

Stochastic sewing in Banach spaces*

Khoa Lê^{†‡}

Abstract

A stochastic sewing lemma which is applicable for processes taking values in Banach spaces is introduced. Applications to additive functionals of fractional Brownian motion of distributional type are discussed.

Keywords: stochastic sewing; stochastic regularization; martingale type; local time.

MSC2020 subject classifications: Primary 60H99; 46N30, Secondary 60H50.

Submitted to EJP on January 26, 2022, final version accepted on February 7, 2023.

1 Introduction

The sewing lemma is an instrumental and versatile tool originated from Lyons’ theory of rough paths [23]. A specific case of the lemma can be traced back at least to the work [27] of Young on the Riemann–Stieltjes integrals. Lyons utilizes Young’s argument in [23] effectively to show unique extension of almost rough paths. In [13], Gubinelli gives a general statement (Proposition 1 therein) of what we call the sewing lemma, his proof uses some elements from geometric measure theory. The connection with abstract Riemann sums and the proof based on dyadic partitions, which we will adopt and develop upon herein, are due to Feyel and de La Pradelle in [8]. The broadly accepted name “sewing lemma” seemingly appears first in [8].

The sewing lemma is purely analytic and provides quantitative estimates on (rough) integrals. The stochastic sewing lemma introduced by the author in [19] is an extension of the sewing lemma which takes into account stochastic cancellations. (See also [9, Section 4.6] for a brief introduction on the stochastic sewing lemma.) Since its introduction, the stochastic sewing lemma has caught some attention and led to new interesting applications; ranging from regularization by noise problems [1, 16, 12, 15], numerical methods for stochastic differential equations [3, 7, 20], rough stochastic differential equations [10] to averaging principle with fractional dynamics [14, 21]. While the sewing

*The research is supported partly by the Alexander von Humboldt Foundation and the European Research Council under the European Union’s Horizon 2020 research and innovation program (grant agreement No. 683164). Supports from the Hausdorff Research Institute for Mathematics (HIM), where the research was initiated during the junior trimester program in 2019, are acknowledged.

[†]TU Berlin.

[‡]Current address: School of Mathematics, University of Leeds, U.K. E-mail: k.le@leeds.ac.uk

lemma is applicable for processes in any Banach spaces, its stochastic version from [19] is only applicable for stochastic processes in \mathbb{R}^d . The main purpose of this article is to extend the stochastic sewing lemma for stochastic processes taking values in any Banach spaces. Our result presented herein reconciles the two sewing lemmas while at the same time, opens up new possibilities for applications. We provide one example on the spatial Besov regularity of additive functionals of fractional Brownian motion of the type $\int_0^1 f_r(B_r + \cdot) dr$ where f is a time-dependent distribution on \mathbb{R}^d . Closely related results along this direction include Berman's condition ([11, Theorem 28.1]) and the work [16] in which spatial Sobolev regularity in $W_2^\alpha(\mathbb{R}^d)$ of local times of Gaussian processes are considered. We are able to deal with generic time dependent distributions on Besov scales with integrability indices in $(1, \infty)$.

Let us discuss on further detail. The sewing method concerns about the convergence of the abstract Riemann sums

$$\sum_{[u,v] \in \pi} A_{u,v} \tag{1.1}$$

where A is a map from the simplex $\Delta := \{(s, t) \in [0, T]^2 : s \leq t\}$ to a Banach space $(\mathcal{X}, |\cdot|_{\mathcal{X}})$ and π is a generic partition of $[0, T]$. We think of $A_{s,t}$ as a generalized increment over the time interval $[s, t]$. For general Banach spaces, the sewing lemma gives the following sufficient condition for the convergence of (1.1) in \mathcal{X} : there exist positive constants ε, C and a control w such that

$$|\delta A_{s,u,t}|_{\mathcal{X}} \leq Cw(s,t)^{1+\varepsilon} \quad \forall 0 \leq s \leq u \leq t \leq T, \tag{1.2}$$

where $\delta A_{s,u,t} = A_{s,t} - A_{s,u} - A_{u,t}$. Hereafter, a control w is a continuous function $w : \Delta \rightarrow [0, \infty)$ such that

$$w(s, u) + w(u, t) \leq w(s, t) \quad \forall s \leq u \leq t. \tag{1.3}$$

The sewing lemma also provides bounds on the limiting object. Let \mathcal{A}_t denote the limit of Riemann sums (1.1) among partitions π of $[0, t]$. Under (1.2), one has

$$|\delta \mathcal{A}_{s,t} - A_{s,t}|_{\mathcal{X}} \leq Cw(s,t)^{1+\varepsilon} \quad \forall (s, t) \in \Delta \tag{1.4}$$

for some constant C , where $\delta \mathcal{A}_{s,t} = \mathcal{A}_t - \mathcal{A}_s$.

When $\mathcal{X} = [L^m(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})]^d$ for some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ and integers $d \geq 1, m \geq 2$, the stochastic sewing lemma from [19] gives the following sufficient condition for the convergence of (1.1) in $[L^m(\Omega)]^d$: $A_{s,t}$ is \mathcal{F}_t -measurable and there are positive constants ε, C such that

$$\|\delta A_{s,u,t}\|_{L^m(\Omega)} \leq C|t-s|^{\frac{1}{2}+\varepsilon} \quad \text{and} \quad \|\mathbb{E}(\delta A_{s,u,t} | \mathcal{F}_s)\|_{L^m(\Omega)} \leq C|t-s|^{1+\varepsilon} \tag{1.5}$$

for every $0 \leq s \leq u \leq t \leq T$. The corresponding estimates on \mathcal{A} supplied by the stochastic sewing lemma are

$$\|\delta \mathcal{A}_{s,t} - A_{s,t}\|_{L^m(\Omega)} \leq C|t-s|^{\frac{1}{2}+\varepsilon} \quad \text{and} \quad \|\mathbb{E}(\delta \mathcal{A}_{s,t} - A_{s,t} | \mathcal{F}_s)\|_{L^m(\Omega)} \leq C|t-s|^{1+\varepsilon}. \tag{1.6}$$

When $w(s, t) = t - s$ and $\mathcal{X} = [L^m(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})]^d$, one can compare the two results: (1.5) requires less regularity on δA but instead impose adaptiveness and an additional regularity condition on the conditional quantity $\mathbb{E}(\delta A_{s,u,t} | \mathcal{F}_s)$. When A is deterministic, (1.5) evidently deduces to (1.2). On the other hand, the stochastic sewing lemma is restricted to the space $[L^m(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})]^d$ whereas the sewing lemma applies

for general Banach spaces. Additionally, a curious feature of the stochastic sewing lemma is the exponent $1/2$ in (1.5). These differences are due to the upper bound in the Burkholder–Davis–Gundy (BDG) inequality in \mathbb{R}^d used in [19].

The current article gives an extension of the stochastic sewing lemma when \mathcal{X} takes the form $L_m(V) := L^m((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P}); V)$ for some separable Banach space $(V, |\cdot|_V)$. The space $L^m((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P}); V)$ contains all V -valued random variables which have finite m -th moment. To state the result, however, we need to introduce another feature on the geometric properties of V .

As explained previously, the exponent $1/2$ in (1.5) is tied to the upper bound in the BDG inequalities in \mathbb{R}^d . This inequality estimates moments of a martingale by the moments of its quadratic variation. While the BDG inequality is easily extended to Hilbert spaces, such inequality is not available for general Banach spaces. Nevertheless, it turns out that many Banach spaces, (including Hilbert spaces, Lebesgue spaces, Besov spaces, Triebel–Lizorkin spaces) possess a certain variant of the BDG upper bound parametrized by a number $p \in [1, 2]$. Such property is called *martingale type*, a concept originated from the study of geometry of Banach spaces. We refer to Definition 2.1 for the precise definition and to Proposition 2.4 for various examples of Banach spaces with non-trivial martingale types $p \in (1, 2]$.

Theorem A. Let w be a control, $m \geq 2$ and $(V, |\cdot|_V)$ be a separable Banach space with martingale type p . Let $A : \Omega \times \Delta \rightarrow V$ be a measurable map such that $A_{s,t}$ is \mathcal{F}_t -measurable for every $(s, t) \in \Delta$. Suppose that there are positive constants ε, C such that

$$\|\mathbb{E}(\delta A_{s,u,t} | \mathcal{F}_s) |V\|_{L^m(\Omega)} \leq Cw(s, t)^{1+\varepsilon} \quad \text{and} \quad \|\delta A_{s,u,t} |V\|_{L^m(\Omega)} \leq Cw(s, t)^{\frac{1}{p}+\varepsilon} \quad (1.7)$$

for every $0 \leq s \leq t \leq T$. Then the Riemann sums (1.1) converge in $L^m(\Omega; V)$.

Theorem A reconciles the sewing lemma and its stochastic version from [19]. Indeed, it is evident that when A is deterministic, condition (1.7) deduces to (1.2) and hence, Theorem A deduces to the sewing lemma in its full generality. On the other hand, the space \mathbb{R}^d has martingale type $p = 2$ and hence Theorem A deduces to the stochastic sewing lemma from [19] when $V = \mathbb{R}^d$. The exponent $1/2$ in (1.5) is revealed in (1.7) as $1/p$, where p is the martingale type of V . Theorem A follows from a more general result, Theorem 3.1 herein.

As an application, we apply the stochastic sewing lemma in Banach spaces to study a class of additive functionals of fractional Brownian motion. To be more precise, let B be a fractional Brownian motion in \mathbb{R}^d and f be a time-dependent distribution in the space $L^\theta([0, T]; \mathcal{B}_{p,1}^\alpha(\mathbb{R}^d))$, $\alpha \in \mathbb{R}$, $\theta, p \in (1, \infty)$. Consider additive functionals of the type

$$(t, x) \mapsto I[f]_t(x) := \int_0^t f_r(B_r + x) dr. \quad (1.8)$$

Such additive functional is related to the local times of fractional Brownian motion and its regularity properties are linked to the regularizing effect of fraction Brownian motion ([4, 19]). The two main issues that need to be addressed are defining $I[f]$ and establishing its spatial regularity. Because f_r is only a distribution for a.e. r , the composition $f_r(B_r + x)$ is not a priori well-defined, so is the functional $f \mapsto I[f]$. To overcome this issue, we define $f \mapsto I[f]$ as the continuous extension on the space of smooth functions with respect to the topology generated by $L^\theta([0, T]; \mathcal{B}_{p,1}^\alpha(\mathbb{R}^d))$. This means that whenever defined, for every sequence of smooth functions $(f^n)_n$ which converges to f in $L^\theta([0, T]; \mathcal{B}_{p,1}^\alpha(\mathbb{R}^d))$, we have

$$\lim_n \int_0^t f_r^n(B_r + x) dr = I[f]_t(x) \quad (1.9)$$

in an appropriate sense. It turns out that the limit in (1.9) holds even if the spatial regularity is greater than α . Hence, (1.9) also quantifies the regularizing effect of B and establishes spatial regularity for $I[f]$, see Theorem 4.2 herein. Although the method can be extended to other stochastic processes, as one has seen from the aforementioned applications of the stochastic sewing lemma, we restrict herein to fractional Brownian motion. An advantage is that it keeps our presentation concise while at the same time exhibit the method in different scales by varying the Hurst parameter.

The structure of the paper is as follows. In Section 2, we briefly summarize necessary facts on vector valued stochastic processes as well as the concept of martingale type. The stochastic sewing lemma in Banach spaces is stated and proved in Section 3. Section 4 discusses an application to additive functional of fractional Brownian motion. Appendix A presents an extension of Kolmogorov continuity criterion with controls which can be used to construct a continuous modification of the process constructed from Theorem A. Appendix B contains some auxiliary estimates.

