

Malliavin calculus for signatures with applications to finance

Eduardo Abi Jaber^{*1}, Clément Rey^{†1}, and Dimitri Sotnikov^{‡1, 2}

¹Ecole Polytechnique, CMAP

²Engie Global Markets

April 27, 2026

Abstract

Malliavin calculus is a powerful and general framework for the analysis of square-integrable random variables, but it often suffers from a lack of tractability and explicit representations. To address this limitation, we focus on a subclass of random variables given by finite linear combinations of time-extended Brownian motion signatures. The class remains rich due to the universal approximation properties of signatures. Leveraging the algebraic structure of signatures, we first derive explicit formulas for the Malliavin derivative of signatures of continuous Itô processes. As a consequence, we obtain closed-form expressions for the Clark–Ocone representation, the Ornstein–Uhlenbeck semigroup and its generator, as well as the integration-by-parts formula within the class of Brownian signature variables. These results provide purely algebraic formulations of the classical operators of Malliavin calculus. As an application, we compute Greeks for general path-dependent options under signature volatility models, and numerically compare different choices of Malliavin weights.

Mathematics Subject Classification (2010): 60H07, 60H30

Keywords: Path signatures, Malliavin calculus, Integration by parts, Greeks.

1 Introduction

A central problem in probability is to establish differentiability of the mapping $\theta \mapsto \mathbb{E}[f(G^\theta)]$, where G^θ is a random variable depending on a parameter $\theta \in \mathbb{R}$, and to derive tractable representations of the associated parameter sensitivity

$$\frac{\partial}{\partial \theta} \mathbb{E}[f(G^\theta)].$$

The [Malliavin \(1978\)](#) calculus¹, which serves as a stochastic analogue of the calculus of variations, provides a powerful tool to address this problem whenever G^θ is an \mathcal{F}_T^W -measurable random variable with W a Brownian motion. Under suitable regularity and non-degeneracy assumptions, the Malliavin integration-by-parts formula yields the representation

$$\frac{\partial}{\partial \theta} \mathbb{E}[f(G^\theta)] = \mathbb{E}[f(G^\theta)\pi_T], \tag{1.1}$$

where π_T denotes the associated Malliavin weight. An important application of such formulas is the derivation of explicit expressions for the density of random variables. In particular, [Bismut \(1981\)](#) introduced a probabilistic proof of Hörmander’s theorem based on integration by parts on the Wiener space, establishing the smoothness and lower bounds of densities for a broad class of diffusion processes ([Kusuoka and Stroock, 1985](#); [Kohatsu-Higa, 2003](#); [Nualart, 2006](#); [Bally, 2006](#); [Nourdin and Viens, 2009](#)). Furthermore, the representation (1) is

^{*}eduardo.abi-jaber@polytechnique.edu. The first author is grateful for the financial support from the Chaires FiME-FDD, Financial Risks at Ecole Polytechnique.

[†]clement.rey@polytechnique.edu.

[‡]dmitrii.sotnikov@polytechnique.edu. The third author is grateful for the financial support provided by Engie Global Markets.

¹We refer to [Nualart \(2006\)](#); [Alós, Lorig, and Gatarek \(2021\)](#) for a comprehensive treatment of Malliavin calculus and its applications.

particularly attractive for Monte Carlo methods, as it avoids differentiating the function f . Integration-by-parts techniques have proven useful in stochastic optimal control (Gobet and Munos, 2005), gradient estimation in machine learning (Oden, 2025), and the improvement of learning procedures (Pidstrigach et al., 2025). In mathematical finance, the computation of sensitivities (the so-called Greeks) via Malliavin integration by parts was first developed within the Black–Scholes framework (Fournié, Lasry, Lions, Lebuchoux, and Touzi, 1999). This approach was later extended to stochastic volatility models (El-Khatib, 2009), including hybrid (Yilmaz, 2018), jump-diffusion (El-Khatib and Privault, 2003; Davis and Johansson, 2006; Petrou, 2008), and non-Markovian settings (Al-Foraih, Pospíšil, and Vives, 2023), and to more complex derivatives such as barrier, lookback (Gobet and Kohatsu-Higa, 2003), and American options (Bally, Caramellino, and Zanette, 2003).

However, in general, the Malliavin weight π_T is either not explicit or difficult to compute numerically, except in very specific configurations where additional structure is available. This naturally leads to the following question:

Can one identify a structural and generic class of random variables G^θ for which the Malliavin weight admits an explicit and numerically tractable form?

A natural solution comes from the theory of path signatures, introduced by Chen (1957) as sequences of iterated integrals, which provide a systematic way to represent functionals of continuous paths. More precisely, the signature $\widehat{\mathbb{W}}$ of the time-augmented Brownian motion $\widehat{W}_t = (t, W_t)$, $t \geq 0$, is defined as the sequence

$$\widehat{\mathbb{W}}_t = \left(1, \binom{t}{W_t}, \left(\int_0^t W_s ds, \int_0^t s dW_s \right), \dots \right),$$

and serves as a universal approximator: it naturally appears in stochastic Taylor expansions of Markovian (Ben Arous, 1989; Kloeden and Platen, 1992) and non-Markovian systems (Fliess, 1983; Abi Jaber, Gérard, and Huang, 2024), and satisfies a Stone–Weierstrass–type theorem for functionals of Brownian motion. This highlights the role of signatures as analogues of polynomials on path space and underlines their foundational importance in the theory of rough paths initiated by Terry Lyons (see, e.g., Lyons, Caruana, and Lévy, 2007).

The interplay between Malliavin calculus and the theory of rough paths and signatures has received considerable attention. For instance, Friz and Victoir (2010, Chapter 20) apply Malliavin calculus to establish the existence of smooth densities for rough differential equations (RDEs). Further developments include the study of Malliavin differentiability of RDE solutions by Inahama (2014), the analysis of Malliavin differentiability and existence of smooth densities for rough SDEs by Bugini, Coghi, and Nilssen (2024), and the computation of the Wiener–Itô chaos expansion of Gaussian-process signatures by Cass and Ferrucci (2024). Moreover, Malliavin calculus was used to establish the densities of the truncated path signatures themselves. See Kusuoka and Stroock (1987) for the case of Brownian motion signature and Baudoin, Feng, and Ouyang (2020) for the fractional Brownian motion case. Regarding the computation of Greeks, Teichmann (2006) proposed an approach relying on the cubature on Wiener space introduced by Lyons and Victoir (2004). Most existing approaches start from an SDE, use a stochastic Taylor expansion to bring out the signature, and then, under conditions such as Hörmander’s, apply Malliavin calculus. A comprehensive framework for applying Malliavin calculus directly to the Brownian signature $\widehat{\mathbb{W}}_t$ is not yet fully developed. We instead isolate the Malliavin/signature component, aiming at a unified and generic treatment that exploits the algebraic structure of signatures and extends beyond the SDE setting to non-Markovian dynamics.

Going back to our question, and motivated by the universal approximation properties of path signatures, we restrict our attention to random variables that are linear in the signature of Brownian motion, of the form

$$G^\theta = \left\langle \ell^\theta, \widehat{\mathbb{W}}_T \right\rangle,$$

where the coefficients ℓ^θ are deterministic. A first crucial observation is that this representation decouples the stochastic noise from the parameter dependence. As a consequence, the sensitivity analysis reduces to first understanding the Malliavin calculus of the generic object $\widehat{\mathbb{W}}$ and then applying integration-by-parts formulas to the class of linear functionals. Compared to the related literature, a key feature of our approach is that all fundamental Malliavin operations on the signature are systematically encoded algebraically as linear operators on the (extended) tensor algebra. This allows us to exploit the rich algebraic structure of signatures to derive explicit expressions for the Malliavin weights, while the computability of $\widehat{\mathbb{W}}_T$ enables the design of numerically tractable implementations.

To implement this strategy, we first develop a probabilistic perspective on the path signature of stochastic processes, which allows us to derive results for the Malliavin derivative of the signature of general Itô processes (Theorem 4.4) echoing the calculus of variations for rough differential equations (Marie, 2015), but using a purely probabilistic approach. In Theorem 4.3, we further identify a class of processes for which the Stratonovich integral is Malliavin differentiable, and we derive a formula for its Malliavin derivative.

By combining this viewpoint with the powerful algebraic properties of the signature, we then show that several fundamental operations from Malliavin calculus admit explicit and tractable representations when applied to Brownian signatures. In particular, we derive explicit expressions for the Skorokhod integral (Proposition 5.7), the Ornstein–Uhlenbeck semigroup (Theorem 5.14), chaos expansions (Proposition 5.9), and ultimately to an integration-by-parts formula (Theorem 6.8) for functionals of the form $f(G^\theta)$, where the Malliavin weights are determined explicitly and belong to a class of random variables stable under integration by parts. More precisely, this stable class turns out to be the class of rational functions of $\widehat{\mathbb{W}}_T$, that is, ratios of two linear functionals of the signature.

As an application, we use our results to derive and compute model sensitivities (Greeks) of path-dependent payoffs in signature volatility models in Section 7. Signature volatility models form a universal and flexible class of volatility models as shown by Arribas, Salvi, and Szpruch (2020); Cuchiero, Gazzani, Möller, and Svaluto-Ferro (2025a); Abi Jaber and Gérard (2025a), capable of representing and approximating a wide range of (possibly non-Markovian) dynamics while preserving computational tractability for pricing and hedging derivatives. Within this setting, our Malliavin integration-by-parts formula yields explicit and reusable expressions for Greeks of a large class of path-dependent payoffs. In particular, any payoff of the form $f(\langle \ell, \mathbb{X}_T \rangle)$, where \mathbb{X}_T denotes the signature of the time-augmented log-price process $(t, \log S_t)$, can be treated within a unified framework. The resulting sensitivity formula depends only on the deterministic coefficient ℓ , without requiring additional model-specific derivations, which makes the approach particularly robust and attractive for practical implementation. For completeness, we perform numerical comparisons of different choices of Malliavin weights for European and Asian options.

The results presented in this paper can be useful in a variety of settings where signatures are employed. Beyond volatility modeling, signatures have found wide-ranging applications across various domains; see Chevyrev and Kormilitzin (2025) for a comprehensive overview of the signature method. In particular, they have been used to solve stochastic control problems (Bayer, Hager, Riedel, and Schoenmakers, 2023; Bayer, Pelizzari, and Schoenmakers, 2025; Abi Jaber, Hainaut, and Motte, 2025b) to capture memory effects and nonlinearities in machine learning for time series (Levin, Lyons, and Ni, 2016; Drobac, Brégère, de Vilmarest, and Wintenberger, 2025; Abi Jaber and Gérard, 2025b), and in finance (Lyons, Nejad, and Arribas, 2020; Cartea, Arribas, and Sánchez-Betancourt, 2022).

We finally mention that an alternative framework to Malliavin calculus is Functional Itô Calculus, introduced by Dupire (2019) and subsequently employed by Dupire and Tissot-Daguette (2022) to prove the functional Taylor expansion. These results were then extended by Cuchiero, Guo, and Primavera (2025b) to functionals of càdlàg rough paths, leveraging signature approximation properties, and by Bielert (2025) using rough integration theory. The connection between the functional (vertical) derivative and the Malliavin derivative was established in Cont and Fournié (2013). Building on these ideas, Jazaerli and F. Saporito (2017) proposed an approach to Greek computation based on integration by parts with functional derivatives. However, their tractable formula applies only to so-called weakly path-dependent functionals, which exclude many path-dependent derivative contracts of practical interest.

Outline. Section 2 provides a summary of our main results. Section 3 introduces the necessary preliminaries from path signatures. In Section 4, we derive the formula for the Malliavin derivative of the signature of Itô processes. Section 5 is devoted to developing the framework for working with the signature of Brownian motion and provides explicit computations of certain Malliavin calculus operators for this case. In Section 6, we prove the integration by parts formula for random variables that depend on the Brownian signature. Finally, in Section 7, we apply these results to the computation of Greeks for possibly path-dependent options under the signature volatility model. Appendix A contains the proof of the Malliavin differentiability of Stratonovich integrals.

2 Summary of main results

We first introduce the notation and basic framework in Section 2.1, and then present streamlined versions of our principal theorems in Section 2.2.

