



HAL
open science

Path-dependent processes from signatures

Eduardo Abi Jaber, Louis-Amand Gérard, Yuxing Huang

► **To cite this version:**

Eduardo Abi Jaber, Louis-Amand Gérard, Yuxing Huang. Path-dependent processes from signatures. 2024. hal-04637402

HAL Id: hal-04637402

<https://hal.science/hal-04637402v1>

Preprint submitted on 6 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Path-dependent processes from signatures

Eduardo Abi Jaber^{*1}, Louis-Amand Gérard^{†2,3}, and Yuxing Huang^{‡1}

¹Ecole Polytechnique, CMAP

²Université Paris 1 Panthéon-Sorbonne, CES

³Gefip

July 6, 2024

Abstract

We provide explicit series expansions to certain stochastic path-dependent integral equations in terms of the path signature of the time augmented driving Brownian motion. Our framework encompasses a large class of stochastic linear Volterra and delay equations and in particular the fractional Brownian motion with a Hurst index $H \in (0, 1)$. Our expressions allow to disentangle an infinite dimensional Markovian structure and open the door to straightforward and simple approximation schemes, that we illustrate numerically.

Mathematics Subject Classification (2010): 60L10, 60L70, 60H20, 60G22

Keywords: Path-signatures, Volterra processes, Stochastic Delay equations, fractional Brownian motion.

1 Introduction

We provide explicit solutions to certain stochastic path-dependent integral equations in the form

$$X_t = X_0 + \int_0^t b(t, s, (X_u)_{u \leq s}) ds + \int_0^t \sigma(t, s, (X_u)_{u \leq s}) dW_s,$$

for some linear coefficients b, σ , in terms of the path signature of the time augmented driving Brownian motion $(t, W_t)_{t \geq 0}$.

We consider two non-Markovian specifications:

- **The Volterra case**, where the coefficients b, σ are of the form

$$g(t, s, (X_u)_{u \leq s}) = K(t - s)(a + bX_s),$$

for some locally integrable deterministic kernel K and $a, b \in \mathbb{R}$. The Riemann-Liouville Fractional Brownian motion and Gaussian Volterra processes constitute a particular case.

- **The Delay case**, where the coefficients b, σ are of the form

$$g(t, s, (X_u)_{u \leq s}) = a + bX_s + \int_0^s K(s - u)X_u du.$$

^{*}eduardo.abi-jaber@polytechnique.edu. The first author is grateful for the financial support from the Chaires FiME-FDD, Financial Risks, Deep Finance & Statistics and Machine Learning and systematic methods in finance at Ecole Polytechnique.

[†]louis-amand.gerard@etu.univ-paris1.fr

[‡]yuxing.math@gmail.com.

The authors would like to thank Xuyang Lin for fruitful discussions.

In both cases, the unique strong solution X is a non-anticipative measurable function of the whole path of the driving Brownian motion, that is

$$X_t = f(t, (W_u)_{u \leq t}), \quad t \geq 0,$$

for some measurable functional f .

Our main contribution is to explicit this functional f . More precisely, we show that the solution X can be written in terms of an infinite linear combination of the signature process $\widehat{\mathbb{W}}_t$ of the time extended Brownian motion $\widehat{W}_t := (t, W_t)$ defined by the infinite sequence of iterated integrals in the Stratonovich sense:

$$\widehat{\mathbb{W}}_t = \left(1, \begin{pmatrix} t \\ W_t \end{pmatrix}, \begin{pmatrix} \frac{t^2}{2!} & \int_0^t s dW_s \\ \int_0^t W_s ds & \frac{W_t^2}{2!} \end{pmatrix}, \begin{pmatrix} \frac{t^3}{3!} & \int_0^t \int_0^s u dW_u ds \\ \int_0^t \int_0^s W_u du ds & \frac{1}{2!} \int_0^t W_s^2 ds \\ \int_0^t \int_0^s s^2 dW_s & \int_0^t \int_0^s u dW_u \circ dW_s \\ \int_0^t \int_0^s W_u du \circ dW_s & \frac{W_t^3}{3!} \end{pmatrix}, \dots \right),$$

that is

$$X_t = \langle \boldsymbol{\ell}_t, \widehat{\mathbb{W}}_t \rangle, \quad t \geq 0, \quad (1.1)$$

where the (possibly) time-dependent deterministic coefficients $\boldsymbol{\ell}_t$, given by a sequence of tensors of increasing order, are explicit.

To highlight our strategy, consider a generalized geometric Brownian motion process X solution to

$$X_t = x + \int_0^t (a + bX_s) ds + \int_0^t (\alpha + \beta X_s) \circ dW_s, \quad \text{with } x, a, b, \alpha, \beta \in \mathbb{R}, \quad (1.2)$$

where \circ denotes Stratonovich integration rule.

Our recipe:

1. Translate the probabilistic equation into an algebraic linear equation on $\boldsymbol{\ell}$:

$$\boldsymbol{\ell} = x\boldsymbol{\emptyset} + (a + b\boldsymbol{\ell})\mathbf{1} + (\alpha + \beta\boldsymbol{\ell})\mathbf{2}, \quad (1.3)$$

where concatenation with the letters $\mathbf{1}$ and $\mathbf{2}$ play respectively the role of integration with respect to dt and dW_t (in the sense of Stratonovich).

2. Solve the linear algebraic equation (1.3) by writing

$$\boldsymbol{\ell} = \boldsymbol{p} + \boldsymbol{\ell}\boldsymbol{q}, \quad (1.4)$$

with

$$\boldsymbol{p} = (x\boldsymbol{\emptyset} + a\mathbf{1} + \alpha\mathbf{2}) \quad \text{and} \quad \boldsymbol{q} = b\mathbf{1} + \beta\mathbf{2},$$

to get

$$\boldsymbol{\ell} = \boldsymbol{p}(\boldsymbol{\emptyset} - \boldsymbol{q})^{-1} = \boldsymbol{p} \sum_{k \geq 0} \boldsymbol{q}^{\otimes k}.$$

3. Argue that the infinite series $\langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \rangle$ is well-defined and solves the stochastic differential equation (1.2).

The recipe is simple yet powerful, as it remains robust across non-Markovian settings and for multi-dimensional Brownian motions, with the same algebraic equation (1.4) but with different \boldsymbol{p} and \boldsymbol{q} . To make the analysis rigorous, we proceed as follows. First, we study the existence of solutions of the algebraic equation (1.3) taking values in the extended tensor algebra together with some estimates for the solutions, see Proposition 2.2 and Theorem 3.9. Second, since $\boldsymbol{\ell}$ can have infinitely many non-zero elements, a delicate analysis is needed to make

sense of the series expansions in (1.1), this is the object of Section 3.3. In particular, we derive a tractable criterion for the global convergence in the supremum norm which is new compared to the related literature on infinite series of signature elements, see for instance Ben Arous [7], Cuchiero, Svaluto-Ferro, and Teichmann [12]. For the final step, a central ingredient for the derivation of our representation formulas in Section 4 is an Itô’s formula for possibly infinite linear combination of signature elements, see Theorem 3.3.

We provide two type of representations: (i) time-independent representations (1.1) for non-Markovian stochastic Volterra and delay equations with smooth enough kernels K where the coefficients ℓ do not depend on time, see Theorems 4.2 and 4.4. Such ℓ would in principal correspond to the coefficients of a stochastic Itô-Taylor expansion applied in a path-dependent setting, except that here we compute ℓ algebraically and establish the convergence of the Taylor series. For more general kernels, we provide approximation results in Corollary 4.3; but also (ii) time-dependent representations (1.1) with ℓ_t depending on time which allowed us to recover more general Gaussian processes in Theorem 4.5 with singular kernels. For instance, for an Ornstein-Uhlenbeck process, a time-dependent representation involves

$$\ell_t^{\text{OU}} = e^{\sqcup - \kappa(t\sigma - 1)} \mathbf{2},$$

which is the algebraic translation of the probabilistic expression

$$X_t = \int_0^t e^{-\kappa(t-s)} dW_s,$$

with e^{\sqcup} the shuffle exponential (2.7).

The representation (1.1) has at least two main features. First, it allows to disentangle an infinite dimensional Markovian structure in terms of the process $\widehat{\mathbb{W}}_t$. Second it opens the door to straightforward and simple approximation schemes by truncating ℓ_t , numerical illustrations are provided in Section 4.4.

Motivation. The elegant framework of path signatures, as introduced by Chen [9] and extensively used in rough path theory [18, 24], not only captures the intricate dynamics of path-dependent processes but also facilitates their integration into various fields of applications. A notable attribute driving the recent prominence of signatures is their inherent ‘linearization property’, or put differently, the universal approximation property of linear functionals of the signature. This property has found many practical applications in Machine Learning (Chevyrev and Kormilitzin [10], Fermanian [15], Perez Arribas, Goodwin, Geddes, Lyons, and Saunders [28]), Mathematical Finance (Abi Jaber and Gérard [3], Bayer, Hager, Riedel, and Schoenmakers [5], Bayer, Pelizzari, and Schoenmakers [6], Cartea, Perez Arribas, and Sanchez-Betancourt [8], Cuchiero, Gazzani, Möller, and Svaluto-Ferro [11], Dupire and Tissot-Daguette [13], Lyons, Nejad, and Arribas [26]). . .

Our motivation for studying exact representation formulae in the form (1.1) for path-dependent processes arises from the importance of non-Markovian stochastic modeling in various fields. Specifically, stochastic Volterra and delay equations capture intricate temporal dependencies in time series data. For such non-Markovian processes, the representation formulae (1.1) not only reveal the underlying Markovian structure in the whole signature process $\widehat{\mathbb{W}}_t$ but also enable us to re-frame them within a rather universal framework of linear functionals of the signature, for which one can develop generic methods. As illustration, recent works by Abi Jaber and Gérard [3], Cuchiero, Svaluto-Ferro, and Teichmann [12] have demonstrated interest in infinite linear combinations of signature elements, unveiling a surprising infinite-dimensional affine structure underlying a broad spectrum of stochastic models. Certain Fourier-Laplace transforms of linear functionals of the signature can be computed in terms of non-standard tensor valued infinite-dimensional Riccati equations and applied in practice for risk management purposes in mathematical finance for instance.

Related literature on series and functional expansions. Stochastic-Taylor expansions of solutions to Markovian stochastic differential equations (SDEs) in terms of the signature of the time augmented driving Brownian motion $(t, W_t)_{t \geq 0}$ have appeared in Kloeden et al. [22, Chapter 5], they have been used to develop a cubature formula on the Wiener space by Lyons [27]; Ben Arous [7] establishes convergence criterion for such stochastic Taylor series for Markovian SDEs with analytic coefficients; the so-called Chen–Fliess [16, 17] approximation that generalizes the classic stochastic Taylor expansion for path dependent functionals of Markovian SDEs is studied by Litterer and Oberhauser [23]; variation of constants formulae for linear forward and

backward stochastic Volterra integral equations using certain sequence of iterated Volterra integrals, adapting Wiener–Itô chaos expansions, are derived by Hamaguchi [21]; a functional Taylor expansion in terms of a finite series of higher order functional derivatives multiplied by iterated integrals is presented by Dupire and Tissot-Daguette [13], a nice overview on series and functional expansions is given there.

Our expansions can be seen as an infinite Stochastic-Taylor series expansion of the solution of path-dependent stochastic equations whose coefficients depend linearly on the whole trajectory. In particular, we believe that the coefficients of our time-independent expansions, obtained in ‘one shot’ by solving an algebraic equation, can be recovered from successive functional derivatives of the coefficients of the stochastic equation. None of the works cited above rely on such algebraic method. When it comes to studying the convergence of the stochastic Taylor series, our work is closest in spirit to Ben Arous [7] which operates in a Markovian setting. Just like Taylor expansions [22, 13], and unlike Volterra and Wiener expansions, the coefficients ℓ that depend on the model parameters and the signature $\widehat{\mathbb{W}}_t$ of the driving noise are disentangled. Combined with the explicit knowledge of the coefficients, this allows for a straightforward and simple numerical implementation of non-Markovian processes.

Outline. Section 2 deals with some preliminaries on path-signatures. Section 3 is concerned with infinite linear combinations of signature elements and their properties. Section 4 collects our representations formulae together with a numerical illustration. The proofs are postponed to later sections and to the Appendices.

2 Preliminaries

In this section, we setup the framework for dealing with signatures of semimartingales. We not only collect some (well-known) properties of finite linear combinations of signature elements, refer also to the first sections in [5, 12, 26], but we also introduce and study the resolvent object in Section 2.2 which plays a crucial role in our main results.

2.1 Tensor algebra

Let $d \in \mathbb{N}$ and denote by \otimes the tensor product over \mathbb{R}^d , e.g. $(x \otimes y \otimes z)_{ijk} = x_i y_j z_k$, for $i, j, k = 1, \dots, d$, for $x, y, z \in \mathbb{R}^d$. For $n \geq 1$, we denote by $(\mathbb{R}^d)^{\otimes n}$ the space of tensors of order n and by $(\mathbb{R}^d)^{\otimes 0} = \mathbb{R}$. In the sequel, we will consider mathematical objects, path signatures, that live on the extended tensor algebra space $T((\mathbb{R}^d))$ over \mathbb{R}^d , that is the space of (infinite) sequences of tensors defined by

$$T((\mathbb{R}^d)) := \{ \ell = (\ell^n)_{n=0}^\infty : \ell^n \in (\mathbb{R}^d)^{\otimes n} \}.$$

Similarly, for $M \geq 0$, we define the truncated tensor algebra $T^M(\mathbb{R}^d)$ as the space of sequences of tensors of order at most M defined by

$$T^M(\mathbb{R}^d) := \{ \ell \in T((\mathbb{R}^d)) : \ell^n = 0, \quad \forall n > M \},$$

and the tensor algebra $T(\mathbb{R}^d)$ as the space of all finite sequences of tensors defined by

$$T(\mathbb{R}^d) := \bigcup_{M \in \mathbb{N}} T^M(\mathbb{R}^d).$$

We clearly have $T(\mathbb{R}^d) \subset T((\mathbb{R}^d))$. For $\ell = (\ell^n)_{n \in \mathbb{N}}, \mathbf{p} = (\mathbf{p}^n)_{n \in \mathbb{N}} \in T((\mathbb{R}^d))$ and $\lambda \in \mathbb{R}$, we define the following operations:

$$\ell + \mathbf{p} := (\ell^n + \mathbf{p}^n)_{n \in \mathbb{N}}, \quad \ell \otimes \mathbf{p} := \left(\sum_{k=0}^n \ell^k \otimes \mathbf{p}^{n-k} \right)_{n \in \mathbb{N}}, \quad \lambda \ell := (\lambda \ell^n)_{n \in \mathbb{N}}.$$

For the rest of the paper, we will use $\ell \mathbf{p}$ and $\ell \otimes \mathbf{p}$ interchangeably. These operations are closed on $T^M(\mathbb{R}^d)$ and $T(\mathbb{R}^d)$.

Important notations. Let $\{e_1, \dots, e_d\} \subset \mathbb{R}^d$ be the canonical basis of \mathbb{R}^d and $A_d = \{\mathbf{1}, \mathbf{2}, \dots, \mathbf{d}\}$ be the corresponding alphabet. To ease reading, for $i \in \{1, \dots, d\}$, we write e_i as the blue letter \mathbf{i} and for $n \geq 1$,

$i_1, \dots, i_n \in \{1, \dots, d\}$, we write $e_{i_1} \otimes \dots \otimes e_{i_n}$ as the concatenation of letters $\mathbf{i}_1 \cdots \mathbf{i}_n$, that we call a word of length n . We note that the family $(e_{i_1} \otimes \dots \otimes e_{i_n})_{(i_1, \dots, i_n) \in \{1, \dots, d\}^n}$ is a basis of $(\mathbb{R}^d)^{\otimes n}$ that can be identified with the set of words of length n defined by

$$V_n := \{\mathbf{i}_1 \cdots \mathbf{i}_n : \mathbf{i}_k \in A_d \text{ for } k = 1, 2, \dots, n\}. \quad (2.1)$$

Moreover, we denote by \emptyset the empty word and we set $V_0 = \{\emptyset\}$ which serves as a basis for $(\mathbb{R}^d)^{\otimes 0} = \mathbb{R}$. It follows that $V := \cup_{n \geq 0} V_n$ represents the standard basis of $T((\mathbb{R}^d))$. In particular, every $\ell \in T((\mathbb{R}^d))$ can be decomposed as

$$\ell = \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \mathbf{v}, \quad (2.2)$$

where $\ell^{\mathbf{v}} \in \mathbb{R}$ is the real coefficient of ℓ at coordinate \mathbf{v} . Representation (2.2) will be frequently used in the paper. We stress again that in the sequel, every blue ‘word’ $\mathbf{v} \in V$ represents an element of the canonical basis of $T((\mathbb{R}^d))$, i.e. there exists $n \geq 0$ such that \mathbf{v} is of the form $\mathbf{v} = \mathbf{i}_1 \cdots \mathbf{i}_n$, which represents the element $e_{i_1} \otimes \dots \otimes e_{i_n}$. The concatenation $\ell \mathbf{v}$ of elements $\ell \in T((\mathbb{R}^d))$ and the word $\mathbf{v} = \mathbf{i}_1 \cdots \mathbf{i}_n$ means $\ell \otimes e_{i_1} \otimes \dots \otimes e_{i_n}$. In addition to the decomposition (2.2) of elements $\ell \in T((\mathbb{R}^d))$, we introduce the projection $\ell|_{\mathbf{u}} \in T((\mathbb{R}^d))$ as

$$\ell|_{\mathbf{u}} := \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{u}\mathbf{v}} \mathbf{v} \quad (2.3)$$

for all $\mathbf{u} \in V$. The projection plays an important role in the space of iterated integrals as it is closely linked to partial differentiation, in contrast with the concatenation that relates to integration. It will be used throughout the paper.

Remark 2.1. The projection allows us to decompose elements of the extended tensor algebra $\ell \in T((\mathbb{R}^d))$ as

$$\ell = \ell^{\emptyset} \emptyset + \sum_{\mathbf{i} \in A_d} \ell|_{\mathbf{i}} \mathbf{i}.$$

For instance, for $\ell = 4 \cdot \emptyset + 3 \cdot \mathbf{1} - 1 \cdot \mathbf{12} + 2 \cdot \mathbf{2212}$, we have that $\ell^{\emptyset} = 4$, $\ell|_{\mathbf{1}} = 3 \cdot \emptyset$, $\ell|_{\mathbf{2}} = -1 \cdot \mathbf{1} + 2 \cdot \mathbf{221}$ and $\ell|_{\mathbf{3}} = 0$. ■

We define the bracket between $\ell \in T((\mathbb{R}^d))$ and $\mathbf{p} \in T((\mathbb{R}^d))$ by

$$\langle \ell, \mathbf{p} \rangle = \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \mathbf{p}^{\mathbf{v}}. \quad (2.4)$$

Notice that it is well defined as ℓ has finitely many non-zero terms. For $\ell \in T((\mathbb{R}^d))$, the series in (2.4) involves infinitely many terms and requires special care, this will be discussed in Section 3 below.

We will also consider another operation on the space of words V , the shuffle product. The shuffle product plays a crucial role for an integration by parts formula on the space of iterated integrals, see Lemma 2.5 below.

Definition 2.1 (Shuffle product). *The shuffle product $\sqcup\sqcup: V \times V \rightarrow T((\mathbb{R}^d))$ is defined inductively for all words \mathbf{v} and \mathbf{w} and all letters \mathbf{i} and \mathbf{j} in A_d by*

$$\begin{aligned} (\mathbf{vi}) \sqcup\sqcup (\mathbf{wj}) &= (\mathbf{v} \sqcup\sqcup (\mathbf{wj}))\mathbf{i} + ((\mathbf{vi}) \sqcup\sqcup \mathbf{w})\mathbf{j} \\ \mathbf{w} \sqcup\sqcup \emptyset &= \emptyset \sqcup\sqcup \mathbf{w} = \mathbf{w}. \end{aligned}$$

With some abuse of notation, the shuffle product on $T((\mathbb{R}^d))$ induced by the shuffle product on V will also be denoted by $\sqcup\sqcup$. The shuffle product is clearly commutative. See [19, 29] for more information on the shuffle product.

The shuffle product corresponds to the shuffling of two decks of cards together, while keeping the order of each single deck as illustrated on the following example: $\mathbf{1} \sqcup\sqcup \mathbf{23} = \mathbf{123} + \mathbf{213} + \mathbf{231}$.

2.2 Resolvent and linear equation

For $n \in \mathbb{N}$ and $\ell \in T((\mathbb{R}^d))$, we define the concatenation power of ℓ by

$$\ell^{\otimes n} := \overbrace{\ell \otimes \ell \otimes \cdots \otimes \ell}^{n \text{ times}},$$

with the convention $\ell^{\otimes 0} = \emptyset$.

For every $\ell \in T((\mathbb{R}^d))$ such that $|\ell^\circ| < 1$, we define its *resolvent* by

$$(\emptyset - \ell)^{-1} = \sum_{n=0}^{\infty} \ell^{\otimes n}. \quad (2.5)$$

Remark 2.2. It is clear that if the resolvent exist then, $(\emptyset - \mathbf{q})(\emptyset - \mathbf{q})^{-1} = (\emptyset - \mathbf{q})^{-1}(\emptyset - \mathbf{q}) = \emptyset$. \blacksquare

The assumption $|\ell^\circ| < 1$ ensures that the resolvent is well-defined, in the sense that every coefficient of $(\emptyset - \ell)^{-1}$ is finite, see Lemma A.1. The purpose of the resolvent is to solve linear algebraic equations as shown in the next proposition.

Proposition 2.2. *Let $\mathbf{p}, \mathbf{q} \in T((\mathbb{R}^d))$ such that $|\mathbf{q}^\circ| < 1$, then the unique solution $\ell \in T((\mathbb{R}^d))$ to the linear algebraic equation*

$$\ell = \mathbf{p} + \ell \mathbf{q} \quad (2.6)$$

is given by

$$\ell = \mathbf{p}(\emptyset - \mathbf{q})^{-1}.$$

Proof. It is easy to verify that $\ell = \mathbf{p}(\emptyset - \mathbf{q})^{-1}$ is a solution of (2.6). On the other hand, for every $\tilde{\ell} \in T((\mathbb{R}^d))$ s.t. $\tilde{\ell} = \mathbf{p} + \tilde{\ell} \mathbf{q}$, we derive $\tilde{\ell}(\emptyset - \mathbf{q}) = \mathbf{p}$. By remark 2.2 we know $\tilde{\ell} = \mathbf{p}(\emptyset - \mathbf{q})^{-1}$, verifying the uniqueness. \square

Interestingly, whenever ℓ is a linear combination of single letters, the resolvent of ℓ is equal to the shuffle exponential $e^{\sqcup \ell}$ defined by

$$e^{\sqcup \ell} := \sum_{n=0}^{\infty} \frac{\ell^{\sqcup n}}{n!} \quad (2.7)$$

with

$$\ell^{\sqcup n} := \overbrace{\ell \sqcup \ell \sqcup \cdots \sqcup \ell}^{n \text{ times}}, \quad n \geq 1, \quad \ell^{\sqcup 0} = \emptyset.$$

Proposition 2.3. *Let $\ell = \sum_{\mathbf{i} \in \mathcal{A}_d} \ell^{\mathbf{i}} \mathbf{i}$, with $\ell^{\mathbf{i}} \in \mathbb{R}$, we have that*

$$(\emptyset - \ell)^{-1} = e^{\sqcup \ell}. \quad (2.8)$$

In particular, this implies

$$e^{\sqcup \ell} = \emptyset + e^{\sqcup \ell} \ell = \emptyset + \ell e^{\sqcup \ell}. \quad (2.9)$$

Proof. We first observe that $\ell^\circ = 0$, so $(\emptyset - \ell)^{-1}$ is well-defined, thanks to Lemma A.1. we then prove that $\frac{1}{k!} \ell^{\sqcup k} = \ell^{\otimes k}$. By induction, it suffices to prove for every $k \in \mathbb{N}_+$,

$$\ell^{\otimes k-1} \sqcup \ell = k \ell^{\otimes k}. \quad (2.10)$$

It is easy to check that (2.10) is satisfied for $k = 1, 2$. Assuming it holds for $k - 1 \geq 1$, then

$$\begin{aligned} \ell^{\otimes k-1} \sqcup \ell &= \sum_{\mathbf{i} \in \mathcal{A}_d} \ell^{\otimes(k-1)} \ell^{\mathbf{i}} \mathbf{i} + \sum_{\mathbf{i} \in \mathcal{A}_d} \left(\ell^{\otimes(k-1)} \Big|_{\mathbf{i}} \sqcup \ell \right) \mathbf{i} \\ &= \ell^{\otimes k} + \sum_{\mathbf{i} \in \mathcal{A}_d} \left(\ell^{\otimes(k-2)} \sqcup \ell \right) \ell^{\mathbf{i}} \mathbf{i} \\ &= \ell^{\otimes k} + \left(\ell^{\otimes(k-2)} \sqcup \ell \right) \ell \\ &= \ell^{\otimes k} + (k-1) \ell^{\otimes(k-1)} \ell \\ &= k \ell^{\otimes k}, \end{aligned}$$

which proves (2.10) by induction, hence proving the equality (2.8). Noticing that $(\emptyset - \ell)^{-1} = \emptyset + \ell(\emptyset - \ell)^{-1} = \emptyset + (\emptyset - \ell)^{-1} \ell$, we derive (2.9). \square

2.3 Signature of a continuous semimartingale

We define the (path) signature of a semimartingale as the sequence of iterated stochastic integrals in the sense of Stratonovich. Throughout the paper, the Itô integral is denoted by $\int_0^\cdot Y_t dX_t$ and the Stratonovich integral by $\int_0^\cdot Y_t \circ dX_t$. If both X and Y are semimartingales then, $\int_0^\cdot Y_t \circ dX_t := \int_0^\cdot Y_t dX_t + \frac{1}{2}[X, Y]$.

