

A stochastic sewing lemma and applications*

Khoa Lê[†]

Abstract

We introduce a stochastic version of Gubinelli’s sewing lemma ([18]), providing a sufficient condition for the convergence in moments of some random Riemann sums. Compared with the deterministic sewing lemma, adaptiveness is required and the regularity restriction is improved by a half. The limiting process exhibits a Doob-Meyer-type decomposition. Relations with Itô calculus are established. To illustrate further potential applications, we use the stochastic sewing lemma in studying stochastic differential equations driven by Brownian motions or fractional Brownian motions with irregular drifts.

Keywords: sewing lemma; Doob-Meyer decomposition; rough paths; regularization by noise; stochastic differential equations; fractional Brownian motion; additive functional; chaos expansion.

AMS MSC 2020: Primary 60H10, Secondary 60H05; 60L20.

Submitted to EJP on October 27, 2018, final version accepted on March 6, 2020.

Supersedes arXiv:1810.10500.

Contents

1	Introduction	2
2	Stochastic sewing lemma	4
2.1	Relations with Itô calculus	10
2.2	Proofs	13
3	Distributive functionals of Markov processes	20
4	Stochastic flows	26
5	Chaos expansions	34
6	SDEs driven by fractional Brownian motions	41

*Supported by the Leverhulme Trust through Martin Hairer’s leadership award.

[†]Imperial College London, London, SW7 2AZ, United Kingdom. E-mail: le@math.tu-berlin.de

7 Averaging along fractional Brownian motions **50****References** **53****1 Introduction**

The sewing lemma was introduced by Gubinelli in [18, Proposition 1]. It generalizes earlier works of Young [45] and Lyons [31], provides a sufficient condition ensuring the convergence of some abstract Riemann sums. This point of view was later highlighted in the work of Feyel and de La Pradelle [11, Lemma 2.1], in which the lemma was called sewing lemma. Known for its use in deriving estimates for rough integrals (see for instance [15, Chapter 4]), the sewing lemma is one of the essential tools in Lyons' rough path theory ([31]).

The current article introduces a stochastic version of the sewing lemma, Theorem 2.1 below. It relaxes the regularity assumption of the original sewing lemma by a half but instead requires a certain adaptiveness of the considered increment processes. In a context of multidimensional parameter spaces, the sewing lemma is called reconstruction theorem and is introduced by Hairer [19, Theorem 3.23]. Needless to say, the reconstruction theorem also plays a fundamental role in the theory of regularity structures. However, it is not clear at the moment of writing if a stochastic reconstruction theorem is available.

We will describe the stochastic sewing lemma in detail in Section 2. While its proof is reminiscent of [11], the new observation that we bring in is the use of the Doob's decomposition ([9]). This approach naturally leads to a unique decomposition of the constructed process into a martingale and a remainder (Theorem 2.2). Such result is reminiscent of the classical Doob-Meyer decomposition, except that the remainder is not necessary a process of finite variation. Relations between the stochastic sewing lemma and Itô calculus are discussed in Examples 2.10, 2.11 and 2.12. In these examples, we show that Itô integrations, quadratic variations of certain martingales and Itô formulas can be formulated and obtained by means of the stochastic sewing lemma. These examples suggest that the essential elements of the stochastic sewing lemma have deep connections with the foundations of stochastic analysis. Therefore, formulating these elements as a single instrumental lemma provides new insights and leads to new applications. To illustrate this point, we have included a few non-trivial applications, which are briefly described below.

(i) Suppose that $\{f_t\}_{t \geq 0}$ is a family of distributions with a certain negative regularity index and X is a Markov process whose transition semigroup maps each f_t to a bounded continuous function. In Section 3, we provide a robust definition for the additive functional $\int_0^\cdot f_s(X_s) ds$ which extends the classical integration in the case f is continuous.

(ii) We consider the stochastic differential equation

$$X_t^x = x + \int_0^t b(s, X_s^x) ds + W_t, \quad \forall t \in [0, T], \quad (1.1)$$

where $x \in \mathbb{R}^d$, $b \in [L^\infty([0, T]; C_b^\alpha(\mathbb{R}^d))]^d$ for some $\alpha \in (0, 1)$ and W is a standard Brownian motion. In [12], the authors show that the map $x \mapsto X_t^x$ is differentiable and its derivatives are Hölder continuous in the spatial variables. However, because b is not differentiable, it is difficult to write down an equation for the process $Y := \nabla X^x$ rigorously. We explain in Section 4 that Y satisfies a system of Young-type differential equations. As a consequence, we show that $t \mapsto \nabla X_t^x$ is Hölder continuous for every fixed x .

(iii) In Section 5, we study weak solutions of the stochastic differential equation (1.1) with drift $b \in [L^q([0, T]; \mathcal{C}^\nu(\mathbb{R}^d))]^d$, for some suitable $q \in [1, \infty]$ and $\nu \in (-1, 0)$.

Here, $\mathcal{C}^\nu(\mathbb{R}^d)$ is the Besov-Hölder space. Depending on each situation, existence and uniqueness in law of weak solutions to (1.1) can be derived. We will not ponder on this problem in the article, but rather refer readers to various examples in the literature, [14, 2, 25, 8, 5, 13]. Starting from a pathwise solution (W, X) defined on a complete probability space, we derive truncated Wiener chaos expansions for $\phi(X_t)$, where ϕ is a regular deterministic test function. Consequently, we obtain a criterion to determine if (W, X) is indeed a strong solution. Verifying this criterion, however, is beyond the scope of the article. This result extends previous works of Krylov and Veretennikov in [42, 41] who considered the cases when the drifts are measurable functions.

(iv) The stochastic sewing lemma is also capable in situations where Markov properties are not apparent at the first sight. In Section 6, we consider the stochastic differential equation

$$X_t = x + \int_0^t b(r, X_r)dr + B_t^H, \quad t \in [0, T], \tag{1.2}$$

where $x \in \mathbb{R}^d$, b is a Borel function in $[L^q([0, T]; L^p(\mathbb{R}^d))]^d$, $p, q \in [1, \infty]$. Here B^H is a standard fractional Brownian motion with Hurst parameter $H \in (0, \frac{1}{2})$. We obtain weak existence and uniqueness in law for (1.2) under the condition

$$H \frac{d}{p} + \frac{1}{q} < \frac{1}{2}.$$

In addition, we show that pathwise uniqueness and strong existence hold if

$$H \frac{d}{p} + \frac{1}{q} < \frac{1}{2} - H.$$

The former result relies on Girsanov transformation, while the later is obtained by means of the stochastic sewing lemma. The results of Section 6 extend earlier works of Nualart and Ouknine in [37, 38] and Baños, Nilssen and Proske in [1].

(v) In Section 7, we investigate the averaging effect of fractional Brownian motions. Namely, for a given distribution f in $L^q([0, T]; \mathcal{C}^\nu(\mathbb{R}))$, $\nu \in \mathbb{R}$, the random field

$$(t, x) \rightarrow \int_0^t f_r(B_r^H + x)dr$$

can be defined and has a joint-Hölder continuous (in the sense of [23]) version. This type of regularity plays a central role in Catellier and Gubinelli’s study on path-by-path uniqueness for stochastic differential equations driven by fractional Brownian motions with distributional drifts ([6]). To obtain joint-Hölder continuity properties for such random fields, the method of [6] involves Fourier transforms, moment estimates and chaining arguments. Here, we obtain these properties by means of the stochastic sewing lemma and the multiparameter Garsia-Rodemich-Rumsey inequality of Hu and Lê in [23]. In comparison with [6], our approach provides explicit regularity exponents in space and time simultaneously.

After the appearance of the first manuscript of this article, Hairer and Li in [21] have used the stochastic sewing lemma introduced herein to study averaging dynamics of slow and fast systems where the slow system is driven by fractional Brownian motion with Hurst parameter $H > \frac{1}{2}$. The stochastic sewing lemma is also used by Butkovsky, Dareiotis and Gerencsér in [4] to obtain convergence rate of the Euler-Maruyama scheme for stochastic differential equations driven by fractional Brownian motions with irregular drifts.

We conclude the introduction with some notation which are used throughout the article. For every $\nu \leq 0$, $\mathcal{C}^\nu(\mathbb{R}^d)$ denotes the Besov-Hölder space $\mathcal{B}_{\infty, \infty}^\nu(\mathbb{R}^d)$. For each

integer $k \geq 0$, $C_b^k(\mathbb{R}^d)$ denotes the classical space of bounded functions with bounded continuous derivatives up to order k . The space of all bounded uniformly continuous functions on \mathbb{R}^d is denoted by $BUC(\mathbb{R}^d)$. Let $\mathcal{S}(\mathbb{R}^d)$ be the space of Schwartz functions on \mathbb{R}^d . The dual of \mathcal{S} is the space of all tempered distributions $\mathcal{S}'(\mathbb{R}^d)$. The notation \lesssim means $\leq C$ for some multiplicative non-negative constant C , whose value can change from one line to another.

2 Stochastic sewing lemma

Hereafter, $d \geq 1$ is a dimension, $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, $m \geq 2$ is a fixed number, L_m denotes $[L^m(\Omega, \mathcal{F}, \mathbb{P})]^d$. Let us begin with the following observation which will be employed several times. Often the case, one would like to estimate moments of a sum of the form

$$S = \sum_{i=1}^n Z_i,$$

where Z_i 's are some random variables in L_m . Without any additional structure, one at least uses triangle inequality to obtain

$$\|S\|_{L_m} \leq \sum_{i=1}^n \|Z_i\|_{L_m}.$$

Indeed, this kind of estimate is used to obtain the deterministic sewing lemma ([11]). Suppose in addition that there is an increasing sequence of σ -algebras $\mathcal{G}_i \subset \mathcal{F}$ such that $Z_1, \dots, Z_{i-1} \in \mathcal{G}_i$ for every i . Then, using the so-called Doob's decomposition ([9]), we can write

$$S = \sum_{i=1}^n \mathbb{E}^{\mathcal{G}_i} Z_i + \sum_{i=1}^n (Z_i - \mathbb{E}^{\mathcal{G}_i} Z_i) =: S_1 + S_2. \tag{2.1}$$

Hereafter, $\mathbb{E}^{\mathcal{G}}$ denotes the expectation conditioned on a given σ -algebra \mathcal{G} . Estimating S_1 by triangle inequality gives

$$\|S_1\|_{L_m} \leq \sum_{i=1}^n \|\mathbb{E}^{\mathcal{G}_i} Z_i\|_{L_m}. \tag{2.2}$$

S_2 is a sum of martingale differences and can be estimated using Burkholder-Davis-Gundy (BDG) inequality ([3]) and Minkowski inequality,

$$\|S_2\|_{L_m} \leq \kappa_{m,d} \left\| \sum_{i=1}^n |Z_i - \mathbb{E}^{\mathcal{G}_i} Z_i|^2 \right\|_{L_{m/2}}^{\frac{1}{2}} \leq \kappa_{m,d} \left(\sum_{i=1}^n \|Z_i - \mathbb{E}^{\mathcal{G}_i} Z_i\|_{L_m}^2 \right)^{\frac{1}{2}}, \tag{2.3}$$

where $\kappa_{m,d}$ is the constant in BDG inequality in L_m . Hence, we have shown that

$$\|S\|_{L_m} \leq \sum_{i=1}^n \|\mathbb{E}^{\mathcal{G}_i} Z_i\|_{L_m} + \kappa_{m,d} \left(\sum_{i=1}^n \|Z_i - \mathbb{E}^{\mathcal{G}_i} Z_i\|_{L_m}^2 \right)^{\frac{1}{2}}. \tag{2.4}$$

In some cases, it is more convenient to estimate the second sum on the right-hand side further by mean of triangle inequality and contraction property of conditional expectation. This yields the following inequality

$$\|S\|_{L_m} \leq \sum_{i=1}^n \|\mathbb{E}^{\mathcal{G}_i} Z_i\|_{L_m} + 2\kappa_{m,d} \left(\sum_{i=1}^n \|Z_i\|_{L_m}^2 \right)^{\frac{1}{2}}. \tag{2.5}$$

The decomposition (2.1) and inequalities (2.2)-(2.5) are certainly well-known. They appear, for instance, in Davie's [7, pg. 18] and in Delarue and Diel's [8] in an attempt to identify the distributional drift of a diffusion. In the current article, we apply the identity (2.1) and inequalities (2.2), (2.3) in the sewing lemma of [18, 11]. This approach yields a stochastic version of the sewing lemma, Theorem 2.1 below.

Before stating the result, let us introduce some notation which are used throughout the article. Let $\{\mathcal{F}_t\}_{t \geq 0}$ be a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$. We always assume that \mathcal{F}_0 contains \mathbb{P} -null sets, which ensures that modifications of $\{\mathcal{F}_t\}$ -adapted processes are still $\{\mathcal{F}_t\}$ -adapted. Let S, T be fixed non-negative numbers such that $S < T$. We denote by $[S, T]_{\leq}^2$ the simplex $\{(s, t) \in [S, T]^2 : s \leq t\}$. Let $(A_{s,t})_{S \leq s \leq t \leq T}$ be a two-parameter stochastic process with values in \mathbb{R}^d . This means that $A_{s,t}$ is a random variable in \mathbb{R}^d for each (s, t) in $[S, T]_{\leq}^2$. For every $S \leq s \leq u \leq t \leq T$, we define

$$\delta A_{s,u,t} = A_{s,t} - A_{s,u} - A_{u,t}.$$

We say that A is *adapted* to $\{\mathcal{F}_t\}$ if $A_{s,t}$ is \mathcal{F}_t -measurable for every $(s, t) \in [S, T]_{\leq}^2$; A is L_m -*integrable* if $A_{s,t}$ belongs to L_m for every $(s, t) \in [S, T]_{\leq}^2$. Similarly, for a one-parameter stochastic process $(\varphi_t)_{S \leq t \leq T}$ in \mathbb{R}^d , we say that φ is adapted to $\{\mathcal{F}_t\}$ if φ_t is \mathcal{F}_t -measurable for every $t \in [S, T]$; φ is L_m -integrable if φ_t belongs to L_m for every $t \in [S, T]$. We now state our main result. Its proof will be presented later in Section 2.2.

Theorem 2.1 (Stochastic sewing lemma). *Let $m \geq 2$ be a real number and $(A_{s,t})_{S \leq s \leq t \leq T}$ be a two-parameter stochastic process with values in \mathbb{R}^d which is L_m -integrable and adapted to $\{\mathcal{F}_t\}$. Suppose that there are non-negative constants Γ_1, Γ_2 and positive constants $\varepsilon_1, \varepsilon_2$ such that*

$$\|\mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t}\|_{L_m} \leq \Gamma_1 |t - s|^{1+\varepsilon_1} \quad \text{for every } S \leq s \leq u \leq t \leq T, \tag{2.6}$$

and

$$\|\delta A_{s,u,t} - \mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t}\|_{L_m} \leq \Gamma_2 |t - s|^{\frac{1}{2}+\varepsilon_2} \quad \text{for every } S \leq s \leq u \leq t \leq T. \tag{2.7}$$

Then, there exists a unique (up to modifications) stochastic process $(\mathcal{A}_t)_{S \leq t \leq T}$ with values in \mathbb{R}^d satisfying the following properties

(2.1a) $\mathcal{A}_S = 0$, \mathcal{A} is $\{\mathcal{F}_t\}$ -adapted and L_m -integrable,

(2.1b) there are non-negative constants C_1, C_2 such that

$$\|\mathcal{A}_t - \mathcal{A}_s - A_{s,t}\|_{L_m} \leq C_1 |t - s|^{1+\varepsilon_1} + C_2 |t - s|^{\frac{1}{2}+\varepsilon_2} \tag{2.8}$$

and

$$\|\mathbb{E}^{\mathcal{F}_s} (\mathcal{A}_t - \mathcal{A}_s - A_{s,t})\|_{L_m} \leq C_1 |t - s|^{1+\varepsilon_1} \tag{2.9}$$

for every $S \leq s \leq t \leq T$.

The least constants C_1, C_2 are at most $\Gamma_1(1 - 2^{-\varepsilon_1})^{-1}$ and $\kappa_{m,d}\Gamma_2(1 - 2^{-\varepsilon_2})^{-1}$ respectively.

Furthermore, for every fixed $t \in [S, T]$ and any partition $\pi = \{S = t_0 < t_1 < \dots < t_N = t\}$ of $[S, t]$, define the Riemann sum

$$A_t^\pi := \sum_{i=0}^{N-1} A_{t_i, t_{i+1}}.$$

Then $\{A_t^\pi\}_\pi$ converges to \mathcal{A}_t in L_m as the mesh size $|\pi| := \max_i |t_{i+1} - t_i|$ goes to 0.

Remark. Up to a modification of the constants Γ_1, Γ_2 , the two conditions (2.6) and (2.7) are equivalent to the two conditions (2.6) and

$$\|\delta A_{s,u,t}\|_{L_m} \leq \Gamma_2 |t - s|^{\frac{1}{2} + \varepsilon_2} \quad \text{for every } S \leq s \leq t \leq T. \quad (2.10)$$

We have favored (2.7) over (2.10) because in this form, it is easier to derive estimates for the martingale decomposition in Theorem 2.2 (below) from Theorem 2.1.

We also note that no continuity assumption on the map $A : [S, T]_{\leq}^2 \rightarrow L_m$ is imposed in Theorem 2.1. In most known references on the sewing lemma, continuity of A (at least on the diagonal) is usually assumed. An extension of sewing lemma without any continuity assumption on A is due to Yaskov's [44]. We will indeed use part of his arguments in the proof of Theorem 2.1 (see Lemma 2.14 below). This allows us to drop all regularity assumption on the map $A : [S, T]_{\leq}^2 \rightarrow L_m$. This fact complements the original ideas of Gubinelli in [18] that the (sewing) map $A \mapsto ((s, t) \mapsto \mathcal{A}_t - \mathcal{A}_s - A_{s,t})$ is actually a function of δA , and hence, depends solely on the properties of δA . Another illustration of this observation appears in Remark 2.17 below, in which we explain that the integrability and adaptiveness of A assumed in Theorem 2.1 can be replaced by those of δA .

With an additional assumption on A (namely (2.11) below), it is possible to decompose \mathcal{A} into the sum of a martingale \mathcal{M} and a remainder process \mathcal{J} . Such a decomposition is similar to the well-known Doob-Meyer's one. However, in our case, the process \mathcal{J} need not be of bounded variation. To ensure that the decomposition $\mathcal{A} = \mathcal{M} + \mathcal{J}$ is unique, the bounded variation property is replaced by qualitative bounds on the increments of the process \mathcal{J} centered about the process $(s, t) \mapsto \mathbb{E}^{\mathcal{F}_s} A_{s,t}$ (see (2.13) below). One can also give other different characterizations of such decomposition and of \mathcal{M} and \mathcal{J} themselves. These findings are formalized in detail in the next theorem.

Theorem 2.2. *Suppose that the hypotheses of Theorem 2.1 holds. In addition, we assume that there are constants $\Gamma_3 \geq 0$ and $\varepsilon_3 > 0$ such that*

$$\|\mathbb{E}^{\mathcal{F}_s} A_{u,t} - \mathbb{E}^{\mathcal{F}_u} A_{u,t}\|_{L_m} \leq \Gamma_3 |t - s|^{\frac{1}{2} + \varepsilon_3} \quad \text{for every } S \leq s \leq u \leq t \leq T. \quad (2.11)$$

Then, there exist stochastic processes \mathcal{M}, \mathcal{J} and non-negative constants C_1, C_2, C_3 satisfying the following properties

(2.2a) \mathcal{M}, \mathcal{J} are $\{\mathcal{F}_t\}$ -adapted, L_m -integrable and $\mathcal{A}_t = \mathcal{M}_t + \mathcal{J}_t$ a.s. for every $t \in [S, T]$,

(2.2b) $(\mathcal{M}_s)_{S \leq s \leq T}$ is an $\{\mathcal{F}_t\}$ -martingale with $\mathcal{M}_S = 0$,

(2.2c) for any $S \leq s \leq t \leq T$,

$$\|\mathcal{M}_t - \mathcal{M}_s - A_{s,t} + \mathbb{E}^{\mathcal{F}_s} A_{s,t}\|_{L_m} \leq C_2 |t - s|^{\frac{1}{2} + \varepsilon_2} + C_3 |t - s|^{\frac{1}{2} + \varepsilon_3}, \quad (2.12)$$

(2.2d) for any $S \leq s \leq t \leq T$,

$$\|\mathcal{J}_t - \mathcal{J}_s - \mathbb{E}^{\mathcal{F}_s} A_{s,t}\|_{L_m} \leq C_1 |t - s|^{1 + \varepsilon_1} + C_3 |t - s|^{\frac{1}{2} + \varepsilon_3}, \quad (2.13)$$

(2.2e) for any $S \leq s \leq t \leq T$,

$$\|\mathbb{E}^{\mathcal{F}_s} (\mathcal{J}_t - \mathcal{J}_s - A_{s,t})\|_{L_m} \leq C_1 |t - s|^{1 + \varepsilon_1}. \quad (2.14)$$

Given A , we have the following characterizations.

(2.2f) *The pair of processes $(\mathcal{M}, \mathcal{J})$ is characterized uniquely by the set of properties (2.2a), (2.2b), (2.2c) or, alternatively by the set of properties (2.2a), (2.2b), (2.2d).*

(2.2g) The process \mathcal{M} is characterized uniquely by (2.2b) and (2.2c).

(2.2h) The process \mathcal{J} is characterized uniquely by (2.2d), (2.2e) and the fact that \mathcal{J} is $\{\mathcal{F}_t\}$ -adapted and $\mathcal{J}_S = 0$.

The least constants C_1, C_2, C_3 are at most $\Gamma_1(1 - 2^{-\varepsilon_1})^{-1}$, $\kappa_{m,d}\Gamma_2(1 - 2^{-\varepsilon_2})^{-1}$ and $\kappa_{m,d}\Gamma_3(1 - 2^{-\varepsilon_3})^{-1}$ respectively.

Furthermore, for every fixed $t \in [S, T]$ and any partition $\pi = \{S = t_0 < t_1 < \dots < t_N = t\}$ of $[S, t]$, define the Riemann sums

$$M_t^\pi := \sum_{i=0}^{N-1} (A_{t_i, t_{i+1}} - \mathbb{E}^{\mathcal{F}_{t_i}} A_{t_i, t_{i+1}}) \quad \text{and} \quad J_t^\pi := \sum_{i=0}^{N-1} \mathbb{E}^{\mathcal{F}_{t_i}} A_{t_i, t_{i+1}}.$$

Then $\{M_t^\pi\}_\pi$ and $\{J_t^\pi\}_\pi$ converge to \mathcal{M}_t and \mathcal{J}_t respectively in L_m as $|\pi|$ goes to 0.

The above result can be regarded as a consequence of Theorem 2.1. Indeed, we define two-parameter processes J, M by

$$J_{s,t} = \mathbb{E}^{\mathcal{F}_s} A_{s,t} \quad \text{and} \quad M_{s,t} = A_{s,t} - \mathbb{E}^{\mathcal{F}_s} A_{s,t}. \tag{2.15}$$

It is evident that J and M are $\{\mathcal{F}_t\}$ -adapted and L_m -integrable. In addition, for every $s \leq u \leq t$, we have

$$\begin{aligned} \delta J_{s,u,t} &= \mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t} + (\mathbb{E}^{\mathcal{F}_s} A_{u,t} - \mathbb{E}^{\mathcal{F}_u} A_{u,t}), \\ \delta M_{s,u,t} &= (\delta A_{s,u,t} - \mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t}) - (\mathbb{E}^{\mathcal{F}_s} A_{u,t} - \mathbb{E}^{\mathcal{F}_u} A_{u,t}), \end{aligned}$$

and hence

$$\mathbb{E}^{\mathcal{F}_s} \delta J_{s,u,t} = \mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t}, \quad \delta J_{s,u,t} - \mathbb{E}^{\mathcal{F}_s} \delta J_{s,u,t} = \mathbb{E}^{\mathcal{F}_s} A_{u,t} - \mathbb{E}^{\mathcal{F}_u} A_{u,t},$$

and

$$\mathbb{E}^{\mathcal{F}_s} \delta M_{s,u,t} = 0. \tag{2.16}$$

The conditions (2.6), (2.7) and (2.11) on A imply that

$$\|\mathbb{E}^{\mathcal{F}_s} \delta J_{s,u,t}\|_{L_m} \leq \Gamma_1 |t - s|^{1+\varepsilon_1}, \tag{2.17}$$

$$\|\delta J_{s,u,t} - \mathbb{E}^{\mathcal{F}_s} \delta J_{s,u,t}\|_{L_m} \leq \Gamma_3 |t - s|^{\frac{1}{2}+\varepsilon_3}, \tag{2.18}$$

$$\|\delta M_{s,u,t} - \mathbb{E}^{\mathcal{F}_s} \delta M_{s,u,t}\|_{L_m} \leq \Gamma_{2,3} |t - s|^{\frac{1}{2}+\varepsilon_2 \wedge \varepsilon_3}, \tag{2.19}$$

for some constant $\Gamma_{2,3}$ which depends on Γ_2, Γ_3, S, T . From here, by applying Theorem 2.1 to the two-parameter processes J and M respectively, the processes \mathcal{J} and \mathcal{M} can be constructed which are $\{\mathcal{F}_t\}$ -adapted and L_m -integrable. From (2.8), (2.9) and (2.17), (2.18) we can derive the estimates (2.13), (2.14). From (2.8), (2.9) and (2.16), (2.19), we can derive the following estimates

$$\|\mathcal{M}_t - \mathcal{M}_s - A_{s,t} + \mathbb{E}^{\mathcal{F}_s} A_{s,t}\|_{L_m} \leq C |t - s|^{\frac{1}{2}+\varepsilon_2 \wedge \varepsilon_3}. \tag{2.20}$$

and

$$\|\mathbb{E}^{\mathcal{F}_s} (\mathcal{M}_t - \mathcal{M}_s - A_{s,t} + \mathbb{E}^{\mathcal{F}_s} A_{s,t})\|_{L_m} \leq 0 \tag{2.21}$$

for every $s \leq t$. Obviously, $\mathbb{E}^{\mathcal{F}_s} (A_{s,t} - \mathbb{E}^{\mathcal{F}_s} A_{s,t}) = 0$, so (2.21) implies that \mathcal{M} is an $\{\mathcal{F}_t\}$ -martingale. Unfortunately, the estimate (2.20) does not imply (2.12) with the optimal constants stated in Theorem 2.3. This is due to the use of (2.19). A better estimate for δM which should be employed here is the following

$$\|\delta M_{s,u,t} - \mathbb{E}^{\mathcal{F}_s} \delta M_{s,u,t}\|_{L_m} \leq \Gamma_2 |t - s|^{\frac{1}{2}+\varepsilon_2} + \Gamma_3 |t - s|^{\frac{1}{2}+\varepsilon_3}. \tag{2.22}$$

However, the right-hand sides of (2.22) and (2.7) are of different forms, preventing us from applying Theorem 2.1 directly. This is a minor issue and will be resolved in Section 2.2 once we have better understanding on the proof of Theorem 2.1.

In most applications considered herein, it is more convenient to verified a simpler condition than (2.7). This is described in the next result.

Theorem 2.3. *Let $m \geq 2$ be a real number and $(A_{s,t})_{S \leq s \leq t \leq T}$ be a stochastic process with values in \mathbb{R}^d which is L_m -integrable and adapted to $\{\mathcal{F}_t\}$. Assume that there are constants $\Gamma_1, \Gamma_4 \geq 0$ and $\varepsilon_1, \varepsilon_4 > 0$ such that (2.6) holds and*

$$\|A_{s,t}\|_{L_m} \leq \Gamma_4 |t - s|^{\frac{1}{2} + \varepsilon_4} \quad \text{for every } (s, t) \in [S, T]_{\leq}^2. \tag{2.23}$$

Then, there exists a unique (up to modifications) stochastic process $(\mathcal{A}_t)_{S \leq t \leq T}$ with values in \mathbb{R}^d satisfying the following properties

(2.3a) $\mathcal{A}_S = 0$ and for every $t \in [S, T]$, \mathcal{A}_t is \mathcal{F}_t -measurable and L_m -integrable,

(2.3b) there are non-negative constants C_1, C_4 such that for every $(s, t) \in [S, T]_{\leq}^2$, \mathcal{A} satisfies the inequality

$$\|\mathcal{A}_t - \mathcal{A}_s\|_{L_m} \leq C_1 |t - s|^{1 + \varepsilon_1} + C_4 |t - s|^{\frac{1}{2} + \varepsilon_4} \tag{2.24}$$

and inequality (2.9).

The least constants C_1, C_4 are at most $\Gamma_1(1 - 2^{-\varepsilon_1})^{-1}$ and $\Gamma_4(6\kappa_{m,d}(1 - 2^{-\varepsilon_4})^{-1} + 1)$ respectively.

Proof. This result is a direct application of Theorem 2.1. We note that condition (2.23) implies condition (2.7) with $\Gamma_2 = 6\Gamma_4$ and $\varepsilon_2 = \varepsilon_4$. Hence, Theorem 2.1 is applied to construct the process \mathcal{A} satisfying (2.1a) and (2.1b). The estimate (2.8) and condition (2.23) imply (2.24). \square

Remark 2.4. We note that if \mathcal{A} satisfies the hypotheses of Theorem 2.3, then it also satisfies the hypotheses of Theorem 2.2. This is because condition (2.23) also implies condition (2.11) with $\Gamma_3 = 2\Gamma_4$ and $\varepsilon_3 = \varepsilon_4$. Consequently, the process \mathcal{A} in Theorem 2.3 is uniquely decomposed into $\mathcal{M} + \mathcal{J}$ where \mathcal{M}, \mathcal{J} are the processes in Theorem 2.2. It turns out that the martingale part of \mathcal{A} , namely \mathcal{M} , has to be 0. This can be seen by directly checking that $\mathcal{M} \equiv 0$ satisfies (2.12) and hence, by uniqueness, property (2.2g), the martingale part of \mathcal{A} has to vanish. Consequently, for each $t \in [S, T]$, the Riemann sums $\{A_t^\pi\}_\pi$ and $\{J_t^\pi\}_\pi$ both converge to \mathcal{A}_t in L_m while $\{M_t^\pi\}_\pi$ converges to 0 in L_m .

In Theorem 2.1 and its proof (in Section 2.2), it is seen that for each $t \in [S, T]$, \mathcal{A}_t is the limit in L_m of certain Riemann sums. Conversely, starting from the estimates (2.8) and (2.9), one can show the convergence of Riemann sums in L_m , which is the content of the following result. The deterministic counterpart of this result can be found, for instance, in [11, Corollary 2.4] and [18, Corollary 1].