2.1 Notations

From now on, let $T > 0$ be fixed. Throughout this paper, we work on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t^W)_{t \in [0, T]}, \mathbb{P})$ supporting a d -dimensional Brownian motion W , where $(\mathcal{F}_t^W)_{t \in [0, T]}$ denotes the filtration generated by W .

We will denote by \circ the Stratonovich integrals. Namely, for two continuous semimartingales $X = (X_t)_{t \in [0, T]}$ and $Y = (Y_t)_{t \in [0, T]}$, the Stratonovich integral is defined by

$$\int_0^\cdot Y_s \circ dX_s = \int_0^\cdot Y_s dX_s + \frac{1}{2} \langle X, Y \rangle, ,$$

where the integral on the right-hand side is the standard Itô integral $\langle X, Y \rangle$ denotes the quadratic covariation between X and Y .

We will denote by $C_p^k(V, W)$ (resp. $C_b^k(V, W)$) the space of k times differentiable functions from V to W with all derivatives of polynomial growth (resp. bounded).

Path signature. We introduce the key objects needed to present the main results of this work. For a more thorough introduction to path signatures, we refer to [Chevyrev and Kormilitzin \(2025\)](#). We start with the definition of the path signature of a continuous \mathbb{R}^d -valued semimartingale $X = (X_t)_{t \in [0, T]}$, which will be denoted by \mathbb{X} . The *path signature* $\mathbb{X}_{s,t}$ of X over $[s, t]$ for $0 \leq s \leq t \leq T$ is the infinite sequence of iterated Stratonovich integrals:

$$\mathbb{X}_{s,t} := \left(\mathbb{X}_{s,t}^{\mathbf{i}_1 \dots \mathbf{i}_n} \right)_{\mathbf{i}_1 \dots \mathbf{i}_n \in V} = \left(\int_{s \leq u_1 \leq \dots \leq u_n \leq t} \circ dX_{u_1}^{i_1} \dots \circ dX_{u_n}^{i_n} \right)_{\mathbf{i}_1 \dots \mathbf{i}_n \in V}. \quad (2.1)$$

Here, V denotes the set of all words in the alphabet $A = \{\mathbf{1}, \dots, \mathbf{d}\}$. We will also use the standard notation $\mathbb{X}_t := \mathbb{X}_{0,t}$.

Remark 2.1. Throughout this paper, we will often consider the *time-augmented* paths $\widehat{X}: t \mapsto (t, X_t) \in \mathbb{R}^{d+1}$ and their signatures, denoted by $\widehat{\mathbb{X}}_t$. In this case, it is convenient to consider the alphabet $\{\mathbf{0}, \mathbf{1}, \dots, \mathbf{d}\}$, where the letter $\mathbf{0}$ corresponds to integration with respect to the time component, while the letters \mathbf{i} correspond to integration with respect to X^i for $i = 1, \dots, d$. \blacksquare

Extended tensor algebra. For $n \geq 0$, each word $\mathbf{v} = \mathbf{i}_1 \dots \mathbf{i}_n \in V$ of length $|\mathbf{v}| = n$, with $\mathbf{i}_k \in A$, $k = 1, \dots, n$, can be identified with a basis vector $e_{i_1} \otimes \dots \otimes e_{i_n}$ of $(\mathbb{R}^d)^{\otimes n}$, where e_1, \dots, e_d form an orthonormal basis of \mathbb{R}^d . We use the convention $(\mathbb{R}^d)^{\otimes 0} \cong \mathbb{R}$ and denote the corresponding empty word by \emptyset . This leads to the definition of the *extended tensor algebra*

$$T((\mathbb{R}^d)) := \prod_{n \geq 0} (\mathbb{R}^d)^{\otimes n} = \{ \ell = (\ell_0, \ell_1, \dots) : \ell_n \in (\mathbb{R}^d)^{\otimes n}, n \geq 0 \}.$$

Since each element $\ell_n \in (\mathbb{R}^d)^{\otimes n}$ admits a representation $\ell_n = \sum_{|\mathbf{v}|=n} \ell^{\mathbf{v}} \mathbf{v}$, each element $\ell \in T((\mathbb{R}^d))$ can be written as the formal sum

$$\ell = \sum_{n \geq 0} \sum_{|\mathbf{v}|=n} \ell^{\mathbf{v}} \mathbf{v}.$$

In particular, the path signature defined in (2.1) belongs to $T((\mathbb{R}^d))$ and admits the representation

$$\mathbb{X}_{s,t} = \sum_{n \geq 0} \sum_{|\mathbf{v}|=n} \mathbb{X}_{s,t}^{\mathbf{v}} \mathbf{v}.$$

Tensor algebra and duality. We denote by $T(\mathbb{R}^d) \subset T((\mathbb{R}^d))$ the *tensor algebra*

$$T(\mathbb{R}^d) := \bigoplus_{n \geq 0} (\mathbb{R}^d)^{\otimes n},$$

i.e. the subspace of sequences $\ell \in T((\mathbb{R}^d))$ having only finitely many non-zero components. Note that $T(\mathbb{R}^d)$ is the topological dual of $T((\mathbb{R}^d))$ (equipped with the product topology), with pairing given by

$$\langle \ell, \mathbb{X} \rangle = \sum_{n \geq 0} \sum_{|\mathbf{v}|=n} \ell^{\mathbf{v}} \mathbb{X}^{\mathbf{v}}, \quad \ell \in T(\mathbb{R}^d), \quad \mathbb{X} \in T((\mathbb{R}^d)),$$

where the sum is well-defined since only finitely many terms are nonzero. Moreover, given a continuous linear operator $\mathbf{A}: T((\mathbb{R}^d)) \rightarrow T((\mathbb{R}^d))$, we call $\mathbf{A}^*: T(\mathbb{R}^d) \rightarrow T(\mathbb{R}^d)$ its *adjoint* if

$$\langle \ell, \mathbf{A}\mathbb{X} \rangle = \langle \mathbf{A}^*\ell, \mathbb{X} \rangle,$$

for all $\ell \in T(\mathbb{R}^d)$ and $\mathbb{X} \in T((\mathbb{R}^d))$.

Algebraic structure. The extended tensor algebra $T((\mathbb{R}^d))$ becomes an algebra with unit element \emptyset when equipped with the standard operations of addition, scalar multiplication, and the tensor product defined by

$$\ell \otimes \mathbf{p} \in T((\mathbb{R}^d)), \quad (\ell \otimes \mathbf{p})_n = \sum_{k=0}^n \ell_k \otimes \mathbf{p}_{n-k}, \quad n \geq 0.$$

A key algebraic property of the path signature, known as the [Chen \(1957\)](#) identity, states that the signature of a path over $[s, t]$ is equal to the product of its signatures over $[s, u]$ and $[u, t]$ for all $s \leq u \leq t$, namely,

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \otimes \mathbb{X}_{u,t}.$$

In particular, it follows from the uniqueness result of [Boedihardjo et al. \(2016\)](#) that the signature of a path X concatenated with its time-reversal $\overleftarrow{X}^{[0,T]}: t \mapsto X_{T-t}$ is trivial. By Chen's identity,

$$\mathbb{X}_T \otimes \overleftarrow{\mathbb{X}}_T^{[0,T]} = \emptyset,$$

and hence the inverse group element of \mathbb{X}_T is given by the signature of the time-reversed path, $\mathbb{X}_T^{-1} = \overleftarrow{\mathbb{X}}_T^{[0,T]}$.

Malliavin derivative of path signatures. We introduce the space of simple functionals as follows:

$$\mathcal{S} := \left\{ F: \Omega \rightarrow \mathbb{R}: F(\omega) = f(W_{t_{k+1}}(\omega) - W_{t_k}(\omega), k = 0, \dots, 2^n - 1), f \in \mathcal{C}_p^\infty((\mathbb{R}^d)^{2^n}, \mathbb{R}), t_k = \frac{kT}{2^n}, n \in \mathbb{N} \right\}.$$

For $F \in \mathcal{S}$, we define the *Malliavin derivative* of F as an element of $L^2(\Omega \times [0, T], \mathbb{R}^d)$ with components

$$\mathbf{D}_t^j F = \partial_{x_l^j} f(W_{t_{k+1}} - W_{t_k}, k = 0, \dots, 2^n - 1), \quad j = 1, \dots, d, \quad (2.2)$$

for $t \in (t_l, t_{l+1}]$, where $\partial_{x_l^j}$ denotes the partial derivative with respect to the j -th component of the argument x_l , associated with the increment $W_{t_{l+1}} - W_{t_l}$. We use the convention $\mathbf{D}_0^j F = 0$.

Higher-order Malliavin derivatives are defined recursively as

$$\mathbf{D}_{t_1} \dots \mathbf{D}_{t_\ell} F \in (\mathbb{R}^d)^{\otimes \ell}.$$

Notice that \mathcal{S} is dense in $L^2(\Omega)$, and \mathbf{D} is a closable operator from $L^2(\Omega)$ to $L^2(\Omega \times [0, T], \mathbb{R}^d)$. For $k = 1, 2, \dots$, the space $\mathbb{D}^{k,2}$ is defined as the closure of \mathcal{S} with respect to the norm

$$\|F\|_{\mathbb{D}^{k,2}}^2 := \mathbb{E} [|F|^2] + \sum_{m=1}^k \mathbb{E} \left[\int_0^T \dots \int_0^T |\mathbf{D}_{t_1} \dots \mathbf{D}_{t_m} F|_{(\mathbb{R}^d)^{\otimes m}}^2 dt_1 \dots dt_m \right].$$

In particular, $\mathbb{D}^{k,2}$ is equipped with the inner product

$$\langle F, G \rangle_{k,2} := \mathbb{E}[FG] + \sum_{m=1}^k \mathbb{E} \left[\int_0^T \dots \int_0^T \langle \mathbf{D}_{t_1} \dots \mathbf{D}_{t_m} F, \mathbf{D}_{t_1} \dots \mathbf{D}_{t_m} G \rangle_{(\mathbb{R}^d)^{\otimes m}} dt_1 \dots dt_m \right].$$

and hence forms a Hilbert space. We also define $\mathbb{D}^{\infty,2} := \bigcap_{k \geq 1} \mathbb{D}^{k,2}$.

The Malliavin derivative satisfies the chain rule. For a random vector $F = (F_1, \dots, F_k)$ such that $F_i \in \mathbb{D}^{1,2}$ for all i , and for a function $\varphi \in C_b^1(\mathbb{R}^k, \mathbb{R})$, the random variable $\varphi(F)$ is also Malliavin differentiable, and

$$\mathbf{D}(\varphi(F)) = \sum_{i=1}^k \partial_{x_i} \varphi(F) \mathbf{D}F_i. \quad (2.3)$$

When the signature components $\mathbb{X}_T^{\mathbf{v}}$ of an \mathbb{R}^m -valued process X belong to $\mathbb{D}^{k,2}$ for all $\mathbf{v} \in V$, we denote by $\mathbf{D}_{t_1}^{j_1} \dots \mathbf{D}_{t_k}^{j_k} \mathbb{X}_T$ the element-wise application of the Malliavin derivative:

$$\mathbf{D}_{t_1}^{j_1} \dots \mathbf{D}_{t_k}^{j_k} \mathbb{X}_T := \left(\mathbf{D}_{t_1}^{j_1} \dots \mathbf{D}_{t_k}^{j_k} \mathbb{X}_T^{\mathbf{v}} \right)_{\mathbf{v} \in V} \in T((\mathbb{R}^m)).$$

Then the element $\underbrace{\mathbf{D} \dots \mathbf{D}}_{k \text{ times}} \mathbb{X}_T$ belongs to $L^2(\Omega \times [0, T]^k, (\mathbb{R}^d)^{\otimes k} \otimes T((\mathbb{R}^m)))$.

2.2 Main results

We now present our main results.