We close the introduction with a list of notation. $\mathcal{S}(\mathbb{R}^d)$ is the Schwartz space, $\mathcal{B}_{p,q}^\alpha(\mathbb{R}^d)$ is the Besov space, $W_p^\alpha(\mathbb{R}^d)$ is the fractional Sobolev space, $\mathcal{C}^\alpha(\mathbb{R}^d) := \mathcal{B}_{\infty,\infty}^\alpha(\mathbb{R}^d)$ is the Hölder–Zygmund space (see [2] for precise definitions). $L^p = L^p(\mathbb{R}^d)$ is the Lebesgue space while $L_m = L^m(\Omega)$ is the moment space. Deterministic norms are denoted by $|\cdot|$. For example, the norms in $\mathcal{B}_{p,q}^\alpha(\mathbb{R}^d)$ and $L^p(\mathbb{R}^d)$ are denoted by $|\cdot|_{\mathcal{B}_{p,q}^\alpha}$ and $|\cdot|_{L^p}$ respectively. The norm in the moment space $L_m(\Omega)$ is denoted by $\|\cdot\|_m$. For an V -valued random variable X , where V is a Banach space, the moment norm of X is denoted by $\|X\|_{V;m} := \| |X|_V \|_m$. For each $q \in [1, \infty]$, q' denotes the Hölder conjugate of q , defined by $\frac{1}{q} + \frac{1}{q'} = 1$. We always use the convention that $1/\infty = 0$. The notation \lesssim means $\leq C$ for some universal constant $C > 0$ whose value can change from one line to another.

2 Vector valued stochastic processes

Throughout the article, $T > 0$ is a fixed finite time and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0,T]}, \mathbb{P})$ is a complete filtered probability space such that \mathcal{F}_0 contains the \mathbb{P} -null sets. The expectation with respect to \mathbb{P} is denoted by \mathbb{E} while the conditional expectation with respect to \mathcal{F}_s is denoted by \mathbb{E}_s . For a topological space S , the Borel σ -field on S is denoted by $\mathcal{B}(S)$.

Let $(V, |\cdot|_V)$ be a separable Banach space and $\mathcal{B}(V)$ be its Borel σ -field. A V -valued random variable Y is a measurable map $Y : (\Omega, \mathcal{F}) \rightarrow (V, \mathcal{B}(V))$. Since V is separable, weakly measurable maps are strongly measurable (this is Pettis measurability theorem [17, Theorem 1.1.20]) and we will henceforth always write “measurable” for “strongly measurable”.

For $m \in [1, \infty)$, we say that a V -valued random variable Y is L_m -integrable if $\mathbb{E}(|Y|_V^m)$ is finite. When $m = 1$, we simply say Y is integrable. In this case, $\mathbb{E}Y$ is a well-defined element in V and one has

$$|\mathbb{E}Y|_V \leq \mathbb{E}(|Y|_V).$$

The space of all L_m -integrable V -valued random variables is denoted by $L_m(V)$. As commonly practiced, the dependence on Ω is omitted. Nevertheless, when it is necessary to emphasize the stochastic basis, we write $L^m((\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}); V)$ or $L^m(\Omega; V)$ for $L_m(V)$. The norm on $L_m(V)$ is defined by

$$Y \mapsto \|Y\|_{V;m} := (\mathbb{E}|Y|_V^m)^{1/m}.$$

Let \mathcal{G} be a sub σ -field of \mathcal{F} and Y be a V -valued random variable. The conditional expectation of Y with respect to \mathcal{G} , whenever it exists, is a (unique) V -valued random variable $Z = \mathbb{E}(Y|\mathcal{G})$ such that

$$\mathbb{E}(Y\mathbf{1}_J) = \mathbb{E}(Z\mathbf{1}_J) \quad \forall J \in \mathcal{G}.$$

If Y is integrable, we have from [17, Lemma 2.6.19] that $|\mathbb{E}(Y|\mathcal{G})|_V \leq \mathbb{E}(|Y|_V|\mathcal{G})$. We also denote

$$\|Y|\mathcal{G}\|_{V;m} := [\mathbb{E}(|Y|_V^m|\mathcal{G})]^{1/m}$$

which is a \mathcal{G} -random variable. When $V = \mathbb{R}^k$ for some integer $k \geq 1$, we simply write $\|\cdot\|_m$ and $\|\cdot|\mathcal{G}\|_m$ respectively for $\|\cdot\|_{\mathbb{R}^k,m}$ and $\|\cdot|\mathcal{G}\|_{\mathbb{R}^k,m}$. For further information on integration and conditional expectation of Banach-valued random variables, we refer to [17, Sections 1.2 and 2.6].

Let $(f_h)_{h=0}^N$ be a sequence of integrable V -valued random variables and $\{\mathcal{F}_h\}_{h=0}^N$ be a non-decreasing sequence of subfields of \mathcal{F} . We say that $(f_h, \mathcal{F}_h)_{h=0}^N$, or simply $(f_h)_{h=0}^N$, is a martingale if

$$\mathbb{E}(df_{h+1}|\mathcal{F}_h) = 0 \quad \forall h \geq 0, \quad \text{where } df_h = f_h - f_{h-1}.$$

When f_h is L_m -integrable for every h , we say that $(f_h)_{h=0}^N$ is an L_m -integrable martingale.

Definition 2.1. We say that V has martingale type $p \in [1, 2]$ if for some (and hence for all, see [17, Proposition 3.5.27]) $m \in (1, \infty)$, there exists a constant $C_{m,p,V}$ such that

$$\|f_N\|_{L_m(V)} \leq C_{m,p,V} \left\| \left(|f_0|_V^p + \sum_{n=1}^N |df_n|_V^p \right)^{1/p} \right\|_{L_m(\mathbb{R})} \tag{2.1}$$

for every L_m -integrable V -valued martingale $(f_n)_{n=0}^N$.

We note that every Banach space has trivial martingale type $p = 1$, in which (2.1) follows from the triangle inequality. Hence, if $p \neq 1$, we say that V has non-trivial martingale type. Although not being used, it is worth noting the following relation between martingale type and smoothness of Banach spaces, due to Pisier.

Proposition 2.2 ([25, Corollary 4.22]). Let V be a Banach space and p be in $[1, 2]$. V has martingale type p if and only if V is (p, C) -smooth for some constant C , i.e. there is an equivalent norm $|\cdot|$ on V such that

$$2^{-1}(|x+y|^p + |x-y|^p) \leq |x|^p + C^p|y|^p, \quad \forall x, y \in V.$$

In our considerations, it is useful to observe that (2.1) also holds with conditional expectations. To be more precise, suppose that $(f_h)_{h=0}^N$ is an L_m -integrable V -valued martingale with respect to a filtration $\{\mathcal{F}_h\}_{h=0}^N$ and \mathcal{G} is a sub σ -field of \mathcal{F}_0 . Then

$$\|f_N|\mathcal{G}\|_{V;m} \leq C_{m,p,V} \left\| \left(|f_0|_V^p + \sum_{n=1}^N |df_n|_V^p \right)^{1/p} \right\|_{L_m} \Big|_{\mathcal{G}} \quad \text{a.s.} \tag{2.2}$$

where $C_{m,p,V}$ is the same constant in (2.1). Indeed, let G be a measurable set in \mathcal{G} . It is easily checked that $(f_h \mathbf{1}_G)_{h=0}^N$ is an L_m -integrable V -valued martingale. Hence, applying (2.1) gives

$$\|f_N|\mathbf{1}_G\|_{L_m} \leq C_{m,p,V} \left\| \left(|f_0|_V^p + \sum_{n=1}^N |df_n|_V^p \right)^{1/p} \right\|_{L_m} \mathbf{1}_G.$$

Since this inequality holds for every G in \mathcal{G} , one can apply [17, Lemma 2.6.15] to obtain (2.2).

Hereafter, the following assumption will be enforced.

Assumption 2.3. V has non-trivial martingale type $p \in (1, 2]$.

Under Assumption 2.3, inequalities (2.1) and (2.2) hold for every $m \in (1, \infty)$, every L_m -integrable V -valued martingale $(f_n, \mathcal{F}_n)_{n=0}^N$ and every subfield \mathcal{G} of \mathcal{F}_0 . We collect a few known examples Banach spaces with non-trivial type \mathfrak{p} .

Proposition 2.4 (Examples of Banach spaces with non-trivial type). *Let $p, q \in (1, \infty)$ and $\alpha \in \mathbb{R}$.*

- (a) *Every Hilbert space has martingale type 2.*
- (b) *The Lebesgue space $L^p(\mathbb{R}^d)$ has martingale type $\mathfrak{p} = \min\{2, p\}$.*
- (c) *The Besov space $\mathcal{B}_{p,q}^\alpha(\mathbb{R}^d)$ has martingale type $\mathfrak{p} = \min\{2, p, q\}$.*
- (d) *The Triebel–Lizorkin space $F_{p,q}^\alpha(\mathbb{R}^d)$ has martingale type $\mathfrak{p} = \min\{2, p, q\}$.*
- (e) *The Sobolev space $W_p^\alpha(\mathbb{R}^d)$ has martingale type $\mathfrak{p} = \min\{2, p\}$.*

Proof. We mostly point to existing literature and refer to [17, 18] for relevant definitions of (Rademacher) type and cotype, martingale cotype and UMD property. (a) comes from the Pythagorean identity

$$\|f_N\|_{L_2(V)}^2 = \|f_0\|_{L_2(V)}^2 + \sum_{n=1}^N \|df_n\|_{L_2(V)}^2$$

which is valid for Hilbert space V and any martingale $(f_n)_{n=0}^N$. (b) is proved in [17, Proposition 3.5.30]. The Besov space $\mathcal{B}_{p,q}^\alpha(\mathbb{R}^d)$ and $F_{p,q}^\alpha(\mathbb{R}^d)$ have UMD-property ([6, Theorem 6.3]) and have (Rademacher) type $\min\{2, p, q\}$. On UMD spaces, type and martingale type are equivalent, [17, Proposition 4.3.13]. This implies (c) and (d). When α is not an integer, (e) follows from (c) because $W_p^\alpha(\mathbb{R}^d) = \mathcal{B}_{p,p}^\alpha(\mathbb{R}^d)$. When α is a non-negative integer, $W_p^\alpha(\mathbb{R}^d)$ has type $\min\{2, p\}$ and cotype $\max\{2, p\}$ ([5]). Additionally, $W_p^\alpha(\mathbb{R}^d)$ is UMD ([17, Example 4.2.18]), hence it has martingale type $\min\{2, p\}$. When α is a negative integer, by duality ([17, Prop. 3.5.29]), the martingale type of $W_p^\alpha(\mathbb{R}^d)$ is the same as the martingale cotype of $W_{p'}^{-\alpha}(\mathbb{R}^d)$, $\frac{1}{p} + \frac{1}{p'} = 1$. We have deduced (e) for all possible values of α . \square

Further examples of Banach spaces of non-trivial martingale type are the Besov spaces of modeled distributions, see [22]. We conclude the section by some useful estimates for adapted sequences of V -valued random variables.