Proposition 2.5. *Let $m \geq 2$ be a real number, $(\varphi_t)_{S \leq t \leq T}$ and $(\mu_{s,t})_{S \leq s \leq t \leq T}$ be $\{\mathcal{F}_t\}$ -adapted and L_m -integrable processes with values in \mathbb{R}^d . Suppose that there are constants $C \geq 0$ and $\varepsilon > 0$ such that*

$$\|\varphi_t - \varphi_s - \mu_{s,t}\|_{L_m} \leq C |t - s|^{\frac{1}{2} + \varepsilon} \tag{2.25}$$

and

$$\|\mathbb{E}^{\mathcal{F}_s}(\varphi_t - \varphi_s - \mu_{s,t})\|_{L_m} \leq C |t - s|^{1 + \varepsilon} \tag{2.26}$$

for every $(s, t) \in [S, T]_{\leq}^2$. For every fixed $(s, t) \in [S, T]_{\leq}^2$ and any partition $\pi = \{s = t_0 < t_1 < \dots < t_N = t\}$ of $[s, t]$, define the Riemann sum

$$\mu_{s,t}^{\pi} := \sum_{i=0}^{N-1} \mu_{t_i, t_{i+1}}.$$

Then $\{\mu_{s,t}^{\pi}\}_{\pi}$ converges to $\varphi_t - \varphi_s$ in L_m as the mesh size $|\pi|$ goes to 0.

Proof. We write

$$\varphi_t - \varphi_s - \mu_{s,t}^{\pi} = \sum_i (\varphi_{t_{i+1}} - \varphi_{t_i} - \mu_{t_i, t_{i+1}})$$

and apply inequality (2.5) to obtain that

$$\begin{aligned} \|\varphi_t - \varphi_s - \mu_{s,t}^{\pi}\|_{L_m} &\leq \sum_i \|\mathbb{E}^{\mathcal{F}_{t_i}}(\varphi_{t_{i+1}} - \varphi_{t_i} - \mu_{t_i, t_{i+1}})\|_{L_m} \\ &\quad + 2\kappa_{m,d} \left(\sum_i \|\varphi_{t_{i+1}} - \varphi_{t_i} - \mu_{t_i, t_{i+1}}\|_{L_m}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

In conjunction with (2.25) and (2.26), the above inequality implies that

$$\|\varphi_t - \varphi_s - \mu_{s,t}^{\pi}\|_{L_m} \lesssim |\pi|^{\varepsilon}.$$

This means that $\lim_{|\pi| \rightarrow 0} \mu_{s,t}^{\pi} = \varphi_t - \varphi_s$ in L_m . □

For later purposes, it is convenient to view the resulting processes in Theorems 2.1, 2.2 and 2.3 as operators whose input is the increment process A . This leads to the following convention.

Definition 2.6. Let t be in $[S, T]$ and let $\pi = \{S = t_0 < \dots < t_N = T\}$ denote a partition of $[S, t]$. Let A_t^{π}, M_t^{π} and J_t^{π} be the Riemann sums defined in Theorems 2.1 and 2.2. We define $\mathcal{I}_t[A], \mathcal{M}_t[A]$ and $\mathcal{J}_t[A]$ respectively by $\lim_{|\pi| \downarrow 0} A_t^{\pi}, \lim_{|\pi| \downarrow 0} M_t^{\pi}$ and $\lim_{|\pi| \downarrow 0} J_t^{\pi}$ whenever these limits exist in probability.

Theorems 2.1, 2.2 and 2.3 provides sufficient conditions for the well-posedness of $\mathcal{I}[A], \mathcal{M}[A]$ and $\mathcal{J}[A]$. In addition, from the discussion succeeding Theorem 2.2, we see that under hypotheses of Theorem 2.2, $\mathcal{J}[A], \mathcal{M}[A]$ and $\mathcal{I}[A]$ are related by

$$\mathcal{J}[A] = \mathcal{J} = \mathcal{I} [(s, t) \mapsto \mathbb{E}^{\mathcal{F}_s} A_{s,t}]$$

and

$$\mathcal{M}[A] = \mathcal{M} = \mathcal{I} [(s, t) \mapsto A_{s,t} - \mathbb{E}^{\mathcal{F}_s} A_{s,t}].$$

It is also possible to write \mathcal{J}, \mathcal{M} as actions of \mathcal{I} on \mathcal{A} . To see this, we need the following observation.

Proposition 2.7. Let $m \geq 2$ be a real number and $(A)_{S \leq s \leq t \leq T}$ be a two-parameter process in \mathbb{R}^d which is $\{\mathcal{F}_t\}$ -adapted and L_m -integrable. Suppose that there are non-negative constants Γ_4, Γ_5 and positive constants $\varepsilon_4, \varepsilon_5$ such that (2.23) holds and

$$\|\mathbb{E}^{\mathcal{F}_s} A_{s,t}\|_{L_m} \leq C_5 |t - s|^{1+\varepsilon_5} \tag{2.27}$$

for every (s, t) in $[S, T]_{\leq}^2$. Then A satisfies all conditions of Theorem 2.1 and $\mathcal{I}[A] = 0$.

Proof. Using (2.23) and (2.27), we can directly verify (2.6), (2.7) and that the zero process satisfies (2.1a) and (2.1b). By uniqueness of Theorem 2.1, this means $\mathcal{I}_t[A] = 0$ a.s. for every $t \in [S, T]$. □

Corollary 2.8. *Suppose that the hypotheses of Theorem 2.2 are satisfied. Let A, M, \mathcal{J} be the processes in Theorem 2.2. For every $s, t \in [S, T]$, define*

$$A_{s,t}^{(1)} = \mathbb{E}^{\mathcal{F}_s}(\mathcal{A}_t - \mathcal{A}_s) \quad \text{and} \quad A_{s,t}^{(2)} = \mathcal{A}_t - \mathcal{A}_s - \mathbb{E}^{\mathcal{F}_s}(\mathcal{A}_t - \mathcal{A}_s).$$

Then $A^{(1)}, A^{(2)}$ satisfy the hypotheses of Theorem 2.1 and $\mathcal{I}[A^{(1)}] = \mathcal{J} = \mathcal{J}[A]$ and $\mathcal{I}[A^{(2)}] = \mathcal{M} = \mathcal{M}[A]$.

Proof. Let J, M be the processes defined in (2.15). This result amounts to verifying conditions (2.23) and (2.27) of Proposition 2.7 for $A^{(1)} - J$ and $A^{(2)} - M$. In both cases, (2.23) and (2.27) are consequences of (2.8) and (2.9). \square

We give an example hinting that the restriction $m \geq 2$ is necessary in Theorem 2.1.

Example 2.9. Let $(N_t, \mathcal{F}_t)_{t \geq 0}$ be a Poisson point process on \mathbb{R} with intensity $\lambda > 0$ (see [28, Definition 3.3]). It is well-known that $\text{Var}(N_t - N_s) = \mathbb{E}(N_t - N_s) = \lambda(t - s)$ and that $(N_t - \lambda t, \mathcal{F}_t)_{t \geq 0}$ is a square integrable martingale. Define $A_{s,t} = N_t - N_s - \lambda(t - s)$. Then $\delta A_{s,u,t} = 0$ and hence we can apply Theorem 2.1, with $m = 2$ for simplicity, to A to find a unique process \mathcal{A} satisfying (2.1a) and (2.1b). Since $\delta A = 0$, we find that C_1 and C_2 can be taken to be 0 in (2.8) and (2.9). This implies that $\mathcal{A}_t = N_t - \lambda t$ for every t .

On the other hand, we also have

$$\mathbb{E}^{\mathcal{F}_s}(A_{s,t}) = 0 \quad \text{and} \quad \|A_{s,t}\|_{L_n} \leq (2\lambda)^{\frac{1}{n}}(t - s)^{\frac{1}{n}} \tag{2.28}$$

for every $s \leq t$ and $n \in [1, 2]$. In the above, the identity is obvious while the inequality is derived by interpolating

$$\mathbb{E}|A_{s,t}|^2 = \lambda(t - s) \quad \text{and} \quad \mathbb{E}|A_{s,t}| \leq 2\lambda(t - s).$$

Note that when $n \in [1, 2)$, $\frac{1}{n} > \frac{1}{2}$. This means that Theorem 2.1 does not hold for $m = n \in [1, 2)$. For if it were true, (2.28) and Proposition 2.7 (a consequence of Theorem 2.1) would imply that $\mathcal{A} \equiv 0$. In other words, we would have $N_t - \lambda t = 0$ for all t , which is a contradiction.

2.1 Relations with Itô calculus

To see that the conditions (2.6), (2.7) and (2.11) are natural, let us see how the stochastic sewing lemma is related to Itô calculus. This is established through the following examples.

Example 2.10. Let B be a standard Brownian motion in \mathbb{R}^d with respect to a filtration $\{\mathcal{F}_t\}$. Assume that $\{\mathcal{F}_0\}$ contains \mathbb{P} -null sets. We wish to define the Itô integral $\int_0^T f(B_s) \otimes dB_s$, where $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a Hölder continuous function with exponent $\tau \in (0, 1]$. We define $A_{s,t} = f(B_s) \otimes (B_t - B_s)$. Then for every $s \leq u \leq t$,

$$\delta A_{s,u,t} = -[f(B_u) - f(B_s)] \otimes (B_t - B_u).$$

It follows that for every $m \geq 2$,

$$\|\delta A_{s,u,t}\|_{L_m} \leq \|f\|_{C^\tau} |t - s|^{\frac{1}{2} + \frac{\tau}{2}} \quad \text{and} \quad \mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t} = 0.$$

In other words, A satisfies conditions (2.6) and (2.7), respectively with $\Gamma_1 = 0$, $\Gamma_2 = \|f\|_{C^\tau}$ and $\varepsilon_2 = \frac{\tau}{2}$. By Theorem 2.1, we can define

$$\int_0^T f(B_s) \otimes dB_s := \lim_{\max_i |t_{i+1} - t_i| \downarrow 0} \sum_i f(B_{t_i}) \otimes (B_{t_{i+1}} - B_{t_i}) \quad \text{in } L_m,$$

where $\{t_i\}$ is any partition of $[0, T]$. The estimates (2.8) and (2.9), respectively, become

$$\left\| \int_s^t f(B_r) \otimes dB_r - f(B_s) \otimes (B_t - B_s) \right\|_{L_m} \leq \kappa_{m,d} (1 - 2^{-\frac{\alpha}{2}})^{-1} \|f\|_{C^r} |t - s|^{\frac{1}{2} + \frac{\alpha}{2}}$$

and

$$\left\| \mathbb{E}^{\mathcal{F}_s} \int_s^t f(B_r) \otimes dB_r - \mathbb{E}^{\mathcal{F}_s} f(B_s) \otimes (B_t - B_s) \right\|_{L_m} \leq 0.$$

Since $\mathbb{E}^{\mathcal{F}_s} f(B_s) \otimes (B_t - B_s) = 0$, the previous estimate implies that $\int_0^\cdot f(B_r) \otimes dB_r$ is a martingale.

Alternatively, we can also see that $\int_0^\cdot f(B_r) \otimes dB_r$ is a martingale from Theorem 2.2. Indeed, we note that $\mathbb{E}^{\mathcal{F}_s} A_{s,t} = 0$ for every $s \leq t$. In particular, (2.11) is satisfied with $\Gamma_3 = 0$. By Theorem 2.2, we have the decomposition $\int_0^t f(B_s) \otimes dB_s = \mathcal{M}_t + \mathcal{J}_t$. We observe that the process $\mathcal{J} \equiv 0$ satisfies (2.13). Hence, by (2.2f), $\int_0^\cdot f(B_s) \otimes dB_s = \mathcal{M}$, which is a martingale.

Lastly, we point out an interesting implication of the uniqueness part of Theorem 2.1. The Itô integral $\int_0^\cdot f(B) \otimes dB$ is the unique $\{\mathcal{F}_t\}$ -adapted process $\varphi : [0, T] \rightarrow L_m$ such that $\varphi_0 = 0$ and

$$\|\varphi_t - \varphi_s - f(B_s) \otimes (B_t - B_s)\|_{L_m} \lesssim |t - s|^{\frac{1}{2} + \varepsilon}, \tag{2.29}$$

$$\|\mathbb{E}^{\mathcal{F}_s}(\varphi_t - \varphi_s)\|_{L_m} \lesssim |t - s|^{1 + \varepsilon} \tag{2.30}$$

for every $s \leq t$ for some $\varepsilon > 0$.

Example 2.11 (Quadratic variation). Let M be a L_4 -integrable $\{\mathcal{F}_t\}$ -martingale in \mathbb{R}^d . Assume that \mathcal{F}_0 contains \mathbb{P} -null sets and M satisfies

$$\|M_{s,u} \otimes M_{u,t}\|_{L_2} \leq C |t - s|^{\frac{1}{2} + \varepsilon} \quad \forall s \leq u \leq t \tag{2.31}$$

for some constants $\varepsilon > 0$ and $C > 0$. Here, we adopt the notation $M_{s,t} = M_t - M_s$. We consider $A_{s,t} = M_{s,t} \otimes M_{s,t}$, which is a random element in $\mathbb{R}^d \otimes \mathbb{R}^d$ for every fixed $s \leq t$. It is straightforward to verify that for every $s \leq u \leq t$

$$\delta A_{s,u,t} = M_{s,u} \otimes M_{u,t} + M_{u,t} \otimes M_{s,u}.$$

It follows that $\|\delta A_{s,u,t}\|_{L_2} \leq 2C |t - s|^{\frac{1}{2} + \varepsilon}$ and $\mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t} = 0$. By Theorem 2.1, there exists a unique adapted process, denoted by $[M] : [0, T] \rightarrow [L^2(\Omega)]^{d \times d}$, such that $[M]_0 = 0$,

$$\|[M]_{s,t} - M_{s,t} \otimes M_{s,t}\|_{L_2} \lesssim |t - s|^{\frac{1}{2} + \varepsilon} \quad \text{and} \quad \mathbb{E}^{\mathcal{F}_s} [M]_{s,t} = \mathbb{E}^{\mathcal{F}_s} (M_{s,t} \otimes M_{s,t}) \tag{2.32}$$

for every $s \leq t$. Here, we use the notation $[M]_{s,t} = [M]_t - [M]_s$. In addition, we have

$$[M]_t = \lim_{\max_i |t_{i+1} - t_i| \downarrow 0} \sum_i M_{t_i, t_{i+1}} \otimes M_{t_i, t_{i+1}} \quad \text{in } L_2$$

for every partition $\{t_i\}$ of $[0, t]$ which shows that $[M]$ (defined via Theorem 2.1 as above) is indeed the quadratic variation of M (see for instance [39, Chapter 1, (2.3)]).

Observing that $\mathbb{E}^{\mathcal{F}_s} (M_{s,t} \otimes M_{s,t}) = \mathbb{E}^{\mathcal{F}_s} (M_t \otimes M_t - M_s \otimes M_s)$, the identity in (2.32) implies that $M \otimes M - [M]$ is a martingale.

Finally, for L_4 -integrable martingales satisfying (2.31), the quadratic variation $[M]$ is the unique adapted process $\varphi : [0, T] \rightarrow [L^2(\Omega)]^{d \times d}$ such that

$$\|\varphi_t - \varphi_s - M_{s,t} \otimes M_{s,t}\|_{L_2} \lesssim |t - s|^{\frac{1}{2} + \varepsilon}, \tag{2.33}$$

$$\|\mathbb{E}^{\mathcal{F}_s}(\varphi_t - \varphi_s - M_{s,t} \otimes M_{s,t})\|_{L_2} \lesssim |t - s|^{1 + \varepsilon} \tag{2.34}$$

for every $s \leq t$ and for some $\varepsilon > 0$.

We illustrate the usefulness of the previous characterization in computing the quadratic variations by considering two specific cases of Brownian motion and compensated Poisson process.

In the case when $M = B$ is a standard Brownian motion in \mathbb{R}^d , we have $\|B_{s,t} \otimes B_{s,t}\|_{L_2} \lesssim \|B_{s,t}\|_{L_4} \lesssim |t-s|$ and $\mathbb{E}^{\mathcal{F}_s}(B_{s,t} \otimes B_{s,t}) = (t-s)I_d$, where I_d is the $d \times d$ -identity matrix. Hence, from (2.33) and (2.34), the quadratic variation of B is the unique adapted process $\varphi : [0, T] \rightarrow [L^2(\Omega)]^{d \times d}$ such that

$$\|\varphi_t - \varphi_s\|_{L_2} \lesssim |t-s|^{\frac{1}{2}+\varepsilon} \quad \text{and} \quad \|\mathbb{E}^{\mathcal{F}_s}(\varphi_t - \varphi_s - (t-s)I_d)\|_{L_2} \lesssim |t-s|^{1+\varepsilon}$$

for every $s \leq t$ and for some $\varepsilon > 0$. It is evident that the process $\varphi_t = tI_d$ satisfies the above inequalities. By uniqueness, $[B]_t = tI_d$.

Consider the case when $M = \bar{N}$ is a compensated Poisson point process on \mathbb{R} with intensity λ . This means that $\bar{N}_t = N_t - \lambda t$ where $(N_t, \mathcal{F}_t)_{t \geq 0}$ is a Poisson point process on \mathbb{R} with intensity $\lambda > 0$ (see [28, Definition 3.3]). By direct calculations, we have $\mathbb{E}(\bar{N}_{s,u}\bar{N}_{u,t})^2 = \lambda^2(u-s)(t-u)$, $\mathbb{E}^{\mathcal{F}_s}(\bar{N}_{s,t}\bar{N}_{s,t}) = \lambda(t-s)$ and $\mathbb{E}(N_{s,t} - \bar{N}_{s,t}\bar{N}_{s,t})^2 = 2\lambda^2(t-s)^2$. We see that condition (2.31) is satisfied and hence, from (2.33) and (2.34), the quadratic variation of \bar{N} is the unique adapted process $\varphi : [0, T] \rightarrow L^2(\Omega)$ such that

$$\|\varphi_t - \varphi_s - N_{s,t}\|_{L_2} \lesssim |t-s|^{\frac{1}{2}+\varepsilon} \quad \text{and} \quad \|\mathbb{E}^{\mathcal{F}_s}(\varphi_t - \varphi_s - \lambda(t-s))\|_{L_2} \lesssim |t-s|^{1+\varepsilon}$$

for every $s \leq t$ and for some $\varepsilon > 0$. It is evident that the process $\varphi_t = N_t$ satisfies the above inequalities. By uniqueness, $[\bar{N}]_t = N_t$.

Example 2.12 (Itô formula). Let M be a martingale such that

$$\|M_{s,t}\|_{L_6} \leq C|t-s|^{\frac{1}{3}+\varepsilon} \quad \forall (s,t) \in [0,1]_{\leq}^2 \tag{2.35}$$

for some constants $\varepsilon > 0$ and $C \geq 0$. We recall the notation $M_{s,t} = M_t - M_s$ used in the previous example. From Cauchy-Schwarz inequality, it is easy to see that (2.35) implies condition (2.31). Hence the quadratic variation of M , denoted by $[M]$, is well-defined by Example 2.11 and satisfies

$$\|[M]_{s,t}\|_{L_2} \lesssim |t-s|^{\frac{2}{3}+2\varepsilon}. \tag{2.36}$$

Let f be a function in $C_b^3(\mathbb{R}^d)$. Let us explain how to obtain the following Itô's formula, which is well-known, by stochastic sewing lemma

$$f(M_1) - f(M_0) = \int_0^1 \langle \nabla f(M_s), dM_s \rangle + \frac{1}{2} \int_0^1 \langle \nabla^2 f(M_s), d[M]_s \rangle.$$

In the right-hand side above, $\langle \cdot, \cdot \rangle$ denotes the inner product, the former integral is an Itô integral and the later integral is a Young integral in $L^1(\Omega)$.

We consider $A_{s,t} = f(M_t) - f(M_s)$. It is obvious that $f(M_1) - f(M_0) = \mathcal{I}[A]$, where \mathcal{I} is defined in Definition 2.6. By Taylor expansion, we have

$$\begin{aligned} A_{s,t} &= \langle \nabla f(M_s), M_{s,t} \rangle + \frac{1}{2} \langle \nabla^2 f(M_s), M_{s,t} \otimes M_{s,t} \rangle + \mathcal{O}(|M_{s,t}|^3) \\ &=: A_{s,t}^{(1)} + A_{s,t}^{(2)} + A_{s,t}^{(3)}. \end{aligned}$$

where $\mathcal{O}(h)$ is some quantity such that $\|\mathcal{O}(h)\|_{L_2} \lesssim h$. Obviously, $\mathcal{I}[A] = \mathcal{I}[A^{(1)}] + \mathcal{I}[A^{(2)}] + \mathcal{I}[A^{(3)}]$ whenever each term on the right-hand side exists, hence, it suffices to show that $\mathcal{I}[A^{(1)}] = \int_0^1 \langle \nabla f(M_s), dM_s \rangle$, $\mathcal{I}[A^{(2)}] = \frac{1}{2} \int_0^1 \langle \nabla^2 f(M_s), d[M]_s \rangle$ and $\mathcal{I}[A^{(3)}] = 0$. Reasoning as in Example (2.10), it is straightforward to see that $\mathcal{I}[A^{(1)}] = \int_0^1 \langle \nabla f(M_s), dM_s \rangle$. To compute $\mathcal{I}[A^{(2)}]$, we write

$$A_{s,t}^{(2)} = \frac{1}{2} \langle \nabla^2 f(M_s), [M]_{s,t} \rangle + \frac{1}{2} \langle \nabla^2 f(M_s), [M]_{s,t} - M_{s,t} \otimes M_{s,t} \rangle =: A_{s,t}^{(4)} + A_{s,t}^{(5)}.$$

From $\delta A_{s,u,t}^{(4)} = -\frac{1}{2} \langle \nabla^2 f(M_u) - \nabla^2 f(M_s), [M]_{u,t} \rangle$, applying Cauchy-Schwarz inequality, (2.35) and (2.36), we obtain

$$\|\delta A_{s,u,t}^{(4)}\|_{L_1} \lesssim |\nabla^3 f|_\infty \|M_{s,u}\|_{L_2} \|[M]_{u,t}\|_{L_2} \lesssim |\nabla^3 f|_\infty |t-s|^{1+3\varepsilon}.$$

Applying the sewing lemma ([11, Lemma 2.1]), we see that $\mathcal{I}[A^{(4)}] = \int_0^\cdot f(M_s) d[M]_s$. From (2.32), we see that

$$\|A_{s,t}^{(5)}\|_{L_2} \lesssim |\nabla^2 f|_\infty |t-s|^{\frac{1}{2}+\bar{\varepsilon}} \quad \text{and} \quad \mathbb{E}^{\mathcal{F}_s} A_{s,t}^{(5)} = 0$$

for some $\bar{\varepsilon} > 0$. By Proposition 2.7, $\mathcal{I}[A^{(5)}] = 0$ and hence $\mathcal{I}[A^{(2)}] = \int_0^\cdot f(M_s) d[M]_s$. Finally, we note that $\|A_{s,t}^{(3)}\|_{L_2} \lesssim \|M_{s,t}\|_{L_6}^3 \lesssim |t-s|^{1+3\varepsilon}$ and apply Proposition 2.7 once again to see that $\mathcal{I}[A^{(3)}] = 0$.

2.2 Proofs

We present herein the proofs of Theorems 2.1 and 2.2. We first show in Lemma 2.13 two uniform estimates for the Riemann sums along dyadic partitions, which are straightforward applications of inequalities (2.2) and (2.3). In Lemma 2.15, similar uniform estimates for Riemann sums along any partitions are obtained using a dyadic allocation procedure adapted from [44, Lemma 2.2]. The procedure is described in Lemma 2.14, which asserts that every partition can be arranged into dyadic partitions so that the estimates for Riemann sums along dyadic partitions can be carried over more or less in the same fashion.

Lemma 2.13. *Let $(A_{s,t})_{S \leq s \leq t \leq T}$ be a two-parameter process in \mathbb{R}^d which is L_m -integrable and adapted to $\{\mathcal{F}_t\}$. We assume that there exist constants $\Gamma_1, \Gamma_2, \Gamma_3 \geq 0$ and $\varepsilon_1, \varepsilon_2, \varepsilon_3 > 0$ such that (2.6) holds and*

$$\|\delta A_{s,u,t} - \mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t}\|_{L_m} \leq \Gamma_2 |t-s|^{\frac{1}{2}+\varepsilon_2} + \Gamma_3 |t-s|^{\frac{1}{2}+\varepsilon_3} \tag{2.37}$$

for every $S \leq s \leq u \leq t \leq T$. For each (s, t) in $[S, T]_{\leq}^2$ and each integer $n \geq 1$, let $\pi_{s,t}^n = \{s = t_0^n < t_1^n < \dots < t_{2^n}^n = t\}$ be the dyadic partition of $[s, t]$ with mesh size $|\pi_{s,t}^n| = |t_i^n - t_{i+1}^n| = 2^{-n}(t-s)$. Define

$$A_{s,t}^n = \sum_{i=0}^{2^n-1} A_{t_i^n, t_{i+1}^n}. \tag{2.38}$$

Then for every $n \geq 1$, we have

$$\|A_{s,t}^n - A_{s,t}\|_{L_m} \leq \frac{\Gamma_1}{1-2^{-\varepsilon_1}} |t-s|^{1+\varepsilon_1} + \frac{\kappa_{m,d} \Gamma_2}{1-2^{-\varepsilon_2}} |t-s|^{\frac{1}{2}+\varepsilon_2} + \frac{\kappa_{m,d} \Gamma_3}{1-2^{-\varepsilon_3}} |t-s|^{\frac{1}{2}+\varepsilon_3}, \tag{2.39}$$

and

$$\|\mathbb{E}^{\mathcal{F}_s}(A_{s,t}^n - A_{s,t})\|_{L_m} \leq \frac{\Gamma_1}{1-2^{-\varepsilon_1}} |t-s|^{1+\varepsilon_1}. \tag{2.40}$$

Proof. For each k, i , let u_i^k be the midpoint of $[t_i^k, t_{i+1}^k]$. We have

$$A_{s,t}^k - A_{s,t}^{k+1} = \sum_{i=0}^{2^k-1} \delta A_{t_i^k, u_i^k, t_{i+1}^k} = I_1 + I_2 \tag{2.41}$$

where

$$I_1 = \sum_{i=0}^{2^k-1} \mathbb{E}^{\mathcal{F}_{t_i^k}} \delta A_{t_i^k, u_i^k, t_{i+1}^k} \quad \text{and} \quad I_2 = \sum_{i=0}^{2^k-1} \left(\delta A_{t_i^k, u_i^k, t_{i+1}^k} - \mathbb{E}^{\mathcal{F}_{t_i^k}} \delta A_{t_i^k, u_i^k, t_{i+1}^k} \right). \tag{2.42}$$

Observe that $A_{s,t}^k - A_{s,t}^{k+1}$, I_1, I_2 correspond respectively to S, S_1, S_2 in (2.1). We use (2.2) and (2.6) to estimate I_1 , which yields

$$\begin{aligned} \|I_1\|_{L_m} &\leq \sum_{i=0}^{2^k-1} \|\mathbb{E}^{\mathcal{F}_{t_i^k}} \delta A_{t_i^k, u_i^k, t_{i+1}^k}\|_{L_m} \\ &\leq \Gamma_1 |t-s|^{1+\varepsilon_1} \sum_{i=0}^{2^k-1} 2^{-k(1+\varepsilon_1)} = \Gamma_1 |t-s|^{1+\varepsilon_1} 2^{-k\varepsilon_1}. \end{aligned} \tag{2.43}$$

Similarly, we use (2.3), (2.37) and Minkowski inequality to estimate I_2 , which gives

$$\begin{aligned} \|I_2\|_{L_m} &\leq \kappa_{m,d} \left(\sum_{i=0}^{2^k-1} \|\delta A_{t_i^k, u_i^k, t_{i+1}^k} - \mathbb{E}^{\mathcal{F}_{t_i^k}} \delta A_{t_i^k, u_i^k, t_{i+1}^k}\|_{L_m}^2 \right)^{\frac{1}{2}} \\ &\leq \kappa_{m,d} \left(\Gamma_2^2 |t-s|^{1+2\varepsilon_2} \sum_{i=0}^{2^k-1} 2^{-k(1+2\varepsilon_2)} \right)^{\frac{1}{2}} + \kappa_{m,d} \left(\Gamma_3^2 |t-s|^{1+2\varepsilon_3} \sum_{i=0}^{2^k-1} 2^{-k(1+2\varepsilon_3)} \right)^{\frac{1}{2}} \\ &= \kappa_{m,d} \Gamma_2 |t-s|^{\frac{1}{2}+\varepsilon_2} 2^{-k\varepsilon_2} + \kappa_{m,d} \Gamma_3 |t-s|^{\frac{1}{2}+\varepsilon_3} 2^{-k\varepsilon_3}. \end{aligned} \tag{2.44}$$

Hence, we have shown that

$$\begin{aligned} \|A_{s,t}^k - A_{s,t}^{k+1}\|_{L_m} &\leq \Gamma_1 |t-s|^{1+\varepsilon_1} 2^{-k\varepsilon_1} \\ &\quad + \kappa_{m,d} \Gamma_2 |t-s|^{\frac{1}{2}+\varepsilon_2} 2^{-k\varepsilon_2} + \kappa_{m,d} \Gamma_3 |t-s|^{\frac{1}{2}+\varepsilon_3} 2^{-k\varepsilon_3}. \end{aligned} \tag{2.45}$$

Using triangle inequality, we have

$$\|A_{s,t}^n - A_{s,t}\|_{L_m} \leq \sum_{k=0}^{\infty} \|A_{s,t}^k - A_{s,t}^{k+1}\|_{L_m}.$$

Taking into account the estimate (2.45) and summing up the resulting geometric series in k , we obtain (2.39) from the above inequality. Similarly, using triangle inequality, we have

$$\|\mathbb{E}^{\mathcal{F}_s}(A_{s,t}^n - A_{s,t})\|_{L_m} \leq \sum_{k=0}^{\infty} \|\mathbb{E}^{\mathcal{F}_s}(A_{s,t}^k - A_{s,t}^{k+1})\|_{L_m}.$$

To estimate the right-hand side above, we use (2.41), (2.43) and the fact that $\mathbb{E}^{\mathcal{F}_s} I_2 = 0$ to obtain that

$$\|\mathbb{E}^{\mathcal{F}_s}(A_{s,t}^k - A_{s,t}^{k+1})\|_{L_m} = \|\mathbb{E}^{\mathcal{F}_s} I_1\|_{L_m} \leq \Gamma_1 |t-s|^{1+\varepsilon_1} 2^{-n\varepsilon_1}.$$

Combining the previous inequalities together and sum up the resulting geometric series in k yields (2.40). □

To obtain similar estimates to (2.39) and (2.40) for arbitrary partitions, the following result, which is adapted from Yaskov’s [44, Lemma 2.2], is needed.

