NONLINEAR YOUNG INTEGRALS
AND DIFFERENTIAL SYSTEMS IN HÖLDER MEDIA

YAOZHONG HU AND KHOA LÊ

Abstract. For H"older continuous random field $W(t, x)$ and stochastic process $\varphi_t$, we define nonlinear integral $\int_a^b W(dt, \varphi_t)$ in various senses, including pathwise and Itô-Skorohod. We study their properties and relations. The stochastic flow in a time dependent rough vector field associated with $\varphi_t = (\partial_t W)(t, \varphi_t)$ is also studied, and its applications to the transport equation $\partial_t u(t, x) - \partial_t W(t, x) \nabla u(t, x) = 0$ in rough media are given. The Feynman-Kac solution to the stochastic partial differential equation with random coefficients $\partial_t u(t, x) + Lu(t, x) + u(t, x) \partial_t W(t, x) = 0$ is given, where $L$ is a second order elliptic differential operator with random coefficients (dependent on $W$). To establish such a formula the main difficulty is the exponential integrability of some nonlinear integrals, which is proved to be true under some mild conditions on the covariance of $W$ and on the coefficients of $L$. Along the way, we also obtain an upper bound for increments of stochastic processes on multi-dimensional rectangles by majorizing measures.

1. Introduction

The Feynman integral is an important tool in quantum physics. The Feynman-Kac formula is a variant of the Feynman integral and plays a very important role in...
the study of (parabolic) partial differential equations (see [20] and [46]). Recently, there have been several successes in extending the Feynman-Kac formula to the following stochastic partial differential equations with noisy (random) potentials on $[0,T] \times \mathbb{R}^d$ (see e.g. [29], [32], and [35]):

$$\partial_t u(t, x) = \frac{1}{2} \Delta u(t, x) + u(t, x) \partial_t W(t, x),$$

where $\Delta$ is the Laplacian with respect to spatial variable and $\{\partial_t W(t, x), 0 \leq t \leq T, x \in \mathbb{R}^d\}$ is a Gaussian noise (the derivatives in the sense of Schwartz distribution of a Gaussian field). As indicated in the aforementioned papers, there are three tasks to accomplish for establishing the Feynman-Kac formula. The first one is to give a meaning to the nonlinear stochastic integral

$$\int_0^T W(ds, x + B_s),$$

to give a meaning to the nonlinear stochastic integral

$$\int_0^T W(ds, x + B_s),$$

and hence the Feynman-Kac expression (which we may call the Feynman-Kac solution) has a rigorous meaning. The final task is to show that the Feynman-Kac expression is indeed a solution to the equation in a certain sense. It should be emphasized that the independence between $B$ and $W$ played a crucial role in previous studies.

In many applications, one needs to study more general stochastic partial differential equations. For example, in modeling of the pressure in an oil reservoir in the Norwegian sea with a log normal stochastic permeability one was led to study the stochastic partial differential equation on some bounded domain in $\mathbb{R}^d$ of the form

$$\text{div}(k(x) \nabla u(x)) = f(x),$$

where the permeability $k(x)$ is the (Wick) exponential of white noise, $\text{div}$ is the divergence operator, and $\nabla$ is the gradient operator; see [27] and in particular the references therein. Recently, there has been a great amount of research on uncertainty quantification. Among the huge literature on this topic let us just mention the books [23], [50], and the references therein. Many different types of stochastic partial differential equations with random coefficients have been studied.

This motivates us to study the Feynman-Kac formula for general stochastic partial differential equations with random coefficients, namely,

$$\partial_t u(t, x) + Lu(t, x) + u(t, x) \partial_t W(t, x) = 0,$$

where

$$Lu(t, x) = \frac{1}{2} \sum_{i,j=1}^d a_{ij}(t, x, W) \partial^2_{x_i x_j} u(t, x) + \sum_{i=1}^d b_i(t, x, W) \partial_{x_i} u(t, x)$$

and for notational simplicity and up to a time change we assume that the terminal condition $u(T, x) = u_T(x)$ is given. The product $u(t, x) \partial_t W(t, x)$ in (1.1) is the ordinary product. If $\sigma(t, x) = (\sigma_{ij}(t, x, W))_{1 \leq i, j \leq d}$ satisfies $a = \sigma \sigma^T$ and if $X^{r,x}_t$ is the solution to the stochastic differential equation

$$dX^{r,x}_t = \sigma(t, X^{r,x}_t, W) \delta B_t + b(t, X^{r,x}_t, W) dt,$$

then $u(r, x) = E^B \left\{ u_T(X^{r,x}_T) \exp \left[ \int_r^T W(ds, X^{r,x}_s) \right] \right\}$ should be the Feynman-Kac solution to (1.1) with $u(T, x) = u_T(x)$. As indicated above, there are three tasks to complete to justify the above claim. The first task to give a meaning to the nonlinear stochastic integral $\int_r^T W(ds, X^{r,x}_s)$ is much more challenging than what has been accomplished before (see for instance [29], [32], and [35]), although the major focus of the work [32] is to give a meaning to the nonlinear integral $\int_r^T W(ds, X^{r,x}_s)$. However, in that paper $X^{r,x}_s = B_s$ is a Brownian motion independent of $W$, and
then we can consider $X^r_{s,x}$ as “deterministic”. In our current situation since $X^r_{s,x}$ and $W$ are correlated, the nonlinear integral is a true stochastic one. In addition, the noise $W$ may enter to $X^r_{s,x}$ in an anticipative way. Thus, the general stochastic calculus for semimartingales cannot be applied in a straightforward way due to the lack of adaptedness.

If $W(t,x)$ is only continuous in $t$ (without any Hölder continuity in $t$) but has certain differentiability on $x$, then we can use the semimartingale structure of $X^r_{t,x}$ plus some new techniques developed in Section 4 to define $\int^T_r W(ds,X^r_{s,x})$ and study the corresponding Feynman-Kac solution to (1.1). This result extends the work of [32] in two aspects. One is that the Laplacian is replaced by a general second order elliptic operator with general and in particular random coefficients. The other one is that in [32], the Hurst parameter $H$ in time is assumed to be greater than $1/4$, while the result of this paper is applicable to fractional Brownian field whose Hurst parameter $H$ in time can be any number between 0 and 1.

When $W(t,x)$ has certain (Hölder) regularity in the time variable, it is natural to see whether one can reduce its regularity in spatial variable $x$ to define $\int^T_r W(ds,X^r_{s,x})$. Having in mind the recent development on rough path analysis and encouraged by the previous success in the case when $X^r_{s,x}$ is the Brownian motion ([29], [32], and [35]), we dedicate ourselves to a systematic study of the nonlinear integral $\int^b_a W(ds,\varphi_s)$, where $W(s,x)$ is a Hölder continuous function on $s$ and $x$, and $\varphi_s$ is also a Hölder continuous function. Some elementary properties of the integral are obtained as well. These results are presented in Section 2. Let us emphasize that this nonlinear integral $\int^b_a W(ds,\varphi_s)$ is defined in a purely deterministic way. In fact, it is an extension of integration of Young type ([51]).

For Gaussian noise a very important (linear) stochastic integral is the Itô (or Itô-Skorohod) integral. It is also called a divergence integral. In probability theory, this integral is a central concept in stochastic analysis. For our stochastic partial differential equation (1.1) it is needed if the product $u(t,x)\partial_t W(t,x)$ there is a Wick product. We shall introduce the nonlinear Itô-Skorohod integral $\int^b_a W(ds,\varphi_s)$ ($\varphi_s$ depends on $W$) by using Malliavin calculus. This is done in Appendix A.

The relation of this integral with other types of integrals is also discussed in this section. Naturally, readers may ask the question how to study the Itô-Skorohod type stochastic differential equation $\partial_t u(t,x) + Lu(t,x) + u(t,x) \partial_t W(t,x) = 0$, where $u(t,x) \circ W(t,x)$ denotes the Wick product between $u(t,x)$ and $\partial_t W(t,x)$. However, this seems to be very complex since $L$ depends on $W$ in a sophisticated way and will not be considered in this work.

When $W(t,x)$ is a semimartingale in $t$ for any fixed $x$ and is smooth in $x$ for any fixed $t$, there have been many studies on stochastic flows which contribute significantly to the study of stochastic partial differential equations (see [32] and the references therein). The important tool there is the nonlinear stochastic integral (with respect to semimartingale) and the corresponding flow. After defining the nonlinear Young integral and motivated by this aspect, we study the pathwise flow associated with time dependent rough vector field $W(t,x)$. That is, we study the differential equation $\varphi_t = x + \int^t_0 W(ds,\varphi_s)$ under joint Hölder continuity assumptions of $W(t,x)$. We shall study the flow and other properties of the solution $\varphi_t$. This is presented in Section 3. The applications to the transport equation in rough media of the form $\partial_t u(t,x) - \partial_t W(t,x)\nabla u(t,x) = 0$ are also investigated in Subsection 3.4.
After completion of the first task of defining the nonlinear integral another major
difficulty (the above-mentioned second task) to overcome in the construction of the
Feynman-Kac solution is the exponential integrability of \( \int_r^T W(ds, X^{r,x}_s) \). In the
previous work of [22], [32], and [35], this is achieved by showing \( \mathbb{E}[u^2(r, x)] \) is finite.
If we continue to follow the idea in the aforementioned papers, then we are led to show
\[
\mathbb{E}_{B, B} \mathbb{E}^W \left\{ u_T(X^{r,x}_T) u_T(\hat{X}^{r,x}_T) \exp \left[ \int_r^T W(ds, X^{r,x}_s) + \int_r^T W(ds, \hat{X}^{r,x}_s) \right] \right\}
\]
is finite, where \( \hat{X}^{r,x}_t \) is the solution to the equation (1.2) with a Brownian motion \( B \),
independent of \( B \) and \( W \). It seems to us that in our situation, due to
the dependence of \( X^{r,x}_t \) on \( W \), it is hard to show the above quantity is finite.
To get around this difficulty, our strategy is then to show that \( u(r, x) = \mathbb{E}^B \left\{ u_T(X^{r,x}_T) \exp \left[ \int_r^T W(ds, X^{r,x}_s) \right] \right\} \) is finite for every fixed path of \( W \), assuming
some mild pathwise conditions on \( W \) (see for instance (4.30)). The third (and the
last) task, to show that the Feynman-Kac solution is indeed a solution to (1.1), is
relatively easier and will be completed by using an approximation technique. All
these will be done in Section 4.

Intentionally, the paper is divided into two parts. The first three chapters can
be read without knowledge of probability theory. A single (rough) sample \( W(t, x) \)
satisfying some joint Hölder continuity and growth conditions is considered. For
instance, the (stochastic) partial differential equation (1.1), the nonlinear Young
integral (Definition 2.1), and the transport equation (3.24) are considered for every
fixed sample path \( W(t, x) \). Since \( W(t, x) \) is fixed, we also drop the dependence of
\( a_{ij}(t, x) \) and \( b_{1}(t, x) \) on \( W \) throughout the paper. So, the integrals and equations are
defined and studied for a (fixed) rough function. The stochastic partial differential
equation considered in Section 4 is for a single rough sample path. But Brownian
motion is used to represent the solution.

As a probabilist, one may ask whether a stochastic process satisfies the joint
Hölder continuity conditions together with the growth conditions assumed throughout
the paper. For instance, condition (1.30) in Section 3 requires the paths of \( W \) to satisfy
\[
(1.3) \quad |W(s, x) - W(s, y) - W(t, x) + W(t, y)| \leq C(1 + |x|^\beta + |y|^\beta)|t - s|^\tau |x - y|^\lambda
\]
for all \( s, t \in [0, T] \) and \( x, y \in \mathbb{R}^d \). We give a partial answer to this problem in Section
5 where an extra assumption \( |x - y| \leq \delta \) for a fixed constant \( \delta \) is imposed.
Pathwise boundedness and pathwise regularity (Hölder continuity) have been extensively
studied in the literature (see Section 5 for more detailed discussions.) However, esti-
mates similar to (1.30) have not been studied thoroughly. Comparing with the ex-
isting literature (e.g. [11], [19]), where estimates for increments over one parameter
interval are obtained, the left side of (1.3) is an increment over a two parameter re-
tangle. Difficulties arise because the increments behave differently when the number
of parameters gets large. For instance, the corresponding entropic volumetric to the
left side of (1.3), \( d((s, x), (t, y)) = (\mathbb{E}|W(s, x) - W(s, y) - W(t, x) + W(t, y)|^2)^{1/2} \),
do not satisfy the triangular inequality. Therefore, classical estimates (such as those
appearing in [49]) are no longer applicable; new tools are needed to prove
(1.30). If in (1.3), \( x, y \) are restricted in a compact set, a similar problem has been
considered by the authors by extending the Garsia-Rodemich-Rumsey inequality.
Nevertheless, the exact growth rate when \( x, y \) get large is not discussed in that paper. Motivated by this requirement, we extend and sharpen our previous work in [30] so that it is applicable to our current situation. Since in many applications, \( W \) will be a Gaussian noise, we focus on the case where \( W \) satisfies normal concentration inequalities to obtain the desirable pathwise property from the covariance structure of the process. As is well-known it is usually hard to obtain properties for each sample path in the theory of stochastic processes. We hope this work will shed some light along this direction.

**Notation.** We collect here some notation that we will use throughout the entire paper. \( A \lesssim B \) means there is a constant \( C \) such that \( A \leq CB \). We represent a vector \( x \) in \( \mathbb{R}^d \) as a matrix of dimension \( d \times 1 \); \( A^T \) represents the transpose of a matrix \( A \). Sometimes we write \( x_i \) for column vector \( x^T \) and \( x^i \) for the row vector \( x \). We use the Einstein convention on summation over repeated indices. For instance, \( b_ic_i \) is the abbreviation for \( \sum_{i=1}^{d} b_ic_i \).

### 2. Nonlinear Young integral

Let \( W \) and \( \varphi \) be \( \mathbb{R}^d \)-valued functions defined on \( \mathbb{R} \times \mathbb{R}^d \) and \( \mathbb{R}^d \) respectively. We define in the current section the nonlinear Young integration \( \int W(ds, \varphi_s) \).

We make the following assumption on the regularity of \( W \):

**\((W)\)** There are constants \( \tau, \lambda \in (0, 1], \beta \geq 0 \) such that for all \( a < b \), the seminorm

\[
\|W\|_{\beta, \tau, \lambda; a, b} := \sup_{a \leq s < t \leq b, x, y \in \mathbb{R}^d; x \neq y} \frac{|W(s, x) - W(t, x) - W(s, y) + W(t, y)|}{(1 + |x| + |y|)^{\beta}|t - s|^\tau|x - y|^\lambda}
\]

(2.1)

is finite.

**About the function \( \varphi \), we assume**

**\((\phi)\)** \( \varphi \) is locally Hölder continuous of order \( \gamma \in (0, 1] \). That is, the seminorm

\[
\varphi_{\gamma; a, b} = \sup_{a \leq s < t \leq b} \frac{|\varphi(t) - \varphi(s)|}{|t - s|^\gamma}
\]

is finite for every \( a < b \).

Throughout the current section, we assume that \( \tau + \lambda \gamma > 1 \). Among three terms appearing in (2.1), we will pay special attention to the first term. Thus, we denote

\[
[W]_{\beta, \tau, \lambda; a, b} = \sup_{a \leq s < t \leq b, x, y \in \mathbb{R}^d; x \neq y} \frac{|W(s, x) - W(t, x) - W(s, y) + W(t, y)|}{(1 + |x| + |y|)^{\beta}|t - s|^\tau|x - y|^\lambda}.
\]

When \( \beta = 0 \), we denote \( \|W\|_{\tau, \lambda; a, b} := \|W\|_{0, \tau, \lambda; a, b} \). If \( a, b \) are clear in the context, we frequently omit the dependence on \( a, b \). For instance, \( \|W\|_{\beta, \tau, \lambda} \) is an abbreviation for \( \|W\|_{\beta, \tau, \lambda; a, b} \). \( \|\varphi\|_{\gamma} \) is an abbreviation for \( \|\varphi\|_{\gamma; a, b} \), and so on. We shall assume that \( a \) and \( b \) are finite. It is easy to see that for any \( c \in [a, b] \),

\[
\sup_{a \leq t \leq b} |\varphi(t)| = \sup_{a \leq t \leq b} |\varphi(c) + \varphi(t) - \varphi(c)| \leq |\varphi(c)| + \|\varphi\|_{\gamma}|b - a|^{-\gamma} < \infty.
\]
Thus assumption \( (\phi) \) also implies that
\[
\| \varphi \|_{\infty; a, b} := \sup_{a \leq t \leq b} |\varphi(t)| < \infty.
\]

For the results presented in this section, the condition \( \boxed{W} \) can be relaxed to:

\( \boxed{W'} \) There are constants \( \tau, \lambda \in (0, 1] \) such that for all \( a < b \) and compact set \( K \) in \( \mathbb{R}^d \), the seminorm
\[
\sup_{a \leq s < t \leq b} \frac{|W(s, x) - W(t, x) - W(s, y) + W(t, y)|}{|t - s|^\tau |x - y|^\lambda} + \sup_{a \leq t \leq b} \frac{|W(t, y) - W(t, x)|}{|x - y|^\lambda}
\]
is finite.

However, the polynomial growth rate is needed in the following sections to solve differential equations.

For later purposes, we denote by \( C^{(\tau, \lambda)}_b(\mathbb{R} \times \mathbb{R}^d) \) (respectively \( C^{(\tau, \lambda)}_{\text{loc}}(\mathbb{R} \times \mathbb{R}^d) \)) the collection of all functions \( W \) satisfying condition \( \boxed{W} \) (respectively \( \boxed{W'} \)). \( \kappa \) denotes a universal generic constant depending only on \( \lambda, \tau, \alpha \) and independent of \( W, \varphi \) and \( a, b \). The value of \( \kappa \) may vary from one occurrence to another.

2.1. Definition. We define the nonlinear integral \( \int W(ds, \varphi_s) \) as follows.

**Definition 2.1.** Let \( a, b \) be two fixed real numbers, \( a < b \). Let \( \pi = \{a = t_0 < t_1 < \cdots < t_m = b\} \) be a partition of \( [a, b] \) with mesh size \( |\pi| = \max_{0 \leq i \leq m-1} |t_{i+1} - t_i| \).

The Riemann sum corresponding to \( \pi \) is
\[
J_\pi = \sum_{i=1}^{m-1} W(t_{i+1}, \varphi_i) - W(t_i, \varphi_i).
\]

If the sequence of Riemann sums \( J_\pi \)'s is convergent when \( |\pi| \) shrinks to 0, we denote the limit as the nonlinear integral \( \int_a^b W(ds, \varphi_s) \).

We observe that in the particular case when \( W(t, x) = g(t)x \) for some functions \( g : \mathbb{R} \to \mathbb{R} \), the nonlinear integral \( \int_a^b W(ds, \varphi_s) \) defined above, if it exists, coincides with the Riemann-Stieltjes integral \( \int_a^b \varphi_s dg(s) \). It is well-known that if \( \varphi \) and \( g \) are Hölder continuous with exponents \( \alpha, \beta \) respectively and \( \alpha + \beta > 1 \), then the Riemann-Stieltjes integral \( \int_a^b \varphi_s dg(s) \) exists and is called the Young integral (\[51]\).

More generally, for each partition \( \pi \) of an interval \([a, b]\), one can consider the (abstract) Riemann sum
\[
J_\pi(\mu) = \sum_{i=1}^{m-1} \mu(t_i, t_{i+1})
\]
where \( \mu \) is a function defined on \([a, b]^2\) with values in a Banach space. A sufficient condition for convergence of the limit \( \lim_{|\pi| \to 0} J_\pi(\mu) \) is obtained by Gubinelli in \[24\] via the so-called sewing map. This point of view has important contributions to Lyons’ theory of rough paths (\[39\][40]). Since we will apply Gubinelli’s sewing lemma, we restate the result as follows.
Lemma 2.2 (Sewing lemma). Let \( \mu \) be a continuous function on \([a, b]^2\) with values in a Banach space \((B, \| \cdot \|)\) and let \( \varepsilon > 0 \). Suppose that \( \mu \) satisfies
\[
\| \mu(s, t) - \mu(s, c) - \mu(c, t) \| \leq K|t - s|^{1+\varepsilon} \quad \forall \ a \leq s \leq c \leq t \leq b.
\]
Then there exists a function \( J\mu(t) \) unique up to an additive constant such that
\[
\|J\mu(t) - J\mu(s) - \mu(s, t)\| \leq K(1 - 2^{-\varepsilon})^{-1}|t - s|^{1+\varepsilon} \quad \forall \ a \leq s \leq t \leq b.
\]
In addition, when \( |\pi| \) shrinks to 0, the Riemann sums \((2.3)\) converge to \( J\mu(b) - J\mu(a) \).

In what follows, we adopt the notation \( J^b_{a}\mu = J\mu(b) - J\mu(a) \). The map \( \mu \mapsto J\mu \) is called the sewing map. The setting of Lemma 2.2 is adopted from [17]. In several instances, one needs to prove a relation between two or more integrals. The following result provides a simple method for this problem.

Lemma 2.3. Suppose \( \mu_1 \) and \( \mu_2 \) are two functions as in Lemma 2.2. In addition, assume that
\[
|\mu_1(s, t) - \mu_2(s, t)| \leq C|t - s|^{1+\varepsilon'} \quad \forall \ a \leq s \leq t \leq b
\]
for some positive constant \( \varepsilon' \). Then \( J\mu_1 \) and \( J\mu_2 \) are different by an absolute constant. That is, \( J^t_{s}\mu_1 = J^t_{s}\mu_2 \) for all \( s, t \).

Proof. From Lemma 2.2, \( J(\mu_1 - \mu_2) = J\mu_1 - J\mu_2 \) and
\[
|J^t_{s}(\mu_1 - \mu_2)| \leq |\mu_1(s, t) - \mu_2(s, t)| + |t - s|^{1+\varepsilon}
\]
\[
\leq |t - s|^{1+\varepsilon'} + |t - s|^{1+\varepsilon}
\]
for all \( s, t \). This implies \( J^t_{s}(\mu_1 - \mu_2) = 0 \) for all \( s, t \).

Returning to our main objective of the current section, we consider
\[
\mu(s, t) = W(t, \varphi_s) - W(s, \varphi_s).
\]
Then the condition in Lemma 2.2 is guaranteed by (W) and (\( \phi \)). Indeed, for every \( s < c < t \),
\[
|\mu(s, t) - \mu(s, c) - \mu(c, t)| = |W(t, \varphi_s) - W(c, \varphi_s) - W(t, \varphi_c) + W(c, \varphi_c)|
\]
\[
\leq |W|_{\beta, \tau, \lambda}(1 + \|\varphi\|^{\beta}_{\infty})(t-s)^{\tau}\|\varphi_s - \varphi_c\|^{\lambda}
\]
\[
\leq |W|_{\beta, \tau, \lambda}(1 + \|\varphi\|^{\beta}_{\infty})\|\varphi\|^{\lambda}_{\tau; a, b}(t-s)^{\tau+\lambda \gamma}.
\]
Hence, by combining the sewing lemma and the previous estimate, we obtain

Proposition 2.4. Assuming the conditions (W), (\( \phi \)) with \( \tau + \lambda \gamma > 1 \), the sequence of Riemann sums \((2.2)\) is convergent when \( |\pi| \) goes to 0. In other words, the nonlinear integral \( \int_{a}^{b} W(ds, \varphi_s) \) is well-defined.

In addition, the following estimate holds:
\[
(2.5) \quad |\int_{s}^{t} W(dr, \varphi_r) - W(t, \varphi_c) + W(s, \varphi_c)|
\]
\[
\leq \kappa\|W\|_{\tau, \lambda; a, b}(1 + \|\varphi\|^{\beta}_{\infty})\|\varphi\|^{\lambda}_{\tau; a, b}(t-s)^{\tau+\lambda \gamma}
\]
for all \( a \leq s \leq c \leq t \leq b \).
Remark 2.5. After the completion of this work, we were made aware of the work [5] (and also [7,8,25]), where a similar nonlinear Young integral is studied. The objective of that paper is to define the averaging of the form \( \int_0^t f(X_u) du \) for some process \( X_u \) and for some irregular function \( f \). The sewing lemma that we follow is from [17], which is after the work of [24].

Remark 2.6. (i) In the particular case when \( W(t,x) = g(t)x \), Proposition 2.4 reduces to the existence of the Young integral \( \int \varphi_s dg(s) \). Hence, from now on we refer to the integral \( \int W(ds, \varphi_s) \) as the nonlinear Young integral. 

(ii) In Proposition 2.4, we can also consider the Riemann sums with right-end points 

\[
J^+_a = \sum_{i=0}^{m-1} [W(t_{i+1}, \varphi_{t_{i+1}}) - W(t_i, \varphi_{t_{i+1}})].
\]

Then the corresponding limit exists and equals \( \int_a^b W(ds, \varphi_s) \). This is a straightforward consequence of Lemma 2.3.

It is evident that 

\[
\int_s^t W(dr, \varphi_r) = \int_s^c W(dr, \varphi_r) + \int_c^t W(dr, \varphi_r) \quad \forall \ s < c < t.
\]

This together with (2.5) implies easily the following.

Proposition 2.7. Assume that \( W \) and \( \varphi \) hold with \( \lambda \gamma + \tau > 1 \). As a function of \( t \), the indefinite integral \( \{ \int_a^t W(ds, \varphi_s), \ a \leq t \leq b \} \) is Hölder continuous of exponent \( \tau \).

Fractional calculus is very useful in the study of the (linear) Young integral. It leads to some detailed properties of the integral and solution of a differential equation (see [33], [34], and the references therein). It is interesting to extend this approach to the nonlinear Young integral. In fact, the authors obtain in [31] the following presentation for the nonlinear Young integral by using fractional calculus. Since this method is not pursued in the current paper, we refer the reader to [31] for further details.

Theorem 2.8. Assume conditions \( W \) and \( \varphi \) are satisfied. In addition, we suppose that \( \lambda \gamma + \tau > 1 \). Let \( \alpha \in (1 - \tau, \lambda \tau) \). Then the following identity holds:

\[
(2.6)
\int_a^b W(dt, \varphi_t) = -\frac{1}{\Gamma(\alpha)\Gamma(1 - \alpha)} \left\{ \int_a^b \frac{W_{b-}(t, \varphi_t)}{(b-t)^{1-\alpha}(t-a)^{\alpha}} dt \right. \\
+ \alpha \int_a^b \int_a^t \frac{W_{b-}(t, \varphi_t) - W_{b-}(r, \varphi_r)}{(b-t)^{1-\alpha}(t-r)^{\alpha+1}} dr dt \\
+ (1 - \alpha) \int_a^b \int_t^b W(t, \varphi_t) - W(s, \varphi_t) ds dt \\
+ \alpha(1 - \alpha) \int_a^b \int_a^t \int_t^b \frac{W(t, \varphi_t) - W(s, \varphi_t) - W(t, \varphi_r) + W(s, \varphi_r)}{(s-t)^{2-\alpha}(t-r)^{\alpha+1}} ds dr dt \left\}
\]

where \( W_{b-}(t,x) = W(t,x) - W(b,x) \).
2.2. Mapping properties. Let \( \mu \) be a function as in Lemma 2.2. Let us define the quality

\[
[\mu]_{1+\varepsilon} = \sup_{s,c,t \in I: s < c < t} \frac{|\mu(s,t) - \mu(s,c) - \mu(c,t)|}{|t-s|^{1+\varepsilon}}.
\]

In several instances, given two functions \( \mu_1 \) and \( \mu_2 \) such that \( [\mu_1]_{1+\varepsilon} \) and \( [\mu_2]_{1+\varepsilon} \) are finite, one would like to compare the integrals \( J_{\mu_1} \) and \( J_{\mu_2} \). The following result answers this question.

**Lemma 2.9.** Let \( \mu_1 \) and \( \mu_2 \) be two continuous functions on \( [a,b]^2 \) such that \( [\mu_1]_\alpha \) and \( [\mu_2]_\alpha \) are finite for some \( \alpha > 1 \). Then for every \( s, t \in [a,b] \),

\[
|J_{\mu_1} - J_{\mu_2}| \leq |\mu_1(s,t) - \mu_2(s,t)| + (1 - 2^{1-\alpha})^{-1}[\mu_1 - \mu_2]_\alpha|s-t|^{\alpha}.
\]

**Proof.** The proof is rather trivial thanks to the linearity nature of Lemma 2.2. Put \( \mu = \mu_1 - \mu_2 \). Notice that \( [\mu]_\alpha \leq [\mu_1]_\alpha + [\mu_2]_\alpha < \infty \). Thus we can apply Lemma 2.2 to \( \mu \). The claim follows after observing that \( J_{\mu} = J_{\mu_1} - J_{\mu_2} \). \( \square \)

As an application, we study the dependence of the nonlinear Young integration \( \int W(ds, \varphi_s) \) with respect to the medium \( W \) and the integrand \( \varphi \).

