\big).\end{aligned} \tag{3.2.24}$$

Notice that, by definition, $\mathbb{W} \in \mathscr{C}_2^{2\alpha,\boldsymbol{\beta}_1^*,\boldsymbol{\beta}_1^{**}}([0,T]^2 \times V^2; V)$ where $\boldsymbol{\beta}_1^* = (\beta_0, \beta_0 \vee \beta_1)$ and $\boldsymbol{\beta}_1^{**} = (\beta_1, \beta_1 \vee \beta_2)$. Combining (3.2.22) - (3.2.24), with (3.2.6) and (3.2.7) and recalling (3.2.11), we obtain the following inequality

$$\begin{aligned}&\|\delta\Xi_{s,u,t}\|_V \\ \le& \|DW_{u,t}(Y_s)\|_{\mathfrak{L}_1(V;V)}\|R^Y_{s,u}\|_V + \frac{1}{2}\|W\|_{\alpha,\boldsymbol{\beta}_2}(1+2\|Y\|_\infty)^{\beta_2}\|Y\|^2_\alpha|t-s|^{3\alpha} \\ &+ \|\mathbb{W}\|_{2\alpha,\boldsymbol{\beta}_1^*,\boldsymbol{\beta}_1^{**}}(1+2\|\dot{Y}\|_\infty)^{\beta_1^*}(1+2\|Y\|_\infty)^{\beta_1^{**}}(\|Y\|_\alpha + \|\dot{Y}\|_\alpha)|t-s|^{3\alpha} \\ \le& \|\mathbf{W}\|_{\mathscr{C}_2}(1+2\|\dot{Y}\|_\infty)^{\beta_0\vee\beta_1}(1+2\|Y\|_\infty)^{\beta_1\vee\beta_2}\big(\|Y\|_\alpha + \|Y\|^2_\alpha + \|\dot{Y}\|_\alpha + \|R^Y\|_{2\alpha}\big)|t-s|^{3\alpha}.\end{aligned} \tag{3.2.25}$$

Thus we complete the proof by applying Lemma 3.1.5. □

Notice that $\mathscr{J}_{s,t}(\Xi)$ in Theorem 3.2.10 can be expressed as the limit of sums over a sequence of partitions π^n as $|\pi^n| \to 0$. As a consequence of this fact, one can show that $\mathscr{I}_{s,t}(\Xi)$ is additive. Therefore, we can define the nonlinear integral of Y against W on any time interval $[s,t] \subset [0,T]$ by $\mathscr{I}_{s,t}(\Xi)$, that is

$$\int_s^t W(dr, Y_r) := \mathscr{I}_{s,t}(\Xi). \tag{3.2.26}$$

By definition, we can easily verify that Ξ in Theorem 3.2.10 is also α-Hölder continuous. Recall that $\boldsymbol{\beta}_1^* = (\beta_0, \beta_0 \vee \beta_1)$ and $\boldsymbol{\beta}_1^{**} = (\beta_1, \beta_1 \vee \beta_2)$. Thus we have the following estimate,

$$\begin{aligned}\|\Xi_{s,t}\|_V \le & \|W_{s,t}(Y_s)\|_V + \|\mathbb{W}_{s,t}(\dot{Y}_s, Y_s)\|_V \\ \le & \|W\|_{\alpha,\boldsymbol{\beta}_2}(1+\|Y\|_\infty)^{\beta_0}|t-s|^\alpha \\ & + \|\mathbb{W}\|_{2\alpha,\boldsymbol{\beta}_1^*,\boldsymbol{\beta}_1^{**}}(1+\|\dot{Y}\|_\infty)^{\beta_0}(1+\|Y\|_\infty)^{\beta_1}|t-s|^{2\alpha}.\end{aligned} \tag{3.2.27}$$

The following estimates follow from (3.2.20) and (3.2.27):

$$\begin{aligned}\left\|\int_s^t W(dr, Y_r)\right\|_V \le & \|\Xi_{s,t}\|_V + \|\mathscr{I}_{s,t}(\Xi) - \Xi_{s,t}\|_V \\ \le & C_2|t-s|^\alpha,\end{aligned} \tag{3.2.28}$$

where

$$C_2 = C_1 T^{2\alpha} + \|W\|_{\alpha,\boldsymbol{\beta}_2}(1+\|Y\|_\infty)^{\beta_0} + T^\alpha \|\mathbb{W}\|_{2\alpha,\boldsymbol{\beta}_1^*,\boldsymbol{\beta}_1^{**}}(1+\|\dot{Y}\|_\infty)^{\beta_0}(1+\|Y\|_\infty)^{\beta_1}.$$

Remark 3.2.11. *To define a nonlinear rough integral, the growth condition on* $(W, \mathbb{W})$ *is not necessary. In fact, let* $\mathscr{C}_{loc}^{\alpha,2}([0,T] \times V; V)$ *be the collection of pairs* $(W, \mathbb{W})$ *such that* $W : [0,T] \times V \to V$ *is* α*-Hölder in time, and twice differentiable in space with locally bounded spatial derivatives,* $\mathbb{W} : [0,T]^2 \times V \to V$ *is* 2α*-Hölder continuous in time, and differentiable in space with locally bounded spatial derivative, and Chen's relation (3.2.10) holds. For any* $\mathbf{W} = (W, \mathbb{W}) \in \mathscr{C}_{loc}^{\alpha,2}([0,T] \times V; V)$*, and* $(Y, \dot{Y}) \in \mathscr{E}_W^{2\alpha}$*, the expression (3.2.19) is still a well-defined nonlinear rough integral. However,* *the growth condition is really needed to consider the global existence of RDEs (see Section 3.3.2).*

3.2.2 Properties of nonlinear rough integrals

In this section, we present some properties of nonlinear rough integrals. The next proposition shows that the nonlinear rough integral is a basic nonlinear controlled rough path (see Proposition 3.1.6 for the linear result).

Proposition 3.2.12. *Let $\mathbf{W} = (W, \mathbb{W}) \in \mathscr{C}^{\alpha, \boldsymbol{\beta}_2}([0,T] \times V; V)$. Suppose that $(Y, \dot{Y}) \in \mathscr{E}_W^{2\alpha}$. Let $Z : [0,T] \to V$ be the nonlinear rough integral of Y against W in the sense of (3.2.26):*

$$Z_t = \int_0^t W(dr, Y_r). \tag{3.2.29}$$

Then, Z is controlled by W: $(Z, \dot{Z}) = (Z, Y) \in \mathscr{E}_W^{2\alpha}$.

Proof. Let $R^Z_{s,t} := Z_{s,t} - W_{s,t}(Y_s)$. Then by (3.2.20), we can write

$$\begin{aligned}
\|R^Z_{s,t}\|_V &= \left\| \int_s^t W(dr, Y_r) - W_{s,t}(Y_s) \right\|_V \\
&\leq \| \mathscr{I}_{s,t}(\Xi) - \Xi_{s,t} \|_V + \| \mathbb{W}_{s,t}(\dot{Y}_s, Y_s) \|_V \\
&\leq C_1 |t-s|^{3\alpha} + \|\mathbb{W}\|_{2\alpha, \boldsymbol{\beta}_2^*, \boldsymbol{\beta}_2^{**}} (1 + \|\dot{Y}\|_\infty)^{\beta_0} (1 + \|Y\|_\infty)^{\beta_1} |t-s|^{2\alpha},
\end{aligned}$$

where C_1 is the constant appearing in (3.2.21). It follows that

$$\begin{aligned}
\|R^Z\|_{2\alpha} \leq & k_\alpha \|\mathbf{W}\|_{\mathscr{C}} (1 + 2\|\dot{Y}\|_\infty)^{\beta_0 \vee \beta_1} (1 + 2\|Y\|_\infty)^{\beta_1 \vee \beta_2} \\
& \times \left[1 + T^\alpha \left(\|Y\|_\alpha + \|Y\|_\alpha^2 + \|\dot{Y}\|_\alpha + \|R^Y\|_{2\alpha} \right) \right].
\end{aligned} \tag{3.2.30}$$

As a consequence, Z is controlled by W with the Gubinelli derivative $\dot{Z} = Y$. □

In the next proposition, we consider the stability of nonlinear rough integrals.

Proposition 3.2.13. *Let $\mathbf{W}, \widetilde{\mathbf{W}} \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_3}([0,T]\times V;V)$. Suppose that $(Y,\dot{Y}) \in \mathscr{E}_W^{2\alpha}$ and $(\widetilde{Y},\dot{\widetilde{Y}}) \in \mathscr{E}_{\widetilde{W}}^{2\alpha}$. Define*

$$Z_t = \int_0^t W(dr,Y_r) \text{ and } \widetilde{Z}_t = \int_0^t \widetilde{W}(dr,\widetilde{Y}_r).$$

Then $(Z,Y) \in \mathscr{E}_W^{2\alpha}$ and $(\widetilde{Z},\widetilde{Y}) \in \mathscr{E}_{\widetilde{W}}^{2\alpha}$ by Proposition 3.2.12. In addition, the following inequality holds:

$$\begin{aligned} d_{\alpha,W,\widetilde{W}}\big((Z,Y),(\widetilde{Z},\widetilde{Y})\big) \leq & C_3\rho_{\alpha,\boldsymbol{\beta}_3}(\mathbf{W},\widetilde{\mathbf{W}}) + C_4\big(\|Y_0-\widetilde{Y}_0\|_V + \|\dot{Y}_0-\dot{\widetilde{Y}}_0\|_V\big) \\ & + C_5 d_{\alpha,W,\widetilde{W}}((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})), \end{aligned} \tag{3.2.31}$$

where

$$\begin{aligned} C_3 =& 2k_\alpha(1+T^\alpha)^2(1+\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3})(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2^{**}}(1+2\|\dot{Y}\|_\infty+2\|\dot{\widetilde{Y}}\|_\infty)^{\beta_2^*+\beta_0\vee\beta_1} \\ & \times\big[1+\|Y\|_\alpha+\|\dot{Y}\|_\alpha+\|\widetilde{Y}\|_\alpha+\|\dot{\widetilde{Y}}\|_\alpha+(\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)^2+\|R^Y\|_{2\alpha}\big], \\ C_4 =& 5k_\alpha(1+T^\alpha)^2(\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3}+\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3}^2)(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2^{**}}(1+2\|\dot{Y}\|_\infty+2\|\dot{\widetilde{Y}}\|_\infty)^{\beta_2^*+\beta_1} \\ & \times\big[1+\|Y\|_\alpha+\|\dot{Y}\|_\alpha+\|\widetilde{Y}\|_\alpha+\|\dot{\widetilde{Y}}\|_\alpha+(\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)^2+\|R^Y\|_{2\alpha}\big] \\ C_5 =& 6k_\alpha T^\alpha(1+T^\alpha)(1+\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3})^2(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2^{**}}(1+2\|\dot{Y}\|_\infty+2\|\dot{\widetilde{Y}}\|_\infty)^{\beta_2^*+\beta_1} \\ & \times\big[1+T^\alpha(\|Y\|_\alpha+\|\dot{Y}\|_\alpha+\|\widetilde{Y}\|_\alpha+\|\dot{\widetilde{Y}}\|_\alpha)+T^{2\alpha}(\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)^2+T^{2\alpha}\|R^Y\|_{2\alpha}\big], \end{aligned}$$

$\beta_2^* = \max\{\beta_0,\beta_1,\beta_2\}$ *and* $\beta_2^{**} = \max\{\beta_1,\beta_2,\beta_3\}$.

Proof. Due to Lemma 3.2.7, it suffices to estimate $\|R^Z - R^{\widetilde{Z}}\|_{2\alpha}$. Let Ξ and $\widetilde{\Xi}$ be the approximations of Z and $\widetilde{Z}$ respectively. That is,

$$\Xi_{s,t} = W_{s,t}(Y_s) + \mathbb{W}_{s,t}(\dot{Y}_s,Y_s) \quad \text{and} \quad \widetilde{\Xi}_{s,t} = \widetilde{W}_{s,t}(\widetilde{Y}_s) + \widetilde{\mathbb{W}}_{s,t}(\dot{\widetilde{Y}}_s,\widetilde{Y}_s).$$

Set $\Delta = \Xi - \widetilde{\Xi}$. Then by Proposition 3.2.12, we know that

$$\|R^Z_{s,t} - R^{\widetilde{Z}}_{s,t}\|_V \leq \|Z_{s,t} - \widetilde{Z}_{s,t} - \Delta_{s,t}\|_V + \|\mathbb{W}_{s,t}(\dot{Y}_s,Y_s) - \widetilde{\mathbb{W}}_{s,t}(\dot{\widetilde{Y}}_s,\widetilde{Y}_s)\|_V. \tag{3.2.32}$$

Due to the Sewing Lemma 3.1.5, to estimate the first term on the right-hand side of (3.2.32), it suffices to estimate $\|\delta\Delta\|_V$. Taking into account formulas (3.2.22) - (3.2.24), we can write

$$\begin{aligned}-\delta\Delta_{s,u,t} =& \big[DW_{u,t}(Y_s)(R^Y_{s,u}) - D\widetilde{W}_{u,t}(\widetilde{Y}_s)(R^{\widetilde{Y}}_{s,u})\big] + \big[\mathscr{R}^W_{u,t}(Y_s,Y_u) - \mathscr{R}^{\widetilde{W}}_{u,t}(\widetilde{Y}_s,\widetilde{Y}_u)\big]\\ &+ \big[\mathscr{D}^{\mathbb{W}}_{u,t}\big((\dot{Y}_s,Y_s),(\dot{Y}_u,Y_u)\big) - \mathscr{D}^{\widetilde{\mathbb{W}}}_{u,t}\big((\dot{\widetilde{Y}}_s,\widetilde{Y}_s),(\dot{\widetilde{Y}}_u,\widetilde{Y}_u)\big)\big]\\ :=& J_1 + J_2 + J_3, \end{aligned} \tag{3.2.33}$$

where $\mathscr{R}^W$, $\mathscr{R}^{\widetilde{W}}$, $\mathscr{D}^W$ and $\mathscr{D}^{\widetilde{W}}$ are defined as in (3.2.4) and (3.2.5), respectively.

Estimates for J_1: Triangular inequality implies that

$$\begin{aligned}\|J_1\|_V \leq& \big\|DW_{u,t}(Y_s)(R^Y_{s,u}) - D\widetilde{W}_{u,t}(Y_s)(R^Y_{s,u})\big\|_V\\ &+ \big\|D\widetilde{W}_{u,t}(Y_s)(R^Y_{s,u}) - D\widetilde{W}_{u,t}(\widetilde{Y}_s)(R^Y_{s,u})\big\|_V\\ &+ \big\|D\widetilde{W}_{u,t}(\widetilde{Y}_s)(R^Y_{s,u}) - D\widetilde{W}_{u,t}(\widetilde{Y}_s)(R^{\widetilde{Y}}_{s,u})\big\|_V\\ \leq& \|W-\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}(1+\|Y\|_\infty)^{\beta_1}\|R^Y\|_{2\alpha}|t-s|^{3\alpha}\\ &+ \|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}(1+\|Y\|_\infty+\|\widetilde{Y}\|_\infty)^{\beta_2}\|Y_s-\widetilde{Y}_s\|_V\|R^Y\|_{2\alpha}|t-s|^{3\alpha}\\ &+ \|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}(1+\|\widetilde{Y}\|_\infty)^{\beta_1}\|R^Y-R^{\widetilde{Y}}\|_{2\alpha}|t-s|^{3\alpha}. \end{aligned} \tag{3.2.34}$$

Plugging (3.2.16) into (3.2.34), we have

$$\begin{aligned}\|J_1\|_V \leq& \Big\{(1+T^\alpha)(1+\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3})(1+\|Y\|_\infty+\|\widetilde{Y}\|_\infty)^{\beta_1\vee\beta_2}(1+\|\dot{Y}\|_\infty)^{\beta_0}\\ &\quad\times\|R^Y\|_{2\alpha}\|W-\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}\\ &+ (1+T^\alpha)(\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}+\|\widetilde{W}\|^2_{\alpha,\boldsymbol{\beta}_3})(1+\|Y\|_\infty+\|\widetilde{Y}\|_\infty)^{\beta_1\vee\beta_2}\\ &\quad\times(1+\|\dot{Y}\|_\infty+\|\dot{\widetilde{Y}}\|_\infty)^{\beta_1}\|R^Y\|_{2\alpha}\big(\|Y_0-\widetilde{Y}_0\|_V+\|\dot{Y}_0-\dot{\widetilde{Y}}_0\|_V\big)\\ &+ (\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}+\|\widetilde{W}\|^2_{\alpha,\boldsymbol{\beta}_3})(1+\|Y\|_\infty+\|\widetilde{Y}\|_\infty)^{\beta_1\vee\beta_2}(1+\|\dot{Y}\|_\infty+\|\dot{\widetilde{Y}}\|_\infty)^{\beta_1}\\ &\quad\times(1+T^{2\alpha}\|R^Y\|_{2\alpha})d_{\alpha,W,\widetilde{W}}\big((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})\big)\Big\}|t-s|^{3\alpha}. \end{aligned} \tag{3.2.35}$$

Estimates for J_2: In order to bound J_2, we decompose J_2 as follows

$$J_2 = \mathscr{R}^{W-\widetilde{W}}_{u,t}(Y_s,Y_u) + [\mathscr{R}^{\widetilde{W}}_{u,t}(Y_s,Y_u) - \mathscr{R}^{\widetilde{W}}_{u,t}(\widetilde{Y}_s,\widetilde{Y}_u)] := J_2^1 + J_2^2.$$

By (3.2.6), we can write

$$\|J_2^1\|_V \leq \frac{1}{2}(1+2\|Y\|_\infty)^{\beta_2}\|Y\|_\alpha^2\|W-\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}|t-s|^{3\alpha}. \tag{3.2.36}$$

Thus, using Taylor's Theorem 3.1.11 and inequality (3.2.8), we have

$$\begin{aligned}
\|J_2^2\|_V \leq & \sup_{\tau\in[0,1]} \left\|D\mathscr{R}^{\widetilde{W}}_{u,t}\big(\tau Y_s+(1-\tau)\widetilde{Y}_s, \tau Y_u+(1-\tau)\widetilde{Y}_u\big)\big(Y_s-\widetilde{Y}_s, Y_u-\widetilde{Y}_u\big)\right\|_V \\
\leq & |t-s|^\alpha\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2\vee\beta_3} \\
& \times \left[\|\tau Y_{s,u}+(1-\tau)\widetilde{Y}_{s,u}\|_V^2\|Y_u-\widetilde{Y}_u\|_V + \|\tau Y_{s,u}+(1-\tau)\widetilde{Y}_{s,u}\|_V\|(Y-\widetilde{Y})_{s,u}\|_V\right] \\
\leq & \|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2\vee\beta_3} \\
& \times \left[b(\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)^2\|Y-\widetilde{Y}\|_\infty + (\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)\|Y-\widetilde{Y}\|_\alpha\right]|t-s|^{3\alpha}.
\end{aligned}$$

Applying (3.2.15) and (3.2.16), and putting together the terms with $\|W-\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}$, $\big(\|Y_0-\widetilde{Y}_0\|_V + \|\dot{Y}_0-\dot{\widetilde{Y}}_0\|_V\big)$ and $d_{\alpha,W,\widetilde{W}}\big((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})\big)$, respectively, we have

$$\begin{aligned}
\|J_2^2\|_V \leq & \Big[F_1\times\|W-\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3} + F_2\times\big(\|Y_0-\widetilde{Y}_0\|_V+\|\dot{Y}_0-\dot{\widetilde{Y}}_0\|_V\big) \\
& + F_3\times d_{\alpha,W,\widetilde{W}}\big((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})\big)\Big]|t-s|^{3\alpha},
\end{aligned} \tag{3.2.37}$$

where

$$\begin{aligned}
F_1 = & \|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2\vee\beta_3}(1+\|\dot{Y}\|_\infty)^{\beta_0} \\
& \times\big((\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)+T^\alpha(\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)^2\big),
\end{aligned}$$

$$
\begin{aligned}
F_2 =&(1+T^{\alpha})(\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}+\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}^2)(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2\vee\beta_3}(1+\|\dot{Y}\|_\infty+\|\dot{\widetilde{Y}}\|_\infty)^{\beta_1}\\
&\times\big((\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)+(\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)^2\big)
\end{aligned}
$$

and

$$
\begin{aligned}
F_3 =&(\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}+\|\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3}^2)(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2\vee\beta_3}(1+\|\dot{Y}\|_\infty+\|\dot{\widetilde{Y}}\|_\infty)^{\beta_1}\\
&\times\big(T^{\alpha}(\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)+T^{2\alpha}(\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)^2\big).
\end{aligned}
$$

Estimates for J_3: Similarly, we decompose J_3 as follows

$$
\begin{aligned}
J_3 =&\mathscr{D}_{u,t}^{\mathbb{W}-\widetilde{\mathbb{W}}}\big((\dot{Y}_s,Y_s),(\dot{Y}_u,Y_u)\big)+\big[\mathscr{D}_{u,t}^{\widetilde{\mathbb{W}}}\big((\dot{Y}_s,Y_s),(\dot{Y}_u,Y_u)\big)-\mathscr{D}_{u,t}^{\widetilde{\mathbb{W}}}\big((\dot{\widetilde{Y}}_s,\widetilde{Y}_s),(\dot{\widetilde{Y}}_u,\widetilde{Y}_u)\big)\big]\\
:=&J_3^1+J_3^2,
\end{aligned}
$$

The estimate for J_3^1 can by obtained by inequality (3.2.7), that is

$$
\begin{aligned}
\|J_3^1\|_V \leq&(1+2\|\dot{Y}\|_\infty)^{\beta_0\vee\beta_1}(1+2\|Y\|_\infty)^{\beta_1\vee\beta_2}(\|\dot{Y}\|_\alpha+\|Y\|_\alpha)\\
&\times\|\mathbb{W}-\widetilde{\mathbb{W}}\|_{2\alpha,\boldsymbol{\beta}_2^*,\boldsymbol{\beta}_2^{**}}|t-s|^{3\alpha}.
\end{aligned}
\tag{3.2.38}
$$

To bound J_3^2, we apply Taylor's Theorem 3.1.11 and get

$$
\|J_3^2\|_V \leq \sup_{0\leq\tau\leq 1}\big\|D\mathscr{D}_{u,t}^{\widetilde{\mathbb{W}}}(\boldsymbol{\xi}(\tau))\big(\dot{Y}_s-\dot{\widetilde{Y}}_s,Y_s-\widetilde{Y}_s,\dot{Y}_u-\dot{\widetilde{Y}}_u,Y_u-\widetilde{Y}_u\big)\big\|_V,
$$

where $\boldsymbol{\xi}(\tau)=\tau(\dot{Y}_s,Y_s,\dot{Y}_u,Y_u)+(1-\tau)(\dot{\widetilde{Y}}_s,\widetilde{Y}_s,\dot{\widetilde{Y}}_u,\widetilde{Y}_u)$. Therefore, using inequalities (3.2.9), (3.2.15) and (3.2.16), we can show that

$$
\begin{aligned}
\|J_3^2\|_V \leq & \Big[(1+T^\alpha)\|\widetilde{\mathbb{W}}\|_{2\alpha,\boldsymbol{\beta}_2^*,\boldsymbol{\beta}_2^{**}}(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2^{**}}(1+2\|\dot{Y}\|_\infty+2\|\dot{\widetilde{Y}}\|_\infty)^{\beta_2^*+\beta_0} \\
& \times(\|Y\|_\alpha+\|\dot{Y}\|_\alpha+\|\widetilde{Y}\|_\alpha+\|\dot{\widetilde{Y}}\|_\alpha)\|W-\widetilde{W}\|_{\alpha,\boldsymbol{\beta}_3} \\
& +3(1+T^\alpha)(\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3}+\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3}^2)(\|Y\|_\alpha+\|\dot{Y}\|_\alpha+\|\widetilde{Y}\|_\alpha+\|\dot{\widetilde{Y}}\|_\alpha) \\
& \times(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2^{**}}(1+2\|\dot{Y}\|_\infty+2\|\dot{\widetilde{Y}}\|_\infty)^{\beta_2^*+\beta_1} \\
& \times\big(\|Y_0-\widetilde{Y}_0\|_V+\|\dot{Y}_0-\dot{\widetilde{Y}}_0\|_V\big) \\
& +2(1+T^\alpha)(\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3}+\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3}^2)\big(1+T^\alpha(\|Y\|_\alpha+\|\dot{Y}\|_\alpha+\|\widetilde{Y}\|_\alpha+\|\dot{\widetilde{Y}}\|_\alpha)\big) \\
& \times(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2^{**}}(1+2\|\dot{Y}\|_\infty+2\|\dot{\widetilde{Y}}\|_\infty)^{\beta_2^*+\beta_1} \\
& \times d_{\alpha,W,\widetilde{W}}\big((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})\big)\Big]|t-s|^{3\alpha}.
\end{aligned}
\tag{3.2.39}
$$

Therefore, combining (3.2.33) and (3.2.35) - (3.2.39), we have

$$
\begin{aligned}
& \|\delta\Delta_{s,u,t}\|_V \\
\leq & \Big\{(1+T^\alpha)(1+\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3})(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2^{**}}(1+2\|\dot{Y}\|_\infty+2\|\dot{\widetilde{Y}}\|_\infty)^{\beta_2^*+\beta_0} \\
& \times\big[\|Y\|_\alpha+\|\dot{Y}\|_\alpha+\|\widetilde{Y}\|_\alpha+\|\dot{\widetilde{Y}}\|_\alpha+(\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)^2+\|R^Y\|_{2\alpha}\big]\rho_{\alpha,\boldsymbol{\beta}_3}(\mathbf{W},\widetilde{\mathbf{W}}) \\
& +4(1+T^\alpha)(\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3}+\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3}^2)(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2^{**}}(1+2\|\dot{Y}\|_\infty+2\|\dot{\widetilde{Y}}\|_\infty)^{\beta_2^*+\beta_1} \\
& \times\big[\|Y\|_\alpha+\|\dot{Y}\|_\alpha+\|\widetilde{Y}\|_\alpha+\|\dot{\widetilde{Y}}\|_\alpha+(\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)^2+\|R^Y\|_{2\alpha}\big] \\
& \times\big(\|Y_0-\widetilde{Y}_0\|_V+\|\dot{Y}_0-\dot{\widetilde{Y}}_0\|_V\big) \\
& +4(1+T^\alpha)(\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3}+\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3}^2)(1+2\|Y\|_\infty+2\|\widetilde{Y}\|_\infty)^{\beta_2^{**}}(1+2\|\dot{Y}\|_\infty+2\|\dot{\widetilde{Y}}\|_\infty)^{\beta_2^*+\beta_1} \\
& \times\big[1+T^\alpha(\|Y\|_\alpha+\|\dot{Y}\|_\alpha+\|\widetilde{Y}\|_\alpha+\|\dot{\widetilde{Y}}\|_\alpha)+T^{2\alpha}(\|Y\|_\alpha+\|\widetilde{Y}\|_\alpha)^2+T^{2\alpha}\|R^Y\|_{2\alpha}\big] \\
& \times d_{\alpha,W,\widetilde{W}}\big((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})\big)\Big\}|t-s|^{3\alpha}.
\end{aligned}
\tag{3.2.40}
$$

On the other hand, by (3.2.7) and (3.2.16), we can show that

$$
\begin{aligned}
&\|\mathbb{W}_{s,t}(\dot{Y}_s,Y_s)-\widetilde{\mathbb{W}}_{s,t}(\dot{\widetilde{Y}}_s,\widetilde{Y}_s)\|_V=\big\|(\mathbb{W}-\widetilde{\mathbb{W}})_{s,t}(\dot{Y}_s,Y_s)-\mathscr{D}^{\widetilde{\mathbb{W}}}_{s,t}\big((\dot{Y}_s,Y_s),(\dot{\widetilde{Y}}_s,\widetilde{Y}_s)\big)\big\|_V\\
\le&\Big\{(1+T^\alpha)(1+\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3})(1+\|\dot{Y}\|_\infty+\|\dot{\widetilde{Y}}\|_\infty)^{2\beta_0\vee\beta_1}(1+\|Y\|_\infty+\|\widetilde{Y}\|_\infty)^{\beta_1\vee\beta_2}\\
&\quad\times\rho_{\alpha,\boldsymbol{\beta}_3}(\mathbf{W},\widetilde{\mathbf{W}})\\
&+2(1+T^\alpha)(\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3}+\|\widetilde{\mathbf{W}}\|^2_{\mathscr{C}_3})(1+\|\dot{Y}\|_\infty+\|\dot{\widetilde{Y}}\|_\infty)^{\beta_0\vee\beta_1+\beta_1}(1+\|Y\|_\infty+\|\widetilde{Y}\|_\infty)^{\beta_1\vee\beta_2}\\
&\quad\times\big(\|Y_0-\widetilde{Y}_0\|_V+\|\dot{Y}_0-\dot{\widetilde{Y}}_0\|_V\big)\\
&+2T^\alpha(1+T^\alpha)(\|\widetilde{\mathbf{W}}\|_{\mathscr{C}_3}+\|\widetilde{\mathbf{W}}\|^2_{\mathscr{C}_3})(1+\|\dot{Y}\|_\infty+\|\dot{\widetilde{Y}}\|_\infty)^{\beta_0\vee\beta_1+\beta_1}(1+\|Y\|_\infty+\|\widetilde{Y}\|_\infty)^{\beta_1\vee\beta_2}\\
&\quad\times d_{\alpha,W,\widetilde{W}}\big((Y,\dot{Y}),(\widetilde{Y},\dot{\widetilde{Y}})\big)\Big\}|t-s|^{2\alpha}.
\end{aligned}
\tag{3.2.41}
$$

Recall inequality (3.2.32). Inequality (3.2.31) follows from (3.2.15), (3.2.40), (3.2.41) and the Sewing Lemma. This completes the proof of the proposition. □

3.3 Rough Differential Equations

Let $\alpha\in(\frac{1}{3},\frac{1}{2}]$, $\boldsymbol{\beta}_3=(\beta_0,\dots,\beta_3)$ where $\beta_k\ge 0$, $k=0,\dots,3$, and let $\mathbf{W}=(W,\mathbb{W})\in\mathscr{C}^{\alpha,\boldsymbol{\beta}_3}([0,T]\times V;V)$. That is, W is α-Hölder in time, and three times differentiable in space with growth multi-index $\boldsymbol{\beta}_3$, $\mathbb{W}$ is 2α-Hölder in time and twice differentiable in space with growth multi-indexes $\boldsymbol{\beta}_2^*=(\beta_0,\beta_0\vee\beta_1,\beta_0\vee\beta_1\vee\beta_2)$ and $\boldsymbol{\beta}_2^{**}=(\beta_1,\beta_1\vee\beta_2,\beta_1\vee\beta_2\vee\beta_3)$, and $(W,\mathbb{W})$ satisfies Chen's relation (3.2.10). In this section, we study the following nonlinear RDE:

$$
Y_t=\xi+\int_0^t W(dr,Y_r). \tag{3.3.1}
$$

Definition 3.3.1. *An α-Hölder continuous function Y is said to be a solution to (3.3.1), if $(Y,Y)\in\mathscr{E}^{2\alpha}_{W,(\xi,\xi)}$, and equality (3.3.1) holds for all $t\in[0,T]$ where the integral on the right-hand side is a nonlinear rough integral in the sense of Theorem 3.2.10.*

3.3.1 Local existence

In this section, we establish the (local) existence of a solution for equation (3.3.1) using Picard iteration method. To this end, we introduce the following notation. Let $\Phi \in \mathscr{C}^{\alpha}([0,T];V)$. For any $0 \le s < t \le T$, we write

$$\|\Phi\|_{\alpha,[s,t]} := \sup_{u \neq v \in [s,t]} \frac{\|\Phi_{u,v}\|_V}{|v-u|^{\alpha}}.$$

We also define $d_{W,\alpha,[s,t]}$ in a similar way, where we recall Remark 3.2.8 for the definition of $d_{W,\alpha}$.