Malliavin derivative of signatures of Itô processes. Consider a continuous \mathbb{R}^m -valued Itô process of the form

$$dX_t = a_t^0 dt + \sum_{k=1}^d a_t^k \circ dW_t^k = A_t \circ d\widehat{W}_t,$$

where $a_t^j \in \mathbb{R}^m$, $j = 0, 1, \dots, d$, are square-integrable adapted processes satisfying additional Malliavin differentiability and regularity conditions specified in Assumption 4.1, which together form the columns of the matrix $A_t \in \mathbb{R}^{m \times (d+1)}$, and $\widehat{W}_t = (t, W_t)$. We show that the Malliavin derivative of \mathbb{X}_t , the signature of X , when exists, is given explicitly by

$$\mathbf{D}_s \mathbb{X}_t = \mathbb{X}_s \otimes A_s \otimes \mathbb{X}_{s,t} + \int_s^t \mathbb{X}_u \otimes \left(\mathbf{D}_s A_u \circ d\widehat{W}_u \right) \otimes \mathbb{X}_{u,t}, \quad s \leq t,$$

see Theorem 4.4. In particular, when $X_t = \widehat{W}_t = (t, W_t)$, this implies

$$\mathbf{D}_{s_1}^{i_1} \dots \mathbf{D}_{s_n}^{i_n} \widehat{W}_t = \widehat{W}_{s_1} \otimes \mathbf{i}_1 \otimes \widehat{W}_{s_1, s_2} \otimes \mathbf{i}_2 \otimes \dots \otimes \mathbf{i}_n \otimes \widehat{W}_{s_n, t}, \quad 0 \leq s_1 \leq \dots \leq s_n \leq t. \quad (2.4)$$

The ‘‘pierced’’ formula (2.2) shows that iterated Malliavin derivatives of the Brownian signature correspond to inserting the letters $\mathbf{i}_1, \dots, \mathbf{i}_n$ into the signature at the differentiation times. The formula admits a natural geometric interpretation, as explained in the next remark.

Remark 2.2. By definition (2.1), the Malliavin derivative \mathbf{D}_s^i , $s \geq 0$ of a functional $F((W_u)_{u \in [0,t]})$ can be formally viewed as the directional derivative of F in the direction of the perturbation $\mathbf{1}_{\{[s, +\infty)\}} \mathbf{i}$, that is,

$$\mathbf{D}_s^i \widehat{W}_t = \frac{d}{d\epsilon} \Big|_{\epsilon=0} F((W_u + \epsilon \mathbf{1}_{\{u \geq s\}} \mathbf{i})_{u \in [0,t]}).$$

Of course, the Malliavin derivative can be rigorously interpreted as a directional derivative only along directions in the Cameron–Martin space, which does not contain indicator functions. Therefore, this remark is purely formal and is intended only as a geometric illustration of the result. Geometrically, this variation corresponds to inserting a short ‘‘vertical’’ segment of length ϵ in the direction \mathbf{i} between the two parts $(W_u)_{u \in [0,s]}$ and

Ornstein–Uhlenbeck semigroup and its generator. We next consider the generalized Ornstein–Uhlenbeck semigroup, which is a family of operators acting on $L^2(\Omega, \mathcal{F}_T^W, \mathbb{P})$. Note that any random variable $F \in L^2(\Omega, \mathcal{F}_T^W, \mathbb{P})$ can be represented as a measurable functional of the Brownian motion, that is, $F = F(W)$. For $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_d) \in \mathbb{R}_+^d$, the corresponding Ornstein–Uhlenbeck semigroup is then defined by

$$\mathbf{T}_\vartheta^\boldsymbol{\kappa} F(W) = \mathbb{E}[F(B^{\vartheta, \boldsymbol{\kappa}}) \mid \mathcal{F}_T^W],$$

where

$$B_t^{\vartheta, \boldsymbol{\kappa}} := \left(e^{-\kappa_i \vartheta} W_t^i + \sqrt{1 - e^{-2\kappa_i \vartheta}} W_t^{\perp, i} \right)_{i=1, \dots, d}, \quad t \in [0, T], \quad \vartheta \geq 0,$$

and W^\perp is a standard Brownian motion on \mathbb{R}^d independent of W . The infinitesimal generator of the semigroup $(\mathbf{T}_\vartheta^\boldsymbol{\kappa})_{\vartheta \geq 0}$ is defined by

$$\mathbf{L}^\boldsymbol{\kappa} = \left. \frac{d}{d\vartheta} \right|_{\vartheta=0} \mathbf{T}_\vartheta^\boldsymbol{\kappa}.$$

Theorem 5.14 shows that, for $F(W) = \widehat{\mathbb{W}}_t$, the Ornstein–Uhlenbeck semigroup and its generator, viewed as linear operators acting on $L^2(\Omega, \mathcal{F}_T^W, \mathbb{P})$, can be identified with linear operators on the tensor algebra:

$$\begin{aligned} \mathbf{T}_\vartheta^\boldsymbol{\kappa} \widehat{\mathbb{W}}_t &= \exp \left(\sum_{i=1}^d \frac{1 - e^{-2\kappa_i \vartheta}}{2} \boldsymbol{\Psi}_{\mathbf{ii}}^0 \right) \left(\prod_{i=1}^d \mathbf{J}_i^{\kappa_i \vartheta} \right) \widehat{\mathbb{W}}_t, \\ \mathbf{L}^\boldsymbol{\kappa} \widehat{\mathbb{W}}_t &= \sum_{i=1}^d \kappa_i (\boldsymbol{\Psi}_{\mathbf{ii}}^0 - \boldsymbol{\Lambda}_i) \widehat{\mathbb{W}}_t, \end{aligned}$$

and that $\mathbf{T}_\vartheta^\boldsymbol{\kappa}$ and $\mathbf{L}^\boldsymbol{\kappa}$ map $\mathcal{P}(\widehat{\mathbb{W}}_t)$ to itself:

$$\begin{aligned} \mathbf{T}_\vartheta^\boldsymbol{\kappa} \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \rangle &= \left\langle \left(\prod_{i=1}^d \mathbf{J}_i^{\kappa_i \vartheta} \right) \exp \left(\sum_{i=1}^d \frac{1 - e^{-2\kappa_i \vartheta}}{2} \boldsymbol{\Psi}_{\mathbf{0}}^{\mathbf{ii}} \right) (\boldsymbol{\ell}), \widehat{\mathbb{W}}_t \right\rangle, \\ \mathbf{L}^\boldsymbol{\kappa} \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \rangle &= \left\langle \sum_{i=1}^d \kappa_i (\boldsymbol{\Psi}_{\mathbf{0}}^{\mathbf{ii}} - \boldsymbol{\Lambda}_i) (\boldsymbol{\ell}), \widehat{\mathbb{W}}_t \right\rangle. \end{aligned}$$

Here, the diagonal operator $\boldsymbol{\Lambda}_i$ multiplies each coordinate $\boldsymbol{\ell}^{\mathbf{v}}$ by the number $|\mathbf{v}|_i$ of occurrences of \mathbf{i} in \mathbf{v} , and \mathbf{J}_i^ϑ multiplies the coordinates $\boldsymbol{\ell}^{\mathbf{v}}$ by $e^{-|\mathbf{v}|_i \vartheta}$.

We also discuss the relation

$$\mathbf{L}^i := \mathbf{L}^{e_i} = -\delta^i \mathbf{D}^i,$$

between the generator, the Malliavin derivative, and the Skorokhod integral (defined in Section 5.2).

Integration by parts. Integration by parts is one of the key applications of Malliavin calculus. Typically, it establishes formulae of the form

$$\mathbb{E}[f'(G)F] = \mathbb{E}[f(G)\pi_T], \quad (2.6)$$

which hold for $f \in C_b^1$ and random variables F, G satisfying suitable integrability conditions. However, the Malliavin weight π_T is often intractable and cannot be computed explicitly. To address this issue, we introduce an operation

$$\diamond : T(\mathbb{R}^{d+1}) \times T(\mathbb{R}^{d+1}) \rightarrow T(\mathbb{R}^{d+1}),$$

which allows for the linearization of the L^2 scalar product between the Malliavin derivatives of random variables $F = \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_T \rangle$ and $F' = \langle \boldsymbol{\ell}', \widehat{\mathbb{W}}_T \rangle$ from $\mathcal{P}(\widehat{\mathbb{W}}_T)$:

$$\langle \mathbf{D}F, \mathbf{D}F' \rangle_{L^2(\Omega \times [0, T], \mathbb{R}^d)} = \langle \boldsymbol{\ell} \diamond \boldsymbol{\ell}', \widehat{\mathbb{W}}_T \rangle.$$

This leads to the integration by parts formula (2.2) established in Theorem 6.8 for the case

$$G = \langle \boldsymbol{\ell}^G, \widehat{\mathbb{W}}_T \rangle, \quad F = \frac{\langle \boldsymbol{\ell}_1^F, \widehat{\mathbb{W}}_T \rangle}{\langle \boldsymbol{\ell}_2^F, \widehat{\mathbb{W}}_T \rangle}, \quad \boldsymbol{\ell}^G, \boldsymbol{\ell}_1^F, \boldsymbol{\ell}_2^F \in T(\mathbb{R}^{d+1}).$$

Furthermore, we show that the Malliavin weight π_T belongs to the same class as F and provide explicit and tractable formula for it, in contrast with the typically implicit weights arising in classical Malliavin calculus.

Finally, we illustrate this result by providing formulas for the sensitivities of path-dependent options within the signature volatility framework of [Abi Jaber and Gérard \(2025a\)](#). More precisely, we consider the class of payoffs representable as linear functionals of the log-price signature, expressing the Malliavin weight as a rational function of the Brownian signature.

3 Properties of path signatures

In this subsection, we present some properties of path signatures that will be used throughout this paper.

Signature dynamics. It follows immediately from the definition (2.1) that for each $n \geq 1$, every coordinate of the signature satisfies

$$\mathbb{X}_t^{i_1 \dots i_n} = \int_0^t \mathbb{X}_s^{i_1 \dots i_{n-1}} \circ dX_s^{i_n}.$$

In compact tensor form, this yields the signature dynamics

$$d\mathbb{X}_t = \mathbb{X}_t \otimes \circ dX_t, \quad \mathbb{X}_0 = \emptyset. \quad (3.1)$$

As already mentioned in Subsection 2.1, the inverse element \mathbb{X}_t^{-1} with respect to the tensor product is the signature of the time-reversed path $\overleftarrow{X}^{[0,t]}: u \mapsto X_{t-u}$, $u \in [0, t]$. The next lemma provides the dynamics of this inverse signature.

Proposition 3.1. *The inverse signature \mathbb{X}_t^{-1} is a continuous semimartingale (elementwise) satisfying*

$$d\mathbb{X}_t^{-1} = -dX_t \otimes \mathbb{X}_t^{-1}, \quad \mathbb{X}_0^{-1} = \emptyset. \quad (3.2)$$

Proof. Differentiating the identity

$$\mathbb{X}_t \otimes \mathbb{X}_t^{-1} = \emptyset$$

and using the product rule yields

$$d\mathbb{X}_t \otimes \mathbb{X}_t^{-1} + \mathbb{X}_t \otimes d\mathbb{X}_t^{-1} = \mathbb{X}_t \otimes dX_t \otimes \mathbb{X}_t^{-1} + \mathbb{X}_t \otimes d\mathbb{X}_t^{-1} = 0.$$

Multiplying on the left by \mathbb{X}_t^{-1} gives (3.1). □

Variation of constants for signatures. Since the group increment

$$\mathbb{X}_{s,t} = \mathbb{X}_s^{-1} \otimes \mathbb{X}_t$$

plays the role of the fundamental solution, it can be used to obtain a variation-of-constants formula for $T((\mathbb{R}^d))$ -valued linear SDEs of the form

$$d\mathbb{Y}_t = \mathbb{Y}_t \otimes \circ dX_t + d\sigma_t. \quad (3.3)$$

We interpret (3) elementwise: for any truncation order N , the truncated equation is a finite-dimensional linear Stratonovich SDE whose coefficients depend only on the first N levels of \mathbb{Y}_t . Thus, if each truncated system admits a unique strong solution, we say that (3) admits a unique strong solution.