**Proposition 2.10.** Let \( W_1 \) and \( W_2 \) be real valued functions on \( \mathbb{R} \times \mathbb{R}^d \) satisfying the condition \( [W]_c \). Let \( \varphi \) be a function in \( C^\gamma(\mathbb{R}; \mathbb{R}^d) \) and let \( \tau + \lambda \gamma > 1 \). Then

\[
\left| \int_a^b W_1(ds, \varphi_s) - \int_a^b W_2(ds, \varphi_s) \right| \leq |W_1(b, \varphi_b) - W_1(a, \varphi_a) - W_2(b, \varphi_b) + W_2(a, \varphi_a)|
\]

\[
+ c(\|\varphi\|_\infty)[W_1 - W_2]_{\beta,\tau,\lambda}\|\varphi\|_\gamma|b-a|^\tau + \lambda \gamma.
\]

**Proof.** Let \( a < c < b \). Put

\[
\mu_1(a,b) = W_1(b, \varphi_b) - W_1(a, \varphi_a),
\]

\[
\mu_2(a,b) = W_2(b, \varphi_b) - W_2(a, \varphi_a),
\]

\[
\mu = \mu_1 - \mu_2.
\]

The argument before Proposition 2.4 shows that

\[
[\mu]_{\tau + \lambda \gamma} \leq [W_1 - W_2]_{\beta,\tau,\lambda}(1 + \|\varphi\|_\infty^\alpha)|\varphi|_\gamma.
\]

The proposition follows from Lemma 2.9. \( \square \)

**Proposition 2.11.** Let \( W \) be a function on \( \mathbb{R} \times \mathbb{R}^d \) satisfying the condition \( [W]_c \). Let \( \varphi^1 \) and \( \varphi^2 \) be two functions in \( C^\gamma(\mathbb{R}; \mathbb{R}^d) \) and let \( \tau + \lambda \gamma > 1 \). Let \( \theta \in (0,1) \) such that \( \tau + \theta \lambda \gamma > 1 \). Then for any \( u < v \),

\[
\left| \int_u^v W(ds, \varphi_s^1) - \int_u^v W(ds, \varphi_s^2) \right|
\]

\[
\leq C_1[W]_{\beta,\tau,\lambda}\|\varphi_1 - \varphi_2\|_\infty^\lambda|v-u|^\tau
\]

\[
+ C_2[W]_{\beta,\tau,\lambda}\|\varphi_1 - \varphi_2\|_\infty^{\lambda(1-\theta)}|v-u|^\tau + \theta \lambda \gamma,
\]

where \( C_1 = 1 + \|\varphi^1\|_\infty^\beta + \|\varphi^2\|_\infty^\beta \) and \( C_2 = 2^{1-\theta}C_1(\|\varphi^1\|_\infty^\lambda + \|\varphi^1\|_\lambda^\gamma)^\theta \).
Proof. We put \( \mu_1(a, b) = W(b, \varphi_1^a) - W(a, \varphi_1^a) \), \( \mu_2(a, b) = W(b, \varphi_2^a) - W(a, \varphi_2^a) \) and \( \mu = \mu_1 - \mu_2 \). Applying Lemma 2.9 we obtain, for any \( \theta \in (0, 1) \) such that \( \tau + \theta \lambda \gamma > 1 \),
\[
| \int_0^v W(ds, \varphi_1^a) - \int_0^v W(ds, \varphi_2^a) | \\
\leq | W(v, \varphi_1^a) - W(u, \varphi_1^a) - W(v, \varphi_2^a) + W(u, \varphi_2^a) | \\
+ [\mu]_{\tau + \theta \lambda \gamma} | v - u |^{\tau + \theta \lambda \gamma} .
\]

Notice that
\[
| W(v, \varphi_1^a) - W(u, \varphi_1^a) - W(v, \varphi_2^a) + W(u, \varphi_2^a) | \leq C_1 | W |_{\beta, \tau, \lambda} | u - v |^{\tau} | \varphi_1^a - \varphi_2^a |_{\infty}^{\lambda} .
\]

It remains to estimate \( [\mu]_{\tau + \theta \lambda \gamma} \). It is obvious that for \( i = 1, 2 \),
\[
[\mu_i]_{\tau + \theta \lambda \gamma} \leq [ W |_{\beta, \tau, \lambda} (1 + | \varphi_i^a |_{\infty}^{\beta}) | \varphi_i^a |_{\gamma}^{\lambda} \leq C_1 | W |_{\beta, \tau, \lambda} | \varphi_i^a |_{\gamma}^{\lambda}
\]
and hence
\[
[\mu]_{\tau + \theta \lambda \gamma} \leq [\mu_1]_{\tau + \lambda \gamma} + [\mu_2]_{\tau + \lambda \gamma} \leq C_1 | W |_{\beta, \tau, \lambda} \sum_{i=1}^{2} | \varphi_i^a |_{\gamma}^{\lambda} .
\]

On the other hand,
\[
| \mu(a, b) - \mu(a, c) - \mu(c, b) | \\
\leq | W(b, \varphi_1^a) - W(b, \varphi_2^a) - W(c, \varphi_1^a) + W(c, \varphi_2^a) | \\
+ | W(b, \varphi_1^a) - W(b, \varphi_2^a) - W(c, \varphi_1^a) + W(c, \varphi_2^a) | \\
\leq 2C_1 | W |_{\beta, \tau, \lambda} | b - c |^{\tau} | \varphi_1^a - \varphi_2^a |_{\infty}^{\lambda} .
\]

Combining the two bounds for \( \mu \) we get for any \( \theta \in (0, 1) \) such that \( \tau + \theta \lambda \gamma > 1 \),
\[
[\mu]_{\tau + \theta \lambda \gamma} \leq C_2 | W |_{\beta, \tau, \lambda} | \varphi_1^a - \varphi_2^a |_{\lambda(1-\theta)}^{\lambda} .
\]

This completes the proof. \( \square \)

Corollary 2.12. Let \( I \) be a nonempty closed, bounded and connected interval. Let \( t_0 \) be in \( I \). Assume condition \( (W) \) with \( \tau + \lambda \gamma > 1 \). Then the map
\[
M : C^\gamma(I) \to C^\tau(I)
\]
\[
M x(t) = \int_{t_0}^{t} W(ds, x_s)
\]
is continuous and compact.

Proof. Continuity follows immediately from Proposition 2.11. For compactness, suppose \( B \) is a bounded subset of \( C^\gamma(I) \). The estimate in Proposition 2.11 implies that \( \{M x\}_{x \in B} \) is bounded in \( C^\tau(I) \). By the Arzelà-Ascoli theorem, the set \( \{M x\}_{x \in B} \) is relatively compact in \( C^\tau(I) \) for every \( \tau' < \tau \). We show that \( \{M x\}_{x \in B} \) is indeed relatively compact in \( C^\tau(I) \). More precisely, suppose \( \{M x^n\} \) is a convergent sequence in \( M(B) \) in the norm of \( C^{\tau'}(I) \); by taking further subsequence, we can assume that the sequence \( \{x^n\} \) converges to \( x \) in \( C^{\gamma'}(I) \), for some \( \gamma' < \gamma \) (this is possible since \( B \) is bounded). It is sufficient to show that \( M x^n \) converges to \( M x \) in \( C^\gamma(I) \). To prove this, we choose \( \theta \in (0, 1) \) and \( \gamma' < \gamma \) such that \( \tau + \theta \lambda \gamma' > 1 \), and then we apply Proposition 2.11 to obtain
\[
\| M x - M x^n \|_{\tau} \leq c \| W \|_{\beta, \tau, \lambda}(\| x - x^n \|_{\infty}^{\lambda} + \| x - x^n \|_{\lambda(1-\theta)}^{\lambda(1-\theta)}) .
\]
3. Differential equations

Let $W : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ satisfy the condition \[ (W) \] stated at the beginning of Section 2 with $\tau(1 + \lambda) > 1$. In this section we consider the following differential equation:

\[ \phi_t = \phi_{t_0} + \int_{t_0}^{t} W(ds, \phi_s). \]

We are concerned with the existence, uniqueness, boundedness and the flow property of the solution. We shall also study the dependence of the solution on the initial conditions. Some related results in this direction were also obtained independently by Catellier and Gubinelli [5]. Applications of the results obtained are represented in Subsections 3.3 and 3.4, where we consider a transport equation of the type

\[ u(dt, x) = \nabla u(t, x)W(dt, x). \]

Literature on transport equations is vast and mostly focuses on irregularity of the spatial variables of the vector field (see for instance [13] for Sobolev vector fields, [2] for BV vector fields and [3] for Besov vector fields). In the case $W$ being a semimartingale, the above equation is treated in [37]. It appears to be new in the context of nonlinear Young integration.

3.1. Existence and uniqueness.

**Theorem 3.1** (Existence). Suppose that $W$ satisfies the assumption \[ (W) \] with $\tau(1 + \lambda) > 1$ and $\beta + \lambda \leq 1$. Then equation (3.1) has a solution in the space of\( H^\beta \) functions $C^\tau([t_0 - T, t_0 + T])$ for any $T > 0$. Moreover, if $\phi$ is a solution in $C^\tau([t_0 - T, t_0 + T])$,

\[ \sup_{t \in [t_0 - T, t_0 + T]} |\phi_t| + \sup_{t_0 - T \leq s < t \leq t_0 + T} |\phi_t - \phi_s| \leq C_{\gamma, \tau} T e^{k_{\gamma, \tau} T} \|W\|_{\gamma, \tau}^{\frac{1}{\gamma + \tau}} (1 \vee |\phi_{t_0}|), \]

where the constant $k_{\gamma, \tau} T$ and $C_{\gamma, \tau}$ depend only on $\gamma, \tau$ and $T$.

**Proof.** Fix $T > 0$; we denote $\|W\| = \|W\|_{\gamma, \tau,[t_0 - T, t_0 + T]}$. We define a mapping $M$ acting on $C^\tau([t_0 - T, t_0 + T])$ as

\[ Mx = x_0 + \int_{t_0}^{t} W(ds, x_s), \quad \forall x \in C^\tau([t_0 - T, t_0 + T]). \]

We shall verify that $M$ satisfies the hypothesis of the Leray-Schauder theorem (see [22], Theorem 11.3]).

**Step 1.** $M$ is well-defined, continuous and compact. This immediately follows from Corollary 2.12.

**Step 2.** Now we explain that the set \[ \{x \in C^\tau([t_0 - T, t_0 + T]) : x = \sigma Mx, \quad 0 \leq \sigma \leq 1 \} \] is bounded. Let $x$ satisfy $x = \sigma Mx$ for some $\sigma \in [0, 1]$. By definition of $M$, we see $x = \sigma Mx$ can be written as

\[ x_b - x_a = \sigma \int_{a}^{b} W(ds, x_s). \]
From (2.5), it follows that for any \( a, b \in [t_0 - T, t_0 + T] \), we have

\[
|x_b - x_a| = \sigma \left| \int_a^b W(ds, x_s) \right|
\leq \sigma \|W\| (1 + \|x\|_{\infty; a,b}^\beta) \|x\|_{\infty; a,b}^\lambda (b - a)^\tau
+ \sigma \kappa \|W\| (1 + \|x\|_{\infty; a,b}^\beta) \|x\|_{\tau, a,b}^\lambda |b - a|^{\tau + \lambda^\tau}.
\]

Since \( \sigma \leq 1 \), this yields

\[
\|x\|_{\tau; a,b} \leq \|W\| (1 + \|x\|_{\infty; a,b}^\beta) \|x\|_{\infty; a,b}^\lambda + \kappa \|W\| (1 + \|x\|_{\infty; a,b}^\beta) \|x\|_{\tau, a,b}^\lambda |b - a|^{\lambda^\tau},
\]

for every \( a, b \in [t_0, t_0 + T] \) with \( a < b \). We emphasize that the constant \( \kappa \) appearing in the previous inequality is independent of \( \sigma \). An application of the Young inequality gives

\[
\|x\|_{\infty; a,b}^\beta \|x\|_{\tau, a,b}^\lambda \leq \|x\|_{\infty; a,b}^{\beta + \lambda} + \|x\|_{\tau, a,b}^{\beta + \lambda},
\]

Thus

\[
\|x\|_{\tau; a,b} \leq \|W\| (\|x\|_{\infty; a,b}^\lambda + \|x\|_{\tau; a,b}^{\beta + \lambda}) + \kappa \|W\| \|x\|_{\tau, a,b}^{\beta + \lambda} |b - a|^{\lambda^\tau},
\]

Applying the inequality \( z^\theta \leq 1 \vee z (~\theta \in [0, 1] ~\text{and} ~z \geq 0) \), we obtain

\[
\|x\|_{\tau; a,b} \leq \|W\| (2 + \kappa |b - a|^{\lambda^\tau}) (1 \vee \|x\|_{\infty; a,b}^\lambda) + \kappa \|W\| (1 \vee \|x\|_{\tau; a,b}) |b - a|^{\lambda^\tau}.
\]

We further use

\[
\|x\|_{\infty; a,b} \leq |x_a| + \|x\|_{\tau; a,b} |b - a|^{\tau}
\]

to obtain

\[
\|x\|_{\tau; a,b} \leq A \|W\| (1 \vee |x_a|) + A \|W\| (1 \vee \|x\|_{\tau; a,b}) |b - a|^{\lambda^\tau},
\]

where \( A \) is a constant depending only on \( \tau, \lambda \) and \( T \). Let \( \Delta \) be a positive number such that

\[
A \|W\| \Delta^{\tau^\lambda} = \frac{1}{2}.
\]

If \( |b - a| \leq \Delta \), then from (3.3)

\[
\|x\|_{\tau; a,b} \leq 2A \|W\| (1 \vee |x_a|).
\]

Hence, we obtain

\[
(1 \vee \|x\|_{\infty; a,b}) \leq (2A \|W\| \Delta^{\tau^\lambda} + 1) (1 \vee |x_a|).
\]

Divide the interval \([t_0, t_0 + T]\) into \( n = \lfloor T/\Delta \rfloor + 1 \) subintervals of length less than or equal to \( \Delta \). Applying the inequality (3.6) on the intervals \([t_0, t_0 + \Delta], [t_0 + \Delta], \ldots, [t_0 + (n - 1)\Delta, t_0 + n\Delta \land T]\), recursively, we obtain

\[
(1 \vee \|x\|_{\infty; t_0, t_0 + T}) \leq (2A \|W\| \Delta^{\tau^\lambda} + 1)^n (1 \vee |x_{t_0}|).
\]

We can also assume that \( \Delta \leq T \). Thus \( n \leq 2T/\Delta \). We use the bound \( 2A \|W\| \Delta^{\tau^\lambda} + 1 \leq \exp(2A \|W\| \Delta^{\tau^\lambda}) \). Then (3.7) yields

\[
(1 \vee \|x\|_{\infty; t_0, t_0 + T}) \leq \exp(2A \|W\| \Delta^{\tau^\lambda} \frac{2T}{\Delta}) (1 \vee |x_{t_0}|).
\]

Using (3.4), namely,

\[
\Delta = (2A \|W\|)^{-\frac{1}{\lambda^\lambda}},
\]
we have

\[ (1 \lor \|x\|_{\infty;[t_0,t_0+T]}) \leq e^{T(2A\|W\|)\frac{1-\tau+\lambda}{\lambda}} (1 \lor |x_{t_0}|), \]

where \( C_{\tau,\lambda} \) and \( \kappa_{\tau,\lambda} \) are uniformly bounded in \( \sigma \in [0,1] \). The argument applies similarly to the other interval \([t_0-T,t_0]\). Thus

\[ (3.8) \quad (1 \lor \|x\|_{\infty;[t_0-T,t_0+T]}) \leq e^{T(2A\|W\|)\frac{1-\tau+\lambda}{\lambda}} (1 \lor |x_{t_0}|). \]

Together with the estimate \((3.5)\), this inequality \((3.8)\) implies that the set

\[ \{x \in C^r([t_0-T,t_0+T]) : x = \sigma Lx, 0 \leq \sigma \leq 1 \} \]

is bounded in \( C^r([t_0-T,t_0+T]) \).

**Step 3.** Applying the Leray-Schauder theorem, we see that the equation \((3.1)\) has a solution \( \{\varphi_t, t \in [t_0-T,t_0+T]\} \) in \( C^r([t_0-T,t_0+T]) \) for every \( T \). The estimate \((3.2)\) comes from \((3.8)\) together with \((3.5)\).

Next, we study some stability result. In particular, we want to know how the solution depends on the initial condition \( x_{t_0} \).

**Theorem 3.2.** Let the condition \((W)\) be satisfied with \( \tau + \tau \lambda > 1 \). In addition, we assume that \( W(t,x) \) is differentiable with respect to \( x \) for every \( t \) and the spatial gradient matrix of \( W \) is denoted by \( \nabla W(t,x) = \left( \frac{\partial W_{i,j}(t,x)}{\partial x_j} \right)_{1 \leq i,j \leq d} \). Suppose

\[
\|\nabla W\|_{\tau,\lambda;[t_0-T,t_0+T] \times K} := \sup_{t_0-T \leq s \leq t \leq t_0+T} \sup_{x \in K} \frac{|\nabla W(t,x) - \nabla W(s,x)|}{|t-s|^\tau} + \sup_{t_0-T \leq s \leq t \leq t_0+T} \sup_{x,y \in K, x \neq y} \frac{|\nabla W(t,x) - \nabla W(s,x) - \nabla W(t,y) + \nabla W(s,y)|}{|t-s|^\tau |x-y|^\lambda}
\]

is finite for all compact sets \( K \) in \( \mathbb{R}^d \). Let \( x_t \) and \( y_t \) be two solutions in \( C^r([t_0-T,t_0+T]) \) to the integral equation \((3.1)\) with initial conditions \( x_0 \) and \( y_0 \) respectively. Then the following estimate holds:

\[ (3.9) \quad \sup_{t \in [t_0-T,t_0+T]} |x_t - y_t| \leq 2e^{TA\frac{1}{\lambda}} |x_0 - y_0|, \]

where \( A \) is a constant depending on \( \nabla W, x, y \) and \( T \) (the precise formula is given in \((3.10)\) below).

**Proof.** We denote \( R = \max\{\|x\|_{\infty;[t_0-T,t_0+T]}, \|y\|_{\infty;[t_0-T,t_0+T]}\} \), \( K = \{ x \in \mathbb{R}^d : |x| \leq R \} \) and \( \|\nabla W\| = \|\nabla W\|_{\tau,\lambda;[t_0-T,t_0+T] \times K} \). We also denote \( z_t = x_t - y_t \), \( \rho_\tau = (|x| + |y|)^\lambda \) and \( \eta_t = \eta x_t + (1-\eta)y_t \) for each \( \eta \in (0,1) \). For every \( s, t \) and \( x \), we use the notation \( W([s,t],x) = W(t,x) - W(s,x) \).

We shall obtain an estimate for \( z \) in \( C([t_0-T,t_0+T]) \). Fix \( a < b \) in \([t_0-T,t_0+T]\).

We then write

\[ z_b - z_a = \int_a^b W(ds, x_s) - \int_a^b W(ds, y_s) = \int_a^b W(ds, y_s) - \int_a^b W(ds, x_s) = \int_0^1 \nabla W([s,t], \eta s) z_s d\eta, \]

where \( \mu \) is the function

\[ \mu(s,t) = W([s,t], x_s) - W([s,t], y_s) = \int_0^1 \nabla W([s,t], \eta s) z_s d\eta. \]
For every \( s \leq c \leq t \) in \([a, b]\), we can write
\[
\mu(s, t) - \mu(s, c) - \mu(c, t) = \int_0^1 ([\nabla W([c, t], \eta_s) - \nabla W([c, t], \eta_c)] z_s + \nabla W([c, t], \eta_c) (z_t - z_c)) \, d\eta.
\]
We note that \(|\eta_t - \eta_s|^\lambda = |\eta(x_t - x_s) + (1 - \eta)(y_t - y_s)|^\lambda \leq \rho^\lambda|u - v|^\lambda\). It follows that
\[
[\mu]_{\tau(1+\lambda);[a,b]} \leq \|\nabla W\| (\rho^\lambda \|z\|_{\infty;a,b} + |b - a|^{\tau(1-\lambda)}\|z\|_{\tau,a,b}).
\]
On the other hand, it is obvious that \(|\mu(a, b)| \leq \|\nabla W\| |b - a|^\tau \|z\|_{\infty;a,b}\). Hence, the estimate (2.4) implies
\[
|z_b - z_a| \leq \|\nabla W\| |b - a|^\tau \|z\|_{\infty;a,b} + \kappa \|\nabla W\| |b - a|^\tau + \lambda \tau (\rho^\lambda \|z\|_{\infty;a,b} + |b - a|^{\tau(1-\lambda)}\|z\|_{\tau,a,b}).
\]
In other words,
\[
\|z\|_{\tau;a,b} \leq A \|z\|_{\infty;a,b} + A \|z\|_{\tau,a,b} (b - a)^\tau,
\]
where
\[
(3.10) \quad A = \kappa \|\nabla W\| [1 + \rho^\lambda T^\lambda].
\]
Therefore, using the bound \(\|z\|_{\infty;a,b} \leq |z_a| + \|z\|_{\tau;a,b}\) one gets
\[
(3.11) \quad \|z\|_{\tau;a,b} \leq A|z_a| + 2A \|z\|_{\tau,a,b} (b - a)^\tau.
\]
Now we shall use the above inequality to show our theorem. Choose \(a, b\) such that
\[
|b - a| \leq \Delta = \left(\frac{1}{4A}\right)^{\frac{1}{\tau}}.
\]
Then inequality (3.11) implies \(\|z\|_{\tau;a,b} \leq 2A|z_a|\) for all \(a < b\). By the definition of the Hölder norm, we see that if \(|b - a| \leq \Delta\), then
\[
\|z\|_{\infty,a,b} \leq |z_a| + \|z\|_{\tau,a,b} (b - a)^\tau
\leq |z_a| + 2A|z_a| \Delta^\tau
\leq 2|z_a|.
\]
Divide the interval \([t_0, t_0 + T]\) into \(n = [T/\Delta] + 1\) subintervals of length less than or equal to \(\Delta\). Applying the previous inequality on the intervals \([t_0, t_0 + \Delta]\), \([t_0 + \Delta, t_0 + 2\Delta]\), \ldots, \([t_0 + (n - 1)\Delta, t_0 + n\Delta]\), recursively, we obtain
\[
\|z\|_{\infty,t_0,t_0+T} \leq 2^n |z_{t_0}|.
\]
We can assume \(\Delta \leq T\). Thus
\[
n = [T/\Delta] + 1 \leq \frac{2T}{\Delta} = 2T (4A)^{\frac{1}{\tau}}.
\]
This implies
\[
\|z\|_{\infty,t_0,t_0+T} \leq 2^{1+2/\tau} TA^{\frac{1}{\tau}} |z_{t_0}|,
\]
which yields the bound (3.3) on the interval \([t_0, t_0 + T]\). Estimates on \([t_0 - T, t_0]\) are analogous.

An immediate consequence of the theorem is the following uniqueness result.

**Corollary 3.3.** Under the hypothesis of Theorem 3.2, equation (3.1) has a unique solution.
3.2. Compositions. Given a function \( G : \mathbb{R}^2 \to \mathbb{R}^d \), we may define the Riemann-Stieltjes integral \( \int_a^b G(ds, s) \) as the limit of Riemann sums

\[
\sum_i G(t_i, t_{i-1}) - G(t_{i-1}, t_i).
\]

The sewing lemma (Lemma 2.2) gives a sufficient condition so that the aforementioned limit exists; namely \( G \) satisfies

\[
|G(s, s) - G(s, t) - G(t, s) + G(t, t)| \lesssim |t - s|^{1+\varepsilon}
\]

for some \( \varepsilon > 0 \). In such case, Lemma 2.3 also allows one to choose Riemann sums with right-end points. In other words, the Riemann sums with right-end points

\[
\sum_i G(t_i, t_i) - G(t_{i-1}, t_i)
\]

also converge to the Riemann-Stieltjes integral \( \int_a^b G(ds, s) \). In what follows, all integrals are understood as Riemann-Stieltjes integration, except for a few occasions, which we will indicate. The following result can be regarded as the Itô formula or chain rule for compositions of functions in the context of nonlinear Young integrals.

**Theorem 3.4.** Let \( F \) be a function in \( C^{(\tau_F, \lambda_F)}_{\text{loc}}(\mathbb{R} \times \mathbb{R}^d) \) (i.e. \( F \) satisfies the condition \((W')\) with \( \tau_F \) and \( \lambda_F \)), and let \( g \) and \( x \) be Hölder continuous functions with exponents \( \tau_g \) and \( \tau \) respectively. We suppose that \( \tau_F + \lambda_F \tau > 1 \) and \( \tau_g + \tau_F > 1 \). The following integration by parts formula holds:

\[
\int_0^T g(t)dF(t, x_t) = \int_0^T g(t)F(dt, x_t) + \int_0^T g(t)\, F(t, dx_t). \tag{3.12}
\]

In particular, suppose that \( F \) belongs to \( C^{(\tau_F, \lambda_F)}_{\text{loc}}(\mathbb{R}; C^{1, \lambda_F}_{\text{loc}}(\mathbb{R}^d)) \), \( x \) is of the form \( x_t = \int_0^t W(ds, \phi_s) \), where \( W \) satisfies the condition \((W')\) with \( \tau \) and \( \lambda, \phi \) satisfies \((\phi)\) with \( \gamma, \tau + \lambda \gamma > 1 \) and \( \tau \lambda_F + \tau > 1 \). Then (3.12) becomes

\[
\int_0^T g(t)dF(t, x_t) = \int_0^T g(t)F(dt, x_t) + \int_0^T g(t)(\nabla F)(t, x_t)W(dt, \phi_t). \tag{3.13}
\]

An important consequence of (3.13) is when \( g \) is a constant function

\[
F(b, x_b) - F(a, x_a) = \int_a^b F(dt, x_t) + \int_a^b (\nabla F)(t, x_t)W(dt, \phi_t). \tag{3.14}
\]

**Proof.** We choose a compact set \( K \) such that \( K \) contains \( \{x_t, 0 \leq t \leq T\} \) and denote \( \|F\| = \|F\|_{\tau_F, \lambda_F; [0, T] \times K} \). We put

\[
\mu(a, b) = g(b)F(b, x_b) - g(b)F(a, x_a),
\nu(a, b) = g(a)F(a, x_b) - g(a)F(a, x_a),
\vartheta(a, b) = g(a)F(b, x_b) - g(a)F(a, x_a).
\]

...
For every $a < c < b$, we have
\[
|\mu(a, b) - \mu(a, c) - \mu(c, a)| = | - g(b) \mathcal{F}(a, x_b) - g(c) \mathcal{F}(c, x_c) + g(c) \mathcal{F}(a, x_c) + g(b) \mathcal{F}(c, x_b)|
\leq |g(c)||\mathcal{F}(a, x_b) - \mathcal{F}(c, x_c) + \mathcal{F}(a, x_c) + \mathcal{F}(c, x_b)| + |g(c) - g(b)||\mathcal{F}(c, x_c) - \mathcal{F}(a, x_c)|
\leq \|g\|_{\infty}\|||x||^{\lambda_F}|b - a|^{\tau_F + \lambda_F\tau} + \|g\|_{\tau_g}\|\mathcal{F}||b - a|^{\tau_g + \tau_F}
\]
and
\[
|\nu(a, b) - \nu(a, c) - \nu(c, a)| = |g(a)\mathcal{F}(a, x_b) - g(a)\mathcal{F}(a, x_a) - g(c)\mathcal{F}(c, x_b) + g(c)\mathcal{F}(c, x_c)|
\leq |g(c)||\mathcal{F}(a, x_b) - \mathcal{F}(a, x_a) - \mathcal{F}(b, x_b) + \mathcal{F}(c, x_c)| + |g(c) - g(a)||\mathcal{F}(a, x_b) - \mathcal{F}(a, x_a)|
\leq \|g\|_{\infty}\|\mathcal{F}||b - a|^{\tau_F + \lambda_F\tau} + \|g\|_{\tau_g}\|\mathcal{F}||x||^{\lambda_F}|b - a|^{\tau_g + \lambda_F\tau}.
\]

Hence, from Lemmas 2.2 and 2.3, $\mathcal{J}_T\mu = \int_0^T g(t)\mathcal{F}(dt, x_t)$ and $\mathcal{J}_T\nu = g(t)\mathcal{F}(t, dx_t)$. On the other hand,
\[
|\vartheta(a, b) - \mu(a, b) - \nu(a, b)| = |\|g(a) - g(b)||\mathcal{F}(b, x_b) - \mathcal{F}(a, x_b)|| \leq \|g\|_{\tau_g}\|\mathcal{F}||b - a|^{\tau_g + \tau_F}.
\]
This together with Lemma 2.3 implies (3.12).

To prove the $\mathcal{J}_T\nu = g(t)\mathcal{F}(t, dx_t)$, it suffices to show
\[
(3.15) \quad \int_0^T g(t)\mathcal{F}(t, dx_t) = \int_0^T g(t)(\nabla \mathcal{F})(t, x_t)W(dt, \phi_t).
\]
We denote
\[
\tilde{\nu}(a, b) = g(a)\nabla \mathcal{F}(a, x_a)[W(b, \phi_a) - W(a, \phi_a)].
\]
Then we write
\[
\nu(a, b) = g(a)\int_0^1 \nabla \mathcal{F}(a, \eta x_a + (1 - \eta)x_b)d\eta(x_a - x_b)
= g(a)\int_0^1 \nabla \mathcal{F}(a, \eta x_a + (1 - \eta)x_b)d\eta \int_a^b W(ds, \phi_s).
\]
Using the estimate (2.5), we obtain
\[
|\nu(a, b) - \tilde{\nu}(a, b)| \leq |g(a)\int_0^1 \nabla \mathcal{F}(a, \eta x_a + (1 - \eta)x_b) - \nabla \mathcal{F}(a, x_a)d\eta \int_a^b W(ds, \phi_s)|
+ |g(a)\nabla \mathcal{F}(a, x_a)\int_a^b W(ds, \phi_s) - W(b, \phi_b) + W(a, \phi_a)||
\leq |b - a|^{\lambda_F\tau + \tau} + |b - a|^{\tau + \lambda_T}.
\]
Identity (3.15) follows from Lemma 2.3 and the previous estimate. \qed
3.3. Regularity of flow. In the rest of the current section, we assume the hypothesis of Theorem 3.2. This assumption guarantees that \( \varphi(t, x) \), the solution to

\[
\varphi(t, x) = x + \int_0^t W(ds, \varphi(s, x)),
\]

is unique. Moreover, by the result in Subsection 3.1, for fixed \( t \), \( \varphi(t, \cdot) \) is an automorphism on \( \mathbb{R}^d \), and its inverse is \( \varphi(t, \cdot)^{-1} = \varphi(-t, \cdot) \). Hence, the family \( \{ \varphi(t, \cdot) : t \in \mathbb{R} \} \) forms a flow of homeomorphisms, i.e. it satisfies the following properties:

- \( \varphi(t+s, \cdot) = \varphi(t, \varphi(s, \cdot)) \) holds for all \( s, t \),
- \( \varphi(0, \cdot) \) is the identity map,
- the map \( \varphi(t, \cdot) : \mathbb{R}^d \to \mathbb{R}^d \) is a homeomorphism for all \( t \).

Moreover, one can show that \( \varphi(t, \cdot) \) is indeed a diffeomorphism.

**Theorem 3.5.** Assume the hypothesis of Theorem 3.2. For any \( t \) in \( \mathbb{R} \), the map \( \varphi(t, \cdot) \) is a diffeomorphism. The following conclusions hold.