Lemma 2.5. *Let m be a real number and n be an extended real number such that $n, m \geq \mathfrak{p}$. Let $\{y_k\}_{k \geq 0}$ be a sequence of random variables in V which is L_m -integrable and adapted to a filtration $\{\mathcal{F}_k\}_{k \geq 0}$. For each $k \geq 1$, we denote $\mathbb{E}_{k-1}y_k = \mathbb{E}(y_k|\mathcal{F}_{k-1})$. Let \mathcal{G} be a σ -field such that $\mathcal{G} \subset \mathcal{F}_0$. Then for every $N \geq 0$*

$$\begin{aligned} \left\| \sum_{k=0}^N y_k | \mathcal{G} \right\|_{V;m} \|n\| &\leq \sum_{k=1}^N \left\| \mathbb{E}_{k-1} y_k | \mathcal{G} \right\|_{V;m} \|n\| \\ &+ C_{m,p,V} \left(\left\| y_0 | \mathcal{G} \right\|_{V;m} \|n\|^{\mathfrak{p}} + \sum_{k=1}^N \left\| y_k - \mathbb{E}_{k-1} y_k | \mathcal{G} \right\|_{V;m} \|n\|^{\mathfrak{p}} \right)^{\frac{1}{\mathfrak{p}}} \end{aligned} \tag{2.3}$$

and

$$\left\| \sum_{k=0}^N y_k | \mathcal{G} \right\|_{V;m} \|n\| \leq \sum_{k=1}^N \left\| \mathbb{E}_{k-1} y_k | \mathcal{G} \right\|_{V;m} \|n\| + 2C_{m,p,V} \left(\sum_{k=0}^N \left\| y_k | \mathcal{G} \right\|_{V;m} \|n\|^{\mathfrak{p}} \right)^{\frac{1}{\mathfrak{p}}}. \tag{2.4}$$

Proof. Define $f_{-1} = 0$, $f_0 = y_0$ and $f_k = y_0 + \sum_{h=1}^k (y_h - \mathbb{E}_{h-1}y_h)$ for each $k \geq 1$. Then $\{f_k, \mathcal{F}_k\}_{k \geq 0}$ is a martingale with values in V which is L_m -integrable. From the Doob's decomposition

$$\sum_{k=0}^N y_k = \sum_{k=1}^N \mathbb{E}_{k-1}y_k + f_N$$

we obtain by triangle inequality that

$$\left\| \left\| \sum_{k=0}^N y_k | \mathcal{G} \right\|_{V;m} \right\|_n \leq \sum_{k=1}^N \left\| \left\| \mathbb{E}_{k-1}y_k | \mathcal{G} \right\|_{V;m} \right\|_n + \left\| \left\| f_N | \mathcal{G} \right\|_{V;m} \right\|_n.$$

Since $m, n \geq p$, from (2.2), we apply Minkowski inequality to obtain that

$$\left\| \left\| f_N | \mathcal{G} \right\|_{V;m} \right\|_n \leq C_{m,p,V} \left\| \left\| \left(\sum_{k=0}^N |df_k|_V^p \right)^{1/p} | \mathcal{G} \right\|_m \right\|_n \leq C_{m,p,V} \left(\sum_{k=0}^N \left\| \left\| df_k | \mathcal{G} \right\|_{V;m} \right\|_n^p \right)^{1/p}.$$

This leads to the following inequality

$$\left\| \left\| \sum_{k=0}^N y_k | \mathcal{G} \right\|_{V;m} \right\|_n \leq \sum_{k=1}^N \left\| \left\| \mathbb{E}_{k-1}y_k | \mathcal{G} \right\|_{V;m} \right\|_n + C_{m,p,V} \left(\sum_{k=0}^N \left\| \left\| df_k | \mathcal{G} \right\|_{V;m} \right\|_n^p \right)^{1/p}.$$

Since $df_0 = y_0$ and $df_k = y_k - \mathbb{E}_{k-1}y_k$ for $k \geq 1$, the previous inequality is equivalent to (2.3). The estimate (2.4) follows from (2.3) if we can show that

$$\left\| \left\| y_k - \mathbb{E}_{k-1}y_k | \mathcal{G} \right\|_{V;m} \right\|_n \leq 2 \left\| \left\| y_k | \mathcal{G} \right\|_{V;m} \right\|_n$$

for every $k \geq 1$. This is a trivial consequence of triangle inequality and the estimate $|\mathbb{E}_{h-1}y_h|_V \leq \mathbb{E}_{h-1}(|y_h|_V)$. \square

3 Stochastic sewing lemma in Banach spaces

3.1 The results

We recall that $(V, |\cdot|_V)$ is a separable Banach space with martingale type $p \in (1, 2]$, Δ is the simplex $\{(s, t) \in [0, T]^2 : s \leq t\}$ for some fixed $T > 0$ and w is a continuous control on Δ . We denote by Δ_2 the set $\{(s, u, t) \in [0, T]^3 : s \leq u \leq t\}$ and by $\mathcal{P}(I)$ the set of partitions of some interval I . For each $\pi \in \mathcal{P}(I)$, we define its mesh size with respect to the control w as $|\pi|_w := \sup_{[u,v] \in \pi} w(u, v)$.

Theorem 3.1 (Stochastic Sewing Lemma in Banach spaces). *Let $p \leq m \leq n \leq \infty$ be fixed, $m < \infty$. Let $A : \Omega \times \Delta \rightarrow V$ be $\mathcal{F} \otimes \mathcal{B}(\Delta) / \mathcal{B}(V)$ -measurable such that $A_{s,s} = 0$ and $A_{s,t}$ is \mathcal{F}_t -measurable for every $(s, t) \in \Delta$. For each $t \in [0, T]$, define*

$$\mathcal{A}_t = \lim_{\pi \in \mathcal{P}([0,t]), |\pi|_w \downarrow 0} \sum_{[u,v] \in \pi} A_{u,v} \tag{3.1}$$

whenever the limit exists in probability. Suppose that there are constants $\Gamma_1, \Gamma_2 \geq 0$ such that for any $(s, u, t) \in \Delta_2$

$$\left\| \left\| \mathbb{E}_s[\delta A_{s,u,t}] \right\|_{V;n} \right\| \leq \Gamma_1 w(s, t)^{1+\varepsilon_1} \tag{3.2}$$

and

$$\left\| \left\| \delta A_{s,u,t} - \mathbb{E}_s \delta A_{s,u,t} | \mathcal{F}_s \right\|_{V;m} \right\|_n \leq \Gamma_2 w(s, t)^{\frac{1}{p} + \varepsilon_2}. \tag{3.3}$$

Then, we have

(a) \mathcal{A}_t is well-defined and $\mathcal{A}_t - A_{0,t}$ is L_m -integrable for each $t \in [0, T]$, \mathcal{A} is $\{\mathcal{F}_t\}$ -adapted,

(b) there exists a constant $C > 0$ such that for every $(s, t) \in \Delta$,

$$\|\mathbb{E}_s(\mathcal{A}_t - \mathcal{A}_s - A_{s,t})\|_{V;n} \leq C\Gamma_1 w(s, t)^{1+\varepsilon_1} \tag{3.4}$$

and

$$\|\|\mathcal{A}_t - \mathcal{A}_s - A_{s,t}|\mathcal{F}_s\|_{V;m}\|_n \leq C\Gamma_1 w(s, t)^{1+\varepsilon_1} + C\Gamma_2 w(s, t)^{\frac{1}{p}+\varepsilon_2}. \tag{3.5}$$

It is implicitly assumed in (3.2) that $\delta A_{s,u,t}$ is integrable for each $(s, u, t) \in \Delta_2$, which does not imply the integrability of $A_{s,t}$ for each $(s, t) \in \Delta$. For this reason, the process \mathcal{A} is not necessary integrable. From (3.4) and (3.5), one can derive a rate of convergence of the Riemann sums.

Corollary 3.2 (Riemann sum approximation). *Assume that the hypotheses of Theorem 3.1 are satisfied. Let $(s, t) \in \Delta$ be fixed and π be a partition of $[s, t]$. Define the Riemann sum*

$$A^\pi = \sum_{[u,v] \in \pi} A_{u,v}.$$

Then, there exists a constant $C > 0$ such that

$$\|\delta \mathcal{A}_{s,t} - A^\pi\|_{V;m} \leq C\Gamma_1 |\pi|_w^{\varepsilon_1} w(s, t) + C\Gamma_2 |\pi|_w^{\varepsilon_2} w(s, t)^{\frac{1}{p}}. \tag{3.6}$$

The following result can be considered as an extension of the Doob–Meyer decomposition for vector valued stochastic processes.

Theorem 3.3. *Suppose that the hypotheses of Theorem 3.1 hold. In addition, assume that A is integrable and there are constants $\Gamma_3 \geq 0, \varepsilon_3 > 0$ such that*

$$\|\|(\mathbb{E}_s - \mathbb{E}_u)A_{u,t}|\mathcal{F}_s\|_{V;m}\|_n \leq \Gamma_3 w(s, t)^{\frac{1}{p}+\varepsilon_3}, \tag{3.7}$$

for every $(s, u, t) \in \Delta_2$. Then, there exist stochastic processes $\mathcal{M}, \mathcal{J} : \Omega \times [0, T] \rightarrow V$ and positive constants C_1, C_2, C_3 satisfying the following properties

- (i) \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 $t \in [0, T]$,
- (ii) $(\mathcal{M}_s)_{0 \leq s \leq T}$ is an $\{\mathcal{F}_t\}$ -martingale with $\mathcal{M}_0 = 0$,
- (iii) for any $(s, t) \in \Delta$,

$$\|\|\mathcal{M}_t - \mathcal{M}_s - A_{s,t} + \mathbb{E}_s A_{s,t}|\mathcal{F}_s\|_{V;m}\|_n \leq C_2 \Gamma_2 w(s, t)^{\frac{1}{p}+\varepsilon_2} + C_3 \Gamma_3 w(s, t)^{\frac{1}{p}+\varepsilon_3}, \tag{3.8}$$

(iv) for any $(s, t) \in \Delta$,

$$\|\|\mathcal{J}_t - \mathcal{J}_s - \mathbb{E}_s A_{s,t}|\mathcal{F}_s\|_{V;m}\|_n \leq C_1 \Gamma_1 w(s, t)^{1+\varepsilon_1} + C_3 \Gamma_3 w(s, t)^{\frac{1}{p}+\varepsilon_3}, \tag{3.9}$$

(v) for any $(s, t) \in \Delta$,

$$\|\mathbb{E}_s(\mathcal{J}_t - \mathcal{J}_s - A_{s,t})\|_{V;n} \leq C_1 \Gamma_1 w(s, t)^{1+\varepsilon_1}. \tag{3.10}$$

Given A , we have the following characterizations.

- (vi) The pair of processes $(\mathcal{M}, \mathcal{J})$ is characterized uniquely by the set of properties (i), (ii), (iii) or, alternatively by the set of properties (i), (ii), (iv).

(vii) The process \mathcal{M} is characterized uniquely by (ii) and (iii).

(viii) The process \mathcal{J} is characterized uniquely by (iv), (v) and the fact that \mathcal{J} is $\{\mathcal{F}_t\}$ -adapted and $\mathcal{J}_0 = 0$.

Furthermore, for every fixed $t \in [0, T]$ and any partition $\pi = \{0 = t_0 < t_1 < \dots < t_N = t\}$ of $[0, t]$, define the Riemann sums

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

Then $\{M_t^\pi - A_{0,t} + \mathbb{E}_0 A_{0,t}\}_\pi$ and $\{J_t^\pi - \mathbb{E}_0 A_{0,t}\}_\pi$ converge to $\mathcal{M}_t - A_{0,t} + \mathbb{E}_0 A_{0,t}$ and $\mathcal{J}_t - \mathbb{E}_0 A_{0,t}$ respectively in L_m as $|\pi|_w$ goes to 0.

Remark 3.4. Relations between Theorems 3.1 and 3.3 and Itô calculus are described in [19]. The conditional norms (with $n = \infty$) in Theorem 3.1 play an indispensable role in [10]. An immediate application of Theorem 3.3 which is described in [10] is the decomposition of a stochastic controlled rough path as the sum of a martingale and a controlled rough path. We refer to the afore-mentioned references for further detail.