Proposition 3.2. *Let $X = (X_t)_{t \in [0, T]}$ be a continuous \mathbb{R}^d -valued semimartingale with signature \mathbb{X}_t . Suppose that each component $\sigma^\mathbf{v}$, $\mathbf{v} \in V$, is a continuous semimartingale. Then (3) admits a unique strong solution \mathbb{Y}_t , and this solution admits the following representation:*

$$\mathbb{Y}_t = \mathbb{Y}_s \mathbb{X}_{s,t} + \int_s^t \circ d\sigma_u \otimes \mathbb{X}_{u,t}. \quad (3.4)$$

Proof. For any \mathbb{Y}_t satisfying (3), observe that

$$\begin{aligned} d(\mathbb{Y}_t \otimes \mathbb{X}_t^{-1}) &= d\mathbb{Y}_t \otimes \mathbb{X}_t^{-1} + \mathbb{Y}_t \otimes d\mathbb{X}_t^{-1} \\ &= (\mathbb{Y}_t \otimes dX_t + d\sigma_t) \otimes \mathbb{X}_t^{-1} - \mathbb{Y}_t \otimes (dX_t \otimes \mathbb{X}_t^{-1}) \\ &= d\sigma_t \otimes \mathbb{X}_t^{-1}. \end{aligned}$$

Hence,

$$\mathbb{Y}_t \otimes \mathbb{X}_t^{-1} = \mathbb{Y}_s \otimes \mathbb{X}_s^{-1} + \int_s^t d\sigma_u \otimes \mathbb{X}_u^{-1}.$$

Multiplying on the right by \mathbb{X}_t and using the identity $\mathbb{X}_u^{-1} \otimes \mathbb{X}_t = \mathbb{X}_{u,t}$ yields (3.2). This proves uniqueness. Existence follows from the fact that the process \mathbb{Y}_t defined by (3.2) satisfies (3), which can be verified by direct differentiation. \square

Shuffle property. Another key feature of signatures is that products of linear functionals is itself a linear functional, as a consequence of the shuffle property. The shuffle product $\sqcup\sqcup$ is a bilinear mapping $T(\mathbb{R}^d) \times T(\mathbb{R}^d) \rightarrow T(\mathbb{R}^d)$. For words $\mathbf{v}, \mathbf{w} \in V$ and letters $\mathbf{i}, \mathbf{j} \in A$, the shuffle product is defined recursively by

$$\mathbf{vi} \sqcup\sqcup \mathbf{wj} = (\mathbf{vi} \sqcup\sqcup \mathbf{w})\mathbf{j} + (\mathbf{v} \sqcup\sqcup \mathbf{wj})\mathbf{i},$$

with the base case $\mathbf{v} \sqcup\sqcup \emptyset = \emptyset \sqcup\sqcup \mathbf{v} = \mathbf{v}$. Since $\sqcup\sqcup$ is bilinear, the definition extends to $T(\mathbb{R}^d)$: for $\ell, \mathbf{p} \in T(\mathbb{R}^d)$,

$$\ell \sqcup\sqcup \mathbf{p} := \sum_{\mathbf{v}, \mathbf{w} \in V} \ell^{\mathbf{v}} \mathbf{p}^{\mathbf{w}} (\mathbf{v} \sqcup\sqcup \mathbf{w}).$$

Proposition 3.3 (Shuffle property). *Let X be a continuous \mathbb{R}^d -valued semimartingale with signature \mathbb{X} . Then for all $\ell, \mathbf{p} \in T(\mathbb{R}^d)$ and for all $t \in [0, T]$,*

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

Proof. See (Lyons et al., 2007, Theorem 2.15). \square

The shuffle product also linearizes certain entire functions. For $\ell \in T(\mathbb{R}^d)$, the shuffle exponential is defined by

$$\exp^{\sqcup\sqcup}(\ell) := \sum_{n \geq 0} \frac{\ell^{\sqcup\sqcup n}}{n!} \in T((\mathbb{R}^d)), \quad (3.5)$$

where $\ell^{\sqcup\sqcup n}$ is the n -fold shuffle product. Using the decomposition

$$\exp^{\sqcup\sqcup}(\ell) = e^{\ell^\emptyset} \sum_{n \geq 0} \frac{(\ell - \ell^\emptyset \emptyset)^{\sqcup\sqcup n}}{n!}, \quad (3.6)$$

with the convention $\mathbf{p}^{\sqcup\sqcup 0} = \emptyset$, we see that each component contains only finitely many terms, and the shuffle exponential is well defined. Formula (3) also yields a more efficient computational method.

Expected signature of Brownian motion. The expected signature

$$\widehat{\mathcal{E}}_t := \mathbb{E}[\widehat{\mathbb{W}}_t], \quad t \in [0, T],$$

of the time-extended Brownian path $\widehat{W}: t \mapsto (t, W_t^1, \dots, W_t^d)$ can be computed explicitly. By Fawcett (2003),

$$\widehat{\mathcal{E}}_t = \exp^{\otimes} \left(\mathbf{0} + \frac{1}{2} \sum_{i=1}^d \mathbf{ii} \right), \quad (3.7)$$

where $\exp^{\otimes}(\ell) := \sum_{n \geq 0} \frac{\ell^{\otimes n}}{n!}$ for $\ell \in T(\mathbb{R}^{d+1})$.

4 Malliavin derivative of an Itô process signature

Let a^0, a^1, \dots, a^d be adapted \mathbb{R}^m -valued adapted square-integrable processes. Let us define the m -dimensional Itô process X by

$$dX_t = a_t^0 dt + \sum_{k=1}^d a_t^k \circ dW_t^k = A_t \circ d\widehat{W}_t, \quad (4.1)$$

where $(a_t^j)_{j=0, \dots, d}$ form the columns of the matrix $A_t \in \mathbb{R}^{m \times (d+1)}$ and $\widehat{W}_t = (t, W_t)$. We first aim to establish the following formula for the Malliavin derivative of X : for $\mathbb{P} \otimes ds \otimes dt$ -almost every $(\omega, s, t) \in \Omega \times [0, T]^2$,

$$\mathbf{D}_s^i X_t = a_s^i \mathbb{1}_{[0, t]}(s) + \int_s^t \mathbf{D}_s^i a_u^0 du + \sum_{k=1}^d \int_s^t \mathbf{D}_s^i a_u^k \circ dW_u^k, \quad (4.2)$$

similar to the corresponding formula in the Itô integral case (Alós et al., 2021, Section 3.2.1). Treating Stratonovich integrals requires stronger regularity assumptions on the integrands a^k than in the Itô case.

Assumption 4.1. *We assume that*

1. $a^0 \in \mathbb{L}^{1,2}$, that is, $a^0 \in L^2(\Omega \times [0, T])$ such that $a_t^0 \in \mathbb{D}^{1,2}$ for a.e. $t \in [0, T]$, and

$$\|a^0\|_{\mathbb{L}^{1,2}}^2 := \mathbb{E} \left[\int_0^T |a_t^0|^2 dt + \int_0^T \int_0^T |\mathbf{D}_s a_t^0|^2 ds dt \right] < \infty.$$

2. For $k = 1, \dots, d$, $a^k \in \mathbb{L}^{2,2}$, that is, $a^k \in L^2(\Omega \times [0, T])$, $a_t^k \in \mathbb{D}^{2,2}$ for a.e. $t \in [0, T]$, and

$$\|a^k\|_{\mathbb{L}^{2,2}}^2 := \mathbb{E} \left[\int_0^T |a_t^k|^2 dt + \int_0^T \int_0^T |\mathbf{D}_s a_t^k|^2 ds dt + \int_0^T \int_0^T \int_0^T |\mathbf{D}_r \mathbf{D}_s a_t^k|^2 dr ds dt \right] < \infty. \quad (4.3)$$

3. For $k = 1, \dots, d$, there exist processes $\mathbf{D}^+ a^k, \mathbf{D}^- a^k \in L^2(\Omega \times [0, T])$ such that

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_0^T \sup_{\substack{s < u < T \\ |u-s| \leq 1/n}} \mathbb{E} \left[|\mathbf{D}_s a_u^k - (\mathbf{D}^+ a^k)_s|^2 \right] ds &= 0, \\ \lim_{n \rightarrow \infty} \int_0^T \sup_{\substack{0 < u < s \\ |s-u| \leq 1/n}} \mathbb{E} \left[|\mathbf{D}_s a_u^k - (\mathbf{D}^- a^k)_s|^2 \right] ds &= 0. \end{aligned} \quad (4.4)$$

4. For $k = 1, \dots, d$, the processes $\mathbf{D}^+ a^k$ and $\mathbf{D}^- a^k$ belong to $\mathbb{L}^{1,2}$, and their Malliavin derivatives satisfy

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_0^T \int_0^T \sup_{\substack{s < u < T \\ |u-s| \leq 1/n}} \mathbb{E} \left[|\mathbf{D}_s \mathbf{D}_r a_u^k - \mathbf{D}_r (\mathbf{D}^+ a^k)_s|^2 \right] ds dr &= 0, \\ \lim_{n \rightarrow \infty} \int_0^T \int_0^T \sup_{\substack{0 < u < s \\ |s-u| \leq 1/n}} \mathbb{E} \left[|\mathbf{D}_s \mathbf{D}_r a_u^k - \mathbf{D}_r (\mathbf{D}^- a^k)_s|^2 \right] ds dr &= 0. \end{aligned} \quad (4.5)$$

Remark 4.2. Condition (4) is a natural extension of the regularity class $\mathbb{L}_2^{1,2}$, that is, $\mathbb{L}^{1,2}$ together with the Malliavin derivative trace regularity condition (3), which is required for the Stratonovich integrability of a^k (see (Nualart, 2006, Section 3.1.1)). In the present setting, we aim to prove not only the Stratonovich integrability of a^k , but also the differentiability of the associated Stratonovich integral, and the well-posedness of the map

$$(r, t) \mapsto \int_0^t \mathbf{D}_r a_u^k \circ dW_u^k$$

as an $L^2(\Omega \times [0, T]^2)$ -valued process rather than merely as a random variable. Accordingly, condition (4) arises naturally as a derivative trace condition for the integrand $\mathbf{D}_r a_t^k$. This extension of the results of Nualart (2006) will be used in Theorem 4.4 to derive the Malliavin derivatives of iterated Stratonovich integrals, for which a process-based point of view appears more natural. ■

Theorem 4.3. Under Assumption 4.1, we have $X \in \mathbb{L}^{1,2}$, and (4) holds.

Proof. The proof is given in Appendix A. □

We next show that a formula analogous to (4) also holds for the signature \mathbb{X} of X .

Theorem 4.4. Let X be an Itô process defined by (4). Assume that $a^{\mathbb{X}} := (\mathbb{X} \otimes a^i)_{i=0,1,\dots,d}$ satisfies Assumption 4.1 elementwise, that is, coordinatewise with respect to each $\mathbf{v} \in V$. Then, $\mathbb{X} \in \mathbb{L}^{1,2}$ elementwise and for $\mathbb{P} \otimes ds \otimes dt$ -almost every $(\omega, s, t) \in \Omega \times [0, T]^2$ with $s \leq t$,

$$\mathbf{D}_s \mathbb{X}_t = \mathbb{X}_s \otimes A_s \otimes \mathbb{X}_{s,t} + \int_s^t \mathbb{X}_u \otimes (\mathbf{D}_s A_u \circ d\widehat{W}_u) \otimes \mathbb{X}_{u,t},$$

or, in coordinate form, for $i = 1, \dots, d$,

$$\mathbf{D}_s^i \mathbb{X}_t = \mathbb{X}_s \otimes a_s^i \otimes \mathbb{X}_{s,t} + \sum_{k=0}^d \int_s^t (\mathbb{X}_u \otimes \mathbf{D}_s^i a_u^k \otimes \mathbb{X}_{u,t}) \circ d\widehat{W}_u^k. \quad (4.6)$$

Proof. The dynamics of the signature \mathbb{X} is given by (3), so that

$$\mathbb{X}_t = \emptyset + \int_0^t \mathbb{X}_u \otimes \circ dX_u = \emptyset + \sum_{k=0}^d \int_0^t \mathbb{X}_u \otimes a_u^k \circ d\widehat{W}_u^k.$$

In particular \mathbb{X} satisfies the assumptions from Theorem 4.3 elementwise and $\mathbb{X} \in \mathbb{L}^{1,2}$. Taking the Malliavin derivative \mathbf{D}_s^i and applying the formula (4) for the Malliavin derivative of Itô processes elementwise yields

$$\mathbf{D}_s^i \mathbb{X}_t = \mathbb{X}_s \otimes a_s^i + \int_s^t \mathbf{D}_s^i \mathbb{X}_u \otimes \circ dX_u + \sum_{k=0}^d \int_s^t \mathbb{X}_u \otimes \mathbf{D}_s^i a_u^k \circ d\widehat{W}_u^k.$$