(i) The gradient of \( \varphi_t \) at \( x \), denoted by \( \nabla \varphi(t, x) = \{ \partial_j \varphi^i(t, x) \}_{i,j} \), satisfies the equation

\[
\partial_i \varphi^i(t, x) = \delta_{ij} + \int_0^t \partial_k W^i(ds, \varphi(s, x)) \partial_k \varphi^j(s, x)
\]

where \( \delta_{ij} \) is the Kronecker symbol. Equation (3.16) can be written in short

\[
\nabla \varphi(t, x) = I_d + \int_0^t \nabla W(ds, \varphi(s, x)) \nabla \varphi(s, x).
\]

(ii) For every \( t \) and \( x \), the matrix \( \nabla \varphi(t, x) \) is invertible and its inverse \( M(t, x) = [\nabla \varphi(t, x)]^{-1} \) satisfies the equation

\[
M(t, x)^j \cdot = \delta_j \cdot - \int_0^t M(s, x)^{ik} \partial_k W^k(ds, \varphi(s, x))
\]

or in short

\[
M(t, x) = I_d - \int_0^t M(s, x) \nabla W(ds, \varphi(s, x)).
\]

(iii) \( \varphi \) is jointly Hölder continuous of order \((\tau, 1)\). That is,

\[
|\varphi(s, x) - \varphi(s, y) - \varphi(t, x) + \varphi(t, y)| \lesssim |t-s|^{\tau} |x-y|.
\]

(iv) Let \( J(t, x) \) denote the determinant of \( \nabla \varphi(t, x) \). Then \( J \) satisfies the following scalar linear equation:

\[
J(t, x) = 1 + \int_0^t J(s, x) \text{Div}(W(ds, \varphi(s, x))).
\]

(v) The flow \( \varphi(t, \cdot) \) is a Lagrangian flow; namely there exists a constant \( L \) such that

\[
\mathcal{L}^d(\varphi(t, \cdot)^{-1}(A)) \leq L \mathcal{L}^d(A) \quad \text{for every Borel set } A \subseteq \mathbb{R}^d
\]

where \( \mathcal{L}^d \) is the Lebesgue measure on \( \mathbb{R}^d \).
Proof. Let $e$ be a unit vector in $\mathbb{R}^d$. For each $h$ in $\mathbb{R}$, we denote

$$\eta^h_t = \frac{1}{h}(\varphi(t, x + he) - \varphi(t, x)).$$

To prove (i), it is sufficient to show that for every sequence $h_n$ converging to 0, there is a subsequence $h_{n_k}$ such that $\eta^{h_{n_k}}$ converges to the solution of the following equation:

\[(3.21) \quad \eta_t = e + \int_0^t \nabla W(ds, \varphi(s, x)) \eta_s.\]

We remark that the equation (3.21) is linear and the existence and uniqueness of the solution in $C^\tau(\mathbb{R})$ follows from our method discussed in Subsection 3.1. From Theorem 3.2 we see that

$$\|\eta^h\|_{\tau; K} \leq \kappa_K$$

uniformly in $h$ for every compact interval $K$ in $\mathbb{R}$. Hence, by the Arzelà-Ascoli theorem, there is a subsequence, still denoted by $h_n$ such that $\eta^{h_n}$ converges to $\eta$ in $C^\tau'(K)$ for any arbitrary $\tau' < \tau$. On the other hand, we notice that $\eta^h$ satisfies

\[(3.22) \quad \eta^h_t = e + \int_0^1 d\tau \int_0^t \nabla W(ds, \tau \varphi(s, x + he) - (1 - \tau) \varphi(s, x)) \eta^h_s.\]

Passing through the limit $h_n \to 0$, we see that $\eta$ satisfies the equation (3.21) and then (i) follows. Assertion (iii) is a consequence of the estimate (3.9) in Theorem 3.2. In fact,\n
$$|\varphi(s, x) - \varphi(s, y) - \varphi(t, x) + \varphi(t, y)| \leq \|\varphi(\cdot, x) - \varphi(\cdot, y)\|_{\tau; [s, t]} |t - s|^\tau \lesssim |t - s|^\tau |x - y|.$$\n
Assertion (iv) follows from the Itô formula (3.14) applied to $J(t, x) = \det(\nabla \varphi(t, x))$ and Jacobi’s formula

$$d \det(M) = \det(M) tr(M^{-1}dM).$$

To prove (v), we notice that the equation (3.19) can be solved explicitly thanks to (3.14):

\[(3.23) \quad J(t, x) = \exp \int_0^t \text{Div} (W(dt, \varphi(t, x))).\]

Therefore, from (2.5), we obtain

$$|J(t, x)|^{-1} \leq e^{\kappa |t|^\tau}.$$\n
Together with the area formula

$$\mathcal{L}^d(\varphi(t, \cdot)^{-1}(A)) = \int_{\varphi(-t, A)} dx = \int_A |\det(\nabla \varphi)(-t, x)| dx$$

this estimate implies (3.20). \qed
3.4. **Transport differential equation.** As an application of the above Itô formula (3.13) and flow property (Theorem 3.5), we study the following transport differential equation in Hölder media. Specifically, let $W : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ satisfy the conditions in Theorem 3.2. Consider the following first order partial differential equations (transport equation in Hölder media $W$):

\begin{equation}
\frac{\partial}{\partial t} u(t, x) + \left( \frac{\partial}{\partial t} W(t, x) \right) \cdot \nabla u(t, x) = 0.
\end{equation}

Here $\nabla$ is the gradient operator (with respect to spatial variables). Since $W$ is only Hölder continuous in time, the equation (3.24) is only formal. We can however define solutions in integral form. More precisely, a continuous function $u : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ is called a solution to (3.24) with the initial condition $u(0, x) = h(x)$ if it is differentiable with respect to $x \in \mathbb{R}^d$ and the following equation holds:

\begin{equation}
 u(t, x) = h(x) - \int_0^t \nabla u(s, x) W(ds, x) \quad \forall \ t \geq 0, \ x \in \mathbb{R}^d.
\end{equation}

**Theorem 3.6.** Assume $W$ satisfies the conditions in Theorem 3.2. Let $h$ be a function in $C^{1+\lambda_0}_{loc}(\mathbb{R}^d)$ where $\lambda_0$ satisfies $(1+\lambda_0)\tau > 1$. Let $\varphi(t, x)$ be the unique solution to

\begin{equation}
 \varphi(t, x) = x + \int_0^t W(ds, \varphi(s, x)) , \ \forall t \geq 0. \nonumber
\end{equation}

Let $\psi(t, x)$ be the inverse of $\varphi$ as a function $x \in \mathbb{R}^d$ to $\mathbb{R}^d$. Namely, $\varphi(t, \psi(t, x)) = x$ for all $t \geq 0, \ x \in \mathbb{R}^d$. Then the function $u$ defined by

\begin{equation}
 u(t, x) = h(\psi(t, x))
\end{equation}

is a solution to the above transport equation.

**Proof.** From Theorem 3.5 such $\psi(t, x)$ exists and both $\varphi(t, x)$ and $\psi(t, x)$ are differentiable with respect to $x$. We differentiate $\varphi(t, \psi(t, x)) = x$ with respect to $x$ and we see that

\begin{equation}
 (\nabla \varphi)(t, \psi(t, x)) \nabla \psi(t, x) = I
\end{equation}

or

\begin{equation}
 (\nabla \psi(t, x))^{-1} = (\nabla \varphi)(t, \psi(t, x)).
\end{equation}

Let $\rho(r) = \varphi(r, \psi(r, x)), 0 \leq r < \infty$. Thanks to Theorem 3.3(iii), the Itô formula (3.13) is applicable. More precisely, for any $C^7$-function $g(r)$, we have

\begin{equation}
 \int_0^t g(r) d\rho(r) = \int_0^t g(r) \varphi(dr, \psi(r, x)) + \int_0^t g(r)(\nabla \varphi)(r, \psi(r, x)) \psi(dr, x). \nonumber
\end{equation}

Since $\rho(r) = x$, we have $d\rho(r) = 0$. Thus

\begin{equation}
 \int_0^t g(r)(\nabla \varphi)(r, \psi(r, x)) \psi(dr, x) = - \int_0^t g(r) \varphi(dr, \psi(r, x)).
\end{equation}
Now the Itô formula \([3.14]\) applied to \(h(\psi(t,x))\) yields

\[
u(t, x) = h(\psi(t,x)) = h(x) + \int_0^t (\nabla h)(\psi(r,x))\psi(dr,x)
\]

\[
= h(x) + \int_0^t \nabla [h(\psi(r,x))] (\nabla \psi(r,x))^{-1} \psi(dr,x)
\]

\[
= h(x) + \int_0^t \nabla u(r,x) (\nabla \psi(r,x))^{-1} \psi(dr,x)
\]

\[
= h(x) + \int_0^t \nabla u(r,x) (\nabla \varphi)(r,\psi(r,x))\psi(dr,x).
\]

Using the equation \([3.26]\) for \(g(r) = \nabla u(r,x)\), we have

\[
u(t, x) = h(x) - \int_0^t \nabla u(r,x) \varphi(dr,\psi(r,x))
\]

\[
= h(x) - \int_0^t \nabla u(r,x)W(dr,\varphi(r,\psi(r,x)))
\]

\[
= h(x) - \int_0^t \nabla u(r,x)W(dr,x).
\]

This completes the proof of the theorem. \(\square\)

We also have the following uniqueness result.

**Theorem 3.7.** Assume \(W\) satisfies the conditions in Theorem \([3.2]\) Let \(\lambda_0\) be in \((0,1]\) such that \((\lambda_0 + 1)\tau > 1\). Equation \([3.25]\) has a unique solution in the class \(C^{(\tau,\lambda_0)}_{\text{loc}}(\mathbb{R} \times \mathbb{R}^d)\). More precisely, suppose \(u\) belongs to \(C^{(\tau,\lambda_0)}_{\text{loc}}(\mathbb{R} \times \mathbb{R}^d)\) and satisfies \([3.25]\). Then \(u\) is uniquely defined by the relation \(u(t,x) = h(\psi(t,x))\), where \(\varphi\) and \(\psi\) are the functions defined in Theorem \([3.6]\).

**Proof.** Let \(u\) be a solution to \([3.25]\). Applying the Itô formula \([3.14]\) for the function \(u(t,\varphi(t,x))\) we have

\[
u(t, \varphi(t,x)) - h(x) = \int_0^t u(ds, \varphi(s,x)) + \int_0^t \nabla u(s, \varphi(s,x))W(ds, \varphi(s,x)).
\]

It suffices to show the right hand side vanishes. In other words the following relation between the two nonlinear Young integrals holds:

\[
\int_0^t u(ds, \varphi(s,x)) = -\int_0^t \nabla u(s, \varphi(s,x))W(ds, \varphi(s,x)).
\]

For clarity, we will omit \(x\) in the notation. We write

\[
\mu_1(a,b) = u(b, \varphi_a) - u(a, \varphi_a),
\]

\[
\mu_2(a,b) = \nabla u(a, \varphi_a)[W(b, \varphi_a) - W(a, \varphi_a)].
\]

Since \(u\) satisfies the equation \([3.25]\), we can write

\[
\mu_1(a,b) = -\int_a^b \nabla u(s, \varphi_a)W(ds, \varphi_a).
\]
Thus
\[ \mu_1(a, b) + \mu_2(a, b) = -\int_a^b \nabla u(s, \varphi_a) W(ds, \varphi_a) + \nabla u(a, \varphi_a)[W(b, \varphi_a) - W(a, \varphi_a)]. \]
The estimate (2.4) (or (2.5)) implies
\[ |\mu_1(a, b) + \mu_2(a, b)| \lesssim |b - a|^{2\tau}. \]
Since \(2\tau > 1\), Lemma 2.3 yields \( J(t) = -J(t) \). This completes the proof after observing that the aforementioned identity is exactly the same as (3.27).

Remark 3.8. In the context of ordinary differential equation of the type
\[ \frac{dX}{dt}(t, x) = b(t, X(t, x)), \]
with nonregular vector field \( b \), existence and uniqueness and stability of regular Lagrangian flows were proved by R.J. DiPerna and P.-L. Lions (13) for Sobolev vector fields with bounded divergence. This result has been extended by L. Ambrosio (2) to BV coefficients with bounded divergence. In [9], it is shown that under slightly relaxed assumptions many of the ODE results of DiPerna-Lions theory can be recovered, from a priori estimates, similar to (3.20). The current paper proposes another extension of this theory, where the vector field is distribution (rough) in time (derivative of a Hölder continuous function) and smooth in space. It is also interesting to extend the results presented here for vector fields which are rougher in time (see e.g. [34] for the linear case) or which are both rough in time and in space.

4. Feynman-Kac formula - A pathwise approach

In this section we shall study the stochastic parabolic equation with Hölder continuous noise in a Hölder random media (see equation (4.1) below). A feature of this problem is that for the noise we don’t assume any Hölder continuity in time variable. To make up for lack of regularity in time, we assume some regularity on spatial variables. In this case, the method presented in this section works for each sample path of the noise.

Throughout the current section, \( T \) is a fixed positive time. To describe the noise, we introduce the following space. Let \( \beta \) be a fixed nonnegative number. We say that \( f \) belongs to \( C(0, T, \mathbb{R}^d) \) and satisfies the following condition:

\[ \left[ f \right]_{\alpha + \beta + 1, \infty} := \sup_{t \in [0, T], x \in \mathbb{R}^d} \frac{|f(t, x)|}{1 + |x|^{\alpha + \beta + 1}} < \infty. \]

We notice that the condition (4.1) implies the growth conditions on \( \nabla f \) and \( f \). More precisely, one has

\[ \left[ \nabla f \right]_{\alpha + \beta, \infty} := \sup_{t \in [0, T], x \in \mathbb{R}^d} \frac{|\nabla f(t, x)|}{1 + |x|^{\alpha + \beta}} < \infty \]

and

\[ \left[ f \right]_{\alpha + \beta + 1, \infty} := \sup_{t \in [0, T], x \in \mathbb{R}^d} \frac{|f(t, x)|}{1 + |x|^{\alpha + \beta + 1}} < \infty. \]
It is easy to see that \( \|f\|_{C^{0,1+\alpha}_\beta} := \|f\|_{\alpha+\beta+1,\infty} + \|\nabla f\|_{\alpha,\infty} + \|\nabla f\|_{\beta,\alpha} \) forms a norm on \( C^{0,1+\alpha}_\beta([0,T] \times \mathbb{R}^d) \). In the rest of this section, we denote

\[
C^{0,1+\alpha}_\beta = \bigcap_{0<\alpha'<\alpha} C^{0,1+\alpha'}_\beta ([0,T] \times \mathbb{R}^d).
\]

Similar to the classical Hölder spaces, the space of smooth functions is not dense in \( C^{0,1+\alpha}_\beta([0,T] \times \mathbb{R}^d) \). However, we can still approximate a function in \( C^{0,1+\alpha}_\beta([0,T] \times \mathbb{R}^d) \) by smooth functions with a little trade off in spatial regularity. More precisely, let \( \eta \) be a function in \( C_\infty^\infty(\mathbb{R}^{d+1}) \) supported in \((-1,1)^{d+1}\) and \( \int \eta(t,x) dtdx = 1 \). For \( \epsilon > 0 \), we put \( \eta_\epsilon(t,x) = \epsilon^{-d-1} \eta(\epsilon^{-1}(t,x)) \). Let \( f \) be in \( C^{0,1+\alpha}_\beta([0,T] \times \mathbb{R}^d) \). We define \( f_\epsilon(t,x) = (f * \eta_\epsilon)(t,x) \). It is clear that \( f_\epsilon \) belongs to \( C_\infty^\infty(\mathbb{R}^{d+1}) \). In addition, we have the following result.

**Lemma 4.1.** For every \( \alpha' < \alpha \), \( \|\nabla f_\epsilon - \nabla f\|_{\beta,\infty} \) and \( \|\nabla f_\epsilon - \nabla f\|_{\beta,\alpha'} \) converge to 0 as \( \epsilon \) goes to 0.

**Proof.** We have

\[
|\nabla f_\epsilon(t,x) - \nabla f(t,x)| \leq \iint |\nabla f(t,z) - \nabla f(t,x)| \eta_\epsilon(t,x-z) dtdz \\
\leq \|\nabla f\|_{\beta,\alpha} \iint |x-z|^{\alpha}(1+|x|^\beta + |z|^\beta) \eta(t,x-z) dtdz \\
\leq \|\nabla f\|_{\beta,\alpha} \epsilon^\alpha(1 + |x|^\beta),
\]

which implies \( |\nabla f_\epsilon - \nabla f\|_{\beta,\infty} \to 0 \). This also implies

\[
|\nabla f_\epsilon(t,x) - \nabla f_\epsilon(t,y) - \nabla f(t,x) + \nabla f(t,y)| \lesssim \|\nabla f\|_{\beta,\alpha} \epsilon^\alpha(1 + |x|^\beta + |y|^\beta).
\]

On the other hand,

\[
|\nabla f_\epsilon(t,x) - \nabla f_\epsilon(t,y)| \leq \iint |\nabla f(t,x-z) - \nabla f(t,y-z)| \eta_\epsilon(t,z) dtdz \\
\leq \|\nabla f\|_{\beta,\alpha} |x-y|^{\alpha} \iint (1 + |x-z|^{\beta} + |y-z|^{\beta}) \eta_\epsilon(t,z) dtdz \\
\lesssim \|\nabla f\|_{\beta,\alpha} |x-y|^{\alpha}(1 + |x|^\beta + |y|^\beta),
\]

thus

\[
|\nabla f_\epsilon(t,x) - \nabla f_\epsilon(t,y) - \nabla f(t,x) + \nabla f(t,y)| \lesssim \|\nabla f\|_{\beta,\alpha} |x-y|^{\alpha}(1 + |x|^\beta + |y|^\beta).
\]

Interpolating these two bounds, we get

\[
|\nabla f_\epsilon(t,x) - \nabla f_\epsilon(t,y) - \nabla f(t,x) + \nabla f(t,y)| \lesssim \|\nabla f\|_{\beta,\alpha} \epsilon^{\alpha-\alpha'} |x-y|^{\alpha'}(1 + |x|^\beta + |y|^\beta)
\]

for every \( \alpha' < \alpha \). This implies \( |\nabla f_\epsilon - \nabla f\|_{\beta,\alpha'} \to 0 \).

In Section 5, we shall give conditions on the covariance of a Gaussian field \( W(t,x) \) such that it is in \( C^{0,1+\alpha}_\beta([0,T] \times \mathbb{R}^d) \).

Assume that \( W \) belongs to the space \( C^{0,1+\alpha}_\beta([0,T] \times \mathbb{R}^d) \); throughout this section, we denote \( W_n = W * \eta_{1/n} \). We consider the following parabolic equation with multiplicative noise:

\[
\partial_t u + Lu + u \partial_t W = 0, \quad u(T,x) = u_T(x),
\]
where the terminal function $u_T$ is assumed to be measurable with polynomial growth and $L$ is a second order differential operator of the form

$$L = \frac{1}{2} \sum_{i,j=1}^{d} a^{ij}(t,x) \partial_{x_i} \partial_{x_j} + \sum_{i=1}^{d} b^i(t,x) \partial_{x_i} .$$

Here the novelty is that we allow the coefficients $a^{ij}(t,x) = a^{ij}(t,x,W)$ and $b^i(t,x) = b^i(t,x,W)$ to depend on $W$. Since we are going to solve the equation and to establish a Feynman-Kac type formula for every sample path of $W$, we omit the explicit dependence of $a^{ij}$ and $b^i$ on $W$. Notice that with a time reversal $t \to T - t$, we can solve the stochastic parabolic equation with initial condition:

$$\partial_t u = Lu - u \partial_t W, \quad u(0, x) = u_0(x) .$$

The stochastic differential equations with random coefficients have been studied in a large number of papers. For example, it has been used in the modeling of the pressure in an oil reservoir with a log normal random permeability in [27] (see in particular the references therein). Recently, there has been a great amount of research work on uncertainty quantization from the numerical computation community. Many different types of stochastic partial differential equations with random coefficients have been studied. Let us only mention the books [23], [50], and the references therein. Since the classical Feynman-Kac formula has already experienced many applications, including the so-called Monte-Carlo particle approximation (see [11][12]), we expect that the Feynman-Kac formula we obtained will be a significant addition to this literature, particularly in the use of the Monte-Carlo method for the computations.

We assume the following conditions on the operator $L$ appearing in the equation (4.4).

(L1) $L$ is uniformly elliptic, that is, there exist positive numbers $\lambda$ and $\Lambda$ such that

$$\lambda|\xi|^2 \leq \sum_{i,j=1}^{d} a^{ij}(t,x) \xi^i \xi^j \leq \Lambda|\xi|^2, \quad \forall \, \xi \in \mathbb{R}^d .$$

(L2) For every $t$, the coefficients $a(t, \cdot)$ belong to $C^{2+\alpha}_b(\mathbb{R}^d)$ with bounded derivatives uniformly in $t$. That is,

$$\sup_t ||a(t, \cdot)||_{C^{2+\alpha}_b(\mathbb{R}^d)} \leq \Lambda .$$

(L3) $b$ is Lipschitz continuous and has linear growth, that is, there exists a positive constant $\kappa(b)$ such that

$$\sup_t |b^i(t, x)| \leq \kappa(b)(1 + |x|), \quad \forall \xi \in \mathbb{R}^d ,$$

$$\sup_t |b^i(t, y) - b^i(t, x)| \leq \kappa(b)|y - x|, \quad \forall \, x, y \in \mathbb{R}^d .$$

Under our conditions on $W$, it turns out that we can define the Feynman-Kac solution to equation (4.4), namely,

$$u(r, x) = \mathbb{E}^B \left[ u_T(X^r_T, x) \exp \left\{ \int_r^T W(ds, X^r_s) \right\} \right] ,$$

where $B$ is a Brownian motion started at $x$.
where \( \{X^{r,x}_s, s \geq r\} \) is the diffusion process generated by \( L \) starting from \( x \) at time \( r \). More precisely, for every \( r \leq t \leq T \) and \( x \in \mathbb{R}^d \), let \( X^{r,x}_t \) be the diffusion process given by the stochastic differential equation

\[
dX^{r,x}_t = \sigma^{ij}(t, X^{r,x}_t) \delta B^j_t + b^i(t, X^{r,x}_t) dt, \quad X^{r,x}_r = x,
\]

where \( \sigma \) is the square root matrix of \( a \), namely, \( a_{ij} = \sum_{k=1}^d \sigma^{ik} \sigma^{jk} \) and \( \delta B^j \) denotes the Itô differential. We will occasionally omit the index \( r, x \) and write \( X_s \) for \( X^{r,x}_s \).

Under conditions \((L1)-(L3)\), it is well-known that the diffusion process \( X^{r,x}_t \) exists and has finite moments of all orders.

Equation (4.4) with \( W \) replaced by \( W_n \) is classic and one can obtain a smooth solution \( u_n \) (see for instance [36] where a more general situation is studied). The main result of the current section is to show that \( u_n \) converges to the Feynman-Kac solution \( u \) defined above. There are three main tasks to be accomplished:

(i) One needs to define the nonlinear integration \( \int W(ds, X_s) \). Since here \( W \) is only continuous in time, this integration is different from the Young integration considered in Section 2.

(ii) One needs to show exponential integrability of \( \int W(ds, X_s) \). In particular, the function \( u \) defined by Feynman-Kac formula is well-defined.

(iii) One needs to show that the exponential functional of this integration is stable under approximations by smooth functions.

The outline of this section is as follows. In Subsection 4.1, we define the nonlinear stochastic integration \( \int W(ds, X_s) \) and show that it has finite moment of all orders. Exponential integrability is obtained if \( W \) has strictly sub-quadratic growth, namely, if \( \alpha \) and \( \beta \) in (4.1)-(4.3) satisfy \( \beta + \alpha < 1 \). In Subsection 4.2 we show that the Feynman-Kac solution is indeed a solution in a certain sense. When \( W \) has more regularity in time such as in the case of Brownian sheets or fractional Brownian sheets, one can use this regularity to reduce the regularity requirement in space. This case is considered in Subsection 4.3 when \( W \) satisfies the conditions in Section 2. Along the way, we will make use of some fundamental estimates for exponential moment of various norms of the diffusion \( X \) on finite intervals. These estimates are stated and proved in Appendix 3.

In what follows, \( \mathbb{E} \) denotes the expectation with respect to a Brownian motion \( B; \| \cdot \|_p \) denotes the \( L^p \) norm corresponding to \( \mathbb{E} \).

4.1. Nonlinear stochastic integral. Let \( X^{r,x}_t \) satisfy (4.6) and let \( W \) be in \( C^{0,1+\alpha}_\beta([0, T] \times \mathbb{R}^d) \). We shall define a new nonlinear integration \( \int_r^T W(ds, X^{r,x}_s) \). If \( W \) is differentiable in time, the natural definition for this type of integration is \( \int_r^T \partial_t W(s, X^{r,x}_s) ds \). If \( W \) satisfies \( \mathbf{(W)} \) then we can define it as in Section 2. However, in this section, Hölder continuity of \( W \) on \( t \) is not required. On the other hand, we shall use the crucial fact that \( \{X^{r,x}_t, t \geq r\} \) is a semimartingale. We first give the following definition.

**Definition 4.2.** Let \( W_n \) be a sequence of smooth functions with compact support converging to \( W \) in \( C^{0,1+\alpha}_\beta([0, T] \times \mathbb{R}^d) \). We define

\[
\int_r^T W(ds, X^{r,x}_s) = \lim_n \int_r^T \partial_s W_n(s, X^{r,x}_s) ds
\]

if the above limit exists in probability.
Of course, at first glance, there is no reason for the limit in (4.7) to converge. We will show, however, that the above definition is well-defined, thanks to the smoothing effect of the diffusion process $X_{s}^{r,x}$. Our first task is to find an appropriate representation for the integration $\int_{r}^{T} \partial_{t} W_{n}(s, X_{s}^{r,x}) ds$. To accomplish this, we consider the partial differential equation

$$(\partial_{t} + L_{0}) v_{n}(r, x) = - \partial_{t} W_{n}(r, x), \quad v(T, x) = - W_{n}(T, x),$$

where we recall that $L$ is defined by (4.5) and

$$L_{0} = L - b \nabla = \frac{1}{2} \sum_{i,j=1}^{d} a^{ij}(t, x) \partial_{x_{i}} \partial_{x_{j}}.$$

We could have chosen $L_{0} = L$, but the above choice of $L_{0}$ will allow us to show exponential integrability later. Since $W_{n}$ is a smooth function, the solution $v_{n}$ is a strong solution which is at least twice differentiable in space and once differentiable in time. We then apply the Itô formula to obtain

$$d v_{n}(s, X_{s}^{r,x}) = (\partial_{t} + L) v_{n}(s, X_{s}^{r,x}) ds + \sigma^{ij}(s, X_{s}^{r,x}) \partial_{x_{i}} v_{n}(s, X_{s}^{r,x}) \delta B_{s}^{j}$$

$$= - \partial_{t} W_{n}(s, X_{s}^{r,x}) ds - b(s, X_{s}^{r,x}) \cdot \nabla v_{n}(s, X_{s}^{r,x}) ds$$

$$+ \sigma^{ij}(s, X_{s}^{r,x}) \partial_{x_{i}} v_{n}(s, X_{s}^{r,x}) \delta B_{s}^{j}.$$

Thus, it follows that

$$\int_{r}^{T} \partial_{t} W_{n}(s, X_{s}^{r,x}) ds$$

$$(4.8) \quad = W_{n}(T, X_{T}^{r,x}) + v_{n}(r, x) - \int_{r}^{T} b(s, X_{s}^{r,x}) \cdot \nabla v_{n}(s, X_{s}^{r,x}) ds$$

$$+ \int_{r}^{T} \sigma^{ij}(s, X_{s}^{r,x}) \partial_{x_{i}} v_{n}(s, X_{s}^{r,x}) \delta B_{s}^{j}.$$

Notice that the time derivative in $W_{n}$ is transferred to the spatial derivative in $v_{n}$. The next task is to show that $v_{n}$ and its derivative $\nabla v_{n}$ converge. This is accomplished by some estimates which are in the same spirit of the well-known Schauder estimates for parabolic equations in Hölder spaces. More precisely, we have

**Lemma 4.3.** Suppose that $W$ belongs to $C^{2}_{\text{loc}}(\mathbb{R}^{d+1})$ and satisfies

$$[W]_{\beta_{1}, \infty} := \sup_{0 \leq t \leq T} \sup_{x \in \mathbb{R}^{d}} \frac{\|\nabla W(t, x)\|}{1 + |x|^\beta_{1}} < \infty$$

and

$$[W]_{\beta_{2}, \alpha} := \sup_{0 \leq t \leq T} \sup_{x \neq y \in \mathbb{R}^{d}} \frac{|\nabla W(t, x) - \nabla W(t, y)|}{|x - y|^\alpha (1 + |x|^\beta_{2} + |y|^\beta_{2})} < \infty$$

for some nonnegative numbers $\beta_{1}, \beta_{2}$. Let $v$ be a strong solution with polynomial growth to the partial differential equation

$$(\partial_{t} + L_{0}) v = - \partial_{t} W, \quad v(T, x) = - W(T, x).$$

Let $t \mapsto \varphi_{t}$ be the diffusion process generated by $L_{0}$, that is,

$$\varphi_{t}^{r,x} = x + \int_{r}^{t} \sigma(s, \varphi_{s}^{r,x}) \delta B_{s}, \quad t \geq r.$$
Then $v$ is uniquely defined and verifies

$$
\tag{4.11} (v + W)(r, x) = -\mathbb{E} \int_r^T L_0 W(s, \varphi_s^{r,x}) ds.
$$

In addition, the following estimates hold:

$$
\tag{4.12} \sup_{x \in \mathbb{R}^d} \frac{|(v + W)(r, x)|}{1 + |x|^{\beta_1}} \leq c(\beta_1, \lambda, \Lambda)[(T - r)^{1/2} + (T - r)] |\nabla W|_{\beta_1, \infty},
$$

$$
\tag{4.13} \sup_{x \in \mathbb{R}^d} \frac{\nabla(v + W)(r, x)}{1 + |x|^{\beta_2}} \leq c(\alpha, \beta_2, \lambda, \Lambda)[(T - r)^{\alpha/2} + (T - r)^{\alpha/2 + 1/2}] |\nabla W|_{\beta_2, \alpha},
$$

and for every $\alpha' \in (0, \alpha)$,

$$
\tag{4.14} \sup_{x \in \mathbb{R}^d} \frac{|\nabla(v + W)(r, x) - \nabla(v + W)(r, y)|}{(1 + |x|^{\beta_2} + |y|^{\beta_2})|x - y|^\alpha'} \leq c(\alpha', \alpha, \beta_2, \lambda, \Lambda)[(T - r)^{(\alpha - \alpha')/2} + (T - r)^{(\alpha - \alpha')/2 + 1/2}] |\nabla W|_{\beta_2, \alpha}.
$$

The proof of this result, even though lengthy, is straightforward and is provided in detail in Appendix C.

**Proposition 4.4.** Suppose that $W$ belongs to $C_{\beta}^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$. Then there exists a $C^1$-generalized solution $v$ to the parabolic partial differential equation

$$
\tag{4.15} (\partial_t + L_0) v = -\partial_t W, \quad v(T, x) = -W(T, x),
$$

such that for every $0 < \alpha' < \alpha$, the following estimates hold:

$$
\tag{4.16} [v + W]_{\alpha + \beta + 1, \infty} \leq c(\alpha, \beta, \lambda, \Lambda) |\nabla W|_{\alpha + \beta, \infty},
$$

$$
\tag{4.17} [\nabla(v + W)]_{\beta, \infty} \leq c(\alpha, \beta, \lambda, \Lambda) |\nabla W|_{\beta, \alpha},
$$

$$
\tag{4.18} [\nabla(v + W)]_{\beta, \alpha'} \leq c(\alpha, \alpha', \beta, \lambda, \Lambda) |\nabla W|_{\beta, \alpha}.
$$

As a consequence, $v$ belongs to the space $C_{\beta}^{0,1+\alpha^-}([0, T] \times \mathbb{R}^d)$.