Lemma 2.14 (Dyadic allocation). *Let $(A_{s,t})_{S \leq s \leq t \leq T}$ be a two-parameter process in \mathbb{R}^d which is L_m -integrable and adapted to $\{\mathcal{F}_t\}$. We assume that $A_{s,s} = 0$ for every $s \in [S, T]$. Let s, t be in $[S, T]$ with $s < t$. For each n , let $\{t_i^n\}_{i=0}^{2^n}$ be the dyadic partition of $[s, t]$ of mesh size $2^{-n}(t-s)$, that is $t_i^n = s + i2^{-n}(t-s)$ for each n, i . Then, for every $N \geq 0$ and every $s \leq t_0 < \dots < t_N \leq t$, there exist a positive integer n_0 and random variables R_i^n , $i = 0, \dots, 2^n - 1$, $n \geq 0$, such that*

- (i) $R_i^n = 0$ for every $n \geq n_0$ and every i ;

(ii) for each n, i , there exist four (not necessarily distinct) points $s_1^{n,i} \leq s_2^{n,i} \leq s_3^{n,i} \leq s_4^{n,i}$ in $[t_i^n, t_{i+1}^n]$ so that

$$R_i^n = A_{s_1^{n,i}, s_2^{n,i}} + A_{s_2^{n,i}, s_3^{n,i}} + A_{s_3^{n,i}, s_4^{n,i}} - A_{s_1^{n,i}, s_4^{n,i}} \tag{2.46}$$

(iii) the following identity holds

$$\sum_{i=0}^{N-1} A_{t_i, t_{i+1}} - A_{t_0, t_N} = \sum_{n \geq 0} \sum_{i=0}^{2^n-1} R_i^n \tag{2.47}$$

Proof. For each collection $\pi = \{s_i\}_{i=0}^K$ we define

$$I(\pi) = \sum_{i=0}^{K-1} A_{s_i, s_{i+1}} - A_{s_0, s_K} \quad \text{if } K \geq 1$$

and $I(\pi) = 0$ if either $K = 0$ or π is empty. For any two finite collections π_1, π_2 , define

$$\delta I(\pi_1, \pi_2) = I(\pi_1 \cup \pi_2) - I(\pi_1) - I(\pi_2) \tag{2.48}$$

Put $\pi_0^0 = \{t_i\}_{i=0}^N$, which is a subset of $[t_0^0, t_1^0]$. The main idea of the proof is to allocate the elements of π_0^0 into the dyadic subintervals of $[s, t]$ while keeping track of the resulting changes in $I(\pi_0^0)$ during the process. For each $n \geq 1$, define

$$\pi_{2^{n-1}}^n = \pi_0^0 \cap [t_{2^{n-1}}^n, t_{2^n}^n] \quad \text{and} \quad \pi_i^n = \pi_0^0 \cap [t_i^n, t_{i+1}^n] \quad \text{for } i = 0, \dots, 2^n - 2.$$

For each $n \geq 0$ and $i = 0, \dots, 2^n - 1$, define

$$R_i^n := \delta I(\pi_{2i}^{n+1}, \pi_{2i+1}^{n+1}) = I(\pi_i^n) - I(\pi_{2i}^{n+1}) - I(\pi_{2i+1}^{n+1}) \tag{2.49}$$

where the second identity comes from the fact that $\pi_i^n = \pi_{2i}^{n+1} \cup \pi_{2i+1}^{n+1}$. We verify that the random variables $\{R_i^n\}_{n,i}$ satisfy (i)-(iii).

Since π_0^0 is a finite set, there exists a finite integer $n_0 \geq 1$ so that $[t_i^n, t_{i+1}^n] \cap \pi_0^0$ contains at most one point for every $n \geq n_0$ and every $i = 0, \dots, 2^n - 1$. Hence, when $n \geq n_0$, we have $I(\pi_i^n) = 0$ and $R_i^n = 0$ for every i . This shows (i).

If either π_{2i}^{n+1} or π_{2i+1}^{n+1} is empty, then $R_i^n = 0$ and (2.46) is satisfied with $s_j^{n,i} = t_i^n$ for $j = 1, 2, 3, 4$. In the case when π_{2i}^{n+1} and π_{2i+1}^{n+1} are not empty, we define

$$s_1^{n,i} = \min \pi_{2i}^{n+1}, \quad s_2^{n,i} = \max \pi_{2i}^{n+1}, \quad s_3^{n,i} = \min \pi_{2i+1}^{n+1} \quad \text{and} \quad s_4^{n,i} = \max \pi_{2i+1}^{n+1}.$$

Here, \min (respectively \max) of a nonempty finite set F is the smallest (respectively largest) element of F . We derive (2.46) from (2.49) and the definition of I at the beginning of the proof. Hence, (ii) is verified.

Lastly, to show (iii), we apply (2.49) recursively to see that

$$\begin{aligned} I(\pi_0^0) &= I(\pi_0^1) + I(\pi_1^1) + R_0^0 \\ &= \sum_{i=0}^{2^n-1} I(\pi_i^n) + \sum_{k=0}^{n-1} \sum_{i=0}^{2^k-1} R_i^k \quad \text{for every } n \geq 1. \end{aligned}$$

Since $I(\pi_i^n) = 0$ as soon as $n \geq n_0$, the previous identity implies (2.47). This completes the proof. \square

When combined with estimates (2.6) and (2.7) on δA , the previous lemma yields uniform estimates on Riemann sums along arbitrary partitions.

Lemma 2.15. *Let A be the process in Theorem 2.1. Then for every $N \geq 0$ and every $S \leq t_0 < \dots < t_N \leq T$, we have*

$$\|\mathbb{E}^{\mathcal{F}_{t_0}} \left(\sum_{i=0}^{N-1} A_{t_i, t_{i+1}} - A_{t_0, t_N} \right)\|_{L_m} \leq \frac{2\Gamma_1}{1 - 2^{-\varepsilon_1}} |t_N - t_0|^{1+\varepsilon_1} \quad (2.50)$$

and

$$\left\| \sum_{i=0}^{N-1} A_{t_i, t_{i+1}} - A_{t_0, t_N} \right\|_{L_m} \leq \frac{2\Gamma_1}{1 - 2^{-\varepsilon_1}} |t_N - t_0|^{1+\varepsilon_1} + \frac{2\kappa_{m,d}\Gamma_2}{1 - 2^{-\varepsilon_2}} |t_N - t_0|^{\frac{1}{2}+\varepsilon_2}. \quad (2.51)$$

Proof. Put $s = t_0$ and $t = t_N$. Applying Lemma 2.14, we can find random variables R_i^n , $i = 0, \dots, 2^n - 1$, $n \geq 0$ which satisfy properties (i)-(iii) stated there.

Let $n \geq 0$ be fixed. Define $\mathcal{G}_{2^n}^n = \mathcal{F}_{t_{2^n}^n}$ and $\mathcal{G}_i^n = \mathcal{F}_{s_1^{n,i}}$ for each $i = 0, \dots, 2^n - 1$. We recall that $s_1^{n,i}$ is defined in (ii). The sequence $\{\mathcal{G}_i^n\}_{i=0}^{2^n}$ forms a filtration such that R_i^n is \mathcal{G}_{i+1}^n -measurable for every $i = 0, \dots, 2^n - 1$. The formula (2.46) can be written as

$$R_i^n = -\delta A_{s_1^{n,i}, s_2^{n,i}, s_3^{n,i}} - \delta A_{s_1^{n,i}, s_3^{n,i}, s_4^{n,i}}$$

Applying the conditions (2.6) and (2.7), we obtain from the previous identity that

$$\|\mathbb{E}^{\mathcal{G}_i^n} R_i^n\|_{L_m} \leq 2\Gamma_1 |t_{i+1}^n - t_i^n|^{1+\varepsilon_1} \quad \text{and} \quad \|R_i^n - \mathbb{E}^{\mathcal{G}_i^n} R_i^n\|_{L_m} \leq 2\Gamma_2 |t_{i+1}^n - t_i^n|^{\frac{1}{2}+\varepsilon_2}. \quad (2.52)$$

For each $n \geq 0$, we apply inequality (2.4) to obtain that

$$\left\| \sum_{i=0}^{2^n-1} R_i^n \right\|_{L_m} \leq \sum_{i=0}^{2^n-1} \|\mathbb{E}^{\mathcal{G}_i^n} R_i^n\|_{L_m} + \kappa_{m,d} \left(\sum_{i=0}^{2^n-1} \|R_i^n - \mathbb{E}^{\mathcal{G}_i^n} R_i^n\|_{L_m}^2 \right)^{\frac{1}{2}}.$$

The bounds in (2.52) are applied to estimate each sum on the right-hand side above

$$\sum_{i=0}^{2^n-1} \|\mathbb{E}^{\mathcal{G}_i^n} R_i^n\|_{L_m} \leq 2\Gamma_1 |t - s|^{1+\varepsilon_1} \sum_{i=0}^{2^n-1} 2^{-n(1+\varepsilon_1)} = 2\Gamma_1 |t - s|^{1+\varepsilon_1} 2^{-n\varepsilon_1} \quad (2.53)$$

and

$$\sum_{i=0}^{2^n-1} \|R_i^n - \mathbb{E}^{\mathcal{G}_i^n} R_i^n\|_{L_m}^2 \leq 4\Gamma_2^2 |t - s|^{1+2\varepsilon_2} \sum_{i=0}^{2^n-1} 2^{-n(1+2\varepsilon_2)} = \left(2\Gamma_2 |t - s|^{\frac{1}{2}+\varepsilon_2} 2^{-n\varepsilon_2} \right)^2.$$

Combining the previous three estimates yields

$$\left\| \sum_{i=0}^{2^n-1} R_i^n \right\|_{L_m} \leq 2\Gamma_1 |t - s|^{1+\varepsilon_1} 2^{-n\varepsilon_1} + 2\kappa_{m,d}\Gamma_2 |t - s|^{\frac{1}{2}+\varepsilon_2} 2^{-n\varepsilon_2}. \quad (2.54)$$

From (2.47), applying triangle inequality and contraction property of conditional expectation (noting that $\mathcal{F}_{t_0} \subset \mathcal{G}_i^n$), we see that

$$\|\mathbb{E}^{\mathcal{F}_{t_0}} \left(\sum_{i=0}^{N-1} A_{t_i, t_{i+1}} - A_{t_0, t_N} \right)\|_{L_m} \leq \sum_{n \geq 0} \sum_{i=0}^{2^n-1} \|\mathbb{E}^{\mathcal{G}_i^n} R_i^n\|_{L_m}$$

and

$$\left\| \sum_{i=0}^{N-1} A_{t_i, t_{i+1}} - A_{t_0, t_N} \right\|_{L_m} \leq \sum_{n \geq 0} \left\| \sum_{i=0}^{2^n-1} R_i^n \right\|_{L_m}.$$

We apply inequalities (2.53) and (2.54) to the previous two estimates and compute the resulting geometric sums in n to obtain (2.50) and (2.51). \square

Remark 2.16. Of course, one could consider a more general situation where conditions (2.6) and (2.7) are replaced by

$$\|\mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t}\|_{L_m} \leq \rho_1(s,t) \quad \text{and} \quad \|\delta A_{s,u,t} - \mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t}\|_{L_m} \leq \rho_2(s,t) \quad (2.55)$$

for every $S \leq s \leq u \leq t \leq T$, where ρ_1, ρ_2 are some non-negative functions with certain properties. As in Lyons' rough path theory ([32, 16]), we expect that such an extension is necessary to enlarge the scope of applications of the stochastic sewing lemma and possibly improve some of the results stated in the current paper. However, we will not consider such general situation on this occasion and defer it for later publications.

Proof of Theorem 2.1. We divide the proof into three steps. In the first step, a process \mathcal{A} is constructed as limit of some Riemann sums using Lemma 2.15. The second step verifies the properties (2.1a) and (2.1b) using Lemma 2.13. In the last step, uniqueness of \mathcal{A} is shown using (2.1a) and (2.1b).

Step 1. Let t be fixed but arbitrary in $[S, T]$. We show that the Riemann sums

$$A_t^\pi = \sum_{i=0}^{n-1} A_{t_i, t_{i+1}}$$

over partitions $\pi = \{t_i\}_{i=0}^n$ of $[0, t]$ has a limit in L_m , denoted by \mathcal{A}_t , as the mesh size $|\pi| := \max_i |t_{i+1} - t_i|$ shrinks to 0. Let $\pi' = \{s_i\}_{i=0}^{n'}$ be another partition of $[0, t]$ and define $\pi'' = \pi \cup \pi'$. We denote the points in π'' by $\{u_i\}_{i=0}^{n''}$ with $u_0 \leq u_1 \leq \dots \leq u_{n''}$ and some integer $n'' \leq n + n'$. Then we have

$$A_t^{\pi''} - A_t^\pi = \sum_{i=0}^{n-1} Z_i \quad \text{where} \quad Z_i = \sum_{j:t_i \leq u_j < t_{i+1}} A_{u_j, u_{j+1}} - A_{t_i, t_{i+1}}.$$

Applying inequality (2.4), we have

$$\|A_t^{\pi''} - A_t^\pi\|_{L_m} \lesssim \sum_{i=0}^{n-1} \|\mathbb{E}^{\mathcal{F}_{t_i}} Z_i\|_{L_m} + \left(\sum_{i=0}^{n-1} \|Z_i\|_{L_m}^2 \right)^{\frac{1}{2}}.$$

Lemma 2.15 is applied to obtain that

$$\|\mathbb{E}^{\mathcal{F}_{t_i}} Z_i\|_{L_m} \lesssim |t_{i+1} - t_i|^{1+\varepsilon_1} \quad \text{and} \quad \|Z_i\|_{L_m} \lesssim |t_{i+1} - t_i|^{\frac{1}{2}+\varepsilon_2}.$$

This implies that

$$\sum_{i=0}^{n-1} \|\mathbb{E}^{\mathcal{F}_{t_i}} Z_i\|_{L_m} \lesssim \sum_{i=0}^{n-1} |t_{i+1} - t_i|^{1+\varepsilon_1} \lesssim |\pi|^{\varepsilon_1}$$

and similarly $\sum_{i=0}^{n-1} \|Z_i\|_{L_m}^2 \lesssim |\pi|^{\varepsilon_2}$. Hence, we have

$$\|A_t^\pi - A_t^{\pi''}\|_{L_m} \lesssim |\pi|^{\varepsilon_1} + |\pi|^{\varepsilon_2}.$$

The same argument is applied to $A_t^{\pi'} - A_t^{\pi''}$ which gives $\|A_t^{\pi'} - A_t^{\pi''}\|_{L_m} \lesssim |\pi'|^{\varepsilon_1} + |\pi'|^{\varepsilon_2}$. Hence, by triangle inequality,

$$\|A_t^\pi - A_t^{\pi'}\|_{L_m} \leq \|A_t^\pi - A_t^{\pi''}\|_{L_m} + \|A_t^{\pi'} - A_t^{\pi''}\|_{L_m} \lesssim |\pi|^{\varepsilon_1} + |\pi|^{\varepsilon_2} + |\pi'|^{\varepsilon_1} + |\pi'|^{\varepsilon_2}.$$

This implies that $\{A_t^\pi\}_\pi$ is Cauchy in L_m and hence $\mathcal{A}_t := \lim_{|\pi| \rightarrow 0} A_t^\pi$ is well-defined in L_m .

Step 2. We show that the process $(\mathcal{A}_t)_{S \leq t \leq T}$ defined in the previous step satisfies (2.1a) and (2.1b).

The condition (2.7) with $s = u = t$ implies that $A_{s,s} = 0$ for every $s \in [S, T]$. Hence, it is evident that $\mathcal{A}_S = 0$. The fact that A is $\{\mathcal{F}_t\}$ -adapted implies that \mathcal{A} is $\{\mathcal{F}_t\}$ -adapted. Obviously, \mathcal{A}_t , being a limit in L_m , belongs to L_m for each t . This shows (2.1a).

Let (s, t) be fixed but arbitrary in $[S, T]_{\leq}^2$. For each integer $n \geq 1$, define $A_{s,t}^n$ as in (2.38). Applying Lemma 2.13, we obtain that

$$\|A_{s,t}^n - A_{s,t}\|_{L_m} \leq \frac{\Gamma_1}{1 - 2^{-\varepsilon_1}} |t - s|^{1+\varepsilon_1} + \frac{\kappa_{m,d}\Gamma_2}{1 - 2^{-\varepsilon_2}} |t - s|^{\frac{1}{2}+\varepsilon_2}, \tag{2.56}$$

and

$$\|\mathbb{E}^{\mathcal{F}_s}(A_{s,t}^n - A_{s,t})\|_{L_m} \leq \frac{\Gamma_1}{1 - 2^{-\varepsilon_1}} |t - s|^{1+\varepsilon_1}. \tag{2.57}$$

From construction of \mathcal{A} in the previous step, we see that

$$\mathcal{A}_t - \mathcal{A}_s = \lim_{n \rightarrow \infty} A_{s,t}^n \text{ in } L_m. \tag{2.58}$$

Hence, passing through the limit $n \rightarrow \infty$ in (2.56) and (2.57), we obtain (2.8) and (2.9) respectively with constants $C_1 = \Gamma_1(1 - 2^{-\varepsilon_1})^{-1}$ and $C_2 = \kappa_{m,d}\Gamma_2(1 - 2^{-\varepsilon_2})^{-1}$. This shows (2.1b).

Step 3. Uniqueness: We show that properties (2.1a) and (2.1b) uniquely determine the process \mathcal{A} . Suppose $\bar{\mathcal{A}}$ is (another) $\{\mathcal{F}_t\}$ -adapted, L_m -integrable process on $[S, T]$ which satisfies $\bar{\mathcal{A}}_S = 0$ and for every $S \leq s \leq t \leq T$

$$\begin{aligned} \|\bar{\mathcal{A}}_t - \bar{\mathcal{A}}_s - A_{s,t}\|_{L_m} &\leq \bar{C}|t - s|^{\frac{1}{2}+\bar{\varepsilon}}, \\ \|\mathbb{E}^{\mathcal{F}_s}(\bar{\mathcal{A}}_t - \bar{\mathcal{A}}_s - A_{s,t})\|_{L_m} &\leq \bar{C}|t - s|^{1+\bar{\varepsilon}} \end{aligned}$$

for some positive constants $\bar{C}, \bar{\varepsilon}$. Then the difference $\tilde{\mathcal{A}} = \mathcal{A} - \bar{\mathcal{A}}$ satisfies

$$\|\tilde{\mathcal{A}}_t - \tilde{\mathcal{A}}_s\|_{L_m} \leq C|t - s|^{\frac{1}{2}+\varepsilon} \text{ and } \|\mathbb{E}^{\mathcal{F}_s}(\tilde{\mathcal{A}}_t - \tilde{\mathcal{A}}_s)\|_{L_m} \leq C|t - s|^{1+\varepsilon} \tag{2.59}$$

for every $S \leq s \leq t \leq T$, for some positive constants C, ε . Let us now fix $t \in [S, T]$, an integer $n \geq 1$ and put $t_i = S + i2^{-n}(t - S)$ for each $i = 0, \dots, 2^n$. We write

$$\tilde{\mathcal{A}}_t = \sum_{i=0}^{2^n-1} (\tilde{\mathcal{A}}_{t_{i+1}} - \tilde{\mathcal{A}}_{t_i})$$

and apply inequality (2.5) (with $\mathcal{G}_i = \mathcal{F}_{t_i}$) to obtain that

$$\|\tilde{\mathcal{A}}_t\|_{L_m} \leq \sum_{i=0}^{2^n-1} \|\mathbb{E}^{\mathcal{F}_{t_i}}(\tilde{\mathcal{A}}_{t_{i+1}} - \tilde{\mathcal{A}}_{t_i})\|_{L_m} + 2\kappa_{m,d} \left(\sum_{i=0}^{2^n-1} \|\tilde{\mathcal{A}}_{t_{i+1}} - \tilde{\mathcal{A}}_{t_i}\|_{L_m}^2 \right)^{\frac{1}{2}}.$$

In view of (2.59), the above inequality yields $\|\tilde{\mathcal{A}}_t\|_{L_m} \lesssim 2^{-n\varepsilon}$ for all integer $n \geq 1$. This means $\mathcal{A}_t = \bar{\mathcal{A}}_t$ a.s. for all $t \in [S, T]$. \square

Remark 2.17. (i) In the second step of the above proof, we can use Lemma 2.15 in place of Lemma 2.13. However, this will produce estimates (2.8) and (2.9) with constants $C_1 = 2\Gamma_1(1 - 2^{-\varepsilon_1})^{-1}$ and $C_2 = 2\kappa_{m,d}\Gamma_2(1 - 2^{-\varepsilon_2})^{-1}$.

(ii) It is evident from the above proof that we can replace the condition that A is L_m -integrable by the condition that δA is L_m -integrable, meaning that $\delta A_{s,u,t}$ is L_m -integrable for every $s \leq u \leq t$. In this case, it is necessary to replace (2.1a) in Theorem 2.1 by

$$(2.1a') \quad \mathcal{A}_S = 0, \mathcal{A} \text{ is } \{\mathcal{F}_t\}\text{-adapted and } \mathcal{A}_t - A_{S,t} \text{ is } L_m\text{-integrable for every } t \in [S, T],$$

and we have $\{A_t^\pi - A_{S,t}\}_\pi$ converges to $\mathcal{A}_t - A_{S,t}$ in L_m as $|\pi| \downarrow 0$. Similarly, the condition that A is $\{\mathcal{F}_t\}$ -adapted can be replaced by the condition that δA is $\{\mathcal{F}_t\}$ -adapted, meaning that $\delta A_{s,u,t}$ is \mathcal{F}_t -measurable for every $s \leq u \leq t$. In Theorem 2.1, if one only assumes that δA is $\{\mathcal{F}_t\}$ -adapted and L_m -integrable, it is necessary to replace (2.1a) by

(2.1a'') $\mathcal{A}_S = 0$, $\mathcal{A}_t - A_{S,t}$ is \mathcal{F}_t -measurable and L_m -integrable for every $t \in [S, T]$.

Proof of Theorem 2.2. We recall the definitions of J, M in (2.15). We have seen from the discussion succeeding (2.15) the existence of the processes \mathcal{J}, \mathcal{M} which are $\{\mathcal{F}_t\}$ -adapted, L_m -integrable and satisfy (2.2b), (2.2d), (2.2e). In addition, the same arguments as Step 1 in the proof of Theorem 2.1 imply that for every $t \in [S, T]$

$$\mathcal{J}_t = \lim_{|\pi| \rightarrow 0} \sum_{i=0}^{N-1} J_{t_i, t_{i+1}} \quad \text{and} \quad \mathcal{M}_t = \lim_{|\pi| \rightarrow 0} \sum_{i=0}^{N-1} M_{t_i, t_{i+1}} \tag{2.60}$$

where the limits are taken in L_m and $\pi = \{0 = t_0 < \dots < t_N = t\}$ denotes a partition of $[0, t]$.

We now show the remaining properties stated in Theorem 2.2. Let (s, t) be fixed but arbitrary in $[S, T]_{\leq}^2$. For each integer $n \geq 1$, let $\pi_{s,t}^n = \{s = t_0^n < t_1^n < \dots < t_{2^n}^n = t\}$ be the dyadic partition of $[s, t]$ as defined in Lemma 2.13. Define

$$M_{s,t}^n = \sum_{i=0}^{2^n-1} M_{t_i^n, t_{i+1}^n}.$$

From (2.16), (2.22) and Lemma 2.13, we obtain that

$$\|M_{s,t}^n - M_{s,t}\|_{L_m} \leq \frac{\kappa_{m,d}\Gamma_2}{1-2^{-\varepsilon_2}} |t-s|^{\frac{1}{2}+\varepsilon_2} + \frac{\kappa_{m,d}\Gamma_3}{1-2^{-\varepsilon_3}} |t-s|^{\frac{1}{2}+\varepsilon_3}.$$

Using the property (2.60), the above inequality implies (2.12) with constants $C_2 = \kappa_{m,d}\Gamma_2(1-2^{-\varepsilon_2})^{-1}$ and $C_3 = \kappa_{m,d}\Gamma_3(1-2^{-\varepsilon_3})^{-1}$. Hence (2.2c) is verified.

To see that $\mathcal{A} = \mathcal{M} + \mathcal{J}$, we note that for any partition $\pi = \{0 = t_0 < \dots < t_N = t\}$ of $[0, t]$, we have

$$\sum_{i=0}^{N-1} A_{t_i, t_{i+1}} = \sum_{i=0}^{N-1} M_{t_i, t_{i+1}} + \sum_{i=0}^{N-1} J_{t_i, t_{i+1}}.$$

In view of (2.58), (2.60), the above identity implies that $\mathcal{A}_t = \mathcal{M}_t + \mathcal{J}_t$ a.s. This shows (2.2a).

Uniqueness: Suppose $(\mathcal{M}, \mathcal{J})$ and $(\bar{\mathcal{M}}, \bar{\mathcal{J}})$ are two pairs of processes which satisfy (2.2a), (2.2b) and (2.2c) with the same processes A and \mathcal{A} . Then the difference processes $\tilde{\mathcal{M}} = \mathcal{M} - \bar{\mathcal{M}}$ and $\tilde{\mathcal{J}} = \mathcal{J} - \bar{\mathcal{J}}$ satisfy

$$\tilde{\mathcal{M}}_t = -\tilde{\mathcal{J}}_t \quad \text{a.s.} \quad \text{and} \quad \|\tilde{\mathcal{J}}_t - \tilde{\mathcal{J}}_s\|_{L_m} \leq C|t-s|^{\frac{1}{2}+\varepsilon} \tag{2.61}$$

for every $S \leq s \leq t \leq T$, for some constants $C \geq 0$ and $\varepsilon > 0$. The process $\tilde{\mathcal{M}}$, being a square integrable martingale with super-diffusive increments, ought to be a constant process. Alternatively, one could use the same arguments used in Step 3 of the proof of Theorem 2.1 to conclude that $\tilde{\mathcal{M}}$ is a constant process. Since $\mathcal{M}_S = \bar{\mathcal{M}}_S = 0$, this implies that $\mathcal{M} = \bar{\mathcal{M}}$ and $\mathcal{J} = \bar{\mathcal{J}}$.

If instead, $(\mathcal{M}, \mathcal{J})$ and $(\bar{\mathcal{M}}, \bar{\mathcal{J}})$ satisfy (2.2a), (2.2b) and (2.2d) with the same processes A and \mathcal{A} , the difference processes $\tilde{\mathcal{M}}$ and $\tilde{\mathcal{J}}$ also satisfy (2.61). Hence, we also have $\mathcal{M} = \bar{\mathcal{M}}$ and $\mathcal{J} = \bar{\mathcal{J}}$. This shows the claim (2.2f).

Since \mathcal{M} is obtained by applying Theorem 2.1 to $(s, t) \mapsto M_{s,t} = A_{s,t} - \mathbb{E}^{\mathcal{F}_s} A_{s,t}$, the characterization of \mathcal{M} in (2.2g) follows from the uniqueness part of Theorem 2.1. One can show the characterization of \mathcal{J} in (2.2h) analogously. \square

3 Distributive functionals of Markov processes

In probability theory and in the study of stochastic differential equations with distributional drifts ([2, 8, 25, 5]), one encounters the integral of the form $\int_0^\cdot f(X_r)dr$ where X is a solution to a stochastic differential equation (for instance a Brownian motion) and f is a distribution. To define and study such integrals, previous methods involve explicit moment computations (as in the study of local times, [34]) or the so-call Zvonkin transformation, also known as the Itô-Tanaka trick (commonly used in [2, 13, 8]). In the present section, a new method is proposed which is based on Theorem 2.3. This method relies on a certain smoothing effect of the transition probability kernel associated to X and does not require explicit moment computations nor Itô calculus to work. The main results are described below in Propositions 3.5 and 3.10.

We proceed with the detail. Let us fix a time horizon $T > 0$ and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space equipped with a filtration $\{\mathcal{F}_t\}$ such that \mathcal{F}_0 contains \mathbb{P} -null sets. For each $m \geq 1$ and $\tau \in (0, 1]$, the space $C_T^\tau L_m = C^\tau([0, T]; [L^m(\Omega, \mathcal{F}, \mathbb{P})]^d)$ contains all stochastic processes ψ with values in \mathbb{R}^d such that

$$\|\psi\|_{C_T^\tau L_m} := \sup_{t \in [0, T]} \|\psi_t\|_{L_m} + \sup_{s, t \in [0, T]; s \neq t} \frac{\|\psi_t - \psi_s\|_{L_m}}{|t - s|^\tau} < \infty.$$

Similarly, the space $C_T L_m = C([0, T]; [L^m(\Omega, \mathcal{F}, \mathbb{P})]^d)$ contains all stochastic processes ψ with values in \mathbb{R}^d such that $\psi : [0, T] \rightarrow L_m$ is continuous. The norm of ψ in $C_T L_m$ is defined by $\|\psi\|_{C_T L_m} := \sup_{t \in [0, T]} \|\psi_t\|_{L_m}$.

Let $X = \{X_t\}_{t \geq 0}$ be a Markov process in \mathbb{R}^d with respect to $\{\mathcal{F}_t\}$ with Lebesgue-measurable sample paths. Let $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ be a normed vector space, which contains $C_b(\mathbb{R}^d)$ and is a subset of $\mathcal{S}'(\mathbb{R}^d)$. Let $U = \{U_{s,t}\}_{0 \leq s \leq t}$ be a two-parameter family of bounded operators on \mathcal{H} . The following condition relates X, U and \mathcal{H} .

Condition 3.1. (X, U, \mathcal{H}) satisfies:

- (i) there are finite numbers $\|U\|_{\mathcal{H}, \nu} \geq 0$ and $\nu \leq 0$ such that for every $h \in \mathcal{H}$ and every $0 \leq s < t \leq T$, $U_{s,t}h$ is a bounded continuous function on \mathbb{R}^d and satisfies

$$\|U_{s,t}h\|_\infty \leq \|U\|_{\mathcal{H}, \nu} \|h\|_{\mathcal{H}} |t - s|^{\frac{\nu}{2}}, \tag{3.1}$$

where $\|\cdot\|_\infty$ is the supremum norm,

- (ii) for every $s \geq 0$, $x \in \mathbb{R}^d$ and $h \in \mathcal{H}$, the map $[s, T] \ni r \rightarrow U_{s,r}h(x) \in \mathbb{R}$ is Lebesgue measurable,
- (iii) for every $s \leq r < t$, $x \in \mathbb{R}^d$ and $h \in \mathcal{H}$, $\mathbb{E}^{\mathcal{F}_s}[U_{r,t}h](X_r) = U_{s,t}h(X_s)$.

The first two conditions (i),(ii) are purely analytic which are necessary for integrations. The last condition (iii) enforces that U is the propagator which corresponds to the transition function of X (see [10, pg. 156]).

We think of \mathcal{H} as a space containing irregular functions, measures or distributions with a certain negative regularity index. To be more specific, we briefly recall the definition of Hölder-Besov spaces. For each $\phi \in \mathcal{S}(\mathbb{R}^d)$, $x \in \mathbb{R}^d$ and $\lambda > 0$, ϕ_x^λ is defined by $\phi_x^\lambda(y) = \lambda^{-d} \phi(\lambda^{-1}(y - x))$. For each integer $r \geq 0$, let \mathfrak{B}^r be the set of all smooth functions ϕ on \mathbb{R}^d which are supported on the unit ball $\{x \in \mathbb{R}^d : |x| \leq 1\}$ and satisfy $\|\phi\|_{C_b^r(\mathbb{R}^d)} \leq 1$.