Applying the variation-of-constants formula from Proposition 3.2 to $\mathbb{Y}_t = \mathbf{D}_s^i \mathbb{X}_t$ yields (4.4). □

Remark 4.5. Theorem 4.4 is primarily a structural result: it identifies the form of $\mathbf{D}\mathbb{X}$ and connects it to the known expression (4) for Itô processes, provided the derivative exists. However, it does not furnish tractable conditions ensuring this existence. Nevertheless, as demonstrated in Corollary 4.7 below, the theorem applies to the time-augmented Brownian motion \widehat{W} , the main process of interest in this paper, since the integrability of the Brownian signature $\widehat{\mathbb{W}}$ and its derivatives can be verified directly. ■

Remark 4.6. In the expression (4.4), the stochastic integrals are to be understood as standard Stratonovich integrals with predictable integrands, namely

$$\left[\int_s^t (\mathbb{X}_u \otimes \mathbf{D}_s^i a_u^k \otimes \mathbb{X}_u^{-1}) \circ d\widehat{W}_u^k \right] \otimes \mathbb{X}_t. \quad (4.7)$$

However, we will often write expressions in which anticipative terms appear inside the integrals, as in (4.4). Such Stratonovich integrals with anticipative integrands can still be defined as limits of Riemann sums (see Appendix A and Nualart (2006, Section 3.1)). Since the resulting integrals coincide, we will use (4.6) and (4.4) interchangeably throughout the paper. ■

Corollary 4.7. If X is a time-augmented Brownian motion $X_t = \widehat{W}_t$, then $a_t^k \equiv \mathbf{k}$, so that $\mathbf{D}_s a_t^k \equiv 0$, and, for $\mathbb{P} \otimes ds \otimes dt$ -almost every $(\omega, s, t) \in \Omega \times [0, T]^2$ with $s \leq t$,

$$\mathbf{D}_s^i \widehat{\mathbb{W}}_t = \widehat{\mathbb{W}}_s \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{s,t}. \quad (4.8)$$

Moreover, $\widehat{\mathbb{W}}_t \in \mathbb{D}^{\infty,2}$ for a.e. $t \in [0, T]$, and, for $\mathbb{P} \otimes dt^{\otimes(n+1)}$ -almost every $(\omega, s_1, \dots, s_n, t) \in \Omega \times [0, T]^{n+1}$ with $s_1 \leq \dots \leq s_n \leq t$, we have

$$\mathbf{D}_{s_1}^{i_1} \cdots \mathbf{D}_{s_n}^{i_n} \widehat{\mathbb{W}}_t = \widehat{\mathbb{W}}_{s_1} \otimes \mathbf{i}_{i_1} \otimes \widehat{\mathbb{W}}_{s_1, s_2} \otimes \mathbf{i}_{i_2} \otimes \cdots \otimes \mathbf{i}_{i_n} \otimes \widehat{\mathbb{W}}_{s_n, t}. \quad (4.9)$$

Proof. The formula (4.7) follows from a recursive application of Theorem 4.4. For $n = 1$, (4.7) is given directly by Theorem 4.4. Suppose now that (4.7) holds for $n - 1$. Then, applying the product rule for the Malliavin derivative, we obtain

$$\begin{aligned} \mathbf{D}_{s_1}^{i_1} \cdots \mathbf{D}_{s_n}^{i_n} \widehat{\mathbb{W}}_t &= \mathbf{D}_{s_1}^{i_1} (\widehat{\mathbb{W}}_{s_2} \otimes \mathbf{i}_2 \otimes \widehat{\mathbb{W}}_{s_2, s_3} \otimes \mathbf{i}_3 \otimes \cdots \otimes \mathbf{i}_n \otimes \widehat{\mathbb{W}}_{s_n, t}) \\ &= (\mathbf{D}_{s_1}^{i_1} \widehat{\mathbb{W}}_{s_2}) \otimes \mathbf{i}_2 \otimes \widehat{\mathbb{W}}_{s_2, s_3} \otimes \mathbf{i}_3 \otimes \cdots \otimes \mathbf{i}_n \otimes \widehat{\mathbb{W}}_{s_n, t} \\ &= \widehat{\mathbb{W}}_{s_1} \otimes \mathbf{i}_1 \otimes \widehat{\mathbb{W}}_{s_1, s_2} \otimes \mathbf{i}_2 \otimes \widehat{\mathbb{W}}_{s_2, s_3} \otimes \mathbf{i}_3 \otimes \cdots \otimes \mathbf{i}_n \otimes \widehat{\mathbb{W}}_{s_n, t}, \end{aligned}$$

where the second equality follows from the fact that $\widehat{\mathbb{W}}_{s_k, s_{k+1}}$ is measurable with respect to $\sigma((W_u - W_{s_k}), u \in [s_k, s_{k+1}])$, so that $\mathbf{D}_{s_1} \widehat{\mathbb{W}}_{s_k, s_{k+1}} = 0$ for all $k \geq 2$; see (Nualart, 2006, Corollary 1.2.1).

It remains to prove the integrability of the Malliavin derivatives. Indeed, for all $\mathbf{v} \in V$,

$$\begin{aligned} \mathbb{E} \left[(\mathbf{D}_{s_1}^{i_1} \cdots \mathbf{D}_{s_n}^{i_n} \widehat{\mathbb{W}}_t^{\mathbf{v}})^2 \right] &= \mathbb{E} \left[\left((\widehat{\mathbb{W}}_{s_1} \otimes \mathbf{i}_1 \otimes \widehat{\mathbb{W}}_{s_1, s_2} \otimes \mathbf{i}_2 \otimes \cdots \otimes \mathbf{i}_n \otimes \widehat{\mathbb{W}}_{s_n, t})^{\mathbf{v}} \right)^2 \right] \\ &= \mathbb{E} \left[\left(\sum_{\substack{\mathbf{w}_0, \dots, \mathbf{w}_n \\ \mathbf{w}_0 \mathbf{i}_1 \mathbf{w}_1 \dots \mathbf{w}_n \mathbf{i}_n = \mathbf{v}}} \widehat{\mathbb{W}}_{s_1}^{\mathbf{w}_0} \widehat{\mathbb{W}}_{s_1, s_2}^{\mathbf{w}_1} \cdots \widehat{\mathbb{W}}_{s_n, t}^{\mathbf{w}_n} \right)^2 \right] \\ &\leq \mathbb{E} \left[\left(\sum_{\substack{\mathbf{w}_0, \dots, \mathbf{w}_n \\ |\mathbf{w}_0|, \dots, |\mathbf{w}_n| < |\mathbf{v}|}} |\widehat{\mathbb{W}}_{s_1}^{\mathbf{w}_0} \widehat{\mathbb{W}}_{s_1, s_2}^{\mathbf{w}_1} \cdots \widehat{\mathbb{W}}_{s_n, t}^{\mathbf{w}_n}| \right)^2 \right] \\ &= \mathbb{E} \left[\left(\sum_{|\mathbf{w}_0| < |\mathbf{v}|} |\widehat{\mathbb{W}}_{s_1}^{\mathbf{w}_0}| \right)^2 \right] \cdots \mathbb{E} \left[\left(\sum_{|\mathbf{w}_n| < |\mathbf{v}|} |\widehat{\mathbb{W}}_{s_n, t}^{\mathbf{w}_n}| \right)^2 \right] \\ &\leq \mathbb{E} \left[\left(\sum_{|\mathbf{w}| < |\mathbf{v}|} |\widehat{\mathbb{W}}_T^{\mathbf{w}}| \right)^2 \right]^{n+1} < \infty, \end{aligned}$$

where, in the last inequality, we used the diffusive scaling of the signature $(\widehat{\mathbb{W}}_t^{\mathbf{v}})_{\mathbf{v} \in V} \stackrel{d}{=} \left(t^{\frac{|\mathbf{v}| + |\mathbf{v}|_0}{2}} \widehat{\mathbb{W}}_1^{\mathbf{v}} \right)_{\mathbf{v} \in V}$, where $|\mathbf{v}|_0$ denotes the number of occurrences of the letter $\mathbf{0}$ in the word \mathbf{v} . Therefore, $\mathbb{E} \left[(\mathbf{D}_{s_1}^{i_1} \cdots \mathbf{D}_{s_n}^{i_n} \widehat{\mathbb{W}}_t^{\mathbf{v}})^2 \right]$ is bounded uniformly in $s_1, \dots, s_n, t \in [0, T]$ and $i_1, \dots, i_n \in \{1, \dots, d\}$, so that $\widehat{\mathbb{W}}_t \in \mathbb{D}^{n, 2}$ for all $n \geq 1$. \square

5 Malliavin calculus for $\widehat{\mathbb{W}}$

We now focus on the signature of the time-augmented Brownian motion $\widehat{W}: t \mapsto (t, W_t)$. This object plays the role of Brownian motion on path space in the sense of (Baudoïn, 2004, Definition 1.2) and hence inherits many of its nice properties as we will show in this section.

5.1 Iterated integrals of Malliavin derivatives of $\widehat{\mathbb{W}}_t$

We have already seen that the higher-order Malliavin derivative of the Brownian signature reduces to the ‘‘insertion’’ of letters in Chen’s identity, see (4.7). Our goal is to show that this operation can be, in some sense, inverted by integrating $\mathbf{D}_{s_1}^{i_1} \cdots \mathbf{D}_{s_n}^{i_n} \widehat{\mathbb{W}}_t$ over the simplex $0 \leq s_1 \leq \cdots \leq s_n \leq t$ with respect to the components of $d\widehat{W}$.

In particular, this will allow us to show in Theorem 5.6 that if $F = \langle \ell, \widehat{\mathbb{W}}_t \rangle \in \mathcal{P}(\widehat{\mathbb{W}}_t)$, recall (2.2), then

$$\int_{0 \leq s_1 \leq \cdots \leq s_n \leq t} \mathbf{D}_{s_1}^{i_1} \cdots \mathbf{D}_{s_n}^{i_n} \langle \ell, \widehat{\mathbb{W}}_t \rangle \circ d\widehat{W}_{s_1}^{j_1} \cdots \circ d\widehat{W}_{s_n}^{j_n} \in \mathcal{P}(\widehat{\mathbb{W}}_t),$$

and its coefficients can be expressed explicitly in terms of ℓ . This property will be crucial for deriving closed-form formulae for the Ornstein–Uhlenbeck semigroup and integration by parts.

We start with the simplest case $\widehat{\mathbb{W}}_s \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{s,t}$, which can be written as

$$\widehat{\mathbb{W}}_s \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{s,t} = \sum_{\mathbf{v}_1, \mathbf{v}_2 \in V} \widehat{\mathbb{W}}_s^{\mathbf{v}_1} \widehat{\mathbb{W}}_{s,t}^{\mathbf{v}_2}(\mathbf{v}_1 \mathbf{i} \mathbf{v}_2). \quad (5.1)$$

The behavior of the products $\widehat{\mathbb{W}}_s^{\mathbf{v}_1} \widehat{\mathbb{W}}_{s,t}^{\mathbf{v}_2}$ under integration with respect to s is described in the following lemma, which holds for general path signatures.