Definition 2.4 (Signature). *Fix $T > 0$. Let $(X_t)_{t \geq 0}$ be a continuous semimartingale on \mathbb{R}^d on a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. The signature of X is defined by*

$$\begin{aligned} \mathbb{X} : \Omega \times [0, T] &\rightarrow T(\mathbb{R}^d) \\ (\omega, t) &\mapsto \mathbb{X}_t(\omega) := (1, \mathbb{X}_t^1(\omega), \dots, \mathbb{X}_t^n(\omega), \dots), \end{aligned}$$

where

$$\mathbb{X}_t^n := \int_{0 < u_1 < \dots < u_n < t} \circ dX_{u_1} \otimes \dots \otimes \circ dX_{u_n}$$

takes value in $(\mathbb{R}^d)^{\otimes n}$, $n \geq 0$. Similarly, the truncated signature of order $M \in \mathbb{N}$ is defined by

$$\begin{aligned} \mathbb{X}^{\leq M} : [0, T] &\rightarrow T^M(\mathbb{R}^d) \\ (\omega, t) &\mapsto \mathbb{X}_t^{\leq M}(\omega) := (1, \mathbb{X}_t^1(\omega), \dots, \mathbb{X}_t^M(\omega), 0, \dots, 0, \dots). \end{aligned} \tag{2.11}$$

The signature plays a similar role as polynomials on path-space. Indeed, in dimension $d = 1$, the signature of X is the sequence of monomials $\left(\frac{(X_t - X_0)^n}{n!}\right)_{n \in \mathbb{N}}$. In particular, for $\ell \in T^M(\mathbb{R})$, $\langle \ell, \mathbb{X}_t \rangle$ is a polynomial in $(X_t - X_0)$ of degree no greater than M .

Example 2.3. For $X_t = (t, W_t^1, W_t^2, \dots, W_t^d)$, where $(W_t^1, W_t^2, \dots, W_t^d)$ is a d -dimensional Brownian motion, we can compute the first level of the signature

$$\mathbb{X}_t^1 = (t, W_t^1, W_t^2, \dots, W_t^d),$$

and the second level

$$\mathbb{X}_t^2 = \begin{pmatrix} \frac{t^2}{2} & \int_0^t s dW_s^1 & \dots & \int_0^t s dW_s^d \\ \int_0^t W_s^1 ds & \frac{(W_t^1)^2}{2} & \dots & \int_0^t W_s^1 \circ dW_s^d \\ \vdots & \vdots & \ddots & \vdots \\ \int_0^t W_s^d ds & \int_0^t W_s^d \circ dW_s^1 & \dots & \frac{(W_t^d)^2}{2} \end{pmatrix}.$$

Remark 2.4. It follows from Definition 2.4 that the level n of the signature $\mathbb{X}_t^n := (\mathbb{X}_t^{\mathbf{i}_1 \dots \mathbf{i}_n})_{(\mathbf{i}_1 \dots \mathbf{i}_n) \in V_n} \in (\mathbb{R}^d)^{\otimes n}$ can be written in the following iterated form

$$\mathbb{X}_t^{\mathbf{i}_1 \dots \mathbf{i}_n} = \int_0^t \mathbb{X}_s^{\mathbf{i}_1 \dots \mathbf{i}_{n-1}} \circ dX_s^{\mathbf{i}_n}, \quad \mathbf{i}_1 \dots \mathbf{i}_n \in V_n. \tag{2.12}$$

2.4 Finite linear combinations of signature elements

In this section, we fix X a continuous semimartingale on \mathbb{R}^d and we let \mathbb{X} be its signature. We recall some important properties of finite linear combinations of signature elements $\langle \ell, \mathbb{X}_t \rangle$ for $\ell \in T(\mathbb{R}^d)$, recall (2.4).

The first one highlights a crucial linearization property of the signature, the product of two linear combinations of the signature is again a linear combination of the signature where the coefficients are given by the shuffle product, recall Definition 2.1. This property will be later extended to infinite linear combinations of signatures in Proposition 3.1 below.

Lemma 2.5 (Shuffle product property). *For $\ell, \mathbf{p} \in T(\mathbb{R}^d)$,*

$$\langle \ell, \mathbb{X}_t \rangle \langle \mathbf{p}, \mathbb{X}_t \rangle = \langle \ell \sqcup \mathbf{p}, \mathbb{X}_t \rangle.$$

Proof. This follows from the standard chain rule of Stratonovich integrals, see Gaines [19, Proposition 2.2]. \square

The shuffle product can be seen as an extension of the Cauchy product to the space of signatures. For example, set $d = 1$, $X_0 = 0$ and $\ell \in T^M(\mathbb{R})$, $\mathbf{p} \in T^N(\mathbb{R})$ for some $N, M \in \mathbb{N}$, i.e. there exists $(a_m)_{m \leq M}, (b_n)_{n \leq N} \in \mathbb{R}$ such that $\ell = \sum_{m=0}^M a_m \mathbf{1}^{\otimes m}$, $\mathbf{p} = \sum_{n=0}^N b_n \mathbf{1}^{\otimes n}$. Then,

$$\langle \ell, \mathbb{X}_t \rangle \langle \mathbf{p}, \mathbb{X}_t \rangle = \sum_{m=0}^M a_m \frac{(X_t)^m}{m!} \sum_{n=0}^N b_n \frac{(X_t)^n}{n!} = \sum_{k=0}^{M+N} c_k \frac{(X_t)^k}{k!}$$

with $c_k = \sum_{i=0}^k \binom{k}{i} a_i b_{k-i}$. In this case $\ell \sqcup \mathbf{p} = \sum_{k=0}^{M+N} c_k \mathbf{1}^{\otimes k}$.

The second property shows that any finite linear combination of signature elements is a semi-martingale with an explicit decomposition given in terms of the projection elements, recall (2.3). This semi-martingality property is no longer always true when infinite linear combinations of signature elements are considered, see Theorem 3.3 below.

Lemma 2.6. *Let $\ell \in T(\mathbb{R}^d)$, then $(\langle \ell, \mathbb{X}_t \rangle)_{t \geq 0}$ is a semimartingale with decomposition*

$$d \langle \ell, \mathbb{X}_t \rangle = \sum_{\mathbf{i} \in A_d} \langle \ell |_{\mathbf{i}}, \mathbb{X}_t \rangle \circ dX_t^{\mathbf{i}}.$$

Proof. It follows from the equivalent definition of the signature (2.12) that

$$\langle \ell |_{\mathbf{i}}, \mathbb{X}_t \rangle = \int_0^t \langle \ell |_{\mathbf{i}}, \mathbb{X}_s \rangle \circ dX_s^{\mathbf{i}}, \quad \mathbf{i} \in A_d.$$

Then, using the projection relation (2.3) and the linearity of the bracket, we get that

$$\langle \ell, \mathbb{X}_t \rangle = \left\langle \ell^\emptyset \emptyset + \sum_{\mathbf{i} \in A_d} \ell |_{\mathbf{i}}, \mathbb{X}_t \right\rangle = \ell^\emptyset + \sum_{\mathbf{i} \in A_d} \int_0^t \langle \ell |_{\mathbf{i}}, \mathbb{X}_s \rangle \circ dX_s^{\mathbf{i}}.$$

\square

Finally, we show how to transform the Stratonovich integral into the Itô integral in the context of finite linear combination of the signature.

Lemma 2.7. *Let $\ell \in T(\mathbb{R}^d)$, then for every $\mathbf{i} \in A_d$:*

$$\langle \ell, \mathbb{X}_t \rangle \circ dX_t^{\mathbf{i}} = \langle \ell, \mathbb{X}_t \rangle dX_t^{\mathbf{i}} + \frac{1}{2} \sum_{\mathbf{j} \in A_d} \langle \ell |_{\mathbf{j}}, \mathbb{X}_t \rangle d[X^{\mathbf{j}}, X^{\mathbf{i}}]_t.$$

Proof. Denote by $Y_t = \langle \ell, \mathbb{X}_t \rangle$ and fix $\mathbf{i} \in A_d$. It follows from Lemma 2.6 that Y is a semimartingale with dynamics

$$dY_t = \sum_{\mathbf{j} \in A_d} \langle \ell |_{\mathbf{j}}, \mathbb{X}_t \rangle \circ dX_t^{\mathbf{j}}.$$

Therefore, by the definition of the Stratonovich integral:

$$Y_t \circ dX_t^{\mathbf{i}} = Y_t dX_t^{\mathbf{i}} + \frac{1}{2} d[Y, X^{\mathbf{i}}]_t = Y_t dX_t^{\mathbf{i}} + \frac{1}{2} \sum_{\mathbf{j} \in A_d} \langle \ell |_{\mathbf{j}}, \mathbb{X}_t \rangle d[X^{\mathbf{j}}, X^{\mathbf{i}}]_t.$$

\square

3 Infinite linear combinations of signature elements

From now on, we fix a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$ satisfying the usual conditions, and $(X_t)_{0 \leq t \leq T}$ a continuous \mathbb{F} -semimartingale on \mathbb{R}^d and we let \mathbb{X} be its signature. In Section 2.3, we recalled well-known properties of finite linear combinations of signature elements $\langle \ell, \mathbb{X}_t \rangle$ for $\ell \in T(\mathbb{R}^d)$. In this section, we study infinite linear combinations $\langle \ell, \mathbb{X}_t \rangle$ for certain admissible $\ell \in T(\mathbb{R}^d)$ for which the infinite series will make sense. Our central results in this section are an Itô formula for such elements in Theorem 3.3 below, a control for their supremum norm in Propositions 3.5 and 3.6, and an estimate of solutions of algebraic linear equations taking values in the extended tensor algebra in Theorem 3.9.

For this, we consider the space $\mathcal{A}(\mathbb{X})$ of admissible elements ℓ using the associated stochastic semi-norm:

$$\|\ell\|_t^{\mathcal{A}(\mathbb{X})} := \sum_{n=0}^{\infty} \left| \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \mathbb{X}_t^{\mathbf{v}} \right|, \quad t \geq 0,$$

recall the definition of V_n in (2.1) and the decomposition (2.2). $\|\ell\|_t^{\mathcal{A}(\mathbb{X})}$ is actually random variables. Whenever $\|\ell\|_t^{\mathcal{A}(\mathbb{X})} < \infty$ a.s., the infinite linear combination

$$\langle \ell, \mathbb{X}_t \rangle := \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \mathbb{X}_t^{\mathbf{v}}$$

is well-defined. This leads to the following definition for the admissible set $\mathcal{A}(\mathbb{X})$:

$$\mathcal{A}(\mathbb{X}) := \left\{ \ell \in T(\mathbb{R}^d) : \|\ell\|_t^{\mathcal{A}(\mathbb{X})} < \infty \text{ for all } t \in [0, T] \text{ a.s.} \right\}.$$

For simplicity, we write \mathcal{A} (resp. $\|\cdot\|_t^{\mathcal{A}}$), in place of $\mathcal{A}(\mathbb{X})$ (resp. $\|\cdot\|_t^{\mathcal{A}(\mathbb{X})}$), when it does not cause ambiguity. Note that $T(\mathbb{R}^d) \subset \mathcal{A}$ and that $\langle \ell, \mathbb{X}_t \rangle$ extends (2.4), as the two bracket operations $\langle \cdot, \cdot \rangle$ coincide whenever $\ell \in T(\mathbb{R}^d)$.

Remark 3.1. The space \mathcal{A} plays a crucial role in this paper. Whenever we talk about $\langle \ell, \mathbb{X}_t \rangle$ for $\ell \in T(\mathbb{R}^d)$, we have to first verify that ℓ belongs to \mathcal{A} to ensure that $\langle \ell, \mathbb{X}_t \rangle$ is well defined. The set \mathcal{A} is a natural set to consider. It has also been studied in Cuchiero, Svaluto-Ferro, and Teichmann [12, Example 3.2 (iii)]. In practice, however, we will use a stronger but more tractable criterion in Section 3.3 below to obtain that a given ℓ belongs to \mathcal{A} . ■

Remark 3.2. If $\ell \in \mathcal{A}$, then $\langle \ell, \mathbb{X}_t \rangle$ is a progressively measurable process, as it is the a.s.-limit of a series of progressively measurable processes. ■

3.1 Closed operations

It is straightforward to see that the sets \mathcal{A} is closed under linear operations, i.e. for $a, b \in \mathbb{R}$ and $\ell, \mathbf{p} \in \mathcal{A}$, then $a\ell + b\mathbf{p} \in \mathcal{A}$.

Moreover, \mathcal{A} is closed under the shuffle product. The next proposition extends the shuffle property, derived in Proposition 2.5, to the case of infinite linear combination of signatures. We note that it can be obtained from a particular instance of the more general result in [12, Lemma 4.4] dealing with so-called shuffle compatible partitions. For completeness we provide a proof for our specific setting in Appendix C.

Proposition 3.1 (Extended Shuffle property). *If $\ell, \mathbf{p} \in \mathcal{A}$, then $\ell \sqcup \mathbf{p} \in \mathcal{A}$ and*

$$\langle \ell, \mathbb{X}_t \rangle \langle \mathbf{p}, \mathbb{X}_t \rangle = \langle \ell \sqcup \mathbf{p}, \mathbb{X}_t \rangle.$$

Proof. The proof is provided in Appendix C. □

3.2 Itô's formula

The aim of this subsection is the derivation of an Itô's formula in Theorem 3.3 below for infinite linear combinations of signature elements. In preparation, we define $\mathcal{Q}_i(X)$, the set of coordinates of X that have a non-zero quadratic covariation with $X^{\mathbf{i}}$:

$$\mathcal{Q}_i(X) := \left\{ \mathbf{j} \in A_d : [X^{\mathbf{j}}, X^{\mathbf{i}}]_t \neq 0, \text{ on a set of non-zero } dt \otimes d\mathbb{P} \text{ measure} \right\}, \quad \mathbf{i} \in A_d.$$

Remark 3.3. $\mathcal{Q}_i(X)$ is introduced in order to get rid of unnecessary assumptions in Theorem 3.2, since we only need $\ell_{|\mathbf{j}} \in \mathcal{A}$ for those $\mathbf{j} \in \mathcal{Q}_i(X)$. For example, in the particular case of $X_t = (t, W_t)$, with W a one-dimensional Brownian motion, assumptions are only needed for the terms $\ell, \ell_{|1}, \ell_{|2}, \ell_{|22} \in \mathcal{A}$, but not for $\ell_{|21}$ and $\ell_{|12}$. \blacksquare

Lemma 3.2. Fix $\mathbf{i} \in A_d$ and let $\ell \in \mathcal{A}$ be such that

$$\ell_{|\mathbf{i}} \in \mathcal{A}, \quad \int_0^T \left(\|\ell\|_t^{\mathcal{A}} \right)^2 d[X^{\mathbf{i}}, X^{\mathbf{i}}]_t < \infty \text{ and } \int_0^T \|\ell\|_t^{\mathcal{A}} d|F_t^{\mathbf{i}}| < \infty, \quad (3.1)$$

where $F^{\mathbf{i}}$ is the finite variation part of $X^{\mathbf{i}}$. We also assume that for every $\mathbf{j} \in \mathcal{Q}_i(X)$,

$$\ell_{|\mathbf{j}} \in \mathcal{A} \quad \text{and} \quad \int_0^T \|\ell_{|\mathbf{j}}\|_t^{\mathcal{A}} d[X^{\mathbf{j}}, X^{\mathbf{i}}]_t < \infty. \quad (3.2)$$

Then,

$$\langle \ell_{|\mathbf{i}}, \mathbb{X}_t \rangle = \int_0^t \langle \ell, \mathbb{X}_s \rangle dX_s^{\mathbf{i}} + \frac{1}{2} \sum_{\mathbf{j} \in A_d} \int_0^t \langle \ell_{|\mathbf{j}}, \mathbb{X}_s \rangle d[X^{\mathbf{j}}, X^{\mathbf{i}}]_s, \quad a.s. \quad (3.3)$$

for every $t \in [0, T]$.

Proof. Fix $t \in [0, T]$. Clearly, under the assumptions (3.1)-(3.2) on ℓ , the right side of equation (3.3) is well defined. For every word $\mathbf{v} \in V$, we can write it as $\mathbf{v} = \sum_{\mathbf{j} \in A_d} \mathbf{v}_{|\mathbf{j}}$. By Lemma 2.7, we obtain:

$$\langle \mathbf{v}_{|\mathbf{i}}, \mathbb{X}_t \rangle = \int_0^t \langle \mathbf{v}, \mathbb{X}_s \rangle \circ dX_s^{\mathbf{i}} = \frac{1}{2} \sum_{\mathbf{j} \in A_d} \int_0^t \langle \mathbf{v}_{|\mathbf{j}}, \mathbb{X}_s \rangle d[X^{\mathbf{j}}, X^{\mathbf{i}}]_s + \int_0^t \langle \mathbf{v}, \mathbb{X}_s \rangle dX_s^{\mathbf{i}}.$$

Combined with the decomposition (2.2), it yields that

$$\langle \ell_{|\mathbf{i}}, \mathbb{X}_t \rangle = \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \langle \mathbf{v}_{|\mathbf{i}}, \mathbb{X}_t \rangle = \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \int_0^t \left(\frac{1}{2} \sum_{\mathbf{j} \in A_d} \langle \mathbf{v}_{|\mathbf{j}}, \mathbb{X}_s \rangle d[X^{\mathbf{j}}, X^{\mathbf{i}}]_s + \langle \mathbf{v}, \mathbb{X}_s \rangle dX_s^{\mathbf{i}} \right). \quad (3.4)$$

It remains to argue, using dominated convergence, that both terms appearing on the right-hand side of (3.4) correspond to the terms in (3.3).

- (i) For the first term, if $\mathbf{j} \notin \mathcal{Q}_i(X)$, $[X^{\mathbf{j}}, X^{\mathbf{i}}]_t = 0$ for all t , so we only care about the letters $\mathbf{j} \in \mathcal{Q}_i(X)$. Since $\ell_{|\mathbf{j}} = \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \mathbf{v}_{|\mathbf{j}} \in \mathcal{A}$ by assumption (3.2), we can apply the dominated convergence theorem, with $\|\ell_{|\mathbf{j}}\|_t^{\mathcal{A}}$ as the dominated function, and get

$$\begin{aligned} \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \int_0^t \langle \mathbf{v}_{|\mathbf{j}}, \mathbb{X}_s \rangle d[X^{\mathbf{j}}, X^{\mathbf{i}}]_s &= \int_0^t \left(\sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \langle \mathbf{v}_{|\mathbf{j}}, \mathbb{X}_s \rangle \right) d[X^{\mathbf{j}}, X^{\mathbf{i}}]_s \\ &= \int_0^t \langle \ell_{|\mathbf{j}}, \mathbb{X}_s \rangle d[X^{\mathbf{j}}, X^{\mathbf{i}}]_s \end{aligned} \quad (3.5)$$

- (ii) For the second term, let us define

$$Y_t^N := \sum_{n=0}^N \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \int_0^t \langle \mathbf{v}, \mathbb{X}_s \rangle dX_s^{\mathbf{i}}.$$

On the one hand, as $\int_0^T \left(\|\ell\|_t^{\mathcal{A}} \right)^2 d[X^{\mathbf{i}}, X^{\mathbf{i}}]_t < \infty$ and $\sum_{n=0}^N \sum_{\mathbf{v} \in V_n} \ell^{\mathbf{v}} \langle \mathbf{v}, \mathbb{X}_t \rangle \leq \|\ell\|_t^{\mathcal{A}}$ for every N and assumption (3.1), by the dominated convergence theorem for stochastic integrals, the random variable Y^N converges in probability to

$$Y_t^\infty := \int_0^t \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^\mathbf{v} \langle \mathbf{v}, \mathbb{X}_s \rangle dX_s^\mathbf{i}$$

as $N \rightarrow \infty$.

On the other hand, since $\ell \mathbf{i} \in \mathcal{A}$, the left-hand side of (3.4) is a.s. finite, and since second term of its right-hand side has been proved to be finite, it follows that $\sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^\mathbf{v} \int_0^t \langle \mathbf{v}, \mathbb{X}_s \rangle dX_s^\mathbf{i}$ exists a.s. and corresponds to an almost sure limit of Y^N as $N \rightarrow \infty$. Which yields

$$\sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \ell^\mathbf{v} \int_0^t \langle \mathbf{v}, \mathbb{X}_s \rangle dX_s^\mathbf{i} = Y_t^\infty = \int_0^t \langle \ell, \mathbb{X}_s \rangle dX_s^\mathbf{i} \quad a.s. \quad (3.6)$$

Combining (3.4), (3.5) and (3.6) yields (3.3) and completes the proof. \square

Remark 3.4. To obtain that equality (3.3) holds for every t almost surely, one could argue that $\langle \ell \mathbf{i}, \mathbb{X}_t \rangle$ is a continuous process for instance. This requires additional conditions on ℓ , see for instance Proposition 3.6-(iii) below. \blacksquare

We are now in place to state the main result of this section which is an Itô's formula for infinite linear combinations of signature elements. For this, we need to define the following set

$$\mathcal{I}(X) := \left\{ \ell \in \mathcal{A} : \int_0^T \left(\|\ell_{|\mathbf{i}}\|_t^{\mathcal{A}} d|F_t^\mathbf{i}| + \|\ell_{|\mathbf{j}}\|_t^{\mathcal{A}} d[X^\mathbf{j}, X^\mathbf{i}]_t + \left(\|\ell_{|\mathbf{i}}\|_t^{\mathcal{A}} \right)^2 d[X^\mathbf{i}, X^\mathbf{i}]_t \right) < \infty, \text{ a.s.} \right\},$$

which can be seen as the analog of the space $C^2(\mathbb{R})$ when one applies Itô's formula on $f(W)$, see Example 3.5 below.

Theorem 3.3 (Itô's formula). *Let $\ell \in \mathcal{I}(X)$, then $(\langle \ell, \mathbb{X}_t \rangle)_{t \leq T}$ is an Itô process such that*

$$\langle \ell, \mathbb{X}_t \rangle = \ell^\emptyset + \sum_{\mathbf{i} \in A_d} \int_0^t \langle \ell_{|\mathbf{i}}, \mathbb{X}_s \rangle dX_s^\mathbf{i} + \frac{1}{2} \sum_{\mathbf{i} \in A_d} \sum_{\mathbf{j} \in A_d} \int_0^t \langle \ell_{|\mathbf{j}}, \mathbb{X}_s \rangle d[X^\mathbf{j}, X^\mathbf{i}]_s, \quad (3.7)$$

for all $t \leq T$. In particular, when we take X_t to be $\widehat{W}_t = (t, W_t^2, \dots, W_t^d)$, with W a $(d-1)$ -Brownian motion then, the set \mathcal{I} reads

$$\mathcal{I}(\widehat{W}) = \left\{ \ell \in \mathcal{A} : \ell_{|\mathbf{1}}, \ell_{|\mathbf{j}}, \ell_{|\mathbf{j}\mathbf{j}} \in \mathcal{A}, \mathbf{j} \in \{2, \dots, d\}, \text{ and } \int_0^T \|\ell_{|\mathbf{1}}\|_t^{\mathcal{A}} dt + \sum_{\mathbf{j} \in \{2, \dots, d\}} \int_0^T \left(\|\ell_{|\mathbf{j}\mathbf{j}}\|_t^{\mathcal{A}} + \left(\|\ell_{|\mathbf{j}}\|_t^{\mathcal{A}} \right)^2 \right) dt < \infty, \text{ a.s.} \right\}, \quad (3.8)$$

and for every $\ell \in \mathcal{I}(\widehat{W})$:

$$\langle \ell, \widehat{W}_t \rangle = \ell^\emptyset + \int_0^t \left\langle \ell_{|\mathbf{1}} + \frac{1}{2} \sum_{\mathbf{j} \in \{2, \dots, d\}} \ell_{|\mathbf{j}\mathbf{j}}, \widehat{W}_s \right\rangle ds + \sum_{\mathbf{j} \in \{2, \dots, d\}} \int_0^t \langle \ell_{|\mathbf{j}}, \widehat{W}_s \rangle dW_s^\mathbf{j}. \quad (3.9)$$

Proof. The proof follows from an application of Lemma 3.2 combined with the decomposition (2.3). \square

The following example highlights that Theorem 3.3 is indeed an extension of the usual Itô's formula on $f(W)$.