3.2 Proofs

Lemma 3.5. Let $w, \eta : \Delta \rightarrow \mathbb{R}_+$ be functions such that w is super-additive and

$$\lim_{|t-s| \rightarrow 0} \eta(s, t) = 0.$$

Let $(R_{s,t})_{(s,t) \in \Delta}$ be a two-parameter stochastic process with values in V such that R is $\{\mathcal{F}_t\}$ -adapted and satisfies

$$\|R_{s,t}\|_{V;p} \leq (w(s, t)\eta(s, t))^{1/p} \quad \text{and} \quad \|\mathbb{E}_s R_{s,t}\|_{V;p} \leq w(s, t)\eta(s, t). \quad (3.11)$$

Then

$$\lim_{|\pi| \rightarrow 0} \sum_{[u,v] \in \pi} R_{u,v} = 0 \quad \text{in} \quad L_p(V), \quad (3.12)$$

where $|\pi| = \sup_{[u,v] \in \pi} |v - u|$. In particular, if additionally R is additive, i.e. $R_{s,u} + R_{u,t} = R_{s,t}$ a.s. for all $(s, u, t) \in \Delta_2$, then R is identically 0.

Proof. Let $\varepsilon \in (0, 1)$ be fixed. Let π be a partition of $[0, T]$ such that $\sup_{[s,t] \in \pi} \eta(s, t) \leq \varepsilon$. Applying estimate (2.4) (with $m = n = p$), we see that

$$\left\| \sum_{[u,v] \in \pi} R_{u,v} \right\|_{V;p} \lesssim \left(\sum_{[u,v] \in \pi} \|R_{u,v}\|_{V;p}^p \right)^{1/p} + \sum_{[u,v] \in \pi} \|\mathbb{E}_u R_{u,v}\|_{V;p}.$$

Using super-additivity of w and condition (3.11), we have

$$\sum_{[u,v] \in \pi} \|R_{u,v}\|_{V;p}^p \leq \varepsilon \sum_{[u,v] \in \pi} w(u, v) \leq \varepsilon w(0, T).$$

In a similar way, we have $\sum_{[u,v] \in \pi} \|\mathbb{E}_u R_{u,v}\|_{V;p} \leq \varepsilon w(0, T)$. It follows that

$$\left\| \sum_{[u,v] \in \pi} R_{u,v} \right\|_{V;p} \lesssim \varepsilon^{1/p}.$$

This implies (3.12).

Suppose additionally that R is additive. For each $(s, t) \in \Delta$ and any partition π of $[s, t]$, we have by additivity $\sum_{[u,v] \in \pi} R_{u,v} = R_{s,t}$ a.s. Hence, (3.12) implies that $R_{s,t} = 0$. \square

Let w be a control. For each $(s, t) \in \Delta$, define

$$u = \inf\{r \in [s, t] : w(s, r) \geq \frac{1}{2}w(s, t)\}$$

and call u as the w -midpoint of $[s, t]$. Since t trivially belongs to the set defining u above, such a point always exists and is uniquely defined. If u is a w -midpoint of $[s, t]$, then it follows from the continuity of w that

$$w(s, u) \leq \frac{1}{2}w(s, t) \quad \text{and} \quad w(u, t) \leq \frac{1}{2}w(s, t). \tag{3.13}$$

Indeed, by definition, there exists a sequence $\{r_j\}_j \subset [s, t]$ decreasing to u such that $w(s, r_j) \geq \frac{1}{2}w(s, t)$. Since w is super-additive, $w(s, r_j) + w(r_j, t) \leq w(s, t)$. This implies that $w(r_j, t) \leq \frac{1}{2}w(s, t)$. Since w is continuous, $\lim_j w(r_j, t) = w(u, t)$. This implies that $w(u, t) \leq \frac{1}{2}w(s, t)$. We show $w(s, u) \leq \frac{1}{2}w(s, t)$ by contradiction. Suppose that $w(s, u) > \frac{1}{2}w(s, t)$. By continuity, there is $\varepsilon > 0$ such that $w(s, r) > \frac{1}{2}w(s, t)$ for every $r \in (u - \varepsilon, u)$. This contradicts with the definition of u .

Let (s, t) be in Δ . For each integer $h \geq 0$, we define h -dyadic points of $[s, t]$ with respect to the control w in the following way. Define $d_0^0(s, t) = s$ and $d_1^0(s, t) = t$. For each integer $i = 0, \dots, 2^{h+1}$, we set $d_i^{h+1}(s, t) = d_{i/2}^h(s, t)$ if i is even and $d_i^{h+1}(s, t)$ equal to the w -midpoint of $[d_{(i-1)/2}^h(s, t), d_{(i+1)/2}^h(s, t)]$ if i is odd. In the specific case when $w(\bar{s}, \bar{t}) = \bar{t} - \bar{s}$, we have $d_i^h = s + i2^{-h}(t - s)$ for every $i = 0, \dots, 2^{h+1}$, which are the standard dyadic points of $[s, t]$. Thus, the collection $D_w^h(s, t) := \{d_i^h(s, t)\}_{i=0}^{2^{h+1}}$ is regarded as h -dyadic points of $[s, t]$ with respect to the control w . As expected, these collections of points have similar properties to the standard dyadic points. Namely, for every integers $h \geq 0$ and $i = 0, \dots, 2^h - 1$, we have $D_w^h(s, t) \subset D_w^{h+1}(s, t)$,

$$[d_i^h(s, t), d_{i+1}^h(s, t)] = [d_{2i}^{h+1}(s, t), d_{2i+1}^{h+1}(s, t)] \cup [d_{2i+1}^{h+1}(s, t), d_{2i+2}^{h+1}(s, t)] \tag{3.14}$$

and

$$w(d_i^h(s, t), d_{i+1}^h(s, t)) \leq 2^{-h}w(s, t). \tag{3.15}$$

Lemma 3.6. *Let $A : \Omega \times \Delta \rightarrow V$ be $\mathcal{F} \otimes \mathcal{B}(\Delta)/\mathcal{B}(V)$ -measurable such that $A_{s,s} = 0$ for every $s \in [0, T]$. Let (s, t) be in Δ . Then, for every $N \geq 0$ and every $s \leq t_0 < \dots < t_N \leq t$, there exist a positive integer h_0 and random variables $R_i^h, i = 0, \dots, 2^h - 1, h \geq 0$, such that*

- (i) $R_i^h = 0$ for every $h \geq h_0$ and every i ;
- (ii) for each h, i , there exist four (not necessarily distinct) points $s_1^{h,i} \leq s_2^{h,i} \leq s_3^{h,i} \leq s_4^{h,i}$ in $[d_i^h(t_0, t_N), d_{i+1}^h(t_0, t_N)]$ so that

$$R_i^h = A_{s_1^{h,i}, s_2^{h,i}} + A_{s_2^{h,i}, s_3^{h,i}} + A_{s_3^{h,i}, s_4^{h,i}} - A_{s_1^{h,i}, s_4^{h,i}}; \tag{3.16}$$

- (iii) the following identity holds

$$\sum_{i=0}^{N-1} A_{t_i, t_{i+1}} - A_{t_0, t_N} = \sum_{h \geq 0} \sum_{i=0}^{2^h-1} R_i^h. \tag{3.17}$$

Proof. For brevity, we abbreviate d_i^h for $d_i^h(t_0, t_N)$. For each collection $\mathcal{P} = \{s_i\}_{i=0}^K$ we define

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

and $I(\mathcal{P}) = 0$ whenever $K = 0$ or \mathcal{P} is empty. For any two finite collections $\mathcal{P}_1, \mathcal{P}_2$, define

$$\delta I(\mathcal{P}_1, \mathcal{P}_2) = I(\mathcal{P}_1 \cup \mathcal{P}_2) - I(\mathcal{P}_1) - I(\mathcal{P}_2). \tag{3.18}$$

Put $\mathcal{P}_0^0 = \{t_i\}_{i=0}^N$, which is a subset of $[d_0^0, d_1^0]$. The main idea of the proof is to allocate the elements of \mathcal{P}_0^0 into the w -dyadic subintervals of $[s, t]$ while keeping track of the resulting changes in $I(\mathcal{P}_0^0)$ during the process. For each $h \geq 1$, define

$$\mathcal{P}_{2^{h-1}}^h = \mathcal{P}_0^0 \cap [d_{2^{h-1}}^h, d_{2^h}^h] \quad \text{and} \quad \mathcal{P}_i^h = \mathcal{P}_0^0 \cap [d_i^h, d_{i+1}^h) \quad \text{for } i = 0, \dots, 2^h - 2.$$

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

$$R_i^h := \delta I(\mathcal{P}_{2i}^{h+1}, \mathcal{P}_{2i+1}^{h+1}) = I(\mathcal{P}_i^h) - I(\mathcal{P}_{2i}^{h+1}) - I(\mathcal{P}_{2i+1}^{h+1}) \tag{3.19}$$

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

Since \mathcal{P}_0^0 is a finite set, there exists a finite integer $h_0 \geq 1$ so that $[d_i^h, d_{i+1}^h] \cap \mathcal{P}_0^0$ contains at most one point for every $h \geq h_0$ and every $i = 0, \dots, 2^h - 1$. Hence, when $h \geq h_0$, we have $I(\mathcal{P}_i^h) = 0$ and $R_i^h = 0$ for every i . This shows (i).

If either \mathcal{P}_{2i}^{h+1} or \mathcal{P}_{2i+1}^{h+1} is empty, then $R_i^h = 0$ and (3.16) is satisfied with $s_j^{h,i} = d_i^h$ for $j = 1, 2, 3, 4$. We assume that \mathcal{P}_{2i}^{h+1} and \mathcal{P}_{2i+1}^{h+1} are not empty. In such case, we define

$$s_1^{h,i} = \min \mathcal{P}_{2i}^{h+1}, \quad s_2^{h,i} = \max \mathcal{P}_{2i}^{h+1}, \quad s_3^{h,i} = \min \mathcal{P}_{2i+1}^{h+1} \quad \text{and} \quad s_4^{h,i} = \max \mathcal{P}_{2i+1}^{h+1}.$$

Here, \min (respectively \max) of a nonempty finite set F is the smallest (respectively largest) element of F . We derive (3.16) from (3.19) and the definition of I at the beginning of the proof. Hence, (ii) is verified.

Lastly, to show (iii), we apply (3.19) recursively to see that

$$\begin{aligned} I(\mathcal{P}_0^0) &= I(\mathcal{P}_0^1) + I(\mathcal{P}_1^1) + R_0^0 \\ &= \sum_{i=0}^{2^h-1} I(\mathcal{P}_i^h) + \sum_{k=0}^{h-1} \sum_{i=0}^{2^k-1} R_i^k \quad \text{for every } h \geq 1. \end{aligned}$$

Since $I(\mathcal{P}_i^h) = 0$ as soon as $h \geq h_0$, the previous identity implies (3.17). This completes the proof. \square

Lemma 3.7. *Let A be the process in Theorem 3.1. Then for every $N \geq 0$ and every $S \leq t_0 < \dots < t_N \leq T$, we have*

$$\left\| \mathbb{E}_{t_0} \left(\sum_{i=0}^{N-1} A_{t_i, t_{i+1}} - A_{t_0, t_N} \right) \right\|_{V;n} \leq C_1 \Gamma_1 w(t_0, t_N)^{1+\varepsilon_1} \tag{3.20}$$

and

$$\left\| \sum_{i=0}^{N-1} A_{t_i, t_{i+1}} - A_{t_0, t_N} \Big| \mathcal{F}_{t_0} \right\|_{V;m} \leq C_1 \Gamma_1 w(t_0, t_N)^{1+\varepsilon_1} + C_2 \Gamma_2 w(t_0, t_N)^{\frac{1}{p}+\varepsilon_2} \tag{3.21}$$

where $C_1 = 2(1 - 2^{-\varepsilon_1})^{-1}$ and $C_2 = 2C_{m,p,V}(1 - 2^{-\varepsilon_2})^{-1}$.

Proof. Put $s = t_0$ and $t = t_N$ and $d_i^h = d_i^h(s, t)$. Applying Lemma 3.6, we can find random variables $R_i^h, i = 0, \dots, 2^h - 1, h \geq 0$ which satisfy properties (i)-(iii) stated there.