**Proof.** We recall that $\eta$ is the bump function defined at the beginning of this section and $W_n = W * \eta_{1/n}$. Lemma 4.3 yields $[W_n - W]_{\beta, \infty}$ and $[W_n - W]_{\beta, \alpha}$ converge to 0 as $n \to \infty$. Thanks to linearity of the equation 4.15, $v_n - v_m$ is a strong solution to

$$
(\partial_t + L_0)(v_n - v_m) = -\partial_t(W_n - W_m), \quad (v_n - v_m)(T, x) = (W_n - W_m)(T, x).
$$

The results in Lemma 4.3 (with $\beta_1 = \beta_2 = \beta$) imply

$$
[(v_n + W_n) - (v_m + W_m)]_{\beta, \infty} \lesssim [\nabla W_n - \nabla W_m]_{\beta, \infty},
$$

$$
[\nabla(v_n + W_n) - \nabla(v_m + W_m)]_{\beta, \infty} \lesssim [\nabla W_n - \nabla W_m]_{\beta, \alpha},
$$

and for every $\alpha' \in (0, \alpha)$,

$$
[\nabla(v_n + W_n) - \nabla(v_m + W_m)]_{\beta, \alpha'} \lesssim [\nabla W_n - \nabla W_m]_{\beta, \alpha}.
$$

As a consequence, $v_n$ is a Cauchy sequence in $C([0, T], C^1(K))$ for every compact set $K$ in $\mathbb{R}^d$. Thus $v_n$ converges to $v$ in $C([0, T], C^1(K))$ for every compact set $K$. It is then straightforward to verify that $v$ is a weak solution to 4.15. The estimates 4.16, 4.17 and 4.18 follow from a limiting argument. \qed
Theorem 4.5. Suppose that \( W \) belongs to \( C^{0,1+\alpha}_{\beta}([0, T] \times \mathbb{R}^d) \). Let \( v \) be the \( C^{0,1+\alpha'}_{\beta} \)-generalized solution to (4.19) constructed in Proposition 4.4. Then for every \( t \in [r, T] \), the integration \( \int_r^t W(ds, X_{s}^{r,x}) \) is well-defined (in the sense of Definition 4.2). Moreover, it has moment of all positive orders and satisfies

\[
\int_r^t W(ds, X_{s}^{r,x}) = v(r, x) - v(t, X_{t}^{r,x}) - \int_r^t b(s, X_{s}^{r,x}) \cdot \nabla v(s, X_{s}^{r,x}) ds + \int_r^t \sigma^{ij}(s, X_{s}^{r,x}) \partial_x v(s, X_{s}^{r,x}) d\mathcal{B}_s^j.
\]

Proof. We consider \( W_n = W \ast \eta_{1/n} \) as in the proof of the previous proposition. It follows from the Itô formula that (see (4.8))

\[
\int_r^t \partial_t W_n(s, X_{s}^{r,x}) ds = v_n(r, x) - v_n(t, X_{t}^{r,x}) - \int_r^t b(s, X_{s}^{r,x}) \cdot \nabla v_n(s, X_{s}^{r,x}) ds + \int_r^t \sigma^{ij}(s, X_{s}^{r,x}) \partial_x v_n(s, X_{s}^{r,x}) d\mathcal{B}_s^j.
\]

Lemma 4.3 and Proposition 4.4 say that \( v_n \) (and its derivatives) has polynomial growth and converges in \( C([0, T]; C^{0,1+\alpha'}_{\beta}L^p)) \) to \( v \) for every \( \alpha' < \alpha \). Hence, the right hand side of the above formula is convergent in \( L^p(\Omega) \) for every \( p > 1 \). Passing through the limit in \( n \) yields the equation (4.19).

Remark 4.6. To define \( \int_r^t W(ds, X_{s}^{r,x}) \), usually one needs some regularity of \( W \) on the temporal variable \( t \). The equation (4.19) states that the requirement of the regularity on \( t \) can be transformed to the one on spatial variable \( x \) of another function \( v \) (defined by (4.9)). The use of \( v \) appears in many situations. If \( L_0 \) is replaced by \( L \) in the definition of \( v \) (e.g. equation (4.9)) and the terminal condition is replaced by \( v(0, x) = \delta(x-y) \) for any fixed \( y \), then \( v \) corresponds to the transition density of the process \( X_s \). This transition density is a fundamental concept in Markov processes and some other fields. It has also been used to simplify the proofs of a number of inequalities (see e.g. [1], [2]). The reason we use \( L_0 \) instead of \( L \) is that we don’t need to assume conditions on \( b \) to define \( v \) and that \( \partial_t v \) will appear in (4.19) even if we use \( L \). The removal of temporal regularity also appears in other contexts. For example, to study the equation \( dX_t = b(X_t) + dB_t \), the transformation \( Y_t = X_t - B_t \) will satisfy \( \dot{Y}_t = b(Y_t + B_t) \). The map \( (t, x) \mapsto \int_0^t b(x + B_s) ds \), averaging along the trajectories of a Brownian motion, then has better regularity than that of \( b \). In the field of stochastic differential equations, this phenomena has been observed by A. M. Davie in [10] and is recently studied in more depth in [5].

As a direct consequence, we obtain

Corollary 4.7. Let \( W \) be in \( C^{0,1+\alpha}_{\beta}([0, T] \times \mathbb{R}^d) \). Then for every \( \alpha' < \alpha, p > 2 \) and \( K \) compact subset of \( \mathbb{R}^d \),

\[
\| \int_r^T W(ds, X_{s}^{r,x}) - \int_r^T W(ds, X_{s}^{r,y}) - \int_r^T W_n(ds, X_{s}^{r,x}) + \int_r^T W_n(ds, X_{s}^{r,y}) \|_p \leq C(\alpha, \alpha', \beta, \lambda, \Lambda, K, T, p)([\nabla(W - W_n)]_{\beta, \infty} + [\nabla(W - W_n)]_{\beta, \alpha})|x - y|^\alpha'.
\]
Proof. Fix \( \alpha' < \alpha, p > 2 \) and \( K \) compact subset of \( \mathbb{R}^d \). We put \( g(r, x) = \int_r^T W(ds, X_s^{r,x}), g_n(r, x) = \int_r^T W_n(ds, X_s^{r,x}) \) and \( h = v - v_n \). From (4.19),

\[
\|g(r, x) - g(r, y) - g_n(r, x) + g_n(r, y)\|_p \leq I_1 + I_2 + I_3 + I_4,
\]

where

\[
egin{align*}
I_1 &= |h(r, x) - h(r, y)| \\
I_2 &= \|h(T, X_T^{r,x}) - h(T, X_T^{r,y})\|_p \\
I_3 &= \int_r^T \|(b \cdot \nabla h)(s, X_s^{r,x}) - (b \cdot \nabla h)(s, X_s^{r,y})\|_p ds \\
I_4 &= \|\int_r^T \sigma \nabla h)(s, X_s^{r,x}) - (\sigma \nabla h)(s, X_s^{r,y}) \cdot \delta B_s\|_p.
\end{align*}
\]

Proposition 4.4 implies

\[
|\nabla h(z)| \lesssim (|\nabla (W - W_n)|_{\beta, \infty} + |\nabla (W - W_n)|_{\beta, \alpha})(1 + |z|^\beta)
\]

and

\[
|\nabla h(x) - \nabla h(y)| \lesssim |\nabla (W - W_n)|_{\beta, \alpha}(1 + |x|^{\beta'} + |y|^{\beta'})|x - y|^{\alpha'}
\]

where \( \beta' = \beta + \alpha - \alpha' \). Therefore we can estimate

\[
I_1 = || \int_0^1 \nabla h(\tau x + (1 - \tau)y)d\tau(x - y)| \lesssim ||W - W_n|||x - y|,
\]

\[
I_2 = || \int_0^1 \nabla h(X_T^{r,x} + (1 - \tau)X_T^{r,y})d\tau(X_T^{r,x} - X_T^{r,y})||_p \lesssim ||W - W_n|||x - y|,
\]

\[
I_3 \leq \int_r^T \|(b(s, X_s^{r,x}) - b(s, X_s^{r,y})\nabla h(s, X_s^{r,x})\|_p ds \\
&+ \int_r^T \|b(s, X_s^{r,y})\nabla h(s, X_s^{r,x}) - \nabla h(s, X_s^{r,y})\|_p ds \\
\lesssim ||W - W_n|||x - y|^{\alpha'},
\]

where we have used the Hölder inequality. Similarly, we can estimate \( I_4 \) using the Burkholder-Davis-Gundy inequality to get \( I_4 \lesssim |x - y|^{\alpha'} \). From these bounds, the result follows. \( \square \)

**Proposition 4.8.** Suppose \( W \) belongs to \( C_{\alpha}^{0,1+\alpha}[0, T] \times \mathbb{R}^d \) with \( \alpha + \beta < 1 \). Then \( \int_r^T W(ds, X_s^{r,x}) \) is exponentially integrable uniformly over compact sets. More precisely, for every \( \gamma > 0, K \) compact subset of \( \mathbb{R}^d \),

\[
(4.20) \sup_{x \in K} \mathbb{E} \exp \left\{ \gamma \int_r^T W(ds, X_s^{r,x}) \right\} < \infty
\]

for all \( \gamma > 0 \).
Proof. From (4.19) it suffices to show that for every $\gamma > 0$,

$$\text{(4.21)} \quad \sup_{x \in K} \mathbb{E} \exp \left\{ \gamma \int_r^t \sigma^{ij}(s,X^{r,x}_s) \partial_i v(s,X^{r,x}_s) dB^j_s \right\} < \infty,$$

$$\text{(4.22)} \quad \sup_{x \in K} \mathbb{E} \exp \left\{ \gamma \int_r^t b(s,X^{r,x}_s) \cdot \nabla v(s,X^{r,x}_s) ds \right\} < \infty,$$

$$\text{(4.23)} \quad \sup_{x \in K} \mathbb{E} \exp \{ \gamma |v(t,X^{r,x}_t)| \} < \infty.$$

Let $0 < \theta < 2$. We claim that

$$\text{(4.24)} \quad \sup_{x \in K} \mathbb{E} \exp \left\{ \gamma \int_r^t |X^{r,x}_s|^\theta ds \right\} < \infty, \quad \forall \gamma > 0.$$

In fact, by the Jensen inequality,

$$\mathbb{E} \exp \left\{ \gamma \int_r^t |X^{r,x}_s|^\theta ds \right\} \leq (T-r)^{-1} \int_r^T \mathbb{E} e^{\gamma(T-r)|X^{r,x}_s|^\theta} ds.$$

The quality on the right hand side is finite thanks to (B.4).

For any martingale $M_t$ with $\mathbb{E} e^{2(M)_t} < \infty$ we have

$$\mathbb{E} e^{M_t} = \mathbb{E} e^{M_{t-(M)_t} e^{(M)_t}} \leq \left\{ \mathbb{E} e^{2M_{t-2(M)_t}} \right\}^{1/2} \left\{ \mathbb{E} e^{2(M)_t} \right\}^{1/2} = \left\{ \mathbb{E} e^{2(M)_t} \right\}^{1/2}.$$

Thus we have

$$\mathbb{E} \exp \left\{ \gamma \int_r^t \sigma^{ij}(s,X^{r,x}_s) \partial_i v(s,X^{r,x}_s) dB^j_s \right\} \leq \left\{ \mathbb{E} \exp \left[ 2\gamma^2 \int_r^t (a^{ij} \partial_i v \partial_j v)(s,X^{r,x}_s) ds \right] \right\}^{1/2}.$$

Taking into account the growth property of $\nabla v$ (see (4.17)) and $a$, we have

$$\sup_{x \in K} \mathbb{E} \exp \left[ 2\gamma^2 \int_r^t (a^{ij} \partial_i v \partial_j v)(s,X^{r,x}_s) ds \right] \leq \sup_{x \in K} \mathbb{E} \exp \left[ c \int_r^t |X^{r,x}_s|^{2(\alpha+\beta)} ds \right],$$

which together with the previous claim shows (4.21) since $2(\alpha + \beta) < 2$. Similarly, since $b$ has linear growth,

$$\sup_{x \in K} \mathbb{E} \exp \left\{ \gamma \int_r^t b(s,X^{r,x}_s) \cdot \nabla v(s,X^{r,x}_s) ds \right\} \leq \sup_{x \in K} \mathbb{E} \exp \left[ c \int_r^t |X^{r,x}_s|^{1+\alpha+\beta} \right],$$

which shows (4.22) since $1 + \alpha + \beta < 2$.

Using the growth property of $v$, i.e. the estimate (4.16),

$$\mathbb{E} \exp \{ \gamma |v(t,X^{r,x}_t)| \} \leq \mathbb{E} \exp \left[ c |X^{r,x}_t|^{1+\alpha+\beta} \right],$$

which shows (4.23). \hfill \Box

Lemma 4.9. Let $W$ be in $C^{0,1+\alpha}_\beta([0,T] \times \mathbb{R}^d)$. Suppose $\alpha + \beta < 1$. For every $\gamma > 0$ and $r \in [0,T]$, we put $u(r,x) = \mathbb{E} \exp \left[ \gamma \int_r^T W(ds,X^{r,x}_s) \right]$ and $u_n(r,x) = \mathbb{E} \exp \left[ \gamma \int_r^T W_n(ds,X^{r,x}_s) \right]$. Then $u_n$ converges to $u$ in $C^{0,\alpha'}([0,T] \times K)$ for every $\alpha' < \alpha$ and $K$ compact in $\mathbb{R}^d$. 

\begin{proof}
\end{proof}
Proof. For a smooth function $f$, using fundamental theorem of calculus, we obtain

$$f(x) - f(a) - f(y) + f(b) = \int_0^1 \int_0^1 f''(\xi) [\tau(x - y) + (1 - \tau)(a - b)] d\eta d\tau (x - a)$$

$$+ \int_0^1 f'(\theta) d\tau (x - a - y - b),$$

where

$$\xi = \tau \eta x + (1 - \tau) \eta a + \tau (1 - \eta) y + (1 - \tau) (1 - \eta) b,$$

$$\theta = \tau y + (1 - \tau) b.$$ 

Thus, for every $x, y$ in $K$, with $f(w) = \exp(\gamma w)$, we have

$$u(r, x) - u_n(r, x) - u(r, y) + u_n(r, y)$$

$$= \gamma^2 \mathbb{E} \int_0^1 \int_0^1 f(\xi) [\tau A(x, y) + (1 - \tau) A_n(x, y)] d\eta d\tau B_n(x)$$

$$+ \gamma \mathbb{E} \int_0^1 f(\theta) d\tau C_n(x, y),$$

(4.25)

where

$$A(x, y) = \int_r^T W(ds, X_s^r, x) - \int_r^T W(ds, X_s^r, y),$$

$$A_n(x, y) = \int_r^T W_n(ds, X_s^r, x) - \int_r^T W_n(ds, X_s^r, y),$$

$$B_n(x) = \int_r^T W(ds, X_s^r, x) - \int_r^T W_n(ds, X_s^r, x),$$

$$C_n(x, y) = A(x, y) - A_n(x, y).$$

The random variables $\xi$ and $\eta$ are linear combinations of these terms. From Proposition 4.8 we know that moments of $f(\xi)$ and $f(\theta)$ are bounded uniformly in $x$ and $\tau, \eta$. On the other hand, from Corollary 4.7 for every $\alpha' < \alpha$ and $p > 2$,

$$\|A(x, y)\|_p \lesssim |x - y|^{\alpha'},$$

$$\sup_n \|A_n(x, y)\|_p \lesssim |x - y|^{\alpha'},$$

$$\lim_{n \to 0} \sup_{x \in K} \|B_n(x)\| = 0,$$

and

$$\|C_n(x, y)\|_p \lesssim (|\nabla(W - W_n)|_{\beta, \infty} + |\nabla(W - W_n)|_{\beta, \alpha}) |x - y|^{\alpha'}.$$ 

From (4.25), applying Hölder inequality and the above estimates for $A, B, C$ we obtain

$$|u(r, x) - u_n(r, x) - u(r, y) + u_n(r, y)|$$

$$\lesssim \left[ \sup_{x \in K} \|B_n(x)\|_p + |\nabla(W - W_n)|_{\beta, \infty} + |\nabla(W - W_n)|_{\beta, \alpha} \right] |y - x|^{\alpha'}$$

for all $x, y$ in $K$ and $\alpha' < \alpha$. This completes the proof. \qed
4.2. **Feynman-Kac formula I.** If \( W \) is a smooth function, then the classical Feynman-Kac formula asserts that

\[
(4.26) \quad u(r, x) = \mathbb{E}^B \left[ u_T(X^r_T) \exp \left( \int_r^T W(ds, X^r_s) \right) \right]
\]

is the unique strong solution to (4.4). Indeed, suppose \( W \) is smooth and \( u \) is a strong solution to (4.4). Applying the Itô formula to the process

\[
t \mapsto u(t, X^r_t) \exp \left\{ \int_r^t \frac{\partial_t}{t} W(s, X^r_s)ds \right\}
\]

we obtain

\[
\delta u(t, X^r_t) \exp \left\{ \int_r^t \frac{\partial_t}{t} W(s, X^r_s)ds \right\} = \exp \left\{ \int_r^t \frac{\partial_t}{t} W(s, X^r_s)ds \right\} (\partial_t + L + \frac{\partial_t}{t})u(t, X^r_t)dt + \exp \left\{ \int_r^t \frac{\partial_t}{t} W(s, X^r_s)ds \right\} \sigma^{ij}(t, X^r_t)\partial_{x_i}u(t, X^r_t)\delta B_i^j.
\]

Taking into account that \((\partial_t + L)u + \frac{\partial_t}{t} W u = 0\) and integrating over \([r, T]\), we have

\[
u_T(X^r_T) \exp \left\{ \int_r^T \frac{\partial_t}{t} W(s, X^r_s)ds \right\} - u(r, x) = \int_r^T \exp \left\{ \int_r^t \frac{\partial_t}{t} W(s, X^r_s)ds \right\} \sigma^{ij}(t, X^r_t)\partial_{x_i}u(t, X^r_t)\delta B_i^j.
\]

Formula (4.26) is deduced by taking the expectation on both sides.

**Theorem 4.10.** Assume \( W \) belongs to \( C^{0,1+\alpha}_{\beta}(\mathbb{R}^d) \) with \( \alpha + \beta < 1 \). Let \( W_n = W * \eta_{1/n} \). Let \( u_n \) be the solution to the parabolic equation

\[
\partial_t u_n + Lu_n + u_n \partial_t W_n = 0, \quad u_n(T, x) = u_T(x).
\]

Let \( u \) be the function defined in equation (4.26). Then \( u_n \) converges to \( u \) in \( C^{0,\alpha'}([0, T] \times K) \) for every \( \alpha' < \alpha \) and \( K \) compact set in \( \mathbb{R}^d \). As a consequence, \( u \) belongs to \( C^{0,\alpha'}_{\text{loc}}([0, T] \times \mathbb{R}^d) \) for all \( \alpha' < \alpha \).

**Proof.** We notice that

\[
u_n(r, x) - u(r, x)
\]

\[
= \mathbb{E} \left\{ u_T(X^r_T) \left\[ \exp \left( \int_r^T W_n(ds, X^r_s) \right) - \exp \left( \int_r^T W(ds, X^r_s) \right) \right\] \right\}.
\]

This together with Lemma 4.9 yields the theorem. \(\square\)

We notice that \( f \) and \( g \) are locally Hölder continuous functions on \( \mathbb{R}^d \) with exponents \( \alpha \) and \( \gamma \) respectively. Suppose that \( f \) has compact support and \( \alpha + \gamma > 1 \). Then we can define the Young integral

\[
\int_{\mathbb{R}^d} f(x)g(d^j x) = \int_{\mathbb{R}^d} f(x)g(x_1, \ldots, x_{j-1}, dx_j, x_{j+1}, \ldots, x_n)d\hat{x}_j
\]

where \( \hat{x}_j = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \).
We now show that if $W$ is sufficiently regular in space, the Feynman-Kac solution $u$ in (4.26) satisfies an equation derived from (4.4) by a change of variable. To better explain our procedure, let us first assume that $W$ is smooth in space and time and $u_T$ is also smooth. In such case, the equation (4.4) has unique smooth solution $u$ such that

$$\partial_t u(t, x) + Lu(t, x) + u \partial_t W(t, x) = 0$$

for every $t \geq 0$ and $x \in \mathbb{R}^d$. We would like to obtain an equation of $u$ such that the time derivative of $W$ does not appear. To this end, we notice that

$$\partial_t u + u \partial_t W = e^{-W} \partial_t (ue^W).$$

Hence, multiplying the equation with $e^W$ and integrating in time, we obtain

(4.27) $$u_t = e^{W_T - W_t} u_T + \int_t^T e^{W_s - W_t} Lu_s ds.$$ 

In contrast with (4.4), the equation (4.27) does not contain the time derivative of $W$. One can also interpret (4.27) in the weak sense. More precisely, the following result holds.

**Theorem 4.11.** Assume $W$ belongs to $C^{0,1+\alpha}_\beta([0, T] \times \mathbb{R}^d)$ with $\alpha + \beta < 1$. Let $u$ be the function defined in (4.26). Then there is a sequence of smooth functions $W_n$ with compact supports convergent to $W$ in $C^{0,1+\alpha}_\beta([0, T] \times \mathbb{R}^d)$ and a sequence of $u_n$ such that $u_n$ converges to $u$ uniformly over all compact sets. Moreover, for every test function $\varphi \in C^\infty_c(\mathbb{R}^d)$ the sequence

$$\int_t^T \int_{\mathbb{R}^d} \partial_i [e^{W_n(s,x) - W_n(t,x)} \varphi(x)] a^{ij}(s,x) \partial_j u_n(s,x) dx ds$$

is convergent. If $\alpha > 1/2$, then we can identify the limit as

$$\int_t^T \int_{\mathbb{R}^d} \partial_i (e^{W(s,x) - W(t,x)} \varphi(x)) a^{ij}(s,x) u(s,d^j x) ds.$$

In such case, $u$ verifies the equation

(4.28) $$\int_{\mathbb{R}^d} u(t,x) \varphi(x) dx = \int_{\mathbb{R}^d} e^{W(T,x) - W(t,x)} u_T(x) \varphi(x) dx$$

$$+ \frac{1}{2} \int_t^T \int_{\mathbb{R}^d} \partial_i (e^{W(s,x) - W(t,x)} \varphi(x)) a^{ij}(s,x) u(s,d^j x) ds$$

$$- \int_t^T \int_{\mathbb{R}^d} \partial_i \left( e^{W(s,x) - W(t,x)} \varphi(x) \left[ b^i(s,x) - \frac{1}{2} \partial_j a^{ij}(s,x) \right] \right) u(s,x) dx ds.$$ 

**Proof.** We recall that $W_n = W * \eta_{1/n}$ defined at the beginning of this section. Let $u_n$ be the solution to the parabolic equation

$$\partial_t u_n + Lu_n + u_n \partial_t W_n = 0 , \quad u_n(T,x) = -W_n(T,x).$$
Then it is easily verified that
\[
\int_{\mathbb{R}^d} u_n(t, x) \varphi(x) dx = \int_{\mathbb{R}^d} e^{W_n(T, x) - W_n(t, x)} u_n(T, x) \varphi(x) dx
\]
\[
+ \frac{1}{2} \int_t^T \int_{\mathbb{R}^d} \partial_i [e^{W_n(s, x) - W_n(t, x)} \varphi(x)] a^{ij}(s, x) \partial_j u_n(s, x) dx ds
\]
\[
+ \int_t^T \int_{\mathbb{R}^d} e^{W_n(s, x) - W_n(t, x)} \varphi(x) \left[ b^i(s, x) - \frac{1}{2} \partial_j a^{ij}(s, x) \right] \partial_i u_n(s, x) dx ds.
\]

In other words,
\[
\frac{1}{2} \int_t^T \int_{\mathbb{R}^d} \partial_i [e^{W_n(s, x) - W_n(t, x)} \varphi(x)] a^{ij}(s, x) \partial_j u_n(s, x) dx ds
\]
\[
= \int_{\mathbb{R}^d} u_n(t, x) \varphi(x) dx - \int_{\mathbb{R}^d} e^{W_n(T, x) - W_n(t, x)} u_n(T, x) \varphi(x) dx
\]
\[
+ \int_t^T \int_{\mathbb{R}^d} \partial_i \left( e^{W_n(s, x) - W_n(t, x)} \varphi(x) \left[ b^i(s, x) - \frac{1}{2} \partial_j a^{ij}(s, x) \right] \right) u_n(s, x) dx ds.
\]

Since \( \varphi \) has compact support, it is clear that all the terms on the right hand side are convergent. This implies that
\[
\int_t^T \int_{\mathbb{R}^d} \partial_i [e^{W_n(s, x) - W_n(t, x)} \varphi(x)] a^{ij}(s, x) \partial_j u_n(s, x) dx ds
\]
is convergent. In case \( \alpha > 1/2 \), by Theorem 4.10, this limit is convergent in the context of Young integrations. Hence, taking the limit yields (4.28). □

Remark 4.12. (i) The use of Itô’s formula in Subsection 4.1 is inspired from the work [19]. In that work, an Itô-Tanaka trick is applied to obtain some estimates to the commutator related to DiPerna-Lions’ theory ([13]).

(ii) In the case \( W \) belongs to \( C_0^{0,2}([0, T] \times \mathbb{R}^d) \), the Itô-Tanaka formula (4.8) is negligible. In fact, using integration by parts, one has
\[
\int_r^T \partial_t W_n(s, X^{r,x}_s) ds = W_n(T, X^{r,x}_T) - W_n(r, x) - \int_r^T \nabla W_n(s, X^{r,x}_s) dX^{r,x}_s
\]
where the last integral is in the Stratonovich sense. By passing through the limit \( n \to \infty \), we obtain
\[
\int_r^T \partial_t W(s, X^{r,x}_s) ds = W(T, X^{r,x}_T) - W(r, x) - \int_r^T \nabla W(s, X^{r,x}_s) dX^{r,x}_s.
\]

Assuming \( \nabla W \) has linear growth in the spatial variable and \( \nabla^2 W \) is globally bounded, one can also show exponential integrability
\[
\mathbb{E}^B \exp \left[ \int_r^T \partial_t W(s, X^{r,x}_s) ds \right] < \infty.
\]
We consider $u$ as in (4.26). Using the approximation as in the proof of Theorem 4.11, we can show that $u$ verifies
\[
\int_{\mathbb{R}^d} u(t, x) \varphi(x) dx = \int_{\mathbb{R}^d} e^{W(T, x) - W(t, x)} u(T, x) \varphi(x) dx + \int_t^T \int_{\mathbb{R}^d} L^*[e^{W(s, x) - W(t, x)} \varphi(x)] u(s, x) dx ds
\]
for all test functions $\varphi$ in $C^\infty_c(\mathbb{R}^d)$, where $L^*$ is the adjoint of $L$.

4.3. Feynman-Kac formula II. In previous subsections, to obtain the Feynman-Kac solution (4.26) (see Theorem 4.11) we assume that $W$ is only continuous in time but satisfies (4.1)-(4.3) for $f = W$. This means that we suppose the first spatial derivatives of $W$ exist and are Hölder continuous in order to compensate the lack of regularity in time. For many other stochastic processes (such as Brownian sheet or fractional Brownian sheets), $W$ is Hölder continuous in time. In this case, we may use this time regularity to relax the regularity requirement on the space variable. In this subsection we obtain a Feynman-Kac formula for the solution to (4.4) when $W$ satisfies the conditions of the type given in Section 2. For example, we do not require $W$ to possess first derivatives. More precisely, we assume $W : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ satisfies the following condition.

\[(FK)\] There are constants $\tau, \lambda \in (0, 1]$ and $\beta > 0$ such that
\[
\tau + \frac{1}{2} \lambda > 1, \quad \beta + \lambda < 2
\]
and such that the seminorm
\[
\|W\|_{\beta, \tau, \lambda} \leq \sup_{0 \leq s < t \leq T} \sup_{x, y \in \mathbb{R}^d; x \neq y} \frac{|W(s, x) - W(t, x) - W(s, y) + W(t, y)|}{(1 + |x| + |y|)^\beta |t - s|^{\tau} |x - y|^{\lambda}}
\]
\[
+ \sup_{0 \leq s < t \leq T} \sup_{x \in \mathbb{R}^d} \frac{|W(s, x) - W(t, x)|}{(1 + |x|)^{\beta + \lambda} |t - s|^{\tau}} + \sup_{0 \leq t \leq T} \sup_{x, y \in \mathbb{R}^d; x \neq y} \frac{|W(t, y) - W(t, x)|}{(1 + |x| + |y|)^\beta |x - y|^{\lambda}}
\]
is finite.

We continue to use the same notation introduced in previous subsections. For example, $X_t = X_t^{r, x}$ denotes the solution to the equation (4.6). The objectives of this subsection are to show that the expression defined by (4.26) is well-defined under the above condition (FK) and is the solution to (4.4).

From $\tau + \frac{1}{2} \lambda > 1$, it follows that there is a $\gamma \in (0, 1/2)$ such that $\tau + \gamma \lambda > 1$. Since $X_t$ is Hölder continuous of exponent $\gamma$, from Proposition 2.2, we know that $\int_r^T W(ds, X_s^{r, x})$ is well-defined and
\[
\left| \int_r^T W(ds, X_s) \right| \leq C(1 + \|X\|_{\beta}^\gamma)(1 + \|X\|_{\lambda}^\gamma).
\]

Since $\beta + \lambda < 2$, Lemma 3.2 yields that
\[
\mathbb{E} \exp \left\{ c \int_r^T W(ds, X_s) \right\} < \infty
\]
for all $c \in \mathbb{R}$. Thus we have the following proposition.
Proposition 4.13. Assume the conditions \((L1)-(L3)\) are satisfied. Let \((L3)-(L3.30)\) be satisfied. If there is an \(\alpha_0 \in (0,2)\) such that \(|u_T(x)| \leq C_2 e^{C_1|x|^\alpha_0}\), then \(u(r,x)\) defined by \((4.26)\) is finite. Namely,

\[
(4.32) \quad u(r, x) = \mathbb{E}^B \left[ u_T(X_t^r,x) \exp \left( \int_r^T W(ds, X_s^r,x) \right) \right]
\]

is well-defined.

Now, let \(W_n(t,x)\) be a sequence of functions in \(C_0^\infty([0,T] \times \mathbb{R}^d)\) convergent to \(W(t,x)\) under the norm \(\|W\|_\infty + \|W\|_{\beta,\tau,\lambda}\). Denote \(v_n(r,x) = \int_r^T W_n(ds, X_s^r,x)\) and \(v(r,x) = \int_r^T W(ds, X_s^r,x)\) and \(\tilde{v}_n(r,x) = v_n(r,x) - v(r,x)\). Thus, for any \(0 \leq r < t \leq T\), we have

\[
|\tilde{v}_n(t,x) - \tilde{v}_n(r,x)| = \left| \int_r^t \tilde{W}_n(ds, X_s^r,x) - \int_r^t \tilde{W}_n(ds, X_s^r,x) \right| \\
\leq \left| \int_r^t \tilde{W}_n(ds, X_s^r,x) \right| + \left| \int_t^T \left[ \tilde{W}_n(ds, X_s^r,x) - \tilde{W}_n(ds, X_s^r,x) \right] \right| =: I_1(r,t) + I_2(r,t).
\]

Applying the estimate in Proposition 2.4 to \(\tilde{W}_n = W_n - W\), we obtain

\[
I_1(r,t) \leq \kappa \mathbb{E} \left( \sup_{0 \leq r < t \leq T} \left| \left( t - r \right) \right| \right) = 0
\]

for any \(p \geq 1\). Thus

\[
\lim_{n \to \infty} \mathbb{E} \left( \sup_{0 \leq r < t \leq T} \left| \left( t - r \right) \right| \right) = 0
\]

for any \(p \geq 1\).

From Proposition 2.11 we have with \(\tau + \theta \lambda \gamma > 1\),

\[
I_2(r,t) \leq C \mathbb{E} \left( \sup_{0 \leq r < t \leq T} \left| \left( t - r \right)^{\gamma \tau / 2} \right| \right)
\]

\[
\leq C \mathbb{E} \left( \sup_{0 \leq r < t \leq T} \left| \left( t - r \right)^{\gamma \tau / 2} \right| \right) \leq C ,
\]

where the last inequality follows from a similar argument as the proof of \((L3.5)\). Combining this with \((4.34)\) implies

\[
\mathbb{E} \left( \sup_{0 \leq r < t \leq T} \left| \left( t - r \right)^{\gamma \tau (1 - \theta) / 2} \right| \right) \leq C .
\]
Assume $\lambda/2 + \tau - 1 > 0$. For any $\tau' \in (0, \lambda/2 + \tau - 1)$ it is possible to find $\theta \in (0, 1)$ and $0 < \gamma < 1/2$ such that $\tau + \theta \lambda \gamma > 1$ and $\tau' < \gamma' \lambda (1 - \theta)/2$. We see that $v(\cdot, x)$ is Hölder continuous of exponent $\tau'$ and

$$
\lim_{n \to \infty} \|v_n(\cdot, x) - v(\cdot, x)\|_{x'} = 0
$$

uniformly in compact set $K$ of $\mathbb{R}^d$. From (4.32) it is easy to see that

$$
\lim_{n \to \infty} \|u_n(\cdot, x) - u(\cdot, x)\|_{x'} = 0
$$

uniformly in compact set $K$ of $\mathbb{R}^d$. Thus we have

**Proposition 4.14.** Let $W_n$ be a sequence of smooth functions such that $W_n$ converges to $W$ in the norm $\|W\|_{\infty} + \|W\|_{\beta, \tau, \lambda}$ and $u_n$ is the solution to (4.4) and $u$ is given by (4.32). Then for any $\tau' < \lambda/2 + \tau - 1$, $u(t, x)$ is Hölder continuous of exponent $\tau'$ in time variable $t$ and on any compact set $K$ of $\mathbb{R}^d$,

$$
\lim_{n \to \infty} \|u_n(\cdot, x) - u(\cdot, x)\|_{x'} = 0
$$

uniformly on $x \in K$.