Example 3.5. Let $d = 2$ and $\widehat{W}_t = (t, W_t)$, and fix an analytic function $f(y) := \sum_{n \geq 0} a_n y^n$ with infinite radius of convergence which we apply to W . It is clear that

$$f(W_t) = \sum_{n \geq 0} a_n W_t^n = \langle \ell, \widehat{W}_t \rangle,$$

where

$$\ell = \sum_{n \geq 0} a_n \mathbf{2}^{\llcorner n} = \sum_{n \geq 0} a_n n! \mathbf{2}^{\otimes n}.$$

In particular, the projections read:

$$\begin{aligned}\ell|_1 &= 0, \\ \ell|_2 &= \sum_{n \geq 0} a_{n+1}(n+1)! 2^{\otimes n} = \sum_{n \geq 0} a_{n+1}(n+1) 2^{\perp\perp n}, \\ \ell|_{22} &= \sum_{n \geq 0} a_{n+2}(n+2)! 2^{\otimes n} = \sum_{n \geq 0} a_{n+2}(n+1)(n+2) 2^{\perp\perp n}.\end{aligned}$$

It is easy to verify that $\ell, \ell|_2, \ell|_{22} \in \mathcal{A}$ since f has infinite radius of convergence. We can thus further derive that

$$f'(W_t) = \langle \ell|_2, \widehat{W}_t \rangle, \quad f''(W_t) = \langle \ell|_{22}, \widehat{W}_t \rangle.$$

On the other hand we can see that $\|\ell|_2\|_t^A = \sum_{n=0}^{\infty} (n+1)|a_{n+1}| \cdot |W_t|^n$, and since $g(x) := \sum_{n \geq 0} (n+1)|a_{n+1}|x^n$ is also analytic and $\|\ell|_2\|_t^A$ has continuous sample path almost surely, then

$$\sup_{t \in [0, T]} \|\ell|_2\|_t^A < \infty.$$

With similar arguments we can show that $\sup_{t \in [0, T]} \|\ell|_{22}\|_t^A < \infty$. This allows us to verify that $\ell \in \mathcal{I}(\widehat{W})$. An application of our Itô's formula (3.9) yields

$$\langle \ell, \widehat{W}_t \rangle = \ell^\circ + \int_0^t \langle \ell|_2, \widehat{W}_s \rangle dW_s + \frac{1}{2} \int_0^t \langle \ell|_{22}, \widehat{W}_s \rangle ds,$$

which translates into

$$f(W_t) = f(0) + \int_0^t f'(W_s) dW_s + \frac{1}{2} \int_0^t f''(W_s) ds,$$

the standard Itô's formula. ■

Remark 3.6. It should be noted that the identity (3.7) only holds almost surely for every fixed t . However, as remark 3.4, if we additionally assume $\langle \ell, \mathbb{X}_t \rangle$ is continuous, then (3.7) holds for every $t \in [0, T]$ almost surely. In particular, when $X_t = \widehat{W}_t = (t, W_t^1, \dots, W_t^d)$, and $\ell \in \mathcal{A}_h$ (where \mathcal{A}_h is introduced in Section 3.3 below), the continuity assumption is satisfied by Proposition (3.6)-(iii). Our applications of Theorem 3.3 in later section are all in this strong sense. ■

We can easily extend Theorem 3.3 to time-dependent linear functionals, i.e. $\ell : [0, T] \rightarrow T((\mathbb{R}^d))$. For this, let us define

$$\mathcal{I}'(X) := \left\{ \ell : [0, T] \rightarrow \mathcal{A} : \begin{array}{l} \text{for every } \mathbf{v} \in V, \ell^{\mathbf{v}} \in C^1([0, T], \mathbb{R}) \text{ and} \\ \text{for every } \mathbf{i} \in A_d, \mathbf{j} \in Q_i(X) \text{ and for all } t \in [0, T], \ell_t|_{\mathbf{i}}, \ell_t|_{\mathbf{j}}, \dot{\ell}_t \in \mathcal{A} \text{ and} \\ \int_0^T \left(\|\ell_t|_{\mathbf{i}}\|_t^A d|F_t^{\mathbf{i}}| + \|\ell_t|_{\mathbf{j}}\|_t^A d[X^{\mathbf{j}}, X^{\mathbf{i}}]_t + \left(\|\ell_t|_{\mathbf{i}}\|_t^A \right)^2 d[X^{\mathbf{i}}, X^{\mathbf{i}}]_t \right) < \infty \text{ a.s.} \end{array} \right\},$$

where $\dot{\ell}_t = \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \frac{d}{dt} \ell_t^{\mathbf{v}} \mathbf{v}$.

Corollary 3.4. *Let $\ell \in \mathcal{I}'(X)$, then $(\langle \ell_t, \mathbb{X}_t \rangle)_{t \leq T}$ is an Itô process such that*

$$\langle \ell_t, \mathbb{X}_t \rangle = \ell_t^\circ + \sum_{\mathbf{i} \in A_d} \int_0^t \langle \ell_s|_{\mathbf{i}}, \mathbb{X}_s \rangle dX_s^{\mathbf{i}} + \frac{1}{2} \sum_{\mathbf{j} \in A_d} \sum_{\mathbf{i} \in A_d} \int_0^t \langle \ell_s|_{\mathbf{j}}, \mathbb{X}_s \rangle d[X^{\mathbf{j}}, X^{\mathbf{i}}]_s + \int_0^t \langle \dot{\ell}_s, \mathbb{X}_s \rangle ds,$$

for all $t \leq T$.

Proof. The proof is similar to the proof of Theorem 3.3. □

Remark 3.7. The example of the fractional Riemann-Liouville Brownian motion treated in Section 4.3 shows that, contrary to the finite case, not all infinite linear combinations of signature elements of \widehat{W} are semimartingales. In particular, Theorem 3.3 and Corollary 3.4 cannot be applied in such cases. ■

3.3 The space \mathcal{A}_h and moment estimates

In practice, it is not an easy task to verify directly whether a given ℓ belongs to \mathcal{A} , even for the case of the time extended $(d-1)$ -dimensional Brownian motion $X_t = \widehat{W}_t = (t, W_t^2, \dots, W_t^d)$. The aim of this subsection is to introduce a subset \mathcal{A}_h of \mathcal{A} , specific to such \widehat{W} with $d \geq 2$, by essentially controlling the expectation of its signature $\widehat{\mathbb{W}}$. We will give a tractable criterion to prove that a given ℓ belongs to \mathcal{A}_h and thus to \mathcal{A} .

In order to motivate the set \mathcal{A}_h , we will derive an upper bound for the moments of the signature elements of $\widehat{\mathbb{W}}$ with the help of Fawcett's formula [14] using a deterministic function h . For that and to ease the notations, for every $\mathbf{v} \in V$ we let $n(\mathbf{v})$ denote length of \mathbf{v} , i.e. $\mathbf{v} \in V_{n(\mathbf{v})}$, and $x(\mathbf{v})$ the number of occurrences of the letter **1** in \mathbf{v} . In the following, if it does not cause ambiguity, we will write n and x in place of $n(\mathbf{v})$ and $x(\mathbf{v})$ respectively for brevity.

Example 3.8. Let $\mathbf{v} = \mathbf{111321}$, then $n = 6$ and $x = 4$. ■

We first provide a tight bound for the second moment:

Proposition 3.5. *Fix $t \in [0, T]$ and $\mathbf{v} \in V$. We have*

$$(i) \quad \mathbb{E} \left| \left\langle \mathbf{v}, \widehat{\mathbb{W}}_t \right\rangle \right|^2 \leq \binom{2n}{n} \frac{t^{n+x}}{(n+x)!} 2^{x-n}, \quad (3.10)$$

(ii) *and there exists a non-negative constant $C \geq 2$ independent of \mathbf{v} and t , such that*

$$\mathbb{E} \left[\sup_{s \in [0, t]} \left| \left\langle \mathbf{v}, \widehat{\mathbb{W}}_s \right\rangle \right|^2 \right] \leq h_t(\mathbf{v})^2, \quad (3.11)$$

where

$$h_t(\mathbf{v}) := \frac{C}{(n+1)^{\frac{1}{4}}} \frac{(2t)^{\frac{1}{2}(n+x)}}{\sqrt{(n+x-1)!}},$$

with the convention that $n! = 1$ for $n \leq 0$. In particular, h satisfies the following sub-multiplicative property

$$h_t(\mathbf{uv}) \leq h_t(\mathbf{u})h_t(\mathbf{v}), \quad \mathbf{u}, \mathbf{v} \in V. \quad (3.12)$$

Proof. The proof is given in Appendix B. □

Based on Proposition 3.5, an application of Jensen's inequality yields

$$\mathbb{E} \left[\sup_{s \in [0, t]} \left| \left\langle \mathbf{v}, \widehat{\mathbb{W}}_s \right\rangle \right| \right] \leq h_t(\mathbf{v}). \quad (3.13)$$

We are now ready to introduce a norm on $T((\mathbb{R}^d))$ using the function h :

$$\|\ell\|_t^{\mathcal{A}_h} := \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} |\ell^{\mathbf{v}}| h_t(\mathbf{v}), \quad \ell \in T((\mathbb{R}^d)).$$

Unlike $\|\cdot\|^{\mathcal{A}}$, $\|\cdot\|^{\mathcal{A}_h}$ is deterministic, thus simplifying the computation. Moreover, notice that $\|\cdot\|_t^{\mathcal{A}_h}$ is an increasing function of t , which naturally leads to the definition of the normed vector space:

$$\mathcal{A}_h := \left\{ \ell \in T((\mathbb{R}^d)) : \|\ell\|_T^{\mathcal{A}_h} < \infty \right\}.$$

Proposition 3.6. *Fix $\ell \in \mathcal{A}_h$. The following properties hold:*

(i) $\mathcal{A}_h \subset \mathcal{A}$, and

$$\mathbb{E} \left[\sup_{s \in [0, t]} \|\ell\|_s^{\mathcal{A}} \right] \leq \|\ell\|_t^{\mathcal{A}_h}, \quad t \leq T.$$

(ii) \mathcal{A}_h induces a natural upper bound

$$\mathbb{E} \left[\sup_{s \in [0, t]} \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_s \rangle^2 \right] \leq \left(\|\boldsymbol{\ell}\|_t^{\mathcal{A}_h} \right)^2, \quad t \leq T.$$

(iii) $t \mapsto \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \rangle$ is continuous and the strong error of convergence of the truncated linear combination of the signature elements satisfies

$$\mathbb{E} \left[\sup_{t \in [0, T]} \left| \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t^{\leq M} \rangle - \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \rangle \right| \right] \leq \sum_{n=M}^{\infty} \sum_{\mathbf{v} \in V_n} |\boldsymbol{\ell}^{\mathbf{v}}| h_T(\mathbf{v}), \quad M \in \mathbb{N}. \quad (3.14)$$

Proof. (i) For all $\boldsymbol{\ell} \in \mathcal{A}_h$, the inequality (3.13) yields

$$\mathbb{E} \left[\sup_{s \in [0, t]} \|\boldsymbol{\ell}\|_s^{\mathcal{A}} \right] \leq \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} |\boldsymbol{\ell}^{\mathbf{v}}| \mathbb{E} \left[\sup_{s \in [0, t]} \left| \langle \mathbf{v}, \widehat{\mathbb{W}}_s \rangle \right| \right] \leq \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} |\boldsymbol{\ell}^{\mathbf{v}}| h_t(\mathbf{v}) \leq \|\boldsymbol{\ell}\|_t^{\mathcal{A}_h} < \infty.$$

Consequently, $\sup_{t \in [0, T]} \|\boldsymbol{\ell}\|_t^{\mathcal{A}} < \infty$ a.s. and thus $\boldsymbol{\ell} \in \mathcal{A}(\widehat{\mathbb{W}})$.

(ii) Starting with

$$\langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_s \rangle^2 = \left(\sum_{\mathbf{v} \in V} \boldsymbol{\ell}^{\mathbf{v}} \langle \mathbf{v}, \widehat{\mathbb{W}}_s \rangle \right)^2 \leq \sum_{\mathbf{u} \in V} \sum_{\mathbf{w} \in V} |\boldsymbol{\ell}^{\mathbf{u}} \boldsymbol{\ell}^{\mathbf{w}}| \cdot \left| \langle \mathbf{u}, \widehat{\mathbb{W}}_s \rangle \langle \mathbf{w}, \widehat{\mathbb{W}}_s \rangle \right|,$$

we get that

$$\sup_{s \in [0, t]} \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_s \rangle^2 \leq \sum_{\mathbf{u} \in V} \sum_{\mathbf{w} \in V} |\boldsymbol{\ell}^{\mathbf{u}} \boldsymbol{\ell}^{\mathbf{w}}| \cdot \sup_{s \in [0, t]} \left| \langle \mathbf{u}, \widehat{\mathbb{W}}_s \rangle \right| \cdot \sup_{s \in [0, t]} \left| \langle \mathbf{w}, \widehat{\mathbb{W}}_s \rangle \right|.$$

Taking the expectation and applying Cauchy-Schwarz inequality yields that

$$\begin{aligned} \mathbb{E} \left[\sup_{s \in [0, t]} \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_s \rangle^2 \right] &\leq \sum_{\mathbf{u} \in V} \sum_{\mathbf{w} \in V} |\boldsymbol{\ell}^{\mathbf{u}} \boldsymbol{\ell}^{\mathbf{w}}| \sqrt{\mathbb{E} \sup_{s \in [0, t]} \langle \mathbf{u}, \widehat{\mathbb{W}}_s \rangle^2} \sqrt{\mathbb{E} \sup_{s \in [0, t]} \langle \mathbf{w}, \widehat{\mathbb{W}}_s \rangle^2} \\ &= \left(\sum_{\mathbf{v} \in V} |\boldsymbol{\ell}^{\mathbf{v}}| \sqrt{\mathbb{E} \sup_{s \in [0, t]} \langle \mathbf{v}, \widehat{\mathbb{W}}_s \rangle^2} \right)^2 \\ &\leq \left(\sum_{\mathbf{v} \in V} |\boldsymbol{\ell}^{\mathbf{v}}| h_t(\mathbf{v}) \right)^2 \\ &= \left(\|\boldsymbol{\ell}\|_t^{\mathcal{A}_h} \right)^2. \end{aligned}$$

(iii) Define

$$X_t := \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \rangle = \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \boldsymbol{\ell}^{\mathbf{v}} \widehat{\mathbb{W}}_t^{\mathbf{v}} \quad \text{and} \quad X_t^M := \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t^{\leq M} \rangle = \sum_{n=0}^M \sum_{\mathbf{v} \in V_n} \boldsymbol{\ell}^{\mathbf{v}} \widehat{\mathbb{W}}_t^{\mathbf{v}}.$$

It follows from inequality (3.13) that

$$\mathbb{E} \left[\sup_{t \in [0, T]} |X_t - X_t^M| \right] \leq \sum_{n=M+1}^{\infty} \sum_{\mathbf{v} \in V_n} |\boldsymbol{\ell}^{\mathbf{v}}| \mathbb{E} \left[\sup_{t \in [0, T]} \langle \mathbf{v}, \widehat{\mathbb{W}}_t \rangle \right] \leq \sum_{n=M+1}^{\infty} \sum_{\mathbf{v} \in V_n} |\boldsymbol{\ell}^{\mathbf{v}}| h_T(\mathbf{v}).$$

Moreover, $\boldsymbol{\ell} \in \mathcal{A}_h$ implies that

$$\lim_{M \rightarrow \infty} \sum_{n=M+1}^{\infty} \sum_{\mathbf{v} \in V_n} |\boldsymbol{\ell}^{\mathbf{v}}| h_T(\mathbf{v}) = 0.$$

Thus

$$\mathbb{E} \left[\sup_{t \in [0, T]} |X_t - X_t^M| \right] \rightarrow 0,$$

implying that

$$\sup_{t \in [0, T]} |X_t - X_t^M| \rightarrow 0 \quad \text{in probability as } M \rightarrow \infty.$$

So there exists a sub-sequence $(M_k)_{k \in \mathbb{N}}$ such that $\sup_{t \in [0, T]} |X_t - X_t^{M_k}| \rightarrow 0$ a.s. Moreover, as X^{M_k} is continuous, X is also continuous. □

Derived from (3.12), the next proposition shows that $\|\cdot\|^{\mathcal{A}_h}$ is sub-multiplicative for the tensor product.

Proposition 3.7. *For every $\mathbf{p}, \ell \in T((\mathbb{R}^d))$,*

$$\|\mathbf{p}\ell\|_t^{\mathcal{A}_h} \leq \|\mathbf{p}\|_t^{\mathcal{A}_h} \|\ell\|_t^{\mathcal{A}_h}.$$

In particular, if $\mathbf{p}, \ell \in \mathcal{A}_h$, then $\mathbf{p}\ell \in \mathcal{A}_h$.

Proof. First notice that for all $k \in \mathbb{N}$ and $\mathbf{w} \in V_k$,

$$\langle \mathbf{p}\ell, \mathbf{w} \rangle = \sum_{n=0}^k \sum_{\mathbf{u} \in V_n} \sum_{\mathbf{v} \in V_{k-n}} \mathbf{p}^{\mathbf{u}} \ell^{\mathbf{v}} \langle \mathbf{u}\mathbf{v}, \mathbf{w} \rangle,$$

so that

$$\begin{aligned} \|\mathbf{p}\ell\|_t^{\mathcal{A}_h} &= \sum_{k=0}^{\infty} \sum_{\mathbf{w} \in V_k} |\langle \mathbf{p}\ell, \mathbf{w} \rangle| h_t(\mathbf{w}) \\ &= \sum_{k=0}^{\infty} \sum_{\mathbf{w} \in V_k} \left| \sum_{n=0}^k \sum_{\mathbf{u} \in V_n} \sum_{\mathbf{v} \in V_{k-n}} \mathbf{p}^{\mathbf{u}} \ell^{\mathbf{v}} \langle \mathbf{u}\mathbf{v}, \mathbf{w} \rangle \right| h_t(\mathbf{w}) \\ &\leq \sum_{k=0}^{\infty} \sum_{n=0}^k \sum_{\mathbf{u} \in V_n} \sum_{\mathbf{v} \in V_{k-n}} \sum_{\mathbf{w} \in V_k} |\mathbf{p}^{\mathbf{u}} \ell^{\mathbf{v}}| \langle \mathbf{u}\mathbf{v}, \mathbf{w} \rangle h_t(\mathbf{w}) \end{aligned}$$

Then, using (3.12)

$$\begin{aligned} \|\mathbf{p}\ell\|_t^{\mathcal{A}_h} &\leq \sum_{k=0}^{\infty} \sum_{n=0}^k \sum_{\mathbf{u} \in V_n} \sum_{\mathbf{v} \in V_{k-n}} |\mathbf{p}^{\mathbf{u}} \ell^{\mathbf{v}}| h_t(\mathbf{u}) h_t(\mathbf{v}) \sum_{\mathbf{w} \in V_k} \langle \mathbf{u}\mathbf{v}, \mathbf{w} \rangle \\ &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \sum_{\mathbf{u} \in V_n} \sum_{\mathbf{v} \in V_m} |\mathbf{p}^{\mathbf{u}} \ell^{\mathbf{v}}| h_t(\mathbf{u}) h_t(\mathbf{v}) \\ &= \|\mathbf{p}\|_t^{\mathcal{A}_h} \|\ell\|_t^{\mathcal{A}_h}. \end{aligned}$$

□

Remark 3.9. Compared to the literature, our estimate for the supremum norm (3.11) appears to be used for the first time and plays a crucial role in our convergence analysis. The closest related results appeared in Ben Arous [7] where moment estimates of $\mathbb{E}|\langle \mathbf{v}, \widehat{\mathbb{W}}_t \rangle|^2$ and certain convergence criterion for infinite series of signatures of $\widehat{\mathbb{W}}$, i.e. iterated sequence of Stratonovich integrals, are established. The estimate in [7, Lemma 2] has the same order as (3.10) given here, although the constant there is not as sharp as ours (since we make use of Fawcett's formula) but has the same order of magnitude (see the remark following [7, Lemma 3]). We also provide estimates for $\mathbb{E}|\langle \mathbf{v}, \widehat{\mathbb{W}}_t \rangle|^p$ for every $p \in \mathbb{N}_+$, see proposition B.1. But most importantly, we provide an estimate for $\mathbb{E}[\sup_{s \in [0, t]} |\langle \mathbf{v}, \widehat{\mathbb{W}}_s \rangle|^2]$ in (3.11), which is necessary to prove convergence for all t almost surely instead of simply for a fixed t , when compared to the convergence criterion in [7, Theorem 2 (1)]. Additionally, the second criteria [7, Theorem 2 (2)] requires the second order norm to be bounded, we only require the first order norm $\|\cdot\|_t^{\mathcal{A}_h}$ to be bounded, which is hence milder. ■

3.4 Back to the linear algebraic equation

We now prove that the solution ℓ to the linear algebraic equation (2.6) belongs to \mathcal{A}_h under suitable conditions on \mathbf{p} and \mathbf{q} , see Theorem 3.9. This is a crucial ingredient in obtaining convergence of our representations formula $\langle \ell, \widehat{\mathbb{W}}_t \rangle$ in Section 4. For this, we start by introducing the following partial order:

Definition 3.8. Let $\mathbf{p}, \mathbf{q} \in T((\mathbb{R}^d))$. We say that \mathbf{q} is dominated by \mathbf{p} if for all $\mathbf{v} \in V$, $|\mathbf{q}^{\mathbf{v}}| \leq \mathbf{p}^{\mathbf{v}}$. We then write

$$\mathbf{q} \preceq \mathbf{p}.$$

Theorem 3.9. Let $\mathbf{p} \in \mathcal{A}_h$ and $\mathbf{q} \in T((\mathbb{R}^d))$ such that $\mathbf{q}^{\emptyset} = 0$. Assume that either

$$\mathbf{q} \preceq \mathbf{a}e^{\sqcup \mathbf{b}} \quad \text{or} \quad \mathbf{q} \preceq e^{\sqcup \mathbf{b}}\mathbf{a}, \quad (3.15)$$

for some $\mathbf{a} = \sum_{\mathbf{i} \in A_d} \mathbf{a}^{\mathbf{i}} \mathbf{i}$ and $\mathbf{b} = \sum_{\mathbf{i} \in A_d} \mathbf{b}^{\mathbf{i}} \mathbf{i}$ with $\mathbf{a}^{\mathbf{i}}, \mathbf{b}^{\mathbf{i}} \in \mathbb{R}$, then the resolvent $(\emptyset - \mathbf{q})^{-1}$ is well defined and belongs to \mathcal{A}_h . Furthermore, the unique solution $\ell \in T((\mathbb{R}^d))$ to the linear equation

$$\ell = \mathbf{p} + \ell \mathbf{q}$$

belongs to \mathcal{A}_h .

The rest of the section is dedicated to the proof of Theorem 3.9.

First, we prove that $e^{\sqcup \mathbf{b}} \in \mathcal{A}_h$.

Lemma 3.10. Let $\mathbf{b} = \sum_{\mathbf{i} \in A_d} \mathbf{b}^{\mathbf{i}} \mathbf{i}$ with $\mathbf{b}^{\mathbf{i}} \in \mathbb{R}$, then

$$e^{\sqcup \mathbf{b}} = (\emptyset - \mathbf{b})^{-1} \in \mathcal{A}_h.$$

Proof. First, the equality is ensured by Proposition 2.3, so it suffices to prove $(\emptyset - \mathbf{b})^{-1} \in \mathcal{A}_h$. For simplicity, we provide a proof for the case where $d = 2$, the extension to the general case follows along the same lines. For a word $\mathbf{v} \in V$, we recall the notations $n(\mathbf{v})$ and $x(\mathbf{v})$ that denote the length of \mathbf{v} and the number of $\mathbf{1}$ in \mathbf{v} respectively. We also recall that $h_T(\mathbf{v})$ depends on \mathbf{v} only through $n(\mathbf{v})$ and $x(\mathbf{v})$, meaning that for any $\mathbf{u} \in V$, $n(\mathbf{v}) = n(\mathbf{u})$ and $x(\mathbf{v}) = x(\mathbf{u})$ imply $h_T(\mathbf{v}) = h_T(\mathbf{u})$. Therefore, assuming $\mathbf{b} = a\mathbf{1} + b\mathbf{2}$ and $\ell = (\emptyset - (a\mathbf{1} + b\mathbf{2}))^{-1}$, we can decompose $\|\ell\|_T^{\mathcal{A}_h}$ as:

$$\begin{aligned} \|\ell\|_T^{\mathcal{A}_h} &= \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} |\ell^{\mathbf{v}}| h_T(\mathbf{v}) \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} |a|^k |b|^{n-k} h_T(\mathbf{1}^{\otimes k} \mathbf{2}^{\otimes n-k}) \\ &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+m}{k} |a|^k |b|^m h_T(\mathbf{1}^{\otimes k} \mathbf{2}^{\otimes m}) \\ &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{(k+m)!}{k! m!} |a|^k |b|^m \frac{C}{(k+m+1)^{\frac{1}{4}}} \frac{(2T)^{k+\frac{m}{2}}}{\sqrt{(2k+m-1)!}} \\ &= C \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{|2Ta|^k}{k!} \frac{|\sqrt{2T}b|^m}{\sqrt{m!}} (k+m+1)^{-\frac{1}{4}} \sqrt{\frac{(k+m)!^2}{m!(2k+m-1)!}}. \end{aligned}$$

After remarking that $(k+m+1)^{-\frac{1}{4}} \leq 1$ for all $k, m \in \mathbb{N}$, we can shift our attention to the square root on the right-hand side:

$$\frac{(k+m)!^2}{m!(2k+m-1)!} = \frac{\prod_{j=1}^k (m+j)}{\prod_{j=1}^{k-1} (k+m+j)} \leq k+m+1 \leq (k+1)(m+1).$$

This gives

$$\|\ell\|_T^{\mathcal{A}_h} \leq C \sum_{m=0}^{\infty} \left(\sqrt{\frac{m+1}{m!}} |\sqrt{2T}b|^m \right) \sum_{k=0}^{\infty} \left(\frac{\sqrt{k+1}}{k!} |2Ta|^k \right).$$

Now, thanks to Stirling's formula,

$$\frac{\sqrt{k+1}}{k!} \sim \left(\frac{e}{k}\right)^k \quad \text{and} \quad \sqrt{\frac{m+1}{m!}} \sim m^{\frac{1}{4}} \left(\frac{e}{m}\right)^{\frac{m}{2}}$$

so it is clear that both $\sum_{k \geq 0} \frac{\sqrt{k+1}}{k!} x^k$ and $\sum_{m \geq 0} \sqrt{\frac{m+1}{m!}} y^m$ have an infinite radius of convergence, hence proving

$$\|\ell\|_T^{\mathcal{A}_h} < \infty.$$

□

This naturally leads to the following lemma:

Lemma 3.11. *Let $\mathbf{a} = \sum_{\mathbf{i} \in A_d} \mathbf{a}^{\mathbf{i}} \mathbf{i}$ and $\mathbf{b} = \sum_{\mathbf{i} \in A_d} \mathbf{b}^{\mathbf{i}} \mathbf{i}$ with $\mathbf{a}^{\mathbf{i}}, \mathbf{b}^{\mathbf{i}} \in \mathbb{R}$, then*

$$(\phi - \mathbf{a}e^{\sqcup \mathbf{b}})^{-1} = (\phi - \mathbf{b})e^{\sqcup(\mathbf{a}+\mathbf{b})} \in \mathcal{A}_h$$

and

$$(\phi - e^{\sqcup \mathbf{b}} \mathbf{a})^{-1} = e^{\sqcup(\mathbf{a}+\mathbf{b})} (\phi - \mathbf{b}) \in \mathcal{A}_h.$$

Proof. Let us start with $\mathbf{a}, \mathbf{b} \in T((\mathbb{R}^d))$ s.t. $\mathbf{a}^\phi = \mathbf{b}^\phi = 0$. Observing that $-\mathbf{a} = -\mathbf{a}(\phi - \mathbf{b})^{-1}(\phi - \mathbf{b})$, we obtain:

$$\begin{aligned} \phi - (\mathbf{a} + \mathbf{b}) &= \phi - \mathbf{b} - \mathbf{a}(\phi - \mathbf{b})^{-1}(\phi - \mathbf{b}) \\ &= \left(\phi - \mathbf{a}(\phi - \mathbf{b})^{-1}\right)(\phi - \mathbf{b}). \end{aligned}$$

Similarly we derive that

$$\phi - (\mathbf{a} + \mathbf{b}) = (\phi - \mathbf{b}) \left(\phi - (\phi - \mathbf{b})^{-1} \mathbf{a}\right),$$

which directly implies

$$\left(\phi - \mathbf{a}(\phi - \mathbf{b})^{-1}\right)^{-1} = (\phi - \mathbf{b}) (\phi - (\mathbf{a} + \mathbf{b}))^{-1}$$

and

$$\left(\phi - (\phi - \mathbf{b})^{-1} \mathbf{a}\right)^{-1} = (\phi - (\mathbf{a} + \mathbf{b}))^{-1} (\phi - \mathbf{b}).$$

Now taking $\mathbf{a} = \sum_{\mathbf{i} \in A_d} \mathbf{a}^{\mathbf{i}} \mathbf{i}$ and $\mathbf{b} = \sum_{\mathbf{i} \in A_d} \mathbf{b}^{\mathbf{i}} \mathbf{i}$ and applying Lemma 3.10 together with Proposition 3.7 ends the proof. □

Finally, we prove that the resolvent retains the partial order of Definition 3.8.