Definition 3.2. Let $\gamma \leq 0$ be a real number and $r > |\gamma|$ be an integer. The homogeneous Hölder-Besov space $C^\gamma(\mathbb{R}^d)$ consists of all tempered distributions g such that

$$\|g\|_{C^\gamma} := \sup_{\lambda \in (0, 1]} \sup_{x \in \mathbb{R}^d} \sup_{\phi \in \mathfrak{B}^r} \lambda^{-\gamma} |g(\phi_x^\lambda)| < \infty.$$

We refer to [35] for more detail on the Besov-Hölder spaces. As a toy example of a tuple (X, U, \mathcal{H}) which satisfies Condition 3.1, one may take $(X, \{\mathcal{F}_t\})$ as a standard Brownian motion in \mathbb{R}^d , $U_{s,t} = P_{s,t}$ which is the operator associated to the Gaussian kernel $(2\pi(t-s))^{-\frac{d}{2}} e^{-|x-y|^2/(2(t-s))}$, and \mathcal{H} as the Besov-Hölder space $\mathcal{C}^\nu(\mathbb{R}^d)$. Since $P_{s,t}$ has a smooth kernel whenever $s < t$, the action of $P_{s,t}$ on any distribution is a smooth function. Combining with known Schauder estimates of heat kernels (see e.g. [20, Lemma 2.8] or [5, Prop. 2.4]), we see that for every $g \in \mathcal{C}^\nu$ and $s < t$, $P_{s,t}g$ is continuous on \mathbb{R}^d and satisfies

$$\|P_{s,t}g\|_\infty \leq C|t-s|^{\frac{\nu}{2}} \|g\|_{\mathcal{C}^\nu} \tag{3.2}$$

for some positive constant C independent from s, t and g . It follows that Condition 3.1 is satisfied for every $\nu \leq 0$. Other examples will appear in subsequent sections.

For each integer $k \geq 0$, let \mathcal{E}_T and \mathcal{E}_T^k be respectively the sets of measurable finite-valued functions from $[0, T]$ to \mathcal{H} and $C_b^k(\mathbb{R}^d)$. A function f in \mathcal{E}_T has the following form

$$f_r = \sum_{i \in F} 1_{I_i}(r) h_i \quad \forall r \in [0, T], \tag{3.3}$$

where F is a finite set, I_i 's are Lebesgue-measurable subsets of $[0, T]$ and h_i 's are elements in \mathcal{H} . Similarly, a function f in \mathcal{E}_T^k has the form (3.3) with $h_i \in C_b^k(\mathbb{R}^d)$ for every i . Functions in \mathcal{E}_T and \mathcal{E}_T^k are also called simple functions.

A function $f : [0, T] \rightarrow \mathcal{H}$ is strongly Lebesgue-measurable if there exists a sequence of simple functions $\{f^n\}$ in \mathcal{E}_T such that $\lim_n f_r^n = f_r$ in \mathcal{H} for Lebesgue-a.e. $r \in [0, T]$. For each $q \in [1, \infty)$, the Bochner space $L^q([0, T]; \mathcal{H})$ is the linear space of all (equivalence classes of) strongly Lebesgue-measurable functions $f : [0, T] \rightarrow \mathcal{H}$ for which the Lebesgue integral $\int_{[0, T]} \|f_r\|_{\mathcal{H}}^q dr$ is finite. We define $L^\infty([0, T]; \mathcal{H})$ as the linear space of all (equivalence classes of) strongly Lebesgue-measurable functions $f : [0, T] \rightarrow \mathcal{H}$ for which there exists a real number $N \geq 0$ such that the set $\{r \in [0, T] : \|f_r\|_{\mathcal{H}} > N\}$ has Lebesgue measure 0. For brevity, we abbreviate $L_T^q \mathcal{H}$ for $L^q([0, T]; \mathcal{H})$ for every $q \in [1, \infty]$. The norm on $L_T^q \mathcal{H}$ is defined by

$$\|f\|_{L_T^q \mathcal{H}} := \left(\int_{[0, T]} \|f_r\|_{\mathcal{H}}^q dr \right)^{\frac{1}{q}} \quad \text{for finite } q$$

and when $q = \infty$ by

$$\|f\|_{L_T^\infty \mathcal{H}} := \operatorname{ess\,sup}_{r \in [0, T]} \|f_r\|_{\mathcal{H}}.$$

For further properties on Bochner spaces and measurability of Banach-valued functions, we refer the reader to [27, Chapter 1].

For a given $f \in L_T^q \mathcal{H}$, measurability and integrability of the map $r \mapsto U_{s,r} f_r$ are discussed in the following two lemmas.

Lemma 3.3. *Let (X, U, \mathcal{H}) satisfy Condition 3.1. For every $f \in L_T^q \mathcal{H}$ and $x \in \mathbb{R}^d$, the map $r \mapsto U_{s,r} f_r(x)$ is Lebesgue-measurable from $[s, T]$ to \mathbb{R} .*

Proof. Since f is strongly Lebesgue-measurable, we can find a sequence of simple functions $\{f^n\}$ in \mathcal{E}_T such that $\lim_n f_r^n = f_r$ in \mathcal{H} for a.e. $r \in [s, T]$. For each n , if f_r^n has the form (3.3), then for every $r \in (s, T]$,

$$U_{s,r} f_r^n = \sum_{i \in F} 1_{I_i}(r) U_{s,r} h_i.$$

In view of Condition 3.1(ii), the above identity implies that $U_{s,\cdot} f^n(x) : [s, T] \rightarrow \mathbb{R}$ is measurable for every $x \in \mathbb{R}^d$. For each $r \in (s, T]$, using the bound (3.1), we see that $\lim_n U_{s,r} f_r^n(x) = U_{s,r} f_r(x)$ for every $x \in \mathbb{R}^d$. This implies, in particular, that $U_{s,\cdot} f(x) : [s, T] \rightarrow \mathbb{R}$ is Lebesgue-measurable. \square

Lemma 3.4. *Suppose that Condition 3.1 holds. Let f be a function in $L_T^q \mathcal{H}$, with $q \in [1, \infty]$ is such that*

$$(\nu, q) = (0, 1) \quad \text{or} \quad \frac{\nu}{2} - \frac{1}{q} > -1. \tag{3.4}$$

Then, for every $s \in [0, T]$ and $x \in \mathbb{R}^d$, the maps $[s, T] \ni r \mapsto U_{s,r} f_r(x) \in \mathbb{R}^d$ and $[s, T] \ni r \mapsto \|U_{s,r} f_r\|_\infty \in \mathbb{R}$ are Lebesgue-measurable and Lebesgue-integrable. In addition, for every $(s, t) \in [0, T]_{\leq}^2$, we have

$$\int_{[s,t]} \|U_{s,r} f_r\|_\infty dr \leq c(\nu, q) \|U\|_{\mathcal{H}, \nu} \|f\|_{L_T^q \mathcal{H}} (t-s)^{1+\frac{\nu}{2}-\frac{1}{q}}, \tag{3.5}$$

where $c(\nu, q) = 1$ if $q = 1$ and $c(\nu, q) = (1 + \frac{\nu}{2} \frac{q}{q-1})^{\frac{1}{q}-1}$ otherwise.

Proof. From Lemma 3.3, we see that $[s, T] \ni r \mapsto U_{s,r} f_r(x) \in \mathbb{R}^d$ is Lebesgue-measurable for every $x \in \mathbb{R}^d$. Since for each $r > s$, $U_{s,r} f_r : \mathbb{R}^d \rightarrow \mathbb{R}$ is continuous, we have

$$\|U_{s,r} f_r\|_\infty = \sup_{x \in \mathbb{Q}^d} |U_{s,r} f_r(x)|$$

where \mathbb{Q} is the set of rationals on \mathbb{R} . This implies that $\|U_{s,\cdot} f\|_\infty : [s, T] \rightarrow \mathbb{R}$ is Lebesgue-measurable.

From Condition 3.1(i), we have

$$\|U_{s,r} f_r\|_\infty \leq \|U\|_{\mathcal{H}, \nu} |r-s|^{\frac{\nu}{2}} \|f_r\|_{\mathcal{H}}.$$

When $q = 1$, (3.4) implies that $\nu = 0$, so we have

$$\int_{[s,t]} \|U_{s,r} f_r\|_\infty \leq \|U\|_{\mathcal{H}, \nu} \|f\|_{L_T^1 \mathcal{H}}.$$

When $q > 1$, we apply Hölder inequality to see that

$$\int_{[s,t]} |r-s|^{\frac{\nu}{2}} \|f_r\|_{\mathcal{H}} dr \leq c(\nu, q) \|f\|_{L_T^q \mathcal{H}} (t-s)^{1+\frac{\nu}{2}-\frac{1}{q}}.$$

The above argument shows that $\|U_{s,\cdot} f\|_\infty$ is integrable and the estimate (3.5) holds. Obviously, for every $x \in \mathbb{R}^d$ and $r > s$, $|U_{s,r} f_r(x)| \leq \|U_{s,r} f_r\|_\infty$, so that $U_{s,\cdot} f(x)$ is also Lebesgue integrable. This completes the proof. \square

For each $t > 0$ and each simple function f in \mathcal{E}_T^0 having the form (3.3), the integration $\int_{[0,t]} f_r(X_r) dr$ is defined by

$$\int_{[0,t]} f_r(X_r) dr := \sum_{i \in F} \int_{I_i \cap [0,t]} h_i(X_r) dr.$$

Since X has measurable sample paths, each $h_i(X(\omega)\cdot) : [0, T] \rightarrow \mathbb{R}$ is a bounded measurable function and each summand in the right-hand side of the above identity is a Lebesgue integral for a.s. $\omega \in \Omega$. For a function f in $L_T^\infty C_b(\mathbb{R}^d)$, one can find (see [27, Corollary 1.1.21]) a sequence of simple functions $\{f_n\}$ in \mathcal{E}_T^0 such that for a.e. $r \in [0, T]$

$$\|f_n(r, \cdot)\|_\infty \leq \|f(r, \cdot)\|_\infty \quad \text{and} \quad \lim_n \|f_n(r, \cdot) - f(r, \cdot)\|_\infty = 0.$$

It follows that $f : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ and $f(\cdot, X_\cdot) : [0, T] \rightarrow \mathbb{R}$ are Lebesgue measurable. Consequently, the integral $\int_{[0,t]} f(r, X_r) dr$ is well-defined as a Lebesgue integral and

$$\int_{[0,t]} f(r, X_r) dr = \lim_n \int_{[0,t]} f_n(r, X_r) dr$$

by dominated convergence theorem. We are going to define a linear map \mathcal{A}^X on $L_T^q \mathcal{H}$ such that \mathcal{A}^X is continuous on $L_T^q \mathcal{H}$ and for every $t \in [0, T]$, $\mathcal{A}_t^X[f] = \int_{[0,t]} f_r(X_r) dr$ whenever $f \in L_T^\infty C_b(\mathbb{R}^d)$. The process $t \mapsto \mathcal{A}_t^X[f]$ is a natural extension to the integral $\int_0^t f_r(X_r) dr$ in situations when f_r is not a function or when the composition $f_r(X_r)$ is a priori ill-posed. The main toolbox we employ here is the stochastic sewing lemma described in Section 2, more specifically Theorem 2.3.

Let f be an element in $L_T^q \mathcal{H}$. For every $s \leq t$, we put

$$A_{s,t}^X[f] = \int_{[s,t]} U_{s,r} f_r(X_s) dr \tag{3.6}$$

Condition 3.1(i) ensures that $U_{s,r} f_r$ is a continuous function for each $s < r$, hence evaluating $U_{s,r} f_r$ at any point makes perfect sense and there is no ambiguity in (3.6). Furthermore, the estimate (3.5) implies that

$$\|A_{s,t}^X[f]\|_{L_m} \leq c(\nu, q) \|U\|_{\mathcal{H}, \nu} \|f 1_{[s,t]}\|_{L_T^q \mathcal{H}} (t - s)^{1 + \frac{\nu}{2} - \frac{1}{q}}. \tag{3.7}$$

If f is a bounded continuous function on $[0, T] \times \mathbb{R}^d$, we can write

$$A_{s,t}^X[f] = \mathbb{E}^{\mathcal{F}_s} \int_s^t f_r(X_r) dr.$$

For every triplet $s \leq u \leq t$, we have

$$\delta A_{s,u,t}^X[f] = \int_{[u,t]} U_{s,r} f_r(X_s) dr - \int_{[u,t]} U_{u,r} f_r(X_u) dr.$$

Applying Fubini's theorem (which is eligible due to (3.5)) and Condition 3.1(iii), we see that

$$\mathbb{E}^{\mathcal{F}_s} \int_{[u,t]} U_{u,r} f_r(X_u) dr = \int_{[u,t]} \mathbb{E}^{\mathcal{F}_s} U_{u,r} f_r(X_u) dr = \int_{[u,t]} U_{s,r} f_r(X_s) dr.$$

It follows that

$$\mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t}^X[f] = 0. \tag{3.8}$$

In view of (3.7) and the above identity, we see that the hypotheses of Theorem 2.3 are satisfied provided that the exponent of $t - s$ in (3.7) is at least a half. This leads to the following result.

Proposition 3.5. *Let $m \geq 2$ be fixed and assume that Condition 3.1 is satisfied. Let q be a number in $(2, \infty]$ such that*

$$\frac{\nu}{2} - \frac{1}{q} > -\frac{1}{2}. \tag{3.9}$$

There exists a linear map \mathcal{A}^X from $L_T^q \mathcal{H}$ to $C_T^{1 + \frac{\nu}{2} - \frac{1}{q}} L_m$ which satisfies the following properties.

- (a) *For every $f \in L_T^q \mathcal{H}$ and every $0 \leq s \leq t \leq T$, $\mathcal{A}_t^X[f]$ is \mathcal{F}_t -measurable,*

$$\mathbb{E}^{\mathcal{F}_s} (\mathcal{A}_t^X[f] - \mathcal{A}_s^X[f]) = \int_{[s,t]} U_{s,r} f_r(X_s) dr \quad \text{a.s.}$$

and

$$\|\mathcal{A}_t^X[f] - \mathcal{A}_s^X[f]\|_{L_m} \lesssim \|U\|_{\mathcal{H}, \nu} \|f 1_{[s,t]}\|_{L_T^q \mathcal{H}} |t - s|^{1 + \frac{\nu}{2} - \frac{1}{q}}. \tag{3.10}$$

(b) For every $f \in L_T^q \mathcal{H}$ and every t in $[0, T]$, $\mathcal{A}_t^X[f]$ is the limit in L_m of the Riemann sum

$$\sum_i \int_{[t_i, t_{i+1}]} U_{t_i, r} f_r(X_{t_i}) dr$$

as $\max_i |t_{i+1} - t_i| \rightarrow 0$. Here $\{t_i\}_i$ is any finite partition of $[0, t]$.

Proof. Let $f \in L_T^q \mathcal{H}$ be fixed and define $\mathcal{A}_{s,t}^X[f]$ as in (3.6). From (3.9) and the estimate (3.7), we infer that $\mathcal{A}^X[f]$ satisfies (2.23) with $\varepsilon_4 = \frac{1}{2} + \frac{\nu}{2} - \frac{1}{q}$. We apply Theorem 2.3 to $\mathcal{A}_{s,t}[f]$ to obtain an $\{\mathcal{F}_t\}$ -adapted and L_m -integrable process $\mathcal{A}^X[f]$ which satisfies properties (a) and (b). Inequality (3.10) also implies that $\mathcal{A}^X[f]$ belongs to $C_T^{1+\frac{\nu}{2}-\frac{1}{q}} L_m$. \square

The next result connects $\mathcal{A}^X[f]$ with the usual integration $\int_0^t f(r, X_r) dr$ for a class of test functions.

Definition 3.6. $\mathfrak{G}_T(X)$ is the set of all functions f in $L_T^\infty C_b(\mathbb{R}^d)$ such that for a.e. $r \in [0, T]$ and a.s. ω , $\mathbb{E}^{\mathcal{F}_s} f_r(X_r(\omega)) = U_{s,r} f_r(X_s(\omega))$ for every $s \in [0, r]$.

Proposition 3.7. Suppose that the hypotheses of Proposition 3.5 are satisfied. Then, $\mathcal{A}_t^X[f] = \int_{[0,t]} f(r, X_r) dr$ a.s. for every $t \in [0, T]$ and every $f \in \mathfrak{G}_T(X)$.

Proof. By Fubini's theorem and the definition of $\mathfrak{G}_T(X)$, for every $s \leq t$, we have

$$\mathbb{E}^{\mathcal{F}_s} \int_{[s,t]} f_r(X_r) dr = \int_{[s,t]} \mathbb{E}^{\mathcal{F}_s} f_r(X_r) dr = \int_{[s,t]} U_{s,r} f_r(X_s) dr \quad \text{a.s.}$$

It follows that

$$\left\| \int_{[s,t]} f_r(X_r) dr - \int_{[s,t]} U_{s,r} f_r(X_s) dr \right\|_{L_m} \leq 2|t - s| \operatorname{ess\,sup}_{r \in [s,t]} \|f_r\|_\infty.$$

Hence, by Proposition 3.5 and the uniqueness part of Theorem 2.3, $\mathcal{A}_t^X[f] = \int_{[0,t]} f_r(X_r) dr$ a.s. for every $t \geq 0$. \square

Remark 3.8. Let $f \in L_T^q \mathcal{H}$ be fixed.

(i) Since the process $\mathcal{A}^X[f]$ is constructed from Theorem 2.3, it is the unique $\{\mathcal{F}_t\}$ -adapted and L_m -integrable process φ such that $\varphi_0 = 0$,

$$\|\varphi_t - \varphi_s\|_{L_m} \lesssim |t - s|^{\frac{1}{2} + \varepsilon} \quad \text{and} \quad \|\mathbb{E}^{\mathcal{F}_s}(\varphi_t - \varphi_s - \int_{[s,t]} U_{s,r} f_r(X_s) dr)\|_{L_m} \lesssim |t - s|^{1 + \varepsilon}$$

for every (s, t) in $[0, T]_{\geq}^2$, for some $\varepsilon > 0$. This provides a characterization of the process $\mathcal{A}^X[f]$ for a given f .

(ii) By choosing $m \geq 2$ sufficiently large in (3.10) and applying Kolmogorov continuity theorem, we can find a modification of $\mathcal{A}^X[f]$ which is a.s. τ -Hölder continuous on $[0, T]$ for every $\tau \in (0, 1 + \frac{\nu}{2} - \frac{1}{q})$. In the remaining of the section, we will always work with this continuous version of $\mathcal{A}^X[f]$.

Corollary 3.9. Assume that the hypotheses of Proposition 3.5 hold. Let τ be any number in $(\frac{1}{2}, 1 + \frac{\nu}{2} - \frac{1}{q})$. There exists a number $m_\tau \geq 2$ such that for every $f \in L_T^q \mathcal{H}$ and every sequence $\{f_n\}_n \subset \mathfrak{G}_T(X)$ convergent to f in $L_T^q \mathcal{H}$,

$$\lim_n \left\| \left\| \int_{[0,\cdot]} f_n(r, X_r) dr - \mathcal{A}^X[f] \right\|_{C^\tau([0,T])} \right\|_{L_m} = 0 \tag{3.11}$$

for every $m \geq m_\tau$.

Proof. From (3.10), we have

$$\|\mathcal{A}_t^X[f_n - f] - \mathcal{A}_s^X[f_n - f]\|_{L_m} \lesssim \|f_n - f\|_{L_T^q \mathcal{H}} |t - s|^{1 + \frac{\nu}{2} - \frac{1}{q}}.$$

for every $s \leq t$. Applying Garsia-Rodemich-Rumsey inequality, there exists $m_\tau \geq 2$ such that for every $m \geq m_\tau$, we have

$$\|\|\mathcal{A}^X[f_n - f]\|_{C^\tau([0, T])}\|_{L_m} \lesssim \|f_n - f\|_{L_T^q \mathcal{H}}.$$

Observing that $\mathcal{A}_t^X[f_n] = \int_{[0, t]} f_n(r, X_r) dr$ (by Proposition 3.7), the previous estimate implies the result. \square

It worths noting that the process $\mathcal{A}^X[f]$ depends on the trajectory of X , the element f and the transition law of $(X, \{\mathcal{F}_t\})$ (via the transition propagator U). Compare to the situation when f is a continuous bounded function, $\int_0^\cdot f_r(X_r) dr$ only depends on the first two factors. In the last part of the current section, we investigate the continuity dependence of process $\mathcal{A}^X[f]$ with respect to the factors: the transition propagator corresponding to $(X, \{\mathcal{F}_t\})$, the trajectory of X and the element f .

Let X, Y be two Markov processes with respect to $\{\mathcal{F}_t\}$ and let $U^X = \{U_{s,t}^X\}_{s \leq t}$, $U^Y = \{U_{s,t}^Y\}_{s \leq t}$ be two families of bounded operators on \mathcal{H} . Assume that (X, U^X, \mathcal{H}) and (Y, U^Y, \mathcal{H}) satisfy Condition 3.1 with the same $\nu \leq 0$. We denote by $\|U^X - U^Y\|_{\mathcal{H}, \nu}$ the smallest number M such that

$$\|(U_{s,t}^X - U_{s,t}^Y)\phi\|_\infty \leq M \|\phi\|_{\mathcal{H}} |t - s|^{\frac{\nu}{2}}$$

for every $\phi \in \mathcal{H}$ and $0 \leq s < t \leq T$. From the bound (3.1), we see that $\|U^X - U^Y\|_{\mathcal{H}, \nu}$ is necessarily finite. For every $\nu_1 \leq 0$ and every $h > 0$, we define

$$\rho_{\nu_1}^X(h) := \sup_{\phi: \|\phi\|_{\mathcal{H}} \leq 1} \sup_{0 \leq s < t \leq T} \sup_{|x-y| \leq h} \frac{|U_{s,t}^X \phi(x) - U_{s,t}^X \phi(y)|}{|t - s|^{\frac{\nu_1}{2}}}. \tag{3.12}$$

We note that by triangle inequality

$$\rho_{\nu_1}^X(h) \leq 2 \sup_{\phi: \|\phi\|_{\mathcal{H}} \leq 1} \sup_{0 \leq s < t \leq T} \frac{\|U_{s,t}^X \phi\|_\infty}{|t - s|^{\frac{\nu_1}{2}}}$$

so that $\rho_{\nu_1}^X(h)$ is finite whenever $\nu_1 \leq \nu$.

Proposition 3.10. *We assume that (X, U^X, \mathcal{H}) and (Y, U^Y, \mathcal{H}) satisfy the Condition 3.1 with the same $\nu \leq 0$. Let $q \in (2, \infty]$ satisfy inequality (3.9) and let $\nu_1 \leq 0$ satisfy $\frac{\nu_1}{2} - \frac{1}{q} > -\frac{1}{2}$. We assume that $\rho_{\nu_1}^X(h)$ is finite for every $h > 0$. Then, for every $f, g \in L_T^q \mathcal{H}$ and for every $(s, t) \in [0, T]_{\leq}^2$,*

$$\begin{aligned} & \|\mathcal{A}_t^X[g] - \mathcal{A}_s^X[g] - \mathcal{A}_t^Y[f] + \mathcal{A}_s^Y[f]\|_{L_m} \\ & \lesssim \left(\|f - g\|_{L_T^q \mathcal{H}} \|U^X\|_{\mathcal{H}, \nu} + \|f\|_{L_T^q \mathcal{H}} D_{\nu, \nu_1}(X, Y) \right) |t - s|^{1 + \frac{\nu \wedge \nu_1}{2} - \frac{1}{q}}, \end{aligned}$$

where

$$D_{\nu, \nu_1}(X, Y) = \|U^X - U^Y\|_{\mathcal{H}, \nu} + \sup_{r \in [0, T]} \|\rho_{\nu_1}^X(|X_r - Y_r|)\|_{L_m}.$$

Proof. We apply Theorem 2.3 to

$$A_{s,t} := \mathcal{A}_{s,t}^X[g] - \mathcal{A}_{s,t}^Y[f] = \int_{[s,t]} U_{s,r}^X g_r(X_s) dr - \int_{[s,t]} U_{s,r}^Y f_r(Y_s) dr.$$

Note that

$$A_{s,t} = \int_{[s,t]} U_{s,r}^X (g_r - f_r)(X_s) dr + \int_{[s,t]} [U_{s,r}^X f_r(X_s) - U_{s,r}^X f_r(Y_s)] dr + \int_{[s,t]} [U_{s,r}^X f_r(Y_s) - U_{s,r}^Y f_r(Y_s)] dr.$$

We use (3.1), (3.7) and (3.12) to estimate these terms, which yields

$$|A_{s,t}| \lesssim \|U^X\|_{\mathcal{H},\nu} \|g - f\|_{L^q_T \mathcal{H}} |t - s|^{1+\frac{\nu}{2}-\frac{1}{q}} + \rho_{\nu_1}^X(|X_s - Y_s|) \|f\|_{L^q_T \mathcal{H}} |t - s|^{1+\frac{\nu_1}{2}-\frac{1}{q}} + \|U^X - U^Y\|_{\mathcal{H},\nu} \|f\|_{L^q_T \mathcal{H}} |t - s|^{1+\frac{\nu}{2}-\frac{1}{q}}.$$

This implies that

$$\|A_{s,t}\|_{L_m} \lesssim \left(\|f - g\|_{L^q_T \mathcal{H}} \|U^X\|_{\mathcal{H},\nu} + \|f\|_{L^q_T \mathcal{H}} D_{\nu,\nu_1}(X, Y) \right) |t - s|^{1+\frac{\nu \wedge \nu_1}{2}-\frac{1}{q}}.$$

Similar to (3.8), we also have $\mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t} = 0$ for every triplet $s \leq u \leq t$. Since $1 + \frac{\nu \wedge \nu_1}{2} - \frac{1}{q} > \frac{1}{2}$ by our assumptions, Theorem 2.3 is applicable and (2.24) yields the stated inequality. \square

4 Stochastic flows

Let α be a fixed number in $(0, 1)$ and b be a function in $[L^\infty([0, T]; C_b^\alpha(\mathbb{R}^d))]^d$. Given $x \in \mathbb{R}^d$, consider the stochastic differential equation (SDE)

$$dX_t = b(t, X_t) dt + dW_t, \quad t \in [0, T], \quad X_0 = x, \tag{4.1}$$

where $(W, \{\mathcal{F}_t\})$ is a standard Brownian motion in \mathbb{R}^d . Let X_t^x denote the solution to (4.1) starting from $X_0 = x$ at time $t = 0$ which is adapted to $\{\mathcal{F}_t\}$. It is shown in [12, Theorem 5(ii)] that equation (4.1) has a $C^{1+\alpha'}$ -stochastic flow for any $\alpha' \in (0, \alpha)$. In particular, the process $Y_t^{ij} = \partial_{x_i} X_t^{j,x}$ is well-defined. (For simplicity, we omit the dependence of x in Y .) Formally, Y satisfies the following equations

$$Y_t^{ij} = \delta_{ij} + \sum_{k=1}^d \int_0^t \partial_k b^j(r, X_r^x) Y_r^{ik} dr, \quad \forall i, j \in \{1, \dots, d\}, \tag{4.2}$$

where δ_{ij} is the Kronecker delta symbol. Since $\nabla b(r, \cdot)$ is only a distribution, the composition $\nabla b(r, X_r^x)$ is a priori ill-posed. Hence, equation (4.2) is not mathematically rigorous and an equation describing the dynamic of Y is missing in the literature. Filling this gap is the main purpose of the current section.

Indeed, we can make sense of the process

$$V_t^{kj} = V_t^{kj}(b, X) = A_t^X[\partial_k b^j] \tag{4.3}$$

for every $t \geq 0$ and $j, k \in \{1, \dots, d\}$ by applying Proposition 3.5. Moreover, it turns out that the process V^{kj} has a modification which belongs to $C^{\frac{1+\alpha'}{2}}([0, T])$ a.s. for every $\alpha' \in (0, \alpha)$. Then, a rigorous interpretation of the system of equations in (4.2) is

$$Y_t^{ij} = \delta_{ij} + \sum_{k=1}^d \int_0^t Y_r^{ki} dV_r^{kj}(b, X), \quad \forall i, j \in \{1, \dots, d\}, \tag{4.4}$$

where the integrals on the right-hand side are Young integrals ([45]). Our main result in the current section can be stated in the following theorem.

Theorem 4.1. *Let $x \in \mathbb{R}^d$ be fixed but arbitrary. With probability one, the process $\partial_{x_i} X^{j,x}$ is the unique solution to the system of Young-type equations (4.4). In addition, the map $t \rightarrow \nabla X_t^x$ is a.s. $\frac{1+\alpha'}{2}$ -Hölder continuous for every $\alpha' \in (0, \alpha)$.*

To show the above stated result, we first show that the process V given by (4.3) is well-defined and has a $\frac{1+\alpha'}{2}$ -Hölder continuous modification. Then we show that $\partial_{x_i} X^{j,x}$ satisfies equation (4.4) using smooth approximations of b . The first step amounts in verifying the hypothesis of Proposition 3.5, Condition 3.1. To pass through the limit in the second step, we rely on stability properties of Young-type differential equations (see e.g. [6, 26]) and of the functional \mathcal{A}^X (Corollary 3.9).

To verify Condition 3.1, we study the transition propagator of X , which is denoted by $U = \{U_{s,t}\}_{s \leq t}$. Each $U_{s,t}$, with $s < t$, takes values in $C_b(\mathbb{R}^d)$ and is defined first on $C_b^{2+\alpha}(\mathbb{R}^d)$, then extended to larger classes of bounded uniformly continuous functions and Hölder-Besov distributions. The definition of $U_{s,t}$ as an operator on $C_b^{2+\alpha}(\mathbb{R}^d)$ is a direct consequence of the regularity theory of Krylov and Priola in [29] for the Kolmogorov backward parabolic differential equations associated to the SDE (4.1). To extend the domain of $U_{s,t}$ to the larger classes of uniformly continuous functions and Hölder-Besov distributions, we rely on a maximum principle shown in Lemma 4.3 and a priori estimates in Lemma 4.8 showing that solutions to Kolmogorov backward equations are more regular away from the terminal time.

Let us proceed with more detail. For each $t \in [0, T]$, $f \in L^\infty([0, T]; C_b^\alpha(\mathbb{R}^d))$ and $g \in C_b^{2+\alpha}(\mathbb{R}^d)$, consider the Kolmogorov backward equation

$$\left(\partial_s + \frac{1}{2} \Delta + b \cdot \nabla \right) u(s, x) = f \quad \forall (s, x) \in [0, t] \times \mathbb{R}^d, \quad u(t, \cdot) = g(\cdot). \quad (4.5)$$

Since the coefficients b, f are only measurable in time, the notion of solutions to (4.5) is non-standard. As in [12, 29], a function $u^t : [0, t] \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a solution to (4.5) if u^t belongs to $L^\infty([0, t]; C_b^{2+\alpha}(\mathbb{R}^d))$ and satisfies

$$u_t^t(x) = g(x) \quad \text{and} \quad u_s^t(x) - u_r^t(x) = \int_r^s [-L_\theta^b u_\theta^t(x) + f_\theta(x)] d\theta \quad (4.6)$$

for every $0 \leq r \leq s \leq t$, $x \in \mathbb{R}^d$. Here for every $\theta \in [0, T]$, $L_\theta^b = \frac{1}{2} \Delta + b_\theta \cdot \nabla$.