Theorem 3.3.2. *For any $\xi \in V$, there exist a positive number h, such that the RDE (3.3.1) has a solution Y on $[0,h]$ with initial condition $Y_0 = \xi$. In addition, the following inequality holds on $[0,h]$:*

$$\|Y\|_{\alpha,[0,h]} \le 5^{2\gamma_1+2} k_{\alpha}(1+\|\mathbf{W}\|_{\mathscr{C}_3})(1+\|\xi\|_V)^{\gamma_1}, \tag{3.3.2}$$

where $\gamma_1 = \beta_0 \vee \beta_1 + \beta_1 \vee \beta_2$.

Proof. Choose $h \in (0,1]$. Let

$$(Y_t^0, \dot{Y}_t^0) := (\xi + W_{0,t}(\xi), \xi), \quad t \in [0,h].$$

Then $(Y^0, \dot{Y}^0) \in \mathscr{E}^{2\alpha}_{W,(\xi,\xi)}$ with the remainder $R^{Y^0}_{s,t} \equiv 0$ for all $(s,t) \in [0,h]^2$. Due to Proposition 3.2.12, for any $n \ge 1$, we can recursively define an element $(Y^n, \dot{Y}^n) \in \mathscr{E}^{2\alpha}_{W,(\xi,\xi)}$ given by

$$Y_t^{n+1} = \xi + \int_0^t W(dr, Y_t^n), \quad t \in [0,h].$$

By (3.2.30), the following inequality holds for all $n \ge 1$

$$\begin{aligned} \|R^{Y^{n+1}}\|_{2\alpha,[0,h]} \le & k_{\alpha}\|\mathbf{W}\|_{\mathscr{C}_3}(1+2\|\xi\|_V)^{\gamma_1}\left[1+2h^{\alpha}\left(\|Y^n\|_{\alpha,[0,h]}+\|Y^{n-1}\|_{\alpha,[0,h]}\right)\right]^{\gamma_1} \\ & \times \left[1+h^{\alpha}\left(\|Y^n\|_{\alpha,[0,h]}+\|Y^n\|^2_{\alpha,[0,h]}+\|Y^{n-1}\|_{\alpha,[0,h]}+\|R^{Y^n}\|_{2\alpha,[0,h]}\right)\right]. \end{aligned} \tag{3.3.3}$$

By iteration, we know that $(Y^{n+1}, Y^n) \in \mathscr{E}_W^{2\alpha,(\xi,\xi)}$, which implies that

$$\|Y^{n+1}\|_{\alpha,[0,h]} \leq \|W\|_{\alpha,\boldsymbol{\beta}_3}(1+\|\xi\|_V + h^\alpha \|Y^n\|_{\alpha,[0,h]})^{\beta_0} + h^\alpha \|R^{Y^{n+1}}\|_{2\alpha,[0,h]}. \tag{3.3.4}$$

Choose $h_1 = [5^{\gamma_1+2} k_\alpha (1+\|\mathbf{W}\|_{\mathscr{C}_3})(1+2\|\xi\|_V)^{\gamma_1}]^{-\frac{1}{\alpha}}$. We claim that for any $h \in [0,h_1] \subset [0,1]$, $\|Y^n\|_{\alpha,[0,h]}$ and $\|R^n\|_{2\alpha,[0,h]}$ are bounded uniformly in n. To prove this claim, for any $h \in [0,h_1]$, let $f_h, g_h : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ be given by

$$f_h(x,y) = k_\alpha \|\mathbf{W}\|_{\mathscr{C}_3}(1+2\|\xi\|_V)^{\gamma_1}(1+4h^\alpha y)^{\gamma_1}[1+h^\alpha(2y+y^2+x)]$$

and

$$g_h(x,y) = \|W\|_{\alpha,\boldsymbol{\beta}_3}(1+\|\xi\|_V + h^\alpha y)^{\beta_0} + h^\alpha x.$$

Then it is easy to see that f and g are both increasing in each argument h, x and y. Let

$$x_1 = \frac{\|\mathbf{W}\|_{\mathscr{C}_3}}{2(1+\|\mathbf{W}\|_{\mathscr{C}_3})} h_1^{-2\alpha} \quad \text{and} \quad y_1 = \frac{\|\mathbf{W}\|_{\mathscr{C}_3}}{1+\|\mathbf{W}\|_{\mathscr{C}_3}} h_1^{-\alpha}.$$

It follows that for any $h \in [0,h_1] \subset [0,1]$, $x \in [0,x_1]$ and $y \in [0,y_1]$, the following inequalities hold

$$\begin{aligned}
f_h(x,y) \leq & f_{h_1}(x_1,y_1) \\
= & k_\alpha \|\mathbf{W}\|_{\mathscr{C}_3}(1+2\|\xi\|_V)^{\gamma_1} 5^{\gamma_1}\left(1+\frac{2\|\mathbf{W}\|_{\mathscr{C}_3}}{1+\|\mathbf{W}\|_{\mathscr{C}_3}} + \frac{\|\mathbf{W}\|_{\mathscr{C}_3}^2 h^{-\alpha}}{(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2} + \frac{\|\mathbf{W}\|_{\mathscr{C}_3} h_1^{-\alpha}}{2(1+\|\mathbf{W}\|_{\mathscr{C}_3})}\right) \\
\leq & k_\alpha \|\mathbf{W}\|_{\mathscr{C}_3}(1+2\|\xi\|_V)^{\gamma_1} 5^{\gamma_1+1} h_1^{-\alpha} = \frac{\|\mathbf{W}\|_{\mathscr{C}_3}}{5(1+\|\mathbf{W}\|_{\mathscr{C}_3})} h_1^{-2\alpha} \leq x_1
\end{aligned}$$

and

$$g_h(x,y) \leq g_{h_1}(x_1,y_1) \leq 2\|W\|_{\alpha,\boldsymbol{\beta}_3}(1+\|\xi\|_V)^{\beta_0} + \frac{\|\mathbf{W}\|_{\mathscr{C}_3} h_1^{-\alpha}}{2(1+\|\mathbf{W}\|_{\mathscr{C}_3})} \leq y_1.$$

From inequalities (3.3.3) and (3.3.4) we can show, by a recursive argument, that

$$\max_{n\geq 0}\{\|Y^n\|_{\alpha,[0,h]}\} \leq g_{h_1}(x_1,y_1) = 5^{\gamma_1+2}k_\alpha\|\mathbf{W}\|_{\mathscr{C}_3}(1+2\|\xi\|_V)^{\gamma_1}$$

$$\leq 5^{2\gamma_1+2}k_\alpha\|\mathbf{W}\|_{\mathscr{C}_3}(1+\|\xi\|_V)^{\gamma_1} \tag{3.3.5}$$

and

$$\max_{n\geq 0}\{\|R^{Y^n}\|_{2\alpha,[0,h]}\} \leq x_1 = \frac{5^{2\gamma_1+4}}{2}k_\alpha^2\|\mathbf{W}\|_{\mathscr{C}_3}(1+\|\mathbf{W}\|_{\mathscr{C}_3})(1+2\|\xi\|_V)^{2\gamma_1}$$

$$\leq 5^{3\gamma_1+4}k_\alpha^2\|\mathbf{W}\|_{\mathscr{C}_3}(1+\|\mathbf{W}\|_{\mathscr{C}_3})(1+\|\xi\|_V)^{2\gamma_1}, \tag{3.3.6}$$

provided that $\|Y^0\|_{\alpha,[0,h]}, \|\dot{Y}^0\|_{\alpha,[0,h]} \leq y_1$ and $\|R^{Y^0}\|_{2\alpha,[0,h]} \leq x_1$. Indeed, by definition, we know that $\|\dot{Y}^0\|_{\alpha,[0,h]} = \|R^{Y^0}\|_{2\alpha,[0,h]} = 0$, and

$$\|Y^0\|_{\alpha,[0,h]} \leq \|W\|_{\alpha,\boldsymbol{\beta}_3}(1+\|\xi\|_V)^{\gamma_1} \leq y_1.$$

As a consequence, we conclude that $\|Y^n\|_{\alpha,[0,h]}$ and $\|R^n\|_{2\alpha,[0,h]}$ are bounded uniformly in n for $h\in(0,h_1]$. This also yields that

$$\max_{n\geq 0}\{\|Y^n\|_{\infty,[0,h]}\} \leq \frac{\|\mathbf{W}\|_{\mathscr{C}_3}}{1+\|\mathbf{W}\|_{\mathscr{C}_3}} + \|\xi\|_V \leq 1+\|\xi\|_V. \tag{3.3.7}$$

By (3.3.5), (3.3.7), Proposition 3.2.13 and the fact that $0 < h \leq h_1 = [5^{\gamma_1+2}k_\alpha(1+\|\mathbf{W}\|_{\mathscr{C}_3})(1+2\|\xi\|_V)^{\gamma_1}]^{-\frac{1}{\alpha}} < 1$, we get the following estimate

$$d_{\alpha,W,[0,h]}\big((Y^{n+1},Y^n),(Y^n,Y^{n-1})\big) \leq C_5 d_{\alpha,W,[0,h]}\big((Y^n,Y^{n-1}),(Y^{n-1},Y^{n-2})\big),$$

where

$$
\begin{aligned}
C_5 =&6k_\alpha h^\alpha(1+h^\alpha)(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2(1+2\|Y^n\|_{\infty,[0,h]}+2\|Y^{n-1}\|_{\infty,[0,h]})^{\beta_2^{**}}\\
&\times(1+2\|Y^{n-1}\|_{\infty,[0,h]}+2\|Y^{n-2}\|_{\infty,[0,h]})^{\beta_2^*+\beta_1}\\
&\times\Big[1+h^\alpha(2\|Y^{n-1}\|_{\alpha,[0,h]}+\|Y^n\|_{\alpha,[0,h]}+\|Y^{n-2}\|_{\alpha,[0,h]})\\
&\quad+h^{2\alpha}(\|Y^{n-1}\|_{\alpha,[0,h]}+\|Y^n\|_{\alpha,[0,h]})^2+h^{2\alpha}\|R^{Y^n}\|_{2\alpha,[0,h]}\Big]\\
\leq&120\times 5^{\beta_2^{**}+\beta_2^*+\beta_1}k_\alpha(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2(1+\|\xi\|_V)^{\beta_2^{**}+\beta_2^*+\beta_1}h^\alpha.
\end{aligned}
$$

Let $\gamma_2=\beta_2^{**}+\beta_2^*+\beta_1=\max\{\beta_0,\beta_1,\beta_2\}+\max\{\beta_1,\beta_2,\beta_3\}+\beta_1$, and let $C_6=120\times 5^{\gamma_2}k_\alpha$. Then, we have

$$
\begin{aligned}
&d_{\alpha,W,[0,h]}\big((Y^{n+1},Y^n),(Y^n,Y^{n-1})\big)\\
\leq&C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2(1+\|\xi\|_V)^{\gamma_2}h^\alpha d_{\alpha,W,[0,h]}\big((Y^n,Y^{n-1}),(Y^{n-1},Y^{n-2})\big).
\end{aligned}\tag{3.3.8}
$$

Choose $h_2=[2C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2(1+\|\xi\|_V)^{\gamma_2}]^{-\frac{1}{\alpha}}\leq h_1\leq 1$, and let $h\in(0,h_2]$. Then by (3.3.8), we have the following inequality

$$
d_{\alpha,W}\big((Y^{n+1},Y^n),(Y^n,Y^{n-1})\big)\leq\frac{1}{2}d_{\alpha,W}\big((Y^n,Y^{n-1}),(Y^{n-1},Y^{n-2})\big).
$$

This yields that

$$
\sum_{n=1}^{\infty}d_{\alpha,W,[0,h]}\big((Y^{n+1},Y^n),(Y^n,Y^{n-1})\big)<\infty.
$$

Due to Lemma 3.2.9, we can conclude that $(Y^n,Y^{n-1})\to(Y,Y)\in\mathscr{E}^{2\alpha}_{W,(\xi,\xi)}$ as $n\to\infty$. By Lemma 3.2.7 and Proposition 3.2.13, we have for any $0\leq s\leq t\leq h$,

$$
\begin{aligned}
\Big\|Y^{n+1}_{s,t}-\int_s^t W(dr,Y_r)\Big\|_V=&\Big\|\int_s^t W(dr,Y^n_r)-\int_s^t W(dr,Y_r)\Big\|_V\\
\leq&Cd_{\alpha,W,[0,h]}\big((Y^n,Y^{n-1}),(Y,Y)\big)|t-s|^\alpha,
\end{aligned}
$$

for some constant $C > 0$ uniformly in n. This implies that equation (3.3.1) holds for all $t \in [0,h]$. Finally, inequality (3.3.2) follows from (3.3.5) and the fact that (Y,Y) is the limit of (Y^n, Y^{n-1}) in $\mathscr{E}^{2\alpha}_{W,(\xi,\xi)}$. □

3.3.2 Uniqueness and global existence

In this section, we prove the uniqueness of a solution for equation (3.3.1). We also present some hypotheses that imply the global existence of a solution for this equation.

Theorem 3.3.3. *For any time interval $[0,T]$ and initial value $\xi \in V$. There exists at most one solution to equation (3.3.1).*

Proof. Suppose that Y and $\widetilde{Y}$ are two solutions to (3.3.1) with initial condition ξ on $[0,T]$. By Proposition 3.2.13, the following inequality holds on $[0,h] \subset [0,T]$, assuming $h \leq 1$.

$$d_{\alpha,W,[0,h]}\big((Y,Y),(\widetilde{Y},\widetilde{Y})\big) \leq C_5 d_{\alpha,W,[0,h]}\big((Y,Y),(\widetilde{Y},\widetilde{Y})\big), \tag{3.3.9}$$

where

$$\begin{aligned} C_5 =& 12k_\alpha h^\alpha (1+\|\mathbf{W}\|_{\mathscr{C}_3})^2 (1+2\|Y\|_\infty + 2\|\widetilde{Y}\|_\infty)^{\beta_2^{**}} (1+2\|\dot{Y}\|_\infty + 2\|\dot{\widetilde{Y}}\|_\infty)^{\beta_2^*+\beta_1} \\ &\times \big[1 + (\|Y\|_\alpha + \|\dot{Y}\|_\alpha + \|\widetilde{Y}\|_\alpha + \|\dot{\widetilde{Y}}\|_\alpha) + (\|Y\|_\alpha + \|\widetilde{Y}\|_\alpha)^2 + \|R^Y\|_{2\alpha}\big]. \end{aligned}$$

Choosing h small enough, (3.3.9) yields that $Y \equiv \widetilde{Y}$ on $[0,h]$. Notice that the choice of h doesn't dependent on the initial value. Therefore, by iteration, we can extend the uniqueness to any time interval $[0,T]$. □

As stated in Section 3.1, the linear growth of the vector field cannot guarantee the global existence of a RDE driven by a linear rough path. This is also true in the case of nonlinear rough paths. In order to obtain the global existence, we introduce the following growth condition of W. Let $\mathbf{W} = (W, \mathbb{W}) \in \mathscr{C}^{\alpha,\beta_3}([0,T] \times V; V)$, let $\gamma_1 = \beta_0 \vee \beta_1 + \beta_1 \vee \beta_2$, and let $\gamma_2 = \max\{\beta_0,\beta_1,\beta_2\} + \max\{\beta_1,\beta_2,\beta_3\} + \beta_1$.

Hypothesis (H). $\frac{\gamma_2}{\alpha} - \gamma_2 + \gamma_1 \leq 1$.

A similar condition in the linear situation can be seen, e.g., in [6, 19, 60].

Theorem 3.3.4. *Under Hypothesis (H), the RDE* (3.3.1) *has a solution on any time interval* $[0,T]$. *By Theorem 3.3.3, this solution is unique.*

Proof. Let $\varepsilon_1 = [2C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2(1+\|\xi\|_V)^{\gamma_2}]^{-\frac{1}{\alpha}}$ where $C_6 = 120\times 5^{\gamma_2}k_\alpha$ is the constant appearing in (3.3.8). Then, by Theorem 3.3.2, the RDE has a solution $Y^{(1)}$ on $[0,\varepsilon_1]$ with initial condition $Y^{(1)}_0 = \xi$. We denote by $\xi_1 = Y^{(1)}_{\varepsilon_1}$ the terminal value of Y. In order to extend the solution to the entire interval $[0,T]$, we consider the following RDE

$$Y_t = Y_s + \int_s^t W(dr, Y_r) + \int_s^t (dr, Y_r). \tag{3.3.10}$$

By Theorem 3.3.2 again, equation (3.3.10) has a solution $Y^{(2)}$ on $[\varepsilon_1, \varepsilon_1+\varepsilon_2]$ with initial condition $Y_{\varepsilon_1} = \xi_1$, where $\varepsilon_2 = [2C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2(1+\|\xi_1\|_V)^{\gamma_2}]^{-\frac{1}{\alpha}}$. By iteration, we have a sequence $\{\varepsilon_n\}_{n\geq 1}$ with values in $(0,1)$, such that the equation (3.3.10) has a solution $Y^{(n+1)}$ on $[\eta_n, \eta_{n+1}] := [\sum_{k=1}^n \varepsilon_k, \sum_{k=1}^{n+1} \varepsilon_{n+1}]$ with initial condition $Y^{(n+1)}_{\eta_n} = \xi_n := Y^{(n)}_{\eta_n}$ and $\varepsilon_{n+1} = [2C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2(1+\|\xi_n\|_V)^{\gamma_2}]^{-\frac{1}{\alpha}}$. By (3.3.2) we have the following inequality

$$\|\xi_{n+1}\|_V \leq \|Y^{(n+1)}\|_\infty \leq \|\xi_n\|_V + \varepsilon_{n+1}^\alpha \|Y^{(n+1)}\|_\alpha \leq \|\xi_n\|_V + \frac{5^{2\gamma_1+2}\|\mathbf{W}\|_{\mathscr{C}_3}}{2C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2}(1+\|\xi_n\|_V)^{\gamma_1-\gamma_2}.$$

Recall the assumption $\frac{\gamma_2}{\alpha} - \gamma_2 + \gamma_1 \leq 1$. By the mean value theorem for real valued functions, there exist $\tau \in [0,1]$, such that

$$\begin{aligned}(1+\|\xi_{n+1}\|_V)^{\frac{\gamma_2}{\alpha}} \leq &\left[1+\|\xi_n\|_V + \frac{5^{2\gamma_1+2}\|\mathbf{W}\|_{\mathscr{C}_3}}{2C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2}(1+\|\xi_n\|_V)^{\gamma_1-\gamma_2}\right]^{\frac{\gamma_2}{\alpha}} \\ =&(1+\|\xi_n\|_V)^{\frac{\gamma_2}{\alpha}} + \left[5^{2\gamma_1+2}(2C_6)^{-1}(1+\|\xi_n\|_V)^{\gamma_1-\gamma_2}\right] \\ &\times \frac{\gamma_2}{\alpha}\left[1+\|\xi_n\|_V + \tau 5^{2\gamma_1+2}(2C_6)^{-1}(1+\|\xi_n\|_V)^{\gamma_1-\gamma_2}\right]^{\frac{\gamma_2}{\alpha}-1}.\end{aligned} \tag{3.3.11}$$

By definition we know that $\gamma_1 \leq \gamma_2$. This implies

$$
\begin{aligned}
&\left[1+\|\xi_n\|_V+\tau 5^{2\gamma_1+2}(2C_6)^{-1}(1+\|\xi_n\|_V)^{\gamma_1-\gamma_2}\right]^{\frac{\gamma_2}{\alpha}-1}\\
&\quad\leq(1+\|\xi_n\|_V)^{\frac{\gamma_2}{\alpha}-1}\times\max\left\{1,\left[1+5^{2\gamma_1+2}(2C_6)^{-1}\right]^{\frac{\gamma_2}{\alpha}-1}\right\}\\
&\quad=\left[1+5^{2\gamma_1+2}(2C_6)^{-1}\right]^{0\vee(\frac{\gamma_2}{\alpha}-1)}(1+\|\xi_n\|_V)^{\frac{\gamma_2}{\alpha}-1}. \qquad (3.3.12)
\end{aligned}
$$

As a consequence of inequalities (3.3.11) and (3.3.12) , under the assumption (H), we can write

$$
\begin{aligned}
(1+\|\xi_{n+1}\|_V)^{\frac{\gamma_2}{\alpha}} &\leq(1+\|\xi_n\|_V)^{\frac{\gamma_2}{\alpha}}+\frac{\gamma_2}{\alpha}\left[1+5^{2\gamma_1+2}(2C_6)^{-1}\right]^{\frac{\gamma_2}{\alpha}\vee 1}(1+\|\xi_n\|_V)^{\gamma_1-\gamma_2+\frac{\gamma_2}{\alpha}-1}\\
&\leq(1+\|\xi_n\|_V)^{\frac{\gamma_2}{\alpha}}+\frac{\gamma_2}{\alpha}\left[1+5^{2\gamma_1+2}(2C_6)^{-1}\right]^{\frac{\gamma_2}{\alpha}\vee 1}. \qquad (3.3.13)
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\varepsilon_{n+1} &\geq[2C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2]^{-\frac{1}{\alpha}}\left[(1+\|\xi_{n-1}\|_V)^{\frac{\gamma_2}{\alpha}}+\frac{\gamma_2}{\alpha}\left[1+5^{2\gamma_1+2}(2C_6)^{-1}\right]^{\frac{\gamma_2}{\alpha}\vee 1}\right]^{-1} \qquad (3.3.14)\\
&=\left[\varepsilon_n^{-1}+\left(2C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2\right)^{\frac{1}{\alpha}}\frac{\gamma_2}{\alpha}\left[1+5^{2\gamma_1+2}(2C_6)^{-1}\right]^{\frac{\gamma_2}{\alpha}\vee 1}\right]^{-1}:=(\varepsilon_n^{-1}+K_0)^{-1}.
\end{aligned}
$$

Observe that the constant K_0 is independent of n. Thus by iteration, the following inequality holds

$$
\sum_{n=1}^{\infty}\varepsilon_n\geq\sum_{n=0}^{\infty}\frac{1}{\varepsilon_1^{-1}+nK_0}=\infty. \qquad (3.3.15)
$$

In other words, we can extend the solution to any time interval $[0,T]$. □

Assume that the derivatives of W are all bounded, that is $\boldsymbol{\beta}_3=(\beta_0,0,0,0)$. Then, Hypothesis (H) is equivalent to $\beta_0\leq\alpha$ and it coincides with Besalú and Nualart's condition for global existence (see Theorem 4.1 of Besalu & Nualart [6]).

3.3.3 Properties of the solutions

Assume Hypothesis (H). In this section, we prove some properties of the solution to the RDE (3.3.1). The first proposition below provides an estimate for the Hölder norm of the solution to (3.3.1). Before stating the proposition, we first prove the following lemma.

Lemma 3.3.5. *Suppose that $X \in \mathscr{C}^{\alpha}([0,T];V)$. Let $\pi = (0 = t_0 < t_1 < t_2 < \dots, t_n = T)$ be a partition. Then,*

$$\|X\|_{\alpha} \le n^{1-\alpha} \max_{1\le k\le n} \|X\|_{\alpha,[t_k - t_{k-1}]} \le (T/|\pi|)^{1-\alpha} \max_{1\le k\le n} \|X\|_{\alpha,[t_k - t_{k-1}]}.$$

Proof. For any $0 \le s < t \le T$. There exists $0 \le k_1 \le k_2 \le n$ such that $s \le k_1 \le k_2 \le t$. Then by Jensen's inequality for convex function $f(x) = |x|^{\frac{1}{\alpha}}$, we have

$$\begin{aligned}
\frac{\|X_{s,t}\|_V}{|t-s|^{\alpha}} &\le \frac{\|X_{s,t_{k_1}}\|_V + \|X_{t_{k_1},t_{k_1+1}}\|_V + \cdots + \|X_{t_{k_2},t}\|_V}{|t-s|^{\alpha}} \\
&\le \max_{1\le k\le n} \|X\|_{\alpha,[t_k - t_{k-1}]} \times \frac{|t_{k_1} - s|^{\alpha} + \cdots + |t - t_{k_2}|^{\alpha}}{|t-s|^{\alpha}} \\
&\le \max_{1\le k\le n} \|X\|_{\alpha,[t_k - t_{k-1}]} \times \frac{n^{1-\alpha}|t-s|^{\alpha}}{|t-s|^{\alpha}}.
\end{aligned}$$

The lemma is then proved. □

Proposition 3.3.6. *Assume Hypothesis (H). Let Y be the solution to the RDE (3.3.1) with initial condition $\xi \in V$. Then the following estimate holds:*

$$\|Y\|_{\alpha} \le c\|\mathbf{W}\|_{\mathscr{C}_3}(1 + \|\xi\|_V)^{\gamma_1 + \frac{1-\alpha}{\alpha}\gamma_2} e^{(\frac{\alpha\gamma_1}{\gamma_2} + 1 - \alpha)K_0 T} \tag{3.3.16}$$

for some c depending on α and $\boldsymbol{\beta}_3$, where

$$K_0 = \left(2C_6(1 + \|\mathbf{W}\|_{\mathscr{C}_3})^2\right)^{\frac{1}{\alpha}} \frac{\gamma_2}{\alpha} \left[1 + 5^{2\gamma_1 + 2}(2C_6)^{-1}\right]^{\frac{\gamma_2}{\alpha} \vee 1}$$

and $C_6 = 120 \times 5^{\gamma_2} k_{\alpha}$ are the same as in (3.3.14) and (3.3.8), respectively.