Lemma 5.1. *Let $X = (X_t)_{t \in [0, T]}$ be an \mathbb{R}^d -valued continuous semimartingale. Then, for all $\mathbf{v}_1, \mathbf{v}_2 \in V$ and $j \in \{1, \dots, d\}$,*

$$\int_0^t \mathbb{X}_s^{\mathbf{v}_1} \mathbb{X}_{s,t}^{\mathbf{v}_2} \circ dX_s^j = \mathbb{X}_t^{\mathbf{v}_1 \mathbf{j} \mathbf{v}_2}, \quad t \in [0, T].$$

Consequently, for $\ell_1, \ell_2 \in T(\mathbb{R}^d)$,

$$\int_0^t \langle \ell_1, \mathbb{X}_s \rangle \langle \ell_2, \mathbb{X}_{s,t} \rangle \circ dX_s^j = \langle \ell_1 \otimes \mathbf{j} \otimes \ell_2, \mathbb{X}_t \rangle.$$

Proof. For $\mathbf{v}_1, \mathbf{v}_2 \in V$, set $\mathbb{Y}_t^{\mathbf{v}_1 \mathbf{j} \mathbf{v}_2} := \int_0^t \mathbb{X}_s^{\mathbf{v}_1} \mathbb{X}_{s,t}^{\mathbf{v}_2} \circ dX_s^j$. We proceed by induction on $|\mathbf{v}_2|$ to show that

$$\mathbb{Y}_t^{\mathbf{v}_1 \mathbf{j} \mathbf{v}_2} = \mathbb{X}_t^{\mathbf{v}_1 \mathbf{j} \mathbf{v}_2}, \quad t \in [0, T]. \quad (5.2)$$

If $\mathbf{v}_2 = \emptyset$, (5.1) holds by definition of the signature. Otherwise, suppose that $\mathbf{v}_2 = \mathbf{v}'_2 \mathbf{k}$ for some $\mathbf{k} \in A$, and that (5.1) holds for $\mathbb{Y}_t^{\mathbf{v}_1 \mathbf{j} \mathbf{v}'_2}$. Then,

$$\begin{aligned} \mathbb{Y}_t^{\mathbf{v}_1 \mathbf{j} \mathbf{v}_2} &= \int_0^t \mathbb{X}_s^{\mathbf{v}_1} \mathbb{X}_{s,t}^{\mathbf{v}_2} \circ dX_s^j = \int_0^t \mathbb{X}_s^{\mathbf{v}_1} \left(\int_s^t \mathbb{X}_{s,u}^{\mathbf{v}'_2} \circ dX_u^k \right) \circ dX_s^j \\ &= \int_0^t \left(\int_0^u \mathbb{X}_s^{\mathbf{v}_1} \mathbb{X}_{s,u}^{\mathbf{v}'_2} \circ dX_s^j \right) \circ dX_u^k = \int_0^t \mathbb{Y}_u^{\mathbf{v}_1 \mathbf{j} \mathbf{v}'_2} \circ dX_u^k \\ &= \int_0^t \mathbb{X}_u^{\mathbf{v}_1 \mathbf{j} \mathbf{v}'_2} \circ dX_u^k = \mathbb{X}_t^{\mathbf{v}_1 \mathbf{j} \mathbf{v}_2}, \end{aligned}$$

where we applied Fubini's theorem and the induction hypothesis. \square

Integrating (5.1) and applying Lemma 5.1, we obtain

$$\int_0^t (\widehat{\mathbb{W}}_s \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{s,t}) \circ d\widehat{\mathbb{W}}_s^j = \sum_{\mathbf{v}_1, \mathbf{v}_2 \in V} \widehat{\mathbb{W}}_t^{\mathbf{v}_1 \mathbf{j} \mathbf{v}_2}(\mathbf{v}_1 \mathbf{i} \mathbf{v}_2).$$

This motivates the following definition.

Definition 5.2. *Fix $n \in \mathbb{N}$ and words $\mathbf{w}_1, \dots, \mathbf{w}_n \in V$ and $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$. We define the switching operator $\Psi_{\mathbf{u}_1, \dots, \mathbf{u}_n}^{\mathbf{w}_1, \dots, \mathbf{w}_n} : T((\mathbb{R}^d)) \rightarrow T((\mathbb{R}^d))$ by*

$$\Psi_{\mathbf{u}_1, \dots, \mathbf{u}_n}^{\mathbf{w}_1, \dots, \mathbf{w}_n} : \mathbb{X} \mapsto \sum_{\mathbf{v}_0, \dots, \mathbf{v}_n} \mathbb{X}^{\mathbf{v}_0 \mathbf{w}_1 \mathbf{v}_1 \dots \mathbf{w}_n \mathbf{v}_n}(\mathbf{v}_0 \mathbf{u}_1 \mathbf{v}_1 \dots \mathbf{u}_n \mathbf{v}_n).$$

Note that $\Psi_{\mathbf{u}_1, \dots, \mathbf{u}_n}^{\mathbf{w}_1, \dots, \mathbf{w}_n}$ can be equivalently defined by its action on the basis vectors:

$$\Psi_{\mathbf{u}_1, \dots, \mathbf{u}_n}^{\mathbf{w}_1, \dots, \mathbf{w}_n}(\mathbf{v}) = \sum_{\substack{\mathbf{v}_0, \dots, \mathbf{v}_n \\ \mathbf{v}_0 \mathbf{w}_1 \mathbf{v}_1 \dots \mathbf{w}_n \mathbf{v}_n = \mathbf{v}}} \mathbf{v}_0 \mathbf{u}_1 \mathbf{v}_1 \dots \mathbf{u}_n \mathbf{v}_n, \quad \mathbf{v} \in V.$$

Hence, the switching operator $\Psi_{\mathbf{u}_1, \dots, \mathbf{u}_n}^{\mathbf{w}_1, \dots, \mathbf{w}_n}$ admits a natural interpretation: it finds all possible sequences of consecutive and non-intersecting sub-strings $\mathbf{w}_1, \dots, \mathbf{w}_n$ in the word \mathbf{v} , and replaces them with $\mathbf{u}_1, \dots, \mathbf{u}_n$.

The result is the sum of all such modified words. In particular, $\Psi_{\mathbf{i}}^{\mathbf{j}}$ replaces each occurrence of the letter \mathbf{j} by the letter \mathbf{i} , one at a time. For example,

$$\Psi_{\mathbf{0}}^{\mathbf{1}}(\mathbf{01101}) = \mathbf{00101} + \mathbf{01001} + \mathbf{01100}.$$

The switching operator allows us to rewrite compactly the iterated integrals we are interested in.

Theorem 5.3. *Let $X = (X_t)_{t \in [0, T]}$ be an \mathbb{R}^d -valued continuous semimartingale and let \mathbb{X} denote its signature. Then,*

$$\int_{0 \leq s_1 \leq \dots \leq s_n \leq t} \left(\mathbb{X}_{s_1} \otimes \mathbf{i}_1 \otimes \mathbb{X}_{s_1, s_2} \otimes \dots \otimes \mathbf{i}_n \otimes \mathbb{X}_{s_n, t} \right) \circ dX_{s_1}^{j_1} \dots \circ dX_{s_n}^{j_n} = \Psi_{\mathbf{i}_1, \dots, \mathbf{i}_n}^{\mathbf{j}_1, \dots, \mathbf{j}_n}(\mathbb{X}_t). \quad (5.3)$$

Proof. It follows directly from the definition of the tensor product that

$$\mathbb{X}_{s_1} \otimes \mathbf{i}_1 \otimes \mathbb{X}_{s_1, s_2} \otimes \dots \otimes \mathbf{i}_n \otimes \mathbb{X}_{s_n, t} = \sum_{\mathbf{v}_0, \dots, \mathbf{v}_n \in V} \mathbb{X}_{s_1}^{\mathbf{v}_0} \mathbb{X}_{s_1, s_2}^{\mathbf{v}_1} \dots \mathbb{X}_{s_n, t}^{\mathbf{v}_n} (\mathbf{v}_0 \mathbf{i}_1 \mathbf{v}_1 \dots \mathbf{i}_n \mathbf{v}_n).$$

Integrating and applying Lemma 5.1 n times concludes the proof. \square

Remark 5.4. Theorem 5.3 still applies when the order of the iterated integral differs from the number of ‘‘holes’’ in the ‘‘pierced’’ signature. More precisely, if the integral is taken only over a subset of consecutive variables $s_{k_0+1} \leq \dots \leq s_{k_0+m}$ for some $0 \leq k_0 \leq \dots \leq k_0 + m \leq n$, we have

$$\begin{aligned} & \int_{s_{k_0} \leq s_{k_0+1} \leq \dots \leq s_{k_0+m} \leq s_{k_0+m+1}} \left(\mathbb{X}_{s_1} \otimes \mathbf{i}_1 \otimes \mathbb{X}_{s_1, s_2} \otimes \dots \otimes \mathbf{i}_n \otimes \mathbb{X}_{s_n, t} \right) \circ dX_{s_{k_0+1}}^{j_{k_0+1}} \dots \circ dX_{s_{k_0+m}}^{j_{k_0+m}} \\ &= \mathbb{X}_{s_1} \otimes \mathbf{i}_1 \otimes \dots \otimes \mathbf{i}_{k_0} \otimes \Psi_{\mathbf{i}_{k_0+1}, \dots, \mathbf{i}_{k_0+m}}^{\mathbf{j}_{k_0+1}, \dots, \mathbf{j}_{k_0+m}}(\mathbb{X}_{s_{k_0}, s_{k_0+m+1}}) \otimes \mathbf{i}_{k_0+m+1} \otimes \dots \otimes \mathbf{i}_n \otimes \mathbb{X}_{s_n, t}, \end{aligned}$$

where we set $s_0 = 0$ and $s_{n+1} = t$. \blacksquare

We will also be interested in computing the linear functionals of objects of the form (5.3) and to write them as linear functionals of \mathbb{X}_t . In other words, we would like to compute the adjoint of $\Psi_{\mathbf{i}_1, \dots, \mathbf{i}_n}^{\mathbf{j}_1, \dots, \mathbf{j}_n}$. The following proposition answers this question.

Proposition 5.5. *The adjoint operator $(\Psi_{\mathbf{u}_1, \dots, \mathbf{u}_n}^{\mathbf{w}_1, \dots, \mathbf{w}_n})^*$ is given by*

$$(\Psi_{\mathbf{u}_1, \dots, \mathbf{u}_n}^{\mathbf{w}_1, \dots, \mathbf{w}_n})^* = \Psi_{\mathbf{w}_1, \dots, \mathbf{w}_n}^{\mathbf{u}_1, \dots, \mathbf{u}_n}.$$

That is, for $\ell \in T(\mathbb{R}^d)$, we have

$$\langle \ell, \Psi_{\mathbf{u}_1, \dots, \mathbf{u}_n}^{\mathbf{w}_1, \dots, \mathbf{w}_n}(\mathbb{X}_t) \rangle = \langle \Psi_{\mathbf{w}_1, \dots, \mathbf{w}_n}^{\mathbf{u}_1, \dots, \mathbf{u}_n}(\ell), \mathbb{X}_t \rangle.$$

Proof. The proof follows from observing that

$$\langle \ell, \Psi_{\mathbf{u}_1, \dots, \mathbf{u}_n}^{\mathbf{w}_1, \dots, \mathbf{w}_n}(\mathbb{X}_t) \rangle = \sum_{\mathbf{v}_0, \dots, \mathbf{v}_n} \ell^{\mathbf{v}_0 \mathbf{u}_1 \mathbf{v}_1 \dots \mathbf{u}_n \mathbf{v}_n} \mathbb{X}_t^{\mathbf{v}_0 \mathbf{w}_1 \mathbf{v}_1 \dots \mathbf{w}_n \mathbf{v}_n} = \langle \Psi_{\mathbf{w}_1, \dots, \mathbf{w}_n}^{\mathbf{u}_1, \dots, \mathbf{u}_n}(\ell), \mathbb{X}_t \rangle.$$

\square

Theorem 5.3 and Proposition 5.5 imply immediately the following result.

Theorem 5.6. *For all $i_1, \dots, i_n \in \{1, \dots, d\}$ and $j_1, \dots, j_n \in \{0, 1, \dots, d\}$, we have*

$$\int_{0 \leq s_1 \leq \dots \leq s_n \leq t} (\mathbf{D}_{s_1}^{i_1} \dots \mathbf{D}_{s_n}^{i_n} \widehat{\mathbb{W}}_t) \circ d\widehat{W}_{s_1}^{j_1} \dots \circ d\widehat{W}_{s_n}^{j_n} = \Psi_{\mathbf{i}_1, \dots, \mathbf{i}_n}^{\mathbf{j}_1, \dots, \mathbf{j}_n}(\widehat{\mathbb{W}}_t).$$

Moreover, for all $\ell \in T(\mathbb{R}^{d+1})$,

$$\int_{0 \leq s_1 \leq \dots \leq s_n \leq t} \mathbf{D}_{s_1}^{i_1} \dots \mathbf{D}_{s_n}^{i_n} \langle \ell, \widehat{\mathbb{W}}_t \rangle \circ d\widehat{W}_{s_1}^{j_1} \dots \circ d\widehat{W}_{s_n}^{j_n} = \langle \Psi_{\mathbf{j}_1, \dots, \mathbf{j}_n}^{\mathbf{i}_1, \dots, \mathbf{i}_n}(\ell), \widehat{\mathbb{W}}_t \rangle.$$

5.2 Skorokhod integration and the representation of random variables in $\mathcal{P}(\widehat{\mathbb{W}}_T)$

We now illustrate how the derived results can be applied to obtain explicit representations of the (anticipative) Skorokhod integral, the Clark–Ocone formula, and the chaos expansion for random variables $F \in \mathcal{P}(\widehat{\mathbb{W}}_T)$.