It is shown in [29, Theorem 2.8 and Remark 2.9] that equation (4.5) has unique solution in $L^\infty([0, t]; C_b^{2+\alpha}(\mathbb{R}^d))$ for every $f \in L^\infty([0, T]; C_b^\alpha(\mathbb{R}^d))$ and $g \in C_b^{2+\alpha}(\mathbb{R}^d)$. In addition, for this solution, we have

$$\sup_{s \in [0, t]} \|u_s^t\|_{C_b^{2+\alpha}} \leq C(\alpha, d) (\|f\|_{L^\infty C_b^\alpha} + \|g\|_{C_b^{2+\alpha}}). \quad (4.7)$$

We define for every $s \leq t$ and $x \in \mathbb{R}^d$, $U_{s,t}g(x) = u_s^t(x)$, where u^t is the solution (4.5) with $f \equiv 0$. Then $\{U_{s,t}\}_{s \leq t}$ is a propagator on $C_b^{2+\alpha}(\mathbb{R}^d)$. This means that $U_{s,t}$ is a bounded linear operator on $C_b^{2+\alpha}(\mathbb{R}^d)$ and

$$U_{s,t}g(x) = U_{s,r}U_{r,t}g(x) \quad (4.8)$$

for every $s \leq r \leq t$, $x \in \mathbb{R}^d$ and $g \in C_b^{2+\alpha}(\mathbb{R}^d)$. The boundedness of $U_{s,t}$ as a linear operator on $C_b^{2+\alpha}(\mathbb{R}^d)$ follows from the bound (4.7) and the identity (4.8) follows from the fact that equation (4.5) is uniquely solvable with terminal data in $C_b^{2+\alpha}(\mathbb{R}^d)$.

At this point, we reason that $\{U_{s,t}\}$ can be extended to a propagator on $BUC(\mathbb{R}^d)$, the Banach space of all bounded uniformly continuous real functions on \mathbb{R}^d equipped with the supremum norm $\|\cdot\|_\infty$. This requires a maximum principle for equation (4.5). In [12, pages 10,11] and in [29, Theorem 4.1], a maximum principle is shown for second-order differential operators with negative potentials. These results can not be applied to

equation (4.5) with a null potential. Nevertheless, the following Itô formula, shown in [12], is useful for our considerations.

Lemma 4.2 ([12, Lemma 3]). *Let $u : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ be a function in $L^\infty([0, T]; C_b^{2+\alpha}(\mathbb{R}^d))$ such that*

$$u_t(x) - u_s(x) = \int_s^t v_r(x) dr$$

for every $s \leq t$, $x \in \mathbb{R}^d$ with $v \in L^\infty([0, T]; C_b^\alpha(\mathbb{R}^d))$. Let $(X_t)_{t \geq 0}$ be a continuous adapted process of the form

$$X_t = X_0 + \int_0^t b_s ds + W_t$$

where b is a progressively measurable process, b is integrable in t with probability one. Then

$$u_t(X_t) = u_s(X_s) + \int_s^t (v_r + b_r \cdot \nabla u_r + \frac{1}{2} \Delta u_r)(X_r) dr + \int_s^t \nabla u_r(X_r) \cdot dW_r.$$

Lemma 4.3 (Maximum principle). *For every $f \in L^\infty([0, T]; C_b^\alpha(\mathbb{R}^d))$, $g \in C_b^{2+\alpha}(\mathbb{R}^d)$ and every $s \leq t$,*

$$u_s^t(x) = \mathbb{E}g(X_{s,t}^x) - \mathbb{E} \int_s^t f_r(X_{s,r}^x) dr, \tag{4.9}$$

and

$$\|u_s^t\|_\infty \leq \|g\|_\infty + (t - s) \|f\|_{L^\infty C_b}. \tag{4.10}$$

Here, u^t is the solution to (4.5).

Proof. Let $s \leq t$ be fixed and $(X_{s,r}^x)_{r \geq s}$ be the solution to (4.1) starting from time s at $X_{s,s}^x = x$. Let u^t be the solution to (4.5). Applying the Itô formula in Lemma 4.2 to $r \mapsto u_r^t(X_{s,r}^x)$, using the relation (4.6), we have

$$g(X_{s,t}^x) = u_s^t(x) + \int_s^t f_r(X_{s,r}^x) dr + \int_s^t \nabla u_r^t(X_{s,r}^x) \cdot dW_r.$$

By (4.7), the stochastic integral above is a square integrable martingale. Taking expectation, we obtain (4.9), which is a probabilistic presentation for the solution to (4.5). The formula (4.9) implies the estimate (4.10). \square

As a consequence of (4.9), we have for every $g \in C_b^{2+\alpha}(\mathbb{R}^d)$,

$$U_{s,t}g(x) = \mathbb{E}g(X_{s,t}^x). \tag{4.11}$$

This implies that for every $s \leq t$ and $g \in C_b^{2+\alpha}(\mathbb{R}^d)$,

$$\|U_{s,t}g\|_\infty \leq \|g\|_\infty. \tag{4.12}$$

We note that $C_b^\gamma(\mathbb{R}^d)$ is dense in $BUC(\mathbb{R}^d)$ for every $\gamma > 0$. By approximation, the above inequality holds for every $s \leq t$ and $g \in BUC(\mathbb{R}^d)$. This means that $U_{s,t}$ can be extended to a bounded operator on $BUC(\mathbb{R}^d)$ for each $s \leq t$. The family $\{U_{s,t}\}$ then forms a C_0 -propagator on $BUC(\mathbb{R}^d)$. The precise meaning of this statement is the content of the following result.

Proposition 4.4. *The two-parameter family $\{U_{s,t}\}_{s \leq t}$ satisfies*

- (a) $U_{s,t}$ is a contraction on $BUC(\mathbb{R}^d)$ for every $s \leq t$,
- (b) $U_{s,t} = U_{s,r}U_{r,t}$ as operators on $BUC(\mathbb{R}^d)$ for every $s \leq r \leq t$,

(c) $\lim_{h \rightarrow 0} \sup_{s,t:|t-s| \leq h} \|U_{s,t}g - g\|_\infty = 0$ for every $g \in BUC(\mathbb{R}^d)$.

Proof. (a) is shown in the discussion prior to the proposition. Let g be in $BUC(\mathbb{R}^d)$ and $\{g^n\}_n$ be a sequence in $C_b^\alpha(\mathbb{R}^d)$ which converges to g in $BUC(\mathbb{R}^d)$. From (4.12), it follows that $\lim_n U_{s,t}g^n = U_{s,t}g$ in $BUC(\mathbb{R}^d)$. From (4.8), we have $U_{s,t}g^n = U_{s,r}U_{r,t}g^n$ for each n . Sending $n \rightarrow \infty$, using the continuity of $U_{s,r}$ and $U_{r,t}$ on $BUC(\mathbb{R}^d)$, yields $U_{s,t}g = U_{s,r}U_{r,t}g$, which implies (b). To show (c), let $u^{t,n}$ be the solution to (4.5) with $f \equiv 0$ and terminal condition $u_t^{t,n} = g^n$. We note that by triangle inequality, contraction property (a) and definition of U ,

$$\begin{aligned} \|U_{s,t}g - g\|_\infty &\leq \|U_{s,t}g^n - g^n\|_\infty + \|U_{s,t}g^n - U_{s,t}g\|_\infty + \|g^n - g\|_\infty \\ &\leq \|u_s^{t,n} - g^n\|_\infty + 2\|g^n - g\|_\infty. \end{aligned}$$

It follows directly from (4.6) and (4.7) that

$$\|u_s^{t,n} - g^n\|_\infty \leq \int_s^t \|L_\theta^b u_\theta^{t,n}\|_\infty d\theta \lesssim (t-s)\|u^{t,n}\|_{L^\infty C_b^2} \lesssim (t-s)\|g^n\|_{C_b^{2+\alpha}}.$$

Combining the previous estimates we obtain that

$$\limsup_{h \rightarrow 0} \sup_{|t-s| \leq h} \|U_{s,t}g - g\|_\infty \leq 2\|g^n - g\|_\infty$$

for each n . Finally, sending $n \rightarrow \infty$ we obtain (c). □

Remark 4.5. Let $\{Q_t\}_{t \geq 0}$ be a family of bounded operators on $BUC(\mathbb{R}^d)$. It is straightforward to verify that $\{U_{s,t}\}_{s \leq t} := \{Q_{t-s}\}_{s \leq t}$ satisfies the properties (a)-(c) listed above if and only if $\{Q_t\}_{t \geq 0}$ is a contraction C_0 -semigroup. This justifies our terminology of C_0 -propagator.

We show that $\{U_{s,t}\}$ is the transition propagator of X^x , the solution of (4.1), in the following sense.

Lemma 4.6. For every g in $BUC(\mathbb{R}^d)$ and every $s \leq t$

$$\mathbb{E}^{\mathcal{F}_s} g(X_t^x) = U_{s,t}g(X_s^x).$$

Proof. In view of (4.12), it suffices to show the result for g in $C_b^{2+\alpha}(\mathbb{R}^d)$. Applying Itô formula in Lemma 4.2 to $r \mapsto U_{r,t}g(X_r^x)$

$$g(X_t^x) = U_{s,t}g(X_s^x) + \int_s^t \nabla U_{r,t}(X_r^x) \cdot dW_r$$

Taking conditional expectation with respect to \mathcal{F}_s , noting that $\nabla U_{r,t}$ is uniformly bounded (by (4.7)), we obtain the claim. □

Next, we show that $U_{s,t}$ maps $C^{\beta-1}(\mathbb{R}^d)$ with $\beta \in (0, 1]$, to $BUC(\mathbb{R}^d)$ provided that $s < t$. For concreteness, we first describe this procedure for the heat propagator. Let $\{P_{s,t}\}$ be the propagator associated to $\frac{1}{2}\Delta$. This means that for each $s < t$, $g \in BUC(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$, $P_{s,s}g = g$ and

$$P_{s,t}g(x) := \mathbb{E}g(W_t - W_s + x) = (2\pi(t-s))^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|x-y|^2}{2(t-s)}} g(y) dy.$$

Let $s < t$ be fixed. We aim to extend the domain of $P_{s,t}$ to Hölder-Besov space $C^{\beta-1}(\mathbb{R}^d)$ without appealing to the explicit formula of the kernel. This approach has the advantage that it can be carried out for $U_{s,t}$ once certain Schauder estimates are available. We proceed by approximation. Let g be in $C^{\beta-1}(\mathbb{R}^d)$ and let $\{g^n\}_n$ be a mollifying sequence of

g. This means that $g^n(x) = \langle g(x - \cdot), \rho_n(\cdot) \rangle$, $\rho_n(y) = n^d \rho(ny)$, ρ is a smooth non-negative function ρ which equals 1 on $\{x \in \mathbb{R}^d : |x| \leq 1\}$ and vanishes outside $\{x \in \mathbb{R}^d : |x| \leq 2\}$. Hölder-Besov spaces are unfortunately not separable. One cannot expect that $\{g^n\}_n$ converges to g in $C^{\beta-1}(\mathbb{R}^d)$. Nevertheless, it is true that $\lim g^n = g$ in $C^{\beta'-1}$ for every $\beta' < \beta$. From (3.2), we have

$$\|P_{s,t}g^n - P_{s,t}g^k\|_\infty \lesssim |t - s|^{\frac{\beta'}{2} - \frac{1}{2}} \|g^n - g^k\|_{C^{\beta'-1}} \tag{4.13}$$

for every n, k . This implies that $\lim_n P_{s,t}g^n$ exists in $BUC(\mathbb{R}^d)$ and we denote the resulting limit by $P_{s,t}g$, which is a function in $BUC(\mathbb{R}^d)$. Using (3.2) once again, we see that

$$\|P_{s,t}g^n\|_\infty \lesssim |t - s|^{\frac{\beta}{2} - \frac{1}{2}} \|g^n\|_{C^{\beta-1}} \lesssim |t - s|^{\frac{\beta}{2} - \frac{1}{2}} \|g\|_{C^{\beta-1}}.$$

Sending $n \rightarrow \infty$, we obtain

$$\|P_{s,t}g\|_\infty \lesssim |t - s|^{\frac{\beta}{2} - \frac{1}{2}} \|g\|_{C^{\beta-1}}.$$

The extension of $U_{s,t}$ on $C^{\beta-1}(\mathbb{R}^d)$ is carried out through the same procedure. To obtain bounds for $U_{s,t}$ similar to (4.13), we rely on some a priori estimates for the solutions of (4.5) obtained through the mild formulation.

Lemma 4.7 (Mild formulation). *For every $f \in L^\infty([0, T]; C_b^\alpha(\mathbb{R}^d))$, $g \in C_b^{2+\alpha}(\mathbb{R}^d)$ and every $s \leq t$, $x \in \mathbb{R}^d$*

$$u_s^t(x) = P_{s,t}g(x) + \int_s^t P_{s,r}(b_r \cdot \nabla u_r^t - f_r)(x)dr. \tag{4.14}$$

Proof. Let $(W_r)_{r \geq 0}$ be a standard Brownian motion in \mathbb{R}^d and let $s \leq t$ be fixed. Define $X_r = W_r - W_s + x$ for every $r \geq s$. We note that $(W_r - W_s)_{r \geq s}$ is a standard Brownian motion starting at time s . Applying the Itô formula in Lemma 4.2 and using (4.6), we have

$$u_t^t(X_t) = u_s^t(X_s) + \int_s^t (f_r - b_r \cdot \nabla u_r^t)(X_r)dr + \int_s^t \nabla u_r^t(X_r) \cdot dW_r.$$

By (4.7), the stochastic integral above is square integrable. Hence, we can take expectation in the above formula and use the identities $u_t^t(X_t) = g(W_t - W_r + x)$, $u_s^t(X_s) = u_s^t(x)$ to obtain that

$$u_s^t(x) = \mathbb{E}g(W_t - W_s + x) + \mathbb{E} \int_s^t (b_r \cdot \nabla u_r^t - f_r)(W_r - W_s + x)dr.$$

By Fubini's theorem, we can interchange the expectation and the integration in dr in the last term of the above identity. The identity (4.14) follows by definition of P . □

Lemma 4.8. *Let g be in $C_b^{2+\alpha}(\mathbb{R}^d)$, f be in $L^\infty([0, t]; C_b^\alpha(\mathbb{R}^d))$ and u^t be the solution to (4.5) with terminal condition g at time t and right-hand side f . Let β be a fixed number in $(0, 1]$. We define for each $s \in [0, t]$,*

$$F_s^{(\beta)} = (t - s)^{1 - \frac{\beta}{2}} \int_s^t (r - s)^{-\frac{1}{2}} \|f_r\|_\infty dr \quad \text{and} \quad F_s^{(\beta),*} = \sup_{r \in [s,t]} F_r^{(\beta)}. \tag{4.15}$$

Then, we have for every $s \in [0, t]$

$$\|\nabla u_s^t\|_\infty \leq C|t - s|^{\frac{\beta}{2} - 1} \left(\|g\|_{C^{\beta-1}} + F_s^{(\beta),*} \right) \tag{4.16}$$

and

$$\|u_s^t\|_\infty \leq C|t-s|^{\frac{\beta}{2}-\frac{1}{2}}(\|g\|_{C^{\beta-1}} + F_s^{(\beta)}) + C(t-s)^{\frac{\beta}{2}}F_s^{(\beta),*} \tag{4.17}$$

for a constant $C = C(\|b\|_{L^\infty C_b}, T, \beta)$. In particular,

$$\|\nabla U_{s,t}g\|_\infty \leq C|t-s|^{\frac{\beta}{2}-1}\|g\|_{C^{\beta-1}} \quad \text{and} \quad \|U_{s,t}g\|_\infty \leq C|t-s|^{\frac{\beta}{2}-\frac{1}{2}}\|g\|_{C^{\beta-1}} \tag{4.18}$$

for every $s < t$.

Proof. From (4.14), we have

$$\|\nabla u_s^t\|_\infty \leq \|\nabla P_{s,t}g\|_\infty + \int_s^t \|\nabla P_{s,r}f_r\|_\infty dr + \int_s^t \|\nabla P_{s,r}(b_r \cdot \nabla u_r^t)\|_\infty dr.$$

From (3.2) and direct estimations, we know that

$$\|\nabla P_{s,t}g\|_\infty \lesssim |t-s|^{\frac{\beta}{2}-1}\|g\|_{C^{\beta-1}}, \quad \|\nabla P_{s,r}f_r\|_\infty \lesssim |r-s|^{-\frac{1}{2}}\|f_r\|_\infty$$

and

$$\|\nabla P_{s,r}(b_r \cdot \nabla u_r^t)\|_\infty \lesssim |r-s|^{-\frac{1}{2}}\|b_r \cdot \nabla u_r^t\|_\infty \lesssim |r-s|^{-\frac{1}{2}}\|b\|_{L^\infty C_b}\|\nabla u_r^t\|_\infty.$$

It follows that

$$\|\nabla u_s^t\|_\infty \lesssim |t-s|^{\frac{\beta}{2}-1}(\|g\|_{C^{\beta-1}} + F_s^{(\beta)}) + \|b\|_{L^\infty C_b} \int_s^t |r-s|^{-\frac{1}{2}}\|\nabla u_r^t\|_\infty dr.$$

Let λ be a positive constant to be chosen later. For every $s \in [0, t]$, we define

$$m_s = e^{-\lambda(t-s)}(t-s)^{1-\frac{\beta}{2}}\|\nabla u_s^t\|_\infty \quad \text{and} \quad m_s^* = \sup_{r \in [s,t]} m_r,$$

which are finite by (4.7). The previous estimate implies that

$$m_s \lesssim \|g\|_{C^{\beta-1}} + F_s^{(\beta)} + m_s^*\|b\|_{L^\infty C_b}(t-s)^{1-\frac{\beta}{2}} \int_s^t |r-s|^{-\frac{1}{2}}|t-r|^{\frac{\beta}{2}-1}e^{-\lambda(r-s)}dr.$$

To estimate the integral on the right-hand side, we split it into two regions

$$\begin{aligned} \int_s^{\frac{s+t}{2}} |r-s|^{-\frac{1}{2}}|t-r|^{\frac{\beta}{2}-1}e^{-\lambda(r-s)}dr &\lesssim |t-s|^{\frac{\beta}{2}-1} \int_s^\infty (r-s)^{-\frac{1}{2}}e^{-\lambda(r-s)}dr \\ &\lesssim |t-s|^{\frac{\beta}{2}-1}\lambda^{-\frac{1}{2}} \end{aligned}$$

and

$$\begin{aligned} \int_{\frac{s+t}{2}}^t |r-s|^{-\frac{1}{2}}|t-r|^{\frac{\beta}{2}-1}e^{-\lambda(r-s)}dr &\lesssim |t-s|^{-\frac{1}{2}}e^{-\frac{\lambda}{2}(t-s)} \int_{\frac{s+t}{2}}^t (t-r)^{\frac{\beta}{2}-1}dr \\ &\lesssim |t-s|^{\frac{\beta}{2}-\frac{1}{2}}e^{-\frac{\lambda}{2}(t-s)}. \end{aligned}$$

Noting that $e^{-\frac{\lambda}{2}(t-s)} \lesssim \lambda^{-\frac{1}{2}}(t-s)^{-\frac{1}{2}}$, the previous two bounds yields

$$\int_s^t |r-s|^{-\frac{1}{2}}|t-r|^{\frac{\beta}{2}-1}e^{-\lambda(r-s)}dr \lesssim |t-s|^{\frac{\beta}{2}-1}\lambda^{-\frac{1}{2}}.$$

Hence, we have

$$m_s \lesssim \|g\|_{C^{\beta-1}} + F_s^{(\beta)} + m_s^*\|b\|_{L^\infty C_b}\lambda^{-\frac{1}{2}}$$

and

$$m_s^* \lesssim \|g\|_{C^{\beta-1}} + F_s^{(\beta),*} + m_s^*\|b\|_{L^\infty C_b}\lambda^{-\frac{1}{2}}$$

for every $s \in [0, t]$. From here, choosing λ sufficiently large, we obtain that $m_s^* \leq C(\|g\|_{\mathcal{C}^{\beta-1}} + F_s^{(\beta),*})$ for every $s \in [0, t]$ for a constant $C > 0$ depending only on β and $\|b\|_{L^\infty C_b}$. This implies (4.16).

To show (4.17), we derive from (4.14) that

$$\|u_s^t\|_\infty \leq \|P_{s,t}g\|_\infty + \int_s^t \|P_{s,r}f_r\|_\infty dr + \int_s^t \|P_{s,r}(b_r \cdot \nabla u_r^t)\|_\infty dr.$$

We estimate the first term by

$$\|P_{s,t}g\|_\infty \lesssim |t - s|^{\frac{\beta}{2} - \frac{1}{2}} \|g\|_{\mathcal{C}^{\beta-1}},$$

the second term by

$$\int_s^t \|P_{s,r}f_r\|_\infty dr \leq \int_s^t \|f_r\|_\infty dr \leq (t - s)^{\frac{\beta}{2} - \frac{1}{2}} F_s^{(\beta)},$$

and the last term by

$$\|P_{s,r}(b_r \cdot \nabla u_r^t)\|_\infty \leq \|b_r\|_\infty \|\nabla u_r^t\|_\infty \leq C(\|b\|_{L^\infty C_b})(\|g\|_{\mathcal{C}^{\beta-1}} + F_s^{(\beta),*})|t - r|^{\frac{\beta}{2} - 1}$$

in which we have used (4.16) to obtain the last inequality. Combining these bounds together, we obtain (4.17). \square

Lemma 4.9. *For every $s < t$, the operator $U_{s,t} : BUC(\mathbb{R}^d) \rightarrow BUC(\mathbb{R}^d)$ can be extended uniquely to a bounded linear operator $U_{s,t} : \mathcal{C}^{\beta-1}(\mathbb{R}^d) \rightarrow BUC(\mathbb{R}^d)$. In addition, there exists a positive constant $C = C(\|b\|_{L^\infty C_b}, T, \beta)$ such that*

$$\|U_{s,t}g\|_\infty \leq C\|g\|_{\mathcal{C}^{\beta-1}}|t - s|^{\frac{\beta-1}{2}} \tag{4.19}$$

for every $g \in \mathcal{C}^{\beta-1}(\mathbb{R}^d)$.

Proof. Let $\{g^n\}_n$ be a sequence in $C_b^3(\mathbb{R}^d)$ such that g^n converges to g in $\mathcal{C}^{\beta'-1}$ for some $\beta' \in (0, \beta)$ and that $\sup_n \|g^n\|_{\mathcal{C}^{\beta-1}} \lesssim \|g\|_{\mathcal{C}^{\beta-1}}$. Let s, t be fixed but arbitrary in $[0, T]$, $s < t$. From (4.18), we have for every n, k

$$\|U_{s,t}g^n - U_{s,t}g^k\|_\infty \lesssim |t - s|^{\frac{\beta'}{2} - \frac{1}{2}} \|g^n - g^k\|_{\mathcal{C}^{\beta'-1}}.$$

This implies that $\lim_n U_{s,t}g^n$ exists in $BUC(\mathbb{R}^d)$ and the limit is denoted by $U_{s,t}g$. It is standard to check that $U_{s,t}g$ is independent from the choices of the approximating sequence $\{g^n\}_n$ and that $U_{s,t} : \mathcal{C}^{\beta-1} \rightarrow BUC(\mathbb{R}^d)$ is linear. In view of the above estimate, such extension is necessary unique in the sense that if $\bar{U}_{s,t} : \mathcal{C}^{\beta-1}(\mathbb{R}^d) \rightarrow BUC(\mathbb{R}^d)$ is another continuous extension of $U_{s,t} : BUC(\mathbb{R}^d) \rightarrow BUC(\mathbb{R}^d)$, then $U_{s,t} = \bar{U}_{s,t}$ on $\mathcal{C}^{\beta-1}(\mathbb{R}^d)$. In addition, from (4.18), we have

$$\|U_{s,t}g\|_\infty = \lim_n \|U_{s,t}g^n\|_\infty \lesssim |t - s|^{\frac{\beta}{2} - \frac{1}{2}} \limsup_n \|g^n\|_{\mathcal{C}^{\beta-1}} \lesssim |t - s|^{\frac{\beta}{2} - \frac{1}{2}} \|g\|_{\mathcal{C}^{\beta-1}},$$

which shows (4.19). \square

Proposition 4.10. *For every β in $(0, 1)$, $(X, U, \mathcal{C}^{\beta-1}(\mathbb{R}^d))$ satisfies Condition 3.1 with $\nu = \beta - 1$.*

Proof. Let $\mathcal{H} = \mathcal{C}^{\beta-1}(\mathbb{R}^d)$ and $\nu = \beta - 1$. It is evident that Lemma 4.9 implies property (i) of Condition 3.1. Let h be in $\mathcal{C}^{\beta-1}(\mathbb{R}^d)$ and let $\{h^n\}_n$ be a sequence in $C_b^3(\mathbb{R}^d)$ which converges to h in $\mathcal{C}^{\beta'-1}(\mathbb{R}^d)$ for some $\beta' \in (0, \beta)$. For each $s \geq 0$ and n , it follows from (4.11) that $t \mapsto U_{s,t}h^n(x)$ is measurable. Since $\lim_n U_{s,t}h^n = U_{s,t}h$ in $BUC(\mathbb{R}^d)$ for every

$t > s$, this implies the property (ii) of Condition 3.1. Finally, since for $r < t$, $U_{r,t}h$ belongs to $BUC(\mathbb{R}^d)$ (by Lemma 4.9), we derive from Lemma 4.6 and Proposition 4.4(b) that $\mathbb{E}^{F_s}U_{r,t}h = U_{s,r}U_{r,t}h = U_{s,t}h$ for every $s \leq r < t$. This shows property (iii) of Condition 3.1. \square

Note that ∇b belongs to $[L_T^\infty C^{\alpha-1}(\mathbb{R}^d; \mathbb{R}^d)]^d$ and $(\nu, q) = (\alpha - 1, \infty)$ satisfies condition (3.9). Applying Proposition 3.5, we can define the process V in (4.3). The estimate (3.10) implies that

$$\|V_t(b, X) - V_s(b, X)\|_{L_m} \lesssim |t - s|^{\frac{\alpha+1}{2}}$$

for every $s < t$ and $m \geq 2$. By choosing m sufficiently large and applying Kolmogorov continuity theorem, we can find a continuous version of V which is a.s. $\frac{\alpha'+1}{2}$ -Hölder continuous for every $\alpha' \in (0, \alpha)$.

Let $\{b^n\}_n$ be a sequence of functions in $L^\infty([0, T]; C_b^4(\mathbb{R}^d; \mathbb{R}^d))$ which converges to b in $L^\infty([0, T]; C^\beta(\mathbb{R}^d; \mathbb{R}^d))$ for every $\beta \in (0, \alpha)$. Let X^n, U^n denote respectively the solution to (4.1) and the transition semigroup defined in (4.5) with b^n in place of b . From [12, Theorem 5], for every $m \geq 2$, we have the following stability estimates

$$\limsup_n \mathbb{E} \sup_{x \in \mathbb{R}^d} \sup_{r \in [0, T]} |X_r^n(x) - X_r(x)|^m = 0, \tag{4.20}$$

$$\limsup_n \mathbb{E} \sup_{x \in \mathbb{R}^d} \sup_{r \in [0, T]} |\nabla X_r^n(x) - \nabla X_r(x)|^m = 0. \tag{4.21}$$

Lemma 4.11. *For every $\alpha' \in (0, \alpha)$, $\{\int_{[0, \cdot]} \nabla b_r^n(X_r^n) dr\}_n$ converges to $V(b, X)$ in $C^{\frac{1+\alpha'}{2}}([0, T])$ in probability.*

Proof. Let j, k be fixed indices in $\{1, \dots, d\}$. Put $f = \partial_k b^j \in L_T^\infty C^{\alpha-1}(\mathbb{R}^d)$, $f^n = \partial_k b^{n,j}$ (the x_k -partial derivative of the j -th component of b^n). Let α' and β be positive numbers such that $\alpha' < \beta < \alpha$. We note that each f^n belongs to $\mathfrak{G}_T(X)$ (by Lemma 4.6 and recalling the Definition 3.6) and that $\{f^n\}_n$ converges to f in $L_T^\infty C^{\beta-1}$. In addition, $(X, U^X, C^{\beta-1}(\mathbb{R}^d))$ satisfies Condition 3.1 with $\nu = \beta - 1$ (Proposition 4.10). Hence, we can apply Corollary 3.9 to see that $\{\int_{[0, \cdot]} f_r^n(X_r) dr\}_n$ converges to $\mathcal{A}^X[f]$ in $C^{(1+\alpha')/2}([0, T])$ in probability. It suffices to show that $\{\int_{[0, \cdot]} f_r^n(X_r) dr - \int_{[0, \cdot]} f_r^n(X_r^n) dr\}_n$ converges to 0 in $C^{(1+\alpha')/2}([0, T])$ in probability. We are going to apply Proposition 3.10 which requires an estimate for $D_{\alpha-1, \beta-1}(X, X^n)$. The following bounds are sufficient for this purpose

$$\| [U_{s,t}^n - U_{s,t}]g \|_\infty \lesssim \|b - b^n\|_{L^\infty C_b} \|g\|_{C^{\alpha-1}} |t - s|^{\frac{\alpha}{2}} \tag{4.22}$$

and

$$|U_{s,t}g(x) - U_{s,t}g(y)| \lesssim |t - s|^{-\frac{1}{2} + \frac{\beta}{2}} \|g\|_{C^{\alpha-1}} |x - y|^{\alpha-\beta} \tag{4.23}$$

for every $g \in C^{\alpha-1}(\mathbb{R}^d)$, every $x, y \in \mathbb{R}^d$ and every $s < t$.

Reasoning as in Lemma 4.9, it suffices to show (4.22) and (4.23) for $g \in C_b^3(\mathbb{R}^d)$. To obtain (4.22), note that the function $w_s(x) := [U_{s,t}^n - U_{s,t}]g(x)$ satisfies $w_t(x) = 0$ and

$$w_s(x) - w_r(x) = \int_r^s [-L^b w_\theta(x) + (b_\theta - b_\theta^n) \nabla U_{\theta,t}^n g(x)] d\theta$$

for every $0 \leq r \leq s \leq t$, $x \in \mathbb{R}^d$. In other words, w is the solution to (4.5) with right-hand side $f_\theta = (b_\theta - b_\theta^n) \cdot \nabla U_{\theta,t}^n g$ and terminal datum 0 at time t . To estimate f , we apply the first inequality in (4.18) to obtain

$$\|f_r\|_\infty \lesssim \|b - b^n\|_{L^\infty C_b} \|\nabla U_{r,t}^n g\|_\infty \lesssim \|b - b^n\|_{L^\infty C_b} \|g\|_{C^{\alpha-1}} |t - r|^{\frac{\alpha}{2}-1}.$$

Recall the definition of $F^{(\alpha)}$ in Lemma 4.8. It follows that

$$F_s^{(\alpha)} := (t - s)^{1 - \frac{\alpha}{2}} \int_s^t |r - s|^{-\frac{1}{2}} \|f_r\|_{\infty} dr \lesssim \|b - b^n\|_{L^{\infty} C_b} \|g\|_{C^{\alpha-1}} |t - s|^{\frac{1}{2}}$$

and hence $F_s^{(\alpha),*} \lesssim \|b - b^n\|_{L^{\infty} C_b} \|g\|_{C^{\alpha-1}} |t - s|^{\frac{1}{2}}$. Applying the estimate (4.17), we have

$$\|w_s\|_{\infty} \lesssim \|b - b^n\|_{L^{\infty} C_b} \|g\|_{C^{\alpha-1}} |t - s|^{\frac{\alpha}{2}}$$

which is equivalent to (4.22). The estimate (4.23) is obtained by interpolating the two inequalities in (4.18).