If $\tau + \tau' > 1$, then for any $\varphi \in C^\infty_0(\mathbb{R}^d)$, we have

$$
\int_0^t \int_{\mathbb{R}^d} u_n(s, x) (s, x) \varphi(x) \frac{\partial}{\partial s} W_n(s, x) ds dx
$$

converges to the Young integral

$$
\int_0^t \int_{\mathbb{R}^d} u(s, x) (s, x) \varphi(x) W(ds, x) dx.
$$

It is obvious that the existence of $\tau' > 0$ such that $\tau + \tau' > 1$ and $\tau' < \lambda/2 + \tau - 1$ is equivalent to $\lambda + 4\tau > 4$. The above argument means that $u(t, x)$ is a weak solution to (4.4), in the sense of the next theorem.

**Theorem 4.15.** Assume the conditions (L1)–(L3) are satisfied and assume there is an $\alpha_0 \in (0, 2)$ such that $|w_T(x)| \leq C_2 e^{C_1|x|^\alpha_0}$. Let $\|W\|_{\beta, \tau, \lambda}$ defined by (4.30) be finite, where the Hölder exponents $\lambda$ and $\tau$ and the growth exponent $\beta$ satisfy

$$
\tau > 1/2, \quad \beta + \lambda < 2, \quad \lambda + 4\tau > 4.
$$

Then $u$ defined by (4.32) is a weak solution to (4.4) in the sense that $u$ satisfies

$$
\int_{\mathbb{R}^d} u(t, x) \varphi(x) dx = \int_{\mathbb{R}^d} u_0(x) \varphi(x) dx + \int_0^t \int_{\mathbb{R}^d} u(s, x) L^* \varphi(x) dx ds
$$

$$
+ \int_0^t \int_{\mathbb{R}^d} u(s, x) \varphi(x) W(ds, x) dx,
$$

where $\varphi$ is any smooth function with compact support and where the last integral is a Young integral.

**Remark 4.16.** Equation (4.37) is the definition of the weak solution used in [32], [34] and [35].
5. ASYMPTOTIC GROWTH OF GAUSSIAN SAMPLE PATHS

In Sections 2, 3 and 4 we assume the pathwise Hölder continuity and pathwise growth conditions on $W$ in order to define and to solve (partial) differential equations related to the nonlinear integral $\int W(ds, \varphi_s)$. For instance, the conditions (2.1), (4.1), (4.2), (4.3) are essential in various parts of the paper. In probability theory, it is usually hard to obtain properties for (almost) every sample path of a stochastic process from its average properties (from its probability law). In this section, we investigate this pathwise Hölder continuity and these pathwise growth problems for a stochastic process. We shall focus on Gaussian random fields. However, our method works well for other processes provided they satisfy some suitable normal concentration inequalities (for instance, see the assumptions in Theorem 5.5).

Let $W$ be a stochastic process on $[0, T] \times \mathbb{R}^d$. An application of our results yields the asymptotic growth of the quantity

$$I(\delta, R) = \sup_{t \in [0, T]} \sup_{|x|, |y| \leq R; |x-y| \leq \delta} \frac{|W(t, \Box[x, y])|}{|x_1 - y_1|^\lambda \cdots |x_d - y_d|^\lambda}$$

as $R \to \infty$, where $W(t, \Box[x, y])$ denotes the $d$-increment of $W(t, \cdot)$ over the rectangle $[x, y]$. A more precise definition is given in Subsection 5.2. If $R$ is fixed, the quality $I(\delta, R)$ is the objective in our previous work [30] via a multiparameter version of Garsia-Rodemich-Rumsey inequality.

Let us mention some historical remarks. (Pathwise) boundedness and continuity for stochastic processes have been studied thoroughly in the literature. One of the central ideas is originated in an important early paper by Garsia, Rodemich and Rumsey (1970) [21]. This was developed further by Preston (1971, 1972) [44, 45], Dudley (1973) [15] and Fernique (1975) [16]. In these considerations, the parameter space $T$ is bounded and treated as a “single-dimension” object. For instance, the well-known Dudley bound

$$\mathbb{E} \sup_{s,t \in T} |W(t) - W(s)| \lesssim \int_0^{d_w(s,t)} \sqrt{\log N(T, d_w, \varepsilon)} d\varepsilon$$

yields modulus of continuity in terms of the entropy number $N(T, d, \varepsilon)$. This is extended to a more precise bound in terms of majorizing measure

$$\mathbb{E} \sup_{s,t \in T, d_w(s,t) \leq \delta} |W(t) - W(s)| \lesssim \sup_{t \in T} \int_0^{\delta} \log^{1/2} \frac{1}{\mu(B_{d_w}(t, u))} du .$$

The majorizing-measure bound turns out to be necessary for processes which satisfy normal concentration inequalities. This result by M. Talagrand is a milestone in the theory of Gaussian processes. We refer the reader to [41 Chapter 6] and references therein for details and more historical facts. See also Talagrand’s monograph [49] in which the role of majorizing measure is replaced by a variational quality called $\gamma_2(T, d_w)$.

Estimates for the $d$-increment of $W$ over a rectangle are quite different. Difficulties arise since $W(\Box[s, t])$ does not behave nicely as increments. In particular, the corresponding entropic “metric”

$$\left(\mathbb{E} W(\Box[s, t])^2\right)^{1/2}$$
does not satisfy the triangle inequality, but rather behaves like a volume metric. To elaborate this point, let us consider the two dimensional case:

\[ W(\square|s, t|) = W(s_2, t_2) - W(s_2, t_1) - W(s_1, t_2) + W(s_1, t_1) \]

\[ = \Delta_{[s_2,t_1]} W(s_2) - \Delta_{[s_2,t_1]} W(s_1) = \Delta_{[s_2,s_1]} \Delta_{[t_2,t_1]} W, \]

where \( W(s) := \Delta_{[t_2,t_1]} W(s) = W(s; t_2) - W(s; t_1) \). This product-like property is essential in our current approach (see for instance inequality (5.5) below). Alternatively, to obtain a sharp bound for the difference, one can repeatedly apply the Garsia-Rodemich-Rumsey inequality first to \( \Delta_{[s_2,s_1]} \) and then to \( \Delta_{[s_2,s_1]} \Delta_{[t_2,t_1]} \). Indeed, for bounded parameter domains equipped with Lebesgue measure, this direction was developed by the authors in [30]. This idea, while it might be feasible, seems to be more complicated in our current setting with general (unbounded) parameter domains equipped with a general measure.

In Subsection 5.1, we will prove a deterministic inequality, which is more precise than the multiparameter Garsia-Rodemich-Rumsey inequality obtained in [30]. We then apply it to obtain a majorizing-measure bound on the \( d \)-increments of stochastic processes in Subsection 5.2. Our formulations benefit from the treatment in [41]. We however did not consider the necessary conditions for these bounds (i.e. lower bounds). Results in these two subsections are applicable to general stochastic processes.

Given a well-developed toolbox to treat the case when \( T \) is bounded (or for example, \( R \) is fixed in \( I(\delta, R) \)), the asymptotic growth for \( I(\delta, R) \) as \( R \to \infty \) can be obtained using concentration inequalities for Gaussian processes. More precise results are given for fractional Brownian fields. This is done in Subsection 5.3.

5.1. A deterministic inequality.

Throughout the current subsection, we put \( \Psi(u) = \exp(u^2) - 1 \). Suppose \( \mu \) is a nonnegative measure on \( T \) and \( X \) is a measurable function on \( T \). We define

\[ [x](T, \mu) := \inf \left\{ \alpha > 0 : \int_T \Psi \left( \frac{X(t)}{\alpha} \right) \mu(dt) \leq 1 \right\}. \]

When the parameter space \( T \) and the measure \( \mu \) are clear from the context, we often suppress them and write \( [x] \) instead. The following result, whose proof is given in [41] pp. 256-258, is an application of the Young inequality

\[ ab \leq \int_0^ag(x)dx + \int_0^bg^{-1}(x)dx, \]

where \( g \) is a real-valued, continuous and strictly increasing function.

Lemma 5.1. Let \( X \) and \( f \) be measurable functions on \( T \), \( \mu \) be a nonnegative measure on \( T \). Assume that \( [X](T, \mu) \) is finite and \( 0 < \int |f|d\mu < \infty \). Then

\[ \int_T |X(t)f(t)|\mu(dt) \leq 3[X](T, \mu) \int_T |f(t)| \log^{1/2} \left( 1 + \frac{|f(t)|}{\int |f(s)|\mu(ds)} \right) \mu(dt). \]

We consider the case when \( T \) has the form \( T = T_1 \times \cdots \times T_\ell \). A parameter \( t \) in \( T \) has \( \ell \) components, \( t = (t_1, \ldots, t_\ell) \). For each \( i = 1, \ldots, \ell \), the space \( T_i \) is equipped with a metric \( d_i \). We also denote \( d^*(s, t) = d_1(s_1, t_1) \ldots d_\ell(s_\ell, t_\ell) \) for every \( s, t \) in \( T \). Let \( X \) be a function on \( T \). We define the \( \ell \)-increment of \( X \) over a “rectangle”
\[ [s, t] \text{ as} \]
\[ X(\Box [s, t]) = \prod_{j=1}^{\ell} (I - V_{j,s})X(t). \]

In the above expression, \( I \) is the identity operator, \( V_{j,s} \) is the substitution operator which substitutes the \( j \)-th component of a function on \( T \) by \( s_j \), more precisely,
\[ V_{j,s}X(t) = X(t_1, \ldots, t_{j-1}, s_j, t_{j+1}, \ldots, t_{\ell}). \]

We refer to [30] for a more detailed description on this \( \ell \)-increment.

For each \( i \), \( B^i(t_i, u) \) denotes the open ball with radius \( u \) in the metric space \((T_i, d_i)\) centered at \( t_i \). For each \( t \) in \( T \), we denote \( B(t, u) = B^1(t_1, u) \times \cdots \times B^\ell(t, u) \).

For each \( j \), put \( D_j = \sup_{s_j, t_j \in T_j} d_j(s_j, t_j) \).

For each \( i = 1, \ldots, \ell \), let \( \mu_i \) be a probability measure on \( T_i \). Let \( k = (k_1, \ldots, k_\ell) \) be a multi-index in \( \mathbb{N}^\ell \). We define
\[ \mu_k^i(t_i) = \mu^i(B^i(t_i, D_i 2^{-k_i})), \quad \rho_k(t, \cdot) = \frac{1}{\mu_k^i(t_i)} B(t, D_i 2^{-k_i})(\cdot), \]
\[ \mu_k(t) = \prod_{i=1}^\ell \mu_k^i(t_i), \quad \rho_k(t, \cdot) = \prod \rho_k(t_i, \cdot), \]
and
\[ (5.1) \quad M_k(t) = \int_T \rho_k(t, u) X(u) \mu(du). \]

We use the notation \( k + 1 = (k_1 + 1, \ldots, k_\ell + 1), k + j = (k_1, \ldots, k_{j-1}, k_j + 1, k_{j+1}, \ldots, k_\ell), \hat{t}_i = (t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_\ell) \) and \( T_i = T_1 \times \cdots \times T_{i-1} \times T_{i+1} \times \cdots \times T_\ell. \)

**Theorem 5.2.** Let \( \{X(t), t \in T\} \) be a measurable function on \( T \). We put \( \mu = \mu^1 \times \cdots \times \mu^\ell \) and
\[ Z = \inf \left\{ \alpha > 0 : \int_{T \times T} \Psi(X(\Box [u, v])) \mu(du) \mu(dv) \leq 1 \right\}. \]

Assume that \( D_j, j = 1, \ldots, d, \) and \( Z \) are finite. Then, for every \( s, t \) in \( T \) such that the integral
\[ \int_0^{d_1(s_1, t_1)} du_1 \cdots \int_0^{d_\ell(s_\ell, t_\ell)} du_\ell \left( \frac{1}{\mu(B(s, u))} + \log^{1/2} \frac{1}{\mu(B(t, u))} \right) \]
is finite, \( M_k(\Box [s, t]) \) converges to a limit, denoted by \( X'(\Box [s, t]) \), as \( k_1, \ldots, k_\ell \) go to infinity. In addition, \( X'(\Box [s, t]) \) satisfies
\[ |X'(\Box [s, t])| \leq C^\ell Z \int_0^{d_1(s_1, t_1)} du_1 \cdots \int_0^{d_\ell(s_\ell, t_\ell)} du_\ell \]
\[ \times \left( \frac{1}{\mu(B(s, u))} + \log^{1/2} \frac{1}{\mu(B(t, u))} \right). \]

**Proof.** Fix \( s, t \) in \( T \). We choose the multi-index \( n \) such that \( D_j 2^{-n_j - 1} \leq d_j(s_j, t_j) \leq D_j 2^{-n_j} \) for each \( j = 1, \ldots, \ell \). It suffices to show that the following series satisfies the bound in [5.2]:
\[ |M_n(\Box [s, t])| \leq \sum_{k \geq n} |M_{k+1}(\Box [s, t]) - M_k(\Box [s, t])|. \]
We estimate the first term. Notice that we can write
\[ M_n(\square[s,t]) = \iint_{T \times T} X(\square[u,v]) \rho_n(s,u) \rho_n(t,v) \mu(du) \mu(dv). \]

We consider the function \( \{Y(u,v), (u,v) \in T \times T\} \) defined by
\[ Y(u,v) = \begin{cases} \frac{X(\square[u,v])}{d^*(u,v)} & \text{when } d^*(u,v) \neq 0, \\ 0 & \text{otherwise}. \end{cases} \]

It is clear that
\[ |M_n(\square[s,t])| \leq \iint_{T \times T} |Y(u,v)| d^*(u,v) \rho_n(s,u) \rho_n(t,v) \mu(du) \mu(dv) \]
\[ \lesssim (D_1 2^{-n_1}) \cdots (D_\ell 2^{-n_\ell}) \iint_{T \times T} |Y(u,v)| \rho_n(s,u) \rho_n(t,v) \mu(du) \mu(dv), \]
since the support of \( \rho_n(s,\cdot) \rho_n(t,\cdot) \), \( d^*(u,v) \leq (D_1 2^{-n_1}) \cdots (D_\ell 2^{-n_\ell}) \). We now apply Lemma 5.1 to the functions \( Y \) and \( \rho_n(s,\cdot) \otimes \rho_n(t,\cdot) \) on the product space \( (T \times T, \mu \otimes \mu) \), observing that \( Z = \{Y\} \), \( \int \rho_n(s,\cdot) \rho_n(t,\cdot) = 1 \) and \( \rho_n(s,u) \rho_n(t,v) \leq (\mu_n(s,u) \mu_n(t,v))^{-1}, \)
\[ |M_n(\square[s,t])| \lesssim Z(D_1 2^{-n_1}) \cdots (D_\ell 2^{-n_\ell}) \iint_{T \times T} \rho_n(s,u) \rho_n(t,v) \log^{1/2} \]
\[ \times (1 + \rho_n(s,u) \rho_n(t,v)) \mu(du) \mu(dv) \]
\[ \lesssim Z(D_1 2^{-n_1}) \cdots (D_\ell 2^{-n_\ell}) \log^{1/2} \left( 1 + \frac{1}{\mu_n(s) \mu_n(t)} \right). \]

Since \( d^*(s,t) \leq (D_1 2^{-n_1}) \cdots (D_\ell 2^{-n_\ell}) \), this shows that
\[ |M_n(\square[s,t])| \]
(5.4) \[ \lesssim Z \int_0^{d_1(s_1,t_1)} du_1 \cdots \int_0^{d_\ell(s_\ell,t_\ell)} du_\ell \log^{1/2} \left( \frac{1}{\mu(B(s,u))} + \frac{1}{\mu(B(t,u))} \right). \]

We now estimate each term in the sum appearing in (5.3). We denote \( \tau_0 k = k \) and recursively \( \tau_j k = \tau_{j-1} k + 1_j \) for each \( j = 1, \ldots, \ell \). For example, \( \tau_1 k = (k_1 + 1, k_2, \ldots, k_\ell) \) and \( \tau_\ell k = k + 1 \). We then write
\[ |M_{k+1}(\square[s,t]) - M_k(\square[s,t])| \leq \sum_{j=1}^\ell |M_{\tau_j k}(\square[s,t]) - M_{\tau_{j-1} k}(\square[s,t])|. \]

Note that the multi-indices \( \tau_j k \) and \( \tau_{j-1} k \) differ by exactly 1 unit at the \( j \)-th component. Without loss of generality, we consider the case
\[ |M_{\tilde{k}}(\square[s,t]) - M_k(\square[s,t])|, \]
where \( \tilde{k} = k + 1_\ell = (k_1, \ldots, k_{\ell-1}, k_\ell + 1) \). We adopt the notation \( w = (w', w_\ell) \) for every \( w \) in \( T \),
\[ \rho'_k(s',u') = \rho_{k_1}(s_1,u_1) \cdots \rho_{k_{\ell-1}}(s_{\ell-1},u_{\ell-1}) \]
and similarly for \( \rho'_k(s',v') \). We then write
\[ M_k(\square[s,t]) = M_k(\square_{\ell-1}[s',t'], s_\ell) - M_k(\square_{\ell-1}[s',t'], t_\ell) \]
and similarly for $M_k(\Box[s,t])$. Thus
\[
|M_k(\Box[s,t]) - M_k(\Box[s,t])| \\
\leq |M_{k+1}(\Box^{\ell-1}[s',t'],s_{\ell}) - M_k(\Box^{\ell-1}[s',t'],s_{\ell})| \\
+ |M_{k+1}(\Box^{\ell-1}[s',t'],t_{\ell}) - M_k(\Box^{\ell-1}[s',t'],t_{\ell})| \\
= I_1 + I_2.
\]
(5.6)

We only need to estimate $I_1$ since $I_2$ is analogous. We have
\[
M_k(\Box^{\ell-1}[s',t'],s_{\ell}) \\
= \int T \times T X(\Box^{\ell-1}[u',v'],v_{\ell})\rho_k(s',u')\rho_k(t',v')\rho_{\ell+1}(s_{\ell},u_{\ell})\rho_{\ell}(s_{\ell},v_{\ell})\mu(du)\mu(dv)
\]
and similarly
\[
M_k(\Box^{\ell-1}[s',t'],s_{\ell}) \\
= \int T \times T X(\Box^{\ell-1}[u',v'],u_{\ell})\rho_k(s',u')\rho_k(t',v')\rho_{\ell+1}(s_{\ell},u_{\ell})\rho_{\ell}(s_{\ell},v_{\ell})\mu(du)\mu(dv).
\]

Note how the dummy variables $v_{\ell}$ and $u_{\ell}$ have been switched between the two formulas. Hence
\[
|M_k(\Box^{\ell-1}[s',t'],s_{\ell}) - M_k(\Box^{\ell-1}[s',t'],s_{\ell})| \\
\leq \int T \times T |X(\Box^{\ell}[u,v])|\rho_k(s',u')\rho_k(t',v')\rho_{\ell+1}(s_{\ell},u_{\ell})\rho_{\ell}(s_{\ell},v_{\ell})\mu(du)\mu(dv).
\]

Similarly to the term $M_n(\Box[s,t])$ one can obtain
\[
|M_k(\Box^{\ell-1}[s',t'],s_{\ell})| \\
\preceq Z(D_12^{-k_1}) \cdots (D_\ell 2^{-k_\ell}) \log^{1/2} \left(1 + \frac{1}{\mu_k(s)}\right)
\preceq Z(D_12^{-k_1}) \cdots (D_\ell 2^{-k_\ell}) \log^{1/2} \left(1 + \frac{1}{\mu_k(s)}\right).
\]

Therefore, combining together (5.5), (5.6) and the previous estimate, we get
\[
|M_{k+1}(\Box[s,t]) - M_k(\Box[s,t])| \preceq Z\ell(D_1 2^{-k_1}) \cdots (D_\ell 2^{-k_\ell}) \log^{1/2} \left(1 + \frac{1}{\mu_k(s)}\right),
\]
and hence,
\[
\sum_{k \geq n} |M_{k+1}(\Box[s,t]) - M_k(\Box[s,t])| \\
\preceq Z\ell(D_1 2^{-n_1}) \cdots (D_\ell 2^{-n_\ell}) \log^{1/2} \left(1 + \frac{1}{\mu_n(s)}\right)
\preceq Z\ell \int_0^{d_1(s_1,t_1)} du_1 \cdots \int_0^{d_\ell(s_\ell,t_\ell)} du_\ell \log^{1/2} \left(\frac{1}{\mu(B(s,u))} + \frac{1}{\mu(B(t,u))}\right).
\]

Together with the bound for $M_n(\Box[s,t])$ (inequality (5.4)) and (5.3), this completes the proof. □
Remark 5.3. In Theorem 5.2, $X'$ may not be defined as a function on $T$; that is, for each $t$ in $T$, there is no a priori reason for $X'(t)$ to be defined. However, in order to keep the representation compact, we have abused notation and denote the limit as $X'([s, t])$. This object is well-defined for every fixed $s, t$ in $T$.

5.2. Majorizing measure. We now suppose that $X$ is a stochastic process with the probability space $(\Omega, \mathcal{F}, P)$. We introduce the $\ell$-fold volumetric

$$d^\ell(s, t) = (\mathbb{E}[X([s, t])]^2)^{1/2}.$$ 

Assume that $\sup_{s, t \in T} d^\ell(s, t)$ is finite. In addition, for each $i$, there exists a metric $d_i$ on $T_i$ such that

$$d^\ell(s, t) \leq d_1(s_1, t_1) \cdots d_\ell(s_\ell, t_\ell).$$

This is not a restriction since such a collection of metrics always exists. For instance, one can choose

$$d_1(s_1, t_1) = \sup_{\hat{s}_1, \hat{t}_1 \in T_1} d^\ell(s, t)$$

and recursively

$$d_k(s_k, t_k) = \sup_{\hat{s}_k, \hat{t}_k \in T_k} \frac{d^\ell(s, t)}{\prod_{i=1}^{k-1} d_i(s_i, t_i)}$$

with the convention $0/0 = 0$.

We denote $Z$ as in Theorem 5.2 that is,

$$Z = \inf \left\{ \alpha > 0 : \int_{T \times T} \Psi \left( \frac{X([u, v])}{\alpha d^\ell(u, v)} \right) \mu(du) \mu(dv) \leq 1 \right\}. \tag{5.7}$$

We assume that $Z$ is finite almost surely.

Example 5.4. Suppose $X$ is a centered Gaussian process. Then $Z$ has exponential tail. More precisely $P(Z > u) \leq (e \log 2)^{1/2} u^{2^{-\alpha}}$ for all $u > (2 + 1/\log 2)^{1/2}$. This comes from a standard argument by Chebyshev inequality and Hölder inequality; see [41, pp. 256-258] for details.

As an application of Theorem 5.2, we have

**Theorem 5.5.** Let $\{X(t), t \in T\}$ be a stochastic process such that $Z$, defined in (5.7), is finite a.s. Then $X$ has a version $X'$ such that for all $\omega \in \Omega$ and $s, t$ in $T$,

$$|X'(\omega, [s, t])| \leq C\ell Z(\omega) \int_0^{d_1(s_1, t_1)} du_1 \cdots \int_0^{d_\ell(s_\ell, t_\ell)} du_\ell \times \left( \log^{1/2} \frac{1}{\mu(B(s, u))} + \log^{1/2} \frac{1}{\mu(B(t, u))} \right).$$

In particular, if $\mathbb{E}Z$ is finite, then

$$\mathbb{E} \sup_{d_i(s_i, t_i) \leq \delta_i, 1 \leq i \leq \ell} |X([s, t])| \leq C\ell \mathbb{E}Z \sup_{s \in T} \int_0^{\delta_1} du_1 \cdots \int_0^{\delta_\ell} du_\ell \log^{1/2} \frac{1}{\mu(B(s, u))}. \tag{5.8}$$

**Proof.** First note that for every $t, v$ in $T$,

$$\left( \mathbb{E}[X(t) - X(v)]^2 \right)^{1/2} \leq \sum_{i=1}^\ell d_i(t_i, v_i).$$
We recall the notation $M_k(t)$ in (5.1). We have
\[
\mathbb{E}[X(t) - M_k(t)] \leq \int_{\mathbb{T}} \mathbb{E}[X(t) - X(v)]\rho_k(t, v)\mu(dv)
\]
\[
\leq \int_{\mathbb{T}} \sum_{i=1}^{\ell} d_i(t_i, v_i)\rho_k(t, v)\mu(dv) \leq \sum_{i=1}^{\ell} D_i 2^{-k_i}.
\]
Together with the Borel-Cantelli lemma, this shows that for all $t \in \mathbb{T}$, $M_k(t)$ converges to $X(t)$ almost surely. On the other hand, Theorem 5.2 shows for all $s, t \in \mathbb{T}$ that $M_k([s, t])$ converges to a limit, denoted by $X'(][s, t])$. This implies $X([s, t]) = X'(][s, t])$ almost surely. The result now follows from Theorem 5.2. □

5.3. Asymptotic growth. Let $W(t, x)$ be a continuous Gaussian process on $[0, T] \times \mathbb{R}^d$ with mean 0. As in the previous subsection, we define the $d$-fold volumetric
\[
d(x, y) = \sup_{t \in [0, T]} (\mathbb{E} [W(t, [x, y])]^2)^{1/2}.
\]
Without loss of generality, we assume there are metrics $d_1, \ldots, d_d$ on $\mathbb{R}$ such that $d^*(x, y) = d_1(x_1, y_1) \ldots d_d(x_d, y_d)$ satisfies $d(x, y) \leq d^*(x, y)$.

Let $\delta = (\delta_1, \ldots, \delta_\ell)$ be in $(0, \infty)^\ell$. The notation $d^*(x, y) \leq \delta$ means $d_i(x_i, y_i) \leq \delta_i$ for all $i = 1, 2, \ldots, \ell$. We denote $|x|^* = \max_{1 \leq i \leq d} d_i(0, x_i)$ for every $x \in \mathbb{R}^d$. We are interested in the asymptotic growth of the process
\[
W^*(\delta, R) = \sup_{t \in [0, T]} \sup_{d^*(x, y) \leq \delta} \sup_{|x|^*, |y|^* \leq R} |W(t, [x, y])|
\]
as $R$ gets large and $\delta$ can range freely in a bounded neighborhood of 0. $W^*$ also depends on $T$. However since $T$ will always be fixed in our consideration, we suppress the dependence on $T$ in our notation. We put
\[
\mathcal{S}_R = \{x \in \mathbb{R}^d : |x|^* \leq R\},
\]
\[
m(\delta, R) = \mathbb{E} W^*(\delta, R),
\]
and
\[
\sigma(\delta, R) = \sup_{t \leq [0, T]} \sup_{d^*(x, y) \leq \delta} \sup_{x, y \in \mathcal{S}_R} (\mathbb{E} [W(t, [x, y])]^2)^{1/2}.
\]
We first prove the following concentration inequality.

**Lemma 5.6.** For any $r > 0$,
\[
P \left( \frac{1}{\sigma(\delta, R)} |W^*(\delta, R) - m(\delta, R)| > r \right) \leq 2e^{-r^2/2}.
\]
As a consequence,
\[
\mathbb{E} \psi_\rho \left( \frac{|W^*(\delta, R) - m(\delta, R)|}{\sigma(\delta, R)} \right) \leq c_\rho < \infty
\]
for every $\rho < 1/2$, where $\psi_\rho = \exp(\rho x^2)$. 