Let $h \geq 0$ be fixed. Define $\mathcal{G}_{2^h}^h = \mathcal{F}_t$ and $\mathcal{G}_i^h = \mathcal{F}_{s_1^{h,i}}$ for each $i = 0, \dots, 2^h - 1$. We recall that $s_1^{h,i}$ is defined in (ii). The sequence $\{\mathcal{G}_i^h\}_{i=0}^{2^h}$ forms a filtration such that R_i^h is \mathcal{G}_{i+1}^h -measurable for every $i = 0, \dots, 2^h - 1$. The formula (3.16) can be written as

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

Applying the conditions (3.2) and (3.3), we obtain from the previous identity that

$$\|\mathbb{E}(R_i^h|\mathcal{G}_i^h)\|_{V;n} \leq 2\Gamma_1 w(d_i^h, d_{i+1}^h)^{1+\varepsilon_1} \tag{3.22}$$

and

$$\| \|R_i^h - \mathbb{E}(R_i^h|\mathcal{G}_i^h)|\mathcal{G}_i^h \|_{V;m} \|_n \leq 2\Gamma_2 w(d_i^h, d_{i+1}^h)^{\frac{1}{2}+\varepsilon_2}. \tag{3.23}$$

From (2.3), we have

$$\begin{aligned} \| \| \sum_{i=0}^{2^h-1} R_i^h | \mathcal{F}_{t_0} \|_{V;m} \|_n &\leq \sum_{i=0}^{2^h-1} \| \| \mathbb{E}(R_i^h|\mathcal{G}_i^h) | \mathcal{F}_{t_0} \|_{V;m} \|_n \\ &\quad + C_{m,p,V} \left(\sum_{i=0}^{2^h-1} \| \| R_i^h - \mathbb{E}(R_i^h|\mathcal{G}_i^h) | \mathcal{F}_{t_0} \|_{V;m} \|_n^p \right)^{\frac{1}{p}}. \end{aligned}$$

Since $m \leq n$ and $\mathcal{F}_{t_0} \subset \mathcal{G}_i^h$, we have

$$\| \| \mathbb{E}(R_i^h|\mathcal{G}_i^h) | \mathcal{F}_{t_0} \|_{V;m} \|_n \leq \mathbb{E}(R_i^h|\mathcal{G}_i^h)\|_{V;n}$$

and

$$\| \| R_i^h - \mathbb{E}(R_i^h|\mathcal{G}_i^h) | \mathcal{F}_{t_0} \|_{V;m} \|_n \leq \| \| R_i^h - \mathbb{E}(R_i^h|\mathcal{G}_i^h) | \mathcal{G}_i^h \|_{V;m} \|_n.$$

Taking into account (3.22) and (3.23), we get

$$\begin{aligned} \| \| \sum_{i=0}^{2^h-1} R_i^h | \mathcal{F}_{t_0} \|_{V;m} \|_n &\leq 2\Gamma_1 \sum_{i=0}^{2^h-1} w(d_i^h, d_{i+1}^h)^{1+\varepsilon_1} \\ &\quad + 2C_{m,p,V}\Gamma_2 \left(\sum_{i=0}^{2^h-1} w(d_i^h, d_{i+1}^h)^{1+p\varepsilon_2} \right)^{\frac{1}{p}}. \end{aligned}$$

From the estimate (3.15), we see that

$$\sum_{i=0}^{2^h-1} w(d_i^h, d_{i+1}^h)^{1+\varepsilon_1} \leq \sum_{i=0}^{2^h-1} 2^{-h(1+\varepsilon_1)} w(s, t)^{1+\varepsilon_1} = 2^{-h\varepsilon_1} w(s, t)^{1+\varepsilon_1}$$

and similarly

$$\sum_{i=0}^{2^h-1} w(d_i^h, d_{i+1}^h)^{1+p\varepsilon_2} \leq 2^{-hp\varepsilon_2} w(s, t)^{1+p\varepsilon_2}.$$

We combine the previous inequalities to obtain that

$$\| \| \sum_{i=0}^{2^h-1} R_i^h | \mathcal{F}_{t_0} \|_{V;m} \|_n \leq 2^{-h\varepsilon_1} 2\Gamma_1 w(s, t)^{1+\varepsilon_1} + 2^{-h\varepsilon_2} 2C_{m,p,V}\Gamma_2 w(s, t)^{\frac{1}{p}+\varepsilon_2}. \tag{3.24}$$

From (3.17), applying triangle inequality and the above estimate, we see that

$$\begin{aligned} \| \| \sum_{i=0}^{N-1} A_{t_i, t_{i+1}} - A_{t_0, t_N} | \mathcal{F}_{t_0} \|_{V;m} \|_n &\leq \left(\sum_{h \geq 0} 2^{-h\varepsilon_1} \right) 2\Gamma_1 w(s, t)^{1+\varepsilon_1} \\ &\quad + \left(\sum_{h \geq 0} 2^{-h\varepsilon_2} \right) 2C_{m,p,V}\Gamma_2 w(s, t)^{\frac{1}{p}+\varepsilon_2}. \end{aligned}$$

This implies (3.21). To show (3.20), we obtain from (3.17) and triangle inequality that

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

We note that $\|\mathbb{E}(R_i^h | \mathcal{F}_{t_0})\|_{V;n} \leq \|\mathbb{E}(R_i^h | \mathcal{G}_i^h)\|_{V;n}$. Hence we can use (3.22) to estimate the above sum analogously as before. This implies (3.20). \square

Proof of Theorem 3.1. We divide the proof into two steps. In the first step, a process \mathcal{A} is constructed as the limit of some Riemann sums. The second step verifies the properties (a) and (b).

Step 1. Let t be fixed but arbitrary in $[0, T]$. 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 $L_m(V)$, denoted by \mathcal{A}_t , as the mesh size $|\pi|_w := \max_i w(t_i, t_{i+1})$ 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 (2.4), we have

$$\|A_t^{\pi''} - A_t^\pi\|_{V;m} \lesssim \sum_{i=0}^{n-1} \|\mathbb{E}_{t_i} Z_i\|_{V;m} + \left(\sum_{i=0}^{n-1} \|Z_i\|_{V;m}^p \right)^{\frac{1}{p}}. \tag{3.25}$$

Lemma 3.7 is applied to obtain that

$$\|\mathbb{E}_{t_i} Z_i\|_{V;m} \lesssim w(t_i, t_{i+1})^{1+\varepsilon_1} \quad \text{and} \quad \|Z_i\|_{V;m} \lesssim w(t_i, t_{i+1})^{\frac{1}{p}+\varepsilon_2}.$$

By super-additivity of w , this implies that

$$\sum_{i=0}^{n-1} \|\mathbb{E}_{t_i} Z_i\|_{V;m} \lesssim \sum_{i=0}^{n-1} w(t_i, t_{i+1})^{1+\varepsilon_1} \lesssim |\pi|_w^{\varepsilon_1}$$

and similarly $\sum_{i=0}^{n-1} \|Z_i\|_{V;m}^p \lesssim |\pi|_w^{p\varepsilon_2}$. Hence, we have

$$\|A_t^\pi - A_t^{\pi''}\|_{V;m} \lesssim |\pi|_w^{\varepsilon_1} + |\pi|_w^{\varepsilon_2}.$$

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

$$\|A_t^\pi - A_t^{\pi'}\|_{V;m} \leq \|A_t^\pi - A_t^{\pi''}\|_{V;m} + \|A_t^{\pi'} - A_t^{\pi''}\|_{V;m} \lesssim |\pi|_w^{\varepsilon_1} + |\pi|_w^{\varepsilon_2} + |\pi'|_w^{\varepsilon_1} + |\pi'|_w^{\varepsilon_2}.$$

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

Step 2. We show that the process $(\mathcal{A}_t)_{0 \leq t \leq T}$ defined in the previous step satisfies (a) and (b).

The condition (3.3) with $s = u = t$ implies that $A_{s,s} = 0$ for every $s \in [0, T]$. Hence, it is evident that $\mathcal{A}_0 = 0$. The fact that A is $\{\mathcal{F}_t\}$ -adapted implies that \mathcal{A} is $\{\mathcal{F}_t\}$ -adapted. Obviously, $\mathcal{A}_t - A_{0,t}$, being a limit in $L_m(V)$, belongs to $L_m(V)$ for each t . This shows (a).

Let (s, t) be fixed but arbitrary in Δ . Let $\pi = \{s = t_0 < \dots < t_N = t\}$ be an arbitrary partition of $[s, t]$. From construction of \mathcal{A} in the previous step, we see that

$$\mathcal{A}_t - \mathcal{A}_s - A_{s,t} = \lim_{|\pi|_w \downarrow 0} (A_{s,t}^\pi - A_{s,t}) \quad \text{in } L_m(V). \tag{3.26}$$

Hence, passing through the limit $|\pi|_w \downarrow 0$ in (3.20) and (3.21), we obtain (3.4) and (3.5) respectively. \square

Proof of Corollary 3.2. We start from a trivial identity

$$\delta \mathcal{A}_{s,t} - A^\pi = \sum_{[u,v] \in \pi} (\delta \mathcal{A}_{u,v} - A_{u,v}).$$

Then by (2.4), we have

$$\begin{aligned} \|\delta \mathcal{A}_{s,t} - A^\pi\|_{V;m} &\leq \sum_{[u,v] \in \pi} \|\mathbb{E}_u(\delta \mathcal{A}_{u,v} - A_{u,v})\|_{V;m} \\ &\quad + 2C_{m,p,V} \left(\sum_{[u,v] \in \pi} \|\delta \mathcal{A}_{u,v} - A_{u,v}\|_{V;m}^p \right)^{1/p}. \end{aligned}$$

Using the estimate (3.4), the first sum on the right-hand side above is bounded above by a constant multiple of

$$\Gamma_1 \sum_{[u,v] \in \pi} w(u,v)^{1+\varepsilon_1} \leq \Gamma_1 |\pi|_w^{\varepsilon_1} w(s,t).$$

The later sum is estimated similarly, using (3.5). This completes the proof. \square

Proof of Theorem 3.3. For every $(s, t) \in \Delta$, we define $J_{s,t} = \mathbb{E}_s A_{s,t}$ and $M_{s,t} = A_{s,t} - \mathbb{E}_s A_{s,t}$. Then for every $(s, u, t) \in \Delta_2$, we have

$$\delta J_{s,u,t} = \delta A_{s,u,t} + (\mathbb{E}_s - \mathbb{E}_u) A_{u,t}, \quad \mathbb{E}_s \delta J_{s,u,t} = \mathbb{E}_s \delta A_{s,u,t}$$

and

$$\delta M_{s,u,t} = -(\mathbb{E}_s - \mathbb{E}_u) A_{u,t}, \quad \mathbb{E}_s \delta M_{s,u,t} = 0.$$

Applying Theorem 3.1 for J and M , we obtain the existence of the processes \mathcal{J} and \mathcal{M} respectively. Since A is integrable, so are \mathcal{J} and \mathcal{M} . The estimate (3.4) yields (3.9). In addition, the estimate (3.5) implies that \mathcal{J} satisfies (3.10) and \mathcal{M} is a martingale. Since $A_{s,t} = J_{s,t} + M_{s,t}$, it is evident that $\mathcal{A} = \mathcal{J} + \mathcal{M}$. Uniqueness follows from Lemma 3.5. \square

4 Additive functionals of fractional Brownian motion

Let $B = (B^1, \dots, B^d)$ be a fractional Brownian motion in \mathbb{R}^d with Hurst parameter $H \in (0, 1)$. Let f be a time-dependent distribution in $L^\theta([0, T]; \mathcal{B}_{p,1}^\alpha(\mathbb{R}^d))$ where θ, p, α are some fixed parameters, $p, \theta \in (1, \infty)$ and $\alpha \in \mathbb{R}$. In the current section, we study the additive functional