In the notation of Proposition 3.10, the inequalities (4.22) and (4.23) imply that

$$\|U^X - U^{X^n}\|_{C^{\alpha-1, \alpha-1}} \lesssim \|b - b^n\|_{L^{\infty} C_b} \quad \text{and} \quad \rho_{\beta-1}^X(h) \lesssim h^{\alpha-\beta}.$$

Hence,

$$D_{\alpha-1, \beta-1}(X, X^n) \lesssim \|b - b^n\|_{L^{\infty} C_b} + \sup_{r \in [0, T]} \|X_r - X_r^n\|_{L^{(\alpha-\beta)_m}}^{\alpha-\beta},$$

which converges to 0 as $n \rightarrow \infty$ by (4.20). From Proposition 4.10, we know that $(X, U^X, C^{\alpha-1}(\mathbb{R}^d))$ and $(X^n, U^n, C^{\alpha-1}(\mathbb{R}^d))$ satisfy Condition 3.1 with $\nu = \alpha - 1$. Hence, Proposition 3.10 is applicable and yields

$$\left\| \int_{[s,t]} f_r^n(X_r) dr - \int_{[s,t]} f_r^n(X_r^n) dr \right\|_{L_m} \lesssim \|f^n\|_{L^{\infty} C^{\beta-1}} D_{\alpha-1, \beta-1}(X, X^n) |t - s|^{\frac{1+\beta}{2}}.$$

Here, we have used the identities $\mathcal{A}_t^X[f^n] = \int_{[0,t]} f_r^n(X_r) dr$ and $\mathcal{A}_t^{X^n}[f^n] = \int_{[0,t]} f_r^n(X_r^n) dr$ from Proposition 3.7. The above inequality combined with Garsia-Rodemich-Rumsey inequality ([17]) and the fact that

$$\lim_n D_{\alpha-1, \beta-1}(X, X^n) = 0$$

implies that

$$\lim_n \left[\int_{[0,\cdot]} f_r^n(X_r) dr - \int_{[0,\cdot]} f_r^n(X_r^n) dr \right] = 0$$

in $C^{(1+\alpha')/2}([0, T])$ in probability. This completes the proof. \square

Proof of Theorem 4.1. In what follows, we fix a trajectory of V which is $\frac{1+\alpha'}{2}$ -Hölder continuous for some $\alpha' \in (0, \alpha)$. The system (4.4) is of Young-type, hence it has a unique solution Y which is $\frac{1+\alpha'}{2}$ -Hölder continuous (see e.g. [6, Theorems 2.9 and 2.16] or [26]). On the other hand, by direct differentiations, it is evident that $Y^{n,ij} := \partial_i X^{n,j}$ satisfies

$$Y_t^{n,ij} = \delta_{ij} + \sum_{k=1}^d \int_0^t Y_r^{n,ki} dV_r^{n,kj}, \quad \forall i, j \in \{1, \dots, d\},$$

where $V^n = V(b^n, X^n) = \int_{[0,\cdot]} \nabla b_r^n(X_r^n) dr$. Sending $n \rightarrow \infty$ in the above identity, applying (4.21), Lemma 4.11 and stability of Young differential equations (see [6, Theorem 2.16] or [26, Theorem 4]), we see that ∇X satisfies equation (4.4). The result follows. \square

5 Chaos expansions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space equipped with a filtration $\{\mathcal{F}_t\}$ such that \mathcal{F}_0 contains \mathbb{P} -null sets. As in Section 3, let $T > 0$ be a fixed time horizon and let $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ be a normed vector space, which contains $C_b(\mathbb{R}^d)$ and is a subset of $S'(\mathbb{R}^d)$. For each $q \in [1, \infty]$, $L_T^q \mathcal{H}$ denotes the Bochner space $L^q([0, T]; \mathcal{H})$. Let $U = \{U_{s,t}\}_{0 \leq s \leq t}$

be a two-parameter family of bounded operators on \mathcal{H} . We recall that the set $\mathfrak{G}_T(X)$ defined in Definition 3.6. Let σ be a continuous bounded function on $\mathbb{R}_+ \times \mathbb{R}^d$ and b be an element in $L_T^q \mathcal{H}$. We also assume that for every $t \geq 0$, $\sigma(t, \cdot)$ is Hölder continuous with some positive exponent. In the current section, we consider chaos expansions for pathwise solutions to the stochastic differential equation

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = x. \tag{5.1}$$

The results obtained herein extend those of [42, 41] and [22] in which SDEs with more regular coefficients were considered. Since $b(t, \cdot)$ is allowed to be a distribution, a solution to (5.1) requires a non-standard definition. A triplet $(W, X, \{\mathcal{F}_t\})$ is a pathwise solution to the stochastic differential equation (5.1) on $[0, T]$ if

- (i) W is a Brownian motion with respect to $\{\mathcal{F}_t\}$,
- (ii) X is adapted with respect to $\{\mathcal{F}_t\}$,
- (iii) there is a sequence $\{b_n\}_n$ of bounded uniformly continuous functions on $[0, T] \times \mathbb{R}^d$ converging to b in $L^q \mathcal{H}$ such that

$$X_t = x + \text{ucp-}\lim_n \int_0^t b_n(r, X_r)dr + \int_0^t \sigma(r, X_r)dW_r \quad \forall t \in [0, T]. \tag{5.2}$$

In the above, $\text{ucp-}\lim$ means convergence uniformly in t over $[0, T]$ in probability. Condition (iii) is inspired by the work [2] of Bass and Chen.

Definition 5.1. A pathwise solution $(W, X, \{\mathcal{F}_t\})$ is called regular if $(X, \{\mathcal{F}_t\})$ is a Markov process, $\mathfrak{G}_T(X)$ contains the set of bounded uniformly continuous functions on $[0, T] \times \mathbb{R}^d$ and (X, U, \mathcal{H}) satisfies Condition 3.1 with $\frac{\nu}{2} > \frac{1}{q} - \frac{1}{2}$.

A trivial example of regular pathwise solution is the SDE in Section 4. Other examples which we have in mind are the SDEs considered in [8, 5] where b is a controlled distribution in $[C([0, T]; \mathcal{C}^\nu(\mathbb{R}^d))]^d$ with $\nu > -\frac{2}{3}$. In what follows, we always assume that pathwise solutions are regular. In such case, we can define the linear map \mathcal{A}^X as in Proposition 3.5. Furthermore, for every $f \in L_T^q \mathcal{H}$, we will always work with the continuous version of $\mathcal{A}^X[f]$ chosen in Remark 3.8(iii). From Corollary 3.9, we see that $\{\int_{[0, \cdot]} b_n(r, X_r)dr\}_n$ converges to $\mathcal{A}^X[b]$ in $C^\tau([0, T])$ in probability for some (in fact any) $\tau \in (\frac{1}{2}, 1 + \frac{\nu}{2} - \frac{1}{q})$. Consequently, the equation (5.2) can be written as

$$X_t = x + \mathcal{A}_t^X[b] + \int_0^t \sigma(r, X_r)dW_r \quad \forall t \geq 0. \tag{5.3}$$

In particular, we see that X is a.s. α -Hölder continuous for every $\alpha \in (0, \frac{1}{2})$.

We define the following formal differential symbols

$$L_t := \sum_{i,j=1}^d \frac{1}{2} (\sigma^\dagger \sigma)^{i,j}(t, \cdot) \partial_{x_i x_j}^2 \quad \text{and} \quad L_t^b = L_t + \sum_{j=1}^d b_j(t, \cdot) \partial_{x_j}.$$

and introduce the notation

$$\partial_j^\sigma g(r, x) = (\sigma^\dagger \nabla)_j g(r, x) = \sum_{i=1}^d \sigma^{i,j}(r, x) \partial_{x_i} g(r, x)$$

and $\nabla^\sigma = (\partial_1^\sigma, \dots, \partial_d^\sigma)$. The following change-of-variable formula is a slight extension of the classical Itô formula.

Lemma 5.2 (Itô formula). *For every u in $C_b^2([0, T] \times \mathbb{R}^d)$ and $t \in [0, T]$, we have*

$$u(t, X_t) = u(0, x) + \int_0^t (\partial_r + L_r)u(r, X_r)dr + \int_0^t \nabla u(r, X_r) \cdot \mathcal{A}^X[b](dr) + \int_0^t \nabla^\sigma u(r, X_r) \cdot dW_r, \quad (5.4)$$

where the integration with respect to $\mathcal{A}^X[b](dr)$ is in Young sense.

Proof. Without loss of generality, assume $t = 1$ in (5.4). For brevity, we put $\mathcal{A} = \mathcal{A}^X[b]$ and $M = \int_0^\cdot \sigma(r, X_r)dW_r$. We know that \mathcal{A} is a.s. Hölder continuous with some exponent $\tau > \frac{1}{2}$ and that M is an L_m -martingale for every $m \geq 2$. In addition, M satisfies condition (2.35) and $[M]_t = \int_0^t (\sigma^\dagger \sigma)(r, X_r)dr$. Heuristically, to show (5.4), we use Taylor expansion and the facts that $[\mathcal{A}] = [\mathcal{A}, M] = 0$. Consider first the case $u \in C_b^3([0, T] \times \mathbb{R}^d)$. For every $s < t$, put $A_{s,t} = u(t, X_t) - u(s, X_s)$. Obviously, $\mathcal{I}_1[A] = u(1, X_1) - u(0, X_0)$. From Taylor expansion, we have

$$\begin{aligned} A_{s,t} &= \partial_t u(s, X_s)(t - s) + \langle \nabla u(s, X_s), X_t - X_s \rangle \\ &\quad + \frac{1}{2} \langle \nabla^2 u(s, X_s), (X_t - X_s) \otimes (X_t - X_s) \rangle + R(t, X_t, s, X_s) \\ &=: A_{s,t}^{(1)} + A_{s,t}^{(2)} + A_{s,t}^{(3)} + A_{s,t}^{(4)}, \end{aligned}$$

where the Taylor remainder $R(t, x, s, y)$ satisfies

$$|R(t, x, s, y)| \leq \|u\|_{C_b^3}(|t - s|^2 + |x - y|^3).$$

It is clear that $\mathcal{I}_1[A^{(1)}] = \int_0^1 \partial_t u(r, X_r)dr$. We write

$$A_{s,t}^{(2)} = \langle \nabla u(s, X_s), A_t - A_s \rangle + \langle \nabla u(s, X_s), M_t - M_s \rangle.$$

The action of \mathcal{I} to each term on the right-hand side above is computed using Young's theory [45] and Itô stochastic theory respectively, which yields

$$\mathcal{I}_1[A^{(2)}] = \int_0^1 \nabla u(r, X_r) \cdot dA_r + \int_0^1 \langle \nabla u(r, X_r), \sigma(r, X_r)dW_r \rangle.$$

For $A^{(3)}$, we define $A_{s,t}^{(3a)} = \frac{1}{2} \langle \nabla^2 u(s, X_s), (M_t - M_s) \otimes (M_t - M_s) \rangle$ and $A_{s,t}^{(3b)} = A_{s,t}^{(3)} - A_{s,t}^{(3a)}$. As in Example 2.12, we have

$$\mathcal{I}_1[A^{(3a)}] = \frac{1}{2} \int_0^1 \langle \nabla^2 u(r, X_r), d[M]_r \rangle = \frac{1}{2} \int_0^1 \langle \nabla^2 u(r, X_r), (\sigma^\dagger \sigma)(r, X_r) \rangle dr.$$

We note that $\|A_{s,t}^{(3b)}\|_{L_2} \lesssim |t - s|^{1+\varepsilon}$ and $\|A_{s,t}^{(4)}\|_{L_2} \lesssim |t - s|^{1+\varepsilon}$ for some $\varepsilon > 0$. Applying Proposition 2.7, we see that $\mathcal{I}[A^{(3b)}] = \mathcal{I}[A^{(4)}] = 0$. This implies (5.4) with u in $C_b^3([0, 1] \times \mathbb{R}^d)$.

In the case $u \in C_b^2([0, T] \times \mathbb{R}^d)$, take a mollifying sequence $\{u^n\}_n$ in $C_b^3([0, T] \times \mathbb{R}^d)$ such that $\partial_t u^n, \nabla u^n, \nabla^2 u^n$ converge uniformly over compact sets respectively to $\partial_t u, \nabla u, \nabla^2 u$. The Itô formula (5.4) for u is obtained by passing through the limit from the Itô formula for each u^n . In such procedure, the convergence of the Young integrals is an almost sure convergence, the convergence of the Itô stochastic integrals is in probability. \square

To proceed further, we need to introduce an additional structure on \mathcal{H} . A normed vector space $\mathcal{V} \subset \mathcal{S}'(\mathbb{R}^d)$ is called an \mathcal{H} -multiplier if the multiplication

$$\cdot : \mathcal{V} \times \mathcal{H} \rightarrow \mathcal{H}$$

is a continuous bilinear map and is an extension of classical multiplication operation. An extension of classical multiplication operation means that if $(g, f) \in \mathcal{V} \times \mathcal{H} \cap C_b(\mathbb{R}^d) \times C_b(\mathbb{R}^d)$ then $g \cdot f = gf$, the product of two continuous functions.

Definition 5.3. A normed vector space $\mathcal{H} \subset \mathcal{S}'(\mathbb{R}^d)$ is admissible if it contains $C_b(\mathbb{R}^d)$ and $C_b^1(\mathbb{R}^d)$ is an \mathcal{H} -multiplier.

We note that if \mathcal{H} is admissible, for every $t \in [0, T]$, $[\partial_t + L_t^b](f_t)$ belongs to \mathcal{H} for every f in $C_b^2([0, T] \times \mathbb{R}^d)$.

An interesting choice of the function u in (5.4) is $u(r, x) = U_{r,t}\phi(x)$, where t is a fixed terminal time and ϕ is a regular function. Heuristically, such u is a solution to the Kolmogorov backward equation

$$(\partial_t + L_r^b)u(r, \cdot) = 0, \quad u(t, \cdot) = \phi(\cdot).$$

However, unlike the cases treated in [29], a satisfactory theory for such equation is not yet available in our general setup. We therefore proceed with an approximation and impose sufficient conditions for its convergence. In the following, we denote $C^+ = \cup_{\alpha \in (0,1)} C_b^\alpha(\mathbb{R}^d)$. Let $\{b_n\}_n$ be a sequence of bounded uniformly continuous functions on $[0, T] \times \mathbb{R}^d$. For each n , $\phi \in C^+$ and $t \in [0, T]$, a function $u : [0, t] \times \mathbb{R}^d \rightarrow \mathbb{R}$, let $U_{r,t}^n \phi(r, x)$ denote a solution of the parabolic backward equation

$$(\partial_t + L + b_n \cdot \nabla)u(r, x) = 0, \quad u(t, x) = \phi(x) \quad \forall (s, x) \in [0, t] \times \mathbb{R}^d. \quad (5.5)$$

Condition 5.4. There exists a sequence of bounded uniformly continuous functions $\{b_n\}$ converging to b in $L^q \mathcal{H}$ such that for every $t' < t$ and every $\phi \in C^+$

- (i) $\nabla U_{t',t} \phi$ is a function in C^+ ;
- (ii) $\lim_{s \uparrow t} U_{s,t} \phi(x) = \phi(x)$ uniformly in x over compact sets of \mathbb{R}^d ;
- (iii) $U^n \phi$ exists, $(s, x) \rightarrow U_{s,t}^n \phi(x)$ belongs to $C_b^2([0, t'] \times \mathbb{R}^d)$ and $\|U_{s,t}^n \phi\|_\infty \leq \|\phi\|_\infty$;
- (iv) $(s, x) \rightarrow U_{s,t}^n \phi(x)$ converges to $(s, x) \rightarrow U_{s,t} \phi(x)$ uniformly over compact sets of $[0, t'] \times \mathbb{R}^d$ and in $L^2([0, t']; L^\infty(\mathbb{R}^d))$;
- (v) $(s, x) \rightarrow \nabla U_{s,t}^n \phi(x)$ converges to $(s, x) \rightarrow \nabla U_{s,t} \phi(x)$ in $L^2([0, t']; L^\infty(\mathbb{R}^d; \mathbb{R}^d))$;
- (vi) $(s, x) \rightarrow 1_{[0,t']}(s)(b - b_n)(s, x) \cdot \nabla U_{s,t}^n \phi(x)$ converges to 0 in $L_T^q \mathcal{H}$.

Remark 5.5. Condition 5.4 implies that $U_{s,t}$ is a contraction map on $BUC(\mathbb{R}^d)$ for every $s < t$. Indeed, for every $\phi \in C^+$, $U_{s,t} \phi$ belongs to $BUC(\mathbb{R}^d)$ (by 5.4(i)) and satisfies (by 5.4(iii),(iv))

$$\|U_{s,t} \phi\|_\infty = \lim_n \|U_{s,t}^n \phi\|_\infty \leq \|\phi\|_\infty.$$

By a density argument, the previous inequality also holds for ϕ in $BUC(\mathbb{R}^d)$.

Theorem 5.6. Assuming Condition 5.4. Let ϕ be a function in C^+ . Then for every $s < t$,

$$\phi(X_t) = U_{s,t} \phi(X_s) + \sum_{j=1}^d \int_s^t \partial_j^\sigma U_{r,t} \phi(X_r) dW_r^j. \quad (5.6)$$

Proof. Let $\{b_n\}_n$ be the sequence given in Condition 5.4 and t' be a fixed number in $[s, t)$. We note that

$$(\partial_r + L_r)U_{r,t}^n \phi(x) = -b_n(r, x) \cdot \nabla U_{r,t}^n \phi(x) \quad \forall (r, x) \in [s, t'] \times \mathbb{R}^d,$$

and that

$$\int_0^{\cdot} b_n(r, X_r) \cdot \nabla U_{r,t}^n \phi(X_r) dr = \mathcal{A}[b_n \cdot \nabla U_{\cdot,t}^n \phi]$$

by uniform continuity of $b_n(r, \cdot) \cdot \nabla U_{r,t}^n \phi(\cdot)$, Definition 5.1 and Proposition 3.7. Hence, applying the Itô formula (5.4) for $u(r, x) = U_{r,t}^n \phi(x)$ where $r \in [s, t']$, we get

$$U_{t',t}^n \phi(X_{t'}) = U_{s,t}^n \phi(X_s) + \mathcal{A}_{t'}^X [(b - b_n) \cdot \nabla U_{\cdot,t}^n \phi] - \mathcal{A}_s^X [(b - b_n) \cdot \nabla U_{\cdot,t}^n \phi] + \sum_{j=1}^d \int_s^{t'} \partial_j^\sigma U_{r,t}^n \phi(X_r) dW_r^j. \quad (5.7)$$

Using Condition 5.4(vi) and Proposition 3.5, we have

$$\lim_n \mathcal{A}_{t'}^X [(b - b_n) \cdot \nabla U_{\cdot,t}^n \phi] = \lim_n \mathcal{A}_s^X [(b - b_n) \cdot \nabla U_{\cdot,t}^n \phi] = 0$$

in probability. In addition, using Condition 5.4(v), it is straightforward to verify that

$$\lim_n \left\| \int_s^{t'} \sigma^{i,j}(r, X_r) (\partial_i U_{r,t}^n \phi(X_r) - \partial_i U_{r,t} \phi(X_r)) dW_r^j \right\|_{L^2(\Omega)} = 0.$$

Hence, we send $n \rightarrow \infty$ in (5.7) to obtain

$$U_{t',t} \phi(X_{t'}) = U_{s,t} \phi(X_s) + \sum_{j=1}^d \int_s^{t'} \partial_j^\sigma U_{r,t} \phi(X_r) dW_r^j. \quad (5.8)$$

Using Condition 5.4(ii) and the fact that X is continuous, we see that

$$\lim_{t' \uparrow t} U_{t',t} \phi(X_{t'}) = \phi(X_t).$$

On the other hand, we rise the relation (5.8) to second power, take conditional expectation with respect to \mathcal{F}_s and use Itô isometry to obtain that

$$\mathbb{E}^{\mathcal{F}_s} |U_{t',t} \phi(X_{t'})|^2 = \mathbb{E}^{\mathcal{F}_s} |U_{s,t} \phi(X_s)|^2 + \int_s^{t'} \mathbb{E}^{\mathcal{F}_s} |\nabla^\sigma U_{r,t} \phi(X_r)|^2 dr.$$

This implies that

$$\int_s^{t'} \mathbb{E}^{\mathcal{F}_s} |\nabla^\sigma U_{r,t} \phi(X_r)|^2 dr \leq \mathbb{E}^{\mathcal{F}_s} |U_{t',t} \phi(X_{t'})|^2 \leq \|\phi\|_\infty^2$$

for every $t' < t$. Hence, by martingale convergence theorem, the limit

$$\lim_{t' \uparrow t} \sum_{j=1}^d \int_s^{t'} \partial_j^\sigma U_{r,t} \phi(X_r) dW_r^j$$

exists a.s. and in L_2 , which we denote by $\sum_{j=1}^d \int_s^t \partial_j^\sigma U_{r,t} \phi(X_r) dW_r^j$. We can send t' to t from below in (5.8) to obtain formula (5.6). \square

For each positive integer n , denote $[d]^n = \{1, \dots, d\}^n$. An element j in $[d]^n$ is an n -tuple $j = (j_1, \dots, j_n)$ such that $j_i \in \{1, \dots, d\}$ for every $i = 1, \dots, n$. For each $j \in [d]^n$, let $I_{s,t}^j$ denote the n -fold iterated integral with respect to W^{j_1}, \dots, W^{j_n} over the interval $[s, t]$. That is for every $f \in L^2([s, t]^n)$

$$I_{s,t}^j(f) = \int_{[s,t]^n_{>}} f(r_1, \dots, r_n) dW_{r_n}^{j_n} \cdots dW_{r_1}^{j_1}$$

where $[s, t]^n_{>} = \{(r_1, \dots, r_n) \in [s, t]^n : t > r_1 > r_2 > \dots > r_n > s\}$.

From Condition 5.4(i), we see that if ϕ is a function in C^+ , then so is $\partial_j^\sigma U_{r,t} \phi$ for every $r < t$. Hence, we can apply (5.6) for $\partial_j^\sigma U_{r,t} \phi$ to obtain

$$\partial_j^\sigma U_{r,t} \phi(X_r) = U_{s,r} \partial_j^\sigma U_{r,t} \phi(X_s) + \sum_{k \in [d]} \int_s^r \partial_k^\sigma U_{r_1,r} \partial_j^\sigma U_{r,t} \phi(X_{r_1}) dW_{r_1}^k$$

and so

$$\begin{aligned} \phi(X_t) &= U_{s,t} \phi(X_s) + \sum_{j \in [d]} \int_s^t U_{s,r} \partial_j^\sigma U_{r,t} \phi(X_s) dW_r^j \\ &\quad + \sum_{(j,k) \in [d]^2} \int_s^t \int_s^r \partial_j^\sigma U_{r_1,r} \partial_k^\sigma U_{r,t} \phi(X_{r_1}) dW_{r_1}^k dW_r^j. \end{aligned} \quad (5.9)$$

It is evident that this procedure can be iterated.

Theorem 5.7. *Assuming Condition 5.4. Let ϕ be a function in C^+ . For every integer $n \geq 1$ and every $0 \leq s \leq t$*

$$\phi(X_t) = U_{s,t} \phi(X_s) + \sum_{k=1}^n \sum_{j \in [d]^k} I_{s,t}^j(f_{s,t}^j) + \sum_{j \in [d]^{n+1}} I_{s,t}^j(g_{s,t}^j) \quad (5.10)$$

where for every $j = (j_1, \dots, j_n) \in [d]^n$ and every $(s_1, \dots, s_n) \in [s, t]^n_{>}$

$$g_{s,t}^j(s_1, \dots, s_n) = \partial_{j_n}^\sigma U_{s_n, s_{n-1}} \cdots \partial_{j_1}^\sigma U_{s_1, t} \phi(X_{s_n}), \quad (5.11)$$

and

$$f_{s,t}^j(s_1, \dots, s_n) = \mathbb{E} \left(g_{s,t}^j(s_1, \dots, s_n) | \mathcal{F}_s \right) = U_{s, s_n} \partial_{j_n}^\sigma U_{s_n, s_{n-1}} \cdots \partial_{j_1}^\sigma U_{s_1, t} \phi(X_s). \quad (5.12)$$

Proof. Straightforward. □

Let $\{\mathcal{F}_t^W\}$ be the augmented filtration generated by W which satisfies the usual conditions.

Theorem 5.8. *Assuming Condition 5.4. For every function ϕ in C^+ , $\phi(X_t)$ is \mathcal{F}_t^W -measurable if and only if*

$$\lim_{n \rightarrow \infty} \sum_{j \in [d]^n} \int_{(s_1, \dots, s_n) \in [0, t]^n_{>}} U_{0, s_n} [\partial_{j_n}^\sigma U_{s_n, s_{n-1}} \cdots \partial_{j_1}^\sigma U_{s_1, t} \phi(x)]^2 ds_1 \dots ds_n = 0. \quad (5.13)$$

In such case, for every $s \leq t$, we have

$$\phi(X_t) = U_{s,t} \phi(X_s) + \sum_{k=1}^{\infty} \sum_{j \in [d]^k} I_{s,t}^j(f_{s,t}^j), \quad (5.14)$$

where f^j 's are defined in (5.12) and the series converges in $L^2(\Omega)$.

Proof. The argument is essentially that of [41, 42]. The formula (5.10) with $s = 0$ gives

$$\phi(X_t) = U_{0,t}\phi(x) + \sum_{k=1}^n \sum_{j \in [d]^k} I_{0,t}^j(f_{0,t}^j) + \sum_{j \in [d]^{n+1}} I_{0,t}^j(g_{0,t}^j). \tag{5.15}$$

We note from (5.12) that each $f_{0,t}^j$ is deterministic. We recall that the space $L^2(\Omega, \mathcal{F}_t^W)$ admits the Wiener-Itô chaos decomposition

$$L^2(\Omega, \mathcal{F}_t^W) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n$$

where for each $n \geq 0$, \mathcal{H}_n is the closed linear space spanned by $\{I_{0,t}^j(f) : f \in L^2([0, t]^n), j \in [d]^n\}$, see [36, Theorems 1.1.1 and 1.1.2]. Each \mathcal{H}_n is called the Wiener chaos of order n . In particular, $I_{0,t}^j(f_{0,t}^j)$ belongs to \mathcal{H}_n for each $j \in [d]^n$.

Suppose that $\phi(X_t)$ is \mathcal{F}_t^W -measurable. Since ϕ is bounded, it is obvious that $\phi(X_t)$ is square integrable. Then by Wiener-Itô decomposition, we have

$$\phi(X_t) = \sum_{k=0}^{\infty} G_k \tag{5.16}$$

where each G_k belongs to \mathcal{H}_k . Comparing (5.16) with (5.15) and taking into account orthogonality reveals that $G_0 = U_{0,t}\phi(x)$, $G_k = \sum_{j \in [d]^k} I_{0,t}^j(f_{0,t}^j)$ for $1 \leq k \leq n$ and

$$\sum_{j \in [d]^{n+1}} I_{0,t}^j(g_{0,t}^j) = \sum_{k=n+1}^{\infty} G_k.$$

Since (5.16) holds in $L^2(\Omega, \mathcal{F}_t^W)$, this implies that

$$\lim_n \mathbb{E} \left| \sum_{j \in [d]^{n+1}} I_{0,t}^j(g_{0,t}^j) \right|^2 = \lim_n \mathbb{E} \left| \sum_{k=n+1}^{\infty} G_k \right|^2 = 0$$

Taking into account (5.11), we see that the above limit is equivalent to (5.13).

On the other hand, if (5.13) holds, then

$$\lim_n \mathbb{E} \left| \sum_{j \in [d]^{n+1}} I_{0,t}^j(g_{0,t}^j) \right|^2 = 0.$$

One can send $n \rightarrow \infty$ in (5.10) to obtain

$$\phi(X_t) = U_{0,t}\phi(x) + \sum_{k=1}^{\infty} \sum_{j \in [d]^k} I_{0,t}^j(f_{0,t}^j).$$

where the series converges in $L^2(\Omega)$. Since each term on the right-hand side of the above identity is \mathcal{F}_t^W -measurable, it is evident that $\phi(X_t)$ is \mathcal{F}_t^W -measurable. The identity (5.14) is obtained analogously to the previous one. \square

We recall that a pathwise solution $(W, X, \{\mathcal{F}_t\})$ is called strong if X_t is \mathcal{F}_t^W -measurable for every $t \geq 0$.

Corollary 5.9. *A regular pathwise solution $(W, X, \{\mathcal{F}_t\})$ is a strong solution if and only if (5.13) holds for every function ϕ in C^+ and every $t \geq 0$.*

Proof. Suppose that (5.13) holds for every function in $\phi \in C^+$ and every $t \geq 0$. Let $\{\phi_n\}$ be a sequence of functions in C^+ which converges to the identity map uniformly over compact sets. Then, for every $t \geq 0$, $\phi_n(X_t)$ is \mathcal{F}_t^W -measurable. This implies that X_t is \mathcal{F}_t^W -measurable for every $t \geq 0$ and hence (X, W) is a strong solution. The other direction is straightforward from Theorem 5.8. \square

6 SDEs driven by fractional Brownian motions

Let B^H be a standard fractional Brownian motion in \mathbb{R}^d with Hurst parameter $H \in (0, \frac{1}{2})$. This means that B^H is a Gaussian process in \mathbb{R}^d with $B_0 = 0$ and covariance given by

$$\mathbb{E}(B_t^H \otimes B_s^H) = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H})I_d. \quad \text{for all } s, t \geq 0. \quad (6.1)$$

We consider the stochastic differential equation driven by fractional Brownian motion

$$X_t = x + \int_0^t b(r, X_r)dr + B_t^H \quad \forall t \in [0, T], \quad (6.2)$$

where $x \in \mathbb{R}^d$, b is a Borel function in $[L^q([0, T]; L^p(\mathbb{R}^d))]^d$, $p, q \in [1, \infty]$. In particular, when $p = q = \infty$, b is uniformly bounded over $[0, T] \times \mathbb{R}^d$. Since b is only measurable in time, the integral $\int_0^t b(r, X_r)$ should be interpreted in Lebesgue sense. We will always view each element of $[L^q([0, T]; L^p(\mathbb{R}^d))]^d$ as a measurable function instead of an equivalent class of measurable functions. From this point of view, there is no ambiguity for writing $b(r, X_r)$.