Skorokhod integration. The Skorokhod integral δ is defined by the dual relationship

$$\mathbb{E} \left[\int_0^T \langle \mathbf{D}_t F, h_t \rangle_{\mathbb{R}^d} dt \right] = \mathbb{E}[F \delta(h)], \quad (5.4)$$

for $F \in \mathbb{D}^{1,2}$ and $h \in \text{Dom}(\delta)$, where

$$\text{Dom}(\delta) := \left\{ h \in L^2(\Omega \times [0, T], \mathbb{R}^d) : \left| \mathbb{E} \left[\int_0^T \langle \mathbf{D}_t F, h_t \rangle_{\mathbb{R}^d} dt \right] \right| \leq C \|F\|_{L^2(\Omega)}, \forall F \in \mathbb{D}^{1,2} \right\}.$$

Moreover, for $h = (h^1, \dots, h^d) \in \text{Dom}(\delta)$ we may write

$$\delta(h) = \sum_{j=1}^d \delta^j(h^j),$$

where, for each $j \in \{1, \dots, d\}$ and all $F \in \mathbb{D}^{1,2}$,

$$\mathbb{E} \left[\int_0^T \mathbf{D}_t^j F h_t^j dt \right] = \mathbb{E} [F \delta^j(h^j)].$$

If $F \in \mathbb{D}^{1,2}$ and $h \in \text{Dom}(\delta)$ are such that $Fh \in L_2(\Omega \times [0, T], \mathbb{R}^d)$, then $Fh \in \text{Dom}(\delta)$ and the following equality (Nualart, 2006, Proposition 1.3.3) holds:

$$\delta(Fh) = F \delta(h) - \int_0^T \langle \mathbf{D}_s F, h_s \rangle_{\mathbb{R}^d} ds, \quad (5.5)$$

provided the right-hand side is square integrable. This equality plays an important role, in particular, for proving the integration by parts results discussed in Section 6.

We show that the Skorokhod integral of $\widehat{\mathbb{W}}_T$, and hence of any random variable in $\mathcal{P}(\widehat{\mathbb{W}}_T)$, can be expressed using the switching operator $\Psi_{\mathbf{i}}^{\mathbf{0}}$.

Proposition 5.7. For $i = 1, \dots, d$, we have

$$\delta^i(\widehat{\mathbb{W}}_T) = W_T^i \widehat{\mathbb{W}}_T - \Psi_{\mathbf{i}}^{\mathbf{0}}(\widehat{\mathbb{W}}_T).$$

Moreover, if $\ell \in T(\mathbb{R}^{d+1})$, then

$$\delta^i \left(\langle \ell, \widehat{\mathbb{W}}_T \rangle \right) = \langle \ell \sqcup \mathbf{i} - \Psi_{\mathbf{0}}^{\mathbf{i}}(\ell), \widehat{\mathbb{W}}_T \rangle. \quad (5.6)$$

Proof. Applying the integration-by-parts formula (5.2) element-wise with $F = \widehat{\mathbb{W}}_T$ and $h_t \equiv 1$ yields

$$\begin{aligned} \delta^i(\widehat{\mathbb{W}}_T) &= W_T^i \widehat{\mathbb{W}}_T - \int_0^T \mathbf{D}_t^i \widehat{\mathbb{W}}_T dt \\ &= W_T^i \widehat{\mathbb{W}}_T - \Psi_{\mathbf{i}}^{\mathbf{0}}(\widehat{\mathbb{W}}_T), \end{aligned}$$

where we used Theorem 5.6 to compute the integral. Equation (5.7) then follows from Proposition 3.3 and Proposition 5.5. \square

Clark–Ocone formula. The classical Clark–Ocone formula refines the martingale representation theorem. It states that any $F \in \mathbb{D}^{1,2}$ admits the representation

$$F = \mathbb{E}[F] + \sum_{i=1}^d \int_0^T \mathbb{E}[\mathbf{D}_t^i F | \mathcal{F}_t^W] dW_t^i. \quad (5.7)$$

If the random variable F lies in $\mathcal{P}(\widehat{\mathbb{W}}_T)$, the integrand in (5.2) can be computed explicitly.

Proposition 5.8. *Let $F = \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_T \rangle$ with $\boldsymbol{\ell} \in T(\mathbb{R}^{d+1})$. Then, $F \in \mathbb{D}^{1,2}$ and*

$$F = \langle \boldsymbol{\ell}, \widehat{\mathcal{E}}_T \rangle + \sum_{i=1}^d \int_0^T \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \otimes \mathbf{i} \otimes \widehat{\mathcal{E}}_{T-t} \rangle dW_t^i,$$

where $\widehat{\mathcal{E}}_t$ denotes the expected signature given by (3).

Proof. Equation (4.7) yields

$$\mathbb{E}[\mathbf{D}_t^i F | \mathcal{F}_t^W] = \mathbb{E} \left[\langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{t,T} \rangle \middle| \mathcal{F}_t^W \right] = \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \otimes \mathbf{i} \otimes \widehat{\mathcal{E}}_{T-t} \rangle,$$

since $\widehat{\mathbb{W}}_t$ is \mathcal{F}_t^W -measurable and $\widehat{\mathbb{W}}_{t,T}$ is independent of \mathcal{F}_t^W , with $\widehat{\mathbb{W}}_{t,T} \stackrel{d}{=} \widehat{\mathbb{W}}_{T-t}$. \square

Chaos expansion. The chaos expansion (Nualart, 2006, Theorem 1.1.2) expresses any square-integrable random variable $F \in L^2(\Omega, \mathcal{F}^W, \mathbb{P})$ as an infinite series of iterated Wiener–Itô integrals:

$$F = \mathbb{E}[F] + \sum_{n \geq 1} n! \sum_{\mathbf{i}_1 \dots \mathbf{i}_n \in V_n} \int \dots \int_{0 < s_1 < \dots < s_n < t} f_{\mathbf{i}_1 \dots \mathbf{i}_n}(s_1, \dots, s_n) dW_{s_1}^{i_1} \dots dW_{s_n}^{i_n}.$$

Moreover, when F is infinitely Malliavin differentiable (i.e. $F \in \mathbb{D}^{\infty,2}$), the kernel functions $f_{\mathbf{i}_1 \dots \mathbf{i}_n}$ can be computed (Nualart, 2006, Exercise 1.2.6) via

$$f_{\mathbf{i}_1 \dots \mathbf{i}_n}(s_1, \dots, s_n) := \frac{1}{n!} \mathbb{E}[\mathbf{D}_{s_1}^{i_1} \dots \mathbf{D}_{s_n}^{i_n} F], \quad \mathbf{i}_1 \dots \mathbf{i}_n \in V \setminus \{\mathbf{0}\}.$$

It follows from Corollary 4.7 that $\mathcal{P}(\widehat{\mathbb{W}}_T) \subset \mathbb{D}^{\infty,2}$. The following proposition shows that for $F \in \mathcal{P}(\widehat{\mathbb{W}}_T)$, the kernel functions can be expressed explicitly.

Proposition 5.9. *Let $F = \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_T \rangle$ with $\boldsymbol{\ell} \in T(\mathbb{R}^d)$. Then $F \in \mathbb{D}^{\infty,2}$ and the kernel functions $f_{\mathbf{i}_1 \dots \mathbf{i}_n}$ in the chaos expansion are given by*

$$f_{\mathbf{i}_1 \dots \mathbf{i}_n}(s_1, \dots, s_n) = \frac{1}{n!} \langle \boldsymbol{\ell}, \widehat{\mathcal{E}}_{s_1} \otimes \mathbf{i}_1 \otimes \widehat{\mathcal{E}}_{s_2-s_1} \otimes \mathbf{i}_2 \otimes \dots \otimes \mathbf{i}_n \otimes \widehat{\mathcal{E}}_{s_n-s_{n-1}} \rangle.$$

In particular, recalling that $|\mathbf{v}|_{\mathbf{i}}$ denotes the number of occurrences of the letter \mathbf{i} in $\mathbf{v} \in V$, we have

$$f_{\mathbf{i}_1 \dots \mathbf{i}_n}(s_1, \dots, s_n) \equiv 0, \quad n > N = \sup \left\{ \sum_{i=1}^d |\mathbf{v}|_{\mathbf{i}} : \mathbf{v} \in V, \boldsymbol{\ell}^{\mathbf{v}} \neq 0 \right\},$$

so that F has a finite chaos expansion of order N .

Proof. The result follows from (4.7) and the independence of $\widehat{\mathbb{W}}_{s_1}, \widehat{\mathbb{W}}_{s_1, s_2}, \dots, \widehat{\mathbb{W}}_{s_{n-1}, s_n}$. \square

5.3 The Ornstein–Uhlenbeck semigroup and its generator

In this section we establish explicit formulas for the action of the Ornstein–Uhlenbeck semigroup and its generator on the Brownian signature. The Ornstein–Uhlenbeck semigroup \mathbf{T}_ϑ (see, for instance, [Nualart \(2006, Section 1.4\)](#)) and its generator are usually introduced using the splitting of the Brownian motion

$$B^\vartheta := e^{-\vartheta}W + \sqrt{1 - e^{-2\vartheta}}W^\perp, \quad \vartheta \geq 0,$$

where W^\perp is an independent copy of W . The semigroup is then defined by

$$\mathbf{T}_\vartheta F(W) = \mathbb{E}[F(B^\vartheta) \mid \mathcal{F}_T^W], \quad \vartheta \geq 0,$$

and its generator is given by

$$\mathbf{L} = \left. \frac{d}{d\vartheta} \right|_{\vartheta=0} \mathbf{T}_\vartheta = - \sum_{i=1}^d \delta^i \mathbf{D}^i,$$

where the last identity follows from ([Nualart, 2006, Proposition 1.4.3](#)).

However, for later applications to integration-by-parts formulae, it may be convenient to split to all components of W , but only some of them. This leads to the following generalization of \mathbf{T}_ϑ .

Definition 5.10. Fix $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_d) \in \mathbb{R}_+^d$ and set

$$B_t^{\vartheta, \boldsymbol{\kappa}} := \left(e^{-\kappa_i \vartheta} W_t^i + \sqrt{1 - e^{-2\kappa_i \vartheta}} W_t^{\perp, i} \right)_{i=1, \dots, d}, \quad t \in [0, T], \quad \vartheta \geq 0.$$

The $\boldsymbol{\kappa}$ -Ornstein–Uhlenbeck semigroup and its generator are defined by

$$\mathbf{T}_\vartheta^{\boldsymbol{\kappa}} F(W) = \mathbb{E}[F(B^{\vartheta, \boldsymbol{\kappa}}) \mid \mathcal{F}_T^W], \quad \mathbf{L}^{\boldsymbol{\kappa}} = \left. \frac{d}{d\vartheta} \right|_{\vartheta=0} \mathbf{T}_\vartheta^{\boldsymbol{\kappa}}. \quad (5.8)$$

In particular, when $\boldsymbol{\kappa} = (1, \dots, 1)$, the definition coincides with the standard Ornstein–Uhlenbeck semigroup ([5.10](#)). When $\boldsymbol{\kappa} = e_i \in \mathbb{R}_+^d$ for some $i = 1, \dots, d$, this corresponds to splitting only the i -th component of W . We will denote the corresponding semigroup and generator by \mathbf{T}_ϑ^i and \mathbf{L}^i . Note that the operators \mathbf{T}_ϑ^i and \mathbf{L}^i can be viewed as the standard Ornstein–Uhlenbeck semigroup and generator associated with the one-dimensional Brownian motion W^i , conditional on the remaining components $W^1, \dots, W^{i-1}, W^{i+1}, \dots, W^d$. In particular, this implies that

$$\mathbf{L}^i = -\delta^i \mathbf{D}^i,$$

which leads to the linear decomposition of the generator:

$$\mathbf{L} = \sum_{i=1}^d \mathbf{L}^i.$$

Remark 5.11. Note, however, that the splitting of the Ornstein–Uhlenbeck semigroup itself is not linear:

$$\mathbf{T}_\vartheta \neq \sum_{i=1}^d \mathbf{T}_\vartheta^i.$$

For instance, one can easily verify that, for $F = W_T^1 W_T^2$,

$$\mathbf{T}_\vartheta(W_T^1 W_T^2) = e^{-2\vartheta} W_T^1 W_T^2 \neq 2e^{-\vartheta} W_T^1 W_T^2 = (\mathbf{T}_\vartheta^1 + \mathbf{T}_\vartheta^2)(W_T^1 W_T^2).$$