Lemma 3.12. *Let $\mathbf{p}, \mathbf{q} \in T((\mathbb{R}^d))$ such that $|\mathbf{p}^\phi|, |\mathbf{q}^\phi| < 1$ and $\mathbf{q} \preceq \mathbf{p}$, then*

$$(\phi - \mathbf{q})^{-1} \preceq (\phi - \mathbf{p})^{-1}.$$

Proof. We will first rewrite the equation of the resolvent (2.5), for $\ell \in T((\mathbb{R}^d))$ such that $|\ell^\phi| < 1$, as

$$(\phi - \ell)^{-1} = \sum_{k=0}^{\infty} \ell^{\otimes k} = \sum_{k=0}^{\infty} \left(\sum_{m=0}^{\infty} \sum_{\mathbf{v} \in V_m} \ell^{\mathbf{v}} \mathbf{v} \right)^{\otimes k} = \sum_{m=0}^{\infty} \sum_{\mathbf{v} \in V_m} f_{\mathbf{v}}(\ell^{\mathbf{u}}, n(\mathbf{u}) \leq m) \mathbf{v}$$

where $f_{\mathbf{v}}$ is a polynomial function of degree n for all $\mathbf{v} \in V_n$, which is independent of ℓ and has $\frac{d^{n+1}-1}{d-1}$ variables $\{\ell^{\mathbf{u}} : |\mathbf{u}| \leq |\mathbf{v}|\}$. We also notice that the coefficients of $f_{\mathbf{v}}$ are all non-negative integers. (For example, $f_{\mathbf{11}}(\ell^{\mathbf{u}}, |\mathbf{u}| \leq 2) = 2\ell^\phi \ell^{\mathbf{11}} + (\ell^{\mathbf{1}})^2$ and $f_{\mathbf{12}}(\ell^{\mathbf{u}}, |\mathbf{u}| \leq 2) = 2\ell^\phi \ell^{\mathbf{12}} + 2\ell^{\mathbf{1}} \ell^{\mathbf{2}}$.) Now recall the Definition 3.8 where $\mathbf{q} \preceq \mathbf{p}$ implies $|\mathbf{q}^{\mathbf{v}}| \leq |\mathbf{p}^{\mathbf{v}}|$ for all $\mathbf{v} \in V$. Combined this with the fact that f has non-negative coefficients, it is easy to see that, for all $\mathbf{v} \in V_n$ we have

$$|f_{\mathbf{v}}(\mathbf{q}^{\mathbf{u}}, n(\mathbf{u}) \leq m)| \leq f_{\mathbf{v}}(\mathbf{p}^{\mathbf{u}}, n(\mathbf{u}) \leq m),$$

which ends the proof. □

We can now prove Theorem 3.9.

Proof of Theorem 3.9. First we assume that $\mathbf{q} \preceq \mathbf{a}e^{\perp\perp b}$. Since $\mathbf{q}^\circ = 0$, Proposition 2.2 yields that $(\circ - \mathbf{q})^{-1}$ is well-defined and

$$\ell = \mathbf{p}(\circ - \mathbf{q})^{-1}.$$

Furthermore, by Lemmas 3.11 and 3.12, we have that $(\circ - \mathbf{q})^{-1} \in \mathcal{A}_h$. Then an application of Proposition 3.7 shows that $\ell \in \mathcal{A}_h$. The case $\mathbf{q} \preceq e^{\perp\perp b}\mathbf{a}$ follows along the same lines. \square

4 Signature representations of path-dependent processes

In this section, we provide our main signature representation formulae for linear path-dependent equations. We consider three categories:

- Linear Volterra equations: exact, time-independent representations for smooth enough kernels, and an approximation result for more general kernels, in Section 4.1,
- Certain linear delay equations: exact, time-independent representations for a weighted sum of exponential kernels, in Section 4.2,
- Gaussian Volterra processes: exact, time-dependent representations for possibly singular kernels, including the Riemann-Liouville fractional Brownian motion for a Hurst index $H \in (0, 1)$, in Section 4.3

In the sequel, we will consider the 2-dimensional semimartingale $\widehat{W}_t = (t, W_t)$, where W is a one-dimensional Brownian motion, we will discuss extensions for a d -dimensional Brownian motion in Remarks 4.6 and 4.8.

4.1 Linear Volterra equations

The first class we consider is linear Volterra equations of the form

$$Y_t = y + \int_0^t K_1(t-s)(a_1 + b_1 Y_s) ds + \int_0^t K_2(t-s)(a_2 + b_2 Y_s) dW_s, \quad (4.1)$$

for some real coefficients a_1, b_1, a_2, b_2, y and locally square-integrable kernels K_1, K_2 . Since the drift and volatility coefficients are linear in Y and hence Lipschitz continuous, the stochastic integral equation (4.1) admits a unique strong solution Y , see for instance [4, Theorem 3.3].

In our next theorem, we provide an explicit infinite linear representation for the solution $Y_t = \langle \ell, \widehat{W}_t \rangle$ in terms of the signature elements of \widehat{W} and time-independent coefficients ℓ . For this, we will need the following structure on the kernels K_1 and K_2 :

$$K_1(u) = \int_{[0, \infty)} e^{-xu} \mu_1(dx) \quad \text{and} \quad K_2(u) = \int_{[0, \infty)} e^{-xu} \mu_2(dx), \quad (4.2)$$

for finite measures μ_1 and μ_2 and such that

$$\int_{[0, \infty)} x^n \mu_1(dx) + \int_{[0, \infty)} x^n \mu_2(dx) < M^n, \quad n \in \mathbb{N}, \quad (4.3)$$

for some constant $M > 0$.

In particular, this implies that K_1, K_2 are infinitely continuously differentiable on $[0, T]$. For such smooth coefficients K_1, K_2 the solution X is even a semimartingale.

Example 4.1. Let us consider $\mu(dx) = f(x)dx$ for a positive bounded measurable function f with compact support. It is easy to see that for such μ , the assumptions (4.2)-(4.3) are satisfied. Another example is the case $\mu(dx) = \sum_{n=1}^k c_n \delta_{x_n}(dx)$ and $K(u) = \sum_{n=1}^k c_n e^{-x_n u}$ for $c_n, x_n \in \mathbb{R}$. \blacksquare

Before we state the theorem, we need to clarify how one can make sense of $\int_{[0,\infty)} \ell(x)\mu(dx)$ when $\ell : x \mapsto \ell(x)$ is a function taking values in $T((\mathbb{R}^d))$.

Definition 4.1 (Integral of parameterized linear functionals). *Given a measure μ on $[0, \infty)$ and a family $(\ell(x))_{x \geq 0}$ such that $\ell(x) = \sum_{\mathbf{v} \in V} \ell^{\mathbf{v}}(x)\mathbf{v} \in T((\mathbb{R}^d))$ for $x \geq 0$, where $\ell^{\mathbf{v}} : \mathbb{R}^+ \rightarrow \mathbb{R}$. We say that $x \mapsto \ell(x)$ is element-wise integrable in terms of the measure μ , if $\int_{[0,\infty)} |\ell^{\mathbf{v}}(x)| \mu(dx) < \infty$ for all $\mathbf{v} \in V$ and we denote the set of all those $x \mapsto \ell(x)$ by $L^1(T((\mathbb{R}^d)), \mu)$. For every $x \mapsto \ell(x) \in L^1(T((\mathbb{R}^d)), \mu)$, we define:*

$$\int_{[0,\infty)} \ell(x)\mu(dx) := \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \left(\int_{[0,\infty)} \ell^{\mathbf{v}}(x)\mu(dx) \right) \mathbf{v}.$$

Furthermore, if $\int_{[0,\infty)} \ell(x)\mu(dx) \in \mathcal{A}$, we say that $\int_{[0,\infty)} \ell(x)\mu(dx)$ is integrable in terms of the measure μ and we denote the set of all those $x \mapsto \ell(x)$ by $L^1(\mathcal{A}, \mu)$.

Example 4.2. Take $\ell(x) = e^{\llbracket -x \mathbf{1} \rrbracket}$ and $\mu(dx) = e^{-\frac{x^2}{2}} dx$, then $\ell \in L^1(T((\mathbb{R}^d)), \mu)$ and thus $\int_{[0,\infty)} \ell(x)\mu(dx)$ is a well-defined element of $T((\mathbb{R}^d))$. Furthermore we can check that $\int_{[0,\infty)} \ell(x)\mu(dx) \in \mathcal{A}$, showing that $\ell \in L^1(\mathcal{A}, \mu)$. Now fixing $\ell(x) = e^{\llbracket -x \mathbf{1} \rrbracket}$, for a general measure μ , we can easily see that $\ell \in L^1(\mathcal{A}, \mu)$ if and only if $\sum_{n=0}^{\infty} \int_{[0,\infty)} \frac{x^n}{n!} \mu(dx) < \infty$. This partially explains why the assumption (4.3) arises. ■

Theorem 4.2. *Fix K_1, K_2 satisfying (4.2)-(4.3). The solution Y to the linear Volterra equation (4.1) admits the time-independent signature representation*

$$Y_t = \left\langle \ell^{\text{VOL}}, \widehat{\mathbb{W}}_t \right\rangle \text{ for every } t \leq T, \text{ a.s., with } \ell^{\text{VOL}} = \mathbf{p}^{\text{VOL}} (\boldsymbol{\phi} - \mathbf{q}^{\text{VOL}})^{-1} \quad (4.4)$$

and

$$\mathbf{p}^{\text{VOL}} := a_1 \mathbf{1} \int_{[0,\infty)} e^{\llbracket -x \mathbf{1} \rrbracket} \mu_1(dx) + a_2 \left(\mathbf{2} - \frac{1}{2} b_2 K_2(0) \mathbf{1} \right) \int_{[0,\infty)} e^{\llbracket -x \mathbf{1} \rrbracket} \mu_2(dx) + y \boldsymbol{\phi}, \quad (4.5)$$

$$\mathbf{q}^{\text{VOL}} := b_1 \mathbf{1} \int_{[0,\infty)} e^{\llbracket -x \mathbf{1} \rrbracket} \mu_1(dx) + b_2 \left(\mathbf{2} - \frac{1}{2} b_2 K_2(0) \mathbf{1} \right) \int_{[0,\infty)} e^{\llbracket -x \mathbf{1} \rrbracket} \mu_2(dx), \quad (4.6)$$

and are such that $\mathbf{p}^{\text{VOL}}, \mathbf{q}^{\text{VOL}} \in \mathcal{A}_h$ and $\ell^{\text{VOL}} \in \mathcal{A}_h \cap \mathcal{I}(\widehat{W})$. In particular, they all belong to \mathcal{A} .

Proof. The proof is given in Section 5. □

For $b_2 = 0$ and $K_1 = K_2 \equiv 1$, one recovers from the representation in Theorem 4.2 a time-independent representation for the Ornstein-Uhlenbeck process as shown in the following example. For a time-dependent representation of the same process, we refer to Example 4.10 below.

Example 4.3 (Ornstein-Uhlenbeck). Set $b_2 = 0$, and $\mu_1(dx) = \mu_2(dx) = \delta_0(dx)$, where δ_0 is the Dirac mass at 0. In this case, $K_1(u) = K_2(u) = 1$ and equation (4.1) reads

$$dY_t = (a_1 + b_1 Y_t) dt + a_2 dW_t, \quad Y_0 = y \in \mathbb{R}.$$

meaning that Y is an Ornstein-Uhlenbeck process. Moreover, in this case, $\int_{[0,\infty)} e^{\llbracket -x \mathbf{1} \rrbracket} \mu_1(dx) = \int_{[0,\infty)} e^{\llbracket -x \mathbf{1} \rrbracket} \mu_2(dx) = \boldsymbol{\phi}$. So that

$$\mathbf{q}^{\text{VOL}} = b_1 \mathbf{1}, \quad \mathbf{p}^{\text{VOL}} = y \boldsymbol{\phi} + a_1 \mathbf{1} + a_2 \mathbf{2}.$$

An application of Theorem 4.2 yields that the Ornstein-Uhlenbeck process Y can be written as

$$Y_t = \left\langle \ell^{\text{OU}}, \widehat{\mathbb{W}}_t \right\rangle$$

with

$$\ell^{\text{OU}} = (y \boldsymbol{\phi} + a_1 \mathbf{1} + a_2 \mathbf{2}) (\boldsymbol{\phi} - b_1 \mathbf{1})^{-1} = (y \boldsymbol{\phi} + a_1 \mathbf{1} + a_2 \mathbf{2}) e^{\llbracket b_1 \mathbf{1} \rrbracket},$$

where the second equality follows from Proposition 2.3. ■

Furthermore, for $a_1 = a_2 = 0$ and $K_1 = K_2 \equiv 1$, one recovers from the representation in Theorem 4.2 the explicit solution of the geometric Brownian motion:

Example 4.4 (Geometric Brownian motion). Take $a_1 = a_2 = 0$, and $\mu_1(dx) = \mu_2(dx) = \delta_0(dx)$, where δ_0 is the Dirac mass at 0. In this case, $K_1(u) = K_2(u) = 1$ and $\int_{[0,\infty)} e^{\sqcup-x} \mathbf{1} \mu_1(dx) = \int_{[0,\infty)} e^{\sqcup-x} \mathbf{1} \mu_2(dx) = \mathfrak{o}$.

Therefore $\mathbf{q}^{\text{VOL}} = (b_1 - \frac{1}{2}b_2^2) \mathbf{1} + b_2 \mathbf{2}$, $\mathbf{p}^{\text{VOL}} = y \mathfrak{o}$ and thus, using Proposition 2.3,

$$\ell^{\text{VOL}} = y(\mathfrak{o} - (b_1 - \frac{1}{2}b_2^2) \mathbf{1} - b_2 \mathbf{2})^{-1} = ye^{\sqcup(b_1 - \frac{1}{2}b_2^2)\mathbf{1} + b_2 \mathbf{2}}.$$

It follows from the shuffle property 3.1 and the definition of the shuffle exponential in (2.7), that

$$Y_t = y \left\langle e^{\sqcup(b_1 - \frac{1}{2}b_2^2)\mathbf{1} + b_2 \mathbf{2}}, \widehat{\mathbb{W}}_t \right\rangle = ye^{\langle (b_1 - \frac{1}{2}b_2^2)\mathbf{1} + b_2 \mathbf{2}, \widehat{\mathbb{W}}_t \rangle} = ye^{(b_1 - \frac{1}{2}b_2^2)t + b_2 W_t},$$

which is the geometric Brownian motion, solution to

$$\frac{dY_t}{Y_t} = b_1 dt + b_2 dW_t, \quad Y_0 = y \in \mathbb{R}.$$

■

Upon initial inspection, our assumptions regarding K_1 and K_2 may appear limiting. However, the stability results of stochastic Volterra equations allow us to approximate solutions for broader classes of kernels. For example, we can approximate solutions for the singular fractional kernel $K(t) = t^{H-1/2}$ with $H \in (0, 1/2)$ using linear combinations of the signature elements, as demonstrated in the following corollary.

Corollary 4.3. *Let K_1, K_2 be locally square-integrable kernels and denote by Y the solution of (4.1). For $n \in \mathbb{N}$, let $\mu_1^n(dx) = \sum_{i=1}^n c_i \delta_{x_i}(dx)$ and $\mu_2^n(dx) = \sum_{i=1}^n d_i \delta_{y_i}(dx)$, for some $c_i, d_i, x_i, y_i \in \mathbb{R}$, and set K_1^n, K_2^n as in (4.2) and $Y_t^n = \langle \ell_n^{\text{VOL}}, \widehat{\mathbb{W}}_t \rangle$ as in (4.4) both with μ_1^n and μ_2^n instead of μ_1 and μ_2 . Assume that*

$$\int_0^T |K_1^n(s) - K_1(s)|^2 ds + \int_0^T |K_2^n(s) - K_2(s)|^2 ds \rightarrow 0, \quad \text{as } n \rightarrow \infty. \quad (4.7)$$

Then,

$$\sup_{t \in [0, T]} \mathbb{E} \left[\left| Y_t - \langle \ell_n^{\text{VOL}}, \widehat{\mathbb{W}}_t \rangle \right|^p \right] \rightarrow 0, \quad \text{as } n \rightarrow \infty, \quad p \in \mathbb{N}.$$

Proof. For $n \in \mathbb{N}$, it follows from Example 4.1 that μ^n and λ^n satisfy the assumptions (4.3). Hence, an application of Theorem 4.2 yields that $Y_t^n = \langle \ell_n^{\text{VOL}}, \widehat{\mathbb{W}}_t \rangle$ solves the stochastic Volterra equation

$$Y_t^n = y + \int_0^t K_1^n(t-s) (a_1 + b_1 Y_s) ds + \int_0^t K_2^n(t-s) (a_2 + b_2 Y_s) dW_s.$$

Routine applications of Jensen's and Grownall's convolution inequalities, see for instance the proof of [4, Lemma 2.4] and [20, Lemma 9.8.2], yield that

$$\sup_{t \in [0, T]} \mathbb{E} [|Y_t - Y_t^n|^p] \leq C_p \left(\left(\int_0^T |K_1^n(s) - K_1(s)|^2 ds \right)^{\frac{p}{2}} + \left(\int_0^T |K_2^n(s) - K_2(s)|^2 ds \right)^{\frac{p}{2}} \right), \quad p \in \mathbb{N},$$

for some $C_p > 0$, which ends the proof. □

Remark 4.5. Different choices exist in the literature on Volterra processes for the parameters $c_k^n, d_k^n, x_k^n, y_k^n$ to ensure the convergence of the weighted sum of exponentials towards the initial kernels in (4.7) whenever K_1 and K_2 are completely monotone functions on $(0, \infty)$, i.e. infinitely differentiable on $(0, \infty)$ such that $(-1)^n K^{(n)} \geq 0$. For instance, for a Hurst index $H \in (0, 1/2)$, the singular fractional kernel $t^{H-1/2}$ is completely monotone and admits the following representation (following Bernstein's theorem) as a Laplace transform:

$$\frac{t^{H-\frac{1}{2}}}{\Gamma(H + \frac{1}{2})} = \int_{[0, \infty)} e^{-xt} \mu(dx) \quad \text{and} \quad \mu(dx) = \frac{x^{-H-\frac{1}{2}}}{\Gamma(H + \frac{1}{2})\Gamma(\frac{1}{2} - H)} dx.$$

For $n \geq 1$ and $r_n > 1$, such that

$$r_n \downarrow 1 \quad \text{and} \quad n \ln r_n \rightarrow \infty, \quad \text{as } n \rightarrow \infty,$$

with the following parametrization for the weights and the mean reversion

$$c_k^n = \frac{(r_n^{1-\alpha} - 1)r_n^{(\alpha-1)(1+\frac{n}{2})}}{\Gamma(\alpha)\Gamma(2-\alpha)} r_n^{(1-\alpha)k} \quad \text{and} \quad x_k^n = \frac{1-\alpha}{2-\alpha} \frac{r_n^{2-\alpha} - 1}{r_n^{1-\alpha} - 1} r_n^{k-1-n/2}, \quad \text{for } k = 1, \dots, n,$$

where $\alpha := H + 1/2$ for some $H \in (0, 1/2)$, one can obtain the convergence of $\sum_{k=1}^n c_k^n e^{-x_k^n t}$ towards the fractional kernel in the sense of (4.7), see [1, 2] for more details. \blacksquare

Remark 4.6. Theorem 4.2 can be readily extended to the case where the Volterra equation is driven by a $(d-1)$ -dimensional Brownian motion W of the form

$$Y_t = y + \int_0^t K_1(t-s)(a_1 + b_1 Y_s) ds + \sum_{\mathbf{j} \in \{2, \dots, d\}} \int_0^t K_j(t-s)(a_j + b_j Y_s) dW_s^{\mathbf{j}}.$$

Here each K_i is of the form $K_i(u) = \int_{[0, \infty)} e^{-xu} \mu_i(dx)$ with a finite measure μ_i satisfying (4.3) with the same constant M , for $\mathbf{i} \in \{1, \dots, d\}$. Then, similar to Theorem 4.2, we obtain the representation

$$Y_t = \left\langle \ell^{\text{VOL}}, \widehat{\mathbb{W}}_t \right\rangle, \quad \ell^{\text{VOL}} = \mathbf{p}^{\text{VOL}} (\boldsymbol{\phi} - \mathbf{q}^{\text{VOL}})^{-1},$$

while \mathbf{p}^{VOL} and \mathbf{q}^{VOL} are updated to:

$$\begin{aligned} \mathbf{p}^{\text{VOL}} &:= \sum_{\mathbf{i} \in \{1, \dots, d\}} a_i \mathbf{i} \int_{[0, \infty)} e^{\llbracket -x \mathbf{1} \rrbracket} \mu_i(dx) - \frac{1}{2} \sum_{\mathbf{j} \in \{2, \dots, d\}} a_j b_j K_j(0) \mathbf{1} \int_{[0, \infty)} e^{\llbracket -x \mathbf{1} \rrbracket} \mu_j(dx) + y \boldsymbol{\phi}, \\ \mathbf{q}^{\text{VOL}} &:= \sum_{\mathbf{i} \in \{1, \dots, d\}} b_i \mathbf{i} \int_{[0, \infty)} e^{\llbracket -x \mathbf{1} \rrbracket} \mu_i(dx) - \frac{1}{2} \sum_{\mathbf{j} \in \{2, \dots, d\}} b_j^2 K_j(0) \mathbf{1} \int_{[0, \infty)} e^{\llbracket -x \mathbf{1} \rrbracket} \mu_j(dx). \end{aligned}$$

4.2 Linear delay equations

A second equation that we consider is a delay equation with linear delays both in the drift and in the volatility parts. To keep notations simple and highlight the main ideas, we restrict to weighted sum of exponential kernels:

$$\begin{aligned} dZ_t &= \left(a_1 + b_1 Z_t + \int_0^t K_1(t-s) Z_s ds \right) dt + \left(a_2 + b_2 Z_t + \int_0^t K_2(t-s) Z_s ds \right) dW_t, \\ Z_0 &= z \in \mathbb{R}, \end{aligned} \quad (4.8)$$

where

$$K_i(t) = \sum_{m=1}^{k_i} c_i^m e^{\alpha_i^m t},$$

with $a_i, b_i, c_i^m, \alpha_i^m \in \mathbb{R}, k_i \in \mathbb{N}$ for $i = 1, 2$. Again, it is straightforward to obtain the existence and uniqueness of a strong solution Z since the coefficients are linear. We stress that Z cannot be written as a Volterra equation in the sense of (4.1) unless K_2 is zero.