We will show in the current section that two adapted solutions defined on the same filtered probability space coincide provided that

$$H \frac{d}{p} + \frac{1}{q} < \frac{1}{2} - H. \quad (6.3)$$

Hereafter, we use the convention $1/\infty = 0$. Existence and uniqueness in law of weak solutions were obtained earlier by Nualart and Ouknine [38] under the condition

$$H \frac{d}{p} + \frac{1}{q} < \frac{1}{2}. \quad (6.4)$$

In fact, the authors of [38] consider (6.2) in one dimension, however, their arguments, which rely on Girsanov transformation, can be easily extended to multi-dimensions, as we will explain below. Dimension one is special, because of the validity of a comparison principle for (6.2). In this case, pathwise uniqueness holds under condition (6.4). To obtain pathwise uniqueness in high dimensions, we rely on stochastic sewing lemma, Theorem 2.3, which leads to the more restricted condition (6.3).

Equation (6.2) was considered earlier in [37, 38, 1] and in [30] when $H = 1/2$. Nualart and Ouknine in [37] obtain strong existence and pathwise uniqueness for (6.2) in one dimension, where b is a measurable function with linear growth. Modulo a localization procedure, this result corresponds to $p = q = \infty$ in our case. In [38], the same authors extend their previous result in [37] to allow drifts in $L^q([0, T]; L^p(\mathbb{R}))$, with $H \frac{1}{p} + \frac{1}{q} < \frac{1}{2}$.

Baños et.al. [1] obtained existence of a unique strong solution for (6.2) in multi-dimensions, $b \in L^1(\mathbb{R}^d; L^\infty([0, T], \mathbb{R}^d)) \cap L^\infty(\mathbb{R}^d; L^\infty([0, T]; \mathbb{R}^d))$ and $H < \frac{1}{2(2+d)}$. This is also a particular case of the results stated here. The approaches of [38, 1], which rely on Girsanov transformation, comparison principle and Malliavin calculus, are different from the one presented here. Although the condition (6.3) is not optimal, it appears to be new and is the best available for stochastic differential equations driven by fractional Brownian motions in multi-dimensions. We point the readers to the end of the current section for further discussions.

To state the main results precisely, let us recall some relevant definitions. For each $t \geq 0$, we denote by $\mathcal{F}_t^{B^H}$ the σ -field generated by $\{B_r^H : r \in [0, t]\}$ and the sets of null probability. In other words, $\{\mathcal{F}_t^{B^H}\}$ is the augmented filtration generated by B^H . One can find a Wiener process W such that

$$B_t^H = \int_0^t K_H(t, s)dW_s, \quad \forall t \geq 0. \quad (6.5)$$

where K_H is the square integrable kernel given by

$$K_H(t, s) = c_H \left[\left(\frac{t}{s}\right)^{H-\frac{1}{2}} (t-s)^{H-\frac{1}{2}} - \left(H - \frac{1}{2}\right) s^{\frac{1}{2}-H} \int_s^t u^{H-\frac{3}{2}} (u-s)^{H-\frac{1}{2}} du \right],$$

for every $s < t$ and $c_H = \left(\frac{2H\Gamma(\frac{3}{2}-H)}{\Gamma(H+\frac{1}{2})\Gamma(2-2H)}\right)^{1/2}$. A procedure to find W is described in [36, Section 5.1.3]. For the formula of K_H above, we refer to [36, Proposition 5.1.3]. In addition, W can be chosen so that B^H and W generate the same filtration ([36, page 280]).

Let $(\mathcal{F}_t)_{t \geq 0}$ be a right-continuous filtration such that \mathcal{F}_0 contains the \mathbb{P} -null sets. Following [37], we say that B^H is an $\{\mathcal{F}_t\}$ -fractional Brownian motion if there exists an $\{\mathcal{F}_t\}$ -Wiener process W such that (6.5) holds. The previous paragraph can be summarized in the statement that every standard fractional Brownian motion B^H is an $\{\mathcal{F}_t^{B^H}\}$ -fractional Brownian motion.

An essential property of $\{\mathcal{F}_t\}$ -fractional Brownian motion which we exploit is the local nondeterminism property. Namely, there exists a positive constant c such that

$$\sigma_H(s, t) \geq c|t-s|^H \quad \forall s < t, \tag{6.6}$$

where

$$\sigma_H^2(s, t) := \frac{1}{d} \mathbb{E}(|B_t^H - \mathbb{E}^{\mathcal{F}_s} B_t^H|^2) = \int_s^t |K_H(t, r)|^2 dr. \tag{6.7}$$

This property was used in [38] without proof. We give a simple proof of (6.6). Without loss of generality, assume $d = 1$. The fractional Brownian motion B^H admits a moving average representation which was introduced by Mandelbrot and Van Ness [33]

$$B_t^H = C(H) \int_{-\infty}^t [(t-r)_+^{H-\frac{1}{2}} - (-r)^{H-\frac{1}{2}}] d\widetilde{W}_r \quad \text{for every } t \geq 0,$$

where $C(H)$ is a suitable normalizing constant and \widetilde{W} is a two-sided Brownian motion. For each $t \geq 0$, let $\widetilde{\mathcal{F}}_t$ the the σ -algebra generated by $\{\widetilde{W}_r : r \leq t\}$ and the sets of null probability. Clearly B_t^H is $\widetilde{\mathcal{F}}_t$ -measurable and hence $\mathcal{F}_t^{B^H} \subset \widetilde{\mathcal{F}}_t$. It follows that

$$\int_s^t |K_H(t, r)|^2 dr = \mathbb{E}(B_t^H - \mathbb{E}^{\mathcal{F}_s^{B^H}} B_s^H)^2 \geq \mathbb{E}(B_t^H - \mathbb{E}^{\widetilde{\mathcal{F}}_s} B_t^H)^2.$$

Using the moving average representation, it is straightforward to see that

$$\mathbb{E}(B_t^H - \mathbb{E}^{\widetilde{\mathcal{F}}_s} B_t^H)^2 = \mathbb{E} \left(C(H) \int_s^t (t-r)^{H-\frac{1}{2}} d\widetilde{W}_r \right)^2 = C_1(H) |t-s|^{2H}$$

where $C_1(H) = (C(H)(H + \frac{1}{2})^{-1})^2$ is a positive number. From here we obtain (6.6).

By a pathwise solution to (6.2) on $(\Omega, \mathcal{F}, \mathbb{P})$, we mean a triplet $(B^H, X, \{\mathcal{F}_t\}_{t \in [0, T]})$ such that

- (i) $\{\mathcal{F}_t\}$ is a right-continuous filtration, \mathcal{F}_0 contains \mathbb{P} -null sets, X and B^H are $\{\mathcal{F}_t\}$ -adapted,
- (ii) B^H is an $\{\mathcal{F}_t\}$ -fractional Brownian motion,
- (iii) X and B^H satisfy (6.2) a.s. for every $t \in [0, T]$.

A pathwise solution is called a strong solution if X is adapted to the augmented filtration generated by B^H , i.e. $\{\mathcal{F}_t^{B^H}\}$. Since B^H and W generate the same filtration, X is a strong solution iff it is adapted to the augmented filtration generated by W .

A weak solution to (6.2) is a tuple $((\Omega, \mathcal{F}, \mathbb{P}), B^H, X, \{\mathcal{F}_t\}_{t \in [0, T]})$ such that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space and $(B^H, X, \{\mathcal{F}_t\}_{t \in [0, T]})$ is a pathwise solution to (6.2) on $(\Omega, \mathcal{F}, \mathbb{P})$.

The following theorems are our main results in the current section.

Theorem 6.1 (Nualart-Ouknine). *Suppose that (6.4) holds. Then equation (6.2) has a weak solution. Any two weak solutions have the same probability law.*

Theorem 6.2 (Pathwise uniqueness). *Suppose that (6.3) holds. Let $(B^H, X, \{\mathcal{F}_t\})$ and $(B^H, \bar{X}, \{\mathcal{F}_t\})$ be two pathwise solutions to (6.2) starting from x on $(\Omega, \mathcal{F}, \mathbb{P})$. Then X and \bar{X} are indistinguishable.*

Theorem 6.3 (Strong existence). *Suppose that (6.3) holds. Then (6.2) has a unique strong solution.*

Theorem 6.1 is an easy extension of [37, 38] using Girsanov transformation. Theorem 6.2 relies on Lipschitz property of the map $\psi \mapsto \int b(r, B_r^H + \psi_r) dr$, which is derived using Theorem 2.3. The notion of $\{\mathcal{F}_t\}$ -fractional Brownian motion and formula (6.5) allow us to view equation (6.2) as a stochastic integral equation driven by a Wiener process. Having established weak existence and pathwise uniqueness for equation (6.2) in Theorems 6.1 and 6.2, an application of Yamada-Watanabe principle [43] yields Theorem 6.3.

Weak existence and uniqueness in probability law

Let us now discuss in more detail. We begin with the following lemma, which extends [38, Lemma 3.2].

Lemma 6.4. *Let B^H be an $\{\mathcal{F}_t\}$ -fractional Brownian motion. Let $(p, q) \in [1, \infty]^2$ satisfy $H \frac{d}{p} + \frac{1}{q} < 1$. For every non-negative measurable function h in $L^q([0, T]; L^p(\mathbb{R}^d))$, we have*

$$\mathbb{E}^{\mathcal{F}_s} \int_s^T h(t, B_t^H) dt \lesssim \|h\|_{L^q L^p} (T - s)^{1 - H \frac{d}{p} - \frac{1}{q}} \tag{6.8}$$

and

$$\mathbb{E} \exp \left(\int_0^T h(t, B_t^H) dt \right) < \infty, \tag{6.9}$$

where $q' = \frac{q}{q-1}$ is the Hölder conjugate of q .

Proof. Note that $q' < \infty$ because $q > 1$. By Tonelli's theorem, we have

$$\mathbb{E}^{\mathcal{F}_s} \int_s^T h(t, B_t^H) dt = \int_s^T \mathbb{E}^{\mathcal{F}_s} h(t, B_t^H) dt = \int_s^T \mathbb{E}^{\mathcal{F}_s} h(t, \mathbb{E}^{\mathcal{F}_s} B_t^H + (B_t^H - \mathbb{E}^{\mathcal{F}_s} B_t^H)) dt.$$

Using the definition of $\{\mathcal{F}_t\}$ -fractional Brownian motion, (6.5) and (6.7), we see that $B_t^H - \mathbb{E}^{\mathcal{F}_s} B_t^H = \int_s^t K_H(t, r) dW_r$, which has a centered normal distribution with covariance matrix $\sigma_H^2(s, t) I_d$ and is independent from \mathcal{F}_s . Hence, we obtain

$$\mathbb{E}^{\mathcal{F}_s} \int_s^T h(t, B_t^H) dt = \int_s^T P_{\sigma_H^2(s, t)} [h(t, \cdot)] (\mathbb{E}^{\mathcal{F}_s} B_t^H) dt,$$

where $\{P_\sigma\}_{\sigma \geq 0}$ is the heat semigroup associated to the kernel $(2\pi\sigma)^{-\frac{d}{2}} e^{-|x-y|^2/(2\sigma)}$. For each $\sigma > 0$, P_σ maps L^p to $C_b(\mathbb{R}^d)$ and for every $\phi \in L^p(\mathbb{R}^d)$, we have

$$\|P_\sigma \phi\|_\infty \lesssim \sigma^{-\frac{d}{2p}} \|\phi\|_{L^p}.$$

This bound follows from the Hölder inequality $\|p_\sigma * \phi\|_\infty \leq \|p_\sigma\|_{L^{p'}} \|\phi\|_{L^p}$, where p_σ is the Gaussian density of P_σ and $\frac{1}{p'} + \frac{1}{p} = 1$. In conjunction with (6.6) and Hölder inequality, we have

$$\begin{aligned} \mathbb{E}^{\mathcal{F}_s} \int_s^T h(t, B_t^H) dt &\lesssim \int_s^T |t-s|^{-H\frac{d}{p}} \|h(t, \cdot)\|_{L^p(\mathbb{R}^d)} dt \\ &\lesssim \|h\|_{L^q L^p} \left(\int_s^T |t-s|^{-q'H\frac{d}{p}} dt \right)^{\frac{1}{q'}}. \end{aligned} \tag{6.10}$$

This yields (6.8). Then, by Taylor’s expansion, we have

$$\mathbb{E} \exp \left(\int_0^T h(t, B_t^H) dt \right) = 1 + \sum_{n=1}^\infty I_n,$$

where

$$I_n = \mathbb{E} \int_0^T \int_{t_1}^T \cdots \int_{t_{n-1}}^T h(t_1, B_{t_1}^H) \cdots h(t_n, B_{t_n}^H) dt_n \cdots dt_1.$$

By conditioning successively on $\mathcal{F}_{t_{n-1}}, \dots, \mathcal{F}_{t_1}$ and using (6.10) we obtain

$$\begin{aligned} I_n &\leq C^n \|h\|_{L^q L^p}^n \\ &\quad \times \left(\int_0^T \int_{t_1}^T \cdots \int_{t_{n-1}}^T (t_n - t_{n-1})^{-q'H\frac{d}{p}} \cdots (t_2 - t_1)^{-q'H\frac{d}{p}} t_1^{-q'H\frac{d}{p}} dt_n \cdots dt_1 \right)^{\frac{1}{q'}}. \end{aligned}$$

Integrating successively starting from dt_n to dt_1 , using the identity

$$\int_s^T (T-t)^{x-1} (t-s)^{y-1} dt = (T-s)^{x+y-1} \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \quad \text{for every } x, y > 0,$$

it is straightforward to obtain

$$I_n \leq \frac{C^n \|h\|_{L^q L^p}^n T^{n(1-H\frac{d}{p}-\frac{1}{q})}}{\Gamma(n(1-H\frac{d}{p}-\frac{1}{q})+1)}$$

for some constant $C > 0$ depending on p, q, H . Thus we have

$$\mathbb{E} \exp \left(\int_0^T h(t, B_t^H) dt \right) \leq \sum_{n=0}^\infty \frac{C^n \|h\|_{L^q L^p}^n T^{n(1-H\frac{d}{p}-\frac{1}{q})}}{\Gamma(n(1-H\frac{d}{p}-\frac{1}{q})+1)},$$

which implies finiteness of the exponential moment. □

In the above proof, using the identity $\mathbb{E} \left(\int_0^T h(t, B_t^H) dt \right)^n = n! I_n$, we obtain the following estimate

$$\mathbb{E} \left(\int_0^T h(t, B_t^H) dt \right)^n \leq \frac{n! C^n \|h\|_{L^q L^p}^n T^{n(1-H\frac{d}{p}-\frac{1}{q})}}{\Gamma(n(1-H\frac{d}{p}-\frac{1}{q})+1)} \tag{6.11}$$

for every positive integer n .

Proposition 6.5 (Girsanov transformation). *Let B^H be an $\{\mathcal{F}_t\}$ -fractional Brownian motion in \mathbb{R}^d with Hurst parameter $H \in (0, 1/2)$. Consider the shifted process*

$$\tilde{B}_t^H = B_t^H + \int_0^t h_s ds$$

where $(h_s)_{s \in [0, T]}$ is an $\{\mathcal{F}_t\}$ -adapted process in \mathbb{R}^d with integrable trajectories. We define

$$v_s = \frac{1}{\Gamma(\frac{1}{2} - H)} s^{H - \frac{1}{2}} \int_0^s (s - r)^{-\frac{1}{2} - H} r^{\frac{1}{2} - H} h_r dr.$$

Assume that

(i) v belongs to $L^2([0, T]; \mathbb{R}^d)$, almost surely,

(ii) $\mathbb{E}(\xi_T) = 1$ where

$$\xi_T = \exp \left\{ - \int_0^T v_s \cdot dW_s - \frac{1}{2} \int_0^T |v_s|^2 ds \right\}.$$

Then the shifted process \tilde{B}^H is an $\{\mathcal{F}_t^{B^H}\}$ -fractional Brownian motion with Hurst parameter H under the new probability $\tilde{\mathbb{P}}$ defined by $d\tilde{\mathbb{P}} = \xi_T d\mathbb{P}$.

Proof. This result is derived from the Girsanov transformation of multidimensional Brownian motion W via the representation (6.5). The specific form of v is chosen so that $\int_0^t K_H(t, s) v_s ds = \int_0^t h_r dr$ (see [37, Eq. (13)]). The proof follows the same arguments as [37, Theorem 2] with obvious modifications to multidimensional setting. \square

Proof of Theorem 6.1. This result can be obtained following the approach of [38]. The only difference here is that we work on arbitrary finite dimension.

In [38], a weak solution solution is constructed using Girsanov transformation, which relies on the validity of the following inequality (Novikov’s condition)

$$\mathbb{E} \exp \left(\theta \int_0^T |v_s|^2 ds \right) < \infty \tag{6.12}$$

where θ is any positive number and

$$v_s = \frac{1}{\Gamma(\frac{1}{2} - H)} s^{H - \frac{1}{2}} \int_0^s (s - r)^{-\frac{1}{2} - H} r^{\frac{1}{2} - H} b(r, B_r^H) dr.$$

See [38, Lemma 3.3 and equation (3.11)]. These formulas are unchanged in multi-dimensions (see Proposition 6.5). As in [37, 38], using the fact that $H < \frac{1}{2}$ and Young’s inequality, we have

$$\int_0^T |v_s|^2 ds \lesssim \int_0^T \left| \int_0^s (s - r)^{-\frac{1}{2} - H} b(r, B_r^H) dr \right|^2 ds \lesssim \int_0^T |b(s, B_s^H)|^2 ds. \tag{6.13}$$

Hence, to obtain (6.12), it suffices to show that

$$\mathbb{E} \exp \left(\theta \int_0^T |b(r, B_r^H)|^2 dr \right) < \infty \tag{6.14}$$

for any $\theta > 0$. Observe that the function $h = |b|^2$ belongs to $L^{q/2}([0, T]; L^{p/2}(\mathbb{R}^d))$. The condition (6.4) ensures that

$$H \frac{d}{p/2} + \frac{1}{q/2} < 1.$$

Hence, we can apply Lemma 6.4 to obtain (6.14). From here, the arguments used in the proofs of [38, Theorems 3.3 and 3.4] are applicable, which yield existence and uniqueness in law of weak solutions to (6.2). \square

Remark 6.6. In (6.13), one could use Hardy-Littlewood inequality (see [40, Theorem 1, p. 119]) to obtain

$$\int_0^T |v_s|^2 ds \lesssim \left(\int_0^T |b(s, B_s^H)|^{\frac{1}{1-H}} ds \right)^{2(1-H)},$$

which is an improvement over (6.13). If this inequality is used instead of (6.13) in the previous proof, the integrability condition (6.14) should be replaced by

$$\mathbb{E} \exp \left\{ \theta \left(\int_0^T |b(r, B_r^H)|^{\frac{1}{1-H}} dr \right)^{2(1-H)} \right\} < \infty.$$

However, verifying this condition using Taylor expansion and the moment estimate (6.11) also leads to condition (6.4).

For later purposes, we state the following result which is analogous to [38, Lemma 3.3].

Lemma 6.7. Let $(B^H, X, \{\mathcal{F}_t\}_{t \in [0, T]})$ be a pathwise solution of (6.2) defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Set

$$v_s = \frac{1}{\Gamma(\frac{1}{2} - H)} s^{H-\frac{1}{2}} \int_0^s (s-r)^{-\frac{1}{2}-H} r^{\frac{1}{2}-H} b(r, X_r) dr. \tag{6.15}$$

Assume that (6.4) holds. Then v belongs to $L^2([0, T]; \mathbb{R}^d)$ almost surely and

$$\xi_T = \exp \left(- \int_0^T v_s \cdot dW_s - \frac{1}{2} \int_0^T |v_s|^2 ds \right) \tag{6.16}$$

defines a random variable such that the measure $\tilde{\mathbb{P}}$ given by $d\tilde{\mathbb{P}} = \xi_T d\mathbb{P}$ is a probability measure equivalent to \mathbb{P} . Moreover, for every $\theta \in \mathbb{R}$, there exists a constant K_θ depending on θ, T, K, H, p, q and on $\|b\|_{L^q L^p}$ such that

$$\mathbb{E}^{\tilde{\mathbb{P}}} \xi_T^\theta + \mathbb{E}^{\mathbb{P}} \xi_T^\theta \leq K_\theta.$$

Proof. In [38, Lemma 3.3], this result relies on Girsanov transformation and the validity of inequality (6.9). Hence it can be carried over in a multi-dimensional setting. \square

Proposition 6.8. Let $(B^H, X, \{\mathcal{F}_t\}_{t \in [0, T]})$ be a pathwise solution of (6.2) defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and (p_1, q_1) be in $[1, \infty]^2$ such that $H \frac{d}{p_1} + \frac{1}{q_1} < 1$. Assume that (6.4) holds. Then for every Borel function f in $L^{q_1}([0, T]; L^{p_1}(\mathbb{R}^d))$, every integer $m \geq 2$ and $s, t \in [0, T]$,

$$\left\| \int_s^t f_r(X_r) dr \right\|_{L^m} \leq C(T, m, p, q, \|b\|_{L^q L^p}) \|f\|_{L^{q_1} L^{p_1}} |t - s|^{1-H \frac{d}{p_1} - \frac{1}{q_1}}.$$

Proof. Define v and ξ_T by (6.15) and (6.16), respectively. By Lemma 6.7, the process v satisfies conditions (i) and (ii) of Proposition 6.5. Hence, under the measure $\tilde{\mathbb{P}}$ given by $d\tilde{\mathbb{P}} = \xi_T d\mathbb{P}$, the process X is a fractional Brownian motion with Hurst parameter H . By Cauchy-Schwarz inequality, we have

$$\begin{aligned} \mathbb{E}^{\mathbb{P}} \left| \int_s^t f_r(X_r) dr \right|^m &= \mathbb{E}^{\tilde{\mathbb{P}}} \left| \int_s^t f_r(X_r) dr \right|^m \xi_T^{-1} \\ &\leq \left(\mathbb{E}^{\tilde{\mathbb{P}}} \xi_T^{-2} \right)^{\frac{1}{2}} \left(\mathbb{E}^{\tilde{\mathbb{P}}} \left| \int_s^t f_r(X_r) dr \right|^{2m} \right)^{\frac{1}{2}}. \end{aligned}$$

We now apply Lemma 6.7 and (6.11) to obtain the result. \square

As an application, we derive a regularity property for pathwise solutions.

Proposition 6.9. *Let $(B^H, X, \{\mathcal{F}_t\}_{t \in [0, T]})$ be a pathwise solution of (6.2) defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Put $\psi = X - B^H$ and assume that (6.4) holds. Then for every $m \geq 2$ and $s, t \in [0, T]$,*

$$\|\psi_t - \psi_s\|_{L_m} \leq C(T, m, p, q, \|b\|_{L^q L^p}) |t - s|^{1 - H \frac{d}{p} - \frac{1}{q}}.$$

Proof. Note that $\psi_t - \psi_s = \int_s^t b_r(X_r) dr$. The result is a direct application of Proposition 6.8. □

Pathwise uniqueness and strong existence

We denote

$$\tau_H(p, q) = 1 - H \frac{d}{p} - \frac{1}{q}. \tag{6.17}$$

It is useful to observe that the conditions (6.4) and (6.3) are respectively equivalent to

$$\tau_H(p, q) > \frac{1}{2} \quad \text{and} \quad \tau_H(p, q) > H + \frac{1}{2}. \tag{6.18}$$

We recall that $\{P_\sigma\}_{\sigma \geq 0}$ is the heat semigroup and σ_H is defined in (6.7). The following lemma will be needed.

Lemma 6.10. *For every $f \in L^q([0, T]; L^p(\mathbb{R}^d))$ and every $s < t$, we have*

$$\int_s^t \|\nabla P_{\sigma_H^2(s,r)} f_r\|_\infty dr \lesssim \|f\|_{L^q L^p} |t - s|^{\tau_H(p,q) - H}, \quad \text{if } \tau_H(p, q) > H, \tag{6.19}$$

$$\int_s^t \|\nabla^2 P_{\sigma_H^2(s,r)} f_r\|_\infty dr \lesssim \|f\|_{L^q L^p} |t - s|^{\tau_H(p,q) - 2H}, \quad \text{if } \tau_H(p, q) > 2H. \tag{6.20}$$

Proof. We recall that for every $\sigma > 0$, P_σ maps $L^p(\mathbb{R}^d)$ to $C_b^2(\mathbb{R}^d)$. In addition, for every $\phi \in L^p(\mathbb{R}^d)$, applying Hölder inequality, we have

$$\|\nabla P_\sigma \phi\|_\infty \lesssim \sigma^{-\frac{d}{2p} - \frac{1}{2}} \|\phi\|_{L^p} \quad \text{and} \quad \|\nabla^2 P_\sigma \phi\|_\infty \lesssim \sigma^{-\frac{d}{2p} - 1} \|\phi\|_{L^p}. \tag{6.21}$$

Applying the former inequality in (6.21) and Hölder inequality yields

$$\begin{aligned} \int_s^t \|\nabla P_{\sigma_H^2(s,r)} f_r\|_\infty dr &\lesssim \int_s^t |\sigma_H(s, r)|^{-\frac{d}{p} - 1} \|f_r\|_{L^p} dr \\ &\lesssim \|f\|_{L^q L^p} \left(\int_s^t |\sigma_H(s, r)|^{-q' \frac{d}{p} - q'} dr \right)^{\frac{1}{q'}}, \end{aligned}$$

where $q' = \frac{q}{q-1}$, the Hölder conjugate of q . Using (6.6), it is evident that

$$\left(\int_s^t |\sigma_H(s, r)|^{-q' \frac{d}{p} - q'} dr \right)^{\frac{1}{q'}} \lesssim \left(\int_s^t |r - s|^{-H q' \frac{d}{p} - H q'} dr \right)^{\frac{1}{q'}} \lesssim |t - s|^{1 - H \frac{d}{p} - \frac{1}{q} - H}.$$

The above two inequalities imply (6.19). The estimate (6.20) is obtained analogously by applying Hölder inequality and the later inequality in (6.21). □

Remark 6.11. In the previous proof, we have used approximations of $L^{q_1} L^{p_1}$ -functions by smooth functions, which requires that q_1, p_1 are finite.

Let us now fix a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \in [0, T]})$ such that \mathcal{F}_0 contains \mathbb{P} -null sets. On this probability space, B^H is an $\{\mathcal{F}_t\}$ -fractional Brownian motion. The following result is a kind of division property.

Lemma 6.12. *Let $(B^H, X, \{\mathcal{F}_t\})$ and $(B^H, \bar{X}, \{\mathcal{F}_t\})$ be two pathwise solutions to (6.2). Let f be in $L^{q_1}([0, T]; L^{p_1}(\mathbb{R}^d))$, $q_1, p_1 \in [1, \infty)$. Assume that (6.4) holds, $\tau_H(p_1, q_1) > H + \frac{1}{2}$ and $\tau_H(p, q) + \tau_H(p_1, q_1) > 1 + 2H$. Putting $\psi = X - B^H$ and $\bar{\psi} = \bar{X} - B^H$, there exist modifications of $\psi, \bar{\psi}$ and a Hölder continuous process $V = V(f, \psi, \bar{\psi})$ such that with probability one,*

$$\int_0^t f_r(X_r)dr - \int_0^t f_r(\bar{X}_r)dr = \int_0^t (\psi_r - \bar{\psi}_r) \cdot dV_r, \quad \forall t \in [0, T]. \tag{6.22}$$

In the above, $\int_0^t (\psi_r - \bar{\psi}_r) \cdot dV_r$ is a (well-defined) Young integral.

Proof. Herein, we put $\tau = \tau_H(p, q)$ and $\tau_1 = \tau_H(p_1, q_1)$. Let $\{f^n\}$ be a sequence of functions in $C_b^2([0, T]; \mathbb{R}^d)$ convergent to f in $L^{q_1}([0, T]; L^{p_1}(\mathbb{R}^d))$. Then, it is evident that

$$\int_0^t f_r^n(X_r)dr - \int_0^t f_r^n(\bar{X}_r)dr = \int_0^t (\psi_r - \bar{\psi}_r) \cdot dV_r[f^n], \quad \forall t \in [0, T], \tag{6.23}$$

where

$$V_t[f^n] = \int_0^t \int_0^1 \nabla f_r^n(B_r^H + \theta\psi_r + (1 - \theta)\bar{\psi}_r) d\theta dr.$$

Our strategy is to pass through the limit $n \rightarrow \infty$ in (6.23). From Proposition 6.8 and Garsia-Rodemich-Rumsey inequality, we see that left-hand side of (6.23) converges in probability uniformly on $[0, T]$ to

$$\int_0^t f_r(B_r^H + \psi_r)dr - \int_0^t f_r(B_r^H + \bar{\psi}_r)dr.$$

From Proposition 6.9, choosing m sufficiently large and applying Kolmogorov continuity theorem, we see that ψ and $\bar{\psi}$ have continuous modifications in $C^{\frac{1}{2}+\varepsilon}([0, T]; \mathbb{R}^d)$ for some $\varepsilon > 0$. Hence, by continuity of Young integrations, to pass through the limit $n \rightarrow \infty$ in the right-hand side of (6.23), it suffices to show that $V[f^n]$ converges to a process $V[f]$ in $C^{\frac{1}{2}+\varepsilon}([0, T])$ in probability for some $\varepsilon > 0$. This is accomplished below via the stochastic sewing lemma, which is similar to Proposition 3.5 and Corollary 3.9.

We first construct an auxiliary process. For each $g \in L^{q_1}([0, T]; L^{p_1}(\mathbb{R}^d))$, consider

$$A_{s,t}[g] = \int_s^t \int_0^1 \nabla(P_{\sigma_H^2(s,r)}g_r)(\mathbb{E}^{\mathcal{F}_s} B_r^H + \theta\psi_s + (1 - \theta)\bar{\psi}_s) d\theta dr.$$

Let $m \geq 2$ be fixed. From (6.19), we have

$$\|A_{s,t}[g]\|_{L_m} \lesssim \|g\|_{L^{q_1} L^{p_1}} |t - s|^{\tau_1 - H}.$$

To simplify the notation, we denote $z_s^\theta = \theta\psi_s + (1 - \theta)\bar{\psi}_s$. Then

$$\begin{aligned} & \mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t}[g] \\ &= \mathbb{E}^{\mathcal{F}_s} \int_u^t \int_0^1 \left(\nabla(P_{\sigma_H^2(u,r)}g_r)(\mathbb{E}^{\mathcal{F}_u} B_r^H + z_s^\theta) - \nabla(P_{\sigma_H^2(u,r)}g_r)(\mathbb{E}^{\mathcal{F}_u} B_r^H + z_u^\theta) \right) d\theta dr. \end{aligned}$$

Applying (6.20), we have

$$\|\mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t}[g]\|_{L_m} \lesssim \|g\|_{L^{q_1} L^{p_1}} (\|\psi\|_{C^\tau L_m} + \|\bar{\psi}\|_{C^\tau L_m}) |t - s|^{\tau_1 + \tau - 2H}.$$

(Recall that $\|\cdot\|_{C^\tau L_m}$ is defined at the beginning of Section 3.) From our assumptions, $\tau_1 - H > \frac{1}{2}$ and $\tau_1 + \tau - 2H > 1$, so by Theorem 2.3, there exists an adapted process in $C^{\tau_H(p_1, q_1) - H} L_m$, denoted by $\mathcal{A}[g]$, such that

$$\|\mathcal{A}_t[g] - \mathcal{A}_s[g]\|_{L_m} \lesssim \|g\|_{L^{q_1} L^{p_1}} (1 + \|\psi\|_{C^\tau L_m} + \|\bar{\psi}\|_{C^\tau L_m}) |t - s|^{\tau_1 - H}, \tag{6.24}$$

for every $s \leq t$ in $[0, T]$. By choosing m sufficiently large and applying Kolmogorov continuity theorem, we see that $\mathcal{A}[g]$ has continuous modification in $C^\alpha([0, T])$ for every $\alpha < \tau_1 - H$.