Proof. Let $\varepsilon_1 = [2C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2(1+\|\xi\|_V)^{\gamma_2}]^{-\frac{1}{\alpha}}$. Theorems 3.3.2 and 3.3.3 imply that there exists a unique solution to (3.3.1) with initial condition $Y_0 = \xi$ on $[0,\varepsilon_1]$. Denote the solution by $Y^{(1)}$. Then, proceeding with a similar argument as in Theorem 3.3.4, we obtain a sequence $\{Y^{(n+1)}\}_{n\geq 1}$, where $Y^{(n+1)}$ is the unique solution to RDE (3.3.10) on $[\eta_n,\eta_{n+1}] = [\sum_{k=1}^n \varepsilon_k, \sum_{k=1}^{n+1}\varepsilon_k]$ with initial condition $Y_{\eta_n}^{(n+1)} := \xi_n = Y_{\eta_n}^{(n)}$ and $\varepsilon_{n+1} = [2C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2(1+\|\xi_n\|_V)^{\gamma_2}]^{-\frac{1}{\alpha}}$. By inequalities (3.3.2), (3.3.13) and an iteration argument, we have the following estimate:

$$\|Y^{(n+1)}\|_\alpha \leq 5^{2\gamma_1+2}k_\alpha\|\mathbf{W}\|_{\mathscr{C}_3}\Big\{(1+\|\xi\|_V)^{\frac{\gamma_2}{\alpha}} + \frac{(n+1)\gamma_2}{\alpha}\big[1+5^{2\gamma_1+2}(2C_6)^{-1}\big]^{\frac{\gamma_2}{\alpha}\vee 1}\Big\}^{\frac{\alpha\gamma_1}{\gamma_2}}. \quad (3.3.17)$$

In order to obtain (3.3.16), we consider the following two cases. Firstly, if $T \leq \varepsilon_1$, then (3.3.16) holds by taking $n=0$ in (3.3.17). On the other hand, for any $T > \varepsilon_1$, there exists a positive integer N, such that $\eta_N \leq T \leq \eta_{N+1}$. Notice that by (3.3.15), we have

$$T \geq \sum_{n=1}^N \varepsilon_n \geq \sum_{n=1}^N (\varepsilon_1^{-1}+K_0 n)^{-1} \geq \frac{1}{K_0}\big(\log(\varepsilon_1^{-1}+K_0N) - \log(\varepsilon_1^{-1})\big).$$

Recall that $\varepsilon_1 = [2C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2(1+\|\xi\|_V)^{\gamma_2}]^{-\frac{1}{\alpha}}$ and

$$K_0 = \big(2C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2\big)^{\frac{1}{\alpha}}\frac{\gamma_2}{\alpha}\big[1+5^{2\gamma_1+2}(2C_6)^{-1}\big]^{\frac{\gamma_2}{\alpha}\vee 1}.$$

It follows that

$$\begin{aligned} N \leq& \frac{1}{K_0}\big(e^{K_0T+\log(\varepsilon_1^{-1})} - \varepsilon_1^{-1}\big) = K_0^{-1}(e^{K_0T}-1)[2C_6(1+\|\mathbf{W}\|_{\mathscr{C}_3})^2(1+\|\xi\|_V)^{\gamma_2}]^{\frac{1}{\alpha}} \\ =& \frac{\alpha}{\gamma_2}\big[1+5^{2\gamma_1+2}(2C_6)^{-1}\big]^{-(\frac{\gamma_2}{\alpha}\vee 1)}(1+\|\xi\|_V)^{\frac{\gamma_2}{\alpha}}(e^{K_0T}-1). \end{aligned} \quad (3.3.18)$$

Let Y be the solution to (3.3.1) on $[0,T]$ with initial condition ξ. Then, combining Lemma 3.3.5, (3.3.17) and (3.3.18), we have

$$\|Y\|_\alpha \leq N^{1-\alpha}\max_{1\leq n\leq N+1}\|Y^{(n)}\|_\alpha \leq c(1+\|\xi\|_V)^{\gamma_1+\frac{1-\alpha}{\alpha}\gamma_2}e^{(\frac{\alpha\gamma_1}{\gamma_2}+1-\alpha)K_0T}$$

for some c depending on α and $\boldsymbol{\beta}_3$. This completes the proof of the proposition. □

The next proposition provides the dependency of the solution to (3.3.1) on the initial condition under Hypothesis (H).

Proposition 3.3.7. *Assume that $\mathbf{W} = (W, \mathbb{W})$ satisfies the conditions in Theorem 3.3.4. Let Y and $\widetilde{Y}$ be the solutions to the RDE (3.3.1) with initial conditions ξ and $\widetilde{\xi}$, respectively. Then the following estimate holds*

$$d_{\alpha,W}\big((Y,Y),(\widetilde{Y},\widetilde{Y})\big) \le c^T(T^{1-\alpha} \vee 1)\|\xi - \widetilde{\xi}\|_V, \tag{3.3.19}$$

where c is a constant depending on α, $\boldsymbol{\beta}_3$, $\|\mathbf{W}\|_{\mathscr{C}_3}$, ξ and $\widetilde{\xi}$.

Proof. By Propositions 3.2.13 and 3.3.6, and the fact that Y and $\widetilde{Y}$ are solutions to (3.3.1), we can write for any $h \in [0,1]$,

$$d_{\alpha,W,[0,h]}\big((Y,Y),(\widetilde{Y},\widetilde{Y})\big) \le c_1\|\xi - \widetilde{\xi}\|_V + c_2 h^{\alpha} d_{\alpha,W,[0,h]}((Y,Y),(\widetilde{Y},\widetilde{Y}))\big], \tag{3.3.20}$$

where c_1, c_2 are constants depending on $\|\mathbf{W}\|_{\mathscr{C}_3}$, α, $\boldsymbol{\beta}_3$ and ξ, $\widetilde{\xi}$. Let $\varepsilon = (2c_1)^{-\frac{1}{\alpha}} \wedge (2c_2)^{-\frac{1}{\alpha}} \wedge 1$. It follows that

$$d_{\alpha,W,[0,\varepsilon]}\big((Y,Y),(\widetilde{Y},\widetilde{Y})\big) \le 2c_1\|\xi - \widetilde{\xi}\|_V \tag{3.3.21}$$

on $[0,\varepsilon]$. By iteration, we have that for any $n \ge 1$,

$$d_{\alpha,W,[n\varepsilon,(n+1)\varepsilon]}\big((Y,Y),(\widetilde{Y},\widetilde{Y})\big) \le 2c_1\|Y_{n\varepsilon} - \widetilde{Y}_{n\varepsilon}\|_V,$$

and

$$
\begin{aligned}
\|Y_{n\varepsilon}-\widetilde{Y}_{n\varepsilon}\|_V \leq& \|Y_{(n-1)\varepsilon}-\widetilde{Y}_{(n-1)\varepsilon}\|_V+\varepsilon^{\alpha}\|Y-\widetilde{Y}\|_{\alpha,[(n-1)\varepsilon,n\varepsilon]} \\
\leq& \|Y_{(n-1)\varepsilon}-\widetilde{Y}_{(n-1)\varepsilon}\|_V+\varepsilon^{\alpha}d_{\alpha,W,[(n-1)\varepsilon,n\varepsilon]}\big((Y,Y),(\widetilde{Y},\widetilde{Y})\big) \\
\leq& 2\|Y_{(n-1)\varepsilon}-\widetilde{Y}_{(n-1)\varepsilon}\|_V.
\end{aligned}
$$

Thus we can write

$$
d_{\alpha,W,[n\varepsilon,(n+1)\varepsilon]}\big((Y,Y),(\widetilde{Y},\widetilde{Y})\big)\leq 2^{n+1}c_1\|\xi-\widetilde{\xi}\|_V.
$$

In order to obtain the global distance, we proceed as follows. If $T\leq\varepsilon$, then (3.3.19) is a direct consequence of (3.3.21). It suffices to consider the case when $T>\varepsilon$. Let N be the positive integer such that $N\varepsilon<T\leq(N+1)\varepsilon$. Due to Lemma 3.3.5, the following inequality holds

$$
d_{\alpha,W}\big((Y,Y),(\widetilde{Y},\widetilde{Y})\big)\leq (T/\varepsilon)^{1-\alpha}\max_{0\leq n\leq N}\big\{d_{\alpha,W,[n\varepsilon,(n+1)\varepsilon]}\big((Y,Y),(\widetilde{Y},\widetilde{Y})\big)\big\}\leq c^T T^{1-\alpha}\|\xi-\widetilde{\xi}\|_V,
$$

for some $c>0$ depending on α, $\boldsymbol{\beta}_3$, $\|\mathbf{W}\|_{\mathscr{C}}$, ξ and $\widetilde{\xi}$. This completes the proof of the proposition.

□

Due to Propositions 3.3.6 and 3.3.7, we can deduce the following corollary.

Corollary 3.3.8. *Assume Hypothesis (H). Write $Y(\xi)$ for the solution to the RDE (3.3.1) with initial condition $\xi\in V$. Let K be any positive constant. Then,*

(i) $\|Y(\xi)\|_\alpha$ is uniformly bounded in the space $\{\xi,\|\xi\|_V\leq K\}$.

(ii) The constant c in (3.3.19) is fixed in the space $\{(\xi,\widetilde{\xi}),\|\xi\|_V+\|\widetilde{\xi}\|_V\leq K\}$.

Remark 3.3.9. *As a consequence of Proposition 3.3.7, we have the following estimates*

$$
\|Y-\widetilde{Y}\|_\alpha\leq c^T(T^{1-\alpha}\vee 1)\|\xi-\widetilde{\xi}\|_V,
$$

and

$$\sup_{t\in[0,T]} \|Y_t - \widetilde{Y}_t\|_V \le [1 + c^T(T\vee 1)]\|\xi - \widetilde{\xi}\|_V.$$

for some constants c depending on α, $\boldsymbol{\beta}_3$, $\|\mathbf{W}\|_{\mathscr{C}}$, ξ and $\widetilde{\xi}$.

3.4 A functional approach to nonlinear rough paths

Let V be a Banach space. In this section, we consider the nonlinear rough path defined in Section 3.2 as a $\mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$-valued linear rough path. We will show that the two approaches are equivalent under some assumptions.

We start this section by defining the space $\mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$:

Definition 3.4.1. *Let $\boldsymbol{\beta}_n = (\beta_0, \dots, \beta_n)$ be a multi-index, where $\beta_k \ge 0$ for all $k \in \{0, 1, \dots, n\}$. The space $\mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$ is the collection of continuously differentiable functions on V with values in V, equipped with the norm:*

$$\|\phi\|_{\boldsymbol{\beta}_n} = \sum_{k=0}^{n} \sup_{x\in V} \frac{\|D^k\phi(x)\|_{\mathfrak{L}_k(V;V)}}{(1+\|x\|_V)^{\beta_k}} < \infty.$$

It is easy to see that $(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V), \|\cdot\|_{\boldsymbol{\beta}_n})$ is a Banach space. In the following lemma, we show the equivalence of the spaces $\mathscr{C}^{\alpha}([0,T];\mathscr{C}^{\boldsymbol{\beta}_n}(V;V))$ and $\mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$ defined in Definition 3.2.1.

Lemma 3.4.2. *(i) Let $\Phi \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$ be defined by (3.2.1) with $\Phi_0 \in \mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$. Then, $\Phi \in \mathscr{C}^{\alpha}([0,T];\mathscr{C}^{\boldsymbol{\beta}_n}(V;V))$.*

(ii) Conversely, if $\Phi \in \mathscr{C}^{\alpha}([0,T];\mathscr{C}^{\boldsymbol{\beta}_n}(V;V))$, then $\Phi \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$.

Proof. (i) Fix $t \in [0,T]$. We can show that

$$\|\Phi_t\|_{\boldsymbol{\beta}_n} \le \|\Phi_0\|_{\boldsymbol{\beta}_n} + \|\Phi_{0,t}\|_{\alpha,\boldsymbol{\beta}_n} \le \|\Phi_0\|_{\boldsymbol{\beta}_n} + T^{\alpha}\|\Phi\|_{\alpha,\boldsymbol{\beta}_n} < \infty.$$

Similarly for any $0 \le s \le t \le T$, we have

$$\|\Phi_{s,t}\|_{\boldsymbol{\beta}_n} \le \|\Phi\|_{\alpha,\boldsymbol{\beta}_n}|t-s|^{\alpha}.$$

It follows that as a $\mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$-valued function, $\|\Phi\|_{\alpha} \le \|\Phi\|_{\alpha,\boldsymbol{\beta}_n} < \infty$.

(ii) We estimate $\|\Phi\|_{\alpha,\boldsymbol{\beta}_n}$ as follows:

$$\|\Phi\|_{\alpha,\boldsymbol{\beta}_n} = \sum_{k=0}^{n} \sup_{\substack{s\neq t\in[0,T]\\ x\in V}} \frac{\|D^k\Phi_{s,t}(x)\|_V}{|t-s|^{\alpha}(1+\|x\|_V)^{\beta_k}} = \sup_{s\neq t\in[0,T]} \frac{\|\Phi_{s,t}\|_{\boldsymbol{\beta}_n}}{|t-s|^{\alpha}} \le \|\Phi\|_{\alpha}.$$

As a consequence, $\Phi \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$. □

Let $n \ge 1$, and let $(W,\mathscr{W}) \in \mathscr{C}^{\alpha}([0,T];\mathscr{C}^{\boldsymbol{\beta}_n}(V;V))$ be a $\mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$-valued linear rough path in the sense of Definition 3.1.2. Then, $\mathscr{W} \in \mathscr{C}_2^{2\alpha}([0,T]^2;\mathscr{C}^{\boldsymbol{\beta}_n}(V;V)^{\otimes 2})$. We define $\mathbb{W} : [0,T]^2 \times V^2 \to V$ as follows:

$$\mathbb{W}_{s,t}(x,y) := \mathscr{D}^{(2)}\mathscr{W}_{s,t}(x,y), \tag{3.4.1}$$

where $\mathscr{D}^{(2)} : \mathscr{C}^{\boldsymbol{\beta}_n}(V;V) \otimes \mathscr{C}^{\boldsymbol{\beta}_n}(V;V) \to \mathscr{C}^{\boldsymbol{\beta}^*_{n-1},\boldsymbol{\beta}^{**}_{n-1}}(V\times V;V)$ with the multi-indexes $\boldsymbol{\beta}^*_{n-1}$ and $\boldsymbol{\beta}^{**}_{n-1}$ defined in (3.2.3), is given by

$$\mathscr{D}^{(2)}(\phi^1,\phi^2)(x,y) := D\phi^2(y)(\phi^1(x)),$$

for all $(\phi^1,\phi^2) \in \mathscr{C}^{\boldsymbol{\beta}_n}(V;V)^2$ and $(x,y) \in V^2$. One should notice that the operator $\mathscr{D}^{(2)}$ can be extended continuously to the tensor product space $\mathscr{C}^{\boldsymbol{\beta}_n}(V;V) \otimes \mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$, and it is a linear operator on this space. We can also define

$$\mathscr{D}^{(1)}(\phi^1,\phi^2)(x,y) := D\phi^1(x)(\phi^2(x)) \tag{3.4.2}$$

for all $(\phi^1,\phi^2) \in \mathscr{C}^{\boldsymbol{\beta}_n}(V;V)^2$, and continuously extend it to $\mathscr{C}^{\boldsymbol{\beta}_n}(V;V) \otimes \mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$. In the next proposition, we show that $(W, \mathbb{W}) \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T] \times V; V)$.

Proposition 3.4.3. *Let* $\mathbf{W} = (W, \mathscr{W}) \in \mathscr{C}^{\alpha}([0,T]; \mathscr{C}^{\boldsymbol{\beta}_n}(V;V))$*, and let* $\mathbb{W} : [0,T]^2 \times V^2 \to V$ *be given by (3.4.1). Then* $(W, \mathbb{W}) \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T] \times V; V)$.

Proof. According to Lemma 3.4.2, we know that $W \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T] \times V; V)$ and this implies that $\mathbb{W} \in \mathscr{C}_2^{2\alpha,\boldsymbol{\beta}_{n-1}^{*},\boldsymbol{\beta}_{n-1}^{**}}([0,T]^2 \times V^2; V)$. It suffices to verify Chen's relation (3.2.10). Recall that $(W, \mathscr{W}) \in \mathscr{C}^{\alpha}([0,T]; \mathscr{C}^{\boldsymbol{\beta}_n}(V;V))$ satisfies Chen's relation (3.1.3), and the operator $\mathscr{D}^{(2)}$ is linear on $\mathscr{C}^{\boldsymbol{\beta}_n}(V;V) \otimes \mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$. It follows that

$$\begin{aligned}\mathbb{W}_{s,t}(x,y) - \mathbb{W}_{s,u}(x,y) - \mathbb{W}_{u,t}(x,y) &= \mathscr{D}^{(2)}(\mathscr{W}_{s,t} - \mathscr{W}_{s,u} - \mathscr{W}_{u,t})(x,y) \\ &= \mathscr{D}^{(2)}(W_{s,u} \otimes W_{u,t})(x,y) = DW_{u,t}(y)(W_{s,u}(x)).\end{aligned}$$

As a consequence, $(W, \mathbb{W}) \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T] \times V; V)$. □

Remark 3.4.4. *Proposition 3.4.3 shows that* $\mathbb{W}$ *can be constructed from* $\mathscr{W}$*. However, generally we are not able to recover* $\mathscr{W}$ *from* $\mathbb{W}$ *satisfying Chen's relation (3.2.10). In other words, the nonlinear integral* $\int_0^t W(dr, Y_r)$ *and the nonlinear RDE (3.3.1) can be studied using the approach of Section 3.2 even if* $\mathscr{W}$ *does not exist.*

Let $\mathbf{W} = (W, \mathscr{W}) \in \mathscr{C}^{\alpha}([0,T]; \mathscr{C}^{\boldsymbol{\beta}_n}(V;V))$. In the theory of linear rough paths, under the assumption that $Y \in \mathscr{D}_W^{2\alpha}(\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);V))$, the rough integral of Y against W is well-defined. The nonlinear rough integral defined in Section 3.2 can be also interpreted as the linear rough integral. In this case, the controlled rough path Y belongs to a proper subset of $\mathscr{D}_W^{2\alpha}(\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);V))$, that is equivalent to $\mathscr{E}_W^{2\alpha}$ in the sense of Definition 3.2.5. To describe this subset, we introduce the following special class of operators in $\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);V)$. For any $x \in V$, let $\widehat{x} : \mathscr{C}^{\boldsymbol{\beta}_n}(V;V) \to V$ be given by

$$\widehat{x}(\phi) := \phi(x). \tag{3.4.3}$$

Then $\widehat{x} \in \mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);V)$ with operator norm bounded by $(1 + \|x\|_V)^{\beta_0}$. Let $n \geq 1$ and let $W \in$

$\mathscr{C}^\alpha([0,T];\mathscr{C}^{\boldsymbol{\beta}_n}(V;V))$. We introduce the space of basic controlled rough paths of a $\mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$-valued rough path as a subspace of $\mathscr{D}_W^{2\alpha}(\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);V))$, where $\mathscr{D}_W^{2\alpha}(\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);V))$ is defined as in Definition 3.1.3. Here, the state space of W is $\mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$. Additionally, assume that $(\mathscr{Y},\mathscr{Y}')\in\mathscr{D}_W^{2\alpha}(\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);V))$. Then the state spaces of $\mathscr{Y}$ and $\mathscr{Y}'$ are $\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);V)$ and $\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);V))$, respectively.

Definition 3.4.5. *A pair of functions $(\mathscr{Y},\mathscr{Y}')\in\mathscr{D}_W^{2\alpha}(\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);V))$ is called a basic rough path controlled by W, if there exists a pair of functions $(Y,\dot{Y})\in\mathscr{C}^\alpha(V;V)\times\mathscr{C}^\alpha(V;V)$, such that for all $t\in[0,T]$, $\mathscr{Y}_t=\widehat{Y}_t$ and for all $(\phi_1,\phi_2)\in\mathscr{C}^{\boldsymbol{\beta}_n}(V;V)^2$*

$$\mathscr{Y}_t'(\phi_1,\phi_2)=\widehat{Y}_t'(\phi_1,\phi_2):=\mathscr{D}^{(2)}(\phi_1,\phi_2)(\dot{Y}_t,Y_t)=D\phi_2(Y_t)(\phi_1(\dot{Y}_t)). \tag{3.4.4}$$

We write $\widetilde{\mathscr{E}}_W^{2\alpha}$ for the collection of such pairs.

The next proposition provides the equivalence between the spaces $\widetilde{\mathscr{E}}_W^{2\alpha}$ and $\mathscr{E}_W^{2\alpha}$.

Proposition 3.4.6. *Let $n\geq 1$ and let $W\in\mathscr{C}^\alpha([0,T];\mathscr{C}^{\boldsymbol{\beta}_n}(V;V))$. Then by Lemma 3.4.2, $W\in\mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$ as well. In addition, the following properties hold:*

(i) Let $(Y,\dot{Y})\in\mathscr{E}_W^{2\alpha}$ in the sense of Definition 3.2.5. Then, $(\widehat{Y},\widehat{Y}')\in\widetilde{\mathscr{E}}_W^{2\alpha}$ in the sense of Definition 3.4.5, where $\widehat{Y}_t$ and $\widehat{Y}_t'$ are given by (3.4.3) and (3.4.4) respectively, for all $t\in[0,T]$.

(ii) Conversely, let $(\widehat{Y},\widehat{Y}')\in\widetilde{\mathscr{E}}_W^{2\alpha}$ with associated pair $(Y,\dot{Y})\in\mathscr{C}^\alpha(V;V)^2$. Then $(Y,\dot{Y})\in\mathscr{E}_W^{2\alpha}$.

Proof. (i) By assumption $Y\in\mathscr{C}^\alpha([0,T];V)$. It follows that

$$\begin{aligned}\|\widehat{Y}\|_\alpha=&\sup_{s\neq t\in[0,T]}\frac{\|\widehat{Y}_{s,t}\|_{\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);V)}}{|t-s|^\alpha}=\sup_{s\neq t\in[0,T]}\sup_{0\neq\phi\in\mathscr{C}^{\boldsymbol{\beta}_n}(V;V)}\frac{\|\phi(Y_t)-\phi(Y_s)\|_V}{|t-s|^\alpha\|\phi\|_{\boldsymbol{\beta}_n}}\\ \leq&(1+\|Y\|_\infty)^{\beta_1}\|Y\|_\alpha.\end{aligned}$$

This implies that $\widehat{Y}\in\mathscr{C}^\alpha([0,T];\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);V))$. Similarly, since $\dot{Y}\in\mathscr{C}^\alpha([0,T];V)$, we can deduce the following inequality:

$$\|\widehat{Y}'\|_\alpha \le (1+\|Y\|_\infty)^{\beta_2}(1+\|\dot{Y}\|_\infty)^{\beta_0}\|Y\|_\alpha + (1+\|Y\|_\infty)^{\beta_1}(1+\|\dot{Y}\|_\infty)^{\beta_1}\|\dot{Y}\|_\alpha.$$

It suffices to estimate the reminder term. Recall that $\widehat{Y}'$ is defined as in (3.4.4). Then, for any $\phi \in \mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$, the remainder $R^{\widehat{Y}}_{s,t}(\phi)$ can be written as follows,

$$R^{\widehat{Y}}_{s,t}(\phi) = \phi(Y_t) - \phi(Y_s) - D\phi(Y_s)W_{s,t}(\dot{Y}_s).$$

Due to Taylor's Theorem 3.1.11 and the fact that $(Y,\dot{Y}) \in \mathscr{E}^{2\alpha}_W$, we have

$$\begin{aligned}
\|R^{\widehat{Y}}_{s,t}(\phi)\|_V \le & \|D\phi(Y_s)Y_{s,t} - D\phi(Y_s)W_{s,t}(\dot{Y}_s)\|_V + \frac{1}{2}\sup_{0\le\tau\le1}\|D^2\phi(\tau Y_t + (1-\tau)Y_s)(Y_{s,t}Y_{s,t})\|_V \\
= & \|D\phi(Y_s)[W_{s,t}(\dot{Y}_s) + R^Y_{s,t}] - D\phi(Y_s)W_{s,t}(\dot{Y}_s)\|_V \\
& + \frac{1}{2}\sup_{0\le\tau\le1}\|D^2\phi(\tau Y_t + (1-\tau)Y_s)(Y_{s,t}Y_{s,t})\|_V \\
\le & \|\phi\|_{\boldsymbol{\beta}_n}\Big[\frac{1}{2}(1+\|Y\|_\infty)^{\beta_2}\|Y\|^2_\alpha + (1+\|Y\|_\infty)^{\beta_1}\|R^Y\|_{2\alpha}\Big]|t-s|^{2\alpha}.
\end{aligned}$$

This implies $R^{\widehat{Y}} \in \mathscr{C}^{2\alpha}_2([0,T];\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_n}(V;V);V))$. As a consequence, we conclude that $(\widehat{Y},\widehat{Y}') \in \widetilde{\mathscr{E}}^{2\alpha}_W$.

(ii) To prove the converse result, it suffices to show that $R^Y \in \mathscr{C}^{2\alpha}_2([0,T];V)$, where

$$R^Y_{s,t} := Y_{s,t} - W_{s,t}(\dot{Y}_s).$$

Let K be the closed convex hull of the set $\{Y_t, t\in[0,T]\}$, and let $\widetilde{K}$ is a compact set in V whose interior contains K. Choose a function $\phi : V \to V$ that is infinitely differentiable and satisfies the following properties:

a) $\phi(x) = x$ for all $x \in K$. That implies $D\phi(x) = I$ and $D^2\phi(x) = 0$ for all $x \in K$, where I denotes the identity operator in $\mathscr{L}(V;V)$.

b) $\phi(x) \equiv x_0 \in V$ for all $x \notin \widetilde{K}$.

c) ϕ and all its derivatives are bounded.

Then, it is easy to check that $\phi \in \mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$ for any multi-index $\boldsymbol{\beta}_n$. In addition, we can show that

$$\|R^Y_{s,t}\|_V = \|\phi(Y_t) - \phi(Y_s) - D\phi(Y_s)[W_{s,t}(\phi_s)]\|_V = \|R^{\widehat{Y}}_{s,t}(\phi)\|_V \leq \|R^{\widehat{Y}}\|_{2\alpha}\|\phi\|_{\beta_n}|t-s|^{2\alpha}.$$

In other words, $R^Y \in \mathscr{C}^{2\alpha}([0,T];V)$, and thus $(Y,\dot{Y}) \in \mathscr{E}^{2\alpha}_W$. □

In the next theorem, we will show the equivalence of two rough integrals.

Theorem 3.4.7. *Let* $\mathbf{W} = (W, \mathscr{W}) \in \mathscr{C}^{\alpha}([0,T];\mathscr{C}^{\boldsymbol{\beta}_2}(V;V))$. *Due to Proposition 3.4.3, we can construct* $(W, \mathbb{W}) \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_2}([0,T] \times V;V)$. *Assume that* $(\widehat{Y}, \widehat{Y}') \in \widetilde{\mathscr{E}}^{2\alpha}_W$ *with associated pair* $(Y,\dot{Y}) \in \mathscr{E}^{2\alpha}_W$ *by Proposition 3.4.6. Then, the following two rough integrals coincide,*

$$\int_s^t W(dr, Y_r) = \int_s^t \widehat{Y}_r d\mathbf{W}_r, \tag{3.4.5}$$

where the integral on the left-hand side is in the sense of (3.2.26), and the integral on the right-hand side is in the sense of Theorem 3.1.4.