Theorem 4.4. *The solution Z to the linear delay equation (4.8) admits the time-independent signature representation*

$$Z_t = \left\langle \ell^{\text{DE}}, \widehat{\mathbb{W}}_t \right\rangle \text{ for every } t \leq T, \text{ a.s., with } \ell^{\text{DE}} = \mathbf{p}^{\text{DE}} (\boldsymbol{\phi} - \mathbf{q}^{\text{DE}})^{-1}, \quad (4.9)$$

and

$$\begin{aligned} \mathbf{p}^{\text{DE}} &:= (a_1 - \frac{1}{2} b_2 a_2) \mathbf{1} + a_2 \mathbf{2} + z \boldsymbol{\phi}, \\ \mathbf{q}^{\text{DE}} &:= (b_1 - \frac{1}{2} b_2^2) \mathbf{1} + b_2 \mathbf{2} + \left(\sum_{n=1}^{k_1} c_1^n \mathbf{1} e^{\llbracket -\alpha_1^n \mathbf{1} \rrbracket} - \frac{1}{2} b_2 \sum_{m=1}^{k_2} c_2^m \mathbf{1} e^{\llbracket -\alpha_2^m \mathbf{1} \rrbracket} \right) \mathbf{1} + \left(\sum_{m=1}^{k_2} c_2^m \mathbf{1} e^{\llbracket -\alpha_2^m \mathbf{1} \rrbracket} \right) \mathbf{2}, \end{aligned}$$

and are such that $\mathbf{p}^{\text{DE}}, \mathbf{q}^{\text{DE}} \in \mathcal{A}_h$ and $\ell^{\text{DE}} \in \mathcal{A}_h \cap \mathcal{I}(\widehat{\mathbb{W}})$. In particular, they all belong to \mathcal{A} .

Proof. The proof is given in Section 6. \square

Remark 4.7. For more general kernels K_1, K_2 a similar stability argument as in Corollary 4.3 allows to approximate the solution to the delay equation by linear combination of the signatures. \blacksquare

Remark 4.8. Similar to Remark 4.6, Theorem 4.4 can also be extended to the case where the Delayed equation is driven by a $(d-1)$ -dimensional Brownian motion W of the form:

$$dZ_t = \left(a_1 + b_1 Z_t + \int_0^t K_1(t-s)Z_s ds \right) dt + \sum_{j \in \{2, \dots, d\}} \left(a_j + b_j Z_t + \int_0^t K_j(t-s)Z_s ds \right) dW_t^j,$$

$$Z_0 = z \in \mathbb{R},$$

where

$$K_i(t) = \sum_{m=1}^{k_i} c_i^m e^{\alpha_i^m t}.$$

Then similar to Theorem 4.4, we have the representation:

$$Z_t = \left\langle \ell^{\text{DE}}, \widehat{W}_t \right\rangle, \quad \ell^{\text{DE}} = \mathbf{p}^{\text{DE}} (\boldsymbol{\phi} - \mathbf{q}^{\text{DE}})^{-1},$$

while \mathbf{p}^{DE} and \mathbf{q}^{DE} are updated to:

$$\mathbf{p}^{\text{DE}} = \sum_{i \in \{1, \dots, d\}} a_i \mathbf{i} - \frac{1}{2} \sum_{j \in \{2, \dots, d\}} a_j b_j \mathbf{1} + z \boldsymbol{\phi},$$

$$\mathbf{q}^{\text{DE}} = \sum_{i \in \{1, \dots, d\}} b_i \mathbf{i} - \frac{1}{2} \sum_{j \in \{2, \dots, d\}} b_j^2 \mathbf{1} + \sum_{i \in \{1, \dots, d\}} \sum_{m=1}^{k_i} c_i^m \mathbf{1} e^{\llbracket \alpha_i^m \rrbracket \mathbf{i}} - \frac{1}{2} \sum_{j \in \{2, \dots, d\}} b_j \sum_{m=1}^{k_j} c_j^m e^{\llbracket \alpha_j^m \rrbracket \mathbf{1}} \mathbf{1} \mathbf{1}.$$

4.3 Gaussian Volterra processes

The final class that we consider are continuous Gaussian Volterra processes of the form

$$X_t = \int_0^t K(t-s) dW_s, \quad (4.10)$$

for a locally square-integrable deterministic kernel K that is infinitely differentiable on $(0, T]$ and satisfies

$$\begin{cases} \sum_{n=0}^{\infty} |K^{(n)}(t)| \frac{t^{n+\frac{1}{2}}}{n! \sqrt{2n+1}} < \infty & \text{and} & \int_0^t \left(\sum_{n=0}^{\infty} |K^{(n)}(t)| \frac{s^n}{n!} \right)^2 ds < \infty, \quad t \in (0, T], \\ \int_0^T \left(\sum_{n=0}^{\infty} |K^{(n)}(t)| \frac{t^{n+\frac{1}{2}}}{n! \sqrt{2n+1}} \right) dt < \infty, \end{cases} \quad (4.11)$$

where $K^{(n)}$ denotes the n -th derivative on $(0, T]$.

We note that here, and in contrast with Theorem 4.2, K can have a singularity at 0, as shown in Example 4.11 for the Riemann-Liouville fractional Brownian motion. For such kernels, Corollary 4.3 already provides an approximation result by infinite linear combinations of signature elements of \widehat{W} with time-independent coefficients ℓ . More interestingly, the following theorem shows that exact infinite linear representation of X are possible but with time-dependent coefficients ℓ_t . Time-dependent representations seemed crucial for us to obtain exact representations for singular kernels.

Theorem 4.5. *Fixed $K : (0, T] \rightarrow \mathbb{R}$ satisfying (4.11), the Gaussian Volterra process X defined in (4.10) admits a time-dependent signature representation*

$$X_t = \left\langle \ell_t^{\text{GV}}, \widehat{W}_t \right\rangle, \quad a.s.$$

for every $t \leq T$, with ℓ_t^{GV} given by

$$\ell_t^{\text{GV}} = \begin{cases} \sum_{n=0}^{\infty} K^{(n)}(t) (-1)^n \mathbf{1}^{\otimes n} \mathbf{2} & \text{if } t \in (0, T], \\ 0 & \text{else.} \end{cases} \quad (4.12)$$

Proof. The proof is given in Section 7. □

Remark 4.9. Unlike the representations in Theorems 4.2 and 4.4, here $\langle \ell_t^{\text{GV}}, \widehat{\mathbb{W}}_t \rangle$ is a modification of X_t and it is not necessarily indistinguishable from the process X .

The following example shows that the exponential kernel $K(t) = e^{-\kappa t}$ for $\kappa > 0$ satisfies the assumption of Theorem 4.5. This provides a time-dependent representation of the Ornstein-Uhlenbeck process, recall Example 4.3 for the time-independent representation.

Example 4.10 (Ornstein-Uhlenbeck process). Let

$$X_t = \int_0^t e^{-\kappa(t-s)} dW_s = \int_0^t K(t-s) dW_s$$

with $K(t) = e^{-\kappa t}$ and $\kappa \in \mathbb{R}$. We first check that K satisfies (4.11). Clearly K is infinitely differentiable on $[0, T]$ such that $|K^{(n)}(t)| = |\kappa|^n K(t)$. Therefore,

$$\int_0^t \left(\sum_{n=0}^{\infty} \frac{|K^{(n)}(t)| s^n}{n!} \right)^2 ds = \int_0^t \left(K(t) \sum_{n=0}^{\infty} \frac{(|\kappa|s)^n}{n!} \right)^2 ds = K(t)^2 \int_0^t e^{2|\kappa|s} ds < \infty.$$

Moreover

$$\sum_{n=0}^{\infty} \frac{|K^{(n)}(t)|}{n! \sqrt{2n+1}} t^{n+\frac{1}{2}} = \sqrt{t} e^{-\kappa t} \sum_{n=0}^{\infty} \frac{(|\kappa|t)^n}{n!} \frac{1}{\sqrt{2n+1}} \leq \sqrt{T} e^{|\kappa|T} \sum_{n=0}^{\infty} \frac{(|\kappa|t)^n}{n!} = \sqrt{T} e^{2|\kappa|T} < \infty.$$

Thus

$$\int_0^T \left(\sum_{n=0}^{\infty} \frac{|K^{(n)}(t)|}{n! \sqrt{2n+1}} t^{n+\frac{1}{2}} \right) dt \leq T^{\frac{3}{2}} e^{2|\kappa|T} < \infty.$$

Finally, Theorem 4.5 yields the signature representation $X_t = \langle \ell_t^{\text{OU}}, \widehat{\mathbb{W}}_t \rangle$, with

$$\ell_t^{\text{OU}} = e^{-\kappa t} \sum_{n=0}^{\infty} \kappa^n \mathbf{1}^{\otimes n} \mathbf{2}.$$

Note that we can more compactly write it as

$$\ell_t^{\text{OU}} = e^{\llbracket -\kappa(t\theta-1) \rrbracket} \mathbf{2},$$

which is the algebraic translation of the probabilistic expression

$$X_t = \int_0^t e^{-\kappa(t-s)} dW_s.$$

■

The following example shows that the fractional kernel $K(t) = t^{H-1/2}$ for $H \in (0, 1)$, with singularity at 0 for $H < 1/2$, satisfies the assumption of Theorem 4.5. This provides a representation of the Riemann-Liouville fractional Brownian motion, and shows that infinite linear combinations of signature elements are not always semimartingales, recall Remark 3.7.

Example 4.11 (Riemann-Liouville fractional Brownian motion). Let

$$X_t = \frac{1}{\Gamma(H + \frac{1}{2})} \int_0^t (t-s)^{H-\frac{1}{2}} dW_s = \int_0^t K(t-s) dW_s,$$

with $K(t) = \frac{t^{H-\frac{1}{2}}}{\Gamma(H+\frac{1}{2})}$. We first check that K satisfies (4.11). Clearly K is analytical on $(0, T]$ for $H \in (0, 1)$:

$$K^{(n)}(t) = K(t) t^{-n} \left(H - \frac{1}{2}\right)^{\underline{n}} = K(t) (-t)^{-n} \left(\frac{1}{2} - H\right)^{\bar{n}}$$

where $x^n = \prod_{k=0}^{n-1} (x - k)$ and $x^{\bar{n}} = \prod_{k=0}^{n-1} (x + k)$ denote the falling and rising factorials respectively.

Furthermore, one can then remark that

$$|K^{(n)}(t)| = \begin{cases} K^{(n)}(t)(-1)^n & \text{if } H < \frac{1}{2}, \\ K^{(n)}(t)(-1)^{n+1} & \text{else.} \end{cases}$$

Therefore

$$\int_0^t \left(\sum_{n=0}^{\infty} \frac{|K^{(n)}(t)| s^n}{n!} \right)^2 ds = \int_0^t \left(K(t) \sum_{n=0}^{\infty} (H - \frac{1}{2})^n \frac{1}{n!} \left(\frac{-s}{t} \right)^n \right)^2 ds = \int_0^t K(t-s)^2 ds < \infty,$$

where the last equality comes from the Taylor expansion of $(1-x)^{H-1/2}$ which converges for $|x| < 1$. Moreover we have

$$|K^{(n)}(t)| = K(t) \frac{\Gamma(n + \frac{1}{2} - H)}{\Gamma(\frac{1}{2} - H)} t^{-n} = O\left(\left(\frac{n}{t}\right)^{\frac{1}{2}-H} (n-1)! t^{-n}\right), \quad n > 0,$$

therefore

$$\sum_{n=0}^{\infty} \frac{|K^{(n)}(t)|}{n! \sqrt{2n+1}} t^{n+\frac{1}{2}} \leq C t^H \sum_{n=1}^{\infty} n^{-1-H} < \infty.$$

And

$$\int_0^T \left(\sum_{n=0}^{\infty} \frac{|K^{(n)}(t)|}{n! \sqrt{2n+1}} t^{n+\frac{1}{2}} \right) dt \leq \frac{CT^{H+1}}{H+1} \sum_{n=1}^{\infty} n^{-1-H} < \infty.$$

Finally, Theorem 4.5 yields the signature representation $X_t = \langle \ell_t^{\text{RL}}, \widehat{\mathbb{W}}_t \rangle$, with

$$\ell_t^{\text{RL}} = \frac{t^{H-\frac{1}{2}}}{\Gamma(H+\frac{1}{2})} \sum_{n=0}^{\infty} t^{-n} \left(\frac{1}{2} - H\right)^{\bar{n}} \mathbf{1}^{\otimes n} \mathbf{2}, \quad (4.13)$$

where $(\cdot)^{\bar{n}}$ is the rising factorial. Remarking that $(\frac{1}{2} - H)^{\bar{n}} = 0$ for all $n > 0$ when $H = \frac{1}{2}$ gives $\ell_t = \mathbf{2}$, the coordinate of W_t in $\widehat{\mathbb{W}}_t$. Note that we can also formally write it as

$$\ell_t^{\text{RL}} = \frac{1}{\Gamma(H+\frac{1}{2})} (t\phi - \mathbf{1})^{\sqcup H - \frac{1}{2}} \mathbf{2},$$

which is the algebraic translation of the probabilistic expression

$$X_t = \frac{1}{\Gamma(H+\frac{1}{2})} \int_0^t (t-s)^{H-\frac{1}{2}} dW_s.$$

■

4.4 Numerical illustrations

In this subsection, we provide numerical implementations of the truncated form of the signature processes $X^{\leq M}$ defined by

$$X_t^{\leq M} := \langle \ell, \widehat{\mathbb{W}}_t^{\leq M} \rangle, \quad M \geq 0,$$

where $\ell \in T((\mathbb{R}^2))$ and $\widehat{\mathbb{W}}_t^{\leq M} : [0, T] \rightarrow T^M(\mathbb{R}^2)$ is defined in (2.11), i.e. its M first levels coincide with $\widehat{\mathbb{W}}_t$ and everything else is set to 0.

4.4.1 Shifted Riemann-Liouville fractional Brownian motion

Figure 1 and Table 1 display simulations of a shifted Riemann-Liouville fractional Brownian motion X^ϵ together with mean-squared errors, i.e.

$$X_t^\epsilon = \frac{1}{\Gamma(H + \frac{1}{2})} \int_0^t (t + \epsilon - s)^{H - \frac{1}{2}} dW_s,$$

and its linear representation counterpart $\ell_{t+\epsilon}^{\text{RL}}$, following from (4.13), that is

$$\ell_{t+\epsilon}^{\text{RL}} = \frac{(t + \epsilon)^{H - \frac{1}{2}}}{\Gamma(H + \frac{1}{2})} \sum_{n=0}^{\infty} (t + \epsilon)^{-n} \left(\frac{1}{2} - H\right)^{\bar{n}} \mathbf{1}^{\otimes n} \mathbf{2} = \frac{1}{\Gamma(H + \frac{1}{2})} ((t + \epsilon)\phi - \mathbf{1})^{\lfloor H - \frac{1}{2} \rfloor} \mathbf{2},$$

for several truncation orders $M \in \mathbb{N}$. The reason for shifting the kernel is to speed up the convergence as it can be quite slow for small values of H when $\epsilon = 0$.

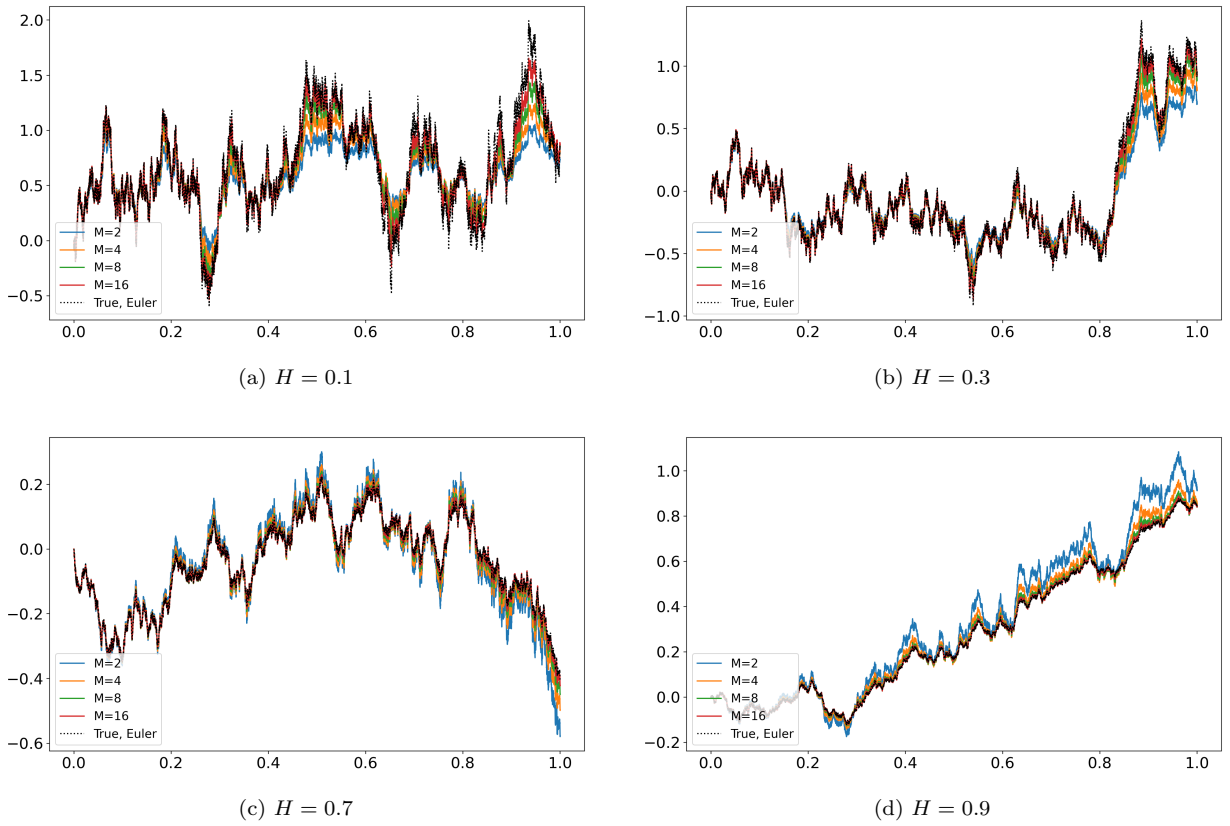


Figure 1: Trajectories of a shifted Riemann-Liouville fractional Brownian motion (black) against their truncated time-dependent linear representation (4.13), i.e. $\langle \ell_{t+\epsilon}^{\text{RL}}, \widehat{\mathbb{W}}_t^{\leq M} \rangle$, for several truncation orders M and $\epsilon = 1/52$.

	H=0.1	H=0.3	H=0.7	H=0.9
M=2	9.915e-02	1.600e-02	4.015e-03	6.217e-03
M=4	4.735e-02	6.197e-03	9.022e-04	9.931e-04
M=8	1.851e-02	1.992e-03	1.819e-04	1.528e-04
M=16	5.298e-03	4.712e-04	2.793e-05	1.848e-05

Table 1: Mean squared error between the shifted Riemann-Liouville fractional Brownian motion and its truncated time-dependent linear representation (4.13), i.e. $\langle \ell_{t+\epsilon}^{\text{RL}}, \widehat{\mathbb{W}}_t^{\leq M} \rangle$, for several truncation orders M and $\epsilon = 1/52$, averaged across 100,000 simulations over 1000 time-steps.

We can see in Table 1 and Figure 1 the clear convergence of the truncated linear representation for several values of H . The smaller H is, the slower the convergence. We can also see that the fit deteriorates with the time variable t .

4.4.2 Linear delay process

Figure 2 and Table 2 display simulations of a delay equation and its truncated representations together with mean-squared errors following Theorem 4.4. This convergence is shown for two sets of parameters: (a) has its b_1 and b_2 coefficients set to 0, contrary to (b) that has even larger coefficients.

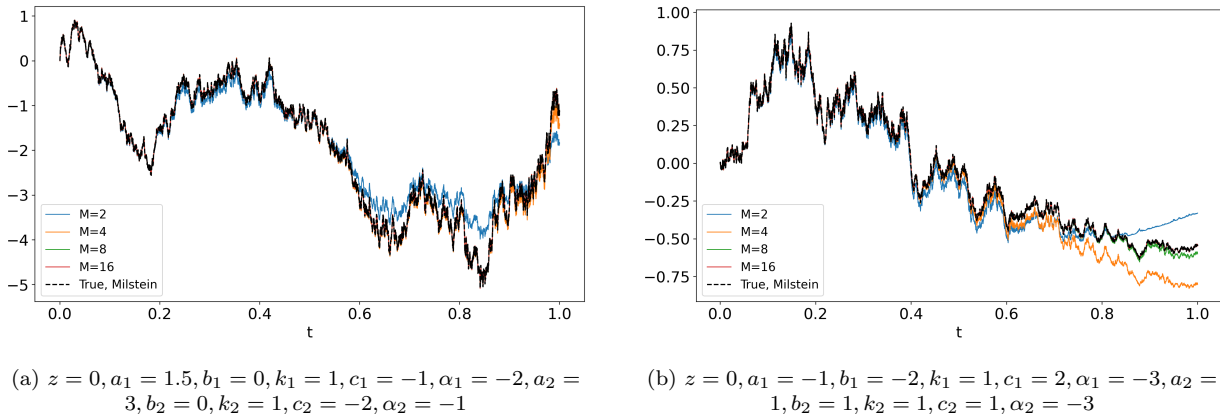


Figure 2: Trajectories of a delay process (black) against their truncated time-independent linear representation (4.9), i.e. $\langle \ell^{\text{DE}}, \widehat{\mathbb{W}}_t^{\leq M} \rangle$, for several truncation orders M .

	(a)	(b)
M=2	4.202e-01	7.984e-01
M=4	9.035e-03	6.742e-01
M=8	1.180e-06	1.554e-02
M=16	4.073e-07	5.744e-07

Table 2: Mean squared error between the delay process and its truncated time-independent linear representation (4.9), i.e. $\langle \ell^{\text{DE}}, \widehat{\mathbb{W}}_t^{\leq M} \rangle$, for several truncation orders M , averaged across 100,000 simulations over 1000 time-steps.

We can see in Table 2 and Figure 2 the clear convergence of the truncated linear representation for the delay equation process.

Recall that here since $\ell^{\text{DE}} \in \mathcal{A}_h$, the inequality (3.14) could in principle yield a control of the strong error, i.e. for every $\epsilon > 0$, there exists $M_\epsilon \in \mathbb{N}$ such that

$$\mathbb{E} \left[\sup_{t \in [0, T]} \left| \langle \ell^{\text{DE}}, \widehat{\mathbb{W}}_t^{\leq M_\epsilon} \rangle - \langle \ell^{\text{DE}}, \widehat{\mathbb{W}}_t \rangle \right| \right] \leq \epsilon.$$

For related numerical illustrations with other processes, we refer to Abi Jaber and Gérard [3].

5 Proof of Theorem 4.2

We start by noting that the element $\ell^{\text{VOL}} = \mathbf{p}^{\text{VOL}}(\phi - \mathbf{q}^{\text{VOL}})^{-1}$ given in (4.4) solves the linear equation

$$\ell^{\text{VOL}} = \mathbf{p}^{\text{VOL}} + \ell^{\text{VOL}} \mathbf{q}^{\text{VOL}},$$

as established by Proposition 2.2 where $\mathbf{p}^{\text{VOL}}, \mathbf{q}^{\text{VOL}}$ are defined in (4.5) and (4.6).

In the first step, we aim to apply Theorem 3.9 to show that $\ell^{\text{VOL}} \in \mathcal{A}_h$. For this, we start by arguing that \mathbf{q}^{VOL} satisfies condition (3.15) and that $\mathbf{p}^{\text{VOL}} \in \mathcal{A}_h$. Indeed, \mathbf{q}^{VOL} is dominated, in the sense of Definition 3.8, by

$$\begin{aligned} \mathbf{q}^{\text{VOL}} &\preceq \left(|b_1| + |b_2| + \frac{1}{2} |b_2^2 K_2(0)| \right) (\mathbf{1} + \mathbf{2}) \sum_{n=0}^{\infty} \mathbf{1}^{\otimes n} \int_{[0, \infty)} x^n (\mu_1(dx) + \mu_2(dx)) \\ &\preceq \left(|b_1| + |b_2| + \frac{1}{2} |b_2^2 K_2(0)| \right) (\mathbf{1} + \mathbf{2}) e^{\llbracket M \mathbf{1} \rrbracket}, \end{aligned}$$

where the first inequality follows from the definition of \mathbf{q}^{VOL} and the second one from assumption (4.3). This shows that \mathbf{q}^{VOL} satisfies the dominating condition (3.15). Similarly,

$$\mathbf{p}^{\text{VOL}} \preceq \left(|a_1| + |a_2| + \frac{1}{2} |a_2 b_2 K_2(0)| \right) (\mathbf{1} + \mathbf{2}) e^{\llbracket M \mathbf{1} \rrbracket} + |y| \mathbf{\emptyset},$$

By the sub-multiplicativity in Proposition 3.7 and lemma 3.10, we derive $\mathbf{q}^{\text{VOL}}, \mathbf{p}^{\text{VOL}} \in \mathcal{A}_h$. An application of Theorem 3.9 yields that $\ell^{\text{VOL}} \in \mathcal{A}_h$. In particular, $\ell^{\text{VOL}} \in \mathcal{A}$ and the process

$$\bar{Y}_t := \left\langle \ell^{\text{VOL}}, \widehat{\mathbb{W}}_t \right\rangle \quad (5.1)$$

is well defined.

In the next step, we would like to apply Itô's formula of Theorem 3.3 on (5.1) to prove that \bar{Y} solves the equation

$$\begin{aligned} d\bar{Y}_t &= K_1(0) (a_1 + b_1 \bar{Y}_t) dt + \int_0^t K_1'(t-s) (a_1 + b_1 \bar{Y}_s) ds dt \\ &\quad + K_2(0) (a_2 + b_2 \bar{Y}_t) dW_t + \int_0^t K_2'(t-s) (a_2 + b_2 \bar{Y}_s) dW_s dt, \\ \bar{Y}_0 &= y. \end{aligned} \quad (5.2)$$

which is the differential version of (4.1) obtained from Itô's lemma on (4.1) since the kernels K_1, K_2 are smooth. Then, we immediately deduce that $Y_t = \bar{Y}_t = \left\langle \ell^{\text{VOL}}, \widehat{\mathbb{W}}_t \right\rangle$ by uniqueness of the strong solution, which will end the proof of Theorem 4.2.