We claim that $V[f^n] = \mathcal{A}[f^n]$. Indeed, since ∇f^n is bounded and Lipschitz, applying Proposition 6.9, it is straightforward to verify that for every $s \leq t$,

$$\|V_t[f^n] - V_s[f^n]\|_{L_m} \leq \|\nabla f^n\|_{C_b([0, T] \times \mathbb{R}^d)} |t - s|$$

and

$$\begin{aligned} & \|\mathbb{E}^{\mathcal{F}_s}(V_t[f^n] - V_s[f^n]) - A_{s,t}[f^n]\|_{L_m} \\ &= \|\mathbb{E}^{\mathcal{F}_s} \int_s^t \int_0^1 [\nabla f_r^n(B_r^H + z_r^\theta) - \nabla f_r^n(B_r^H + z_s^\theta)] d\theta dr\|_{L_m} \\ &\leq \|f^n\|_{C_b^2} (\|\psi\|_{C^\tau L_m} + \|\bar{\psi}\|_{C^\tau L_m}) |t - s|^{1+\tau}. \end{aligned}$$

Hence, by uniqueness of Theorem 2.3, we must have $V_t[f^n] = \mathcal{A}_t[f^n]$ a.s. for every $t \in [0, T]$.

We now define $V[f]$ as the continuous modification of $\mathcal{A}[f]$. Using the estimate (6.24), we see that $\lim_n V[f^n] = V[f]$ in $C_T^{\tau_1 - H} L_m$. By choosing m sufficiently large and applying Garsia-Rodemich-Rumsey inequality, it follows that the sequence $\{V[f^n]\}$ converges to $V[f]$ in $C^{\frac{1}{2} + \varepsilon}([0, T]; \mathbb{R}^d)$ in probability for some $\varepsilon > 0$. This completes the proof. \square

We now have enough material to show pathwise uniqueness for (6.2).

Proof of Theorem 6.2. Suppose $(B^H, X, \{\mathcal{F}_t\}_{t \in [0, T]})$, $(B^H, \bar{X}, \{\bar{\mathcal{F}}_t\}_{t \in [0, T]})$ are two pathwise solutions to (6.2) defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $X_0 = \bar{X}_0 = x$. From Proposition 6.9, using Kolmogorov continuity theorem, we see that $\psi, \bar{\psi}$ are a.s. Hölder continuous with exponent $\frac{1}{2} + \varepsilon$ for some $\varepsilon > 0$. Then with probability one,

$$\psi_t - \bar{\psi}_t = \int_0^t b_r(B_r^H + \psi_r) dr - \int_0^t b_r(B_r^H + \bar{\psi}_r) dr \quad \forall t \in [0, T].$$

Since $[0, T]$ has finite length, we can assume without loss of generality that $q < \infty$. We consider two cases.

Case 1: $p < \infty$. From Lemma 6.12, (for a.s. ω) we can rewrite the above equation to

$$\psi_t^i - \bar{\psi}_t^i = \int_0^t (\psi_r - \bar{\psi}_r) \cdot dV_r^i, \quad \forall t \in [0, T], \forall i = 1, \dots, d, \tag{6.25}$$

where for each i , $\psi^i, \bar{\psi}^i$ are respectively the i -th components of $\psi, \bar{\psi}$, $V^i = V(b^i, \psi, \bar{\psi})$ is the process defined in Lemma 6.12. Note that each integral on the right-hand side (6.25) is a Young integral. Equation (6.25) is a Young differential equation, which has uniqueness ([26]). This implies that $\psi = \bar{\psi}$ and hence $X = \bar{X}$.

Case 2: $p = \infty$. For each n , we denote $\chi_n(x) = \mathbf{1}_{\{|x| \leq n\}}$. Choose $p_1 \in [1, \infty)$ such that $H \frac{d}{p_1} + \frac{1}{q} < \frac{1}{2} - H$. Let $\Omega^* \in \mathcal{F}$ be such that on Ω^* , $\psi, \bar{\psi}$ are Hölder continuous with exponent $\frac{1}{2} + \varepsilon$ for some $\varepsilon > 0$ and

$$\int_0^t [b\chi_n](r, X_r) dr - \int_0^t [b\chi_n](r, \bar{X}_r) dr = \int_0^t (\psi_r - \bar{\psi}_r) \cdot dV_r^{(n)} \quad \forall t \in [0, T], \forall n \geq 1.$$

Here, $V^{(n)} = (V^{(n),1}, \dots, V^{(n),d})$ and for each $j = 1, \dots, d$, $V^{(n),j} = V(b^j \chi_n, \psi, \bar{\psi})$ is the process constructed in Lemma 6.12. It is clear that we can find $\varepsilon > 0$ so that $\mathbb{P}(\Omega^*) = 1$.

Let us fix an arbitrary ω in Ω^* . We define

$$\sigma_n(\omega) = \inf\{t \in [0, T] : |\psi_t(\omega)| > n \text{ or } |\bar{\psi}_t(\omega)| > n\}.$$

For every $t \in [0, \sigma_n(\omega)]$, we have

$$\begin{aligned} \psi_t(\omega) - \bar{\psi}_t(\omega) &= \int_0^t [b\chi_n](r, X_r(\omega))dr - [b\chi_n](r, \bar{X}_r(\omega))dr \\ &= \int_0^t (\psi_r(\omega) - \bar{\psi}_r(\omega)) \cdot dV_r^{(n)}(\omega). \end{aligned}$$

As in the first case, uniqueness of Young differential equations implies that $\psi_t(\omega) = \bar{\psi}_t(\omega)$ for every $t \in [0, \sigma_n(\omega)]$. It is obvious that $\lim_n \sigma_n(\omega) = T$. Hence, we conclude that $\psi(\omega) = \bar{\psi}(\omega)$ on $[0, T]$, which implies $X = \bar{X}$. \square

Proof of Theorem 6.3. Uniqueness of strong solutions is a consequence of pathwise uniqueness obtained in Theorem 6.2. To show existence of a strong solution, we rely on Yamada-Watanabe’s result [43, Corollary 1]. We recall the identity (6.5) and observe that $(B^H, X, \{\mathcal{F}_t\})$ is a solution to (6.2) if and only if $(W, X, \{\mathcal{F}_t\})$ satisfies (i) W is an $\{\mathcal{F}_t\}$ -Wiener process, (ii) X is $\{\mathcal{F}_t\}$ -adapted and (iii) with probability one

$$X_t = x + \int_0^t b(r, X_r)dr + \int_0^t K_H(t, s)dW_s, \quad \forall t \in [0, T]. \tag{6.26}$$

The results of Theorems 6.1 and 6.3 are transferred in obvious ways to equation (6.26). The argument of [43] is then applicable to equation (6.26) as it relies only on properties of Wiener processes, weak existence and pathwise uniqueness. In particular, it does not depend on a particular form of the equation. This shows existence of a strong solution to (6.26). Since B^H and W generate the same filtration, a strong solution to (6.26) is also a strong solution to (6.2). This completes the proof. \square

When $H = 1/2$, a well-known result of Krylov and Röckner [30] states that (6.2) has unique strong solution if

$$\frac{d}{p} + \frac{2}{q} < 1 \quad \Leftrightarrow \quad \tau_{1/2}(p, q) > \frac{1}{2}. \tag{6.27}$$

Our condition (6.3) is therefore not optimal because it forces $p, q \rightarrow \infty$ when $H \rightarrow \frac{1}{2}^-$. The fraction $\frac{1}{2}$ which appears in (6.27) is the same as the one in (2.7) and seems to be fixed for any $H < \frac{1}{2}$ in order to apply Girsanov transformation. It is then natural to expect the following result.

Conjecture 6.13. *Equation (6.2) has a unique strong solution provided that*

$$\tau_H(p, q) > \frac{1}{2}, \tag{6.28}$$

which is indeed equivalent to (6.4).

The validity of this conjecture is out of reach at the moment of writing and is an interesting problem by its own.

7 Averaging along fractional Brownian motions

Let B^H be an $\{\mathcal{F}_t\}$ -fractional Brownian motion in \mathbb{R}^d with $H \in (0, 1)$ and f be an element in $L_T^q \mathcal{C}^\nu(\mathbb{R}^d)$ for some $\nu < 1$. We recall Definition 3.2 of $\mathcal{C}^\nu(\mathbb{R}^d)$ for $\gamma \leq 0$. For $\nu \in (0, 1)$, $\mathcal{C}^\nu(\mathbb{R}^d)$ is the same as the space of bounded Hölder continuous functions $C_b^\nu(\mathbb{R}^d)$. The space $L_T^q \mathcal{C}^\nu(\mathbb{R}^d)$ is defined at the beginning of Section 3 with $\mathcal{H} = \mathcal{C}^\nu(\mathbb{R}^d)$. In the current section, we investigate the almost sure joint-Hölder regularity of the following map

$$(t, x) \mapsto \int_0^t f(r, B_r^H + x)dr.$$

In particular, we seek for almost sure estimates of the type

$$|\int_s^t f(r, B_r^H + x)dr - \int_s^t f(r, B_r^H + y)dr| \lesssim |t - s|^\gamma |x - y|^\alpha \tag{7.1}$$

uniformly for s, t, x, y in compact sets, for some suitable exponents $\gamma, \alpha > 0$. This type of estimates plays a key role in the work [6] of Catellier and Gubinelli in obtaining path-by-path uniqueness for stochastic differential equations of the type

$$dX_t = b(t, X_t)dt + dB_t^H, \quad b \in L_T^\infty \mathcal{C}^\nu. \tag{7.2}$$

They observe that the remainder process $\psi = X - B^H$ satisfies a non-linear Young integral equation in which the driving space-time vector field is $(t, x) \mapsto \int_0^t b(r, B_r^H + x)dr$. The a.s. estimate (7.1) then provides necessary regularity for this space-time vector field so that the equation for ψ has a unique solution. Let us point out that the approach of [6] is an extension of the earlier work [7] of Davie. Non-linear Young integral equations are also considered in [24] although connections with equation (7.2) and Davie’s work were not observed there.

Well-posedness for non-linear Young differential equations are well-understood (see [6, Section 2] or [24, Section 3]). Hence we will not consider them here. We rather provide another way of obtaining estimates of the type (7.1), using stochastic sewing lemma (more specifically Theorem 2.3). Our approach allows for time-dependent vector fields, which are not considered in [6]. Comparing the two papers, most of estimates in [6] can be recovered from ours by simply setting the parameter q to be ∞ . In addition, our treatment provides explicit Hölder exponents in space and time simultaneously.

Let $\{\mathcal{F}_t\}_{t \geq 0}$ be the natural filtration generated by B^H . Let us consider for each $s < t$ and $f \in L_T^q \mathcal{C}^\nu(\mathbb{R}^d)$, the increment process

$$A_{s,t}^x[f] = \int_{[s,t]} [P_{\sigma_H^2(s,r)} f_r](\mathbb{E}^{\mathcal{F}_s} B_r^H + x)dr,$$

where $\{P_\sigma\}_{\sigma \geq 0}$ is the heat semigroup associated to the kernel $(2\pi\sigma)^{-\frac{d}{2}} e^{-|x-y|^2/(2\sigma)}$, σ_H is defined in (6.7). If f is a bounded Borel function, we can also write $A_{s,t}^x[f] = \int_{[s,t]} \mathbb{E}^{\mathcal{F}_s} f_r(B_r^H + x)dr$. We recall that for every $\sigma > 0$, P_σ maps $\mathcal{C}^\nu(\mathbb{R}^d)$ to $C_b^2(\mathbb{R}^d)$. In addition, for every $\alpha \in (0, 1]$ and for every $\phi \in \mathcal{C}^\nu(\mathbb{R}^d)$, we have (see e.g. [5, Prop. 2.4])

$$\|P_\sigma \phi\|_\infty \lesssim \sigma^{\frac{\nu}{2}} \|\phi\|_{\mathcal{C}^\nu} \quad \text{and} \quad \|P_\sigma \phi\|_{\mathcal{C}^\alpha} \lesssim \sigma^{\frac{\nu-\alpha}{2}} \|\phi\|_{\mathcal{C}^\nu}. \tag{7.3}$$

In analogy with (6.19) and (6.20), it is straightforward to obtain following estimates

$$\int_{[s,t]} \|P_{\sigma_H^2(s,r)} f_r\|_\infty dr \lesssim \|f\|_{L^q \mathcal{C}^\nu} |t - s|^{1+H\nu-\frac{1}{q}}, \tag{7.4}$$

$$\int_{[s,t]} \|P_{\sigma_H^2(s,r)} f_r\|_{\mathcal{C}^\alpha} dr \lesssim \|f\|_{L^q \mathcal{C}^\nu} |t - s|^{1+H(\nu-\alpha)-\frac{1}{q}}, \tag{7.5}$$

for every $f \in L_T^q \mathcal{C}^\nu(\mathbb{R}^d)$ and every $s < t$, provided that $1 + H(\nu - \alpha) - \frac{1}{q} > 0$. These estimates are summarized in the following result.

Lemma 7.1. *Let $s < u < t$, $f \in L_T^q \mathcal{C}^\nu(\mathbb{R}^d)$ and $x, y \in \mathbb{R}^d$ be fixed. Suppose that $\alpha \in (0, 1]$ and $1 + H(\nu - \alpha) - \frac{1}{q} > 0$, then*

$$|A_{s,t}^x[f]| \lesssim \|f\|_{L_T^q \mathcal{C}^\nu} |t - s|^{1+H\nu-\frac{1}{q}} \tag{7.6}$$

and

$$|A_{s,t}^x[f] - A_{s,t}^y[f]| \lesssim \|f\|_{L_T^q \mathcal{C}^\nu} |x - y|^\alpha |t - s|^{1+H(\nu-\alpha)-\frac{1}{q}}. \tag{7.7}$$

We define

$$\gamma_H(\nu, q) = 1 + H\nu - \frac{1}{q}. \tag{7.8}$$

The following result is an analogue of Proposition 3.5.

Proposition 7.2. *Let $m \geq 2$ and $x \in \mathbb{R}^d$ be fixed. Suppose that*

$$\gamma_H(\nu, q) > \frac{1}{2} \Leftrightarrow \nu > \frac{1}{H} \left(\frac{1}{q} - \frac{1}{2} \right). \tag{7.9}$$

There exists a linear map \mathcal{A}^x from $L_T^q \mathcal{C}^\nu(\mathbb{R}^d)$ to $C_T^{1+H\nu-\frac{1}{q}} L_m$ such that

(a) For every $f \in L_T^q \mathcal{C}^\nu(\mathbb{R}^d)$ and every $0 \leq s \leq t \leq T$,

$$\|\mathcal{A}_t^x[f] - \mathcal{A}_s^x[f]\|_{L_m} \lesssim \|f\|_{L_T^q \mathcal{C}^\nu} |t - s|^{1+H\nu-\frac{1}{q}}. \tag{7.10}$$

(b) For every $f \in L_T^q \mathcal{C}^\nu(\mathbb{R}^d)$ and every $t \geq 0$, $\mathcal{A}_t[f]$ is the limit in L_m of the Riemann sum

$$\sum_i \int_{[t_i, t_{i+1}]} P_{\sigma_H^2(t_i, r)} f_r(\mathbb{E}^{\mathcal{F}_{t_i}} B_r^H + x) dr$$

as $\max_i |t_{i+1} - t_i| \rightarrow 0$. Here $\{t_i\}_i$ is any partition of $[0, t]$.

(c) If f is a bounded measurable function, then $\mathcal{A}_t^x[f] = \int_{[0, t]} f_r(B_r^H + x) dr$ a.s. for every $t \geq 0$.

Proof. Note that $\mathbb{E}^{\mathcal{F}_s} \delta A_{s,u,t}^x[f] = 0$ for every triplet $s < u < t$. Similar to Proposition 3.5, we can obtain (a) and (b) by using (7.6) and Theorem 2.3. Suppose that f is a bounded measurable function. Then we have $A_{s,t}^x[f] = \int_{[s,t]} \mathbb{E}^{\mathcal{F}_s} f_r(B_r^H + x) dr$ and

$$\left\| \int_{[s,t]} f_r(X_r) dr - A_{s,t}^x[f] \right\|_{L_m} \leq 2 \sup_{r,x} |f(r, x)| |t - s|.$$

Similar to Proposition 3.7, we obtain (c) by uniqueness of Theorem 2.3. □

Hereafter, we fix an element f in $L_T^q \mathcal{C}^\nu(\mathbb{R}^d)$ and write $\int_0^t f_r(B_r^H + x) dr$ for $\mathcal{A}_t^x[f]$ whenever the hypotheses of Proposition 7.2 are met.

Proposition 7.3. *For every $\alpha \in (0, 1]$ such that*

$$\gamma_H(\nu, q) > \frac{1}{2} + H\alpha \Leftrightarrow \alpha < \nu - \frac{1}{H} \left(\frac{1}{q} - \frac{1}{2} \right), \tag{7.11}$$

then

$$\left\| \int_s^t f_r(B_r^H + x) dr - \int_s^t f_r(B_r^H + y) dr \right\|_{L_m} \lesssim \|f\|_{L_T^q \mathcal{C}^\nu} |t - s|^{\gamma_H(\nu, q) - H\alpha} |x - y|^\alpha. \tag{7.12}$$

Proof. This is a direct consequence of (7.7) and Theorem 2.3. □

By choosing m sufficiently large and applying the Kolmogorov continuity theorem, there exists a continuous modification of the random field $\{\int_0^t f_r(B_r^H + x) dr\}_{t,x}$. We then apply the multiparameter Garsia-Rodemich-Rumsey inequality in [23] (choosing m sufficiently large) to see that there are arbitrarily small positive constants $\varepsilon_1, \varepsilon_2$ such that with probability one,

$$\left| \int_s^t f_r(B_r^H + x) dr - \int_s^t f_r(B_r^H + y) dr \right| \lesssim \|f\|_{L_T^q \mathcal{C}^\nu} |t - s|^{\gamma_H(\nu, q) - H\alpha - \varepsilon_1} |x - y|^{\alpha - \varepsilon_2} \tag{7.13}$$

for uniformly in $s, t \in [0, T]$ and x, y in a compact set. This procedure provides a method to obtain almost sure estimates of the type (7.1). It can be applied to most of the moment estimates stated below.

Proposition 7.4. *Suppose that $q \in [1, \infty)$ and*

$$\gamma_H(\nu, q) > \frac{1}{2} + H \Leftrightarrow \nu > 1 + \frac{1}{H} \left(\frac{1}{q} - \frac{1}{2} \right), \quad (7.14)$$

then the map $x \rightarrow \int_0^t f_r(B_r^H + x)$ is differentiable and for every fixed t, x

$$\nabla \int_0^t f_r(B_r^H + x) dr = \int_0^t \nabla f_r(B_r^H + x) dr \quad \text{a.s.} \quad (7.15)$$

In addition, the map $(t, x) \rightarrow \nabla \int_0^t f_r(B_r^H + x)$ has a continuous version. For every $\alpha \in (0, 1]$ such that

$$\gamma_H(\nu, q) > \frac{1}{2} + H + H\alpha \Leftrightarrow \alpha < \nu - 1 - \frac{1}{H} \left(\frac{1}{q} - \frac{1}{2} \right), \quad (7.16)$$

we have the following estimates for every $s < t$ and every $x, y \in \mathbb{R}^d$

$$\|\nabla \int_s^t f_r(B_r^H + x)\|_{L_m} \lesssim \|f\|_{L_T^q \mathcal{C}^\nu} |t - s|^{\gamma_H(\nu, q) - H} \quad (7.17)$$

and

$$\|\nabla \int_s^t f_r(B_r^H + x) dr - \nabla \int_s^t f_r(B_r^H + y) dr\|_{L_m} \lesssim \|f\|_{L_T^q \mathcal{C}^\nu} |t - s|^{\gamma_H(\nu, q) - H - H\alpha} |x - y|^\alpha. \quad (7.18)$$

Proof. We observe that ∇f belongs to $L_T^q \mathcal{C}^{\nu-1}(\mathbb{R}^d)$ with $\|\nabla f\|_{L_T^q \mathcal{C}^{\nu-1}} \leq \|f\|_{L_T^q \mathcal{C}^\nu}$ and that condition (7.14) is equivalent to $\gamma_H(\nu - 1, q) > \frac{1}{2}$. By Proposition 7.2, the random field $\{\int_0^t \nabla f_r(B_r^H + x) dr\}_{t,x}$ can be defined and has a continuous modification. Since $q < \infty$, we can choose a sequence $\{f^n\}$ in \mathcal{E}_T^2 which converges to f in $L_T^q \mathcal{C}^\nu$ ([27, Lemma 1.2.19]). Then from Proposition 7.2, we have for each $t \in [0, T]$,

$$\int_0^t \nabla f_r(B_r^H + x) dr = \lim_n \int_{[0,t]} \nabla f_r^n(B_r^H + x) dr \quad \text{in } L_m$$

and for each n

$$\int_{[0,t]} \nabla f_r^n(B_r^H + x) dr = \nabla \int_{[0,t]} f_r^n(B_r^H + x) dr.$$

The later identity is due to the fact that each f_n is a finite linear combination of $C_b^2(\mathbb{R}^d)$ -functions. This shows that $x \rightarrow \int_0^t f_r(B_r^H + x) dr$ is differentiable and (7.15) holds. The estimates (7.17) and (7.18) follow from identity (7.15) and estimates (7.10), (7.12) with f replaced by ∇f . \square

References

- [1] D. Baños, T. Nilssen, and F. Proske, *Strong existence and higher order Fréchet differentiability of stochastic flows of fractional Brownian motion driven SDE's with singular drift*, J. Dyn. Diff. Equat. (2019), 1–48.
- [2] R. F. Bass and Z.-Q. Chen, *Brownian motion with singular drift*, Ann. Probab. **31** (2003), no. 2, 791–817. MR-1964949
- [3] D. L. Burkholder, B. J. Davis, and R. F. Gundy, *Integral inequalities for convex functions of operators on martingales*, (1972), 223–240. MR-0400380
- [4] O. Butkovsky, K. Dareiotis, and M. Gerencsér, *Approximation of sdes – a stochastic sewing approach*, 2019.

- [5] G. Cannizzaro and K. Chouk, *Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential*, Ann. Probab. **46** (2018), no. 3, 1710–1763. MR-3785598
- [6] R. Catellier and M. Gubinelli, *Averaging along irregular curves and regularisation of ODEs*, Stochastic Process. Appl. **126** (2016), no. 8, 2323–2366. MR-3505229
- [7] A. M. Davie, *Uniqueness of solutions of stochastic differential equations*, Int. Math. Res. Not. IMRN (2007), no. 24, Art. ID rnm124, 26. MR-2377011
- [8] F. Delarue and R. Diel, *Rough paths and 1d SDE with a time dependent distributional drift: application to polymers*, Probab. Theory Related Fields **165** (2016), no. 1-2, 1–63. MR-3500267
- [9] J. L. Doob, *Stochastic processes*, John Wiley & Sons, Inc., New York; Chapman & Hall, Limited, London, 1953. MR-0058896
- [10] S. N. Ethier and T. G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986, Characterization and convergence. MR-0838085
- [11] D. Feyel and A. de La Pradelle, *Curvilinear integrals along enriched paths*, Electron. J. Probab. **11** (2006), no. 34, 860–892. MR-2261056
- [12] F. Flandoli, M. Gubinelli, and E. Priola, *Well-posedness of the transport equation by stochastic perturbation*, Invent. Math. **180** (2010), no. 1, 1–53. MR-2593276
- [13] F. Flandoli, E. Issoglio, and F. Russo, *Multidimensional stochastic differential equations with distributional drift*, Trans. Amer. Math. Soc. **369** (2017), no. 3, 1665–1688. MR-3581216
- [14] F. Flandoli, F. Russo, and J. Wolf, *Some SDEs with distributional drift. I. General calculus*, Osaka J. Math. **40** (2003), no. 2, 493–542. MR-1988703
- [15] P. K. Friz and M. Hairer, *A course on rough paths*, Universitext, Springer, Cham, 2014, With an introduction to regularity structures. MR-3289027
- [16] P. K. Friz and N. B. Victoir, *Multidimensional stochastic processes as rough paths*, Cambridge Studies in Advanced Mathematics, vol. 120, Cambridge University Press, Cambridge, 2010, Theory and applications. MR-2604669
- [17] A. M. Garsia, E. Rodemich, and H. Rumsey, Jr., *A real variable lemma and the continuity of paths of some Gaussian processes*, Indiana Univ. Math. J. **20** (1970/1971), 565–578. MR-0267632
- [18] M. Gubinelli, *Controlling rough paths*, J. Funct. Anal. **216** (2004), no. 1, 86–140. MR-2091358
- [19] M. Hairer, *A theory of regularity structures*, Invent. Math. **198** (2014), no. 2, 269–504. MR-3274562
- [20] M. Hairer and C. Labbé, *A simple construction of the continuum parabolic Anderson model on \mathbf{R}^2* , Electron. Commun. Probab. **20** (2015), no. 43, 11. MR-3358965
- [21] M. Hairer and X.-M. Li, *Averaging dynamics driven by fractional brownian motion*, arXiv preprint arXiv:1902.11251 (2019).
- [22] Y. Hu, *Itô-Wiener chaos expansion with exact residual and correlation, variance inequalities*, J. Theoret. Probab. **10** (1997), no. 4, 835–848. MR-1481650
- [23] Y. Hu and K. Le, *A multiparameter Garsia-Rodemich-Rumsey inequality and some applications*, Stochastic Process. Appl. **123** (2013), no. 9, 3359–3377. MR-3071383
- [24] Y. Hu and K. Lê, *Nonlinear Young integrals and differential systems in Hölder media*, Trans. Amer. Math. Soc. **369** (2017), no. 3, 1935–2002. MR-3581224
- [25] Y. Hu, K. Lê, and L. Mytnik, *Stochastic differential equation for Brox diffusion*, Stochastic Process. Appl. **127** (2017), no. 7, 2281–2315. MR-3652414
- [26] Y. Hu and D. 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. MR-2397797
- [27] T. Hytönen, J. van Neerven, M. Veraar, and L. Weis, *Analysis in Banach spaces. Vol. I. Martingales and Littlewood-Paley theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, vol. 63, Springer, Cham, 2016. MR-3617205

- [28] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, second ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991. MR-1121940
- [29] 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. MR-2748616
- [30] N. V. Krylov and M. Röckner, *Strong solutions of stochastic equations with singular time dependent drift*, Probab. Theory Related Fields **131** (2005), no. 2, 154–196. MR-2117951
- [31] T. J. Lyons, *Differential equations driven by rough signals*, Rev. Mat. Iberoamericana **14** (1998), no. 2, 215–310. MR-1654527
- [32] T. J. Lyons, M. Caruana, and T. Lévy, *Differential equations driven by rough paths*, Lecture Notes in Mathematics, vol. 1908, Springer, Berlin, 2007, Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004. MR-2314753
- [33] B. B. Mandelbrot and J. W. Van Ness, *Fractional Brownian motions, fractional noises and applications*, SIAM Rev. **10** (1968), 422–437. MR-0242239
- [34] M. B. Marcus and J. Rosen, *Markov processes, Gaussian processes, and local times*, Cambridge Studies in Advanced Mathematics, vol. 100, Cambridge University Press, Cambridge, 2006. MR-2250510
- [35] Y. Meyer, *Wavelets and operators*, Cambridge Studies in Advanced Mathematics, vol. 37, Cambridge University Press, Cambridge, 1992, Translated from the 1990 French original by D. H. Salinger. MR-1228209
- [36] D. Nualart, *The Malliavin calculus and related topics*, second ed., Probability and its Applications (New York), Springer-Verlag, Berlin, 2006. MR-2200233
- [37] D. Nualart and Y. Ouknine, *Regularization of differential equations by fractional noise*, Stochastic Process. Appl. **102** (2002), no. 1, 103–116. MR-1934157
- [38] D. Nualart and Y. Ouknine, *Stochastic differential equations with additive fractional noise and locally unbounded drift*, Stochastic inequalities and applications, Progr. Probab., vol. 56, Birkhäuser, Basel, 2003, pp. 353–365. MR-2073441
- [39] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999. MR-1725357
- [40] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970. MR-0290095
- [41] A. J. Veretennikov, *On strong solutions and explicit formulas for solutions of stochastic integral equations*, Sbornik: Mathematics **39** (1981), no. 3, 387–403. MR-0568986
- [42] A. J. Veretennikov and N. V. Krylov, *On explicit formulas for solutions of stochastic equations*, Mathematics of the USSR-Sbornik **29** (1976), no. 2, 239. MR-0410921
- [43] T. Yamada and S. Watanabe, *On the uniqueness of solutions of stochastic differential equations*, J. Math. Kyoto Univ. **11** (1971), 155–167. MR-0278420
- [44] P. Yaskov, *Extensions of the sewing lemma with applications*, Stochastic Process. Appl. **128** (2018), no. 11, 3940–3965. MR-3860015
- [45] L. C. Young, *An inequality of the Hölder type, connected with Stieltjes integration*, Acta Math. **67** (1936), no. 1, 251–282. MR-1555421

Acknowledgments. The author thanks Leonid Mytnik, Oleg Butkovsky, Siva Athreya and Giuseppe Cannizzaro for interesting and motivating discussions. He also thanks the editor, an associated editor and several anonymous referees for numerous constructive suggestions which help improve the presentation of the article.

Electronic Journal of Probability

Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS¹)
- Easy interface (EJMS²)

Economical model of EJP-ECP

- Non profit, sponsored by IMS³, BS⁴, ProjectEuclid⁵
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund⁶ (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

¹LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

²EJMS: Electronic Journal Management System <http://www.vtex.lt/en/ejms.html>

³IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

⁴BS: Bernoulli Society <http://www.bernoulli-society.org/>

⁵Project Euclid: <https://projecteuclid.org/>

⁶IMS Open Access Fund: <http://www.imstat.org/publications/open.htm>