Proof. Let $\Xi_{s,t}$ and $\widetilde{\Xi}_{s,t}$ be the approximations of the integral on the left and right-hand side, respectively. That is,

$$\Xi_{s,t} = W_{s,t}(Y_s) + \mathbb{W}_{s,t}(\dot{Y}_s, Y_s) \text{ and } \widetilde{\Xi}_{s,t} = \widehat{Y}_s W_{s,t} + \widehat{Y}'_s \mathscr{W}_{s,t}.$$

Here $\widehat{Y}'_s$ acting on $\mathscr{W}_{s,t}$ is a continuous extension of formula (3.4.4) to the tensor product space $\mathscr{C}^{\boldsymbol{\beta}_2}(V;V)^{\otimes 2}$. By definition of $\mathbb{W}$ and $(\widehat{Y}, \widehat{Y}')$, we have

$$\widehat{Y}_s W_{s,t} + \widehat{Y}'_s \mathscr{W}_{s,t} = W_{s,t}(Y_s) - \mathscr{D}^{(2)}\mathscr{W}_{s,t}(\dot{Y}_s, Y_s) = W_{s,t}(Y_s) + \mathbb{W}_{s,t}(\dot{Y}_s, Y_s).$$

This implies the equality (3.4.5). □

At the end of this section, we provide an alternative approach to study the nonlinear RDE introduced in Section 3.3. Let $\mathbf{W} = (W, \mathscr{W}) \in \mathscr{C}^{\alpha}([0,T]; \mathscr{C}^{\boldsymbol{\beta}_3}(V;V))$. Then, the RDE (3.3.1) can be also understood as the following equation:

$$Y_t = \xi + \int_0^t \delta(Y_r) d\mathbf{W}_r, \tag{3.4.6}$$

where δ denotes the Dirac delta operator, that is $\delta : V \to \mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_3}(V;V);V)$ is given by $\delta(x) = \widehat{x}$. A function $Y \in \mathscr{C}^{\alpha}([0,T];V)$ is said to be a solution to (3.4.6), if $(Y, \delta(Y)) \in \mathscr{D}_{\mathrm{W}}^{2\alpha}(V)$ and the equality holds. On the other hand, suppose that Y is a solution to (3.4.6). Then, $(\widehat{Y}, \widehat{Y}') \in \widetilde{\mathscr{E}}_{W}^{2\alpha}$ with associated pair $(Y, Y) \in \mathscr{E}_{W}^{2\alpha}$. Therefore, Y is a solution to the equation (3.3.1) in the sense of Definition 3.3.1.

On the other hand, notice that as an $\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_3}(V;V);V)$-valued operator, δ is three times differentiable. More precisely, the derivatives of δ can be written as follows $D^k\delta(x)(\phi) = D^k\phi(x)$ for $k = 1,2,3$. Thus $\|D^k\delta(x)\| \leq (1 + \|x\|_V)^{\beta_k}$ for all $k = 0,1,2,3$. Then the (global) existence and uniqueness of equation (3.3.1) can be derived by the theory of linear rough paths (c.f. Lejay [60]). For other conditions that implies global existence, we refer the reader to the papers of Lejay [60, 61]. We did not consider Lejay's condition for global existence in Section 3.3, because we doubt whether it is applicable in our setting. Under the basic assumptions in Section 3.3, there may not exists $\mathscr{W}$ such that $\mathbb{W} = \mathscr{D}^{(2)}\mathscr{W}$. In this case, the result of linear rough path cannot be directly applied without any changes.

3.5 Some applications of nonlinear rough paths

3.5.1 An Itô-type formula for controlled rough paths

In this section, we follow the idea of Section 3.4 to consider the nonlinear rough path as a $\mathscr{C}^{\boldsymbol{\beta}_n}(V;V)$-valued rough path. Then, we aim to generalize the Itô-type formula (3.12) in Hu & Lê [42] proved in the nonlinear Young's case.

Theorem 3.5.1. *Let* $\mathbf{W} = (W, \mathscr{W}) \in \mathscr{C}^{\alpha}([0,T]; \mathscr{C}^{\beta_3}(V;V))$. *Assume that* $(Y, Y') \in \mathscr{D}_W^{2\alpha}(V)$ *and* $(Z, Z') \in \mathscr{D}_W^{2\alpha}(\mathscr{L}(V;K))$. *Then, the following Itô-type formula holds*

$$\begin{aligned}\int_s^t Z_r dW(r, Y_r) &= \int_s^t Z_r W(dr, Y_r) + \int_s^t Z_r DW(r, Y_r) dY_r \\ &+ \frac{1}{2}\Big[\int_s^t Z_r D^2 W(r, Y_r) d\langle Y\rangle_r + \int_s^t Z_r d\langle\!\langle X, Y\rangle\!\rangle_r + \int_s^t Z_r d\langle\!\langle Y, X\rangle\!\rangle_r\Big],\end{aligned} \tag{3.5.1}$$

where

$$X_t := \int_0^t DW(dr, Y_r) = \lim_{|\pi|\to 0}\big[DW_{t_{k-1},t_k}(Y_{t_{k-1}}) + (\mathscr{D}^{(2)})^2 Y'_{t_{k-1}} \mathscr{W}_{t_{k-1},t_k}\big], \tag{3.5.2}$$

$$(\mathscr{D}^{(2)})^2 Y'_t(\phi_1, \phi_2) := D^2\phi_2(Y'_t\phi_1) \in \mathscr{L}(V;V).$$

The first three integrals in (3.5.1) *are rough integrals in the sense of Proposition 3.1.7 (ii), while the last three integrals on the second line are Young's integral. In the above expressions,* $\langle Y\rangle$, $\langle\!\langle X, Y\rangle\!\rangle$ *and* $\langle\!\langle Y, X\rangle\!\rangle$ *are* 2α*-continuous functions defined in Definition 3.1.9 and Remark 3.1.10.*

Formula (3.5.1) provides the total differential $dW(t, Y_t)$ of $W(t, Y_t)$, that means, heuristically, $dW(t, Y_t) = \frac{d}{dt} W(t, Y_t) dt$. Comparing with the classical Itô lemma, the function W in Theorem 3.5.1 is not differentiable, but only α-Hölder continuous in time. In this case, the assumption that Y is controlled by W ensures that $W(dt, Y_t)$ is well-defined as the differential of the rough path $G_t = \int_0^t W(dr, Y_r)$ controlled by W.

In order to prove Theorem 3.5.1, we should make each integral in (3.5.1) to be well-defined. The first lemma below shows that $F_t = W(t, Y_t)$ is controlled by W.

Lemma 3.5.2. *Let* $W \in \mathscr{C}^{\alpha}([0,T]; \mathscr{C}^{\beta_2}(V;V))$, *and let* $(Y, Y') \in \mathscr{D}_W^{2\alpha}(V)$. *Denote* $F_t = W(t, Y_t)$. *Then,* $F \in \mathscr{D}_W^{2\alpha}(V)$.

Proof. By Taylor's Theorem 3.1.11 and the fact that $(Y,Y') \in \mathscr{D}_W^{2\alpha}(V)$, we get

$$\begin{aligned} F_{s,t} =& F_t - F_s = W_{s,t}(Y_s) + [W_{s,t}(Y_t) - W_{s,t}(Y_s)] + W_s(Y_t) - W_s(Y_s) \\ =& \widehat{Y}_s W_{s,t} + DW_s(Y_s)[Y_s' W_{s,t} + R_{s,t}^Y] + O(\|Y_{s,t}\|_V^2|). \end{aligned}$$

This yields that $(F,F') \in \mathscr{D}_W^{2\alpha}(V)$, where $F' := \widehat{Y} + DW(Y)Y' \in \mathscr{L}(\mathscr{C}^{\beta_2};V)$. □

Suppose that $\mathbf{W} = (W,\mathscr{W}) \in \mathscr{C}^\alpha([0,T];\mathscr{C}^{\beta_3}(V;V))$. As a consequence of Lemma 3.5.2, the integral $\int_s^t Z_r dW(r,Y_r) = \int_s^t Z_r dF_r$ is well-defined as the integral of two controlled rough paths in the sense of (3.1.12). Additionally, by Taylor's Theorem 3.1.11, we can approximate this integral in the following way:

$$\begin{aligned} \int_s^t Z_r dW(r,Y_r) =& Z_s F_{s,t} + Z_s' F_s' \mathscr{W}_{s,t} + O(|t-s|^{3\alpha}) \\ =& Z_s W_{s,t}(Y_s) + Z_s DW_{s,t}(Y_s) Y_{s,t} + Z_s DW_s(Y_s) Y_{s,t} + \frac{1}{2} Z_s D^2 W_s(Y_s)(Y_{s,t}, Y_{s,t}) \\ &+ Z_s' \widehat{Y}_s \mathscr{W}_{s,t} + Z_s' DW(s,Y_s) Y_s' \mathscr{W}_{s,t} + O(|t-s|^{3\alpha}), \end{aligned} \tag{3.5.3}$$

where

$$Z_s' \widehat{Y}_s(\phi_1, \phi_2) = Z_s'(\phi_1)[\phi_2(Y_s)],$$

and

$$Z_s' DW(s,Y_s) Y_s'(\phi_1,\phi_2) = Z_s'(\phi_1)[DW(s,Y_s)(Y_s'(\phi_2))],$$

for all $(\phi_1,\phi_2) \in \mathscr{C}^{\beta_3}(V;V)^2$.

The next lemma provides a generalized version of Theorem 3.2.10. The proof is similar and we omit it.

Lemma 3.5.3. *Let* $(W,\mathscr{W}) \in \mathscr{C}^\alpha([0,T];\mathscr{C}^{\beta_2}(V;V))$*, and let* $(Y,Y') \in \mathscr{D}_W^{2\alpha}(V)$*. Then, the following* *limit exists and defines an additive function:*

$$\int_s^t W(dr, Y_r) := \lim_{|\pi|\to 0} \sum_{k=1}^{n} \left[W_{t_{k-1},t_k}(Y_s) + Y'_{t_{k-1}} \widehat{Y}_{t_{k-1}} \mathscr{D}^{(2)} \mathscr{W}_{t_{k-1},t_k} \right],$$

where $Y'_t \widehat{Y}_t \mathscr{D}^{(2)}(\phi_1, \phi_2) := D\phi_2(Y_t)[Y'_t(\phi_1)]$ *for any* $(\phi_1, \phi_2) \in \mathscr{C}^{\beta_2}(V;V)$.

For all $t \in [0,T]$, let $G_t := \int_0^t W(dr, Y_r)$. Then, a similar argument as in Proposition 3.2.12 implies that $(G,Y) \in \mathscr{E}_W^{2\alpha}$ or equivalently $(G, \widehat{Y}) \in \mathscr{D}_W^{2\alpha}(V)$. Therefore, the integral $\int_s^t Z_r W(dr, Y_r) = \int_s^t Z_r dG_r$, defined as in (3.1.12), can be approximated in the following way,

$$\begin{aligned} \int_s^t Z_r W(dr, Y_r) =& Z_s G_{s,t} + Z'_s \widehat{Y}_s \mathscr{W}_{s,t} + O(|t-s|^{3\alpha}) \\ =& Z_s W_{s,t}(Y_s) + Z_s Y'_s \widehat{Y}_s \mathscr{D}^{(2)} \mathscr{W}_{s,t} + Z'_s \widehat{Y}_s \mathscr{W}_{s,t} + O(|t-s|^{3\alpha}). \end{aligned} \tag{3.5.4}$$

Assume that $(W, \mathscr{W}) \in \mathscr{C}^{\alpha}([0,T]; \mathscr{C}^{\beta_3}(V;V))$. Let $H_t = Z_t DW(t, Y_t) \in \mathscr{L}(V;V)$ for all $t \in [0,T]$. By a similar argument as in Lemma 3.5.2, we can show that

$$H_{s,t} = Z'_s W_{s,t} DW(s, Y_s) + Z_s \widehat{Y}_s DW_{s,t} + Z_s D^2 W(s, Y_s) Y'_s W_{s,t} + O(|t-s|^{2\alpha}).$$

In other words, H is controlled by W. This allows us to define $\int_s^t Z_r DW(r, Y_r) dY_r = \int_s^t H_r dY_r$ by (3.1.12). In addition, we can approximate this integral as follows,

$$\begin{aligned} \int_s^t Z_r DW(r, Y_r) dY_r =& Z_s DW(s, Y_s) Y_{s,t} + Z'_s DW(s, Y_s) Y'_s \mathscr{W}_{s,t} \\ & + Z_s \widehat{Y}_s \mathscr{D}^{(1)} Y'_s \mathscr{W}_{s,t} + Z_s D^2 W(s, Y_s) Y'_s Y'_s \mathscr{W}_{s,t} + O(|t-s|^{3\alpha}), \end{aligned} \tag{3.5.5}$$

where $\mathscr{D}^{(1)}$ is defined as in (3.4.2),

$$Z'_s DW(s, Y_s) Y'_s(\phi_1, \phi_2) = Z'_s(\phi_1)[DW(s, Y_s)(Y'_s(\phi_2))],$$

$$Z_s\widehat{Y}_s\mathscr{D}^{(1)}Y_s'(\phi_1,\phi_2) = Z_s[D\phi_1(Y_s)Y_s'(\phi_2)],$$

and

$$Z_sD^2W(s,Y_s)Y_s'Y_s'(\phi_1,\phi_2) = Z_s[D^2W(s,Y_s)(Y_s'(\phi_1),Y_s'(\phi_2))],$$

for all $(\phi_1,\phi_2) \in \mathscr{C}^{\beta_3}(V;V)$.

By a similar argument as in Theorem 3.2.10 and the Sewing Lemma, we can show that the limit in (3.5.2) uniquely exists. It allows us to define X_t to be the limit. In addition, we can verify that $X \in \mathscr{D}_W^{2\alpha}(\mathscr{L}(V;V))$. Thus the three quadratic compensator terms on the second line of (3.5.1) are all well-defined, and according to Remark 3.1.10 (iii), $\langle Y\rangle \in \mathscr{C}_2^{2\alpha}([0,T];V\otimes V)$ and $\langle\!\langle X,Y\rangle\!\rangle, \langle\!\langle Y,X\rangle\!\rangle \in \mathscr{C}_2^{2\alpha}([0,T];V)$. Therefore, the integrals on the second line of (3.5.1) can be interpreted as Young's integrals. We can approximate them as follows:

$$\int_s^t Z_r\langle\!\langle X,Y\rangle\!\rangle_r = Z_sDW_{s,t}(Y_s)Y_{s,t} - 2Z_s\widehat{Y}_s\mathscr{D}^{(1)}Y_s'\mathscr{W}_{s,t} + O(|t-s|^{3\alpha}), \tag{3.5.6}$$

$$\int_s^t Z_r\langle\!\langle Y,X\rangle\!\rangle_r = Z_sDW_{s,t}(Y_s)Y_{s,t} - 2Z_sY_s'\widehat{Y}_s\mathscr{D}^{(2)}\mathscr{W}_{s,t} + O(|t-s|^{3\alpha}), \tag{3.5.7}$$

and

$$\int_s^t Z_rD^2W(r,Y_r)d\langle Y\rangle_r = Z_sD^2W(s,Y_s) + O(|t-s|^{3\alpha}).$$

Notice that, by definition,

$$\begin{aligned}\langle Y\rangle_{s,t} =& Y_{s,t}\otimes Y_{s,t} - 2\mathbb{Y}_{s,t} = Y_{s,t}\otimes Y_{s,t} - 2\Big(\int_s^t Y_r\otimes dY_r - Y_s\otimes Y_{s,t}\Big)\\ =& Y_{s,t}\otimes Y_{s,t} - 2\big(Y_s\otimes Y_{s,t} + Y_s'Y_s'\mathscr{W}_{s,t} - Y_s\otimes Y_{s,t}\big) + O(|t-s|^{3\alpha}).\end{aligned}$$

This allows us to write

$$\int_s^t Z_r = Z_s D^2 W(s,Y_s)\left[Y_{s,t}\otimes Y_{s,t} - 2Y_s'Y_s'\mathscr{W}_{s,t}\right] + O(|t-s|^{3\alpha}) \tag{3.5.8}$$

As we approximated all the integrals in (3.5.1), the proof of Theorem 3.5.1 is straightforward.

Proof of Theorem 3.5.1. Denote by LHS and RHS the left and right-hand side of equation (3.5.1) respectively. Recall equality (3.5.3), that is,

$$\begin{aligned} LHS =& Z_s W_{s,t}(Y_s) + Z_s DW_{s,t}(Y_s)Y_{s,t} + Z_s DW_s(Y_s)Y_{s,t} + \frac{1}{2}Z_s D^2 W_s(Y_s)(Y_{s,t},Y_{s,t}) \\ &+ Z_s'\widehat{Y}_s\mathscr{W}_{s,t} + Z_s DW_s(Y_s)Y_s'\mathscr{W}_{s,t} + O(|t-s|^{3\alpha}). \end{aligned}$$

On the other hand, combining (3.5.4) - (3.5.8), we have

$$\begin{aligned} RHS =& Z_s W_{s,t}(Y_s) + Z_s DW_{s,t}(Y_s)Y_{s,t} + Z_s DW_s(Y_s)Y_{s,t} + \frac{1}{2}Z_s D^2 W_s(Y_s)(Y_{s,t},Y_{s,t}) \\ &+ Z_s'\widehat{Y}_s\mathscr{W}_{s,t} + Z_s DW_s(Y_s)Y_s'\mathscr{W}_{s,t} + O(|t-s|^{3\alpha}), \end{aligned}$$

as well. Since $\alpha \in (\frac{1}{3},\frac{1}{2}]$, it follows that equality (3.5.1) holds for all $0 \le s \le t \le T$. □

3.5.2 RDEs with spatial parameters

Let $(W,\mathscr{W}) \in \mathscr{C}^\alpha([0,T];\mathscr{C}^{\beta_3}(\mathbb{R}^d;\mathbb{R}^d))$, and let $\mathbb{W}$ be given by (3.4.1). Assume Hypothesis (H). Then, due to Theorem 3.3.4, for any fixed $x \in \mathbb{R}^d$, the following equation

$$Y_t(x) = x + \int_0^t W(dr, Y_r(x)), \tag{3.5.9}$$

has a unique solution $Y(x)$ on $[0,T]$. In this section, by studying the gradient in x of $Y_t(x)$, we will show that $Y_t(x)$ is invertible in x, and the inverse is controlled by W as well.

In the next theorem, we follow the idea of Hu & Lê [42] to show that $Y_t(x)$ is differentiable in x. Before presenting the theorem, we introduce some notations. Let M be a $d \times d$ matrix. We define the operators $M^L, M^M : (\mathbb{R}^d \otimes \mathbb{R}^d)^{\otimes 2} \to \mathbb{R}^d \otimes \mathbb{R}^d$ as follows, for any $(A,B) \in (\mathbb{R}^d \otimes \mathbb{R}^d)^2$,

$$M^L(A \otimes B) = M \cdot A \cdot B \text{ and } M^M(A \otimes B) = A \cdot M^M \cdot B. \tag{3.5.10}$$

For any $d \times d$ matrices M_1, M_2, we define the operator $\{M_1, M_2\} : \mathbb{R}^d \otimes \mathbb{R}^d \to \mathbb{R}$ by

$$\{M_1, M_2\}A = \sum_{k_1,k_2,k_3} M_1^{k_1k_2} M_2^{k_1k_3} A^{k_2k_3} \text{ for all } A \in \mathbb{R}^d \otimes \mathbb{R}^d. \tag{3.5.11}$$

These operators appear when we approximate matrix-valued rough integrals.

Theorem 3.5.4. *Let $(W, \mathscr{W}) \in \mathscr{C}^\alpha([0,T]; \mathscr{C}^{\beta_3}(\mathbb{R}^d; \mathbb{R}^d))$. Assume Hypothesis (H). Let $Y = \{Y_t(x), t \in [0,T], x \in \mathbb{R}^d\}$ be the unique solution to (3.5.9). Then for any $t \in [0,T]$, Y_t is differentiable, and the gradient DY_t satisfies the following equation:*

$$DY_t(x) = I + \int_0^t dF_r(x) DY_r(x), \tag{3.5.12}$$

where I denotes the $d \times d$ identity matrix and $F(x)$ is a $d \times d$ matrix-valued function given by

$$F_t(x) := \int_0^t DW(dr, Y_r(x))$$

that is defined in the sense of (3.5.2). Moreover, for every $t \in [0,T]$ and $x \in \mathbb{R}^d$, $DY_t(x)$ is invertible, and its inverse $(DY_t(x))^{-1} =: M_t(x)$ satisfies the following equation:

$$M_t(x) = I - \int_0^t M_r(x) dF_r(x) + \int_0^t [M_r(x)]^L d\langle F(x) \rangle_r. \tag{3.5.13}$$

where $\langle F(x) \rangle_r$ is the quadratic compensator of $F(x)$, which is an $(\mathbb{R}^d \otimes \mathbb{R}^d)^{\otimes 2}$-valued 2α-Hölder continuous function on $[0,t]$, and $[M_r(x)]^L : (\mathbb{R}^d \otimes \mathbb{R}^d)^{\otimes 2} \to \mathbb{R}^d \otimes \mathbb{R}^d$ is defined as in (3.5.10).

Proof. Fix $x \in \mathbb{R}^d$. Let e be a unit vector in $\mathbb{R}^d$. For any $h \in (0,1)$, we write

$$\eta_t^h := \frac{1}{h}[Y_t(x+he) - Y_t(x)].$$

We claim that as $h \downarrow 0$, η_t^h converges to the solution to the following equation

$$\eta_t = e + \int_0^t dF_r(x)\eta_r = e + \int_0^t DW(dr, Y_r(x))\eta_r. \tag{3.5.14}$$

Firstly, we show that (3.5.14) has a unique solution. Notice that $F(x)$ is defined as a nonlinear rough integral. Then, by Proposition 3.2.12, $F(x)$ is controlled by DW and thus by W. That is,

$$F_{s,t}(x) = DW_{s,t}(Y_s(x)) + O(|t-s|^{2\alpha}) := (\widehat{Y}_s(x)D)W_{s,t} + O(|t-s|^{2\alpha}),$$

where $\widehat{Y}_\cdot(x)D$ is considered as an α-Hölder continuous function on $[0,T]$ that takes values in $\mathscr{L}(\mathscr{C}^{\beta_3}(\mathbb{R}^d;\mathbb{R}^d);\mathscr{L}(\mathbb{R}^d;\mathbb{R}^d))$. Here $\widehat{Y}$ is defined in (3.4.3). We can also directly define the operator $\widehat{Y}_s(x)D$ by the former expression. $DW_{s,t}(Y_s(x))$ is just an approximation of the integral without the double integral term, thus the error is $O(|t-s|^{2\alpha})$. By Proposition 3.1.7 (ii), $F(x)$ can be interpreted as a linear rough path. Thus, equation (3.5.14) is a linear RDE. According to the theory of linear RDE (c.f. Theorem 2 of Lejay [60]), this equation has a unique solution.

On the other hand, by Corollary 3.3.8, $\|\eta^h\|_\alpha$ is uniformly bounded in $h \in (0,1)$. As a consequence of the Arzelà-Ascoli theorem, there exists a sequence $\{h_n\}_{n\geq 1}$, such that, as $n \to \infty$, $h_n \downarrow 0$, and $\eta_t^{h_n}$ converges to some function η_t in $\mathscr{C}^{\alpha'}([0,T];\mathbb{R}^d)$ for any fixed $\alpha' \in (0,\alpha)$. In addition, by the Sewing Lemma, η^{h_n} satisfies the following estimate

$$\eta_{s,t}^{h_n} = DW_{s,t}(Y_s(x))\eta_s^{h_n} + D\mathscr{W}_{s,t}(Y_s(x),Y_s(x))(\eta_s^{h_n},\eta_s^{h_n}) + O(|t-s|^{3\alpha}) + O(h_n), \tag{3.5.15}$$

for all $0 \leq s < t \leq T$. Let $n \to \infty$. The estimate (3.5.15) implies that η_t satisfies the RDE (3.5.14). Therefore, $DY_t(x)$ exists and is the unique solution to (3.5.12).

To prove the invertibility of $DY_t(x)$, we follow Stroock's idea (see Chapter 8 of Stroock [76]). Let $M_t(x)$ be the unique solution to the linear RDE (3.5.13). By (3.1.16) and Itô's formula for linear rough paths (c.f. Theorem 3.4 of Keller & Zhang [53]), we can deduce the following equation:

$$\begin{aligned}DY_t(x)M_t(x) =& I+\int_0^t dF_r(x)DY_r(x)M_r(x)-\int_0^t DY_r(x)M_r(x)dF_r(x)\\&+\int_0^t [DY_r(x)M_r(x)]^L d\langle F(x)\rangle_r-\int_0^t [DY_r(x)M_r(x)]^M d\langle F(x)\rangle_r,\end{aligned}$$

where $[DY_r(x)M_r(x)]^M$ is a linear operator on $(\mathbb{R}^d\otimes\mathbb{R}^d)^{\otimes 2}$ defined as in (3.5.10). Notice that $DY_t(x)M_t(x)\equiv I$ solves this equation. Thus the uniqueness of linear RDEs implies that $M_t = (DY_t)^{-1}$. □

Remark 3.5.5. *By taking further spatial derivatives on both sides of* (3.5.12) *and* (3.5.13)*, we can show that DY_t and M_t are both twice spatial differentiable with locally bounded derivatives. On the other hand, since Theorem 3.5.4 shows that $DY_t(x)$ is invertible in x for all $(t,x)\in[0,T]\times\mathbb{R}^d$, by the implicit function theorem, we deduce that for any fixed $t\in[0,T]$, Y_t has an inverse Z_t such that $Z_t(Y_t(x)) = Y_t(Z_t(x)) = x$.*

In the next lemma, we prove that fix $x\in\mathbb{R}^d$, $Z(x)$ is controlled by W.