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

as the continuous extension of the map

$$\mathcal{S}([0, T] \times \mathbb{R}^d) \ni f \mapsto \left((t, x) \mapsto \int_0^t f_r(B_r + x) dr \right).$$

For each $i \in \{1, \dots, d\}$, B^i has Mandelbrot–Van Ness representation ([24]) with respect to a standard two sided Wiener process W^i on \mathbb{R} , namely

$$B_u^i = \int_{-\infty}^u [(u-r)_+^{H-\frac{1}{2}} - (-r)_+^{H-\frac{1}{2}}] dW_r^i \tag{4.1}$$

where $(x)_+ = \max\{x, 0\}$. We assume that W^i 's (and hence B^i 's) are independent. For every $i \in \{1, \dots, d\}$ and $0 \leq u \leq v$, we have

$$B_v^i = \int_{-\infty}^u [(v-r)^{H-\frac{1}{2}} - (-r)_+^{H-\frac{1}{2}}] dW_r^i + \int_u^v (v-r)^{H-\frac{1}{2}} dW_r^i, \tag{4.2}$$

which also yields that $\mathbb{E}_u B_v^i = \int_{-\infty}^u [(v-r)^{H-\frac{1}{2}} - (-r)_+^{H-\frac{1}{2}}] dW_r^i$. We define

$$\rho(u, v) = \mathbb{E} \left(\int_u^v (v-r)^{H-\frac{1}{2}} dW_r^i \right)^2 = \frac{1}{2H} (v-u)^{2H}. \tag{4.3}$$

For a $d \times d$ -symmetric positive definite matrix Σ , p_Σ denotes the density of a normal random variable in \mathbb{R}^d with mean 0 and variance Σ , i.e.

$$p_\Sigma(x) = (2\pi)^{-\frac{d}{2}} (\det \Sigma)^{\frac{1}{2}} \exp \left(-\frac{1}{2} x^* \Sigma^{-1} x \right). \tag{4.4}$$

P_Σ denotes the spatial convolution operator with p_Σ . The $d \times d$ identity matrix is denoted by I .

Proposition 4.1. *Let p, q be fixed numbers in $(1, \infty)$. Let f be a function in $L^\theta([0, T]; \mathcal{S}(\mathbb{R}^d) \cap \mathcal{B}_{p,1}^\alpha)$, $\theta \in (1, \infty)$. Let γ be a positive number satisfying*

$$\gamma < \frac{1}{H} \left(1 - \frac{1}{\min(2, \theta, p, q)} \right). \tag{4.5}$$

Then for any $\alpha \in \mathbb{R}$, $m \geq 2$ and $(s, t) \in \Delta$,

$$\left\| \int_s^t f_r(B_r + \cdot) dr \Big| \mathcal{F}_s \right\|_{\mathcal{B}_{p,q}^{\alpha+\gamma}; m} \leq C |\mathbf{1}_{[s,t]} f|_{L^\theta \mathcal{B}_{p,\infty}^\alpha} |t-s|^{1-H\gamma-\frac{1}{\theta}}. \tag{4.6}$$

Proof. We observe that because $\mathcal{B}_{p,1}^\alpha \hookrightarrow \mathcal{B}_{p,\infty}^\alpha$, the quantity $|\mathbf{1}_{[s,t]} f|_{L^\theta \mathcal{B}_{p,\infty}^\alpha}$ on the right-hand side of (4.6) is finite. Define $A_{s,t}(x) = \int_s^t P_{\rho(s,r)} I f_r(\mathbb{E}_s B_r + x) dr$. We note that the Besov norm $|\cdot|_{\mathcal{B}_{p,q}^\alpha}$ is translation invariant, i.e. $|f(y + \cdot)|_{\mathcal{B}_{p,q}^{\alpha+\gamma}} = |f|_{\mathcal{B}_{p,q}^{\alpha+\gamma}}$ for every $y \in \mathbb{R}^d$. Note that

$$\langle \rho(s, r) I \xi, \xi \rangle \geq \frac{1}{2} (r-s)^{2H} |\xi|^2.$$

Using triangle inequality and Lemma B.2, we see that

$$\begin{aligned} |A_{s,t}|_{\mathcal{B}_{p,q}^{\alpha+\gamma}} &\leq \int_s^t |P_{\rho(s,r)} I f_r(\mathbb{E}_s B_r + \cdot)|_{\mathcal{B}_{p,q}^{\alpha+\gamma}} dr \\ &= \int_s^t |P_{\rho(s,r)} I f_r|_{\mathcal{B}_{p,q}^{\alpha+\gamma}} dr \lesssim \int_s^t |r-s|^{-H\gamma} |f_r|_{\mathcal{B}_{p,\infty}^\alpha} dr. \end{aligned}$$

By Hölder inequality, we have

$$|A_{s,t}|_{\mathcal{B}_{p,q}^{\alpha+\gamma}} \lesssim |\mathbf{1}_{[s,t]} f|_{L^\theta \mathcal{B}_{p,\infty}^\alpha} |t-s|^{1-H\gamma-\frac{1}{\theta}},$$

where we have used (4.5) to ensure that the integral in time is finite. Define w by

$$|w(s, t)|^{1-H\gamma} = |\mathbf{1}_{[s,t]} f|_{L^\theta \mathcal{B}_{p,\infty}^\alpha} |t-s|^{1-H\gamma-\frac{1}{\theta}}. \tag{4.7}$$

Note that $w = w_1^a w_2^{1-a}$ with $a = 1/(\theta(1 - H\gamma))$ and controls $w_1(s, t) = |\mathbf{1}_{[s,t]} f|_{L^\theta \mathcal{B}_{p,\infty}^\alpha}$, $w_2(s, t) = (t - s)$, one sees that w is a continuous control. Hence, the previous estimate yields

$$|A_{s,t}|_{\mathcal{B}_{p,q}^{\alpha+\gamma}} \lesssim w(s, t)^{1-H\gamma}. \tag{4.8}$$

We recall from Proposition 2.4 that $\mathcal{B}_{p,q}^{\alpha+\gamma}$ has martingale type $\mathfrak{p} = \min\{2, p, q\}$. The condition (4.5) also implies that $1 - H\gamma > \frac{1}{\mathfrak{p}}$. In view of (4.8), condition (3.3) is satisfied. It is straightforward to verify that $\mathbb{E}_s \delta A_{s,u,t} = 0$ for every $s \leq u \leq t$, hence condition (3.2) holds trivially. Hence, we can apply the stochastic sewing lemma, Theorem 3.1, to define the process (\mathcal{A}_t) as in (3.1). The estimates (3.5) and (4.8) imply that

$$\| \delta \mathcal{A}_{s,t} | \mathcal{F}_s \|_{\mathcal{B}_{p,q}^{\alpha+\gamma}; m} \| \infty \lesssim |\mathbf{1}_{[s,t]} f|_{L^\theta \mathcal{B}_{p,\infty}^\alpha} |t - s|^{1-H\gamma-\frac{1}{\mathfrak{p}}}.$$

Hence, to obtain (4.6), it remains to show that $\mathcal{A}_t = \int_0^t f_r(B_r + \cdot) dr$.

We put

$$R_{s,t}(x) = \int_s^t f_r(B_r + x) dr - A_{s,t}(x).$$

It is evident that $\mathbb{E}_s R_{s,t}(x) = 0$. In addition, by Minkowski inequality

$$\| \int_s^t f_r(B_r + \cdot) dr \|_{\mathcal{B}_{p,q}^{\alpha+\gamma}} \leq \int_s^t \| f_r \|_{\mathcal{B}_{p,q}^{\alpha+\gamma}} dr.$$

Combining with (4.8) gives

$$|R_{s,t}|_{\mathcal{B}_{p,q}^{\alpha+\gamma}} \lesssim w(s, t)^{1-H\gamma} + \int_s^t \| f_r \|_{\mathcal{B}_{p,q}^{\alpha+\gamma}} dr.$$

Applying Lemma 3.5, we have that for each $t \in [0, T]$,

$$\int_0^t f_r(B_r + \cdot) dr = \lim_{\pi \in \mathcal{P}([0,t]), |\pi| \rightarrow 0} \sum_{[u,v] \in \pi} A_{u,v},$$

which shows that $\mathcal{A}_t = \int_0^t f_r(B_r + \cdot) dr$ and hence, completes the proof. □

For each $f \in L^\theta([0, T]; \mathcal{S})$, define

$$I[f]_t(x) = \int_0^t f_r(B_r + x) dr. \tag{4.9}$$

Proposition 4.1 shows that the map

$$I : L^\theta([0, T]; \mathcal{S} \cap \mathcal{B}_{p,1}^\alpha) \rightarrow C([0, T]; L_m \mathcal{B}_{p,q}^{\alpha+\gamma})$$

is bounded. By a density argument, we can extend I to $L^\theta([0, T]; \mathcal{B}_{p,1}^\alpha)$.

Theorem 4.2. *For any $\theta, p, q \in (1, \infty)$ and γ satisfying (4.5), the map*

$$I : L^\theta([0, T]; \mathcal{B}_{p,1}^\alpha(\mathbb{R}^d)) \rightarrow C([0, T]; L_m \mathcal{B}_{p,q}^{\alpha+\gamma}(\mathbb{R}^d))$$

is well-defined as the unique continuous extension of (4.9). In addition, one has for any $f \in \mathcal{B}_{p,1}^\alpha(\mathbb{R}^d)$

$$\| \delta I[f]_{s,t} | \mathcal{F}_s \|_{\mathcal{B}_{p,q}^{\alpha+\gamma}; m} \| \infty \leq C |\mathbf{1}_{[s,t]} f|_{L^\theta \mathcal{B}_{p,\infty}^\alpha} |t - s|^{1-H\gamma-\frac{1}{\mathfrak{p}}}. \tag{4.10}$$

Proof. Using the embedding $\mathcal{B}_{p,1}^\alpha \hookrightarrow \mathcal{B}_{p,\infty}^\alpha$ and the trivial estimate $\|\delta I[f]_{s,t}\|_{\mathcal{B}_{p,q}^{\alpha+\gamma};m} \leq \|\|\delta I[f]_{s,t}|_{\mathcal{F}_s}\|_{\mathcal{B}_{p,q}^{\alpha+\gamma};m}\|_\infty$, we obtain from (4.6) that

$$\|\delta I[f]_{s,t}\|_{\mathcal{B}_{p,q}^{\alpha+\gamma};m} \leq C|\mathbf{1}_{[s,r]}f|_{L^\theta([0,T];\mathcal{B}_{p,1}^\alpha)}|t-s|^{1-H\gamma-\frac{1}{\theta}}$$

for every $f \in \mathcal{S}$. Since \mathcal{S} is dense in $\mathcal{B}_{p,1}^\alpha$ and the action $f \mapsto I[f]$ is linear, this implies that I has a unique continuous extension $I : L^\theta([0, T]; \mathcal{B}_{p,1}^\alpha) \rightarrow C([0, T]; L_m \mathcal{B}_{p,q}^{\alpha+\gamma})$.