Therefore, it remains to

(i) **Goal 1:** argue that $\ell^{\text{VOL}} \in \mathcal{I}(\widehat{\mathbb{W}})$ to apply Theorem 3.3 and get

$$d \left\langle \ell^{\text{VOL}}, \widehat{\mathbb{W}}_t \right\rangle = \left\langle \ell^{\text{VOL}}|_{\mathbf{1}} + \frac{1}{2} \ell^{\text{VOL}}|_{\mathbf{22}}, \widehat{\mathbb{W}}_t \right\rangle dt + \left\langle \ell^{\text{VOL}}|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \right\rangle dW_t,$$

(ii) **Goal 2:** identify the coefficients to obtain (5.2).

In order to achieve both goals, we start by computing the projections $\ell^{\text{VOL}}|_{\mathbf{1}}, \ell^{\text{VOL}}|_{\mathbf{2}}, \ell^{\text{VOL}}|_{\mathbf{22}}$. First recall

$$\begin{aligned} \mathbf{p}^{\text{VOL}} &= a_1 \mathbf{1} \int_{[0, \infty)} e^{\llbracket \mathbf{1} - x \mathbf{1} \rrbracket} \mu_1(dx) + a_2 \left(\mathbf{2} - \frac{1}{2} K_2(0) b_2 \mathbf{1} \right) \int_{[0, \infty)} e^{\llbracket \mathbf{1} - x \mathbf{1} \rrbracket} \mu_2(dx) + y \mathbf{\emptyset}, \\ \mathbf{q}^{\text{VOL}} &= b_1 \mathbf{1} \int_{[0, \infty)} e^{\llbracket \mathbf{1} - x \mathbf{1} \rrbracket} \mu_1(dx) + b_2 \left(\mathbf{2} - \frac{1}{2} K_2(0) b_2 \mathbf{1} \right) \int_{[0, \infty)} e^{\llbracket \mathbf{1} - x \mathbf{1} \rrbracket} \mu_2(dx). \end{aligned}$$

We compute that

$$\begin{aligned} \mathbf{p}^{\text{VOL}}|_{\mathbf{1}} &= a_1 \int_{[0, \infty)} e^{\llbracket \mathbf{1} - x \mathbf{1} \rrbracket} \mu_1(dx) - a_2 \int_{[0, \infty)} \left(x \mathbf{2} + \frac{1}{2} K_2(0) b_2 \mathbf{\emptyset} \right) e^{\llbracket \mathbf{1} - x \mathbf{1} \rrbracket} \mu_2(dx), \\ \mathbf{q}^{\text{VOL}}|_{\mathbf{1}} &= b_1 \int_{[0, \infty)} e^{\llbracket \mathbf{1} - x \mathbf{1} \rrbracket} \mu_1(dx) - b_2 \int_{[0, \infty)} \left(x \mathbf{2} + \frac{1}{2} K_2(0) b_2 \mathbf{\emptyset} \right) e^{\llbracket \mathbf{1} - x \mathbf{1} \rrbracket} \mu_2(dx), \\ \mathbf{p}^{\text{VOL}}|_{\mathbf{2}} &= \int_{[0, \infty)} \mu_2(dx) a_2 \mathbf{\emptyset} = K_2(0) a_2 \mathbf{\emptyset}, \quad \mathbf{p}^{\text{VOL}}|_{\mathbf{22}} = 0, \\ \mathbf{q}^{\text{VOL}}|_{\mathbf{2}} &= \int_{[0, \infty)} \mu_2(dx) b_2 \mathbf{\emptyset} = K_2(0) b_2 \mathbf{\emptyset}, \quad \mathbf{q}^{\text{VOL}}|_{\mathbf{22}} = 0. \end{aligned} \quad (5.3)$$

Then, using Lemma A.2 and noticing $(\mathbf{q}^{\text{VOL}})^\circ = 0$, it immediately follows that

$$\ell^{\text{VOL}}|_2 = \mathbf{p}^{\text{VOL}}|_2 + \ell^{\text{VOL}}(\mathbf{q}^{\text{VOL}}|_2) = K_2(0)(a_2\circ + b_2\ell^{\text{VOL}}), \quad (5.4)$$

$$\ell^{\text{VOL}}|_{22} = K_2(0)b_2\ell^{\text{VOL}}|_2 = K_2^2(0)b_2(a_2\circ + b_2\ell^{\text{VOL}}), \quad (5.5)$$

and that

$$\begin{aligned} \ell^{\text{VOL}}|_1 &= \mathbf{p}^{\text{VOL}}|_1 + \ell^{\text{VOL}}(\mathbf{q}^{\text{VOL}}|_1) \\ &= (a_1\circ + b_1\ell^{\text{VOL}}) \int_{[0,\infty)} e^{\llbracket -x\mathbf{1} \rrbracket} \mu_1(dx) - (a_2\circ + b_2\ell^{\text{VOL}}) \int_{[0,\infty)} \left(x\mathbf{2} + \frac{1}{2}K_2(0)b_2\circ \right) e^{\llbracket -x\mathbf{1} \rrbracket} \mu_2(dx) \\ &= K_1(0) (a_1\circ + b_1\ell^{\text{VOL}}) - \frac{1}{2}K_2^2(0)b_2 (a_2\circ + b_2\ell^{\text{VOL}}) \\ &\quad + (a_1\circ + b_1\ell^{\text{VOL}}) \mathbf{1} \int_{[0,\infty)} -xe^{\llbracket -x\mathbf{1} \rrbracket} \mu_1(dx) \\ &\quad + (a_2\circ + b_2\ell^{\text{VOL}}) \left(\mathbf{2} - \frac{1}{2}K_2(0)b_2\mathbf{1} \right) \int_{[0,\infty)} -xe^{\llbracket -x\mathbf{1} \rrbracket} \mu_2(dx). \end{aligned} \quad (5.6)$$

The following proposition establishes our first goal: $\ell^{\text{VOL}} \in \mathcal{I}(\widehat{W})$.

Lemma 5.1. *We have that $\{\ell^{\text{VOL}}, \ell^{\text{VOL}}|_1, \ell^{\text{VOL}}|_2, \ell^{\text{VOL}}|_{22}\} \subset \mathcal{A}_h$. In particular, $\ell^{\text{VOL}} \in \mathcal{I}(\widehat{W})$.*

Proof. At the beginning of Section 5, we have shown that $\ell^{\text{VOL}} \in \mathcal{A}_h$, then by (5.4)-(5.5), it follows that $\ell^{\text{VOL}}|_2, \ell^{\text{VOL}}|_{22} \in \mathcal{A}_h$. For $\ell^{\text{VOL}}|_1$, we remind that for $i = 1, 2$:

$$\int_{[0,\infty)} e^{\llbracket -x\mathbf{1} \rrbracket} \mu_i(dx) \preceq e^{\llbracket M\mathbf{1} \rrbracket}, \quad \int_{[0,\infty)} xe^{\llbracket -x\mathbf{1} \rrbracket} \mu_i(dx) \preceq Me^{\llbracket M\mathbf{1} \rrbracket}.$$

Combined it with (5.3) and the sub-multiplicative property, we derive $\mathbf{p}^{\text{VOL}}|_1, \mathbf{q}^{\text{VOL}}|_1 \in \mathcal{A}_h$. Recalling that $\ell^{\text{VOL}}|_1 = \mathbf{p}^{\text{VOL}}|_1 + \ell^{\text{VOL}}(\mathbf{q}^{\text{VOL}}|_1)$ in (5.6), an application of the sub-multiplicative property of the \mathcal{A}_h -norm in Proposition 3.7 yields that

$$\|\ell^{\text{VOL}}|_1\|_T^{\mathcal{A}_h} \leq \|\mathbf{p}^{\text{VOL}}|_1\|_T^{\mathcal{A}_h} + \|\ell^{\text{VOL}}\|_T^{\mathcal{A}_h} \|\mathbf{q}^{\text{VOL}}|_1\|_T^{\mathcal{A}_h} < \infty,$$

thus proving that $\ell^{\text{VOL}}|_1 \in \mathcal{A}_h$. Finally, by Proposition 3.6, for every $\ell \in \{\ell^{\text{VOL}}|_1, \ell^{\text{VOL}}|_2, \ell^{\text{VOL}}|_{22}\}$,

$$\mathbb{E} \left[\sup_{t \in [0, T]} \|\ell\|_t^{\mathcal{A}} \right] \leq \|\ell\|_T^{\mathcal{A}_h} < \infty.$$

This proves that $\sup_{t \in [0, T]} \|\ell\|_t^{\mathcal{A}} < \infty$ a.s. and thus ℓ^{VOL} satisfies the integral assumption in (3.8). This shows that $\ell^{\text{VOL}} \in \mathcal{I}(\widehat{W})$ and completes the proof. \square

Thanks to Lemma 5.1, $\ell^{\text{VOL}} \in \mathcal{I}(\widehat{W})$, so an application of Itô's formula in Theorem 3.3 and Remark 3.6 yields that the process $\bar{Y}_t := \langle \ell^{\text{VOL}}, \widehat{W}_t \rangle$ is an Itô process with decomposition

$$\bar{Y}_t = (\ell^{\text{VOL}})^\circ + \int_0^t \langle \ell^{\text{VOL}}|_1 + \frac{1}{2}\ell^{\text{VOL}}|_{22}, \widehat{W}_s \rangle ds + \int_0^t \langle \ell^{\text{VOL}}|_2, \widehat{W}_s \rangle dW_s.$$

It remains to achieve our second goal: identify the coefficients, which is the object of the next lemma.

Lemma 5.2. *We have that*

$$\begin{aligned} \langle \ell^{\text{VOL}}|_1 + \frac{1}{2}\ell^{\text{VOL}}|_{22}, \widehat{W}_t \rangle &= K_1(0) (a_1 + b_1\bar{Y}_t) \\ &\quad + \int_0^t K_1'(t-s) (a_1 + b_1\bar{Y}_s) ds + \int_0^t K_2'(t-s) (a_2 + b_2\bar{Y}_s) dW_s, \\ \langle \ell^{\text{VOL}}|_2, \widehat{W}_t \rangle &= K_2(0)(a_2 + b_2\bar{Y}_t). \end{aligned}$$

Proof. The proof is given in Subsection 5.1 below. \square

We are now ready to complete the proof of Theorem 4.2.

Proof of Theorem 4.2. Lemma 5.1 combined with Itô's formula in (4.5) yields that \bar{Y} , given in (5.1), is well-defined and satisfies (5.2), with $\ell^{\text{VOL}} \in \mathcal{A}_h$. An application of Lemma 5.2, after observing that $(\ell^{\text{VOL}})^\circ = y$, yields that \bar{Y}_t solves (5.2) as needed. This shows that $Y_t = \bar{Y}_t = \langle \ell^{\text{VOL}}, \widehat{\mathbb{W}}_t \rangle$ by strong uniqueness of the solution to (4.1) and ends the proof. \square

5.1 Proof of Lemma 5.2

We start by deriving a transformation formula, which is simple but crucial to simplify the algebraic expressions.

Lemma 5.3. *Let $\ell \in T((\mathbb{R}^d))$, $\mathbf{i} \in A_d$ and $\mathbf{b} := \sum_{\mathbf{j} \in A_d} b^{\mathbf{j}} \mathbf{j}$ for $b^{\mathbf{j}} \in \mathbb{R}$, then*

$$\ell \mathbf{i} e^{\sqcup \mathbf{b}} = e^{\sqcup \mathbf{b}} \sqcup ((e^{\sqcup -\mathbf{b}} \sqcup \ell) \mathbf{i}).$$

Proof. Let $\gamma := \ell \mathbf{i} e^{\sqcup \mathbf{b}} - e^{\sqcup \mathbf{b}} \sqcup ((e^{\sqcup -\mathbf{b}} \sqcup \ell) \mathbf{i})$, then

$$\begin{aligned} \gamma &= \ell \mathbf{i} (\phi + e^{\sqcup \mathbf{b}} \mathbf{b}) - (\phi + e^{\sqcup \mathbf{b}} \mathbf{b}) \sqcup ((e^{\sqcup -\mathbf{b}} \sqcup \ell) \mathbf{i}) \\ &= \ell \mathbf{i} + \ell \mathbf{i} e^{\sqcup \mathbf{b}} \mathbf{b} - [\phi \sqcup (e^{\sqcup -\mathbf{b}} \sqcup \ell)] \mathbf{i} - [(e^{\sqcup \mathbf{b}} \mathbf{b}) \sqcup (e^{\sqcup -\mathbf{b}} \sqcup \ell)] \mathbf{i} - [e^{\sqcup \mathbf{b}} \sqcup (e^{\sqcup -\mathbf{b}} \sqcup \ell) \mathbf{i}] \mathbf{b} \\ &= [\ell - (e^{\sqcup \mathbf{b}} \sqcup e^{\sqcup -\mathbf{b}} \sqcup \ell)] \mathbf{i} + [\ell \mathbf{i} e^{\sqcup \mathbf{b}} - e^{\sqcup \mathbf{b}} \sqcup ((e^{\sqcup -\mathbf{b}} \sqcup \ell) \mathbf{i})] \mathbf{b} \\ &= 0 + \gamma \mathbf{b}, \end{aligned}$$

implying that $\gamma = 0$. \square

Now, as we will be using integral of elements of $T((\mathbb{R}^2))$, recall the Definition 4.1, we need to make sure its bracket with the signature and concatenation make sense. This is the goal of the next two lemmas:

Lemma 5.4. *Let $\mathbf{p} \in L^1(T((\mathbb{R}^2)), \mu)$, then $\ell \mathbf{p}, \mathbf{p} \ell \in L^1(T((\mathbb{R}^2)), \mu)$ for every $\ell \in T((\mathbb{R}^2))$,*

$$\int_{[0, \infty)} \ell \mathbf{p}(x) \mu(dx) = \ell \left(\int_{[0, \infty)} \mathbf{p}(x) \mu(dx) \right)$$

and

$$\int_{[0, \infty)} \mathbf{p}(x) \ell \mu(dx) = \left(\int_{[0, \infty)} \mathbf{p}(x) \mu(dx) \right) \ell.$$

Proof. Obvious. \square

Lemma 5.5. *Let $\mathbf{p} \in L^1(T((\mathbb{R}^2)), \mu)$ such that $\int_{[0, \infty)} \|\mathbf{p}(x)\|_t^A \mu(dx) < \infty$ for every $t \in [0, T]$, then*

$$\left\langle \int_{[0, \infty)} \mathbf{p}(x) \mu(dx), \widehat{\mathbb{W}}_t \right\rangle = \int_{[0, \infty)} \langle \mathbf{p}(x), \widehat{\mathbb{W}}_t \rangle \mu(dx).$$

Proof. By assumption

$$\int_{[0, \infty)} \|\mathbf{p}(x)\|_t^A \mu(dx) = \int_{[0, \infty)} \sum_{n=0}^{\infty} \left| \sum_{\mathbf{v} \in V_n} \mathbf{p}^{\mathbf{v}}(x) \widehat{\mathbb{W}}_t^{\mathbf{v}} \right| \mu(dx) < \infty.$$

This implies that $\mathbf{p}(x) \in \mathcal{A}$ μ -a.e., so by the dominated convergence theorem

$$\begin{aligned} \left\langle \int_{[0, \infty)} \mathbf{p}(x) \mu(dx), \widehat{\mathbb{W}}_t \right\rangle &= \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \left(\int_{[0, \infty)} \mathbf{p}^{\mathbf{v}}(x) \mu(dx) \right) \widehat{\mathbb{W}}_t^{\mathbf{v}} \\ &= \int_{[0, \infty)} \sum_{n=0}^{\infty} \sum_{\mathbf{v} \in V_n} \mathbf{p}^{\mathbf{v}}(x) \widehat{\mathbb{W}}_t^{\mathbf{v}} \mu(dx) \\ &= \int_{[0, \infty)} \langle \mathbf{p}(x), \widehat{\mathbb{W}}_t \rangle \mu(dx). \end{aligned}$$

\square

We now introduce two important propositions, which establish the commutativity of the kernel integral and the bracket.

Proposition 5.6. *Let $\ell \in \mathcal{A}_h$. Assume that μ is a measure on $[0, \infty)$ satisfying $\int_{[0, \infty)} x^n \mu(dx) < M^n$ for every $n \in \mathbb{N}$ and a positive constant M . We then have*

$$\left\langle \boldsymbol{\ell} \mathbf{i} \left(\int_{[0, \infty)} x e^{\sqcup-x\mathbf{1}} \mu(dx) \right), \widehat{\mathbb{W}}_t \right\rangle = \left\langle \int_{[0, \infty)} \boldsymbol{\ell} \mathbf{i} x e^{\sqcup-x\mathbf{1}} \mu(dx), \widehat{\mathbb{W}}_t \right\rangle \quad (5.7)$$

$$= \int_{[0, \infty)} \left\langle \boldsymbol{\ell} \mathbf{i} x e^{\sqcup-x\mathbf{1}}, \widehat{\mathbb{W}}_t \right\rangle \mu(dx), \quad \mathbf{i} = \mathbf{1}, \mathbf{2}. \quad (5.8)$$

Proof. We start by proving the first equality (5.7). Fix $\mathbf{i} \in \{\mathbf{1}, \mathbf{2}\}$. By assumption of μ , it is easy to show that $\int_{[0, \infty)} x e^{\sqcup-x\mathbf{1}} \mu(dx) \in \mathcal{A}_h$ by direct computation. In particular $x e^{\sqcup-x\mathbf{1}} \in L^1(\mathcal{A}, \mu)$. A direct application of Lemma 5.4 shows that

$$\boldsymbol{\ell} \mathbf{i} \left(\int_{[0, \infty)} x e^{\sqcup-x\mathbf{1}} \mu(dx) \right) = \int_{[0, \infty)} \boldsymbol{\ell} \mathbf{i} x e^{\sqcup-x\mathbf{1}} \mu(dx) \in T((\mathbb{R}^2)).$$

Moreover, Proposition 3.7 ensures that it actually belongs to \mathcal{A}_h so that the bracket is well defined, which proves the first equality.

For the second equality (5.8), it suffices to verify that $\int_{[0, \infty)} \|\boldsymbol{\ell} \mathbf{i} x e^{\sqcup-x\mathbf{1}}\|_t^{\mathcal{A}} \mu(dx) < \infty$ for every $t \in [0, T]$:

$$\int_{[0, \infty)} \|\boldsymbol{\ell} \mathbf{i} x e^{\sqcup-x\mathbf{1}}\|_t^{\mathcal{A}} \mu(dx) \leq \int_{[0, \infty)} \|\boldsymbol{\ell} \mathbf{i} x e^{\sqcup-x\mathbf{1}}\|_t^{\mathcal{A}_h} \mu(dx) \leq \int_{[0, \infty)} \|\boldsymbol{\ell}\|_t^{\mathcal{A}_h} \|\mathbf{i}\|_t^{\mathcal{A}_h} \|x e^{\sqcup-x\mathbf{1}}\|_t^{\mathcal{A}_h} \mu(dx),$$

where the second equality comes from Proposition 3.7. It then follows

$$\begin{aligned} \int_{[0, \infty)} \|\boldsymbol{\ell} \mathbf{i} x e^{\sqcup-x\mathbf{1}}\|_t^{\mathcal{A}} \mu(dx) &\leq \|\boldsymbol{\ell}\|_t^{\mathcal{A}_h} \|\mathbf{i}\|_t^{\mathcal{A}_h} \int_{[0, \infty)} \|x e^{\sqcup-x\mathbf{1}}\|_t^{\mathcal{A}_h} \mu(dx) \\ &\leq C \|\boldsymbol{\ell}\|_t^{\mathcal{A}_h} \|\mathbf{i}\|_t^{\mathcal{A}_h} \int_{[0, \infty)} \left(\sum_{n=0}^{\infty} x^{n+1} (n+1)^{-\frac{1}{4}} \frac{(2t)^n}{\sqrt{(2n-1)!}} \right) \mu(dx) \\ &\leq C \|\boldsymbol{\ell}\|_t^{\mathcal{A}_h} \|\mathbf{i}\|_t^{\mathcal{A}_h} \sum_{n=0}^{\infty} M^{n+1} (n+1)^{-\frac{1}{4}} \frac{(2t)^n}{\sqrt{(2n-1)!}} \\ &\leq C \|\boldsymbol{\ell}\|_t^{\mathcal{A}_h} \|\mathbf{i}\|_t^{\mathcal{A}_h} M \left(\sum_{n=1}^{\infty} M^n \frac{(2t)^n}{(n-1)!} + 1 \right) \\ &= C \|\boldsymbol{\ell}\|_t^{\mathcal{A}_h} \|\mathbf{i}\|_t^{\mathcal{A}_h} M (2tM e^{2tM} + 1) \\ &< \infty. \end{aligned}$$

An application of Lemma 5.5 completes the proof. \square

Proposition 5.7. *μ is a measure on $[0, \infty)$ satisfying $\int_{[0, \infty)} x^n \mu(dx) < M^n$ for every $n \in \mathbb{N}$ and a positive constant M , then*

$$\int_{[0, \infty)} \left\langle \boldsymbol{\ell}^{\text{VOL}} \mathbf{1} x e^{\sqcup-x\mathbf{1}}, \widehat{\mathbb{W}}_t \right\rangle \mu(dx) = - \int_0^t J'(t-s) \bar{Y}_s ds \quad (5.9)$$

and

$$\int_{[0, \infty)} \left\langle \boldsymbol{\ell}^{\text{VOL}} \mathbf{2} x e^{\sqcup-x\mathbf{1}}, \widehat{\mathbb{W}}_t \right\rangle \mu(dx) = - \int_0^t J'(t-s) \bar{Y}_s dW_s - \frac{1}{2} K_2(0) \int_0^t J'(t-s) (a_2 + b_2 \bar{Y}_s) ds$$

where $J(t) = \int_{[0, \infty)} e^{-xt} \mu(dx)$.

Proof. By Lemma 5.3 and the shuffle property, we derive

$$\langle \ell^{\text{VOL}} \mathbf{1} x e^{\sqcup-x \mathbf{1}}, \widehat{\mathbb{W}}_t \rangle = x \langle e^{\sqcup-x \mathbf{1}}, \widehat{\mathbb{W}}_t \rangle \int_0^t \langle e^{\sqcup x \mathbf{1}} \sqcup \ell^{\text{VOL}}, \widehat{\mathbb{W}}_s \rangle ds = x \int_0^t e^{-x(t-s)} \bar{Y}_s ds, \quad (5.10)$$

Similarly, using Lemma 5.3 and the shuffle property,

$$\begin{aligned} \langle \ell^{\text{VOL}} \mathbf{2} x e^{\sqcup-x \mathbf{1}}, \widehat{\mathbb{W}}_t \rangle &= x \langle e^{\sqcup-x \mathbf{1}}, \widehat{\mathbb{W}}_t \rangle \langle (e^{\sqcup x \mathbf{1}} \sqcup \ell^{\text{VOL}}) \mathbf{2}, \widehat{\mathbb{W}}_t \rangle \\ &= x \langle e^{\sqcup-x \mathbf{1}}, \widehat{\mathbb{W}}_t \rangle \int_0^t \langle e^{\sqcup x \mathbf{1}} \sqcup \ell^{\text{VOL}}, \widehat{\mathbb{W}}_s \rangle dW_s \\ &\quad + \frac{1}{2} x \langle e^{\sqcup-x \mathbf{1}}, \widehat{\mathbb{W}}_t \rangle \int_0^t \langle e^{\sqcup x \mathbf{1}} \sqcup \ell^{\text{VOL}}|_{\mathbf{2}}, \widehat{\mathbb{W}}_s \rangle ds \\ &= x \int_0^t e^{-x(t-s)} \bar{Y}_s dW_s + \frac{1}{2} K_2(0) x \int_0^t e^{-x(t-s)} (a_2 + b_2 \bar{Y}_t) ds, \end{aligned}$$

where the second and third equalities come from Theorem 3.2 and (5.4) respectively.