Lemma 3.5.6. *Let $Y(x) = \{Y_t(x), t\in[0,T]\}$ be the solution to the RDE (3.5.9), and let $Z_t = Y_t^{-1}$ be the inverse of Y_t. Fix $x\in\mathbb{R}^d$. Then $Z(x)$ is controlled by W.*

Proof. Recall that for any $t\in[0,T]$, Z_t is the inverse of Y_t and $DY_tM_t = I$. Therefore, we can deduce that

$$I = Dx = DY_t(Z_t(x)) = DY_t(Z_t(x))DZ_t(x).$$

This yields that

$$DZ_t(x) = M_t(Z_t(x)). \tag{3.5.16}$$

Fix $(t,x)\in(0,T]\times\mathbb{R}^d$. Let $y = Z_t(x)$. Then $x = Y_t(y)$. Notice that a similar argument as in Theorem 3.5.4 implies that $M_t(x)$ is differentiable in x and the derivative is locally bounded. Thus by Taylor's Theorem 3.1.11, the following equality holds for all $s \in [0,t)$

$$\begin{aligned} Z_{s,t}(x) =& Z_s(Y_s(y)) - Z_s(Y_t(y)) \\ =& - DZ_s(Y_s(y))Y_{s,t}(y) + O(|t-s|^{2\alpha}) \\ =& - M_s(Z_s(x))Y_{s,t}(Z_s(x)) + O(|t-s|^{2\alpha}). \end{aligned}$$

On the other hand, by Proposition 3.2.12, we have

$$Y_{s,t}(x) = W_{s,t}(Y_s(x)) + O(|t-s|^2).$$

Combining above two inequalities, we can write

$$Z_{s,t}(x) = -M_s(Z_s(x))W_{s,t}(x) + O(|t-s|^{2\alpha}). \tag{3.5.17}$$

Let $Z'(x) = \{Z'_t(x), t \in [0,T]\}$ where $Z'_t(x) : \mathscr{C}^{\beta_3}(\mathbb{R}^d;\mathbb{R}^d) \to \mathbb{R}^d$ is given by

$$Z'_t(x)\Phi := -M_t(Z_t(x))\Phi(x).$$

Then it is easy to check that $Z'_s(x) \in \mathscr{L}(\mathscr{C}^{\beta_3}(\mathbb{R}^d;\mathbb{R}^d);\mathbb{R}^d)$, and thus $(Z(x), Z'(x)) \in \mathscr{D}^{2\alpha}_W(\mathbb{R}^d)$. □

Remark 3.5.7. *(i) One may find that $Z'_t(x) = -DZ_t(x)$. But they are totally different objects. $Z'_t(x)$ is the Gubinelli derivative that represents the proportional changing rate to W of $Z_t(x)$ with respect to the time argument, while $DZ_t(x)$ is the spatial derivative of Z_t for fixed t.*

(ii) By taking derivative on both sides of (3.5.16), we have

$$D^2 Z_t(x) = DM_t(Z_t(x))M_t(Z_t(x)).$$

Recall that Y is the solution to RDE (3.5.9), thus

$$Y_{s,t}(x) = W_{s,t}(Y_s(x)) + \mathbb{W}_{s,t}(Y_s(x), Y_s(x)) + O(|t-s|^{3\alpha}).$$

This allow us to deduce an estimate, which is more precise than (3.5.17) and will be used in Section 3.5.3 below. We start with the following equation

$$\begin{aligned} Z_{s,t}(x) =& -M_s(Z_t(x))Y_{s,t}(Z_t(x)) - \frac{1}{2}DM_s(Z_t(x))M_s(Z_t(x))Y_{s,t}(Z_t(x))^{\otimes 2} \\ &+ O(|t-s|^{3\alpha}) \\ =& -M_s(Z_t(x))W_{s,t}(Y_s(Z_t(x))) - M_s(Z_t(x))\mathbb{W}_{s,t}(Y_s(Z_t(x)), Y_s(Z_t(x))) \\ &- \frac{1}{2}DM_s(Z_t(x))M_s(Z_t(x))Y_{s,t}(Z_t(x))^{\otimes 2} + O(|t-s|^{3\alpha}). \end{aligned} \tag{3.5.18}$$

Notice that

$$M_s(Z_t(x))\mathbb{W}_{s,t}(Y_s(Z_t(x)), Y_s(Z_t(x))) - M_s(Z_s(x))\mathbb{W}_{s,t}(x,x) = O(|t-s|^{3\alpha}) \tag{3.5.19}$$

and

$$\begin{aligned} & M_s(Z_t(x))W_{s,t}(Y_s(Z_t(x))) - M_s(Z_s(x))W_{s,t}(Y_s(Z_s(x))) \\ =& DM_s(Z_s(x))Z_{s,t}(x)W_{s,t}(Y_s(Z_t(x))) \qquad (3.5.20) \\ & + M_s(Z_s(x))DW_{s,t}(Y_s(Z_s(x)))DY_s(Z_s(x))Z_{s,t}(x) + O(|t-s|^{2\alpha}) \\ =& -DM_s(Z_s(x))M_s(Z_s(x))W_{s,t}(x)^{\otimes 2} \\ & - M_s(Z_s(x))DW_{s,t}(x)DY_s(Z_s(x))M_s(Z_s(x))W_{s,t}(x) + O(|t-s|^{2\alpha}), \qquad (3.5.21) \end{aligned}$$

where for all $i = 1, 2, \ldots, d$,

$$\begin{aligned}
&\left[DM_s(Z_s(x))M_s(Z_s(x))W_{s,t}(x)^{\otimes 2}\right]^i \\
&= \sum_{k_1,k_2,k_3=1}^{d} \frac{\partial M^{ik_2}(x)}{\partial x_{k_1}}(Z_s(x))M_s^{k_1k_3}(Z_s(x))W_{s,t}^{k_2}(x)W_{s,t}^{k_3}.
\end{aligned}$$

Therefore, combining formulas (3.5.18) - (3.5.20), we have

$$\begin{aligned}
Z_{s,t}(x) =& \frac{1}{2}DM_s(Z_s(x))M_s(Z_s(x))W_{s,t}(x)^{\otimes 2} + M_s(Z_s(x))DW_{s,t}(x)W_{s,t}(x) \\
&- M_s(Z_s(x))W_{s,t}(x) - M_s(Z_s(x))\mathbb{W}_{s,t}(x,x) + O(|t-s|^{3\alpha}). \qquad (3.5.22)
\end{aligned}$$

3.5.3 Rough partial differential equations

Let $\mathscr{C}^3_{loc}(\mathbb{R}^d;\mathbb{R})$ be the space of functions that are locally bounded and have locally bounded first, second and third derivatives. Suppose that $h \in \mathscr{C}^3_{loc}(\mathbb{R}^d;\mathbb{R})$. In this section, we will show that $u = \{u(t,x) = h(Z_t(x)), (t,x) \in [0,T] \times \mathbb{R}^d\}$, where $Z_t(x)$ is defined in Section 3.5.1, is a solution to equation (1.2.7). Moreover, the solution is unique if $h \in \mathscr{C}^4_{loc}(\mathbb{R}^d;\mathbb{R})$).

Definition 3.5.8. *Let $(W, \mathscr{W}) \in \mathscr{C}^\alpha([0,T]; \mathscr{C}^{\beta_3}(\mathbb{R}^d;\mathbb{R}^d))$, let $\mathbb{W} : [0,T]^2 \times (\mathbb{R}^d)^2 \to \mathbb{R}^d$ be given by (3.4.1), and let h be a real-valued function on $\mathbb{R}^d$. A function $u = \{u(t,x), (t,x) \in [0,T] \times \mathbb{R}^d\}$ is called a solution to equation (1.2.7) with initial condition h, if the following properties are satisfied:*

(i) $u(0,x) = h(x)$ *for all* $x \in \mathbb{R}^d$.

(ii) *u is twice spatially differentiable everywhere, and $Du(\cdot,x)$ is controlled by W for all $x \in \mathbb{R}^d$.*

(iii) *The following equality is true for all $(t,x) \in [0,T] \times \mathbb{R}^d$*

$$\begin{aligned}
u(t,x) =& h(x) - \int_0^t Du(r,x)W(dr,x) + \frac{1}{2}\int_0^t Du(r,x)d\langle\!\langle DW(x), W(x)\rangle\!\rangle_r \\
&+ \frac{1}{2}\int_0^t Du(r,x)d\langle\!\langle W(x), DW(x)\rangle\!\rangle_r + \frac{1}{2}\int_0^t D^2u(r,x)d\langle W(x)\rangle_r, \qquad (3.5.23)
\end{aligned}$$

where the first integral is defined as follows,

$$\int_0^t Du(r,x)W(dr,x) := \int_0^t Du(r,x)dW_r(\xi)\Big|_{\xi=x},$$

the quadratic compensators

$$\langle\!\langle DW(x), W(x)\rangle\!\rangle_{s,t} := \langle\!\langle DW, W\rangle\!\rangle_{s,t}(\xi_1,\xi_2)\big|_{(\xi_1,\xi_2)=(x,x)},$$

$$\langle\!\langle W(x), DW(x)\rangle\!\rangle_{s,t} := \langle\!\langle W, DW\rangle\!\rangle_{s,t}(\xi_1,\xi_2)\big|_{(\xi_1,\xi_2)=(x,x)},$$

and

$$\langle W(x)\rangle_{s,t} := \langle W\rangle_{s,t}(\xi_1,\xi_2)\big|_{(\xi_1,\xi_2)=(x,x)}$$

are defined by (3.1.15)*,* (3.1.17) *and* (3.1.18) *respectively, $D^2u(r,x)$ is considered as a linear operator from $\mathbb{R}^d \otimes \mathbb{R}^d \to \mathbb{R}$, that is*

$$D^2u(t,x)M = \sum_{i,j=1}^{d} \frac{\partial^2 u(t,x)}{\partial x_i \partial x_j} M^{ij},$$

for any $d \times d$ matrix $M = (M^{ij})_{i,j=1}^d$, and the last three integrals are in Young's sense.

In the next theorem, we will show that $h(Z_t)$, where Z_t is defined as in Lemma 3.5.6, is a solution to equation (1.2.7).

Theorem 3.5.9. *Let $(W, \mathscr{W}) \in \mathscr{C}^\alpha([0,T]; \mathscr{C}^{\beta_3}(\mathbb{R}^d;\mathbb{R}^d))$, and let $\mathbb{W}$ be given by* (3.4.1)*. Assume Hypothesis (H). Let Y be the solution to the equation (3.5.9), and let $Z_t = Y_t^{-1}$ for all $t \in [0,T]$. Suppose that $h \in \mathscr{C}^3_{loc}(\mathbb{R}^d;\mathbb{R})$. Then, $u(t,x) = h(Z_t(x))$ is a solution to* (1.2.7) *in the sense of Definition 3.5.8.*

Proof. We prove this theorem by checking every property in Definition 3.5.8. By assumption, we know that $u(0,x) = h(Z_0(x)) = h(x)$. In addition, since $h \in \mathscr{C}^3_{loc}(\mathbb{R}^d;\mathbb{R})$ and $Z_t(x)$ is twice spatial differentiable, we can show that

$$D[h(Z_t(x))] = (Dh)(Z_t(x))M_t(Z_t(x)) \tag{3.5.24}$$

and

$$D^2[h(Z_t(x))] = (D^2h)(Z_t(x))M_t(Z_t(x))^2 + (Dh)(Z_t(x))DM_t(Z_t(x)),$$

where $(Dh)(Z_t(x))DM_t(x)$ is a $d \times d$ matrix with components

$$[(Dh)(Z_t(x))DM_t(Z_t(x))]^{ij} = \sum_{k=1}^{d} \frac{\partial}{\partial x_k} h(Z_t(x)) \frac{\partial}{\partial x_j} M^{ki}(Z_t(x)).$$

Recall that $M_t(x)$ is the solution to the linear RDE (3.5.13). Then we can write

$$M_{s,t}(x) = -M_s(x)F_{s,t}(x) + O(|t-s|^{2\alpha}) = -M_s(x)DW_{s,t}(Y_s(x)) + O(|t-s|^{2\alpha}).$$

Combining this fact with (3.5.17), we can deduce that

$$\begin{aligned} M_t(Z_t(x)) - M_s(Z_s(x)) &= M_t(Z_t(x)) - M_s(Z_t(x)) + M_s(Z_t(x)) - M_s(Z_s(x)) \\ &= M_{s,t}(Z_t(x)) - DM_s(Z_s(x))Z_{s,t}(x) + O(|t-s|^{2\alpha}) \\ &= M_{s,t}(Z_s(x)) + [M_{s,t}(Z_t(x)) - M_{s,t}(Z_s(x))] \\ &\quad - DM_s(Z_s(x))M_s(Z_s(x))W_{s,t}(x) + O(|t-s|^{2\alpha}) \\ &= -M_s(Z_s(x))DW_{s,t}(x) - DM_s(Z_s(x))M_s(Z_s(x))W_{s,t}(x) + O(|t-s|^{2\alpha}). \end{aligned} \tag{3.5.25}$$

Let $M_t'(Z_t(x)) : \mathscr{C}^{\beta_3}(\mathbb{R}^d;\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^d$ be given by

$$M_t'(Z_t(x))\Phi := -M_t(Z_t(x))D\Phi(x) - DM_t(Z_t(x))M_t(Z_t(x))\Phi(x), \tag{3.5.26}$$

where

$$[DM_t(Z_t(x))M_t(Z_t(x))\Phi(x)]^{ij} = \sum_{k_1,k_2} \frac{\partial}{\partial x_{k_1}} M_t^{ij}(Z_t(x)) M_t^{k_1k_2}(Z_t(x))\Phi^{k_2}(x)$$

for any $(t,x) \in [0,T] \times \mathbb{R}^d$. We can show that

$$M'(Z(x)) \in \mathscr{C}^{\alpha}([0,T];\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_3}(\mathbb{R}^d;\mathbb{R}^d);\mathbb{R}^d \otimes \mathbb{R}^d)). \tag{3.5.27}$$

Thus formulas (3.5.25) - (3.5.27) imply that

$$(M(Z(x)),M'(Z(x))) \in \mathscr{D}_W^{2\alpha}(\mathscr{L}(\mathscr{C}^{\boldsymbol{\beta}_3}(\mathbb{R}^d;\mathbb{R}^d);\mathbb{R}^d \otimes \mathbb{R}^d)).$$

In a similar way, recalling (3.5.24) and (3.5.25), we can also deduce that

$$\begin{aligned}
D[h(Z_t(x))] - D[h(Z_s(x))] &= (Dh)(Z_t(x))M_t(Z_t(x)) - (Dh)(Z_s(x))M_t(Z_s(x)) \\
=&(Dh)(Z_t(x))[M_t(Z_t(x)) - M_s(Z_s(x))] + [(Dh)(Z_t(x)) - Dh(Z_s(x))]M_s(Z_s(x)) \\
=&(Dh)(Z_s(x))\big[-M_s(Z_s(x))DW_{s,t}(x) - DM_s(Z_s(x))M_s(Z_s(x))W_{s,t}(x)\big] \\
&- D^2h(Z_s(x))M_s(Z_s(x))W_{s,t}(x)M_s(Z_s(x)) + O(|t-s|^{2\alpha}).
\end{aligned}$$

As a consequence, $D(h(Z(x))) \in \mathscr{D}_W^{2\alpha}(\mathbb{R}^d)$ where the Gubinelli derivative

$$[Dh(Z(x))]' : \mathscr{C}^{\boldsymbol{\beta}_3}(\mathbb{R}^d;\mathbb{R}^d) \to \mathbb{R}^d$$

is given by

$$\begin{aligned}
[D(h(Z(x)))]'\Phi =& - (Dh)(Z_t(x))DM_t(Z_t(x))M_t(Z_t(x))\Phi(x) \\
&- (Dh)(Z_t(x))M_t(Z_t(x))D\Phi(x) \\
&- (D^2h)(Z_t(x))M_t(Z_t(x))\Phi(x)M_t(Z_t(x))
\end{aligned} \tag{3.5.28}$$

As a consequence, properties (i) and (ii) of Definition 3.5.8 are satisfied.

In the next step, we will prove equality (3.5.23) by a similar argument as in Theorem 3.5.1. For any $0 \le s \le t \le T$, as a consequence of Taylor's Theorem 3.1.11, we can write

$$\begin{aligned} h(Z_t(x)) - h(Z_s(x)) =& (Dh)(Z_s(x))Z_{s,t}(x) + \frac{1}{2}(D^2h)(Z_s(x))Z_{s,t}(x)^{\otimes 2} + O(|t-s|^{3\alpha}) \\ :=& I_1 + I_2 + O(|t-s|^{3\alpha}). \end{aligned} \tag{3.5.29}$$

By (3.5.22), we have

$$\begin{aligned} I_1 =& -(Dh)(Z_s(x))M_s(Z_s(x))W_{s,t}(x) - (Dh)(Z_s(x))M_s(Z_s(x))\mathbb{W}(x,x) \\ &+ \frac{1}{2}(Dh)(Z_s(x))DM_s(Z_s(x))M_s(Z_s(x))W_{s,t}(x)^{\otimes 2} \\ &+ (Dh)(Z_s(x))M_s(Z_s(x))DW_{s,t}(x)W_{s,t}(x) + O(|t-s|^{3\alpha}), \end{aligned} \tag{3.5.30}$$

and

$$I_2 = \frac{1}{2}\left[(D^2h)(Z_s(x))M_s(Z_s(x))W_{s,t}(x)\right] \cdot \left[M_s(Z_s(x))W_{s,t}(x)\right] + O(|t-s|^{3\alpha}), \tag{3.5.31}$$

where

$$\begin{aligned} &(Dh)(Z_s(x))DM_s(Z_s(x))M_s(Z_s(x))W_{s,t}(x)^{\otimes 2} \\ &= \sum_{k_1,\dots,k_4=1}^{d} \frac{\partial h}{\partial x_{k_1}}(Z_s(x))\frac{\partial M_s^{k_1k_2}}{\partial k_3}(Z_s(x))M_s^{k_3k_4}(Z_s(x))W_{s,t}^{k_2}(x)W_{s,t}^{k_4}(x). \end{aligned}$$

Recall that $D[h(Z(x))]$ is controlled by W with Gubinelli derivative given by (3.5.28). Due to Theorem 3.1.4, the integral $\int_s^t D[h(Z_r(x))]W(dr,x)$ is well-defined and it can be approximated as follows

$$
\begin{aligned}
\int_s^t D[h(Z_r(x))]W(dr,x) &= (Dh)(Z_s(x))M_s(Z_s(x))W_{s,t}(x) \\
&\quad - \left\{(D^2h)(Z_s(x))M_s(Z_s(x)), M_s(Z_s(x))\right\}\mathscr{W}_{s,t}(x,x) \\
&\quad - (Dh)(Z_s(x))M_s(Z_s(x))\mathbb{W}^*_{s,t}(x,x) \\
&\quad - (Dh)(Z_s(x))DM_s(Z_s(x))M_s(Z_s(x))\mathscr{W}_{s,t}(x,x) + O(|t-s|^{3\alpha}),
\end{aligned}
\tag{3.5.32}
$$

where $\left\{(D^2h)(Z_s(x))M_s(Z_s(x)), M_s(Z_s(x))\right\}$ is defined as in (3.5.11),

$$
\mathbb{W}^*_{s,t}(x,x) = \int_s^t DW_{s,r}(x)W(dr,x) = \mathscr{D}^{(1)}\mathscr{W}_{s,t}(x,x)
$$

and

$$
\begin{aligned}
&(Dh)(Z_s(x))DM_s(Z_s(x))M_s(Z_s(x))\mathscr{W}_{s,t}(x,x) \\
&\quad = \sum_{k_1,\dots,k_4=1}^{d} \frac{\partial h}{\partial x_{k_1}}(Z_s(x))\frac{\partial M_s^{k_1k_3}}{\partial x_{k_2}}(Z_s(x))M_s^{k_2k_4}(Z_s(x))\mathscr{W}_{s,t}^{k_3k_4}(x,x)
\end{aligned}
$$

Taking into account Definition 3.1.9 and Remark 3.1.10, we can write

$$
\langle W(x)\rangle_{s,t} = W_{s,t}(x)\otimes W_{s,t}(x) - 2\mathscr{W}_{s,t}(x,x),
$$

$$
\langle\!\langle DW(x), W(x)\rangle\!\rangle_{s,t} = DW_{s,t}(x)W_{s,t}(x) - 2\mathbb{W}^*_{s,t}(x,x)
$$

and

$$
\langle\!\langle W(x), DW(x)\rangle\!\rangle_{s,t} = DW_{s,t}(x)W_{s,t}(x) - 2\mathbb{W}_{s,t}(x,x).
$$

Therefore, combining (3.5.29) - (3.5.32), we have

$$\begin{aligned}
h(Z_t(x)) - h(Z_s(x)) + \int_s^t D[h(Z_r(x))]W(dr,x) \\
=& \frac{1}{2}(Dh)(Z_s(x))M_s(Z_s(x))\left[\langle\!\langle DW(x),W(x)\rangle\!\rangle_{s,t} + \langle\!\langle W(x),DW(x)\rangle\!\rangle_{s,t}\right] \\
&+ \frac{1}{2}\left\{(D^2h)(Z_s(x))M_s(Z_s(x)),M_s(Z_s(x))\right\}\langle W(x)\rangle_{s,t} \\
&+ \frac{1}{2}(Dh)(Z_s(x))DM_s(Z_s(x))M_s(Z_s(x))\langle W(x)\rangle_{s,t} + O(|t-s|^{3\alpha}).
\end{aligned}$$

On the other hand, by the theory of Young's integral, we can show that

$$\begin{aligned}
&\int_0^t D[h(Z_r(x))]d\langle\!\langle DW(x),W(x)\rangle\!\rangle_r + \int_0^t D[h(Z_r(x))]d\langle\!\langle W(x),DW(x)\rangle\!\rangle_r \\
&+ \int_0^t D^2[h(Z_r(x))]d\langle W(x)\rangle_r \\
=& (Dh)(Z_s(x))M_s(Z_s(x))\left[\langle\!\langle DW(x),W(x)\rangle\!\rangle_{s,t} + \langle\!\langle W(x),DW(x)\rangle\!\rangle_{s,t}\right] \\
&+ \left\{(D^2h)(Z_s(x))M_s(Z_s(x)),M_s(Z_s(x))\right\}\langle W(x)\rangle_{s,t} \\
&+ (Dh)(Z_s(x))DM_s(Z_s(x))M_s(Z_s(x))\langle W(x)\rangle_{s,t} + O(|t-s|^{3\alpha}).
\end{aligned}$$

It follows that (3.5.23) holds if $u(t,x) = h(Z_t(x))$ for all $(t,x) \in [0,T]\times\mathbb{R}^d$. □

Remark 3.5.10. *The formulation of equation (1.2.7) looks odd. We will provide a Brownian example to this equation in Section 3.6.2, and will see that it is an application of Itô's formula.*

In the next theorem, we will show that the solution is unique in the space $\mathscr{C}^{\alpha,3}_{loc}([0,T]\times\mathbb{R}^d)$ provided that $(W,\mathscr{W}) \in \mathscr{C}^{\alpha}([0,T];\mathscr{C}^{\beta_4}(\mathbb{R}^d;\mathbb{R}^d))$ and $h \in \mathscr{C}^4_{loc}(\mathbb{R}^d;\mathbb{R})$.

Theorem 3.5.11. *Let $(W,\mathscr{W}) \in \mathscr{C}^{\alpha}([0,T];\mathscr{C}^{\beta_3}(\mathbb{R}^d;\mathbb{R}^d))$, and let $\mathbb{W}$ be given by (3.4.1). Assume Hypothesis (H). Let $h \in \mathscr{C}^4_{loc}(\mathbb{R}^d;\mathbb{R})$. The solution to the RPDE (1.2.7) exists and is unique in the space $\mathscr{C}^{\alpha,3}_{loc}([0,T]\times\mathbb{R}^d;\mathbb{R})$.*

Proof. Firstly, we show the existence of the equation (1.2.7) in the space $\mathscr{C}^{\alpha,3}_{loc}([0,T]\times\mathbb{R}^d;\mathbb{R})$. Due to Theorem 3.5.9, it suffice to show that $h(Z) \in \mathscr{C}^{\alpha,3}_{loc}([0,T]\times\mathbb{R}^d;\mathbb{R})$.

Notice that $DZ_t(x) = M_t(Z_t(x))$, $D^2Z_t(x) = DM_t(Z_t(x))M_t(Z_t(x))$, and

$$\begin{aligned} D^3Z_t(Z_t(x)) =& D^2M_t(Z_t(x))M_t(Z_t(x))M_t(Z_t(x)) \\ &+ DM_t(Z_t(x))DM_t(Z_t(x))M_t(Z_t(x)) \end{aligned}$$

for all $(t,x) \in [0,T] \times \mathbb{R}^d$. Fix $x \in \mathbb{R}^d$, the functions $M_t(x)$, $DM_t(x)$, $D^2M_t(x)$ and $D^3M_t(x)$ are all solutions to corresponding linear RDEs driven by α-Hölder linear rough paths. Thus $M_t(x)$, $DM_t(x)$, $D^2M_t(x)$ and $D^3M_T(x)$ are all α-Hölder in time and locally bounded in space. Recall that $h \in \mathscr{C}^4_{loc}(\mathbb{R}^d;\mathbb{R})$. As a consequence $h(Z_t(x))$, $D[h(Z_t(x))]$, $D^2[h(Z_t(x))]$ and $D^3[h(Z_t(x))]$ are all α-Hölder in time and locally bounded in space. In other words, we can conclude that $h(Z) \in \mathscr{C}^{\alpha,3}_{loc}([0,T] \times \mathbb{R}^d;\mathbb{R})$.

In the next step, we will prove the uniqueness of RPDE (1.2.7). Suppose that $u \in \mathscr{C}^{\alpha,3}([0,T] \times \mathbb{R}^d;\mathbb{R})$ is a solution to (1.2.7). Let Y be the solution to RDE (3.5.9). Then, by Taylor's Theorem 3.1.11, we can write

$$\begin{aligned} u(t,Y_t(x)) - u(s,Y_s(x)) =& u_{s,t}(Y_s(x)) + Du_{s,t}(Y_s(x))Y_{s,t} + Du_s(Y_s(x))Y_{s,t}(x) \\ &+ \frac{1}{2}D^2u_s(Y_s(x))Y_{s,t}(x)^{\otimes 2} + O(|t-s|^{3\alpha}). \end{aligned} \tag{3.5.33}$$

Notice that as a solution to (1.2.7), u satisfies the following equality for all $x \in \mathbb{R}^d$,

$$u_{s,t}(x) = -Du_s(x)W_{s,t}(x) + O(|t-s|^{2\alpha}).$$

It follows that fix $x \in \mathbb{R}^d$, $u(x)$ is controlled by $W(x)$. As a consequence, $Du(x)$ is also controlled by $W(x)$ with the Gubinelli derivative $-D^2u_s(x)$. Therefore, the following estimate holds

$$\begin{aligned} u_{s,t}(Y_s(x)) = &-Du(s,Y_s(x))W_{s,t}(Y_s(x)) + D^2u(s,Y_s(x))\mathscr{W}_{s,t}(Y_s(x),Y_s(x)) \\ &+ \frac{1}{2}Du(s,x)\big[\langle\!\langle DW(Y_s(x)),W(Y_s(x))\rangle\!\rangle_{s,t} + \langle\!\langle W(Y_s(x)),DW(Y_s(x))\rangle\!\rangle_{s,t}\big] \\ &+ \frac{1}{2}D^2u(s,x)\langle DW(Y_s(x))\rangle_{s,t} + O(|t-s|^{3\alpha}). \end{aligned} \tag{3.5.34}$$

In addition, recall that Y is the solution to (3.5.9). Then, (3.5.34) implies that

$$\begin{aligned} Du_{s,t}(Y_s(x))Y_{s,t}(x) = &- D^2u(s,Y_s(x))W_{s,t}(Y_s(x))W_{s,t}(Y_s(x)) \\ &- Du(s,Y_s(x))DW_{s,t}(Y_s(x))W_{s,t}(Y_s(x)) + O(|t-s|^{3\alpha}). \end{aligned} \tag{3.5.35}$$

Also, we have the following estimates

$$\begin{aligned} Du_s(Y_s(x))Y_{s,t}(x) = &Du_s(Y_s(x))W_{s,t}(Y_s(x)) \\ &+ Du_s(Y_s(x))\mathbb{W}_{s,t}(Y_s(x),Y_s(x)) + O(|t-s|^{3\alpha}), \end{aligned} \tag{3.5.36}$$

and

$$D^2u_s(Y_s(x))Y_{s,t}(x)^{\otimes 2} = D^2u_s(Y_s(x))W_{s,t}(Y_s(x))^{\otimes 2} + O(|t-s|^{3\alpha}). \tag{3.5.37}$$

Combining (3.5.33) - (3.5.37), we have

$$u(t,Y_t(x)) - u(s,Y_s(x)) = O(|t-s|^{3\alpha}).$$

Because $\alpha \in (\frac{1}{3},\frac{1}{2}]$, it follows that $u(t,Y_t(x)) \equiv u(0,Y_0(x)) = h(x)$. In other words, $u(t,x) = u(t,Y_t(Z_t(x))) = h(Z_t(x))$ for all $(t,x) \in [0,T]\times\mathbb{R}^d$. This completes the proof of the theorem. □

3.6 Examples

In this section, we provide some examples for nonlinear rough paths and transport equation with compensators. More precisely, in section 3.6.1, we construct a nonlinear rough path as the composition of a smooth function and a linear rough path. Then, we prove the double integral term of the nonlinear rough path can be represented as the rough integral in the sense of the classical theory of linear rough paths by using Itô's formula. In Section 3.6.2, we give an example for the transport equation in the Brownian case, and show that the equation is just an application of Itô's formula.

3.6.1 Nonlinear rough paths as compositions of linear rough paths

In this section, we consider a special class of nonlinear rough paths that are constructed by compositions of some nonlinear functions and linear rough paths.