Moving on to (4.10). Let f be in $L^\theta([0, T]; \mathcal{B}_{p,1}^\alpha)$. For each $n \geq 1$, define $f^n = P_{1/n} I f$ which belongs to $L^\theta([0, T]; \mathcal{S})$. We have $\lim_n f^n = f$ in $L^\theta([0, T]; \mathcal{B}_{p,1}^\alpha)$ and $|f^n|_{\mathcal{B}_{p,\infty}^\alpha} \leq |f_r|_{\mathcal{B}_{p,\infty}^\alpha} \leq |f_r|_{\mathcal{B}_{p,1}^\alpha}$. Then from (4.6), we get that

$$\begin{aligned} \|\|\delta I[f^n]_{s,t}|_{\mathcal{F}_s}\|_{\mathcal{B}_{p,q}^{\alpha+\gamma};m}\|_\infty &\leq C|\mathbf{1}_{[s,t]}f^n|_{L^\theta \mathcal{B}_{p,\infty}^\alpha}|t-s|^{1-H\gamma-\frac{1}{\theta}} \\ &\leq C|\mathbf{1}_{[s,t]}f|_{L^\theta \mathcal{B}_{p,\infty}^\alpha}|t-s|^{1-H\gamma-\frac{1}{\theta}}. \end{aligned}$$

Passing through the limit $n \rightarrow \infty$, using the continuity of I on $L^\theta([0, T]; \mathcal{B}_{p,1}^\alpha)$ we obtain (4.10). □

Remark 4.3. Using Besov-Sobolev embeddings

$$\begin{aligned} \mathcal{B}_{p,1}^\alpha &\hookrightarrow W_p^\alpha \hookrightarrow \mathcal{B}_{p,\infty}^\alpha \hookrightarrow \mathcal{B}_{p,1}^{\alpha-\varepsilon} \quad \text{for } \varepsilon > 0, \\ \mathcal{B}_{p,q}^\alpha &\hookrightarrow \mathcal{B}_{p_1,q_1}^\beta \quad \text{for } \alpha - \frac{d}{p} = \beta - \frac{d}{p_1}, \quad p \leq p_1, \quad q \leq q_1, \end{aligned}$$

and the isomorphism $\mathcal{B}_{p,p}^\alpha = W_p^\alpha$ when α is not an integer, Theorem 4.2 can be applied to distributions f in $L^\theta([0, T]; \mathcal{B}_{p,\infty}^\alpha)$ and $L^\theta([0, T]; W_p^\alpha)$.

Remark 4.4. In view of Proposition A.1 and (4.10), for each $f \in \mathcal{B}_{p,1}^\alpha(\mathbb{R})$, $I[f]$ has a continuous modification (as a process taking values in $\mathcal{B}_{p,q}^{\alpha+\gamma}(\mathbb{R}^d)$).

Corollary 3.2 provides an alternative approximation for $I[f]$ by Riemann sums.

Corollary 4.5. Let $p, q, \theta \in (1, \infty)$, $\alpha \in \mathbb{R}$ and f be in $L^\theta([0, T]; \mathcal{B}_{p,1}^\alpha)$. Let $(s, t) \in \Delta$ and π be a partition of $[s, t]$ and define the Riemann sum

$$I^\pi[f]_{s,t}(x) = \sum_{[u,v] \in \pi} \int_u^v P_{\rho(u,r)} I f_r(\mathbb{E}_u B_r + x) dr.$$

Let w be the control defined by the relation (4.7). Then for any γ satisfying (4.5),

$$\|\delta I[f]_{s,t} - I^\pi[f]_{s,t}\|_{\mathcal{B}_{p,q}^{\alpha+\gamma}(\mathbb{R}^d);m} \lesssim |\pi|_w^{1-H\gamma-\frac{1}{\min(2,p,q)}} w(s,t)^{\frac{1}{\min(2,p,q)}}. \tag{4.11}$$

Proof. Straightforward from Theorem 4.2 and Corollary 3.2. □

Using embeddings between Besov spaces and Hölder-Zygmund spaces \mathcal{C}^β , we can derive from Theorem 4.2 the almost sure continuity the additive functional $I[f]$.

Corollary 4.6. Let $p, \theta \in (1, \infty)$, $\alpha \in \mathbb{R}$ and f be in $L^\theta([0, T]; \mathcal{B}_{p,1}^\alpha)$.

(i) Assume that

$$\alpha + \frac{1}{H} \left(1 - \frac{1}{\min(2, \theta, p)}\right) > \frac{d}{p}. \tag{4.12}$$

Then for every β satisfying

$$0 < \beta < \alpha - \frac{d}{p} + \frac{1}{H} \left(1 - \frac{1}{\min(2, \theta, p)}\right), \tag{4.13}$$

we have

$$\|\|\delta I[f]_{s,t}|_{\mathcal{F}_s}\|_{\mathcal{C}^\beta;m}\|_\infty \lesssim |\mathbf{1}_{[s,t]}f|_{L^\theta \mathcal{B}_{p,\infty}^\alpha}|t-s|^{1-H(\beta-\alpha+\frac{d}{p})-\frac{1}{\theta}} \tag{4.14}$$

for every $(s, t) \in \Delta$ and every $m \geq 2$.

(ii) Assume that

$$\alpha + \frac{1}{H} \left(1 - \frac{1}{\min(2, \theta, p)} \right) > 0. \tag{4.15}$$

Then for any $v \in [p, \infty]$ satisfying

$$\alpha + \frac{1}{H} \left(1 - \frac{1}{\min(2, \theta, p)} \right) > \frac{d}{p} - \frac{d}{v}, \tag{4.16}$$

there exists $\gamma = \gamma(v)$ satisfying (4.5) such that

$$\|\delta I[f]_{s,t} | \mathcal{F}_s\|_{L^v; m} \|\infty \lesssim |\mathbf{1}_{[s,t]} f|_{L^\theta \mathcal{B}_{p,\infty}^\alpha} |t - s|^{1-H\gamma - \frac{1}{\theta}} \tag{4.17}$$

for every $(s, t) \in \Delta$ and every $m \geq 2$.

Proof. (i). Assume that β satisfy (4.13) and define γ by the relation $\beta = \gamma + \alpha - \frac{d}{p}$. Then γ satisfies the condition (4.5) and we have the embedding $\mathcal{B}_{p,p}^{\alpha+\gamma} \hookrightarrow \mathcal{C}^\beta$. From Theorem 4.2, we deduce that $I[f]_t$ belongs to \mathcal{C}^β and the estimate (4.14) follows from (4.6).

(ii) From the conditions (4.15) and (4.16), we can choose γ satisfying (4.5) (with $q = p$) such that $\gamma + \alpha > \frac{d}{p} - \frac{d}{v}$. It suffices to apply Theorem 4.2 and the embeddings $\mathcal{B}_{p,p}^{\gamma+\alpha} \hookrightarrow \mathcal{B}_{v,v}^\varepsilon \hookrightarrow L^v(\mathbb{R}^d)$, $\gamma + \alpha - \frac{d}{p} = \varepsilon - \frac{d}{v}$. \square

Remark 4.7. We do not require $\beta \in (0, 1)$ in Corollary 4.6(i). This means that when $k := \alpha - \frac{d}{p} + \frac{1}{H} \left(1 - \frac{1}{\min(2, \theta, p)} \right)$ is larger than 1, the functional $I[f]_t$ is n -times differentiable in the spatial variables for any integer $n < k$.

The class $L^\theta([0, T]; \mathcal{B}_{1,\infty}^0(\mathbb{R}^d))$ contains Dirac distributions and the corresponding functional I is directly related to the local time of fractional Brownian motion. For this class, Theorem 4.2 is still applicable through the Besov embedding $\mathcal{B}_{1,\infty}^0 \hookrightarrow \mathcal{B}_{2,2}^{-d/2}$.

Corollary 4.8. Let f be a distribution in $L^\theta([0, T]; \mathcal{B}_{1,\infty}^0(\mathbb{R}^d))$ with $\theta \geq 2$.

(i) (Small H) When $0 < Hd < \frac{1}{2}$, for every $t \in [0, T]$, $I[f]_t$ belongs to $L^v(\mathbb{R}^d)$ almost surely for every $v \in [2, \infty]$.

(ii) (Large H) When $\frac{1}{2} \leq Hd < 1$, for every $t \in [0, T]$, $I[f]_t$ belongs to $L^v(\mathbb{R}^d)$ almost surely for every $v \in [2, \frac{2Hd}{2Hd-1})$. Here we use the convention that $\frac{2Hd}{2Hd-1} = \infty$ if $Hd = \frac{1}{2}$.

Proof. Straightforward from the embedding $\mathcal{B}_{1,\infty}^0 \hookrightarrow \mathcal{B}_{2,2}^{-d/2}$ and Corollary 4.6. \square

A A continuity criterion with controls

We give an extension of the classical Kolmogorov continuity theorem with generic controls.

Proposition A.1. Let V be a Banach space, w be a continuous control which is w strictly increasing, i.e. $w(u, v) < w(s, t)$ whenever $[u, v] \subsetneq [s, t]$. Let $m \geq 1$ and $\alpha \in (0, 1]$ be such that $\beta_0 := \alpha - 1/m > 0$. Let \mathcal{A} be a V -valued process such that

$$\|\delta \mathcal{A}_{s,t}\|_{V; m} \leq w(s, t)^\alpha \quad \forall (s, t) \in \Delta. \tag{A.1}$$

Then \mathcal{A} has continuous modification $\tilde{\mathcal{A}}$ and for every $\beta \in (0, \beta_0)$, there is a finite constant $C(\beta, \beta_0)$ such that

$$\left\| \sup_{(s,t) \in \Delta, s < t} \frac{|\delta \tilde{\mathcal{A}}_{s,t}|_V}{w(s, t)^\beta} \right\|_m \leq C(\beta, \beta_0) w(0, T)^{\alpha-\beta}. \tag{A.2}$$

Without the monotonicity condition, one can always replace w in (A.1) by the strictly increasing control $(s, t) \mapsto w(s, t) + \varepsilon(t - s)$ for some $\varepsilon > 0$. By using a time change argument in combination with the standard Kolmogorov continuity theorem, it is easy to derive that \mathcal{A} has a continuous version. Below, we provide a direct proof, modifying the proof of the standard Kolmogorov continuity theorem ([26, pg. 26]) to replace the standard dyadic points by the dyadic points with respect to the control w (defined in Section 3.2). Although less elegant than the time-change argument, it may be useful for other situations.

Proof of Proposition A.1. Without loss of generality, we assume that $w(0, T) = 1$. Recall the definition of $d_i^h(0, T)$ from Section 3.2. Let $d_i^h = d_i^h(0, T)$, $D^h = \{d_i^h\}_{i=0}^{2^h-1}$ and $D = \cup_{h \geq 0} D^h$. Because of the monotonicity of w , we have $d_i^h < d_{i+1}^h$ and that D is a dense subset of $[0, T]$. Define

$$K_h = \sup_{s, t \in D^h : s \leq t, w(s, t) \leq 2^{1-h}} |\mathcal{A}_t - \mathcal{A}_s|_V, \quad h \geq 0,$$

and recall that $\beta_0 = \alpha - 1/m > 0$. Then by (A.1),

$$\mathbb{E}|K_h|^m \leq \sum_{s, t \in D^h : s \leq t, w(s, t) \leq 2^{1-h}} \mathbb{E}|\delta\mathcal{A}_{s,t}|_V^m \leq 2^{h+1} 2^{(1-h)m\alpha} = 2^{1+m\alpha} 2^{-hm\beta_0}.$$

Let s, t be in D , $s < t$. For each n , define

$$s_n = \inf\{r \in D^n : r \geq s\} \quad \text{and} \quad t_n = \sup\{r \in D^n : r \leq t\}.$$

It is straightforward to see that $(s_n)_n$ is decreasing and $s_n = s$ for some n on; $(t_n)_n$ is increasing and $t_n = t$ for some n on; and for every n

$$w(s_{n+1}, s_n) \leq 2^{-n}, \quad w(t_n, t_{n+1}) \leq 2^{-n}.$$

To see this, let s'_n be the dyadic point adjacent to s_n to the left, i.e. $s'_n = \max\{r \in D^n : r < s_n\}$. Then we have $s'_n \leq s \leq s_n$, so that $w(s_{n+1}, s_n) \leq w(s, s_n) \leq w(s'_n, s_n) \leq 2^{-n}$ by (3.13). Similarly, let t'_n be the dyadic point adjacent to t_n to the right. One has $t_n \leq t \leq t'_n$, which implies the estimate for $w(t_n, t_{n+1})$.