...
Proof. It suffices to show (5.9). Let \( \{X(u), u \in T\} \) be a Gaussian process. Assume that \( T \) is finite. The following concentration inequality is standard:

\[
\Pr \left( \frac{1}{\sigma} \left| \sup_{u \in T} |X(u)| - \mathbb{E} \left[ \sup_{u \in T} |X(u)| \right] \right| > r \right) \leq 2e^{-r^2/2},
\]

for every \( \sigma \geq \sup_{u \in T} (EX^2(u))^{1/2} \). We refer to [38] or [41, Theorem 5.4.3] for a proof of (5.11). We now fix \((t_1, x_1), \ldots, (t_m, x_m)\) in \([0, T] \times \mathbb{R}^d \) such that \(d^*(x_j, x_k) \leq \delta \) and \(|x_j|^*, |x_k|^* \leq R \) for all \( j, k \). We denote by \( x_j \cup x_k \) the collection of points \( z \in \mathbb{R}^d \) such that each component of \( z \) is the corresponding component of either \( x_j \) or \( x_k \). We consider the centered Gaussian random process \( X(t_i, x_j \cup x_k) := W(t_i, x_j \cup x_k) \) indexed by the parameters \( \{t_i\}_{1 \leq i \leq m} \) and \( \{x_j \cup x_k\}_{1 \leq j, k \leq m} \). It is clear that

\[\mathbb{E}X^2(t_i, x_j \cup x_k) \leq \sigma^2(\delta, R).\]

Thus, the inequality (5.11) becomes

\[
P \left( \frac{1}{\sigma(\delta, R)} \sup_{i,j,k \leq m} |W(t_i, \square [x_j, x_k])| - \mathbb{E} \left[ \sup_{i,j,k \leq m} |W(t_i, \square [x_j, x_k])| \right] > r \right) \leq 2e^{-r^2/2}.
\]

An approximation procedure yields (5.9). \( \square \)

**Theorem 5.7.** With probability one,

\[
\sup_{\delta \in (0,1]^\ell} \limsup_{R \to \infty} \frac{|W^*(\delta, R) - m(2\delta, 2R)|}{\sigma(2\delta, 2R) \sqrt{\log(\delta_1^{-1} \cdots \delta_\ell^{-1} \log R)}} \leq \sqrt{2}.
\]

Proof. We put \( p(\delta, R) = \delta_1^{-1} \cdots \delta_\ell^{-1} (\log R)^2 \) and consider the random variable

\[
\Theta = \sup_{\delta \in (0,1]^\ell, R \geq 1} \frac{1}{p(\delta, R)} \psi_\rho \left( \frac{1}{\sigma(2\delta, 2R)} |W^*(\delta, R) - m(2\delta, 2R)| \right).
\]

For each multi-index \( j = (j_1, \ldots, j_\ell) \) in \( \mathbb{N}^\ell \), we denote \( 2^{-j} = (2^{-j_1}, \ldots, 2^{-j_\ell}) \). The notation \( \delta \leq 2^{-j} \) means \( \delta_i \leq 2^{-j_i} \) for all \( i = 1, 2, \ldots, \ell \). Then using the monotonicity of \( p, \psi_\rho, W^* \) and \( \sigma \), and (5.10) we have

\[
\mathbb{E}\Theta \leq \sum_{k \in \mathbb{N}, j \in \mathbb{N}^\ell} \mathbb{E} \sup_{2^{-j_1-1} \leq \delta \leq 2^{-j}} \frac{1}{p(\delta, R)} \psi_\rho \left( \frac{1}{\sigma(2\delta, 2R)} |W^*(\delta, R) - m(2\delta, 2R)| \right)
\]

\[
\leq \sum_{k \in \mathbb{N}, j \in \mathbb{N}^\ell} \frac{1}{p(2^{-j} - 2^{-k-1}, 2^{-k})} \mathbb{E}\psi_\rho \left( \frac{1}{\sigma(2^{-j}, 2^k)} |W^*(2^{-j}, 2^k) - m(2^{-j}, 2^k)| \right)
\]

\[
\leq c_\rho \sum_{k \in \mathbb{N}, j \in \mathbb{N}^\ell} \frac{1}{p(2^{-j} - 2^{-k-1})} < \infty.
\]

Hence, with probability one, \( \Theta \) is finite and

\[
\psi_\rho \left( \frac{1}{\sigma(2\delta, 2R)} |W^*(\delta, R) - m(2\delta, 2R)| \right) \leq \Theta p(2\delta, R), \quad \forall \delta > 0, \forall R \geq 1.
\]

In particular,

\[
\frac{|W^*(\delta, R) - m(2\delta, 2R)|}{\sigma(2\delta, 2R)} \leq \sqrt{\frac{\log[\Theta p(2\delta, R)]}{\rho}}, \quad \forall \delta > 0, \forall R \geq 1.
\]
We then use the trivial estimate
\[ \sqrt{\log(\Theta p)} \leq \sqrt{\log \Theta} + \sqrt{\log p} \]
to get
\[ \frac{|W^*(\delta, R) - m(2\delta, 2R)|}{\sigma(2\delta, 2R) \sqrt{\log(\delta_1^{-1} \cdots \delta_\ell^{-1} \log R)}} \leq \sqrt{\frac{\log \Theta}{\rho |\log \log R|}} + \sqrt{\frac{\log(\delta_1^{-1} \cdots \delta_\ell^{-1} \log R)^2}{\rho \log(\delta_1^{-1} \cdots \delta_\ell^{-1} \log R)}}. \]
for all \( \delta > 0 \) and \( R \geq 1 \). Since \( \rho \) can be chosen to be any constant less than \( 1/2 \), we can choose a sequence \( \rho_n \) convergent to \( 1/2 \). Since countable unions of events with probability zero still have probability zero, we can pass through the limit \( n \to \infty \) to get, with probability one,
\[ \frac{|W^*(\delta, R) - m(2\delta, 2R)|}{\sigma(2\delta, 2R) \sqrt{\log(\delta_1^{-1} \cdots \delta_\ell^{-1} \log R)}} \leq \sqrt{\frac{2 \log \Theta}{\log \log R}} + \sqrt{\frac{2 \log(\delta_1^{-1} \cdots \delta_\ell^{-1} \log R)^2}{\log(\delta_1^{-1} \cdots \delta_\ell^{-1} \log R)}}, \]
for all \( \delta > 0 \) and \( R \geq 1 \). Finally, let \( R \to \infty \) to complete the proof. \( \square \)

In general, it is hard to say anything about the growth of \( m(\delta, R) \) as \( R \) gets large. In what follows, we restrict ourselves to a particular (but still sufficiently large) class of Gaussian random fields. To be more precise, for each \( i = 1, \ldots, \ell \), let \( \phi_i \) be a majorant for \( d_i \), that is, \( \phi_i \) is strictly increasing with \( \phi_i(0) = 0 \) and
\[ d_i(x_i, y_i) \leq \phi_i(|y_i - x_i|). \]
Define
\[ \tilde{\omega}_i(\delta_i) = \delta_i \log \frac{1}{\phi_i^{-1}(\delta_i)} + \int_0^{\phi_i^{-1}(\delta_i)} \frac{\phi_i(u)}{u \log^{1/2}(1/u)} \, du. \]
We will always presume \( \tilde{\omega}_i \)'s are finite wherever they appear.

**Proposition 5.8.** Denote \( \tilde{\delta}_i = \prod_{j \neq i} \delta_j \). Then we have
\[ m(\delta, R) \leq \delta_1 \cdots \delta_\ell \log^{1/2} \left( \prod_{i=1}^\ell 2\phi_i^{-1}(R) \right) + \sum_{i=1}^\ell \tilde{\delta}_i \tilde{\omega}_i(\delta_i) \]
where the implied constant is independent of \( R \) and \( \delta \).

**Proof.** We take for the majorizing measure \( \mu_i = \lambda/(2\phi_i^{-1}(R)) \), where \( \lambda \) is the Lebesgue measure. By (5.13), the ball \( B^i(x_i, u_i) \) contains the interval \( (x_i - \phi_i^{-1}(u_i), x_i + \phi_i^{-1}(u_i)) \cap \{ z_i : d_i(z_i, 0) \leq R \} \), thus,
\[ \mu_i(B^i(x_i, u_i)) \geq \frac{\phi_i^{-1}(u_i)}{2\phi_i^{-1}(R)}. \]
Hence, for all \( x \) in \( \mathcal{S}_R \),
\[ \log \frac{1}{\mu(B(x, u))} \leq \log \left( \prod_{i=1}^d 2\phi_i^{-1}(R) \phi_i^{-1}(u_i) \right). \]
Therefore, for $\delta$ sufficiently small,
\[
\int_0^{\delta_1} du_1 \cdots \int_0^{\delta_\ell} du_\ell \log^{1/2} \frac{1}{\mu(B(x, u))} \leq \int_0^{\delta_1} du_1 \cdots \int_0^{\delta_\ell} du_\ell \log^{1/2} \left( \prod_{i=1}^d 2\phi_i^{-1}(R) \phi_i^{-1}(u_i) \right) \leq \delta_1 \cdots \delta_\ell \log^{1/2} \left( \prod_{i=1}^\ell 2\phi_i^{-1}(R) \right) + \int_0^{\delta_1} du_1 \cdots \int_0^{\delta_\ell} du_\ell \log^{1/2} \left( \prod_{i=1}^d \frac{1}{\phi_i^{-1}(u_i)} \right).
\]
The last integral in the above formula can be estimated as
\[
\int_0^{\delta_1} du_1 \cdots \int_0^{\delta_\ell} du_\ell \log^{1/2} \left( \prod_{i=1}^d \frac{1}{\phi_i^{-1}(u_i)} \right) = \int_0^{\phi_i^{-1}(\delta_1)} d\phi_1(u_1) \cdots \int_0^{\phi_\ell^{-1}(\delta_\ell)} d\phi_\ell(u_\ell) \log^{1/2} \left( \prod_{i=1}^d \frac{1}{u_i} \right) \leq \sum_{i=1}^\ell \int_0^{\phi_i^{-1}(\delta_i)} \log^{1/2} (1/u_i) d\phi_i(u_i).
\]
Using integration by parts, $\int_0^{\phi_i^{-1}(\delta_i)} \log^{1/2} (1/u_i) d\phi_i(u_i) \leq \tilde{\omega}(\delta_i)$, which completes the proof.

Example 5.9. Let $W = (W(x), x \in \mathbb{R}^d)$ be a factional Brownian sheet with Hurst parameter $H = (H_1, \ldots, H_d) \in (0, 1)^d$. In particular, the covariance of $W$ is given by
\[
\mathbb{E} W(x) W(y) = \prod_{i=1}^d R_{H_i}(x_i, y_i)
\]
where
\[
R_{H_i}(s, t) = \frac{1}{2} |s|^{2H_i} + |t|^{2H_i} - |s - t|^{2H_i}.
\]
We see that
\[
(\mathbb{E} |W(\Box[x, y])|^2)^{1/2} = \prod_{i=1}^d |x_i - y_i|^{H_i},
\]
thus $\phi_i(\delta) = |\delta_i|^{H_i}$ and $\sigma(\delta, R) = \delta_1 \cdots \delta_d$. We put
\[
m(\delta, R) = \mathbb{E} \sup |W(\Box[x, y])|.
\]
where the supremum is taken over the domain $\{x, y : |x_i|^{H_i}, |y_i|^{H_i} \leq R$ and $|x_i - y_i|^{H_i} \leq \delta_i, \forall 1 \leq i \leq d\}$. Note that
\[
\tilde{\omega}_i(\delta) \leq \delta_i \log^{1/2} \frac{1}{\phi_i^{-1}(\delta_i)}.
\]
The bound (5.14) yields
\[
m(\delta, R) \lesssim \delta_1 \cdots \delta_d \sqrt{\log(R\delta_1^{-1} \cdots \delta_d^{-1})}.
\]
Theorem 5.5 yields

\begin{equation}
\sup_{|x|^{H_{1}},|y|^{H_{1}} \leq R: d^*(x,y) \leq \delta} |W(\square[x,y])| \lesssim \delta_1 \cdots \delta_d \sqrt{\log(R\delta_1^{-1} \cdots \delta_d^{-1})},
\end{equation}

when \( R \) gets large. This implies the inequality of the form (2.1) for \( W \).

Remark 5.10. A fractional Brownian sheet belongs to a larger class of random fields called anisotropic random fields. That is,

\begin{equation}
\left( \mathbb{E} |W(y) - W(x)|^2 \right)^{1/2} \asymp \sum_i |y_i - x_i|^{H_i}.
\end{equation}

These random fields may have different behavior along different directions. In [42], the authors investigate the global moduli of continuity for anisotropic Gaussian random fields. As a result, they establish a sharp result for the global modulus of continuity for fractional Brownian sheets. The conditions considered in the current paper are somewhat more general. For instance, the estimate (5.8) implies the upper bound in the anisotropic condition (5.16). We believe our method (Theorems 5.2, 5.5) provides similar results as [42] though we do not report them here.

**Appendix A. Other types of nonlinear stochastic integrals**

The Itô integral is a fundamental concept in stochastic analysis. This integral can be defined under fewer conditions than the Stratonovich one and has a completely different feature such as the famous Itô formula. From the modeling point of view, Itô type stochastic differential equations are more popular since all terms in the Itô equation \( dx_t = b(x_t)dt + \sigma(x_t)\delta B_t \) (see also (4.6)) have a clear meaning: \( b(x_t) \) represents the mean rate of change, and \( \sigma(x_t)\delta B_t \) represents the fluctuation (it has zero mean contribution).

In this section, we will introduce the nonlinear Itô-Skorohod integral. This integral is a probabilistic one and is defined for almost every sample path, while the nonlinear Young integral is defined for every sample path. The relation between these two integrals is through the nonlinear symmetric (Stratonovich) integral.

**A.1. Nonlinear Itô-Skorohod integral.** Let \( H \in (\frac{1}{2},1) \) and denote by \( R_H(s,t) = \frac{1}{2} (s^{2H} + t^{2H} - |s-t|^{2H}) \) the covariance function of a fractional Brownian motion of Hurst parameter \( H \). Let \( q(x,y) \) be a continuous and positive definite function; namely, for any \( x_i \in \mathbb{R}^d, i = 1,2,\ldots,m \) and complex numbers \( \xi_i, i = 1,2,\ldots,m, \) not all 0, we have

\[ \sum_{i,j=1}^m q(x_i,x_j)\bar{\xi}_i\xi_j \geq 0, \]

where \( \bar{\xi}_i \) is the conjugate number of \( \xi_i \). For every \( s,t \geq 0 \) and \( x,y \in \mathbb{R}^d \), we denote

\[ Q(s,t,x,y) = \frac{\partial^2 R_H}{\partial s \partial t}(s,t)q(x,y) = \alpha_H |s-t|^{2H-2}q(x,y), \]

where \( \alpha_H = H(2H-1) \). Let \( \mathcal{S} \) be the set of all smooth functions \( f : [0,T] \times \mathbb{R}^d \to \mathbb{R} \) such that \( f(t,\cdot) \) has compact support for every \( t \in [0,T] \). We introduce a scalar product on \( \mathcal{S} \) in the following way:

\[ \langle \phi, \psi \rangle_{\mathcal{H}} = \int_{[0,T]^2 \times \mathbb{R}^d} \phi(s,x)\psi(t,y)Q(s,t,x,y)dxdydsdt. \]
We denote by $\mathcal{H}$ the Hilbert space of the closure of $S$ with respect to this inner product. Let $T$ be a bijective Hilbert-Schmidt operator on $\mathcal{H}$. Define the Banach space (in fact, it is a Hilbert space) $\Omega$ as the completion of $\mathcal{H}$ with respect to the norm $\|x\|_\Omega := \sqrt{\langle Tx, Tx \rangle_\mathcal{H}}$. Then, it follows from the Bochner-Minlos theorem (see [26], Theorem 3.1) that there is a probability measure $P$ on $(\Omega, \mathcal{F})$ such that $\langle h, \omega \rangle$ is a centered Gaussian random variable with covariance $E[(h, \cdot)(h', \cdot)] = \langle h, h' \rangle_\mathcal{H}$, $\forall h, h' \in \Omega$, where $\Omega$ is the Banach space of all continuous linear functionals on $\Omega$, $\mathcal{F}$ is the Borel $\sigma$-algebra generated by the open sets of $\Omega$, and $\langle h, \omega \rangle$ the pairing between $h \in \Omega' \subset \mathcal{H}$ and $\Omega$. We identify $\Omega' = \mathcal{H}$ so that the embeddings $\Omega' \subset \mathcal{H} = \mathcal{H} \subset \Omega$ are continuous. We can define the Gaussian random variable $\langle h, \omega \rangle$ for all $h \in \mathcal{H}$ by a limiting argument.

First we give some specific elements in $\mathcal{H}$. For any $x \in \mathbb{R}^d$, we denote by $\delta_x$ the Dirac function on $\mathbb{R}^d$. Namely, $\delta_x$ is defined by $\int_{\mathbb{R}^d} \delta_x(y) f(y) dy = f(x)$ for any smooth function of compact support on $\mathbb{R}^d$.

**Proposition A.1.** For any $s > 0$ and $x \in \mathbb{R}^d$, $I_{(0,s]} \delta_x$ is an element in $\mathcal{H}$ and

\[
\langle I_{(0,s]} \delta_x, I_{(0,t]} \delta_y \rangle_\mathcal{H} = R_H(s, t) q(x, y)
\]

and

\[
\| I_{(0,s]} \delta_x - I_{(0,t]} \delta_y \|^2_\mathcal{H} = s^{2H} q(x, x) + t^{2H} q(y, y) - 2R_H(s, t) q(x, y).
\]

**Proof.** For every $\varepsilon > 0$ and $x \in \mathbb{R}^d$, we denote the elementary function

$\delta_x^\varepsilon = (2\varepsilon)^{-d} I_{(x-\varepsilon, x+\varepsilon)}$.

If $\varepsilon$ tends to 0, the function $\delta_x^\varepsilon$ converges in $\mathcal{H}$ to the generalized function $\delta_x$. Indeed, fix $(s, x)$ and $(t, y)$ in $[0, T] \times \mathbb{R}^d$. For any positive number $\varepsilon$ and $\varepsilon'$, we have

$$
\left\langle I_{(0,s]} \delta_x^\varepsilon, I_{(0,t]} \delta_y^{\varepsilon'} \right\rangle_\mathcal{H} = R_H(s, t) (4\varepsilon\varepsilon')^{-d} \int_{y-\varepsilon'}^{y+\varepsilon'} \int_{x-\varepsilon}^{x+\varepsilon} q(x', y') dx' dy'.
$$

Since $q(\cdot, \cdot)$ is continuous, the above right hand side converges to $q(x, y)$ as $\varepsilon$ and $\varepsilon'$ tend to 0. This shows easily that $I_{(0,s]} \delta_x^\varepsilon$ is a Cauchy sequence in $\mathcal{H}$ when $\varepsilon \to 0$. The limit of $I_{(0,s]} \delta_x^\varepsilon$ in $\mathcal{H}$ as $\varepsilon \to 0$ is $I_{(0,s]} \delta_x$. The equations (A.2) and (A.3) are immediate. \hfill $\Box$

Since $I_{(0,s]} \delta_x \in \mathcal{H}$, we can define

\[
W(x, s, \omega) = \langle I_{(0,s]} \delta_x, \omega \rangle, \; \omega \in \Omega.
\]

Thus $\{W(s, x), t \geq 0, x \in \mathbb{R}^d\}$ is a multi-parameter centered Gaussian process with the covariance

$$
E[W(s, x)W(t, y)] = \langle I_{(0,s]} \delta_x, I_{(0,t]} \delta_y \rangle_\mathcal{H} = R_H(s, t) q(x, y).
$$

We also denote

$$
W(\phi) := \int_0^T \int_{\mathbb{R}^d} \phi(s, x) W(ds, dx) := \langle \phi, \omega \rangle \; \forall \phi \in \mathcal{H}.
$$

We denote by $\mathcal{P}$ the set of smooth and cylindrical random variables of the form

\[
F = f(W(\phi_1), \ldots, W(\phi_n)),
\]
\( \phi_i \in \mathcal{H}, f \in C^\infty_0(\mathbb{R}^n) \) (\( f \) and all its partial derivatives have polynomial growth). \( D \) denotes the Malliavin derivative. That is, if \( F \) is of the form \( (A.5) \), then \( DF \) is the \( \mathcal{H} \)-valued random variable defined by

\[
    DF = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(W(\phi_1), \ldots, W(\phi_n)) \phi_j.
\]

The operator \( D \) is closable from \( L^2(\Omega) \) into \( L^2(\Omega; \mathcal{H}) \), and we define the Sobolev space \( \mathbb{D}^{1,2} \) as the closure of \( \mathcal{P} \) under the norm

\[
    \|F\|_{1,2} = \sqrt{\mathbb{E}(F^2) + \mathbb{E}(\|DF\|_{\mathcal{H}}^2)}.
\]

\( D \) can be extended uniquely to an operator from \( \mathbb{D}^{1,2} \) into \( L^2(\Omega; \mathcal{H}) \). The divergence operator \( \delta \) is the adjoint of the Malliavin derivative operator \( D \). We say that a random variable \( u \) in \( L^2(\Omega; \mathcal{H}) \) belongs to the domain of the divergence operator, denoted by \( \text{Dom} \delta \), if there is a constant \( c_u \in (0, \infty) \) such that

\[
    |\mathbb{E}(\langle DF, u \rangle_{\mathcal{H}})| \leq c_u \|F\|_{L^2(\Omega)} \quad \forall \, F \in \mathbb{D}^{1,2}.
\]

In this case \( \delta(u) \) is defined by the duality relationship

\[
    (A.6) \quad \mathbb{E}(\delta(u)F) = \mathbb{E}(\langle DF, u \rangle_{\mathcal{H}}) \quad \forall \, F \in \mathbb{D}^{1,2}.
\]

The following are two basic properties of the divergence operator \( \delta \).

(i) \( \mathbb{D}^{1,2}(\mathcal{H}) \subset \text{Dom} \delta \) and for any \( u \in \mathbb{D}^{1,2}(\mathcal{H}) \),

\[
    (A.7) \quad \mathbb{E}(\delta(u)^2) = \mathbb{E}(\|u\|_{\mathcal{H}}^2) + \mathbb{E}(\langle Du, (Du)^* \rangle_{\mathcal{H} \otimes \mathcal{H}}),
\]

where \( (Du)^* \) is the adjoint of \( Du \) in the Hilbert space \( \mathcal{H} \otimes \mathcal{H} \).

(ii) For any \( F \) in \( \mathbb{D}^{1,2}(\mathcal{H}) \) and any \( u \) in the domain of \( \delta \) such that \( Fu \) and \( F\delta(u) - \langle DF, u \rangle_{\mathcal{H}} \) are square integrable, \( Fu \) is in the domain of \( \delta \) and

\[
    (A.8) \quad \delta(Fu) = F\delta(u) - \langle DF, u \rangle_{\mathcal{H}}.
\]

The operator \( \delta \) is also called the Skorokhod integral because in the case of Brownian motion, it coincides with the generalization of the Itô stochastic integral to anticipating integrands introduced by Skorokhod [17]. On the relation between \( \delta \) and \( D \), we have the identity

\[
    (A.9) \quad D\delta(u) = u + \delta(Du) .
\]

We refer to Nualart’s book [43] for a detailed account of the Malliavin calculus with respect to a Gaussian process. Using the specific definition of our \( \mathcal{H} \), we also denote \( \delta(u) = \int_0^T \int_{\mathbb{R}^d} u(t, x) W(\delta t, x) dx \). In addition, we can write the identity \( (A.7) \) as

\[
    (A.10) \quad \mathbb{E} \left[ \int_0^T \int_{\mathbb{R}^d} u(t, x) W(\delta t, x) dx \right]^2 = \int_{[0,T]^2 \times \mathbb{R}^{2d}} \mathbb{E}[u(t, x)u(s, y)] Q(s, t, x, y) ds dt dx dy
    + \int_{[0,T]^4 \times \mathbb{R}^{4d}} \mathbb{E}[D_{t_2, x_2} u(t_1, x_1) D_{s_2, y_2} u(s_1, y_1)]
    \times Q(t_1, s_2, x_2, y_1) Q(t_2, s_1, x_1, y_2) ds dt dx dy ,
\]
where in the rest of the paper we shall use \( ds = ds_1 \cdots ds_k \), \( dx = dx_1 \cdots dx_m \) and so on, the \( k \) and \( m \) being clear in the context.

Let \( \{W(t, x), t \geq 0, x \in \mathbb{R}^d\} \) be the Gaussian field introduced in Section A.1 whose mean is 0 and whose covariance is

\[
\mathbb{E}(W(s, x)W(t, y)) = R_H(s, t)q(x, y).
\]

Let \( \varphi = \{\varphi_t, t \in [0, T]\} \) be an \( \mathbb{R}^d \)-valued stochastic process. Our aim in this section is to introduce and study the nonlinear stochastic integral

\[
\int_0^T W(\delta t, \varphi_t).
\]

This stochastic integral was studied earlier in order to establish the Feynman-Kac formula when \( \varphi_t \) is a Brownian motion, independent of \( W \). The case \( H > 1/2 \) is discussed in [35] and the case \( H < 1/2 \) is discussed in [32]. When \( \{W(t, x), t \geq 0\} \) is a semimartingale with respect to \( t \) (for fixed \( x \in \mathbb{R}^d \)), this type of stochastic integral has been studied extensively and generalized Itô formulas have been established. It has been applied to solve some stochastic partial differential equations. See for instance Kunita’s book [37] and the references therein.

In this section, we will define the stochastic integral \( \int W(\delta t, \varphi_t) \) based on the covariance structure of \( W \). This method is closely tied to the nature of \( W \) as a Gaussian process. In particular, we introduce here two types of stochastic integrals, namely, the divergence type and symmetric type. We also study their properties and relation. The divergence type integral turns out to have zero mean; thus one can think of it as a generalization of the Itô-Skorohod integral. The symmetric integral does not have vanishing mean and differs from the divergence type integral by a correction term, related to the Malliavin derivative of some random variable. One can also view the symmetric integral as a generalization of the Stratonovich integral.

We shall define the (nonlinear) Itô-Skorohod (divergence) type integral \( \int_0^T W(\delta t, \varphi_t) \) by the (linear) multi-parameter integral \( \int_0^T \int_{\mathbb{R}^d} \delta(\varphi_t - y)W(\delta t, y)dy \). Here and in the remaining part of the paper, the symbol \( \delta \) carries two meanings: the Itô-Skorohod integral and the Dirac delta function. The difference between the two meanings will be clear from the context.

Since \( \delta(\varphi_t - y) \) is a distribution valued random process, to define its stochastic integral we need to approximate the Dirac delta function \( \delta \). To define such sequence \( \eta \), we denote by \( \eta \) the bump function

\[
\eta(x) = c_d \exp\{(|x|^2 - 1)^{-1}\}1_{\{|x|<1\}}, x \in \mathbb{R}^d,
\]

where \( |x| \) is the Euclidean distance in \( \mathbb{R}^d \) and \( c_d \) is the positive constant so that

\[
\int_{\mathbb{R}^d} \eta(x)dx = 1.
\]

The function \( \eta \) is smooth and compactly supported. Its corresponding mollifier is

\[
\eta_\varepsilon(x) = \varepsilon^{-d} \eta\left(\frac{x}{\varepsilon}\right), \quad (A.11)
\]
Here is our definition.

**Definition A.2.** Let $\varphi : [0, T] \times \Omega \to \mathbb{R}^d$ be a measurable stochastic process. If $I_\varepsilon = \int_0^T \int_{\mathbb{R}^d} \eta_e(\varphi_t - y) W(\delta t, y) dy$ is well-defined and it has a limit in $L^2(\Omega, \mathcal{F}, P)$ as $\varepsilon \to 0$, then we define $\int_0^T W(\delta t, \varphi_t)$ as the aforementioned limit.

Next, we shall give conditions to ensure the existence of the stochastic integral $\int_0^T W(\delta t, \varphi_t)$, namely, to ensure the existence of the limit of $I_\varepsilon$ in $L^2(\Omega, \mathcal{F}, P)$. To express the conditions in a more concise way we introduce the following notation:

$$q_\varphi(x, y) = \alpha \int_0^T \int_0^T \mathbb{E}q(x + \varphi_s, y + \varphi_t)|s - t|^{2H-2} dsdt$$

and

$$q_{D\varphi}^*(x, y) = \alpha^2_H \int_{[0, T]^4 \times \mathbb{R}^{2d}} \mathbb{E}D_{s_1, x'} q(x + \varphi_{s_2}, y') D_{t_2, y'} q(x', y + \varphi_{t_1})$$

$$\times |s_1 - t_1|^{2H-2}|s_2 - t_2|^{2H-2} ds_1 ds_2 dt_1 dt_2 dx'dy'$$

whenever the integrals on the right hand side make sense. We make the following assumptions on the process $\varphi_t$.

(A1) $\varphi_t$ belongs to $\mathbb{D}^{1,2}$ for all $t$, and for almost every $\omega \in \Omega$, the sample path $\varphi_t$ is continuous in $t \in [0, T]$.

(A2) $|q|_\varphi$ is integrable on a neighborhood of $(0, 0)$; that is, there exists an open set $U$ in $\mathbb{R}^{2d}$ containing $(0, 0)$ such that

$$\int_U \int_0^T \int_0^T \mathbb{E}|Q(s, t, x + \varphi_s, y + \varphi_t)| dsdt dx dy < \infty.$$

(A3) $q_\varphi(x, y)$ is well-defined in the neighborhood of $(0, 0)$ and it is continuous at $(0, 0)$.

(A4) There exists an open set $U$ in $\mathbb{R}^{2d}$ containing $(0, 0)$ such that

$$\int_U \int_0^T \int_0^T \mathbb{E}|D_{s_1, x'} q(x + \varphi_{s_2}, y') D_{t_2, y'} q(x', y + \varphi_{t_1})|$$

$$\times |s_1 - t_1|^{2H-2}|s_2 - t_2|^{2H-2} ds_1 ds_2 dt_1 dt_2 dx'dy' dx dy < \infty .$$

(A5) $q_{D\varphi}^*(x, y)$ is well-defined in the neighborhood of $(0, 0)$ and it is continuous at $(0, 0)$.

**Theorem A.3.** We assume the conditions (A1) (A5) are satisfied. Then $\int_0^T W(\delta t, \varphi_t)$ is well-defined and

$$\mathbb{E} \left[ \int_0^T W(\delta t, \varphi_t) \right]^2 = q_{D\varphi}^*(0, 0) + q_\varphi(0, 0)$$

$$= \int_{[0, T]^4} \int_{\mathbb{R}^{2d}} \mathbb{E}D_{s_1, x} Q(s_1, t_1, \varphi_{s_2}, y) D_{t_2, y} Q(s_2, t_2, x, \varphi_{t_1}) dx dy ds_1 ds_2 dt_1 dt_2$$

$$+ \int_0^T \int_0^T \mathbb{E}Q(s, t, \varphi_s, \varphi_t) dsdt .$$
Before proceeding to the proof, let us make the following remark which we will use several times in the future.

**Remark A.4.** Suppose that $f$ and $g$ are smooth functions, $f$ has compact support, and $\varphi$ is a random variable in $\mathbb{D}^{1, 2}$. Then the following integration by parts formula holds almost surely:

$$
\int_{\mathbb{R}^d} Df(x - \varphi)g(x)dx = -\int_{\mathbb{R}^d} f(x)Dg(x + \varphi)dx .
$$

Indeed, the integration on the left hand side is

$$
\int_{\mathbb{R}^d} \nabla f(x - \varphi) \cdot D\varphi g(x)dx .
$$

Integrating by parts yields

$$
-\int_{\mathbb{R}^d} f(x - \varphi)D\varphi \cdot \nabla g(x)dx .
$$

With the change of the variable $x \mapsto x + \varphi$,

$$
-\int_{\mathbb{R}^d} f(x)D\varphi \cdot \nabla g(x)dx = -\int_{\mathbb{R}^d} f(x)Dg(x + \varphi)dx .
$$

**Proof of Theorem A.3** For any $\varepsilon > 0$, the $\mathcal{H}$-valued random variable $\eta_\varepsilon(\cdot - \varphi_\varepsilon)$ belongs to $\mathbb{D}^{1, 2}(\mathcal{H})$, hence belongs to Dom $\delta$. Thus, applying (A.7), for all positive numbers $\varepsilon$ and $\varepsilon'$, we obtain

$$
\mathbb{E}(\delta(\eta_\varepsilon(\cdot - \varphi_\varepsilon))\delta(\eta_{\varepsilon'}(\cdot - \varphi_{\varepsilon'}))) = \mathbb{E} \langle \eta_\varepsilon(\cdot - \varphi_\varepsilon), \eta_{\varepsilon'}(\cdot - \varphi_{\varepsilon'}) \rangle_{\mathcal{H}} \\
+ \mathbb{E} \langle D\eta_\varepsilon(\cdot - \varphi_\varepsilon), (D\eta_{\varepsilon'}(\cdot - \varphi_{\varepsilon'}))^* \rangle_{\mathcal{H} \otimes \mathcal{H}} =: E_1 + E_2 .
$$

Using a change of variable, we have

$$
E_1 = \alpha_{\mathcal{H}} \mathbb{E} \int_0^T \int_0^T \int_{\mathbb{R}^{2d}} \eta_\varepsilon(x - \varphi_s)\eta_{\varepsilon'}(y - \varphi_t)q(x,y)|t - s|^{2H-2}dxdydsdt \\
= \alpha_{\mathcal{H}} \mathbb{E} \int_0^T \int_0^T \int_{\mathbb{R}^{2d}} \eta_\varepsilon(x)\eta_{\varepsilon'}(y)q(x + \varphi_s,y + \varphi_t)|t - s|^{2H-2}dxdydsdt \\
= \alpha_{\mathcal{H}} \int_0^T \int_0^T \mathbb{E}q(x + \varphi_s,y + \varphi_t)|t - s|^{2H-2}dxdydsdt .
$$

When $\varepsilon$ and $\varepsilon'$ tend to 0, using the conditions (A2) (A3) this quantity converges to

$$
\alpha_{\mathcal{H}} \int_0^T \int_0^T \mathbb{E}q(\varphi_s,\varphi_t)|t - s|^{2H-2}dsdt = q_\varphi(0, 0) .
$$

Hence, when $\varepsilon$ tends to zero, $\eta_\varepsilon(\cdot - \varphi_\varepsilon)$ converges in $L^2(\Omega; \mathcal{H})$ to an $\mathcal{H}$-valued random variable, denoted by $\delta_\varphi = \delta(\varphi_t - y)$.