We will now show that we can apply the deterministic and stochastic Fubini theorem. First, noticing that for all $t \in [0, T]$,

$$\mathbb{E} |\bar{Y}_t| = \mathbb{E} \left| \langle \ell^{\text{VOL}}, \widehat{\mathbb{W}}_t \rangle \right| \leq \| \ell^{\text{VOL}} \|_t^{\mathcal{A}_h},$$

and thus

$$\sup_{t \in [0, T]} \mathbb{E} |\bar{Y}_t| < \infty.$$

We can now apply Tonelli's theorem:

$$\begin{aligned} \mathbb{E} \left[\int_{[0, \infty)} \int_0^t e^{-x(t-s)} |\bar{Y}_s| ds \mu(dx) \right] &= \int_{[0, \infty)} \int_0^t e^{-x(t-s)} \mathbb{E} |\bar{Y}_s| ds \mu(dx) \\ &\leq \int_{[0, \infty)} \int_0^t \mathbb{E} |\bar{Y}_s| ds \mu(dx) \\ &\leq \mu([0, \infty)) \int_0^t \mathbb{E} |\bar{Y}_s| ds \\ &\leq \mu([0, \infty)) T \sup_{t \in [0, T]} \mathbb{E} |\bar{Y}_t| < \infty. \end{aligned}$$

This leads to

$$\int_{[0, \infty)} \int_0^t e^{-x(t-s)} |\bar{Y}_s| ds \mu(dx) < \infty \text{ a.s..}$$

Now we can apply Fubini theorem on equation (5.10) and derive:

$$\int_{[0, \infty)} \langle \ell^{\text{VOL}} \mathbf{1} x e^{\sqcup-x \mathbf{1}}, \widehat{\mathbb{W}}_t \rangle \mu(dx) = \int_0^t \left(\int_{[0, \infty)} x e^{-x(t-s)} \mu(dx) \right) \bar{Y}_s ds = - \int_0^t J'(t-s) \bar{Y}_s ds,$$

thus proving (5.9). On the other hand, by Proposition 3.6 (ii),

$$\sup_{t \in [0, T]} \mathbb{E} [\bar{Y}_t^2] \leq \left(\| \ell^{\text{VOL}} \|_T^{\mathcal{A}_h} \right)^2 < \infty$$

and

$$\begin{aligned} \mathbb{E} \left[\int_{[0, \infty)} \sqrt{\int_0^t e^{-2x(t-s)} |\bar{Y}_s|^2 ds} \mu(dx) \right] &\leq \mu([0, \infty)) \mathbb{E} \left[\sqrt{\int_0^t |\bar{Y}_s|^2 ds} \right] \\ &\leq \mu([0, \infty)) \sqrt{\mathbb{E} \left[\int_0^t |\bar{Y}_s|^2 ds \right]} \\ &\leq \mu([0, \infty)) \sqrt{T \sup_{t \in [0, T]} \mathbb{E} |\bar{Y}_t|^2} < \infty, \end{aligned}$$

which leads to

$$\int_{[0,\infty)} \sqrt{\int_0^t e^{-2x(t-s)} |\bar{Y}_s|^2 ds} \mu(dx) < \infty \text{ a.s.} \quad (5.11)$$

Note that (5.11) is actually the condition for stochastic Fubini theorem to hold, see [30, Theorem 2.2]. Applying it, we derive

$$\int_{[0,\infty)} \left\langle \ell^{\text{VOL}} \mathbf{2} x e^{\llbracket -x \mathbf{1} \rrbracket}, \widehat{\mathbb{W}}_t \right\rangle \mu(dx) = - \int_0^t J'(t-s) \bar{Y}_s dW_s - \frac{1}{2} K_2(0) \int_0^t J'(t-s) (a_2 + b_2 \bar{Y}_s) ds,$$

thus completing the proof. \square

Now, taking $\mu = \mu_1, \mu_2$, we combine Proposition 5.6 with Proposition 5.7 and derive, for $i = 1, 2$:

$$\left\langle \ell^{\text{VOL}} \mathbf{1} \int_{[0,\infty)} x e^{\llbracket -x \mathbf{1} \rrbracket} \mu_i(dx), \widehat{\mathbb{W}}_t \right\rangle = - \int_0^t K'_i(t-s) \bar{Y}_s ds. \quad (5.12)$$

$$\left\langle \ell^{\text{VOL}} \mathbf{2} \int_{[0,\infty)} x e^{\llbracket -x \mathbf{1} \rrbracket} \mu_i(dx), \widehat{\mathbb{W}}_t \right\rangle = - \int_0^t K'_i(t-s) \bar{Y}_s dW_s - \frac{1}{2} K_2(0) \int_0^t K'_i(t-s) (a_2 + b_2 \bar{Y}_s) ds. \quad (5.13)$$

Proof of Lemma 5.2. Following directly from (5.4)-(5.5), we get

$$\left\langle \ell^{\text{VOL}}|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \right\rangle = K_2(0) (a_2 + b_2 \bar{Y}_t) \quad \text{and} \quad \left\langle \ell^{\text{VOL}}|_{\mathbf{2}\mathbf{2}}, \widehat{\mathbb{W}}_t \right\rangle = K_2^2(0) b_2 (a_2 + b_2 \bar{Y}_t).$$

It now only remains to compute $\left\langle \ell^{\text{VOL}}|_{\mathbf{1}}, \widehat{\mathbb{W}}_t \right\rangle$. For this, recall (5.6)

$$\begin{aligned} \ell^{\text{VOL}}|_{\mathbf{1}} &= K_1(0) (a_1 \boldsymbol{\phi} + b_1 \ell^{\text{VOL}}) - \frac{1}{2} K_2^2(0) b_2 (a_2 \boldsymbol{\phi} + b_2 \ell^{\text{VOL}}) \\ &\quad + (a_1 \boldsymbol{\phi} + b_1 \ell^{\text{VOL}}) \mathbf{1} \int_{[0,\infty)} -x e^{\llbracket -x \mathbf{1} \rrbracket} \mu_1(dx) \\ &\quad + (a_2 \boldsymbol{\phi} + b_2 \ell^{\text{VOL}}) \left(\mathbf{2} - \frac{1}{2} K_2(0) b_2 \mathbf{1} \right) \int_{[0,\infty)} -x e^{\llbracket -x \mathbf{1} \rrbracket} \mu_2(dx). \end{aligned}$$

A direct application of equation (5.12) and (5.13) yields:

$$\begin{aligned} \left\langle \ell^{\text{VOL}}|_{\mathbf{1}}, \widehat{\mathbb{W}}_t \right\rangle &= K_1(0) (a_1 + b_1 \bar{Y}_t) - \frac{1}{2} K_2^2(0) b_2 (a_2 + b_2 \bar{Y}_t) \\ &\quad + \int_0^t K'_1(t-s) (a_1 + b_1 \bar{Y}_s) ds + \int_0^t K'_2(t-s) (a_2 + b_2 \bar{Y}_s) dW_s, \end{aligned}$$

which concludes the proof. \square

6 Proof of Theorem 4.4

We first recall from Lemma A.2 that for all $\mathbf{p}, \mathbf{q} \in T((\mathbb{R}^d))$ such that $\mathbf{q}^\boldsymbol{\phi} = 0$,

$$\left(\mathbf{p}(\boldsymbol{\phi} - \mathbf{q})^{-1} \right) |_{\mathbf{i}} = \mathbf{p}|_{\mathbf{i}} + \mathbf{p}(\boldsymbol{\phi} - \mathbf{q})^{-1} (\mathbf{q}|_{\mathbf{i}}).$$

Returning to the definition of \mathbf{q}^{DE} , \mathbf{p}^{DE} and ℓ^{DE} :

$$\begin{aligned} \ell^{\text{DE}} &= \mathbf{p}^{\text{DE}} (\boldsymbol{\phi} - \mathbf{q}^{\text{DE}})^{-1}, \\ \mathbf{p}^{\text{DE}} &= (a_1 - \frac{1}{2} b_2 a_2) \mathbf{1} + a_2 \mathbf{2} + z \boldsymbol{\phi}, \\ \mathbf{q}^{\text{DE}} &= (b_1 - \frac{1}{2} b_2^2) \mathbf{1} + b_2 \mathbf{2} + \left(\sum_{m=1}^{k_1} c_1^m \mathbf{1} e^{\llbracket \alpha_1^m \mathbf{1} \rrbracket} - \frac{1}{2} b_2 \sum_{m=1}^{k_2} c_2^m \mathbf{1} e^{\llbracket \alpha_2^m \mathbf{1} \rrbracket} \right) \mathbf{1} + \left(\sum_{m=1}^{k_2} c_2^m \mathbf{1} e^{\llbracket \alpha_2^m \mathbf{1} \rrbracket} \right) \mathbf{2}, \end{aligned}$$

it follows that

$$\begin{aligned}\ell^{\text{DE}}|_{\mathbf{1}} &= \ell^{\text{DE}} \left((b_1 - \frac{1}{2}b_2^2) \boldsymbol{\vartheta} + \sum_{m=1}^{k_1} c_1^m \mathbf{1} e^{\llcorner \alpha_1^m \mathbf{1}} - \frac{1}{2}b_2 \sum_{m=1}^{k_2} c_2^m \mathbf{1} e^{\llcorner \alpha_2^m \mathbf{1}} \right) + (a_1 - \frac{1}{2}b_2 a_2) \boldsymbol{\vartheta}, \quad (6.1) \\ \ell^{\text{DE}}|_{\mathbf{2}} &= \ell^{\text{DE}} \left(b_2 \boldsymbol{\vartheta} + \sum_{m=1}^{k_2} c_2^m \mathbf{1} e^{\llcorner \alpha_2^m \mathbf{1}} \right) + a_2 \boldsymbol{\vartheta}, \\ \ell^{\text{DE}}|_{\mathbf{22}} &= b_2 \ell^{\text{DE}}|_{\mathbf{2}}.\end{aligned}$$

Finally,

$$\ell^{\text{DE}}|_{\mathbf{1}} + \frac{1}{2} \ell^{\text{DE}}|_{\mathbf{22}} = \ell^{\text{DE}} \left(b_1 \boldsymbol{\vartheta} + \sum_{m=1}^{k_1} c_1^m \mathbf{1} e^{\llcorner \alpha_1^m \mathbf{1}} \right) + a_1 \boldsymbol{\vartheta}.$$

Lemma 6.1. *We have that $\{\ell^{\text{DE}}, \ell^{\text{DE}}|_{\mathbf{1}}, \ell^{\text{DE}}|_{\mathbf{2}}, \ell^{\text{DE}}|_{\mathbf{22}}\} \subset \mathcal{A}_h$. In particular, $\ell^{\text{DE}} \in \mathcal{I}(\widehat{W})$.*

Proof. First, observe that $\mathbf{i} e^{\llcorner \gamma \mathbf{i}} \preceq \mathbf{i} e^{\llcorner \hat{\gamma} \mathbf{i}} = \frac{1}{\hat{\gamma}} (e^{\llcorner \hat{\gamma} \mathbf{i}} - \boldsymbol{\vartheta})$ where $\hat{\gamma} := \max(|\gamma|, 1)$ and \preceq is defined in Definition 3.8. Thus we obtain that

$$\begin{aligned}\mathbf{q}^{\text{DE}} &\preceq \left| b_1 - \sum_{m=1}^{k_1} \frac{c_1^m}{\hat{\alpha}_1^m} - \frac{1}{2}b_2 \left(b_2 - \sum_{m=1}^{k_2} \frac{c_2^m}{\hat{\alpha}_2^m} \right) \right| \mathbf{1} + \left| b_2 - \sum_{m=1}^{k_2} \frac{c_2^m}{\hat{\alpha}_2^m} \right| \mathbf{2} \\ &\quad + \sum_{m=1}^{k_1} \left| \frac{c_1^m}{\hat{\alpha}_1^m} e^{\llcorner \hat{\alpha}_1^m \mathbf{1}} - \frac{1}{2}b_2 \frac{c_2^m}{\hat{\alpha}_2^m} e^{\llcorner \hat{\alpha}_2^m \mathbf{1}} \right| \mathbf{1} + \sum_{m=1}^{k_2} \left| \frac{c_2^m}{\hat{\alpha}_2^m} e^{\llcorner \hat{\alpha}_2^m \mathbf{1}} \right| \mathbf{2},\end{aligned}$$

where $\hat{\alpha}_i^m := \max(|\alpha_i^m|, 1)$, $i = 1, 2$. We can then further dominate it with

$$\mathbf{q}^{\text{DE}} \preceq M e^{\llcorner (\sum_{m=1}^{k_1} \hat{\alpha}_1^m + \sum_{m=1}^{k_2} \hat{\alpha}_2^m) \mathbf{1}} (\mathbf{1} + \mathbf{2}),$$

where

$$M = \left| b_1 - \sum_{m=1}^{k_1} \frac{c_1^m}{\hat{\alpha}_1^m} - \frac{1}{2}b_2 \left(b_2 - \sum_{m=1}^{k_2} \frac{c_2^m}{\hat{\alpha}_2^m} \right) \right| + \left| b_2 - \sum_{m=1}^{k_2} \frac{c_2^m}{\hat{\alpha}_2^m} \right| + \left| \sum_{m=1}^{k_1} \frac{c_1^m}{\hat{\alpha}_1^m} - \frac{1}{2}b_2 \frac{c_2^m}{\hat{\alpha}_2^m} \right| + \left| \sum_{m=1}^{k_2} \frac{c_2^m}{\hat{\alpha}_2^m} \right|.$$

Thus by Theorem 3.9

$$\sup_{t \in [0, T]} \|\ell^{\text{DE}}\|_t^{A_h} < \infty.$$

With representation (6.1), we apply Proposition 3.7 and proving $\{\ell^{\text{DE}}, \ell^{\text{DE}}|_{\mathbf{1}}, \ell^{\text{DE}}|_{\mathbf{2}}, \ell^{\text{DE}}|_{\mathbf{22}}\} \subset \mathcal{A}_h$. Similarly to lemma 5.1, we directly derive that $\ell^{\text{DE}} \in \mathcal{I}(\widehat{W})$. \square

We are now ready to apply Theorem 3.3 on ℓ^{DE} and get

$$\begin{aligned}\langle \ell^{\text{DE}}, \widehat{W}_t \rangle &= (\ell^{\text{DE}})^{\boldsymbol{\vartheta}} + \int_0^t \langle \ell^{\text{DE}}|_{\mathbf{1}} + \frac{1}{2} \ell^{\text{DE}}|_{\mathbf{22}}, \widehat{W}_s \rangle ds + \int_0^t \langle \ell^{\text{DE}}|_{\mathbf{2}}, \widehat{W}_s \rangle dW_s \\ &= z + \int_0^t \left\langle \ell^{\text{DE}} \left(b_1 \boldsymbol{\vartheta} + \sum_{m=1}^{k_1} c_1^m \mathbf{1} e^{\llcorner \alpha_1^m \mathbf{1}} \right) + a_1 \boldsymbol{\vartheta}, \widehat{W}_s \right\rangle ds \\ &\quad + \int_0^t \left\langle \ell^{\text{DE}} \left(b_2 \boldsymbol{\vartheta} + \sum_{m=1}^{k_2} c_2^m \mathbf{1} e^{\llcorner \alpha_2^m \mathbf{1}} \right) + a_2 \boldsymbol{\vartheta}, \widehat{W}_s \right\rangle dW_s \\ &= z + \int_0^t \left(a_1 + b_1 \langle \ell^{\text{DE}}, \widehat{W}_s \rangle + \sum_{m=1}^{k_1} c_1^m \langle \ell^{\text{DE}} \mathbf{1} e^{\llcorner \alpha_1^m \mathbf{1}}, \widehat{W}_s \rangle \right) ds \\ &\quad + \int_0^t \left(a_2 + b_2 \langle \ell^{\text{DE}}, \widehat{W}_s \rangle + \sum_{m=1}^{k_2} c_2^m \langle \ell^{\text{DE}} \mathbf{1} e^{\llcorner \alpha_2^m \mathbf{1}}, \widehat{W}_s \rangle \right) dW_s.\end{aligned}$$

Finally, using Lemma 5.3, i.e. $\ell \mathbf{i} e^{\llbracket \gamma \mathbf{j} \rrbracket} = e^{\llbracket \gamma \mathbf{j} \rrbracket} \llbracket (e^{\llbracket -\gamma \mathbf{j} \rrbracket} \llbracket \ell \rrbracket \mathbf{i}) \rrbracket$ for all $\gamma \in \mathbb{R}$ and $\mathbf{i}, \mathbf{j} \in A_d$, and the fact that $\ell^{\text{DE}} \mathbf{1} e^{\llbracket \alpha_i^m \mathbf{1} \rrbracket} \in \mathcal{A}$ for all m and $i = 1, 2$, it is then straightforward to see that

$$\left\langle \ell^{\text{DE}} \mathbf{1} e^{\llbracket \alpha_i^m \mathbf{1} \rrbracket}, \widehat{\mathbb{W}}_t \right\rangle = \int_0^t e^{\alpha_i^m(t-s)} \left\langle \ell^{\text{DE}}, \widehat{\mathbb{W}}_s \right\rangle ds.$$

This concludes the proof of Theorem 4.4 that $\left\langle \ell^{\text{DE}}, \widehat{\mathbb{W}}_t \right\rangle$ is solution to (4.8) and hence $Z_t = \left\langle \ell^{\text{DE}}, \widehat{\mathbb{W}}_t \right\rangle$ by strong uniqueness.

7 Proof of Theorem 4.5

We first illustrate that $\left\langle \ell_t^{\text{GV}}, \widehat{\mathbb{W}}_t \right\rangle$ with ℓ_t^{GV} defined in (4.12) is a well-defined stochastic process.

$$\begin{aligned} \mathbb{E} \left[\int_0^T \sum_{n=0}^{\infty} \left| K^{(n)}(t) \left\langle \mathbf{1}^{\otimes n} \mathbf{2}, \widehat{\mathbb{W}}_t \right\rangle \right| dt \right] &= \int_0^T \sum_{n=0}^{\infty} \left| K^{(n)}(t) \right| \mathbb{E} \left[\left| \left\langle \mathbf{1}^{\otimes n} \mathbf{2}, \widehat{\mathbb{W}}_t \right\rangle \right| \right] dt \\ &= \int_0^T \sum_{n=0}^{\infty} \left| K^{(n)}(t) \right| \mathbb{E} \left[\left| \int_0^t \frac{s^n}{n!} dW_s \right| \right] dt \\ &\leq \int_0^T \sum_{n=0}^{\infty} \left| K^{(n)}(t) \right| \frac{t^{n+\frac{1}{2}}}{n! \sqrt{2n+1}} dt \end{aligned}$$

The last inequality is derived by Jensen inequality. Then by assumption we know that

$$\mathbb{E} \left[\int_0^T \sum_{n=0}^{\infty} \left| K^{(n)}(t) \left\langle \mathbf{1}^{\otimes n} \mathbf{2}, \widehat{\mathbb{W}}_t \right\rangle \right| dt \right] < \infty,$$

thus

$$\sum_{n=0}^{\infty} \left| K^{(n)}(t) \left\langle \mathbf{1}^{\otimes n} \mathbf{2}, \widehat{\mathbb{W}}_t \right\rangle \right| < \infty, \mathbb{P}(d\omega) \otimes dt - a.e.$$

Consequently, $\sum_{n=0}^{\infty} K^{(n)}(t) \left\langle \mathbf{1}^{\otimes n} \mathbf{2}, \widehat{\mathbb{W}}_t \right\rangle$ converges $\mathbb{P}(d\omega) \otimes dt - a.e.$, ensuring $\left\langle \ell_t^{\text{GV}}, \widehat{\mathbb{W}}_t \right\rangle$ to be a $\mathbb{P}(d\omega) \otimes dt - a.e.$ -well-defined map. Notice that for every $N \in \mathbb{N}$, $\sum_{n=0}^N K^{(n)}(t) \left\langle \mathbf{1}^{\otimes n} \mathbf{2}, \widehat{\mathbb{W}}_t \right\rangle$ is progressively measurable, so $\left\langle \ell_t^{\text{GV}}, \widehat{\mathbb{W}}_t \right\rangle$ is also progressively measurable. It remains to prove that for every fixed t ,

$$\left\langle \ell_t^{\text{GV}}, \widehat{\mathbb{W}}_t \right\rangle = \int_0^t K(t-s) dW_s. \quad (7.1)$$

For this we write

$$\left\langle \ell_t^{\text{GV}}, \widehat{\mathbb{W}}_t \right\rangle = \sum_{n=0}^{\infty} K^{(n)}(t) (-1)^n \left\langle \mathbf{1}^{\otimes n} \mathbf{2}, \widehat{\mathbb{W}}_t \right\rangle = \sum_{n=0}^{\infty} K^{(n)}(t) \int_0^t \frac{(-s)^n}{n!} dW_s.$$

Since $\int_0^t \left(\sum_{n=0}^{\infty} \left| K^{(n)}(t) \right| \frac{s^n}{n!} \right)^2 ds < \infty$ by assumption (4.11), the dominated convergence theorem of stochastic integrals yields

$$\sum_{n=0}^N K^{(n)}(t) \int_0^t \frac{(-s)^n}{n!} dW_s \longrightarrow \int_0^t \sum_{n=0}^{\infty} K^{(n)}(t) \frac{(-s)^n}{n!} dW_s = \int_0^t K(t-s) dW_s,$$

in probability, as $N \rightarrow \infty$. While the previous analysis ensures that

$$\sum_{n=0}^N K^{(n)}(t) \int_0^t \frac{(-s)^n}{n!} dW_s \xrightarrow{\text{a.s.}} \sum_{n=0}^{\infty} K^{(n)}(t) \int_0^t \frac{(-s)^n}{n!} dW_s = \left\langle \ell_t^{\text{GV}}, \widehat{\mathbb{W}}_t \right\rangle,$$

which proves (7.1) and completes the proof.

A Properties of the resolvent

Lemma A.1. For $\ell \in T((\mathbb{R}^d))$ such that $|\ell^\emptyset| < 1$, $\sum_{n=0}^{\infty} \ell^{\otimes n}$ is well defined, i.e. $\sum_{n=0}^{\infty} \langle \ell^{\otimes n}, \mathbf{v} \rangle < \infty$ for every $\mathbf{v} \in V$.

Proof. Let $m \in \mathbb{N}$ and $\mathbf{v} \in V_m$ a word of length m . Then,

$$\sup_{n \in \mathbb{N}} \sup_{\substack{\mathbf{w}_1, \dots, \mathbf{w}_n \in V \setminus \{\emptyset\} \\ \mathbf{w}_1 \cdots \mathbf{w}_n = \mathbf{v}}} |\ell^{\mathbf{w}_1} \cdots \ell^{\mathbf{w}_n}| < \infty,$$

since the supremum is taken over a finite set. Consequently let us define

$$M(\mathbf{v}) := \sup_{n \in \mathbb{N}} \sum_{\substack{\mathbf{w}_1, \dots, \mathbf{w}_n \in V \setminus \{\emptyset\} \\ \mathbf{w}_1 \cdots \mathbf{w}_n = \mathbf{v}}} |\ell^{\mathbf{w}_1} \cdots \ell^{\mathbf{w}_n}| < \infty.$$

We then have for all $n \in \mathbb{N}$ that

$$\begin{aligned} |\langle \ell^{\otimes n}, \mathbf{v} \rangle| &\leq \sum_{\substack{\mathbf{w}_1, \dots, \mathbf{w}_n \in V \\ \mathbf{w}_1 \cdots \mathbf{w}_n = \mathbf{v}}} |\ell^{\mathbf{w}_1} \cdots \ell^{\mathbf{w}_n}| \\ &\leq \sum_{k=0}^n \binom{n}{k} |\ell^\emptyset|^k \left(\sum_{\substack{\mathbf{w}_1, \dots, \mathbf{w}_{n-k} \in V \setminus \{\emptyset\} \\ \mathbf{w}_1 \cdots \mathbf{w}_{n-k} = \mathbf{v}}} |\ell^{\mathbf{w}_1} \cdots \ell^{\mathbf{w}_{n-k}}| \right) \\ &\leq M(\mathbf{v}) \sum_{k=(n-m)^+}^n \binom{n}{k} |\ell^\emptyset|^k. \end{aligned}$$

The last inequality follows from the fact that the set $\{\mathbf{w}_1, \dots, \mathbf{w}_n \in V \setminus \{\emptyset\} : \mathbf{w}_1 \cdots \mathbf{w}_n = \mathbf{v}\}$ is empty when $n - k > m$. Recall m is the size of the word \mathbf{v} . Therefore

$$\sum_{n=0}^{\infty} |\langle \ell^{\otimes n}, \mathbf{v} \rangle| \leq M(\mathbf{v}) \sum_{n=0}^{\infty} \sum_{k=(n-m)^+}^n \binom{n}{k} |\ell^\emptyset|^k = M(\mathbf{v}) \sum_{n=0}^{m-1} \sum_{k=0}^n \binom{n}{k} |\ell^\emptyset|^k + M(\mathbf{v}) \sum_{n=m}^{\infty} \sum_{k=n-m}^n \binom{n}{k} |\ell^\emptyset|^k.$$

While the first part of the right-hand side is obviously finite, the second needs more work, i.e. using Pascal's rule and the upper bound for the binomial coefficients we have

$$\begin{aligned} \sum_{n=m}^{\infty} \sum_{k=n-m}^n \binom{n}{k} |\ell^\emptyset|^k &= \sum_{n=m}^{\infty} |\ell^\emptyset|^{n-m} \sum_{k=0}^m \binom{n}{n-m+k} |\ell^\emptyset|^k \\ &= \sum_{n=m}^{\infty} |\ell^\emptyset|^{n-m} \sum_{k=0}^m \binom{n}{n-m+k} \\ &\leq |\ell^\emptyset|^{-m} \sum_{n=m}^{\infty} |\ell^\emptyset|^n \binom{n+m}{n}, \end{aligned}$$

where the first inequality follows from the assumption that $|\ell^\emptyset| < 1$. We can finally remark that the right-hand side has the same asymptotics as $\sum_{n=0}^{\infty} n^m |\ell^\emptyset|^n$, which is finite, and conclude the proof. \square

Lemma A.2. Let $\mathbf{p}, \mathbf{q} \in T((\mathbb{R}^d))$ such that $|\mathbf{q}^\emptyset| < 1$ and $\mathbf{i} \in A_d$. Then

$$\left(\mathbf{p}(\emptyset - \mathbf{q})^{-1} \right) |_{\mathbf{i}} = \frac{1}{1 - \mathbf{q}^\emptyset} \left[\mathbf{p} |_{\mathbf{i}} + \mathbf{p}(\emptyset - \mathbf{q})^{-1}(\mathbf{q} |_{\mathbf{i}}) \right].$$

Proof. First recall the decomposition in Remark 2.1, i.e.

$$\mathbf{q} = \mathbf{q}^\emptyset \emptyset + \sum_{\mathbf{i} \in A_d} \mathbf{q} |_{\mathbf{i}},$$

and Proposition 2.2, i.e.

$$(\mathfrak{o} - \mathfrak{q})^{-1} = \mathfrak{o} + (\mathfrak{o} - \mathfrak{q})^{-1} \mathfrak{q}.$$

One can then easily get that

$$(\mathfrak{o} - \mathfrak{q})^{-1} = \mathfrak{o} + (\mathfrak{o} - \mathfrak{q})^{-1} \left(\mathfrak{q}^{\mathfrak{o}} \mathfrak{o} + \sum_{\mathbf{i} \in A_d} \mathfrak{q} |\mathbf{i}| \right) = \frac{1}{1 - \mathfrak{q}^{\mathfrak{o}}} \left(\mathfrak{o} + (\mathfrak{o} - \mathfrak{q})^{-1} \sum_{\mathbf{i} \in A_d} \mathfrak{q} |\mathbf{i}| \right).$$

It then follows that

$$\mathfrak{p}(\mathfrak{o} - \mathfrak{q})^{-1} = \frac{1}{1 - \mathfrak{q}^{\mathfrak{o}}} \left(\mathfrak{p} + \mathfrak{p}(\mathfrak{o} - \mathfrak{q})^{-1} \sum_{\mathbf{i} \in A_d} \mathfrak{q} |\mathbf{i}| \right).$$

Taking the projection ends the proof. \square

B Proof of Proposition 3.5

In this section, we are specifically interested in the case $X_t = \widehat{W}_t = (t, W_t^2, \dots, W_t^d)$ where $W = (W^2, \dots, W^d)$ is a $(d-1)$ -dimensional Brownian motion. We will denote $n(\mathbf{v})$ the size of \mathbf{v} , i.e. its number of letters, and $x(\mathbf{v})$ the number of $\mathbf{1}$ s specifically. If it does not cause ambiguity, we will write n for $n(\mathbf{v})$ and x for $x(\mathbf{v})$.