Definition 3.6.1. *Let m be a positive integer. The space $\mathscr{C}^{m,\boldsymbol{\beta}_n}_{loc}(V^2;V)$ is the collection of function $f : V^2 \to V$ that is m times differentiable in the first argument with locally bounded derivatives and n times differentiable in the second argument with the growth given by $\boldsymbol{\beta}_n$. That is, for any compact set $K \subset V$*

$$\|f\|_{K,m,\boldsymbol{\beta}_n} := \sum_{j=0}^{m}\sum_{k=0}^{n}\sup_{\substack{x\in K\\ y\in V}}\frac{\|D_2^k D_1^j f(x,y)\|_{\mathfrak{B}_{k+j}}}{(1+\|y\|_V)^{\beta_k}} < \infty \tag{3.6.1}$$

where D_1 and D_2 are the partial derivatives of the first and second argument, respectively, and $\mathfrak{B}_{k+j}$ is the corresponding linear space of derivatives.

Let $f \in \mathscr{C}^{m,\boldsymbol{\beta}_n}_{loc}(V^2;V)$, and let $\mathbf{X} = (X,\mathbb{X}) \in \mathscr{C}^{\alpha}([0,T];V)$ be a V-valued linear rough path. We aim to interpret $W(t,x) = f(X_t,x)$ as a nonlinear rough path with suitable parameters $m,n \in \mathbb{N}$. Due to Definition 3.2.3, an α-Hölder nonlinear rough path contains a α-Hölder continuous function W and a 2α-Hölder continuous function $\mathbb{W}$ that defines a version of following double integral:

$$\int_s^t DW(dr,y)W_{s,r}(x) := \mathbb{W}_{s,t}.$$

As $W(t,x) = f(X_t,x)$, we expect that $\mathbb{W}$ is defined via the theory of linear rough paths by the following expression

$$\mathbb{W}_{s,t}(x,y) := \int_s^t g(dr,y)(f(X_r,x)) - g_{s,t}(y)(f(X_s,x)), \tag{3.6.2}$$

where $g(t,y) = D_2 f(X_t,y)$ and $g_{s,t}(y) = g(t,y) - g(s,y)$. Applying Itô's formula for linear rough pahts (c.f. Theorem 3.4 of Keller & Zhang [53] for finite dimensional cases), the integral on the right-hand side of (3.6.2) can be defined as follows

$$\begin{aligned} \int_s^t g(dr,y)(f(X_r,x)) := &\int_s^t D_{21} f(X_r,y) f(X_r,x) d\mathbf{X}_r \\ &+ \frac{1}{2}\int_s^t D_{211} f(X_r,y) f(X_r,x) d\langle X\rangle_r. \end{aligned} \tag{3.6.3}$$

In the next proposition, we will show that $(W,\mathbb{W})$ is a nonlinear rough path where $W(t,x) = f(X_t,x)$ and $\mathbb{W}_{s,t}(x,y)$ is defined in (3.6.2).

Proposition 3.6.2. *Assume that $n \geq 1$, and $f \in \mathscr{C}^{3,\boldsymbol{\beta}_n}_{loc}(V^2;V)$. Suppose that $\mathbf{X} = (X,\mathbb{X}) \in \mathscr{C}^{\alpha}([0,T];V)$. Let $W(t,x) = f(X_t,x)$, and let $\mathbb{W}$ be defined by (3.6.2) and (3.6.3). Then $\mathbf{W} := (W,\mathbb{W}) \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T] \times V;V)$.*

Proof. We prove this proposition by checking the properties in Definition 3.2.3. Let K be the closed convex hull of the set $\{X_t, t \in [0,T]\}$. Then K is a compact subset in V.

(i) For any $k \in \{0,\ldots,n\}$ and $\mathbf{z}_k = (z_1,\ldots,z_k) \in V^k$, by Taylor's theorem 3.1.11, we can show that

$$\begin{aligned} \|D^k W_{s,t}(x)(\mathbf{z}_k)\|_V =& \|D_2^k f(X_t,x)(\mathbf{z}_k) - D_2^k f(X_s,x)(\mathbf{z}_k)\|_V \\ \leq& \sup_{\tau\in[0,1]} \|D_1 D_2^k f(\tau X_t + (1-\tau)X_s,x)(\mathbf{z}_k)(X_{s,t})\|_V \\ \leq& \|f\|_{K,3,\boldsymbol{\beta}_n}\Big(\prod_{i=1}^k \|z_i\|_V\Big)\|X\|_\alpha (1+\|x\|_V)^{\beta_k}|t-s|^\alpha. \end{aligned}$$

This implies that $W \in \mathscr{C}^{\alpha,\boldsymbol{\beta}_n}([0,T]\times V;V)$.

(ii) a) Fix $(x,y) \in V^2$. Set $h(z) = h^{x,y}(z) := D_{21}f(z,y)(f(z,x))$ for all $z \in V$. Then, h is an $\mathscr{L}(V;V)$-valued function on V. It is easy to verify that $h \in \mathscr{C}^2_{loc}(V;V)$. Let $Y_t = h(X_t)$, and let

$$Y_t' = Dh(X_t) = D_{211}f(X_t,y)(f(X_t,x)) + D_{21}f(X_t,y)D_1f(X_t,y),$$

for all $t \in [0,T]$, where $D_{21}f(X_t,y)D_1f(X_t,y)$ is considered as an operator on $V \times V$ with values in V, that is

$$D_{21}f(X_t,y)D_1f(X_t,y)(x_1,x_2) = D_{21}f(X_t,y)(D_1f(X_t,y)(x_2),x_1).$$

By Lemma 7.3 of Friz & Hairer [36], $(Y,Y') \in \mathscr{D}_X^{2\alpha}(\mathscr{L}(V;V))$. In addition, by Taylor's theorem 3.1.11, we can easily show that

$$\|Y'\|_\alpha \leq 2\|f\|^2_{K,3,\boldsymbol{\beta}_n}(1+\|x\|_V)^{\beta_0}(1+\|y\|_V)^{\beta_1}\|X\|_\alpha \tag{3.6.4}$$

and

$$\|R^Y\|_{2\alpha} \leq 2\|f\|^2_{K,3,\boldsymbol{\beta}_n}(1+\|x\|_V)^{\beta_0}(1+\|y\|_V)^{\beta_1}\|X\|^2_\alpha. \tag{3.6.5}$$

Let $\Xi_{s,t} := Y_sX_{s,t} + Y_s'\mathbb{X}_{s,t}$ for any $(s,t) \in [0,T]^2$. The following estimate follows from (3.6.4), (3.6.5) and Theorem 3.1.4:

$$\begin{aligned}\left\| \int_s^t Y_r d\mathbf{X}_r - \Xi_{s,t} \right\|_V \leq & k_\alpha(\|X\|_\alpha \|R^Y\|_{2\alpha} + \|\mathbb{X}\|_{2\alpha}\|Y'\|_\alpha)|t-s|^{3\alpha} \\ \leq & 2k_\alpha \big[\|f\|^2_{K,3,\boldsymbol{\beta}_n}(1+\|x\|_V)^{\beta_0}(1+\|y\|_V)^{\beta_1}\|X\|^3_\alpha \\ & + \|f\|^2_{K,3,\boldsymbol{\beta}_n}(1+\|x\|_V)^{\beta_0}(1+\|y\|_V)^{\beta_1}\|X\|_\alpha\|\mathbb{X}\|_{2\alpha}\big]|t-s|^{3\alpha}. \end{aligned} \tag{3.6.6}$$

On the other hand, by Taylor's theorem, there exists $\xi = cX_s + (1-c)X_t$ for some $c \in [0,1]$ such that

$$
\begin{aligned}
&\Xi_{s,t} - g_{s,t}(y)(f(X_s,x)) \\
=&D_{21}f(X_s,y)(f(X_s,y),X_{s,t}) + D_{211}f(X_s,y)(f(X_s,x),\mathbb{X}_{s,t}) \\
&+D_{21}f(X_s,y)D_1f(X_s,y)\mathbb{X}_{s,t} - \big(D_2f(X_t,y)(f(X_s,x)) - D_2f(X_s,y)(f(X_s,x))\big) \\
=&D_{211}f(X_s,y)(f(X_s,x),\mathbb{X}_{s,t}) + D_{21}f(X_s,y)D_1f(X_s,y)\mathbb{X}_{s,t} \\
&-\frac{1}{2}D_{211}f(\xi,y)(f(X_s,y),X_{s,t},X_{s,t}).
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\|\Xi_{s,t} - g_{s,t}(y)(f(X_s,x))\|_V \leq &\Big[2\|f\|^2_{K,3,\boldsymbol{\beta}_n}(1+\|x\|_V)^{\beta_0}(1+\|y\|_V)^{\beta_1}\|\mathbb{X}\|_{2\alpha} \\
&+\frac{1}{2}\|f\|^2_{K,3,\boldsymbol{\beta}_n}(1+\|x\|_V)^{\beta_0}(1+\|y\|_V)^{\beta_1}\|X\|^2_{\alpha}\Big]|t-s|^{2\alpha}.
\end{aligned} \tag{3.6.7}
$$

Finally, by definition

$$
\langle X\rangle_{s,t} = X_{s,t}\otimes X_{s,t} - 2\mathbb{X}_{s,t},
$$

which implies that

$$
\|\langle X\rangle\|_{2\alpha} \leq \|X\|^2_{\alpha} + 2\|\mathbb{X}\|_{\alpha}.
$$

Therefore, Young's integral term can be estimated as follows

$$
\begin{aligned}
\Big\|\int_s^t D_{211}f(X_r,y)f(X_r,x)d\langle X\rangle_r\Big\|_V \leq &\sup_{z\in K}\big\|D_{211}f(z,y)f(z,x)\langle X\rangle_{s,t}\big\|_V \\
\leq&\|f\|^2_{K,3,\boldsymbol{\beta}_n}(1+\|x\|_V)^{\beta_0}(1+\|y\|_V)^{\beta_1}(\|X\|^2_{\alpha}+2\|\mathbb{X}\|_{2\alpha}).
\end{aligned} \tag{3.6.8}
$$

Recall that

$$
\begin{aligned}
\mathbb{W}_{s,t}(x,y) &= \int_s^t g(dr,y)(f(X_r,x)) - g_{s,t}(y)(f(X_s,x)) \\
&= \int_s^t Y_r d\mathbf{X}_r - g_{s,t}(f(X_s,x)) + \frac{1}{2}\int_s^t D_{211}f(X_r,y)f(X_r,x)d\langle X\rangle_r.
\end{aligned}
$$

Thus by combining (3.6.6) - (3.6.8), we have

$$\|\mathbb{W}(x,y)\|_{2\alpha} \leq C(1+\|x\|_V)^{\beta_0}(1+\|y\|_V)^{\beta_1},$$

where the constant C depends on α, $\|f\|_{K,3,\boldsymbol{\beta}_n}$, $\|X\|_\alpha$ and $\|\mathbb{X}\|_{2\alpha}$.

(ii) b) The next step is to estimate the spatial derivatives of $\mathbb{W}$. Observe that $\mathbb{W}$ consists of three terms: the rough integral, Young's integral, and $g_{s,t}(y)(f(X_s,x))$. Consider $g_{s,t}(y)(f(X_s,x))$ as a function of $(x,y) \in V^2$. Then, for any $(z_1,z_2) \in V^2$,

$$\begin{aligned} Dg_{s,t}(y)(f(X_s,x))(z_1,z_2) =& [D_2 f(X_t,y) - D_2 f(X_s,y)](D_2 f(X_s,y)(z_1)) \\ &+ [D_{22} f(X_t,y) - D_{22} f(X_s,y)](f(X_s,y),z_2). \end{aligned}$$

For the rough integral term, we compute the derivative of its approximation. That is, for all $(z_1,z_2) \in V^2$,

$$\begin{aligned} D\Xi_{s,t}(z_1,z_2) = & D_{21} f(X_s,y)(D_2 f(X_s,x)(z_1),X_{s,t}) + D_{212} f(X_s,y)(f(X_s,x),X_{s,t},z_2) \\ & + D_{211} f(X_s,y)(D_2 f(X_s,x)(z_1),\mathbb{X}_{s,t}) D_{2112} f(X_s,y)(f(X_s,x),\mathbb{X}_{s,t},z_2) \\ & + D_{21} f(X_s,y) D_{12} f(X_s,x)(\mathbb{X}_{s,t},z_1) + D_{212} f(X_s,y) D_1 f(X_s,x)(\mathbb{X}_{s,t},z_2), \end{aligned}$$

where

$$D_{21} f(X_s,y) D_{12} f(X_s,x)(z_1,z_2,z_3) = D_{21} f(X_s,y)(z_1, D_{12} f(X_s,x)(z_2,z_3))$$

and

$$D_{212} f(X_s,y) D_1 f(X_s,x)(z_1,z_2,z_3) = D_{212} f(X_s,y)(D_1 f(X_s,x)(z_2),z_1,z_3).$$

By the sewing lemma, we can show that for all $0 \leq s \leq t \leq T$,

$$\sum_{|\pi|\to 0} D\Xi_{s,t} \to \mathscr{I}_{s,t}(D\Xi),$$

in $\mathscr{L}(V^2;V)$ uniformly on compact sets in $(x,y) \in V^2$. Therefore,

$$\mathscr{J}_{s,t}(D\Xi) = D\mathscr{J}_{s,t}(\Xi) = D\Big[\int_s^t D_{21}f(X_r,y)f(X_r,x)d\mathbf{X}_r\Big].$$

By a similar argument in (ii) a), we can show that

$$D\Big[\int_s^t D_{21}f(X_r,y)f(X_r,x)d\mathbf{X}_r\Big] - Dg_{s,t}(y)(f(X_s,x))$$

is 2α-Hölder continuous in time. Moreover, the growth is of order $\beta_0 \vee \beta_1$ in x, and $\beta_1 \vee \beta_2$ in y. Young's integral term can be also estimated by using the sewing lemma and get the same result. Finally, by iteration, we conclude that $\mathbb{W} \in \mathscr{C}_2^{2\alpha,\boldsymbol{\beta}^*_{n-1},\boldsymbol{\beta}^{**}_{n-1}}([0,T]^2 \times V^2;V)$.

(iii) Notice that the linear rough integral on the right hand side of (3.6.2) is additive, Chen's relation follows immediately. □

Let $W(t,x) = f(X_t,x)$ for all $(t,x) \in [0,T] \times V$. In the next lemma, we show that a rough function controlled by W is also controlled by X.

Lemma 3.6.3. *Suppose that $f \in \mathscr{C}^{3,\boldsymbol{\beta}_1}_{loc}(V^2;V)$ and $X \in \mathscr{C}^\alpha([0,T];V)$. Let $W(t,x) = f(X_t,x)$, and let $(Y,\dot{Y}) \in \mathscr{E}^{2\alpha}_W$ in the sense of Definition 3.2.5. Then $(Y,Y') \in \mathscr{D}^{2\alpha}_W(V)$ in the sense of Definition 3.1.3 for some $Y' \in \mathscr{C}^\alpha(V;\mathscr{L}(V;V))$.*

Proof. Let $R^Y : [0,T]^2 \to V$ be given by

$$R^Y_{s,t} = Y_{s,t} - W_{s,t}(\dot{Y}_s) = Y_{s,t} - [f(X_t,\dot{Y}_s) - f(X_s,\dot{Y}_s)]$$

for all $0 \leq s \leq t \leq T$. Then, $R^Y \in \mathscr{C}^{2\alpha}([0,T];V)$. Additionally, applying Taylor's theorem 3.1.11, one get

$$\|f(X_t,\dot{Y}_s) - f(X_s,\dot{Y}_s)\|_V \leq \|D_1 f(X_s,\dot{Y}_s)X_{s,t}\|_V + \sup_{\tau\in[0,1]} \frac{1}{2}\|D_1 1 f(\tau x_s + (1-\tau)X_t,\dot{Y}_s)X^{\otimes 2}_{s,t}\|_V$$

Let $Y':[0,T]\to\mathscr{L}(V;V)$ be given by $Y'_t := D_1 f(X_t,\dot{Y}_t)$ for any $t\in[0,T]$. Then it follows that

$$\|Y_{s,t}-Y'_sX_{s,t}\|_V \leq \frac{1}{2}\|D_11f(\tau x_s+(1-\tau)X_t,\dot{Y}_s)X_{s,t}^{\otimes 2}\|_V+\|R^Y_{s,t}\|_V. \tag{3.6.9}$$

On the other hand, by using Taylor's theorem 3.1.11 again, we have

$$\begin{aligned}\|Y'_{s,t}\|_{\mathscr{L}(V;V)} =&\|D_1f(X_t,\dot{Y}_t)-D_1f(X_s,\dot{Y}_s)\|_{\mathscr{L}(V;V)}\\ =&\sup_{\tau\in[0,1]}\big[\|D_{11}f(\tau X_t+(1-\tau)X_s,\tau\dot{Y}_t+(1-\tau)\dot{Y}_s)X_{s,t}\|_{\mathscr{L}(V;V)}\\ &+\|D_{12}f(\tau X_t+(1-\tau)X_s,\tau\dot{Y}_t+(1-\tau)\dot{Y}_s)\dot{Y}_{s,t}\|_{\mathscr{L}(V;V)}\big].\end{aligned} \tag{3.6.10}$$

Similarly as in Proposition 3.6.2, let K be the closed convex hull of $\{X_t, 0\leq t\leq T\}$. The equalities (3.6.9) and (3.6.10) yield that

$$\|Y'\|_\alpha \leq \|f\|_{K,2,\boldsymbol{\beta}_1}\big[(1+\|\dot{Y}\|_\infty)^{\beta_1}\|\dot{Y}\|_\alpha+(1+\|\dot{Y}\|_\infty)^{\beta_0}\|X\|_\alpha\big]$$

and

$$\|\widetilde{R}^Y\|_{2\alpha}\leq\frac{1}{2}\|f\|_{K,3,\boldsymbol{\beta}_n}(1+\|\dot{Y}\|_\infty)^{\beta_0}\|X\|_\alpha^2+\|R^Y\|_{2\alpha},$$

where $\widetilde{R}^Y_{s,t}:=Y_{s,t}-Y'_sX_{s,t}$ for all $0\leq s\leq t\leq T$. This completes the proof. □

In the next theorem, we prove the equivalence of linear and nonlinear rough integrals, provided that $(W,\mathbb{W})$ is given in Proposition 3.6.2.

Theorem 3.6.4. *Suppose that $f\in\mathscr{C}^{3,\boldsymbol{\beta}_2}_{loc}(V^2;V)$ and $\mathbf{X}=(X,\mathbb{X})\in\mathscr{C}^\alpha([0,T];V)$. Let $(W,\mathbb{W})$ be defined in Proposition 3.6.2, and let $(Y,\dot{Y})\in\mathscr{E}^{2\alpha}_W$. Then by Lemma 3.6.3, there exits $Y'=D_1f(X,\dot{Y})$, such that $(Y,Y')\in\mathscr{D}^{2\alpha}_X(V)$. In addition, the following equality holds for all $0\leq s\leq t\leq T$,*

$$\int_s^t W(dr,Y_r)=\int_s^t D_1f(X_r,Y_r)d\mathbf{X}_r+\frac{1}{2}\int_s^t D_{11}f(X_r,Y_r)d\langle X\rangle_r, \tag{3.6.11}$$

where the integral on the left hand side is the nonlinear rough integral in the sense of Theorem 3.2.10, the first integral on the right hand side is the linear rough integral in the sense of Theorem 3.1.4, and the last integral is Young's integral.

Proof. Let Ξ and $\widetilde{\Xi}$ be the approximation of left hand and right hand sides of (3.6.4) respectively. That is

$$\begin{aligned}\Xi_{s,t} =& W_{s,t}(Y_s) + D_{21}f(X_s,Y_s)f(X_s,\dot{Y}_s)X_{s,t} + D_{211}f(X_s,Y_s)f(X_s,\dot{Y}_s)\mathbb{X}_{s,t} \\ &+ D_{21}f(X_s,Y_s)D_1f(X_s,\dot{Y}_s)\mathbb{X}_{s,t} + \frac{1}{2}D_{211}f(X_s,Y_s)f(X_s,\dot{Y}_s)\langle X\rangle_{s,t} \\ &- [D_2f(X_t,Y_s)f(X_s,\dot{Y}_s) - D_2f(X_s,Y_s)f(X_s,\dot{Y}_s)]. \end{aligned} \tag{3.6.12}$$

and

$$\begin{aligned}\widetilde{\Xi}_{s,t} =& D_1f(X_s,Y_s)X_{s,t} + D_{11}f(X_s,Y_s)\mathbb{X}_{s,t} + D_{12}f(X_s,Y_s)D_1f(X_s,\dot{Y}_s)\mathbb{X}_{s,t} \\ &+ \frac{1}{2}D_{11}f(X_s,Y_s)\langle X\rangle_{s,t},\end{aligned}$$

where

$$D_{21}f(X_s,Y_s)D_1f(X_s,\dot{Y}_s)(z_1,z_2) = D_{21}f(X_s,Y_s)(D_1f(X_s,\dot{Y}_s)(z_2),z_1),$$

and

$$D_{12}f(X_s,\dot{Y}_s)D_1f(X_s,\dot{Y}_s)(z_1,z_2) = D_{12}f(X_s,Y_s)(z_1,D_1f(X_s,\dot{Y}_s)(z_2)).$$

By Theorem 3.1.4, 3.2.10 and Proposition 3.6.2, it is not hard to verify that

$$\|Z_{s,t} - \Xi_{s,t}\|_V + \|\widetilde{Z}_{s,t} - \widetilde{\Xi}_{s,t}\|_V = O(|t-s|^{3\alpha}),$$

where $Z_{s,t}$ and $\widetilde{Z}_{s,t}$ denotes the left and right hand side of (3.6.4). On the other hand, note that by definition $\langle X\rangle_{s,t} = X_{s,t}\otimes X_{s,t} - 2\mathbb{X}_{s,t}$. Thus by Taylor's theorem 3.1.11, we can show that

$$\begin{aligned}\Xi_{s,t} =& D_1 f(X_s,Y_s)X_{s,t} + \frac{1}{2}D_{11}f(X_s,Y_s)X_{s,t}\otimes X_{s,t} \\ &+ D_{21}f(X_s,Y_s)D_1 f(X_s,\dot{Y}_s)\mathbb{X}_{s,t} + O(|t-s|^{3\alpha}),\end{aligned}$$

and

$$\begin{aligned}\widetilde{\Xi}_{s,t} =& D_1 f(X_s,Y_s)X_{s,t} + \frac{1}{2}D_{11}f(X_s,Y_s)X_{s,t}\otimes X_{s,t} \\ &+ D_{12}f(X_s,\dot{Y}_s)D_1 f(X_s,\dot{Y}_s)\mathbb{X}_{s,t} + O(|t-s|^{3\alpha}).\end{aligned}$$

This yields that $Z_{s,t} = \widetilde{Z}_{s,t}$ for all $0\le s\le t\le T$. □

3.6.2 A Brownian example for the transport equation

In this section, we give a Brownian example for transport equation (1.2.7). Let $B=\{B_t, t\in[0,T]\}$ be a one-dimensional Brownian motion, and let $\mathbb{B}_{s,t} = \frac{1}{2}B_{s,t}^2 - \frac{1}{2}(t-s)$. In other words, $\mathbb{B}$ is a pathwise representation of the iterated Itô-Wiener integral

$$\mathbb{B}_{s,t} = \int_s^t B_{s,r}dB_r.$$

Let $W(t,x) = xB_t$, and let $\mathscr{W}_{s,t}(x,y) = xy\mathbb{B}_{s,t}$. Then, a similar argument as in Section 3.6.1 implies that $(W,\mathscr{W})$ is a $\mathscr{C}(\mathbb{R};\mathbb{R})$-valued rough path. Also, we can easily deduce that

$$\mathbb{W}_{s,t}(x,x) = x\mathbb{B}_{s,t}, \langle W(x,x)\rangle_{s,t} = x^2B_{s,t}^2 - 2x^2\mathbb{B}_{s,t} = x^2(t-s),$$

and

$$\langle\!\langle DW(x), W(x)\rangle\!\rangle_{s,t} = \langle\!\langle W(x), DW(x)\rangle\!\rangle_{s,t} = x(t-s).$$

This allows us to rewrite the transport equation (1.2.7) as follows

$$\frac{\partial}{\partial t}u(t,x)+\frac{\partial}{\partial x}u(t,x)x\dot{B}_t=\frac{\partial}{\partial x}u(t,x)x+\frac{1}{2}\frac{\partial^2}{\partial x^2}u(t,x)x^2. \tag{3.6.13}$$

On the other hand, RDE (3.5.9) in this Brownian setting can be reformulated as follows:

$$Y_t(x)+x+\int_0^t Y_r(x)dB_r.$$

By solving this linear SDE, we get

$$Y_t(x)=xe^{B_t-\frac{1}{2}t},$$

and its inverse

$$Z_t(x)=Y_t^{-1}(x)=xe^{-B_t+\frac{1}{2}t}.$$

For any $h\in\mathscr{C}_b^3(\mathbb{R};\mathbb{R})$, let $u(t,x)=h(Z_t(x))$. Then, Itô's formula yields that

$$u(t,x)=h(x)-\int_0^t h'(Z_s(x))Z_s(x)dB_s+\int_0^t h'(Z_s(x))Z_s(x)ds \tag{3.6.14}$$
$$+\frac{1}{2}\int_0^t h''(Z_s(x))Z_s(x)^2ds. \tag{3.6.15}$$

Using chain rule, we can write

$$\frac{\partial}{\partial x}u(t,x)=\frac{\partial}{\partial x}h(Z_t(x))=h'(Z_t(x))Z_t(x)x^{-1},$$

$$\frac{\partial^2}{\partial x^2}u(t,x)=h''(Z_t(x))Z_t(x)^2x^{-2}.$$

Therefore, (3.6.14) can be reformulated as follows

$$u(t,x)=h(x)-\int_0^t\frac{\partial}{\partial x}u(s,x)xdB_s+\int_0^t\frac{\partial}{\partial x}u(s,x)xds$$
$$+\frac{1}{2}\int_0^t\frac{\partial^2}{\partial x^2}u(s,x)x^2ds,$$

and thus u is a solution to equation (3.6.13).

As we have seen in the example, quadratic compensators $\mathbb{W}_{s,t}(x,x)$, $\langle\langle DW(x), W(x)\rangle\rangle_{s,t}$ and $\langle\langle W(x), DW(x)\rangle\rangle_{s,t}$ come from the quadratic variation of the Brownian motion. If we define the stochastic integral in the Stratonovich sense, namely, $\mathbb{B}_{s,t} = \frac{1}{2}B_{s,t}^2$, such quadratic compensators disappear. In other words, the solution formula provides a solution to the classical transport equation in the Stratonovich sense (c.f. [32, 34, 56] for similar equations).

Chapter 4

Parabolic Anderson model of Skorohod type

In this chapter, we study the following parabolic Anderson model of Skorohod type

$$\frac{\partial}{\partial t}u(t,x) = \frac{1}{2}\Delta u(t,x) + u(t,x) \diamond \frac{\partial}{\partial t}W(t,x), \tag{4.0.1}$$

where $\diamond$ is the Wick product, and W is a Gaussian random field that is fractional Brownian in time with Hurst parameter $H \in (0, \frac{1}{2}]$ and has correlation Q in space. By using the Feynman-Kac representation for the $L^p(\Omega)$ moments of the solution, we find the upper and lower bounds for the moments.

4.1 Preliminaries

Let $W = \{W(t,x), (t,x) \in \mathbb{R}_+ \times \mathbb{R}^d\}$ be a Gaussian random field defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with correlation

$$\mathbb{E}[W(t,x)W(s,y)] = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right)Q(x,y),$$

for all $s,t \in \mathbb{R}_+$ and $x,y \in \mathbb{R}^d$. We assume that the covariance function Q satisfies the following conditions:

Hypothesis (H1). *There exist constants* $\alpha \in (1-2H, 1]$ *and* $C_1 > 0$ *such that*

$$Q(x,x) + Q(y,y) - 2Q(x,y) \le C_1|x-y|^{2\alpha}, \tag{4.1.1}$$

for all $x, y \in \mathbb{R}^d$.