For the second expectation in (A.14), we use (A.10) to obtain

$$
E_2 = \alpha_{\mathcal{H}}^2 \mathbb{E} \int_{[0,T]^4 \times \mathbb{R}^{4d}} D_{s_1,x_1} \eta_\varepsilon(x_2 - \varphi_{s_2})D_{t_2,y_2} \eta_{\varepsilon'}(y_1 - \varphi_{t_1})q(x_1,y_1)q(x_2,y_2) \\
\times |s_1 - t_1|^{2H-2}|s_2 - t_2|^{2H-2}dtdxdy .
$$
An application of \((A.13)\) yields
\[
E_2 = \alpha_H^2 \mathbb{E} \int_{[0,T]^4 \times \mathbb{R}^{4d}} \eta_\varepsilon(x_2)D_{s_1,x_1}q(x_2 + \varphi_{s_2}, y_2) \eta_\varepsilon(y_1)D_{t_1,y_1}q(x_1, y_1 + \varphi_{t_1}) \\
\times |s_1 - t_1|^{2H-2}|s_2 - t_2|^{2H-2}dsdttxdy.
\]
When \(\varepsilon\) and \(\varepsilon'\) tend to 0, this converges to \(q_{D\varphi}^\ast(0,0)\) by using conditions \((A.4)\) \((A.5)\).

Remark A.5. Under the hypothesis of the above theorem, the \(\mathcal{H}\)-valued random variable \(\eta_\varepsilon(\cdot - \varphi)\) converges in \(L^2(\Omega;\mathcal{H})\) to \(\delta_\varphi = \delta(\varphi_t - y)\) as \(\varepsilon\) tends to zero. Moreover, \(\delta_\varphi\) also belongs to the domain of the divergence operator and the convergence also holds under the divergence \(\delta\). Hence, in this case, the stochastic integral in Definition \((A.2)\) can be viewed as \(\delta(\delta_\varphi)\), the divergence of \(\delta_\varphi\).

A.2. Nonlinear symmetric stochastic integral. We introduce and study the symmetric type stochastic integral by using appropriate approximation. This stochastic integral will be different from the Itô-Skorohod type integral introduced in the previous subsection.

Recall that \(W = \{W(s, x, \omega), \omega \in \Omega\}\) is the Gaussian random field (indexed by \((s, x)\)) defined in the previous subsection. Throughout this subsection, we assume that \(W\) is almost surely continuous with respect to \(s \geq 0\) and \(x \in \mathbb{R}^d\). We define the composition of the random field \(W\) and an \(\mathbb{R}^d\)-valued process \(\varphi = \{\varphi_s, s \in [0, T]\}\) by
\[
W(s, \varphi_s) : \Omega \to \mathbb{R} \\
\omega \mapsto W(s, \varphi_s(\omega), \omega).
\]
(A.15)

By convention, we will assume that all processes and functions vanish outside the interval \([0, T]\).

Definition A.6. The symmetric integral \(\int_a^b W(\mathfrak{d}^{\text{sym}}s, \varphi_s)\) is defined as the limit as \(\varepsilon\) tends to zero of
\[
(2\varepsilon)^{-1} \int_a^b (W(s + \varepsilon, \varphi_s) - W(s - \varepsilon, \varphi_s)) \, ds,
\]
providing this limit exists in probability.

Example A.7. In the particular case when \(W(s, x) = B_s f(x)\), where \(f\) is a nice deterministic function and \(\{B_s, s \geq 0\}\) is a Brownian motion, the symmetric integral defined above coincides with the Stratonovich integral. That is, \(\int_0^T W(\mathfrak{d}^{\text{sym}}s, \varphi_s) = \int_0^T f(\varphi_s) \, d\mathbb{B}_s\).

In the following proposition we will see that for a suitable class of \(\mathbb{R}^d\)-valued processes \(\{\varphi_t\}\), the symmetric stochastic integral \(\int_0^T W(\mathfrak{d}^{\text{sym}}s, \varphi_s)\) exists almost surely. This result is an extension of \([1]\) Proposition 3].
Proposition A.8. Let $\varphi$ be an $\mathbb{R}^d$-valued process satisfying assumptions \[\text{(A5)}\] In addition, suppose that $\varphi$ satisfies

\[\text{(A.17)}\] $\int_0^T \int_{|x|<1} [\mathbb{E}q(x+\varphi_s,x+\varphi_s)]^{1/2} dx ds < \infty,$

\[\text{(A.18)}\] $\int_0^T \int_{|x|<1} \left[ \mathbb{E} \left\{ \sum_{i,j=1}^d \langle D\varphi_s^i, D\varphi_s^j \rangle_{\mathcal{H}} \right\} q(x+\varphi_s,x+\varphi_s) \right]^{1/2} \ dx ds < \infty$

and the function

\[\text{(A.19)}\] $x \mapsto \int_0^T \int_0^T \int_{\mathbb{R}^d} |D_{t,y} q(x+\varphi_s,y)| |s-t|^{2H-2} dt ds dy$

is a.s. well-defined and continuous on a neighborhood of 0. Assume also that the Gaussian field $W$ has continuous sample path. Then the symmetric integral \[\text{(A.16)}\] exists and the following formula holds almost surely:

\[\text{(A.20)}\] $\int_0^T W(\text{d}^{\text{sym}} s, \varphi_s) = \int_0^T W(\delta s, \varphi_s) + \alpha H \int_0^T \int_0^T \int_{\mathbb{R}^d} \delta(I_{|s-\varepsilon,s+\varepsilon|} \delta_x) \eta_{s'}(x-\varphi_s) \ dx ds.$

Proof. We shall show the convergence in $L^2$ of \[\text{(A.16)}\]. For every positive $\varepsilon$, since $W$ has continuous sample path, we can write

\[\text{(A.21)}\] $W(s+\varepsilon, \varphi_s) - W(s-\varepsilon, \varphi_s) = \lim_{\varepsilon' \to 0} \int_{\mathbb{R}^d} [W(s+\varepsilon, x) - W(s-\varepsilon, x)] \eta_{s'}(x-\varphi_s) \ dx$

almost surely, where we have used \[\text{(A.4)}\] in the last equality. We notice $\eta_{s'}(x-\varphi_s)$ belongs to $\mathcal{D}^{1,2}$ for every $s$ and $x$. Using \[\text{(A.8)}\], we see that the integrand on the right hand side of \[\text{(A.21)}\] can be written as

$$\delta \left( I_{|s-\varepsilon,s+\varepsilon|} \delta_x \eta_{s'}(x-\varphi_s) \right) + \langle D\eta_{s'}(x-\varphi_s), I_{|s-\varepsilon,s+\varepsilon|} \delta_x \rangle_{\mathcal{H}}.$$

Taking integration with respect to $x$ and $s$, we obtain

$$\int_0^T \int_{\mathbb{R}^d} [W(s+\varepsilon, x) - W(s-\varepsilon, x)] \eta_{s'}(x-\varphi_s) \ dx ds$$

$$= (2\varepsilon)^{-1} \int_0^T \int_{\mathbb{R}^d} \delta \left( I_{|s-\varepsilon,s+\varepsilon|} \delta_x \eta_{s'}(x-\varphi_s) \right) \ dx ds$$

$$+ (2\varepsilon)^{-1} \int_0^T \int_{\mathbb{R}^d} \langle D\eta_{s'}(x-\varphi_s), I_{|s-\varepsilon,s+\varepsilon|} \delta_x \rangle_{\mathcal{H}} \ dx ds$$

$$= I_1 + I_2.$$

The proof is now decomposed into several steps.

**Step 1.** Let us show that the integration with respect to $dx ds$ in $I_1$ can be interchanged with the divergence operator to obtain

$$I_1 = \delta \left( (2\varepsilon)^{-1} \int_0^T \int_{\mathbb{R}^d} I_{|s-\varepsilon,s+\varepsilon|} \delta_x \eta_{s'}(x-\varphi_s) \ dx ds \right).$$
In fact, one can view the integral in $I_1$ in the Bochner sense, that is, integration with $L^2$-valued integrand. In this setting, we have

$$
\int_0^T \int_{\mathbb{R}^d} \delta(u(s, x))dxds = \delta \left( \int_0^T \int_{\mathbb{R}^d} u(s, x)dxds \right)
$$

provided that

(A.23) \[ \int_0^T \int_{\mathbb{R}^d} \|u(s, x)\|_{\mathcal{D}^{1,2}}dxds < \infty \]

and $\delta$ is a bounded operator from $\mathbb{D}^{1,2}$ to $L^2$. The latter fact is automatically guaranteed by (A.17). It remains to check that $u(s, x) = I_{(s-\varepsilon, s+\varepsilon)} \delta_x \eta_{\varepsilon'}(x - \varphi_s)$ satisfies (A.23):

\[
\|u(s, x)\|_{\mathcal{H}}^2 = \int_{s-\varepsilon}^{s+\varepsilon} \int_{s-\varepsilon}^{s+\varepsilon} \frac{\partial^2}{\partial s \partial t} R_H(s', t')ds'dt' q(x, x) \mathbb{E}[\eta_{\varepsilon'}^2(x - \varphi_s)] \\
\leq R_H([0, T]^2)q(x, x) \mathbb{E}[\eta_{\varepsilon'}^2(x - \varphi_s)].
\]

Thus by a change of variable, we obtain

\[
\int_0^T \int_{\mathbb{R}^d} \|u(s, x)\|_{\mathcal{H}}dxds \leq R_H^{1/2}([0, T]^2) \int_0^T \int_{\mathbb{R}^d} (\mathbb{E}q(x, x) \eta_{\varepsilon'}^2(x - \varphi_s))^{1/2}dxds \\
= R_H^{1/2}([0, T]^2) \int_0^T \int_{\mathbb{R}^d} (\mathbb{E}(q(x + \varphi_s, x + \varphi_s) \eta_{\varepsilon'}^2(x)))^{1/2}dxds \\
\leq c(\varepsilon', T) \int_0^T \int_{|x| < 1} (\mathbb{E}(q(x + \varphi_s, x + \varphi_s)))^{1/2}dxds.
\]

The last integral is finite thanks to the condition (A.17). Similarly

\[
\|Du(s, x)\|_{\mathcal{H}^d}^2 = \mathbb{E} \int_{s-\varepsilon}^{s+\varepsilon} \int_{s-\varepsilon}^{s+\varepsilon} \frac{\partial^2}{\partial s \partial t} R_H(s', t')ds'dt' q(x, x) \partial_i \partial_j \eta_{\varepsilon'}(x - \varphi_s)\partial_i \partial_j \eta_{\varepsilon'}(x - \varphi_s) \langle D\varphi_s^i, D\varphi_s^j \rangle_{\mathcal{H}} \\
\leq R_H([0, T]^2)q(x, x) \mathbb{E} \sum_{i, j} \partial_i \eta_{\varepsilon'}(x - \varphi_s)\partial_j \eta_{\varepsilon'}(x - \varphi_s) \langle D\varphi_s^i, D\varphi_s^j \rangle_{\mathcal{H}}.
\]

Thus, by a change of variable and by using the condition (A.18), we obtain

\[
\int_0^T \int_{\mathbb{R}^d} \|Du(s, x)\|_{\mathcal{H}^d}dxds \\
\leq c(T) \int_0^T \int_{\mathbb{R}^d} (\mathbb{E}(q(x, x) \sum_{i, j} \partial_i \eta_{\varepsilon'}(x - \varphi_s)\partial_j \eta_{\varepsilon'}(x - \varphi_s) \langle D\varphi_s^i, D\varphi_s^j \rangle_{\mathcal{H}}))^{1/2}dxds \\
\leq c(\varepsilon', T) \int_0^T \int_{|x| < 1} (\mathbb{E}(q(x + \varphi_s, x + \varphi_s) \sum_{i, j} \langle D\varphi_s^i, D\varphi_s^j \rangle_{\mathcal{H}}))^{1/2}dxds < \infty.
\]
Step 2. We show that
\[
\delta \left( (2\varepsilon)^{-1} \int_0^T \int_{\mathbb{R}^d} I_{(s-\varepsilon,s+\varepsilon)} \delta_x \eta'(x-\varphi_s) \, dx \, ds \right)
= \delta \left( (2\varepsilon)^{-1} \int_0^T \int_{(s-\varepsilon,s+\varepsilon)} \eta'(\cdot - \varphi_s) \, ds \right).
\]
It suffices to show that for every smooth function \( \phi \) with compact support,
\[(A.24) \quad \phi = \int_{\mathbb{R}^d} \phi(y) \delta_y \, dy
\]
is in \( \mathcal{H} \), since with the choice \( \phi = \eta \), \( (A.24) \) will yield the desired identity. Recall that \( \mathcal{S} \) is the space defined in Subsection A.1 and is dense in \( \mathcal{H} \). Thus to show \( (A.24) \), we verify
\[
\langle \phi, \psi \rangle_{\mathcal{H}} = \langle \int_{\mathbb{R}^d} \phi(y) \delta_y \, dy, \psi \rangle_{\mathcal{H}}
\]
for every \( \psi \in \mathcal{S} \). Indeed, we have
\[
\int_{\mathbb{R}^d} \langle \phi(y) \delta_y \, dy, \psi \rangle_{\mathcal{H}} \, dy = \int_{\mathbb{R}^d} \phi(y) \langle \delta_y, \psi \rangle_{\mathcal{H}} \, dy
= \int_{\mathbb{R}^d} \int_0^T \int_0^T \int_{\mathbb{R}^d} \phi(y) \psi(x) Q(s,t,y,x) \, dx \, ds \, dt \, dy = \langle \phi, I_{(0,s)} \delta_x \rangle_{\mathcal{H}},
\]
by Fubini’s theorem.

Step 3. Combining the previous two steps, we obtain
\[
I_1 = \delta \left( (2\varepsilon)^{-1} \int_0^T \int_{(s-\varepsilon,s+\varepsilon)} \eta'(\cdot - \varphi_s) \, ds \right).
\]
It is straightforward to check that when \( \varepsilon' \) and \( \varepsilon \) tend to zero, \( I_1 \) converges to \( \int_0^T W(\delta s, \varphi_s) \) in \( L^2 \).

Step 4. We now show the convergence of \( I_2 \). A direct computation shows that
\[
|I_2| = (2\varepsilon)^{-1} \left| \int_0^T \int_{\mathbb{R}^d} \int_{-\varepsilon}^{\varepsilon} \int_0^T \int_{\mathbb{R}^d} D_{t,y} \eta' \, q(x,y) \, dx \, ds \, dt \, dr \right|
\leq d_H \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left| D_{t,y} \eta' \, q(x,y) \right| \, |t-s|^{2H-2} \, dy \, dt \, dx,
\]
where we have used the inequality
\[
(2\varepsilon)^{-1} \int_{-\varepsilon}^{\varepsilon} |t-r-s|^{2H-2} \, dr \leq d_H |t-s|^{2H-2}
\]
for some constant \( d_H \), independent \( \varepsilon \in (0,1) \) and \( s, t \in \mathbb{R} \). By a change of variable \( x-\varphi_s \to x \), we obtain
\[
I_2 \leq d_H \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left| \eta(x) D_{t,y} q(x+\varphi_s,y) \right| \, |t-s|^{2H-2} \, dy \, dx \, dt.
\]
Hence, by the dominated convergence theorem, when \( \varepsilon' \) and \( \varepsilon \) tend to zero, \( I_2 \) goes to \( \int_0^T \int_{\mathbb{R}^d} D_{t,y} q(x+\varphi_s,y) |t-s|^{2H-2} \, dy \, ds \, dt \). Therefore, passing through the limits in \( (A.22) \), we obtain \( (A.20) \). \qed
If the limit in Definition A.6 exists for almost every sample path of \( W \), then the symmetric integral can also be defined pathwise for a function \( (W(t, x), t \geq 0, x \in \mathbb{R}^d) \). We also call such integral the symmetric integral and denote it by the same symbol \( \int_0^T W(\text{sym}_s, \varphi_s) \).

The following proposition establishes the relation between the symmetric integral and the nonlinear Young integral introduced in Section 2.

**Proposition A.9.** Assume the hypothesis of Proposition 2.4 Then the symmetric integral exists and the following relation holds:

\[
\int_0^T W(\text{sym}_s, \varphi_s) = \int_0^T W(ds, \varphi_s).
\]

**Proof.** Fix \( \epsilon > 0 \). We write

\[ W_\epsilon(s, x) = (2\epsilon)^{-1} \int_{-\epsilon}^{\epsilon} W(s + \eta, x) d\eta. \]

We recall that \( \int_0^T W(\text{sym}_s, \varphi_s) = \lim_{\epsilon \to 0} \int_0^T \partial_t W_\epsilon(s, \varphi) ds \). We write

\[
\begin{align*}
\mu_k(a, b) &= W_{\epsilon k}(b, \varphi_a) - W_{\epsilon k}(a, \varphi_a), \\
\mu(a, b) &= W(b, \varphi_a) - W(a, \varphi_a).
\end{align*}
\]

Since \( W_\epsilon \) is continuously differentiable in time, the integral \( \int W_\epsilon(ds, \varphi_s) \) is understood in the classical sense and is equal to \( \int \partial_t W_\epsilon(s, \varphi_s) ds \). Hence, applying Proposition 2.10 we obtain, for any \( \theta \in (0, 1) \) such that \( \theta \tau + \lambda \gamma > 1 \),

\[
\begin{align*}
|\int_0^T W(ds, \varphi_s) - \int_0^T \partial_t W_\epsilon(s, \varphi_s) ds| &
\leq |W(T, \varphi_0) - W(0, \varphi_0) - W_\epsilon(T, \varphi_0) + W_\epsilon(0, \varphi_0)| \\
&\quad + c(\varphi)|W - W_\epsilon|_{\beta, \tau, \lambda}|b - a|^{\theta \tau \lambda}.
\end{align*}
\]

It remains to estimate the terms on the right side and show that they all converge to 0 when \( \epsilon \) goes to 0. For the first term,

\[
\begin{align*}
|W(T, \varphi_0) - W(0, \varphi_0) - W_\epsilon(T, \varphi_0) + W_\epsilon(0, \varphi_0)| &
\leq (2\epsilon)^{-1} \int_{-\epsilon}^{\epsilon} |W(T, \varphi_0) - W(0, \varphi_0) - W(T + \eta, \varphi_0) + W(\eta, \varphi_0)| d\eta \\
&\lesssim (2\epsilon)^{-1} \int_{-\epsilon}^{\epsilon} |\eta|^\tau d\eta \lesssim \epsilon^\tau.
\end{align*}
\]

For the second term, we put \( F = W - W_\epsilon \) and notice that

\[
\begin{align*}
|W_\epsilon(s, x) - W_\epsilon(s, y) - W_\epsilon(t, x) + W_\epsilon(t, y)| &
\leq (2\epsilon)^{-1} \int_{-\epsilon}^{\epsilon} |W(s + \eta, x) - W(s + \eta, y) - W(t + \eta, x) + W(t + \eta, y)| d\eta \\
&\leq |W|(1 + |x|^\beta + |y|^\beta)(2\epsilon)^{-1} \int_{-\epsilon}^{\epsilon} |s - t|^\tau |x - y|^\lambda d\eta \\
&\leq [W](1 + |x|^\beta + |y|^\beta)|s - t|^\tau |x - y|^\lambda.
\end{align*}
\]

Thus

\[
|F(s, x) - F(s, y) - F(t, x) + F(t, y)| \leq 2|W|(1 + |x|^\beta + |y|^\beta)|s - t|^\tau |x - y|^\lambda.
\]
On the other hand, 

\[ |W_\varepsilon(s, x) - W_\varepsilon(s, y) - W_\varepsilon(t, x) + W_\varepsilon(t, y)| \]

\[ \leq (2\varepsilon)^{-1} \int_{-\varepsilon}^{\varepsilon} d\eta |W(s + \eta, x) - W(s, x) - W(s + \eta, y) + W(s, y)| \]

\[ + |W(t, x) - W(t + \eta, x) - W(t, y) + W(t + \eta, y)| \]

\[ \leq 2|W|(1 + |x|^\beta + |y|^\beta)|x - y|^\lambda(2\varepsilon)^{-1} \int_{-\varepsilon}^{\varepsilon} |\eta|^\gamma d\eta \]

\[ \leq 2(1 + \varepsilon)^{-1}|W|(1 + |x|^\beta + |y|^\beta)|x - y|^\lambda \varepsilon^\gamma. \]

Hence, combining these two bounds, we get

\[ |F(s, x) - F(s, y) - F(t, x) + F(t, y)| \]

\[ \leq |W|(1 + |x|^\beta + |y|^\beta)|s - t|^\theta \varepsilon \|x - y|^\lambda \varepsilon^{(1 - \theta)}. \]

Thus \[W - W_\varepsilon]_{\beta, \theta, r, \lambda} \lesssim \varepsilon^{\gamma(1 - \theta)}, \]

which converges to 0 as \( \varepsilon \to 0. \)

**Appendix B. Estimates for diffusion process**

In this section, we prove the exponential integrability of the Hölder norm and the supremum norm of a diffusion process which is needed in proving the existence of the Feynman-Kac solution in Section 4. The results obtained here are known in the literature (see for instance [6], [18], [43]). However, it is difficult to find a single-source treatment that suits our purpose. Besides, our method is straightforward and unified. We present it here.

We recall that \( X_t^{r,x} \) satisfies the equation (4.6). We denote

\[ (B.1) \quad M_t^{r,x} = \sum_{j=1}^{d} \int_{r}^{t} \sigma^{i,j}(s, X_t^{r,x}) \delta B^j_s. \]

Since \( \sigma \) is bounded (by condition (1.1)), \( (M_t^{r,x}; t \geq r) \) is a continuous \( L^2 \) martingale. In addition, we have the following properties.

**Lemma B.1.** Let \( \alpha \) be a number in \((0, 1/2)\). There exist some positive constants \( \gamma_0 \) and \( \gamma_\alpha \) such that

\[ (B.2) \quad \mathbb{E} \exp \left\{ \gamma_0 \sup_{r \leq t \leq T} |M_t^{r,x}|^2 \right\} \leq C(T - r, \Lambda) < \infty \]

and

\[ (B.3) \quad \mathbb{E} \exp \left\{ \gamma_\alpha \left( \sup_{r \leq s, t \leq T} \frac{|M_t^{r,x} - M_s^{r,x}|}{|t - s|^{\alpha}} \right)^2 \right\} \leq C(T - r, \Lambda, \alpha) < \infty. \]

**Proof.** (B.2) is well-known and is a direct application of Doob’s maximal inequality and the Burkholder-Davis-Gundy inequality. (B.3) is proved in [6] Lemma 2]. However, for the reader’s convenience, we present a proof of (B.3) in the following. We will omit the upper indices \( r, x \). Applying the Garsia-Rodemich-Rumsey theorem (see [21] and [30], specifically [38 Theorem 2.1.3]) with \( \Psi(x) = x^p \) and \( p(x) = x^{\alpha + 2/p} \), we have

\[ |M_t - M_s| \leq 8(1 + \frac{2}{\alpha p}) 4^{1/p} |t - s|^{\alpha} \left\{ \int_{r}^{T} \int_{r}^{T} \left( \frac{|M_u - M_v|}{|u - v|^{\alpha + 2/p}} \right)^p dudv \right\}^{1/p}. \]
Dividing both sides by \(|t-s|^\alpha\) and taking the sup on \(r \leq s < t \leq T\), we see that there is a constant \(C = C(\alpha)\), independent of \(p \geq 1\), such that

\[
\mathbb{E} \left( \sup_{r \leq s < t \leq T} \frac{|M_t - M_s|}{|t-s|^\alpha} \right)^p \leq C^p \int_r^T \mathbb{E} \left| M_u - M_v \right|^p |u-v|^\alpha p + 2 \, du \, dv.
\]

An application of the Burkholder-Davis-Gundy inequality gives

\[
\|M_u - M_v\|_p \leq 2p^{1/2} \int_u^v \alpha u(s, X_s^r, x) ds \|t-s\|^{1/2} / p^{1/2} \leq 2\Lambda^{1/2} p^{1/2} (t-r)^{1/2}.
\]

It follows that there is a constant \(C\), which may be different from the above one, such that the \(p\)-moments of \(\sup_{r \leq s < t \leq T} \frac{|M_t-M_s|}{|t-s|^\alpha}\) is at most \(C^p p^{p/2} (T-r)^{\frac{1}{2} - \alpha} p\) for all \(p > \frac{1}{2} - \alpha\)^{-1}, which yields (B.3).

\[\square\]

**Lemma B.2.** Fix \(\alpha \in (0, 1/2)\). There exist positive constants \(C_0, \gamma_0\) and \(\gamma_\alpha\) such that

(B.4) \[\mathbb{E} \exp \left\{ \gamma_0 \sup_{r \leq t \leq T} |X_t^r, x| \right\} \leq e^{C_0 |x|^2}\]

and

(B.5) \[\mathbb{E} \exp \left\{ \gamma_\alpha \left( \sup_{r \leq s, t \leq T} \frac{|X_t^r, x| - X_s^r, x|}{|t-s|^{\alpha}} \right)^2 \right\} \leq e^{C_0 |x|^2} .\]

**Proof.** We denote \(X_t^* = \sup_{r \leq s \leq t} |X_t^r, x|\) and \(M_t^* = \sup_{r \leq s \leq t} |M_s^r, x|\). We first prove (B.4). Since \(b\) has linear growth (by (L.3)), from equation (4.6), we see that

\[|X_t| \leq |M_t| + |x| + \kappa(b) \int_r^t |X_s| ds .\]

An application of Gronwall’s inequality yields

\[|X_t| \leq |M_t| + |x| + \kappa(b) \int_r^t (|M_s| + |x|)e^{\kappa(b)(t-s)} ds .\]

Hence, for all \(p \geq 0\), applying Jensen’s inequality,

\[
\exp\{pX_t^*\} \leq \exp\{p(M_t^* + |x|)\} \exp\{p\kappa(b) \int_r^T (|M_s| + |x|)e^{\kappa(b)(t-s)} ds\}
\]

\[
\leq \exp\{p(M_t^* + |x|)\} \exp\{p\kappa(b)(T-r) - 1\} \int_r^T \exp\{p(e^{\kappa(b)(t-r)} - 1)(|M_s| + |x|)e^{\kappa(b)(T-s)} ds\}
\]

\[
\leq \exp\{p(M_t^* + |x|)\} \int_r^T \exp\{CP(|M_s| + |x|) ds\}
\]

for some constant \(C\) depending on \(T-r \) and \(\kappa(b)\). We then apply the Cauchy-Schwartz inequality

\[
\mathbb{E} e^{pX_t^*} \leq \mathbb{E} e^{2p(M_t^* + |x|)} + \int_r^T \mathbb{E} e^{2Cp(|M_s| + |x|) ds}
\]

\[
\leq \mathbb{E} e^{2Cp(M_t^* + |x|)} ,
\]

where the constants (including the implied constant) are independent of \(p\). Now we choose \(p\) according to the distribution \(|N(0, a)|\) with \(a\) sufficiently small, where \(N(0, a)\) is a normal distribution independent of \(B\). Using (B.2), the elementary
estimate \( \frac{1}{2} \sigma^2 \tau^2 A^2 \leq E N e^{pA} \leq 2e^{\frac{1}{2} \sigma^2 \tau^2 A^2} \) (with \( A > 0 \)), and the previous estimate, we obtain (B.4).

From (4.6), we have
\[
\frac{X_t - X_s}{(t-s)^\alpha} = \int_s^t b(u, X_u) du - M_t - M_s =: I_1 + I_2.
\]
Since \( b \) has linear growth, \( \sup_{r \leq s < t \leq T} |I_1| \leq c(\kappa(b), T, \alpha)(1 + X_t^r) \). (B.5) follows from (B.4) and (B.3).

Appendix C. Schauder Estimates

We present the proof of Lemma 4.3. The estimates (4.12)–(4.14) are similar to Schauder estimates in the classical theory of parabolic equations. Beside the results obtained in Appendix B, the method adopted here also makes use of Malliavin calculus. For this purpose, we need some preparation.

It is well-known (see e.g. [43]) that \( X_t^{r,x} \) is differentiable (in the Malliavin sense) with respect to the Brownian motion \( B_t \). We denote the Malliavin derivative of \( X \) with respect to \( B \) by \( D^j X \). It is shown in [43] Theorem 2.2.1 that \( DX = (D^1 X, \cdots, D^d X)^T \) has finite moments of all orders and satisfies
\[
dD_t X_t^{r,x} = \sigma_{ij}(t)D_{tr,x} D_tX_t^{r,x} \delta B_i^t + b_i(t) D_tX_t^{r,x} dt, \quad dD_X X_t^{r,x} = \delta ij(\tau, X_t^{r,x})
\]
for \( t \geq \tau \geq t \), \( D_t X_t^{r,x} = 0 \) if \( t < \tau \leq T \). In the above equation, we have used the notation
\[
\sigma_{ij}(t) = \partial_{x_1} \sigma_{ij}(t, X_t^{r,x}), \quad b_i(t) = \partial_{x_1} b_i(t, X_t^{r,x}).
\]
The matrix \( DX \) is understood as \( [DX]_{ij} = D^i X^j \). Following the proof of [43] Theorem 2.2.1, one can show that the map \( x \mapsto X_t^{r,x} \) is differentiable. We denote \( Y(t; r, x) = \frac{\partial}{\partial r} X_t^{r,x} \), the Jacobian of \( x \mapsto X_t^{r,x} \). The matrix \( Y \) is understood as \( [Y]_{ij} = Y_{ij} = \partial_j X^i \). It follows that the \( d \times d \)-matrix valued process \( t \mapsto Y(t; r, x) \) satisfies
\[
\begin{align*}
dY^i(t; r, x) &= \sigma_{ij}^{ij}(t) Y^j(t; r, x) \delta B_i^t + b_i(t) Y^i(t; r, x) dt, \\
Y(r; r, x) &= I_{d \times d}.
\end{align*}
\]
Let \( Z(t) = \) the \( d \times d \) matrix-valued process defined by
\[
dZ^i(t) = -Z_0^i(t) \sigma_{ij}^{\alpha \beta} (B_i^t - Z_0^i(t)) \left[ b_\alpha^i(t) - \sigma_{ij}^{\alpha \beta}(t) \sigma_{ij}^{\alpha \beta}(t) \right] dt, \quad Z(r; r, x) = I_{d \times d}.
\]
By means of Itô’s formula, we have
\[
d(Z_t^i Y_j^k) = -Y_j^k Z_0^i \sigma_{ij}^{\alpha \beta} \delta B_i^t - Y_j^i Z_0^k \sigma_{ij}^{\alpha \beta} \delta B_i^t dt + Z_0^k Y_j^i \sigma_{ij}^{\alpha \beta} \sigma_{ij}^{\alpha \beta} dt + Z_t^k \sigma_{ij}^{\alpha \beta} \delta B_i^t + Z_t^k Y_{ij}^\beta b_i^\beta dt - Z_0^k Y_j^i \sigma_{ij}^{\alpha \beta} \sigma_{ij}^{\alpha \beta} dt = 0
\]
and similarly for \( Y_t Z_t \). Thus we obtain \( Y_t Z_t = Y_t Z_t = I \). As a consequence, for every \( t \geq r \), the matrix \( Y(t; r, x) \) is invertible and its inverse is \( Z(t; r, x) \). It is a standard fact that \( Y \) and \( Z \) have finite moments of all orders. More precisely, one has
\[
\sup_{t \in [r, T], x \in \mathbb{R}^d} E \left[ |Y(t; r, x)|^p + |Y^{-1}(t; r, x)|^p \right] \leq c(p, T).
\]
Since the coefficients of \( L \) are twice differentiable with bounded derivatives, \( DY \) exists and has finite moment of all orders and
\[
\sup_{t \in [r, T], x \in \mathbb{R}^d} E \sup_{\tau \in [r, T]} \left[ |D_t Y(t; r, x)|^p + |D_t Y^{-1}(t; r, x)|^p \right] \leq c(p, T).
\]
Moreover, it is well-known that the following representation holds (see, for instance [43, p. 126]):

\begin{equation}
D_{\tau}X_{t}^{r,x} = Y(t; r, x)Z(\tau; r, x)\sigma(\tau, X_{\tau}^{r,x}), \quad \forall \tau \in [r, t].
\end{equation}

As a consequence, if \( f \) is a smooth function, we have

\begin{equation}
D_{\tau}f(s, X_{s}^{r,x})^{T} = \nabla f(s, X_{s}^{r,x})^{T}Y(s; r, x)Y^{-1}(\tau; r, x)\sigma(\tau, X_{\tau}^{r,x})
\end{equation}

(where and in what follows we denote \( \nabla f(s, X_{s}^{r,x}) = (\nabla f)(s, X_{s}^{r,x}) \)). Later on, we occasionally make use of its variant

\begin{equation}
\nabla f(s, X_{s}^{r,x})^{T}Y(s; r, x) = D_{\tau}f(s, X_{s}^{r,x})^{T}\sigma^{-1}(\tau, X_{\tau}^{r,x})Y(\tau; r, x), \quad \forall \tau \in [r, t].
\end{equation}

**Lemma C.1** (Bismut formula). Suppose \( f \) belongs to \( C^{2}(\mathbb{R}^{d+1}) \) and suppose \( f \) and its derivatives have polynomial growth. Then

\begin{equation}
\mathbb{E}[(\partial_{\tau}f)(s, X_{s}^{r,x})]
= \frac{1}{s - r} \mathbb{E}\left[f(s, X_{s}^{r,x}) \int_{r}^{s} [\sigma^{-1}(\tau, X_{\tau}^{r,x})Y(\tau; r, x)Y^{-1}(s; r, x)]^{ji}\delta B_{t}^{j}\right]
\end{equation}

and

\begin{equation}
\partial_{s}\mathbb{E}f(s, X_{s}^{r,x}) = \frac{1}{s - r} \mathbb{E}\left[f(s, X_{s}^{r,x}) \int_{r}^{s} [\sigma^{-1}(\tau, X_{\tau}^{r,x})Y(\tau; r, x)]^{ji}\delta B_{t}^{j}\right].
\end{equation}

**Proof.** Fix \( \tau \in [r, s] \). The identity (C.5) yields

\[ \nabla f(s, X_{s})^{T} = [D_{\tau}f(s, X_{s})^{T}]^{T} \sigma^{-1}(\tau)Y(\tau)^{-1}(s). \]

Integrating with respect to \( \tau \) from \( r \) to \( s \) and taking the expectation give

\[ \mathbb{E}\nabla f(s, X_{s})^{T} = \frac{1}{s - r} \mathbb{E}\left[\int_{r}^{s} [D_{\tau}f(s, X_{s})^{T}]^{T} \sigma^{-1}(\tau)Y(\tau)^{-1}(s)d\tau\right]. \]

Formula (C.7) then follows from the dual relationship (A.6) between the divergence operator \( \delta \) and the Malliavin derivative \( D \).