If $h \geq 0$ is an integer satisfying $w(s, t) \leq 2^{-h}$, then we have additionally that

$$\mathbf{1}_{(s_h \leq t_h)} w(s_h, t_h) + \mathbf{1}_{(t_h \leq s_h)} w(t_h, s_h) \leq 2^{-h}.$$

Indeed, if $s_h \leq t_h$, then $[s_h, t_h] \subset [s, t]$ and we have $w(s_h, t_h) \leq w(s, t) \leq 2^{-h}$. If $t_h < s_h$ then one has $t_h \leq s \leq t \leq s_h$. In this case, we must have the identity $t_h = s'_h$, where s'_h is defined previously (namely, the dyadic point adjacent to s_h to the left). Then $w(t_h, s_h) = w(s'_h, s_h) \leq 2^{-h}$ by (3.13).

Let s, t be in D , $s < t$ and $w(s, t) \leq 2^{-h}$. We have

$$\delta\mathcal{A}_{s,t} = \delta\mathcal{A}_{s_h, t_h} + \sum_{i=h}^{\infty} \delta\mathcal{A}_{s_i, s_{i+1}} + \sum_{i=h}^{\infty} \delta\mathcal{A}_{t_i, t_{i+1}},$$

where the series are actually finite sums. From the definition of K_n and properties of s_n, t_n described previously, it follows that

$$|\delta\mathcal{A}_{s,t}|_V \leq 2K_h + 2 \sum_{i=h}^{\infty} K_{i+1} \leq 2 \sum_{i=h}^{\infty} K_i.$$

Consequently, setting $M_\beta = \sup\{|\delta\mathcal{A}_{s,t}|_V/w(s,t)^\beta; s, t \in D, s < t\}$ for $\beta \in (0, \beta_0)$, we have

$$\begin{aligned} M_\beta &\leq \sup_{h \geq 0} \sup_{s, t \in D: s < t, w(s,t) \leq 2^{-h}} 2^{(h+1)\beta} |\delta\mathcal{A}_{s,t}|_V \\ &\leq \sup_{h \geq 0} 2^{1+(h+1)\beta} \sum_{i=h}^\infty K_i \leq 2^{1+\beta} \sum_{i=0}^\infty 2^{i\beta} K_i. \end{aligned}$$

It follows that

$$\|M_\beta\|_m \leq 2^{1+\beta} \sum_{i=0}^\infty 2^{i\beta} \|K_i\|_m \leq 2^{3+\beta} \sum_{i=0}^\infty 2^{i(\beta-\beta_0)} < \infty.$$

In particular, for a.e. ω , $\mathcal{A}(\omega)$ is uniformly continuous on D and it makes sense to define for every $t \in [0, T]$,

$$\tilde{\mathcal{A}}_t(\omega) = \lim_{s \rightarrow t, s \in D} \mathcal{A}_s(\omega).$$

It is now standard to verify that $\tilde{\mathcal{A}}$ is the desired modification and that (A.2) holds with the constant $C(\beta, \beta_0) = 2^{3+\beta} \sum_{i=0}^\infty 2^{i(\beta-\beta_0)}$. □

B Auxiliary estimates

To obtain various properties of Besov spaces, we will make use of the following Bernstein’s inequalities. Let f be a function in $L^p(\mathbb{R}^d)$ and let $q \geq p$, $p, q \in [1, \infty]$. For every integer $k \geq 0$, every $\lambda > 0$ and $t > 0$ we have ([2, Lemma 2.1])

$$\text{supp } Ff \subset \lambda\mathcal{B} \Rightarrow \|\nabla^k f\|_{L^q(\mathbb{R}^d)} \leq C^{k+1} \lambda^{k+d(\frac{1}{p}-\frac{1}{q})} \|f\|_{L^p(\mathbb{R}^d)}, \tag{B.1}$$

where Ff denotes the Fourier transform of f and \mathcal{B} is a ball centered at 0 in \mathbb{R}^d .

Lemma B.1. *Let \mathcal{N} be an annulus. Let a be a $d \times d$ -matrix such that*

$$\langle a\xi, \xi \rangle \geq \kappa_1 |\xi|^2$$

for some $\kappa_1 \in (0, M)$ and let P_a be the spatial convolution operator with p_a (defined in (4.4)). Then there exist positive constants $c, C = C(d, M)$ such that for any $\lambda > 0$, $p \in [1, \infty]$ and any function g whose Fourier transform is supported in $\lambda\mathcal{N}$

$$|P_a g|_{L^p} \leq C e^{-c\kappa_1 \lambda^2} |g|_{L^p}. \tag{B.2}$$

Proof. [2, Lemma 2.4]. □

Lemma B.2. *For $\gamma > 0$, $(p, q) \in [1, \infty]^2$, $|P_a g|_{\mathcal{B}_{p,q}^{\alpha+\gamma}} \lesssim (1 + \kappa^{-\gamma/2}) |g|_{\mathcal{B}_{p,\infty}^\alpha}$.*

Proof. From the embedding $\mathcal{B}_{p,1}^{\alpha+\gamma} \hookrightarrow \mathcal{B}_{p,q}^{\alpha+\gamma}$, it suffices to consider the case $q = 1$. We denote by Δ_j , $j \geq -1$, the (nonhomogeneous) Littlewood-Paley blocks ([2, page 61]). We have from (B.2) and (B.1), for $j \geq 0$,

$$|P_a(\Delta_j g)|_{L^p} \lesssim e^{-2^{2j}\kappa} |\Delta_j g|_{L^p} \lesssim e^{-2^{2j}\kappa} 2^{-\alpha j} |g|_{\mathcal{B}_{p,\infty}^\alpha}.$$

Noting that $\Delta_j P_a g = P_a(\Delta_j g)$, we obtain

$$2^{j(\alpha+\gamma)} |\Delta_j(P_a g)|_{L^p} \lesssim |g|_{\mathcal{B}_{p,\infty}^\alpha} e^{-2^{2j}\kappa} 2^{\gamma j}.$$

Since $\gamma > 0$, it is easy to check that

$$\sup_{\kappa > 0} \kappa^{\frac{\gamma}{2}} \sum_{j \geq 0} e^{-2^{2j}\kappa} 2^{\gamma j} < \infty$$

which implies that

$$\sum_{j \geq 0} |\Delta_j(P_a g)|_{L^p} \lesssim |g|_{\mathcal{B}_{p,\infty}^\alpha} \kappa^{-\frac{\gamma}{2}}.$$

For $j = -1$, we have

$$|\Delta_{-1}(P_a g)|_{L^p} = |P_a(\Delta_{-1}g)|_{L^p} \leq |\Delta_{-1}g|_{L^p} \lesssim |g|_{\mathcal{B}_{p,\infty}^\alpha}.$$

Combining the previous two estimates, we obtain the result. \square

References

- [1] Siva Athreya, Oleg Butkovsky, Khoa Lê, and Leonid Mytnik, *Well-posedness of stochastic heat equation with distributional drift and skew stochastic heat equation*, to appear on Communications on Pure and Applied Mathematics.
- [2] Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin, *Fourier analysis and nonlinear partial differential equations*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 343, Springer, Heidelberg, 2011. MR2768550
- [3] Oleg Butkovsky, Konstantinos Dareiotis, and Máté Gerencsér, *Approximation of SDEs: a stochastic sewing approach*, Probability Theory and Related Fields (2021), 1–60. MR4344136
- [4] R. Catellier and M. Gubinelli, *Averaging along irregular curves and regularisation of ODEs*, Stochastic Process. Appl. **126** (2016), no. 8, 2323–2366. MR3505229
- [5] Fernando Cobos, *Clarkson’s inequalities for Sobolev spaces*, Math. Japon. **31** (1986), no. 1, 17–22. MR835828
- [6] Fernando Cobos and Dicesar Lass Fernandez, *Hardy-Sobolev spaces and Besov spaces with a function parameter*, Function spaces and applications (Lund, 1986), Lecture Notes in Math., vol. 1302, Springer, Berlin, 1988, pp. 158–170. MR942266
- [7] Konstantinos Dareiotis, Máté Gerencsér, and Khoa Lê, *Quantifying a convergence theorem of Gyöngy and Krylov*, Annals of Applied Probability, to appear.
- [8] Denis Feyel and Arnaud de La Pradelle, *Curvilinear integrals along enriched paths*, Electron. J. Probab. **11** (2006), no. 34, 860–892. MR2261056
- [9] Peter K. Friz and Martin Hairer, *A course on rough paths*, second ed., Universitext, Springer, Cham, 2020, With an introduction to regularity structures. MR4174393
- [10] Peter K Friz, Antoine Hocquet, and Khoa Lê, *Rough stochastic differential equations*, arXiv preprint arXiv:2106.10340v2 (2021).
- [11] Donald Geman and Joseph Horowitz, *Occupation densities*, Ann. Probab. **8** (1980), no. 1, 1–67. MR556414
- [12] Máté Gerencsér, *Regularisation by regular noise*, Stochastics and Partial Differential Equations: Analysis and Computations (2022), 1–16.
- [13] M. Gubinelli, *Controlling rough paths*, J. Funct. Anal. **216** (2004), no. 1, 86–140. MR2091358
- [14] Martin Hairer and Xue-Mei Li, *Averaging dynamics driven by fractional Brownian motion*, Ann. Probab. **48** (2020), no. 4, 1826–1860. MR4124526
- [15] Fabian A Harang and Chengcheng Ling, *Regularity of local times associated with Volterra-Lévy Processes and Path-Wise Regularization of Stochastic Differential Equations*, Journal of Theoretical Probability (2021), 1–30. MR4488556
- [16] Fabian Andsem Harang and Nicolas Perkowski, *C-infinity regularization of odes perturbed by noise*, Stochastics and Dynamics **0** (2021), no. 0, 2140010. MR4342752
- [17] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz 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 [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 63, Springer, Cham, 2016. MR3617205

- [18] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis, *Analysis in Banach spaces. Vol. II*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 67, Springer, Cham, 2017, Probabilistic methods and operator theory. MR3752640
- [19] Khoa Lê, *A stochastic sewing lemma and applications*, Electron. J. Probab. **25** (2020), Paper No. 38, 55. MR4089788
- [20] Khoa Lê and Chengcheng Ling, *Taming singular stochastic differential equations: A numerical method*, arXiv preprint arXiv:2110.01343 (2021).
- [21] Xue-Mei Li and Julian Sieber, *Slow-fast systems with fractional environment and dynamics*, Annals of Applied Probability. **32** (2022), no. 5, 3964–4003. MR4498200
- [22] Chong Liu, David J. Prömel, and Josef Teichmann, *Stochastic analysis with modelled distributions*, Stoch. Partial Differ. Equ. Anal. Comput. **9** (2021), no. 2, 343–379. MR4261328
- [23] Terry J. Lyons, *Differential equations driven by rough signals*, Rev. Mat. Iberoamericana **14** (1998), no. 2, 215–310. MR1654527
- [24] Benoit B. Mandelbrot and John W. Van Ness, *Fractional Brownian motions, fractional noises and applications*, SIAM Rev. **10** (1968), 422–437. MR242239
- [25] Gilles Pisier, *Martingales in Banach spaces*, Cambridge Studies in Advanced Mathematics, vol. 155, Cambridge University Press, Cambridge, 2016. MR3617459
- [26] Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999. MR1725357
- [27] L. C. Young, *An inequality of the Hölder type, connected with Stieltjes integration*, Acta Math. **67** (1936), no. 1, 251–282. MR1555421

Acknowledgments. I thank Pavel Zorin-Kranich for various discussions about Proposition 2.4 and the referees for their valuable suggestions.