Hypothesis (H2). *There exist constants* $\beta \in [0,1)$ *and* $C_2 > 0$ *such that for any* $M > 0$,

$$\inf_{\min_{i=1,\dots,d}(|x_i| \wedge |y_i|) > M} Q(x,y) \geq C_2 M^{2\beta}. \tag{4.1.2}$$

On the other hand, let $\mathfrak{H}$ be the Hilbert space defined as the completion of the linear span of the indicator functions of rectangles of $\mathbb{R}_+ \times \mathbb{R}^d$ with respect to the inner product

$$\langle \mathbf{1}_{[0,t]\times[0,x]}, \mathbf{1}_{[0,s]\times[0,y]} \rangle_{\mathfrak{H}} = \frac{1}{2}\big(t^{2H} + s^{2H} - |t-s|^{2H}\big) Q(x,y),$$

for all $s,t \in \mathbb{R}_+$ and $x = (x_1, \dots, x_d), y = (y_1, \dots, y_d) \in \mathbb{R}^d$, where $\mathbf{1}_{[0,x]} = \prod_{i=1}^d \mathbf{1}_{[0,x_i]}$ and $\mathbf{1}_{[0,x_i]} = -\mathbf{1}_{[x_i,0]}$ if $x_i < 0$. For any function $h \in \mathfrak{H}$, we write

$$W(h) := \int_0^\infty \int_{\mathbb{R}^d} h(t,x) W(dt,dx),$$

where the integral is the Itô-Wiener integral. Then, $\{W(h), h \in \mathfrak{H}\}$ is an isonormal Gaussian process on $\mathfrak{H}$, that is, a centered Gaussian family with covariance

$$\mathbb{E}[W(h)W(\widehat{h})] = \langle h, \widehat{h} \rangle_{\mathfrak{H}},$$

for all $h, \widehat{h} \in \mathfrak{H}$. For any positive integer n, we write H_n for the Hermite polynomial on $\mathbb{R}$, that is,

$$H_n(x) = \frac{(-1)^n}{n!} e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

Let $\mathbf{H}_n$ be the closed linear subspace of $L^2(\Omega)$ generated by the set of random variables $\{H_n(W(h)), h \in \mathfrak{H}, \|h\|_{\mathfrak{H}} = 1\}$. The space $\mathbf{H}_n$ is called the n-th Wiener chaos. Denote by $\mathfrak{H}^{\otimes n}$ the n-fold tensor product space of $\mathfrak{H}$. We write I_n for the isometry map between $\mathfrak{H}^{\otimes n}$ (with the modified norm $\sqrt{n!}\|\cdot\|_{\mathfrak{H}^{\otimes n}}$) and $\mathbf{H}_n$, given by $I_n(h^{\otimes n}) = H_n(W(h))$. It is known (c.f. Lemma 1.1.1 and Theorem

1.1.2 of Nualart [71]) that

(i) $\mathbf{H}_n$ and $\mathbf{H}_m$ are orthogonal if $n \neq m$. That is

$$\mathbb{E}(FG) = 0, \quad \forall F \in \mathbf{H}_n, G \in \mathbf{H}_m, n \neq m.$$

(ii) Any square integrable W-measurable random variable F can by uniquely represented as the following orthogonal Wiener chaos expansion

$$F = \mathbb{E}(F) + \sum_{n=1}^{\infty} I_n(f_n), \tag{4.1.3}$$

where $f_n \in \mathfrak{H}^{\otimes n}$ are symmetric.

By above properties and the isometry between $\mathfrak{H}^{\otimes n}$ and $\mathbf{H}_n$, for any $F \in L^2(\Omega)$ has the chaos expansion (4.1.3), the following equality holds

$$\mathbb{E}(F^2) = \mathbb{E}(F)^2 + \sum_{n=1}^{\infty} n! \|f_n\|^2_{\mathfrak{H}^{\otimes n}}.$$

Let $F, G \in L^2(\Omega)$. Suppose that $F = \mathbb{E}(F) + \sum_{n=1}^{\infty} I_n(f_n)$ and $G = \mathbb{E}(G) + \sum_{m=1}^{\infty} I_m(g_m)$. Then, by definition, the Wick product of F and G can be written as the following expression, if the last series is convergent in $L^2(\Omega)$,

$$F \diamond G = E(F) \sum_{m=1}^{\infty} I_m(g_m) + E(G) \sum_{n=1}^{\infty} I_n(f_n) + \sum_{n,m=1}^{\infty} I_{n+m}(f_n \widetilde{\otimes} g_m),$$

where $f_n \widetilde{\otimes} g_m$ is the symmetrization of $f_n \otimes g_m$ in $\mathfrak{H}^{\otimes(n+m)}$.

Remark 4.1.1. *The assumption $F, G \in L^2(\Omega)$ does not imply the convergence of $F \diamond G$. We refer the readers to the book of Hu [40] for a detailed account on the Wick product and sufficient conditions for the existence of $F \diamond G$.*

Let $u = \{u(t,x), (t,x) \in \mathbb{R}_+ \times \mathbb{R}^d\}$ be a W-measurable random field. Suppose that $\mathbb{E}[u(t,x)^2] <$

∞ for all $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$. Then, $u(t,x)$ has a Wiener chaos expansion as follows

$$u(t,x) = \mathbb{E}(u(t,x)) + \sum_{n=1}^{\infty} I_n(h_n(\cdot,t,x)). \tag{4.1.4}$$

In the following, we define the Skorohod integral and the solution to the Skorohod type stochastic partial differential equation (SPDE) (4.0.1). For more details on this topic, we refer the readers to Hu and Nualart [46].

Definition 4.1.2. *A square integrable random field u of the the form (4.1.4) is called to be Skorohod integrable, if $\mathbb{E}(u) \in \mathfrak{H}$, $h_n \in \mathfrak{H}^{\otimes(n+1)}$ for all $n \geq 1$ and the series*

$$\delta(u) = \int_0^\infty \int_{\mathbb{R}^d} u(t,x)\delta W(t,x) := W(\mathbb{E}(u)) + \sum_{n=1}^{\infty} I_{n+1}(\widetilde{h}_n)$$

converges in $L^2(\Omega)$, where $\widetilde{h}_n$ is the symmetrization of h_n as an element in $\mathfrak{H}^{\otimes(n+1)}$. The collection of all such random fields is denoted by $Dom(\delta)$.

Definition 4.1.3. *Let u_0 be a bounded measurable function on $\mathbb{R}^d$. A random field $u = \{u(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d\}$ is said to be a (mild) solution to the SPDE (4.0.1) with initial condition u_0, if for any $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$ the random field*

$$\left\{ \mathbf{1}_{[0,t]}(s) \int_{\mathbb{R}^d} p_{t-s}(x-z)u(s,z)\mathbf{1}_{[0,z]}(y)dz, (s,y) \in \mathbb{R}_+ \times \mathbb{R}^d \right\}$$

is an element of $Dom(\delta)$, and the following equality holds almost surely,

$$u(t,x) = \int_{\mathbb{R}} p_t(x-y)u_0(y)dy + \int_0^\infty \int_{\mathbb{R}^d} \left(\mathbf{1}_{[0,t]}(s) \int_{\mathbb{R}^d} p_{t-s}(x-z)u(s,z)\mathbf{1}_{[0,z]}(y)dz \right) \delta W(s,y),$$

where $p_t(x) = (2\pi t)^{-\frac{d}{2}} e^{-\frac{|x|^2}{2t}}$ denotes the heat kernel on $\mathbb{R}^d$ and the last integral is the Skorohod integral in the sense of Definition 4.1.2.

4.2 Feynman-Kac formula, chaos expansion and the upper bound

Let B be a standard d-dimensional Brownian motion independent of W. For any $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, let $B_t^x = x + B_t$, and let $g_{t,x}^B : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ be given by

$$g_{t,x}^B(r,z) := \mathbf{1}_{[0,t]}(r)\mathbf{1}_{[0,B_{t-r}^x]}(z). \tag{4.2.1}$$

Then due to Theorem 2.2 of Chen et al. [14], we know that $g_{t,x} \in \mathfrak{H}$. Since the Feynman-Kac representation for the Stratonovich type equation has been already established in [14], then by the same argument as in Section 6 of Hu et al. [43], we can immediately derive the following theorem.

Theorem 4.2.1. *Suppose that Q satisfies Hypothesis (H1). Let B be a standard d-dimensional Brownian motion independent of W. For any $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, let $g_{t,x}^B$ be defined in (4.2.1). Then for any bounded measurable function u_0 on $\mathbb{R}^d$, the process $u = \{u(t,x), (t,x) \in \mathbb{R}_+ \times \mathbb{R}^d\}$ given by*

$$u(t,x) = \mathbb{E}\big[u_0(B_t^x)\exp\big(W(g_{t,x}^B) - \frac{1}{2}\|g_{t,x}^B\|_{\mathfrak{H}}^2\big)\big] \tag{4.2.2}$$

is the unique (mild) solution to (4.0.1) with initial condition u_0.

Remark 4.2.2. *We can further deduce that $u(t,x)$ has the following chaos expansion,*

$$u(t,x) = \sum_{n=0}^{\infty} I_n(h_n(t,x)),$$

with

$$h_n(t,x)(\mathbf{r},\mathbf{z}) = \frac{1}{n!}\mathbb{E}\big[u_0(B_t^x)g_{t,x}^{B^1}(r_1,z_1)\ldots g_{t,x}^{B^n}(r_n,z_n)\big],$$

where $\{B^k\}_{k\geq 1}$ are independent copies of B, $\mathbf{r} = (r_1,\ldots,r_n) \in \mathbb{R}_+^n$ and $\mathbf{z} = (z_1,\ldots,z_n) \in (\mathbb{R}^d)^n$.

The next theorem provides an upper bound for moments of the solution to (4.0.1).

Theorem 4.2.3. *Suppose that u_0 is bounded and Q satisfies Hypothesis (H1). Let u be the solution to equation (4.0.1). Then for all positive integer n, $t \geq 1$ and $x \in \mathbb{R}^d$, the following inequality holds,*

$$\mathbb{E}\left[u(t,x)^n\right] \leq C_x \exp\left(Cn^{\frac{2-\alpha}{1-\alpha}} t^{\frac{2H+\alpha}{1-\alpha}}\right),$$

where $C > 0$ depends on $d, H, \alpha, \|u_0\|_\infty$ and $C_x > 0$ depends on $d, H, \alpha, \|u_0\|_\infty$ and x.

Proof. Recall that $\{B^k\}_{k\geq 1}$ are independent d-dimensional Brownian motions and $g_{t,x}^{B^k}$ is defined in (4.2.1). By the Feynman-Kac formula (4.2.2), we can write the moment formula for the solution as follows

$$\mathbb{E}[u(t,x)^n] = \mathbb{E}^B\Big[\prod_{k=1}^n u_0(B_t^{k,x}) \exp\Big(\frac{1}{2} \sum_{1\leq i\neq j\leq n} \langle g_{t,x}^{B^i}, g_{t,x}^{B^j}\rangle_{\mathfrak{H}}\Big)\Big]. \tag{4.2.3}$$

Combining (4.2.3) and Theorem 3.1 in [14], we can deduce that

$$\mathbb{E}[u(t,x)^n] \leq \mathbb{E}^B\Big[\prod_{k=1}^n u_0(B_t^{k,x}) \exp\Big(\frac{1}{2} \sum_{1\leq i,j\leq n} \langle g_{t,x}^{B^i}, g_{t,x}^{B^j}\rangle_{\mathfrak{H}}\Big)\Big] \leq C_x \exp\left(Cn^{\frac{2-\alpha}{1-\alpha}} t^{\frac{2H+\alpha}{1-\alpha}}\right).$$

The proof of this theorem is completed. □

Remark 4.2.4. *An alternative proof of Theorem 4.2.3 can be established by the chaos expansion of the solution to the SPDE (4.0.1) and the hypercontractivity property of fixed Wiener chaos (c.f. Hu et al. [41]).*

4.3 Lower bound for the moments

In this section, we prove the following theorem, which provides a lower bound for the moments of the solution to the SPDE (4.0.1).

Theorem 4.3.1. *Suppose that u_0 is bounded, $\inf_{x\in\mathbb{R}^d} u_0 > 0$, and Q satisfies Hypotheses (H1) and (H2) with $\alpha = \beta$. Let u be the solution to equation (4.0.1). Then there exists a positive integer N*

depending on d, H and α, such that for all $n \geq N$, $t \geq 1$ and $x \in \mathbb{R}^d$, the following inequality holds,

$$\mathbb{E}\big[u(t,x)^n\big] \geq C_x \exp\Big(Cn^{\frac{2-\alpha}{1-\alpha}} t^{\frac{2H+\alpha}{1-\alpha}}\Big), \tag{4.3.1}$$

where $C > 0$ depends on $d, H, \alpha, \|u_0\|_\infty, \inf_{x\in\mathbb{R}^d} u_0$ and $C_x > 0$ depends on $d, H, \alpha, \|u_0\|_\infty$, $\inf_{x\in\mathbb{R}^d} u_0$ and x.

Proof. We follow the ideas of Chen et al. [14] and Hu et al. [41] to prove this theorem. Without loss of generality, we assume that $u_0 \equiv 1$. Recall that $\{B^k\}_{k\geq 1}$ are independent d-dimensional Brownian motions and $g_{t,x}^{B^k}$ is defined in (4.2.1). By the moment formula (4.2.3) and Lemma 4.2 of [14], there exist a Gaussian process $X = \{X(x), x \in \mathbb{R}^d\}$ with correlation $\mathbb{E}[X(x)X(y)] = Q(x,y)$ and an independent fractional Brownian motion $\widehat{B} = \{\widehat{B}_t, t \in \mathbb{R}\}$ with Hurst parameter H, such that

$$\mathbb{E}[u(t,x)^n] = \mathbb{E}^B \exp\left\{\mathbb{E}^{X,\widehat{B}}\Big[\frac{1}{2}\Big(\int_0^t \sum_{i=1}^n X(B_{t-s}^{i,x})d\widehat{B}_s\Big)^2\Big] - \frac{1}{2}\sum_{i=1}^n \|g_{t,x}^{B^i}\|_{\mathfrak{H}}^2\right\}. \tag{4.3.2}$$

Due to Lemma 4.3 of [14], we know that there exists a constant $C_H > 0$ depending on H such that

$$\mathbb{E}^{X,\widehat{B}}\Big[\frac{1}{2}\Big(\int_0^t \sum_{i=1}^n X(B_{t-s}^{i,x})d\widehat{B}_s\Big)^2\Big] \geq C_H\Big[\int_0^t \Big(\sum_{i,j=1}^n Q(B_s^{i,x}, B_s^{j,x})\Big)^{\frac{1}{2H}} ds\Big]^{2H}. \tag{4.3.3}$$

On the other hand, by (2.2) and (2.12) of [14], we have

$$\begin{aligned}\|g_{t,x}^{B^{i,x}}\|_{\mathfrak{H}}^2 = \mathbb{E}\big[I_1(g_{t,x}^{B^i,x})^2\big] =& H\int_0^t \theta^{2H-1}[Q(B_\theta^{i,x}, B_\theta^{i,x}) + Q(B_{t-\theta}^{i,x}, B_{t-\theta}^{i,x})]d\theta \\ &+ |\alpha_H| \int_0^t \int_0^\theta r^{2H-2}\widehat{Q}(\theta, \theta - r, B^{i,x}, B^{i,x})drd\theta,\end{aligned} \tag{4.3.4}$$

where $\alpha_H = 2H(2H-1)$ and

$$\widehat{Q}(u,v,\phi,\psi) = \frac{1}{2}[Q(\phi_u,\psi_u) + Q(\phi_v,\psi_v) - Q(\phi_u,\psi_v) - Q(\phi_v,\psi_u)].$$

Recall that Q satisfies Hypothesis (H1). Thus it is easy to deduce that

$$\widehat{Q}(\theta, \theta - r, B^{i,x}, B^{i,x}) \leq \frac{C_1}{2} |B^i_\theta - B^i_{\theta - r}|^{2\alpha} \tag{4.3.5}$$

and

$$|Q(x,y)| \leq (C_1^{1/2}|x|^\alpha + Q(0,0)^{1/2})(C_1^{1/2}|y|^\alpha + Q(0,0)^{1/2}). \tag{4.3.6}$$

To simplify the computations, we assume that $Q(0,0) = 0$. In the general case, the proof can be done in a similar way without significant differences. Let $M > 0$ and let $\varepsilon \in (0, \frac{1}{2})$. Consider the following events

$$G_0^1(M) = \Big\{ \inf_{\substack{1 \leq i \leq n, 1 \leq j \leq d \\ s \in [t/2, t]}} |B_s^{i,x,j}| \geq M \Big\}, \quad G_0^2(M) = \Big\{ \sup_{\substack{1 \leq i \leq n, 1 \leq j \leq d \\ s \in [0,t]}} |B_s^{i,x,j}| \leq 4M \Big\},$$

and

$$G_0^3(M) = \Big\{ \sup_{\substack{1 \leq i \leq n, 1 \leq j \leq d \\ 0 \leq v < u \leq t}} \frac{|B_u^{i,x,j} - B_v^{i,x,j}|}{|u - v|^{\frac{1}{2} - \varepsilon}} \leq \frac{16M}{t^{\frac{1}{2} - \varepsilon}} \Big\},$$

where $B^{i,x,j}$ denotes the j-th component of $B^{i,x}$ for $j = 1, \ldots, d$.

On $G_0^1(M)$, by Hypothesis (H2) and using the assumption that $\alpha = \beta$, we have the inequality

$$\begin{aligned} \Big[\int_0^t \Big(\sum_{i,j=1}^n Q(B_s^{i,x}, B_s^{j,x}) \Big)^{\frac{1}{2H}} ds \Big]^{2H} \geq & \Big[\int_{\frac{t}{2}}^t \big(C_2 n^2 |M|^{2\alpha} \big)^{\frac{1}{2H}} ds \Big]^{2H} \\ = & 2^{-2H} C_2 n^2 M^{2\alpha} t^{2H}. \end{aligned} \tag{4.3.7}$$

On $G_0^2(M)$, using (4.3.6), we get that

$$\int_0^t \theta^{2H-1}[Q(B_\theta^{i,x},B_\theta^{i,x})+Q(B_{t-\theta}^{i,x},B_{t-\theta}^{i,x})]d\theta \le \int_0^t \theta^{2H-1} 2C_1|4\sqrt{d}M|^{2\alpha}d\theta$$
$$=2^{4\alpha}d^\alpha H^{-1}C_1 M^{2\alpha}t^{2H}. \tag{4.3.8}$$

Finally, on $G_0^3(M)$, using (4.3.5), we get

$$\int_0^t\int_0^\theta r^{2H-2}\widehat{Q}(\theta,\theta-r,B^{i,x},B^{i,x})drd\theta \le \int_0^t\int_0^\theta r^{2H-2}\frac{C_1}{2}\Big(\frac{16\sqrt{d}M}{t^{\frac12-\varepsilon}}r^{\frac12-\varepsilon}\Big)^{2\alpha}drd\theta$$
$$=\frac{2^{8\alpha-1}d^\alpha C_1 M^{2\alpha}t^{2H}}{(2H+\alpha-2\alpha\varepsilon-1)(2H+\alpha-2\alpha\varepsilon)}. \tag{4.3.9}$$

Set $G_0(M)=\bigcap_{k=1}^3 G_0^k(M)$. Due to inequalities (4.3.2) - (4.3.4) and (4.3.7) - (4.3.9), we obtain

$$\mathbb{E}[u(t,x)^n] \ge \exp\big[(c_1n^2-c_2n)M^{2\alpha}t^{2H}\big]\mathbb{P}[G_0(M)], \tag{4.3.10}$$

where

$$c_1 = 2^{-2H}C_2C_H \quad \text{and} \quad c_2 = 2^{4\alpha-1}d^\alpha C_1 + \frac{2^{8\alpha-2}d^\alpha C_1|\alpha_H|}{(2H+\alpha-2\alpha\varepsilon-1)(2H+\alpha-2\alpha\varepsilon)}.$$

For any $x=(x_1,\dots,x_d)\in\mathbb{R}^d$, let $\{\widetilde{B}^{j,x_j}\}_{1\le j\le d}$ be independent one-dimensional Brownian motions such that $\widetilde{B}^{j,x_j}$ starts from x_j for all $j=1,\dots,d$. For any j, let $G^j(M)$ be the event given by

$$G^j(M) := \Big\{\inf_{s\in[t/2,t]}|\widetilde{B}_s^{j,x_j}|\ge M,\ \sup_{s\in[0,t]}|\widetilde{B}_s^{j,x_j}|\le 4M,\ \sup_{0\le v<u\le t}\frac{|\widetilde{B}_u^{j,x_j}-\widetilde{B}_v^{j,x_j}|}{|u-v|^{\frac12-\varepsilon}}\le\frac{16M}{t^{\frac12-\varepsilon}}\Big\}, \tag{4.3.11}$$

and denote $G(M)=\bigcap_{j=1}^d G^j(M)$. Since $\{B^{i,x}\}_{1\le i\le n}$ are independent d-dimensional Brownian motions starting at $x=(x_1,\dots,x_d)$, the following equality holds

$$\mathbb{P}[G_0(M)] = \mathbb{P}[G(M)]^n = \prod_{j=1}^d \mathbb{P}[G^j(M)]^n.$$

This allows us to rewrite (4.3.10) in the following way,

$$\mathbb{E}[u(t,x)^n] \geq \exp\left[(c_1 n^2 - c_2 n) M^{2\alpha} t^{2H}\right] \prod_{j=1}^d \mathbb{P}[G^j(M)]^n. \tag{4.3.12}$$

In order to estimate $\mathbb{P}[G^j(M)]$, we pin the Brownian motion $\widetilde{B}^{j,x_j}$ at $t/2$, and obtain that

$$\begin{aligned} \mathbb{P}[G^j(M)] &= \int_M^{4M} \mathbb{P}\left[G^j(M) \middle| \widetilde{B}^{j,x_j}_{t/2} = r\right] q_{t/2}(r - x_j) dr \\ &\geq \int_{2M}^{3M} \mathbb{P}\left[G^j(M) \middle| \widetilde{B}^{j,x_j}_{t/2} = r\right] q_{t/2}(r - x_j) dr, \end{aligned} \tag{4.3.13}$$

where $q_t(x) = (2\pi t)^{-\frac{1}{2}} \exp[-x^2/(2t)]$ is the one-dimensional heat kernel. Notice that conditioned on $\widetilde{B}^{j,x_j}_{t/2} = r$, the process $\{\widetilde{B}^{j,x_j}_s, s \in [0,t/2]\}$ is a Brownian bridge, denoted by $Y = \{Y_s, s \in [0,t/2]\}$, such that $Y_0 = x_j$ and $Y_{t/2} = r$. In addition, the process $\{\widetilde{B}^{j,x_j}_{t/2+s} - r, s \in [0,t/2]\}$, denoted by $Z = \{Z_s, s \in [0,t/2]\}$, is a standard Brownian motion independent of Y. Let $A_1,\dots,A_4$ be the events given by

$$A_1 = \left\{ \sup_{s\in[0,t/2]} |Z_s| \leq M \right\}, \quad A_2 = \left\{ \sup_{s\in[0,t/2]} |Y_s| \leq 4M \right\},$$

$$A_3 = \left\{ \sup_{0\leq u<v\leq t/2} \frac{|Z_u - Z_v|}{|u-v|^{\frac{1}{2}-\varepsilon}} \leq \frac{8M}{t^{\frac{1}{2}-\varepsilon}} \right\}, \quad A_4 = \left\{ \sup_{0\leq u<v\leq t/2} \frac{|Y_u - Y_v|}{|u-v|^{\frac{1}{2}-\varepsilon}} \leq \frac{8M}{t^{\frac{1}{2}-\varepsilon}} \right\}.$$

Observe that for any $0 \leq v < t/2 < u \leq t$, it is easy to see that

$$\frac{|\widetilde{B}^{j,x_j}_u - \widetilde{B}^{j,x_j}_v|}{|u-v|^{\frac{1}{2}-\varepsilon}} \leq 2\max\left\{ \frac{|\widetilde{B}^{j,x_j}_{t/2} - \widetilde{B}^{j,x_j}_v|}{|t/2-v|^{\frac{1}{2}-\varepsilon}}, \frac{|\widetilde{B}^{j,x_j}_u - \widetilde{B}^{j,x_j}_{t/2}|}{|u-t/2|^{\frac{1}{2}-\varepsilon}} \right\}.$$

It follows that conditional on $\widetilde{B}^{j,x_j}_{t/2} = r$,

$$\sup_{0\le v<u\le t} \frac{|\widetilde{B}^{j,x_j}_u - \widetilde{B}^{j,x_j}_v|}{|u-v|^{\frac{1}{2}-\varepsilon}} \le 2\max\left\{ \sup_{0\le v<u\le t/2} \frac{|Y_u - Y_v|}{|u-v|^{\frac{1}{2}-\varepsilon}}, \sup_{0\le v<u\le t/2} \frac{|Z_u - Z_v|}{|u-v|^{\frac{1}{2}-\varepsilon}} \right\},$$

and thus

$$\left\{ \sup_{0\le v<u\le t} \frac{|\widetilde{B}^{j,x_j}_u - \widetilde{B}^{j,x_j}_v|}{|u-v|^{\frac{1}{2}-\varepsilon}} \le \frac{16M}{t^{\frac{1}{2}-\varepsilon}} \right\} \supset A_3 \cap A_4. \tag{4.3.14}$$

Moreover, if we restrict $r \in [2M, 3M]$ as in (4.3.13), the following inclusion is true,

$$\left\{ \inf_{s\in[t/2,t]} |\widetilde{B}^{j,x_j}_s| \ge M, \sup_{s\in[0,t]} |\widetilde{B}^{j,x_j}_s| \le 4M \right\} \supset A_1 \cap A_2. \tag{4.3.15}$$

Therefore, by (4.3.11), (4.3.14) and (4.3.15), we have for $r \in [2M, 3M]$,

$$\mathbb{P}\big[G^j(M)\big|B^{j,x_j}_{t/2} = r\big] \ge \mathbb{P}\Big(\bigcap_{k=1}^4 A_k\Big).$$

Because Y and Z are independent, we can write

$$\begin{aligned}\mathbb{P}\Big(\bigcap_{k=1}^4 A_k\Big) =& 1 - \mathbb{P}\Big(\bigcup_{k=1}^4 A_k^c\Big) \ge 1 - \mathbb{P}\big(A_1^c \bigcup A_2^c\big) - \mathbb{P}\big(A_3^c \bigcup A_4^c\big) \\ =& \mathbb{P}(A_1)\mathbb{P}(A_2) + \mathbb{P}(A_3)\mathbb{P}(A_4) - 1.\end{aligned} \tag{4.3.16}$$

Estimation of $\mathbb{P}(A_1)$: It follows from Doob's martingale inequality that

$$\mathbb{P}(A_1) = 1 - \mathbb{P}(A_1^c) \ge 1 - M^{-2}\mathbb{E}(|Z_{t/2}|^2) = 1 - \frac{t}{2M^2}. \tag{4.3.17}$$

Estimation of $\mathbb{P}(A_3)$: Recall that $\varepsilon \in (0, 1/2)$. By Kolmogorov's continuity criterion (c.f. Theorem 3.1 of Friz and Hairer [36]), there exists a modification of Z, denoted by $\widetilde{Z}$, and a random variable K_ε, such that

$$\sup_{0\le u<v\le t/2}\frac{|\widetilde{Z}_u-\widetilde{Z}_v|}{|u-v|^{\frac{1}{2}-\varepsilon}}\le K_\varepsilon \quad \text{and} \quad \mathbb{E}(|K_\varepsilon|^{\frac{2}{\varepsilon}})\le C_\varepsilon t^2,$$

where $C_\varepsilon>0$ is a constant depending only on ε. Combining this fact with Chebyshev's inequality, we have

$$\mathbb{P}(A_3)=1-\mathbb{P}(A_3^c)\ge 1-\Big(\frac{8M}{t^{\frac{1}{2}-\varepsilon}}\Big)^{-\frac{2}{\varepsilon}}\mathbb{E}(|K_\varepsilon|^{\frac{2}{\varepsilon}})\ge 1-2^{-\frac{6}{\varepsilon}}C_\varepsilon\frac{t^{\frac{1}{\varepsilon}}}{M^{\frac{2}{\varepsilon}}}. \tag{4.3.18}$$