To show (C.8), we use (C.6). We integrate with respect to \( \tau \) from \( r \) to \( s \) and then take the expectation to obtain

\[ \nabla \mathbb{E}f(s, X_{s}) = \frac{1}{s - r} \mathbb{E}\left[\int_{r}^{s} [D_{\tau}f(s, X_{s})^{T}]^{T} \sigma^{-1}(\tau)Y(\tau)d\tau\right]. \]

Formula (C.8) follows from the dual relationship (A.6) between \( \delta \) and \( D \).

**Lemma C.2.** Suppose that \( f \) is differentiable and satisfies

\[ \sup_{s \in [0, T], x \in \mathbb{R}^{d}} \frac{|f(s, x)|}{1 + |x|^\beta} \leq \kappa \]

for some nonnegative constants \( \kappa \) and \( \beta \). Then we have

\begin{equation}
|\mathbb{E} (\nabla f)(s, X_{s}^{r,x})| \leq c(T, \Lambda, \lambda)\kappa(1 + |x|^\beta)[1 + (s - r)^{-1/2}]
\end{equation}

and

\begin{equation}
|\nabla \mathbb{E}f(s, X_{s}^{r,x})| \leq c(T, \Lambda, \lambda)\kappa(1 + |x|^\beta)(s - r)^{-1/2}.
\end{equation}
Proof. We only provide details for the proof of (C.9). The estimate (C.10) is proved similarly, perhaps in an easier manner. Motivated by the formula (C.7), we first estimate the moment of \( \int_r^s [\sigma^{-1}(\tau) Y(\tau) Y^{-1}(s)]^{j i} \delta B_r^j \). From (A.8), we see that
\[
\int_r^s [\sigma^{-1}(\tau) Y(\tau) Y^{-1}(s)]^{j i} \delta B_r^j = \int_r^s [\sigma^{-1}(\tau) Y(\tau)]^{j k} \delta B_r^j [Y^{-1}(s)]^{k i}
- \int_r^s [\sigma^{-1}(\tau) Y(\tau)]^{j k} D_r^j [Y^{-1}(s)]^{k i} d\tau.
\]
From (C.2) and (C.3), it follows that
\[
\sup_{s \in [r, t], x \in \mathbb{R}^d} \mathbb{E} \left| \int_r^s [\sigma^{-1}(\tau) Y(\tau) Y^{-1}(s)]^{j i} \delta B_r^j \right|^p \leq c(p, T) [(s - r)^{1/2} + (s - r)]^p.
\]
Hence, applying the Hölder inequality in (C.7),
\[
|\mathbb{E} \nabla f(s, X_{r,x}^s)| \lesssim (1 + (s - r)^{-1/2}) |\mathbb{E}(1 + |X_{r,x}^s|^\beta)|^{1/2}.
\]
Together with (B.4), this completes the proof of (C.9). \(\Box\)

Proof of Lemma 4.3. Throughout the proof, we denote \( \kappa_1 = [\nabla W]_{\beta_1, \infty} \), \( \kappa_2 = [\nabla W]_{\beta_2, \alpha} \), \( Y = \nabla \varphi \).

Uniqueness: Suppose \( v \) is a solution in \( C^1([0, T]; C^2(\mathbb{R}^d)) \). We apply Itô’s formula to the process \( s \mapsto (v + W)(s, \varphi_{r,x}^s) \), taking into account the fact that \( L_0 \) is the generator of \( \varphi_{r,x}^s \):
\[
d(v + W)(s, \varphi_{r,x}^s) = (\partial_t + L_0)(v + W)(s, \varphi_{r,x}^s) ds
+ \sigma^{ij}(s, \varphi_{r,x}^s) \partial_i (v + W)(s, \varphi_{r,x}^s) \delta B_r^j.
\]
Since \( v \) is a strong solution, we see that \( v + W \) satisfies
\[
(\partial_t + L_0)(v + W) = L_0 W, \quad (v + W)(T, x) = 0.
\]
Thus, integrating (C.11) from \( r \) to \( T \) yields
\[
-(v + W)(r, x) = \int_r^T L_0 W(s, X_{r,x}^s) ds + \int_r^T \sigma^{ij}(s, X_{r,x}^s) \partial_i (v + W)(s, X_{r,x}^s) \delta B_r^j.
\]
Taking the expectation in the above identity, we obtain (4.11), which also shows the uniqueness of \( v \).

\( C^0 \)-estimate: To prove the estimate (4.12), we write \( L_0 W = \partial_i \left( \frac{1}{2} a^{ij} \partial_j W \right) + c^j \partial_j W \) where \( c^j = -1/2 \partial_i a^{ij} \). Then
\[
\mathbb{E} \int_r^T L_0 W(s, X_{r,x}^s) ds = I_1 + I_2
\]
where
\[
I_1 = \mathbb{E} \int_r^T \partial_i \left( \frac{1}{2} a^{ij}(s, X_{r,x}^s) \partial_j W(s, X_{r,x}^s) \right) ds
\]
and
\[
I_2 = \mathbb{E} \int_r^T c^j (s, X_{r,x}^s) \partial_j W(s, X_{r,x}^s) ds.
\]
It follows from our conditions on \( L_0 \) and \( W \) that
\[
\sup_{t \in [0, T], x \in \mathbb{R}^d} \frac{|a^{ij}(t, x) \partial_i W(t, x)|}{1 + |x|^{\beta_i}} \leq \Lambda \kappa_1 \text{ and } \sup_{t \in [0, T], x \in \mathbb{R}^d} \frac{|c^j(t, x) \partial_j W(t, x)|}{1 + |x|^{\beta_j}} \leq \Lambda \kappa_1.
\]
Applying Lemma C.2, we obtain
\[ I_1 \lesssim \kappa_1 \int_r^T ((s - r)^{-1/2} + 1)ds(1 + |x|^{\beta_1}) \lesssim \kappa_1[(T - r)^{1/2} + (T - r)](1 + |x|^{\beta_1}). \]

For the second term, we use (3.7):
\[ I_2 \lesssim \kappa_1 \int_r^T \mathbb{E}(1 + |\varphi_s^{r,x}|^{\beta_1})ds \lesssim \kappa_1(T - r)(1 + |x|^{\beta_1}). \]

These inequalities together imply (4.12).

**Proof**

We now estimate the moment \( G^{ij}(s; r, x) \) using (A.8)
\[ \frac{\partial}{\partial t} \mathbb{E} L_0 W(s, \varphi_s^{r,x}) = (s - r)^{-1} \mathbb{E}[L_0 W(s, \varphi_s^{r,x})H(s, x)] \]
where
\[ H(s, x) = \int_r^s \left[ \sigma^{-1}(\tau, \varphi^{r,x}_\tau)Y(\tau; r, x) \right]^T \delta B_\tau. \]
We denote
\[ A(\tau, x) = \sigma^{-1}(\tau, X^{r,x}_\tau)Y(\tau; r, x). \]
From (C.5), we see that
\[ \partial_{ij}^2 W(s, \varphi_s^{r,x}) = D^k_\tau[\partial_j W(s, \varphi_s^{r,x})][A(\tau)Y^{-1}(s)]^{kj}, \quad \forall \tau \in [r, s]. \]
Thus
\[ L_0 W(s, \varphi_s^{r,x}) = \frac{1}{2} a^{ij}(s, X^{r,x}_s)\partial_{ij}^2 W(s, \varphi_s^{r,x}) \]
\[ = \frac{1}{2} D^k_\tau[\partial_j W(s, \varphi_s^{r,x})][A(\tau)Y^{-1}(s)]^{kj} a^{ij}(s, X^{r,x}_s) \]
\[ = \frac{1}{2}(s - r)^{-1} \int_r^s D^k_\tau[\partial_j W(s, \varphi_s^{r,x})][A(\tau)Y^{-1}(s)a(s, X^{r,x}_s)]^{kj} d\tau. \]

Hence, applying (A.6),
\[ \frac{\partial}{\partial t} \mathbb{E} L_0 W(s, \varphi_s^{r,x}) \]
\[ = \frac{1}{2}(s - r)^{-2} \mathbb{E} \int_r^s D^k_\tau[\partial_j W(s, \varphi_s^{r,x})][A(\tau)Y^{-1}(s)a(s, X^{r,x}_s)]^{kj} H^l(s, x)d\tau \]
\[ = \frac{1}{2}(s - r)^{-2} \mathbb{E} \partial_j W(s, \varphi_s^{r,x}) \int_r^s [A(\tau)Y^{-1}(s)a(s, X^{r,x}_s)]^{kj} H^l(s, x)\delta B^k_\tau. \]

Furthermore, since the random variable
\[ G^{ij}(s; r, x) := \int_r^s [A(\tau)Y^{-1}(s)a(s, X^{r,x}_s)]^{kj} H^l(s, x)\delta B^k_\tau \]
has mean zero, we can write
\[ \frac{\partial}{\partial t} \mathbb{E} L_0 W(s, \varphi_s^{r,x}) = \frac{1}{2}(s - r)^{-2} \mathbb{E}[\partial_j W(s, \varphi_s^{r,x}) - \partial_j W(s, x)]G^{ij}(s; r, x). \]
We now estimate the moment \( G(s; r, x) \).

Applying (A.8), we have
\[ G^{ij}(s; r, x) = \int_r^s [A(\tau)]^{km}[Y^{-1}(s)a(s, X^{r,x}_s)]^{mj} H^l(s, x)\delta B^k_\tau \]
\[ = [Y^{-1}(s)a(s, X^{r,x}_s)]^{mj} H^l(s, x) \int_r^s [A(\tau)]^{km} \delta B^k_\tau \]
\[ - \int_r^s D^k_\tau([Y^{-1}(s)a(s, X^{r,x}_s)]^{mj} H^l(s, x))[A(\tau)]^{km} d\tau. \]
Using properties of the Malliavin derivative, we have
\[
D^k_r([Y^{-1}(s)a(s, X^{r,x}_s)]^mH^l(s, x))
= D^k_r[Y^{-1}(s)a(s, X^{r,x}_s)]^mH^l(s, x) + [Y^{-1}(s)a(s, X^{r,x}_s)]^m D^k_r H^l(s, x).
\]
Hence
\[
G^{jl}(s; r, x) = [Y^{-1}(s)a(s, X^{r,x}_s)]^mH^l(s, x) \int_r^s [A(\tau)]^{km}d\tau
- \int_r^s D^k_r[Y^{-1}(s)a(s, X^{r,x}_s)]^mH^l(s, x)[A(\tau)]^{km}d\tau
- \int_r^s [Y^{-1}(s)a(s, X^{r,x}_s)]^m D^k_r H^l(s, x)[A(\tau)]^{km}d\tau.
\]
(C.13)

Since \(a\) belongs to \(C^2\), estimate \((C.3)\) is valid, the moments of \(A(\tau)\) are also uniformly bounded (because \(a\) is strictly elliptic), and all the terms appearing in \(G^{jl}\) have finite moments of all orders. In addition, observe that
\[
D^i_r H^l(s, x) = 1_{\{r \leq \tau\}}A(\tau)^{il} + \int_r^s D^i_r A(u)^{kl}d\tau.
\]
Thus, the \(L^p\)-norm of \(H(s, x)\) and \(DH(s, x)\) will contribute a factor \((r - s)^{1/2}\). Therefore, it follows from the Burkholder-Davis-Gundy inequality and the Hölder inequality that
\[
\sup_{x \in \mathbb{R}^d} \|G^{jl}(s; r, x)\|_p \leq c(p, \lambda, \Lambda)[(s - r) + (s - r)^{3/2}], \forall p \geq 1.
\]
(C.14)

Using the Hölder continuity of \(W\), for every \(p \geq 1\), we have
\[
\|\nabla W(s, \varphi^{r,x}_s) - \nabla W(s, x)\|_p \leq \kappa_2 \|1 + |\varphi^{r,x}_s|^{\beta_2} + |x|^{\beta_2}|\varphi^{r,x}_s - x|^{\alpha}\|_p.
\]
Taking into account the moment estimate \((B.4)\) and Hölder inequality, this gives
\[
\|\nabla W(s, \varphi^{r,x}_s) - \nabla W(s, x)\|_p \leq c(\alpha, \beta_2, p, \Lambda)\kappa_2(1 + |x|^{\beta_2})(s - r)^{\alpha/2}.
\]
(C.15)

Thus, applying the Cauchy-Schwartz inequality in \((C.12)\) yields
\[
|\partial \mathbb{E}L_0 W(s, \varphi^{r,x}_s)| \leq c(\lambda, \Lambda)(s - r)^{-2}\|\nabla W(s, \varphi^{r,x}_s) - \nabla W(s, x)\|_2\|G(s; r, x)\|_2.
\]
Applying the moment estimate for \(G\) and \((C.15)\), we obtain
\[
|\partial \mathbb{E}L_0 W(s, \varphi^{r,x}_s)| \leq c(\alpha, p, \Lambda)\kappa_2(1 + |x|^{\beta_2})(s - r)^{\alpha/2 - 1} + (s - r)^{\alpha/2 - 1/2} \kappa_2(1 + |x|^{\beta_2}),
\]
which together with \((C.12)\) implies \((C.13)\).

**\(C^{1,\alpha'}\)-estimate:** This is the only place where we use the fact that the second derivatives of \(a\) are Hölder continuous. Each term appearing on the right hand side of \((C.13)\) is either differentiable or Hölder continuous in the \(x\)-variable. Thus, we obtain easily the estimate
\[
\|G(s; r, x) - G(s; r, y)\|_p \leq c(p, \lambda, \Lambda)[(s - r) + (s - r)^{3/2}]|x - y|^{\alpha}.
\]
From \((C.15)\), we see that
\[
\|\nabla W(s, \varphi^{r,x}_s) - \nabla W(s, x) - \nabla W(s, \varphi^{r,y}_s) + \nabla W(s, y)\|_p
\leq \|\nabla W(s, \varphi^{r,x}_s) - \nabla W(s, x)\|_p + \|\nabla W(s, \varphi^{r,y}_s) - \nabla W(s, y)\|_p
\leq c(\alpha, p, \Lambda)\kappa_2(1 + |x|^{\beta_2} + |y|^{\beta_2})(s - r)^{\alpha/2}.
\]
On the other hand, we also have
\[
\| \nabla W(s, \varphi^{r,x}_s) - \nabla W(s, x) - \nabla W(s, \varphi^{r,y}_s) + \nabla W(s, y) \|_p \\
\leq \| \nabla W(s, \varphi^{r,x}_s) - \nabla W(s, \varphi^{r,y}_s) \|_p + \| \nabla W(s, x) - \nabla W(s, y) \|_p \\
\leq k_2(W)(1 + |\varphi^{r,x}_s|^{\beta_2} + |\varphi^{r,y}_s|^{\beta_2})|\varphi^{r,x}_s - \varphi^{r,y}_s|^{\alpha} \|_p \\
+ k_2(W)(1 + |x|^{\beta_2} + |y|^{\beta_2})|x - y|^{\alpha},
\]
where the last estimate comes from (B.4) and the fact that the derivative of the map 
\( x \mapsto \varphi^{r,x}_s \) has finite moments uniformly in \( x \). Interpolating these two inequalities we obtain
\[
(C.17) \quad \| \nabla W(s, \varphi^{r,x}_s) - \nabla W(s, x) - \nabla W(s, \varphi^{r,y}_s) + \nabla W(s, y) \|_p \\
\leq c(\alpha, p, \Lambda)k_2(1 + |x|^{\beta_2} + |y|^{\beta_2})|x - y|^{\beta(p-1)/2}\alpha / 2
\]
for any \( \vartheta \in [0, 1] \). Thus, from (C.12), applying the Cauchy-Schwartz inequality we see that
\[
|\nabla \mathbb{E}L_0W(s, \varphi^{r,x}_s) - \nabla \mathbb{E}L_0W(s, \varphi^{r,y}_s)| \\
\leq (s - r)^{-\vartheta/2}\| \nabla W(s, \varphi^{r,x}_s) - \nabla W(s, x) - \nabla W(s, \varphi^{r,y}_s) + \nabla W(s, y) \|_p \| G(s; r, x) \|_2 \\
+ (s - r)^{-\vartheta/2}\| \nabla W(s, \varphi^{r,y}_s) - \nabla W(s, y) \|_p \| G(s; r, x) - G(s; r, y) \|_2.
\]
Using (C.17), (C.15), (C.14) and (C.16), we obtain
\[
|\nabla \mathbb{E}L_0W(s, \varphi^{r,x}_s) - \nabla \mathbb{E}L_0W(s, \varphi^{r,y}_s)| \\
\leq c(\alpha, \lambda, \Lambda)k_2(1 + |x|^{\beta_2} + |y|^{\beta_2})|x - y|^{\beta(\alpha - 1)/2} - (s - r)^{\beta\alpha/2 - 2/1} + (s - r)^{\beta\alpha/2 - 1/2} \\
+ c(\alpha, \lambda, \Lambda)k_2(1 + |y|^{\beta_2})|x - y|^{\alpha\beta - 1/2} + (s - r)^{\alpha/2 - 1/2}.
\]
Therefore, choosing \( \vartheta < 1 \), this estimate together with (4.11) implies (4.14). \( \square \)

Acknowledgments

The authors sincerely thank the referee for the many constructive and detailed comments which were of great help in revising the manuscript. They also thank Samy Tindel and Massimiliano Gubinelli for their interest in this paper.

References

[1] Elisa Alòs and David Nualart, *Stochastic integration with respect to the fractional Brownian motion*, Stoch. Stoch. Rep. 75 (2003), no. 3, 129–152, DOI 10.1080/1045112031000078917. MR1978896 (2004b:60138)

[2] Luigi Ambrosio, *Transport equation and Cauchy problem for BV vector fields*, Invent. Math. 158 (2004), no. 2, 227–260, DOI 10.1007/s00222-004-0367-2. MR2096794 (2005f:35127)

[3] Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin, *Fourier analysis and nonlinear partial differential equations*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 343, Springer, Heidelberg, 2011. MR2768550 (2012m:35004)

[4] G. Ben Arous and M. Ledoux, *Grandes déviations de Freidlin-Wentzell en norme hölderienne* (French, with English and French summaries), Séminaire de Probabilités, XXVIII, Lecture Notes in Math., vol. 1613, Springer, Berlin, 1994, pp. 293–299, DOI 10.1007/BFb0073852. MR1329119 (96c:60034)

[5] R. Catellier and M. Gubinelli, *Averaging along irregular curves and regularisation of odes*, preprint arXiv:1205.1735, 2012.
[6] Sandra Cerrai, Second order PDE's in finite and infinite dimension: A probabilistic approach, Lecture Notes in Mathematics, vol. 1762, Springer-Verlag, Berlin, 2001. MR1840644 (2002j:35327)

[7] K. Chouk and M. Gubinelli, Nonlinear PDEs with modulated dispersion I: Nonlinear Schrödinger equations, Comm. Partial Differential Equations 40 (2015), no. 11, 2047–2081. MR3418825

[8] K. Chouk and M. Gubinelli, Nonlinear pdes with modulated dispersion II: Korteweg–de Vries equation, preprint arXiv:1406.7675, 2014.

[9] Gianluca Crippa and Camillo De Lellis, Estimates and regularity results for the DiPerna–Lions flow, J. Reine Angew. Math. 616 (2008), 15–46, DOI 10.1515/CRELLE.2008.016. MR2369485 (2008m:34085)

[10] A. M. Davie, Uniqueness of solutions of stochastic differential equations, Int. Math. Res. Not. IMRN 24 (2007), Art. ID rnm124, 26, DOI 10.1093/imrn/rnm124. MR2377011 (2008m:60099)

[11] Pierre Del Moral, Feynman-Kac formulae: Genealogical and interacting particle systems with applications, Probability and its Applications (New York), Springer-Verlag, New York, 2004. MR2044973 (2005f:60003)

[12] Pierre Del Moral, Mean field simulation for Monte Carlo integration, Monographs on Statistics and Applied Probability, vol. 126, CRC Press, Boca Raton, FL, 2013. MR3060290

[13] R. J. DiPerna and P.-L. Lions, Ordinary differential equations, transport theory and Sobolev spaces, Invent. Math. 98 (1989), no. 3, 511–547, DOI 10.1007/BF01393835. MR1022305 (90j:34004)

[14] Bruce K. Driver and Yaozhong Hu, On heat kernel logarithmic Sobolev inequalities, Stochastic analysis and applications (Powys, 1995), World Sci. Publ., River Edge, NJ, 1996, pp. 189–200. MR1453132 (98h:58183)

[15] Pierre Del Moral, Feynman-Kac formulae: Genealogical and interacting particle systems with applications (New York, 1995), World Scientific Publishing Co., Inc., River Edge, NJ, 1996, pp. 189–200. MR1453132 (98h:60183)

[16] R. M. Dudley, Sample functions of the Gaussian process, Ann. Probability 1 (1973), no. 1, 66–103. MR0346884 (54 #1355)

[17] Denis Feyel and Arnaud de La Pradelle, Curvilinear integrals along enriched paths, Electron. J. Probab. 11 (2006), no. 34, 860–892 (electronic), DOI 10.1214/EJP.v11-356. MR2261056 (2007k:60112)

[18] F. Flandoli, M. Gubinelli, and E. Priola, Flow of diffeomorphisms for SDEs with unbounded Hölder continuous drift (English, with English and French summaries), Bull. Sci. Math. 134 (2010), no. 4, 405–422, DOI 10.1016/j.bulsci.2010.02.003. MR2651899 (2011f:60133)

[19] F. Flandoli, M. Gubinelli, and E. Priola, Well-posedness of the transport equation by stochastic perturbation, Invent. Math. 180 (2010), no. 1, 1–53, DOI 10.1007/s00222-009-0224-4. MR2593276 (2011f:35583)

[20] Mark Freidlin, Functional integration and partial differential equations, Annals of Mathematics Studies, vol. 109, Princeton University Press, Princeton, NJ, 1985. MR833742 (87g:60066)

[21] A. M. Garsia, E. Rodemich, and H. Rumsey Jr., A real variable lemma and the continuity of some Gaussian processes, Indiana Univ. Math. J. 20 (1970/1971), 565–578. MR0267632 (42 #2534)

[22] David Gilbarg and Neil S. Trudinger, Elliptic partial differential equations of second order, Classics in Mathematics, Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition. MR1814364 (2001k:35004)

[23] Mircea Grigoriu, Stochastic systems: Uncertainty quantification and propagation, Springer Series in Reliability Engineering, Springer, London, 2012. MR2917674

[24] M. Gubinelli, Controlling rough paths, J. Funct. Anal. 216 (2004), no. 1, 86–140, DOI 10.1016/j.jfa.2004.01.002. MR2091358 (2005k:60169)

[25] Massimiliano Gubinelli and Samy Tindel, Rough evolution equations, Ann. Probab. 38 (2010), no. 1, 1–75, DOI 10.1214/08-AOP437. MR2599193 (2011f:60216)

[26] Takeyuki Hida, Brownian motion, Applications of Mathematics, vol. 11, Springer-Verlag, New York-Berlin, 1980. Translated from the Japanese by the author and T. P. Speed. MR562914 (81a:60089)

[27] Helge Holden and Yaozhong Hu, Finite difference approximation of the pressure equation for fluid flow in a stochastic medium—a probabilistic approach, Comm. Partial Differential Equations
Equations 21 (1996), no. 9-10, 1367–1388, DOI 10.1080/03605309608821231. MR1410834 (97k:65303)

[28] Yaozhong Hu, Itô-Wiener chaos expansion with exact residual and correlation, variance inequalities, J. Theoret. Probab. 10 (1997), no. 4, 835–848, DOI 10.1023/A:1022654314791. MR1481650 (99b:60079)

[29] Yaozhong Hu, Jingyu Huang, David Nualart, and Samy Tindel, Stochastic heat equations with general multiplicative Gaussian noises: Hölder continuity and intermittency, Electron. J. Probab. 20 (2015), no. 55, 50, DOI 10.1214/EJP.v20-3316. MR3354615

[30] Yaozhong Hu and Khoa Le, A multiparameter Garsia-Rodemich-Rumsey inequality and some applications, Stochastic Process. Appl. 123 (2013), no. 9, 3359–3377, DOI 10.1016/j.spa.2013.04.019. MR3071383

[31] Yaozhong Hu and Khoa N. Lê, Nonlinear Young integrals via fractional calculus, Stochastics of environmental and financial economics—Centre of Advanced Study, Oslo, Norway, 2014–2015, Springer Proc. Math. Stat., vol. 138, Springer, Cham, 2016, pp. 81–99, DOI 10.1007/978-3-319-23425-0_3. MR3451169

[32] Yaozhong Hu, Fei Lu, and David Nualart, Feynman-Kac formula for the heat equation driven by fractional noise with Hurst parameter $H < 1/2$, Ann. Probab. 40 (2012), no. 3, 1041–1068, DOI 10.1214/11-AOP649. MR2962086

[33] Yaozhong Hu and David Nualart, Differential equations driven by Hölder continuous functions of order greater than $1/2$, Stochastic analysis and applications, Abel Symp., vol. 2, Springer, Berlin, 2007, pp. 399–413, DOI 10.1007/978-3-540-70847-6_17. MR2397797

[34] Yaozhong Hu and David Nualart, Rough path analysis via fractional calculus, Trans. Amer. Math. Soc. 361 (2009), no. 5, 2689–2718, DOI 10.1090/S0002-9947-08-04631-X. MR2471936

[35] Yaozhong Hu, David Nualart, and Jian Song, Feynman-Kac formula for heat equation driven by fractional white noise, Ann. Probab. 39 (2011), no. 1, 291–326, DOI 10.1214/10-AOP547. MR2778803 (2012b:60208)

[36] N. V. Krylov and E. Priola, Elliptic and parabolic second-order PDEs with growing coefficients, Comm. Partial Differential Equations 35 (2010), no. 1, 1–22, DOI 10.1080/03605300903424700. MR2748616 (2011m:35142)

[37] Hiroshi Kunita, Stochastic flows and stochastic differential equations, Cambridge Studies in Advanced Mathematics, vol. 24, Cambridge University Press, Cambridge, 1990. MR1070361 (91m:60107)

[38] Michel Ledoux, The concentration of measure phenomenon, Mathematical Surveys and Monographs, vol. 89, American Mathematical Society, Providence, RI, 2001. MR1849347 (2003k:28019)

[39] Terry Lyons, Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young, Math. Res. Lett. 1 (1994), no. 4, 451–464, DOI 10.4310/MRL.1994.v1.n4.a5. MR1302388 (96b:60150)

[40] Terry J. Lyons, Differential equations driven by rough signals, Rev. Mat. Iberoamericana 14 (1998), no. 2, 215–310, DOI 10.4171/RMI/140. MR1654527 (2000e:60089)

[41] Michael B. Marcus and Jay Rosen, Markov processes, Gaussian processes, and local times, Cambridge Studies in Advanced Mathematics, vol. 100, Cambridge University Press, Cambridge, 1999. MR1703611 (91m:60107)

[42] Barry Simon, Functional integration and quantum physics, 2nd ed., AMS Chelsea Publishing, Providence, RI, 2005. MR2105995 (2005f:81003)

[43] A. V. Skorohod, On a generalization of the stochastic integral (Russian, with English summary), Teor. Verojatnost. i Primenen. 20 (1975), no. 2, 223–238. MR0391258
[48] Daniel W. Stroock and S. R. Srinivasa Varadhan, *Multidimensional diffusion processes*, Classics in Mathematics, Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition. MR2190038 (2006f:60005)

[49] Michel Talagrand, *The generic chaining: Upper and lower bounds of stochastic processes*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2005. MR2133757 (2006b:60006)

[50] Jack Xin, *An introduction to fronts in random media*, Surveys and Tutorials in the Applied Mathematical Sciences, vol. 5, Springer, New York, 2009. MR2527020 (2010h:35331)

[51] L. C. Young, *An inequality of the Hölder type, connected with Stieltjes integration*, Acta Math. 67 (1936), no. 1, 251–282, DOI 10.1007/BF02401743. MR1555421

Department of Mathematics, The University of Kansas, Lawrence, Kansas 66045

E-mail address: yhu@ku.edu

Department of Mathematics, The University of Kansas, Lawrence, Kansas 66045

E-mail address: khoale@ku.edu