Let $g^p : [0, T] \times V \rightarrow \mathbb{R}$ be defined for all even $p \in \mathbb{N}^*$ by

$$g_t^p(\mathbf{v}) = \left[\prod_{j=2}^p \binom{jn}{n} \right] \frac{t^{\frac{p}{2}(n+x)}}{\left(\frac{p}{2}(n+x)\right)!} 2^{\frac{p}{2}(x-n)}$$

and for all odd $p \in \mathbb{N}$ by

$$g_t^p(\mathbf{v}) = \left(g_t^{p+1}(\mathbf{v}) \right)^{\frac{p}{p+1}},$$

for all $t \in [0, T]$ and all $\mathbf{v} \in V$. Notice that $g_t^p(\mathbf{v})$ only depends on \mathbf{v} through its size and the number of its $\mathbf{1}$ s, i.e. $n(\mathbf{v})$ and $x(\mathbf{v})$.

Proposition B.1. Fix $t \in [0, T]$. We have

$$\mathbb{E} \left[\left| \left\langle \mathbf{v}, \widehat{W}_t \right\rangle \right|^p \right] \leq g_t^p(\mathbf{v}), \quad \mathbf{v} \in V, \quad p \in \mathbb{N}.$$

Proof. We first use Fawcett's formula [25, Proposition 4.10] to write

$$\mathbb{E} \left[\widehat{W}_t \right] = \sum_{m=0}^{\infty} \frac{t^m}{m!} \left(\mathbf{1} + \frac{1}{2} \sum_{\mathbf{j} \in \{2, \dots, d\}} \mathbf{j} \right)^{\otimes m}.$$

We then define for all $n \in \mathbb{N}$

$$\widehat{V}_n := \{ \mathbf{v} \in V_n : \exists j \in \mathbb{N} \text{ such that } \mathbf{v} = \mathbf{u}_1 \cdots \mathbf{u}_j \text{ for some } \mathbf{u}_k \in \{ \mathbf{1}, \mathbf{22}, \dots, \mathbf{d} \} \text{ with } k = 1, 2, \dots, j \}.$$

This set corresponds to the coordinate of every signature element with non-zero expectation, i.e. for every $n \in \mathbb{N}$ and $\mathbf{v} \in V_n$,

$$\mathbb{E} \left[\left\langle \mathbf{v}, \widehat{W}_t \right\rangle \right] = \begin{cases} \frac{t^{\frac{1}{2}(n+x)}}{\left(\frac{1}{2}(n+x)\right)!} \cdot 2^{\frac{1}{2}(x-n)} & \text{if } \mathbf{v} \in \widehat{V}_n, \\ 0 & \text{else.} \end{cases}$$

We are now ready to compute even moments of every signature element. First, remark that for all $\mathbf{u} \in V_k$ and $\mathbf{v} \in V_m$, $\mathbf{u} \sqcup \mathbf{v}$ is a sum of $\binom{k+m}{k}$ elements $\mathbf{w} \in V_{k+m}$ of length $n(\mathbf{w}) = k + m$ each, having exactly $x(\mathbf{w}) = x(\mathbf{u}) + x(\mathbf{v})$ number of $\mathbf{1}$ s, so that

$$\mathbb{E} \left[\left\langle \mathbf{u} \sqcup \mathbf{v}, \widehat{W}_t \right\rangle \right] \leq \binom{k+m}{k} \max_{\substack{\mathbf{w} \in V_{k+m} \\ \text{s.t. } x(\mathbf{w}) = x(\mathbf{u}) + x(\mathbf{v})}} \mathbb{E} \left\langle \mathbf{w}, \widehat{W}_t \right\rangle.$$

Consequently, for every $p \in \mathbb{N}$, we can iteratively apply the shuffle product p -times, and obtain the claimed upper bound for the even powers of the signature:

$$\mathbb{E} \left[\left\langle \mathbf{v}, \widehat{\mathbb{W}}_t \right\rangle^{2p} \right] = \mathbb{E} \left\langle \mathbf{v} \sqcup^{2p}, \widehat{\mathbb{W}}_t \right\rangle \leq \left[\prod_{j=2}^{2p} \binom{jn}{n} \right] \max_{\substack{\mathbf{w} \in V_{2pn} \\ \text{s.t. } x(\mathbf{w})=2px(\mathbf{v})}} \mathbb{E} \left\langle \mathbf{w}, \widehat{\mathbb{W}}_t \right\rangle = \left[\prod_{j=2}^{2p} \binom{jn}{n} \right] \frac{t^{p(n+x)}}{(p(n+x))!} 2^{p(x-n)}.$$

We can finally treat the odd moments by applying Jensen's inequality on the convex function $y \rightarrow y^{\frac{2p}{2p-1}}$ on \mathbb{R}^+ :

$$\mathbb{E} \left[\left| \left\langle \mathbf{v}, \widehat{\mathbb{W}}_t \right\rangle \right|^{2p-1} \right]^{\frac{1}{2p-1}} \leq \mathbb{E} \left[\left\langle \mathbf{v}, \widehat{\mathbb{W}}_t \right\rangle^{2p} \right]^{\frac{1}{2p}} \leq \left(h_t^{2p}(\mathbf{v}) \right)^{\frac{1}{2p}} = \left(h_t^{2p-1}(\mathbf{v}) \right)^{\frac{1}{2p-1}},$$

concluding the proof of the proposition. \square

We are now ready to prove Proposition 3.5:

Proof of Proposition 3.5. Proposition 3.5-(i) follows directly from an application of the above Proposition B.1. In order to prove (3.11) we will consider two cases: either when word \mathbf{v} ends with a $\mathbf{1}$ or it ends with a \mathbf{j} in $\{\mathbf{2}, \dots, \mathbf{d}\}$.

(1) **Case $\mathbf{v} = \mathbf{u}\mathbf{1}$ for some word $\mathbf{u} \in V$.** We apply Cauchy-Schwarz inequality to get

$$\begin{aligned} \mathbb{E} \left[\sup_{s \in [0, t]} \left\langle \mathbf{v}, \widehat{\mathbb{W}}_s \right\rangle^2 \right] &= \mathbb{E} \left[\sup_{s \in [0, t]} \left(\int_0^s \left\langle \mathbf{u}, \widehat{\mathbb{W}}_u \right\rangle du \right)^2 \right] \\ &\leq \mathbb{E} \left[\left(\int_0^t \left| \left\langle \mathbf{u}, \widehat{\mathbb{W}}_u \right\rangle \right| du \right)^2 \right] \\ &\leq \mathbb{E} \left[t \int_0^t \left\langle \mathbf{u}, \widehat{\mathbb{W}}_u \right\rangle^2 du \right] \\ &= t \int_0^t \mathbb{E} \left[\left\langle \mathbf{u}, \widehat{\mathbb{W}}_u \right\rangle^2 \right] du. \end{aligned}$$

Thus, using (3.10) with $n = n(\mathbf{v}) = n(\mathbf{u}) + 1$ and $x = x(\mathbf{v}) = x(\mathbf{u}) + 1$, we have

$$\begin{aligned} \mathbb{E} \left[\sup_{s \in [0, t]} \left\langle \mathbf{v}, \widehat{\mathbb{W}}_s \right\rangle^2 \right] &\leq \binom{2(n-1)}{n-1} \frac{t}{2^{n-x}(n+x-2)!} \int_0^t u^{n+x-2} du \\ &= \binom{2(n-1)}{n-1} \frac{t^{n+x}}{2^{n-x}(n+x-1)!} \\ &\leq \frac{2^{n+x}}{\sqrt{n}} \frac{t^{n+x}}{(n+x-1)!}, \end{aligned} \tag{B.1}$$

where the last inequality comes from an application of Stirling's approximation that reads $\binom{2m}{m} \leq \frac{2^{2m}}{\sqrt{m+1}}$ for all $m \in \mathbb{N}$.

(2) **Case $\mathbf{v} = \mathbf{u}\mathbf{j}$ for some word $\mathbf{u} \in V$ and some letter $\mathbf{j} \in \{\mathbf{2}, \dots, \mathbf{d}\}$.** First we apply Theorem 3.3 to get

$$\begin{aligned} \left\langle \mathbf{v}, \widehat{\mathbb{W}}_s \right\rangle^2 &= \left(\int_0^s \left\langle \mathbf{u}, \widehat{\mathbb{W}}_u \right\rangle dW_u^{\mathbf{j}} + \frac{1}{2} \int_0^s \left\langle \mathbf{u}\mathbf{j}, \widehat{\mathbb{W}}_u \right\rangle du \right)^2 \\ &\leq 2 \left(\int_0^s \left\langle \mathbf{u}, \widehat{\mathbb{W}}_u \right\rangle dW_u^{\mathbf{j}} \right)^2 + \frac{1}{2} \left(\int_0^s \left\langle \mathbf{u}\mathbf{j}, \widehat{\mathbb{W}}_u \right\rangle du \right)^2. \end{aligned} \tag{B.2}$$

Now, since $\int_0^s \langle \mathbf{u}, \widehat{\mathbb{W}}_u \rangle dW_u^{\mathbf{j}}$ is a true martingale, we apply Doob's inequality followed by the previous result (B.1) with $n = n(\mathbf{v}) = n(\mathbf{u}) + 1$ and $x = x(\mathbf{v}) = x(\mathbf{u})$ and get

$$\mathbb{E} \left[\sup_{s \in [0, t]} \left(\int_0^s \langle \mathbf{u}, \widehat{\mathbb{W}}_u \rangle dW_u^{\mathbf{j}} \right)^2 \right] \leq 4 \mathbb{E} \left[\int_0^t \langle \mathbf{u}, \widehat{\mathbb{W}}_u \rangle^2 du \right] = 4 \int_0^t \mathbb{E} \left[\langle \mathbf{u}, \widehat{\mathbb{W}}_u \rangle^2 \right] du \leq 4 \frac{2^{n+x}}{\sqrt{n}} \frac{t^{n+x}}{(n+x)!}.$$

Then, again applying (B.1) to the right-most element in (B.2), we get

$$\mathbb{E} \left[\sup_{s \in [0, t]} \left(\int_0^s \langle \mathbf{u}_{\mathbf{j}}, \widehat{\mathbb{W}}_u \rangle du \right)^2 \right] = \mathbb{E} \left[\sup_{s \in [0, t]} \langle \mathbf{u}_{\mathbf{j}} \mathbf{1}, \widehat{\mathbb{W}}_s \rangle^2 \right] \leq \frac{2^{n+x}}{\sqrt{n}} \frac{t^{n+x}}{(n+x-1)!}.$$

Note that if $\mathbf{u}_{\mathbf{j}} = 0$, the left-hand side is 0 and the equation still holds.

Combining the above inequalities, we derive

$$\mathbb{E} \left[\sup_{s \in [0, t]} \langle \mathbf{v}, \widehat{\mathbb{W}}_s \rangle^2 \right] \leq C_1 \frac{2^{n+x-1}}{\sqrt{n}} \frac{t^{n+x}}{(n+x-1)!}, \quad (\text{B.3})$$

for some $C_1 \in \mathbb{R}$ independent of t and \mathbf{v} .

Combining the two equalities (B.1)-(B.3) above, we have that for every $\mathbf{v} \in V_n$, $n \geq 1$,

$$\mathbb{E} \left[\sup_{s \in [0, t]} \left| \langle \mathbf{v}, \widehat{\mathbb{W}}_s \rangle \right|^2 \right] \leq C_2 \frac{2^{n+x-1}}{\sqrt{n}} \frac{t^{n+x}}{(n+x-1)!}.$$

for some $C_2 \in \mathbb{R}$ independent of t and \mathbf{v} .

Moreover, as $n+1 \sim n$, there exists a constant C such that

$$\mathbb{E} \left[\sup_{s \in [0, t]} \left| \langle \mathbf{v}, \widehat{\mathbb{W}}_s \rangle \right|^2 \right] \leq C \frac{2^{n+x-1}}{\sqrt{n+1}} \frac{t^{n+x}}{(n+x-1)!}.$$

This includes the case $\mathbf{v} = \emptyset \in V_0$ and proves (3.11).

It remains to prove the sub-multiplicative property (3.12). First, in the case $\mathbf{u} = \emptyset$ or $\mathbf{v} = \emptyset$, (3.12) holds trivially. Now we can easily see that

$$\left(\frac{h_t(\mathbf{u})h_t(\mathbf{v})}{h_t(\mathbf{u}\mathbf{v})} \right)^2 = C^2 A(\mathbf{u}, \mathbf{v}) B(\mathbf{u}, \mathbf{v}),$$

where

$$A(\mathbf{u}, \mathbf{v}) := \frac{(n(\mathbf{u}) + x(\mathbf{u}) + n(\mathbf{v}) + x(\mathbf{v}) - 1)!}{(n(\mathbf{u}) + x(\mathbf{u}) - 1)!(n(\mathbf{v}) + x(\mathbf{v}) - 1)!} \quad \text{and} \quad B(\mathbf{u}, \mathbf{v}) := \sqrt{\frac{n(\mathbf{u}) + n(\mathbf{v}) + 1}{(n(\mathbf{u}) + 1)(n(\mathbf{v}) + 1)}}.$$

We now only need to show that $C^2 A(\mathbf{u}, \mathbf{v}) B(\mathbf{u}, \mathbf{v}) \geq 1$ for $\mathbf{u}, \mathbf{v} \neq \emptyset$. Without loss of generality, we assume $\mathbf{u} \neq \emptyset$, i.e. $n(\mathbf{u}) > 0$, otherwise we could simply swap \mathbf{u} and \mathbf{v} in the following equations. Let us then remark that

$$A(\mathbf{u}, \mathbf{v}) = \binom{n(\mathbf{u}) + x(\mathbf{u}) + n(\mathbf{v}) + x(\mathbf{v}) - 1}{n(\mathbf{u}) + x(\mathbf{u})} (n(\mathbf{u}) + x(\mathbf{u}))$$

and that

$$\begin{aligned} \sqrt{n(\mathbf{u}) + 1} &\leq C^2 (n(\mathbf{u}) + x(\mathbf{u})) \\ \sqrt{n(\mathbf{v}) + 1} \sqrt{n(\mathbf{u}) + 1} &\leq C^2 \sqrt{n(\mathbf{v}) + 1} (n(\mathbf{u}) + x(\mathbf{u})) \leq C^2 \sqrt{n(\mathbf{u}) + n(\mathbf{v}) + 1} (n(\mathbf{u}) + x(\mathbf{u})), \end{aligned}$$

which implies

$$B(\mathbf{u}, \mathbf{v}) \geq \frac{1}{C^2(n(\mathbf{u}) + x(\mathbf{u}))}$$

and thus

$$\left(\frac{h_t(\mathbf{u})h_t(\mathbf{v})}{h_t(\mathbf{uv})} \right)^2 \geq \left(\frac{n(\mathbf{u}) + x(\mathbf{u})}{n(\mathbf{u}) + x(\mathbf{u})} + \frac{n(\mathbf{v}) + x(\mathbf{v}) - 1}{n(\mathbf{u}) + x(\mathbf{u})} \right) \geq 1.$$

□

C Proof of Proposition 3.1

Let $n \leq k \in \mathbb{N}$, then for all $\mathbf{u} \in V_n$, $\mathbf{v} \in V_{k-n}$, each element of $\mathbf{u} \sqcup \mathbf{v}$ is of size k . This leads to the fact that $\langle \mathbf{u} \sqcup \mathbf{v}, \mathbf{w} \rangle = 0$ whenever $\mathbf{w} \notin V_k$, i.e.

$$\langle \mathbf{u} \sqcup \mathbf{v}, \mathbb{X}_t \rangle = \sum_{\mathbf{w} \in V_k} \langle \mathbf{u} \sqcup \mathbf{v}, \mathbf{w} \rangle \mathbb{X}_t^{\mathbf{w}}.$$

Moreover, by bilinearity of the shuffle product and for $\mathbf{p}, \ell \in T((\mathbb{R}^d))$,

$$\langle \mathbf{p} \sqcup \ell, \mathbf{w} \rangle = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \sum_{\mathbf{u} \in V_n} \sum_{\mathbf{v} \in V_m} \mathbf{p}^{\mathbf{u}} \ell^{\mathbf{v}} \langle \mathbf{u} \sqcup \mathbf{v}, \mathbf{w} \rangle = \sum_{n=0}^k \sum_{\mathbf{u} \in V_n} \sum_{\mathbf{v} \in V_{k-n}} \mathbf{p}^{\mathbf{u}} \ell^{\mathbf{v}} \langle \mathbf{u} \sqcup \mathbf{v}, \mathbf{w} \rangle.$$

Hence, it is easy to see that for all $k \in \mathbb{N}$

$$\sum_{\mathbf{w} \in V_k} \langle \mathbf{p} \sqcup \ell, \mathbf{w} \rangle \mathbb{X}_t^{\mathbf{w}} = \sum_{n=0}^k \sum_{\mathbf{u} \in V_n} \sum_{\mathbf{v} \in V_{k-n}} \mathbf{p}^{\mathbf{u}} \ell^{\mathbf{v}} \langle \mathbf{u} \sqcup \mathbf{v}, \mathbb{X}_t \rangle.$$

It then follows that the semi-norm of $\mathbf{p} \sqcup \ell$ is of the form

$$\begin{aligned} \|\mathbf{p} \sqcup \ell\|_t^A &= \sum_{k=0}^{\infty} \left| \sum_{\mathbf{w} \in V_k} \langle \mathbf{p} \sqcup \ell, \mathbf{w} \rangle \mathbb{X}_t^{\mathbf{w}} \right| \\ &= \sum_{k=0}^{\infty} \left| \sum_{n=0}^k \sum_{\mathbf{u} \in V_n} \sum_{\mathbf{v} \in V_{k-n}} \mathbf{p}^{\mathbf{u}} \ell^{\mathbf{v}} \langle \mathbf{u} \sqcup \mathbf{v}, \mathbb{X}_t \rangle \right| \\ &\leq \sum_{k=0}^{\infty} \sum_{n=0}^k \left| \sum_{\mathbf{u} \in V_n} \sum_{\mathbf{v} \in V_{k-n}} \mathbf{p}^{\mathbf{u}} \ell^{\mathbf{v}} \langle \mathbf{u} \sqcup \mathbf{v}, \mathbb{X}_t \rangle \right|. \end{aligned}$$

Finally, using the shuffle product property on the right-hand side, recall Lemma 2.5, the semi-norm can be bounded

$$\|\mathbf{p} \sqcup \ell\|_t^A \leq \sum_{k=0}^{\infty} \sum_{n=0}^k \left| \sum_{\mathbf{u} \in V_n} \sum_{\mathbf{v} \in V_{k-n}} \mathbf{p}^{\mathbf{u}} \ell^{\mathbf{v}} \mathbb{X}_t^{\mathbf{u}} \mathbb{X}_t^{\mathbf{v}} \right| \leq \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \left| \sum_{\mathbf{u} \in V_n} \mathbf{p}^{\mathbf{u}} \mathbb{X}_t^{\mathbf{u}} \sum_{\mathbf{v} \in V_m} \ell^{\mathbf{v}} \mathbb{X}_t^{\mathbf{v}} \right| \leq \|\mathbf{p}\|_t^A \cdot \|\ell\|_t^A < \infty.$$

It can then be argued very similarly that $\langle \mathbf{p}, \mathbb{X}_t \rangle \langle \ell, \mathbb{X}_t \rangle = \langle \mathbf{p} \sqcup \ell, \mathbb{X}_t \rangle$, ending the proof.

References

- [1] Eduardo Abi Jaber. Lifting the heston model. *Quantitative finance*, 19(12):1995–2013, 2019.
- [2] Eduardo Abi Jaber and Omar El Euch. Multifactor approximation of rough volatility models. *SIAM journal on financial mathematics*, 10(2):309–349, 2019.

- [3] Eduardo Abi Jaber and Louis-Amand Gérard. Signature volatility models: pricing and hedging with Fourier. arXiv preprint arXiv:2402.01820, 2024.
- [4] Eduardo Abi Jaber, Martin Larsson, and Sergio Pulido. Affine Volterra processes. The Annals of Applied Probability, 29(5):3155–3200, 2019.
- [5] Christian Bayer, Paul P Hager, Sebastian Riedel, and John Schoenmakers. Optimal stopping with signatures. The Annals of Applied Probability, 33(1):238–273, 2023.
- [6] Christian Bayer, Luca Pelizzari, and John Schoenmakers. Primal and dual optimal stopping with signatures. arXiv preprint arXiv:2312.03444, 2023.
- [7] Gérard Ben Arous. Flots et séries de taylor stochastiques. Probability Theory and Related Fields, 81: 29–77, 1989.
- [8] Alvaro Cartea, Imanol Perez Arribas, and Leandro Sanchez-Betancourt. Optimal execution of foreign securities: A double-execution problem, 2020. URL <https://ssrn.com/abstract=3562251>.
- [9] Kuo-Tsai Chen. Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. Annals of Mathematics, 65(1):163–178, 1957.
- [10] Ilya Chevyrev and Andrey Kormilitzin. A primer on the signature method in machine learning. arXiv preprint arXiv:1603.03788, 2016.
- [11] Christa Cuchiero, Guido Gazzani, Janka Möller, and Sara Svaluto-Ferro. Joint calibration to SPX and VIX options with signature-based models. arXiv preprint arXiv:2301.13235, 2023.
- [12] Christa Cuchiero, Sara Svaluto-Ferro, and Josef Teichmann. Signature sdes from an affine and polynomial perspective, 2023.
- [13] Bruno Dupire and Valentin Tissot-Daguette. Functional expansions. arXiv preprint arXiv:2212.13628, 2022.
- [14] Thomas Fawcett. Problems in stochastic analysis. Connections between rough paths and non-commutative harmonic analysis. PhD thesis, University of Oxford, 2003.
- [15] Adeline Fermanian. Embedding and learning with signatures. Computational Statistics & Data Analysis, 157:107148, 2021.
- [16] Michel Fliess. On the concept of derivatives and taylor expansions for nonlinear input-output systems. In The 22nd IEEE Conference on Decision and Control, pages 643–646. IEEE, 1983.
- [17] Michel Fliess. Vers une notion de dérivation fonctionnelle causale. In Annales de l’Institut Henri Poincaré C, Analyse non linéaire, volume 3, pages 67–76. Elsevier, 1986.
- [18] Peter K. Friz and Nicolas B. Victoir. Multidimensional Stochastic Processes as Rough Paths: Theory and Applications. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2010. doi: 10.1017/CBO9780511845079.
- [19] J. G. Gaines. The algebra of iterated stochastic integrals. Stochastics and Stochastic Reports, 49(3-4): 169–179, 1994. doi: 10.1080/17442509408833918. URL <https://doi.org/10.1080/17442509408833918>.
- [20] Gustaf Gripenberg, Stig-Olof Londen, and Olof Staffans. Volterra integral and functional equations. Cambridge University Press, 1990.
- [21] Yushi Hamaguchi. Variation of constants formulae for forward and backward stochastic volterra integral equations. Journal of Differential Equations, 343:332–389, 2023.
- [22] Peter E Kloeden, Eckhard Platen, Peter E Kloeden, and Eckhard Platen. Stochastic differential equations. Springer, 1992.
- [23] Christian Litterer and Harald Oberhauser. On a chen–fliess approximation for diffusion functionals. Monatshefte für Mathematik, 175:577–593, 2014.

- [24] Terry Lyons. Rough paths, signatures and the modelling of functions on streams. arXiv preprint arXiv:1405.4537, 2014.
- [25] Terry Lyons and Nicolas Victoir. Cubature on wiener space. Proceedings: Mathematical, Physical and Engineering Sciences, 460(2041):169–198, 2004. ISSN 13645021. URL <http://www.jstor.org/stable/4143098>.
- [26] Terry Lyons, Sina Nejad, and Imanol Perez Arribas. Non-parametric pricing and hedging of exotic derivatives. Applied Mathematical Finance, 27(6):457–494, 2020. doi: 10.1080/1350486X.2021.1891555. URL <https://doi.org/10.1080/1350486X.2021.1891555>.
- [27] Terry J. Lyons. Differential equations driven by rough signals. Revista Matemática Iberoamericana, 14(2):215–310, 1998. URL <http://eudml.org/doc/39555>.
- [28] Imanol Perez Arribas, Guy M Goodwin, John R Geddes, Terry Lyons, and Kate EA Saunders. A signature-based machine learning model for distinguishing bipolar disorder and borderline personality disorder. Translational psychiatry, 8(1):274, 2018.
- [29] Rimhak Ree. Lie elements and an algebra associated with shuffles. Annals of Mathematics, 68(2):210–220, 1958. ISSN 0003486X. URL <http://www.jstor.org/stable/1970243>.
- [30] Mark Veraar. The stochastic fubini theorem revisited. Stochastics An International Journal of Probability and Stochastic Processes, 84(4):543–551, 2012.