Estimation of $\mathbb{P}(A_2)$: Let $\widetilde{B}$ be a one-dimensional standard Brownian motion. Then the Brownian bridge Y has the same distribution as the process $\widetilde{Y}=\{\widetilde{Y}_s, 0\le s\le t/2\}$ where

$$\widetilde{Y}_s=x_j+\widetilde{B}_s-\frac{2s}{t}(\widetilde{B}_{t/2}-r+x_j). \tag{4.3.19}$$

Thus, we can deduce that

$$\begin{aligned}\mathbb{P}(A_2^c)=&\mathbb{P}\Big[\sup_{0\le s\le t/2}\Big|\Big(1-\frac{2s}{t}\Big)x_j+\frac{2sr}{t}+\widetilde{B}_s-\frac{2s}{t}\widetilde{B}_{t/2}\Big|>4M\Big]\\ \le&\mathbb{P}\Big(\sup_{0\le s\le t/2}|\widetilde{B}_s|+|\widetilde{B}_{t/2}|>4M-r-|x_j|\Big)\le\mathbb{P}\Big(\sup_{0\le s\le t/2}|\widetilde{B}_s|>2M-\frac{r+|x_j|}{2}\Big).\end{aligned}$$

Assume that $\frac{M}{2}>\max\{|x_1|,\dots,|x_n|\}$ and recall that $r\in[2M,3M]$. It follows that

$$\mathbb{P}(A_2)=1-\mathbb{P}(A_2^c)\ge 1-\mathbb{P}\Big(\sup_{0\le s\le t/2}|\widetilde{B}_s|>\frac{M}{4}\Big)\ge 1-\frac{8t}{M^2}. \tag{4.3.20}$$

Estimation of $\mathbb{P}(A_4)$: Due to (4.3.19) and the fact $r\in[2M,3M]$, we have

$$\begin{aligned}\frac{|\widetilde{Y}_u-\widetilde{Y}_v|}{|u-v|^{\frac{1}{2}-\varepsilon}}\le&\frac{|\widetilde{B}_u-\widetilde{B}_v|}{|u-v|^{\frac{1}{2}-\varepsilon}}+\frac{2|u-v|^{\frac{1}{2}+\varepsilon}}{t}(|\widetilde{B}_{t/2}|+r+|x_j|)\\ \le&K_\varepsilon^{\widetilde{B}}+\frac{2^{\frac{1}{2}-\varepsilon}(|\widetilde{B}_{t/2}|+\frac{7}{2}M)}{t^{\frac{1}{2}-\varepsilon}},\end{aligned}$$

for all $0 \le v < u \le t/2$, where $K_\varepsilon^{\widetilde{B}}$ is the almost surely upper bound of the $(\frac{1}{2}-\varepsilon)$-Hölder norm of $\widetilde{B}$ on $[0,t/2]$ and $\mathbb{E}[|K_\varepsilon^{\widetilde{B}}|^{\frac{2}{\varepsilon}}] \le C_\varepsilon t^2$. Therefore,

$$\begin{aligned}\mathbb{P}(A_4) =& 1-\mathbb{P}(A_4^c) = 1-\mathbb{P}\Big(\sup_{0\le u<v\le t/2}\frac{|\widetilde{Y}_u-\widetilde{Y}_v|}{|u-v|^{\frac{1}{2}-\varepsilon}} \ge \frac{8M}{t^{\frac{1}{2}-\varepsilon}}\Big)\\ \ge& 1-\mathbb{P}\Big(K_\varepsilon^{\widetilde{B}}+\frac{2^{\frac{1}{2}-\varepsilon}|\widetilde{B}_{t/2}|}{t^{\frac{1}{2}-\varepsilon}} \ge \frac{M}{t^{\frac{1}{2}-\varepsilon}}\Big) \ge 1-\big[2^{\frac{2}{\varepsilon}-1}C_\varepsilon+2^{\frac{2}{\varepsilon}-3}\mathbb{E}|\widetilde{B}_1|^{\frac{2}{\varepsilon}}\big]\frac{t^{\frac{1}{\varepsilon}}}{M^{\frac{2}{\varepsilon}}}.\end{aligned} \tag{4.3.21}$$

According to inequalities (4.3.17), (4.3.18), (4.3.20) and (4.3.21), and choosing

$$M \ge C_{1,\varepsilon}t^{\frac{1}{2}} := \max\Big\{\Big(\frac{8}{1-\sqrt{3}/2}\Big)^{1/2}, \Big(\frac{2^{\frac{2}{\varepsilon}-1}C_\varepsilon+2^{\frac{2}{\varepsilon}-3}\mathbb{E}|\widetilde{B}_1|^{\frac{2}{\varepsilon}}}{1-\sqrt{3}/2}\Big)^{\varepsilon/2}\Big\}t^{\frac{1}{2}}, \tag{4.3.22}$$

we can make $\mathbb{P}(A_k) \ge \sqrt{3}/2$ for all $k=1,\dots,4$. Thus by (4.3.16), we have

$$\mathbb{P}\big[G^j(M)|\widetilde{B}_{t/2}^{j,x_j}=r\big] = \mathbb{P}\Big(\bigcap_{k=1}^{4}A_k\Big) \ge \frac{1}{2}. \tag{4.3.23}$$

Plugging (4.3.23) into inequality (4.3.13) and recalling that M satisfies (4.3.22) and $\frac{M}{2} \ge |x_j|$, we can write

$$\mathbb{P}(G^j(M)) \ge \frac{1}{2}\int_{2M}^{3M} dr q_{t/2}(x^j-r) \ge \frac{M}{\sqrt{4\pi t}}e^{-\frac{16M^2}{t}} \ge \frac{C_{1,\varepsilon}}{\sqrt{4\pi}}e^{-\frac{16M^2}{t}} \ge e^{-\frac{16M^2}{t}}. \tag{4.3.24}$$

Combining (4.3.12) and (4.3.24), we have

$$\mathbb{E}[u(t,x)^n] \ge \exp\big[(c_1n^2-c_2n)M^{2\alpha}t^{2H}-c_3nM^2t^{-1}\big], \tag{4.3.25}$$

where $c_3 = 16d$. Let N be the smallest integer such that $c_1n-c_2>0$. Then, for any $n \ge N$, by maximizing the function

$$f(M) = (c_1n^2-c_2n)M^{2\alpha}t^{2H}-c_3nM^2t^{-1},$$

we find

$$M_0 = \big(\alpha(c_1 n - c_2)c_3^{-1} t^{2H+1}\big)^{\frac{1}{2-2\alpha}}, \tag{4.3.26}$$

such that

$$\begin{aligned}\sup_{M\geq 0} f(M) = f(M_0) =& (1-\alpha)\alpha^{\frac{\alpha}{1-\alpha}} c_3^{-\frac{\alpha}{1-\alpha}} n(c_1 n - c_2)^{\frac{1}{1-\alpha}} t^{\frac{2H+\alpha}{1-\alpha}} \\ \geq & (1-\alpha)\alpha^{\frac{\alpha}{1-\alpha}} c_3^{-\frac{\alpha}{1-\alpha}} (c_1 - c_2/N) n^{\frac{2-\alpha}{1-\alpha}} t^{\frac{2H+\alpha}{1-\alpha}}. \end{aligned} \tag{4.3.27}$$

Notice that for any $t \geq 1$ and $n \geq N$, the number M_0 given by (4.3.26) satisfies the following inequality

$$M_0 \geq \max\{\big(\alpha(c_1 N - c_2)c_3^{-1} t^{2H+\alpha}\big)^{\frac{1}{2-2\alpha}} t^{\frac{1}{2}}, \big(\alpha(c_1 n - c_2)c_3^{-1}\big)^{\frac{1}{2-2\alpha}} t^{\frac{1}{2}}\}. \tag{4.3.28}$$

Let

$$n_0(x) := \max\Big\{N, \frac{c_2\alpha + c_3[2\max\{|x_1|,\ldots,|x_d|\}]^{2-\alpha}}{c_1\alpha}, \frac{c_2\alpha + c_3 C_{1,\varepsilon}^{2-2\alpha}}{c_1\alpha}\Big\}$$

and let

$$t_0(x) := \max\Big\{1, \Big(\frac{c_3[2\max\{|x_1|,\ldots,|x_d|\}]^{2-2\alpha}}{\alpha(c_1 N - c_2)}\Big)^{\frac{1}{2H+\alpha}}, \Big(\frac{c_3 C_{1,\varepsilon}^{2-2\alpha}}{\alpha(c_1 N - c_2)}\Big)^{\frac{1}{2H+\alpha}}\Big\}.$$

Then, for any

$$(t,n) \in L_1 := \{(s,m) \in \mathbb{R}_+ \times \mathbb{N}, s \geq 1, m \geq n_0(x)\}, \tag{4.3.29}$$

or

$$(t,n) \in L_2 := \{(s,m) \in \mathbb{R}_+ \times \mathbb{N}, s \geq t_0(x), m \geq N\}, \tag{4.3.30}$$

by using (4.3.28), we have $\frac{M_0}{2} \geq \max\{|x_1|, \dots, |x_d|\}$ and $M_0 \geq C_{1,\varepsilon} t^{\frac{1}{2}}$. This implies that if $(t,n) \in L_1 \cup L_2$, inequality (4.3.25) is true when M is replaced by M_0. In this case, it follows from (4.3.27) that

$$\mathbb{E}[u(t,x)^n] \geq e^{f(M_0)} \geq \exp\left[(1-\alpha)\alpha^{\frac{\alpha}{1-\alpha}} c_3^{-\frac{\alpha}{1-\alpha}} (c_1 - c_2/N) n^{\frac{2-\alpha}{1-\alpha}} t^{\frac{2H+\alpha}{1-\alpha}}\right]. \tag{4.3.31}$$

On the other hand, let $M_1 = \max\{2|x_1|, \dots, 2|x_d|, C_{1,\varepsilon} t_0(x)^{\frac{1}{2}}\}$. Then for any

$$(t,n) \in L_3 := \{(s,m) \in \mathbb{R}_+ \times \mathbb{N}, 1 \leq s \leq t_0(x), N \leq m \leq n_0(x)\}, \tag{4.3.32}$$

inequality (4.3.25) is true when M is replaced by M_1. In this case, we can deduce that

$$\begin{aligned}
\mathbb{E}[u(t,x)^n] \geq & \exp\left[(c_1 n^2 - c_2 n) M_1^{2\alpha} t^{2H} - c_3 n M_1^2 t^{-1}\right] \\
\geq & \inf_{\substack{1 \leq t \leq t_0(x) \\ N \leq n \leq n_0(x)}} \left\{ \exp\left[(c_1 n^2 - c_2 n) M_1^{2\alpha} t^{2H} - c_3 n M_1^2 t^{-1} - C_0 n^{\frac{2-\alpha}{1-\alpha}} t^{\frac{2H+\alpha}{1-\alpha}}\right] \right\} \\
& \times \exp\left(C_0 n^{\frac{2-\alpha}{1-\alpha}} t^{\frac{2H+\alpha}{1-\alpha}}\right) \\
:= & C_x \exp\left(C_0 n^{\frac{2-\alpha}{1-\alpha}} t^{\frac{2H+\alpha}{1-\alpha}}\right),
\end{aligned} \tag{4.3.33}$$

where $C_0 = (1-\alpha)\alpha^{\frac{\alpha}{1-\alpha}} c_3^{-\frac{\alpha}{1-\alpha}} (c_1 - c_2/N)$. Notice that $\{t \geq 1, n \geq N\} = L_1 \cup L_2 \cup L_3$ where L_1, L_2 and L_3 are defined in (4.3.29), (4.3.30) and (4.3.32) respectively. Therefore, by (4.3.31) and (4.3.33), we have inequality (4.3.1). This completes the proof of this theorem. □

References

[1] ANDERSON, P. W. Absence of diffusion in certain random lattices. *Phys. Rev. 109*, 5 (1958), 1492. 11

[2] BAILLEUL, I., AND CATELLIER, R. Rough flows and homogenization in stochastic turbulence. *J. Differ. Equ. 263*, 8 (2017), 4894–4928. 9

[3] BAILLEUL, I., AND RIEDEL, S. Rough flows. *J. Math. Soc. Japan 71*, 3 (2019), 915–978. 9

[4] BAILLEUL, I., RIEDEL, S., AND SCHEUTZOW, M. Random dynamical systems, rough paths and rough flows. *J. Differ. Equ. 262*, 12 (2017), 5792–5823. 9

[5] BALLY, V., AND CARAMELLINO, L. Riesz transform and integration by parts formulas for random variables. *Stochastic Process. Appl. 121*, 6 (2011), 1332–1355. 4, 89, 90, 91, 92, 94

[6] BESALÚ, M., AND NUALART, D. Estimates for the solution to stochastic differential equations driven by a fractional Brownian motion with Hurst parameter $H \in (1/3, 1/2)$. *Stochastics and Dynamics 11*, 02n03 (2011), 243–263. 8, 126, 127

[7] BISWAS, N., ETHERIDGE, A., AND KLIMEK, A. The spatial Lambda-Fleming-Viot process with fluctuating selection. *arXiv preprint arXiv:1802.08188* (2018). 42

[8] BRAULT, A., AND LEJAY, A. The non-linear sewing lemma ii: Lipschitz continuous formulation. *arXiv preprint arXiv:1810.11988* (2018). 9

[9] BRAULT, A., AND LEJAY, A. The non-linear sewing lemma i: weak formulation. *Electron. J. Probab. 24*, 59 (2019), 1–24. 9

[10] BRAULT, A., AND LEJAY, A. The non-linear sewing lemma iii: Stability and generic properties. *HAL Preprint hal-2265268* (2019). 9

[11] CARMONA, R. A., AND MOLCHANOV, S. A. Parabolic Anderson problem and intermittency. *Mem. Amer. Math. Soc.*, 108 (1994), 1–125. 11

[12] CATELLIER, R. Rough linear transport equation with an irregular drift. *Stoch. Partial Differ. Equ. Anal. Comput. 4*, 3 (2016), 477–534. 11

[13] CATELLIER, R., AND GUBINELLI, M. Averaging along irregular curves and regularisation of odes. *Stochastic Process. Appl. 126*, 8 (2016), 2323–2366. 5, 6

[14] CHEN, L., HU, Y., KALBASI, K., AND NUALART, D. Intermittency for the stochastic heat equation driven by a rough time fractional gaussian noise. *Probab. Theory Related Fields 171*, 1-2 (2018), 431–457. 12, 173, 174, 175

[15] CHEN, X., HU, Y., NUALART, D., AND TINDEL, S. Spatial asymptotics for the parabolic Anderson model driven by a Gaussian rough noise. *Electron. J. Probab. 22* (2017). 42

[16] CHOUK, K., AND GUBINELLI, M. Nonlinear pdes with modulated dispersion ii: Korteweg–de Vries equation. *arXiv preprint arXiv:1406.7675* (2014). 5

[17] CHOUK, K., AND GUBINELLI, M. Nonlinear PDEs with modulated dispersion i: Nonlinear Schrödinger equations. *Comm. Partial Differ. Equ. 40*, 11 (2015), 2047–2081. 5

[18] DALANG, R. C., AND QUER-SARDANYONS, L. Stochastic integrals for spde's: a comparison. *Expo. Math. 29*, 1 (2011), 67–109. 45

[19] DAVIE, A. M. Differential equations driven by rough paths: an approach via discrete approximation. *Appl. Math. Res. Express 2008* (2008). 126

[20] DAWSON, D. A. Stochastic evolution equations and related measure processes. *J. Multivariate Anal. 5*, 1 (1975), 1–52. 2

[21] DAWSON, D. A. Infinitely divisible random measures and superprocesses. In *Stochastic analysis and related topics*, vol. 31 of *Progress in Probability*. Birkhäuser, Boston, 1992, pp. 1–129. 2, 29, 30

[22] DAWSON, D. A. Measure-valued Markov processes. In *École d'Eté de Probabilités de Saint-Flour XXI-1991*, vol. 1541 of *Lecture notes in Mathematics*. Springer-Verlag, Berlin, Heidelberg, 1993, pp. 1–260. 2

[23] DAWSON, D. A., AND HOCHBERG, K. J. The carrying dimension of a stochastic measure diffusion. *Ann. Probab. 7*, 4 (1979), 693–703. 4

[24] DAWSON, D. A., AND KURTZ, T. G. Applications of duality to measure-valued diffusion processes. In *Advances in filtering and optimal stochastic control*. Springer-Verlag, Berlin, Heidelberg, 1982, pp. 91–105. 36

[25] DAWSON, D. A., VAILLANCOURT, J., AND WANG, H. Stochastic partial differential equations for a class of interacting measure-valued diffusions. *Ann. Inst. Henri. Poincaré Probab. Stat. 36*, 2 (2000), 167–180. 3

[26] DELLACHERIE, C., AND MEYER, P.-A. *Probabilities and potential. B*, vol. 72 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1982. 31

[27] DYNKIN, E. B. *Markov Processes*. Springer-Verlag, Berlin, Heidelberg, 1965. 29

[28] DYNKIN, E. B. *An introduction to branching measure-valued processes*. No. 6 in CRM Monograph Series. Amer. Math. Soc., Providence, 1994. 2

[29] DYNKIN, E. B. *Diffusions, Superdiffusions and Partial Differential Equations*. No. 50 in Colloquium Publications. Amer. Math. Soc., Providence, 2002. 2

[30] ETHERIDGE, A. M. *An Introduction to Superprocesses*. No. 20 in University Lecture Series. Amer. Math. Soc., Providence, 2000. 2

[31] ETHIER, S. N., AND KURTZ, T. G. *Markov processes: characterization and convergence*. John Wiley & Sons, New York, 1986. 36

[32] FEDRIZZI, E., AND FLANDOLI, F. Noise prevents singularities in linear transport equations. *J. Funct. Anal. 264*, 6 (2013), 1329–1354. 168

[33] FEYEL, D., AND DE LA PRADELLE, A. Curvilinear integrals along enriched paths. *Electron. J. Probab. 11*, 34 (2006), 860–892. 97

[34] FLANDOLI, F., GUBINELLI, M., AND PRIOLA, E. Well-posedness of the transport equation by stochastic perturbation. *Invent. Math. 180*, 1 (2010), 1–53. 6, 168

[35] FRIEDMAN, A. *Partial differential equations of parabolic type*. Dover, Mineola, 2008. 37, 38, 40, 41

[36] FRIZ, P., AND HAIRER, M. *A course on rough paths*. Springer, Switzerland, 2014. 8, 95, 96, 97, 98, 99, 107, 160, 179

[37] GUBINELLI, M. Controlling rough paths. *J. Funct. Anal. 216*, 1 (2004), 86–140. 8, 97

[38] HAIRER, M. Solving the KPZ equation. *Ann. Math.* (2013), 559–664. 11

[39] HAVLIN, S., AND BEN-AVRAHAM, D. Diffusion in disordered media. *Adv. Phys. 36*, 6 (1987), 695–798. 11

[40] HU, Y. *Analysis on Gaussian spaces*. World Scientific, Singapore, 2016. 171

[41] HU, Y., HUANG, J., LÊ, K., NUALART, D., AND TINDEL, S. Parabolic Anderson model with rough dependence in space. In *The Abel Symposium* (2018), Springer, Cham, pp. 477–498. 42, 174, 175

[42] HU, Y., AND LÊ, K. Nonlinear Young integrals and differential systems in Hölder media. *Trans. Amer. Math. Soc. 369*, 3 (2017), 1935–2002. 5, 6, 138, 143

[43] HU, Y., LU, F., AND NUALART, D. Feynman-Kac formula for the heat equation driven by fractional noise with Hurst parameter $H < \frac{1}{2}$. *Ann. Probab. 40*, 3 (2012), 1041–1068. 6, 12, 173

[44] HU, Y., LU, F., AND NUALART, D. Hölder continuity of the solutions for a class of nonlinear SPDE's arising from one dimensional superprocesses. *Probab. Theory Related Fields 156*, 1-2 (2013), 27–49. 4

[45] HU, Y., AND NUALART, D. Rough path analysis via fractional calculus. *Trans. Amer. Math. Soc. 361*, 5 (2009), 2689–2718. 8

[46] HU, Y., AND NUALART, D. Stochastic heat equation driven by fractional noise and local time. *Probab. Theory Related Fields 143*, 1-2 (2009), 285–328. 172

[47] HU, Y., NUALART, D., AND SONG, J. Feynman-Kac formula for heat equation driven by fractional white noise. *Ann. Probab. 39*, 1 (2011), 291–326. 6

[48] HU, Y., NUALART, D., AND XIA, P. Hölder continuity of the solutions to a class of SPDE's arising from branching particle systems in a random environment. *Electron. J. Probab. 24* (2019). 1

[49] IKEDA, N., AND WATANABE, S. *Stochastic differential equations and diffusion processes*. North-Holland Publishing Co., Amsterdam, 1989. 26, 51, 75

[50] JACOD, J., AND SHIRYAEV, A. N. *Limit theorems for stochastic processes*, vol. 288 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, Heidelberg, 2003. 30

[51] JIŘINA, M. Branching processes with measure-valued states. In *Trans. Third Prague Conf. Information Theory, Statist. Decision Functions, Random Processes* (1964), Publ. House Czech Acad. Sci., Prague, pp. 333–357. 2

[52] KALLIANPUR, G., AND XIONG, J. *Stochastic differential equations in infinite dimensional spaces*, vol. 26 of *Lecture Notes-Monograph Series*. Institute of Mathematical Statistics, Hayward, 1995. 43, 44

[53] KELLER, C., AND ZHANG, J. Pathwise Itô calculus for rough paths and rough PDEs with path dependent coefficients. *Stochastic Process. Appl. 126*, 3 (2016), 735–766. 100, 146, 159

[54] KIRSCH, W. An invitation to random Schrödinger operators. *arXiv preprint arXiv:0709.3707* (2007). 11

[55] KÖNIG, W. *The parabolic Anderson model: Random Walk in Random Potential*. Pathways in Mathematics. Birkhäuser, Basel, 2016. 11

[56] KUNITA, H. *Stochastic flows and stochastic differential equations*, vol. 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge university press, 1990. 168

[57] KURTZ, T. G. The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities. *Electron. J. Probab 12*, 33 (2007), 951–965. 45

[58] KURTZ, T. G. Equivalence of stochastic equations and martingale problems. In *Stochastic analysis 2010*. Springer-Verlag, Berlin, Heidelber, 2011, pp. 113–130. 42

[59] LE GALL, J.-F. *Spatial Branching Processes, Random Snakes and Partial Differential Equations*. Springer, Basel AG, 1999. 2

[60] LEJAY, A. On rough differential equations. *Electron. J. Probab. 14* (2009), 341–364. 126, 138, 145

[61] LEJAY, A. Global solutions to rough differential equations with unbounded vector fields. In *Séminaire de probabilités XLIV*. Springer-Verlag, Berlin, Heidelber, 2012, pp. 215–246. 138

[62] LI, Z. *Measure-Valued Branching Markov Processes*. Springer-Verlag, Berlin, Heidelberg, 2011. 2

[63] LI, Z., WANG, H., XIONG, J., AND ZHOU, X. Joint continuity of the solutions to a class of nonlinear SPDEs. *Probab. Theory Related Fields 153*, 3-4 (2012), 441–469. 4, 69

[64] LYONS, T. Differential equations driven by rough signals (i): An extension of an inequality of LC Young. *Math. Res. Letters 1*, 4 (1994), 451–464. 7

[65] LYONS, T. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana 14*, 2 (1998), 215–310. 7

[66] LYONS, T., AND QIAN, Z. *System control and rough paths*. Oxford University Press, Oxford, 2002. 8

[67] MA, N., NUALART, D., AND XIA, P. Intermittency for the parabolic Anderson model of Skorohod type driven by a rough noise. *Electron. Commun. Probab. 25*, 48 (2020), 1 – 10. 1

[68] MALLIAVIN, P., AND THALMAIER, A. *Stochastic calculus of variations in mathematical finance*. Springer-Verlag, Berlin, Heidelberg, 2006. 90

[69] MITOMA, I. Tightness of probabilities on $C([0,1];\mathscr{S}')$ and $D([0,1];\mathscr{S}')$. *Ann. Probab. 11*, 4 (1983), 989–999. 30

[70] MYTNIK, L. Superprocesses in random environments. *Ann. Probab. 24*, 4 (1996), 1953–1978. 2, 3, 19

[71] NUALART, D. *The Malliavin Calculus and Related Topics*. Springer-Verlag, Berlin, Heidelberg, 2006. 46, 62, 80, 88, 171

[72] NUALART, D., AND NUALART, E. *Introduction to Malliavin calculus*. Cambridge University Press, Cambridge, 2018. 92

[73] NUALART, D., AND XIA, P. On nonlinear rough paths. *ALEA Lat. Am. J. Probab. Math. Stat. 17* (2020), 545 – 587. 1

[74] PERKINS, E. Dawson-Watanabe superprocesses and measure-valued diffusions. In *Lectures on probability theory and statistics*, vol. 1781 of *Lecture notes in Mathematics*. Birkhäuser, Basel, 2002, pp. 125–329. 2, 32

[75] SKOULAKIS, G., AND ADLER, R. J. Superprocesses over a stochastic flow. *Ann. Appl. Probab. 11*, 2 (2001), 488–543. 2, 33

[76] STROOCK, D. Some applications of stochastic calculus to partial differential equations. In *Ecole d'Eté de Probabilités de Saint-Flour XI-1981*, vol. 976 of *Lecture notes in Mathematics*. Springer-Verlag, Berlin, Heidelber, 1983, pp. 267–382. 47, 49, 145

[77] STROOCK, D. W., AND VARADHAN, S. R. On the support of diffusion processes with applications to the strong maximum principle. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 3. University of California Press, Berkeley, 1972, pp. 333–359. 42

[78] STURM, A. On convergence of population processes in random environments to the stochastic heat equations with colored noise. *Electron. J. Probab. 8*, 6 (2003), 1–39. 2, 3, 19, 25, 26, 27, 29, 30, 32

[79] WALSH, J. B. An introduction to stochastic partial differential equations. In *École d'Été de Probabilités de Saint Flour XIV-1984*. Springer-Verlag, Berlin, Heidelber, 1986, pp. 265–439. 23

[80] WANG, H. State classification for a class of measure-valued branching diffusions in a Brownian medium. *Probab. Theory Related Fields 109*, 1 (1997), 39–55. 2, 3, 14, 18, 39, 41, 42

[81] WANG, H. A class of measure-valued branching diffusions in a random medium. *Stochastic Anal. Appl. 16*, 4 (1998), 753–786. 2, 3, 42

[82] WATANABE, S. A limit theorem of branching processes and continuous state branching processes. *J. Math. Kyoto Univ. 8*, 1 (1968), 141–167. 2

[83] XIONG, J. A stochastic log-Laplace equation. *Ann. Probab. 32*, 3B (2004), 2362–2388. 33

[84] XIONG, J. *Three classes of nonlinear stochastic partial differential equations*. World Scientific, Singapore, 2013. 34

[85] YAMADA, T., AND WATANABE, S. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ. 11*, 1 (1971), 155–167. 45

[86] YOUNG, L. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math. 67*, 1 (1936), 251–282. 5

[87] ZANCO, G. A brief introduction to rough paths. http://pub.ist.ac.at/~gzanco/A_brief_introduction_to_rough_paths.pdf, 2016. 7

[88] ZEIDLER, E. *Applied functional analysis: main principles and their applications*, vol. 109. Springer-Verlag, New York, 1995. 101

[89] ZEL'DOVICH, Y. B., MOLCHANOV, S., RUZMAĬKIN, A., AND SOKOLOV, D. D. Intermittency in random media. *Soviet Phys. Uspekhi 30*, 5 (1987), 353–369. 11

[90] ZEL'DOVICH, Y. B., MOLCHANOV, S., RUZMAĬKIN, A., AND SOKOLOV, D. D. Intermittency, diffusion, and generation in a nonstationary random medium. *Sov. Sci. Rev. C. Math.Phys.*, 7 (1988), 1–110. 11