■

Our goal is to compute $\mathbf{T}_\vartheta^{\boldsymbol{\kappa}} F(W)$ and $\mathbf{L}^{\boldsymbol{\kappa}} F(W)$ in the special case $F(W) = \widehat{\mathbb{W}}_t$. We recall that in this case $\mathbf{T}_\vartheta^{\boldsymbol{\kappa}} \widehat{\mathbb{W}}_t$ and $\mathbf{L}^{\boldsymbol{\kappa}} \widehat{\mathbb{W}}_t$ are understood as elements of $T((\mathbb{R}^d))$, with the operators applied component-wise:

$$\mathbf{T}_\vartheta^{\boldsymbol{\kappa}} \widehat{\mathbb{W}}_t = \left(\mathbf{T}_\vartheta^{\boldsymbol{\kappa}} \widehat{\mathbb{W}}_t^{\mathbf{v}} \right)_{\mathbf{v} \in V}, \quad \mathbf{L}^{\boldsymbol{\kappa}} \widehat{\mathbb{W}}_t = \left(\mathbf{L}^{\boldsymbol{\kappa}} \widehat{\mathbb{W}}_t^{\mathbf{v}} \right)_{\mathbf{v} \in V}.$$

Remark 5.12. The domain of \mathbf{L} coincides with the space $\mathbb{D}^{2,2}$ of random variables that are twice Malliavin differentiable (Nualart, 2006, Section 1.4.2). Therefore, $\mathbf{L}\widehat{\mathbb{W}}_t$ is well-defined, since $\widehat{\mathbb{W}}_t \in \mathbb{D}^{\infty,2}$ by Corollary 4.7. We will show that

$$\mathbf{L}^\kappa = - \sum_{i=1}^d \kappa_i \delta^i \mathbf{D}^i,$$

so that $\text{Dom}(\mathbf{L}^\kappa) = \text{Dom}(\mathbf{L})$ and $\mathbf{L}^\kappa \widehat{\mathbb{W}}_t$ is well-defined as well. \blacksquare

We introduce two linear operators on $T((\mathbb{R}^{d+1}))$ that will be useful for computing $\mathbf{T}_\vartheta^\kappa \widehat{\mathbb{W}}_t$ and $\mathbf{L}^\kappa \widehat{\mathbb{W}}_t$.

Definition 5.13. Fix $\vartheta \geq 0$ and a letter $\mathbf{i} \in A$. We define the operators $\mathbf{J}_\mathbf{i}^\vartheta: T((\mathbb{R}^{d+1})) \rightarrow T((\mathbb{R}^{d+1}))$ and $\mathbf{\Lambda}_\mathbf{i}: T((\mathbb{R}^{d+1})) \rightarrow T((\mathbb{R}^{d+1}))$ by

$$\begin{aligned} \mathbf{J}_\mathbf{i}^\vartheta: \sum_{n \geq 0} \sum_{|\mathbf{v}|=n} \ell^{\mathbf{v}} \mathbf{v} &\mapsto \sum_{n \geq 0} \sum_{|\mathbf{v}|=n} e^{-|\mathbf{v}|\vartheta} \ell^{\mathbf{v}} \mathbf{v}, \\ \mathbf{\Lambda}_\mathbf{i}: \sum_{n \geq 0} \sum_{|\mathbf{v}|=n} \ell^{\mathbf{v}} \mathbf{v} &\mapsto \sum_{n \geq 0} \sum_{|\mathbf{v}|=n} |\mathbf{v}|\mathbf{i} \ell^{\mathbf{v}} \mathbf{v}, \end{aligned}$$

where $|\mathbf{v}|\mathbf{i}$ denotes the number of occurrences of the letter \mathbf{i} in the word \mathbf{v} .

We note that both operators are diagonal, and they satisfy $\mathbf{\Lambda}_\mathbf{i} = -\frac{d}{d\vartheta} \mathbf{J}_\mathbf{i}^\vartheta|_{\vartheta=0}$.

Theorem 5.14. The κ -Ornstein–Uhlenbeck semigroup operator of the Brownian signature is given by

$$\mathbf{T}_\vartheta^\kappa \widehat{\mathbb{W}}_t = \exp\left(\sum_{i=1}^d \frac{1 - e^{-\kappa_i \vartheta}}{2} \Psi_{\mathbf{ii}}^0\right) \left(\prod_{i=1}^d \mathbf{J}_\mathbf{i}^{\kappa_i \vartheta}\right) \widehat{\mathbb{W}}_t. \quad (5.9)$$

Moreover, we have $\widehat{\mathbb{W}}_t \in \text{Dom}(\mathbf{L}^\kappa)$ and

$$\mathbf{L}^\kappa \widehat{\mathbb{W}}_t = \sum_{i=1}^d \kappa_i (\Psi_{\mathbf{ii}}^0 - \mathbf{\Lambda}_\mathbf{i}) \widehat{\mathbb{W}}_t. \quad (5.10)$$

Their adjoint operators are

$$(\mathbf{T}_\vartheta^\kappa)^* = \left(\prod_{i=1}^d \mathbf{J}_\mathbf{i}^{\kappa_i \vartheta}\right) \exp\left(\frac{1 - e^{-2\kappa_i \vartheta}}{2} \Psi_{\mathbf{ii}}^0\right), \quad (\mathbf{L}^\kappa)^* = \sum_{i=1}^d \kappa_i (\Psi_{\mathbf{ii}}^0 - \mathbf{\Lambda}_\mathbf{i}). \quad (5.11)$$

Proof. Step 1. Semigroup $\mathbf{T}_\vartheta^\kappa$. We denote the time-augmented signature of $B^{\vartheta, \kappa}$ by $\widehat{\mathbb{B}}_t^{\vartheta, \kappa}$ and write the dynamics equation (3) for it:

$$d\widehat{\mathbb{B}}_t^{\vartheta, \kappa} = (\widehat{\mathbb{B}}_t^{\vartheta, \kappa} \otimes \mathbf{0}) dt + \sum_{i=1}^d (\widehat{\mathbb{B}}_t^{\vartheta, \kappa} \otimes \mathbf{i}) \circ (e^{-\kappa_i \vartheta} dW_t^i + \sqrt{1 - e^{-2\kappa_i \vartheta}} dW_t^{\perp, i}). \quad (5.12)$$

Writing the dynamics of $(\widehat{\mathbb{B}}_t^{\vartheta, \kappa} \otimes \mathbf{i})$ in a similar way, one can easily verify that

$$d\langle \widehat{\mathbb{B}}_t^{\vartheta, \kappa} \otimes \mathbf{i}, \sqrt{1 - e^{-2\kappa_i \vartheta}} W^{\perp, i} \rangle_t = (1 - e^{-2\kappa_i \vartheta}) (\widehat{\mathbb{B}}_t^{\vartheta, \kappa} \otimes \mathbf{ii}) dt, \quad i = 1, \dots, d.$$

Using the relationship between Itô and Stratonovich integrals for $(\widehat{\mathbb{B}}_t^{\vartheta, \kappa} \otimes \mathbf{i}) \circ \sqrt{1 - e^{-2\kappa_i \vartheta}} dW_t^{\perp, i}$ in (5.3) and rearranging the terms, we obtain

$$d\widehat{\mathbb{B}}_t^{\vartheta, \kappa} = \widehat{\mathbb{B}}_t^{\vartheta, \kappa} \otimes \left(\mathbf{0} + \sum_{i=1}^d \frac{1 - e^{-2\kappa_i \vartheta}}{2} \mathbf{ii}\right) dt + \sum_{i=1}^d (\widehat{\mathbb{B}}_t^{\vartheta, \kappa} \otimes \mathbf{i}) \circ e^{-\kappa_i \vartheta} dW_t^i + \sum_{i=1}^d (\widehat{\mathbb{B}}_t^{\vartheta, \kappa} \otimes \mathbf{i}) \sqrt{1 - e^{-2\kappa_i \vartheta}} dW_t^{\perp, i}.$$

Taking the conditional expectation $\mathbb{E}[\cdot | \mathcal{F}_T^W]$ yields a linear SDE for $\mathbf{T}_\vartheta^\kappa \widehat{\mathbb{W}}_t = \mathbb{E}[\widehat{\mathbb{B}}_t^{\vartheta, \kappa} | \mathcal{F}_T^W]$:

$$d\mathbf{T}_\vartheta^\kappa \widehat{\mathbb{W}}_t = \mathbf{T}_\vartheta^\kappa \widehat{\mathbb{W}}_t \otimes \left(\mathbf{0} + \sum_{i=1}^d \frac{1 - e^{-2\kappa_i \vartheta}}{2} \mathbf{ii}\right) dt + \sum_{i=1}^d \mathbf{T}_\vartheta^\kappa \widehat{\mathbb{W}}_t \otimes (e^{-\kappa_i \vartheta} \mathbf{i}) \circ dW_t^i.$$

Applying Picard iteration, one obtains a solution linear in $\widehat{\mathbb{W}}_t$:

$$\mathbf{T}_\vartheta^\kappa \widehat{\mathbb{W}}_t = \sum_{\mathbf{v}=\mathbf{i}_1 \dots \mathbf{i}_n} (\mathbf{u}^{\vartheta, \kappa}(\mathbf{i}_1) \otimes \dots \otimes \mathbf{u}^{\vartheta, \kappa}(\mathbf{i}_n)) \widehat{\mathbb{W}}_t^{\mathbf{i}_1 \dots \mathbf{i}_n},$$

where

$$\mathbf{u}^{\vartheta, \kappa}(\mathbf{0}) = \mathbf{0} + \sum_{i=1}^d \frac{1 - e^{-\kappa_i \vartheta}}{2} \mathbf{i}_i, \quad \mathbf{u}^{\vartheta, \kappa}(\mathbf{i}) = e^{-\kappa_i \vartheta} \mathbf{i}, \quad i = 1, \dots, d. \quad (5.13)$$

Hence, we can define an operator $\mathbf{T}_\vartheta^\kappa: T((\mathbb{R}^{d+1})) \rightarrow T((\mathbb{R}^{d+1}))$ by

$$\mathbf{T}_\vartheta^\kappa(\mathbf{i}_1 \dots \mathbf{i}_n) = \bigotimes_{j=1}^n \mathbf{u}^{\vartheta, \kappa}(\mathbf{i}_j),$$

so that $\mathbf{T}_\vartheta^\kappa(\widehat{\mathbb{W}}_t)$ coincides with the action of the Ornstein–Uhlenbeck semigroup.

Since for all $\mathbf{j} \in \{\mathbf{0}, \mathbf{1}, \dots, \mathbf{d}\}$,

$$\Psi_{\mathbf{ii}}^{\mathbf{0}}(\mathbf{j}) = \mathbb{1}_{\{j=0\}} \mathbf{i}_i, \quad \mathbf{J}_i^{\kappa_i \vartheta}(\mathbf{j}) = \begin{cases} e^{-\kappa_i \vartheta} \mathbf{j}, & j = i, \\ \mathbf{j}, & j \neq i, \end{cases}$$

and

$$\exp\left(\sum_{i=1}^d \frac{1 - e^{-\kappa_i \vartheta}}{2} \Psi_{\mathbf{ii}}^{\mathbf{0}}\right)(\mathbf{j}) = \begin{cases} \mathbf{0} + \sum_{i=1}^d \frac{1 - e^{-\kappa_i \vartheta}}{2} \mathbf{i}_i, & j = 0, \\ \mathbf{j}, & j \neq 0, \end{cases}$$

equation (5.3) reads

$$\mathbf{u}^{\vartheta, \kappa}(\mathbf{j}) = \exp\left(\sum_{i=1}^d \frac{1 - e^{-\kappa_i \vartheta}}{2} \Psi_{\mathbf{ii}}^{\mathbf{0}}\right) \left(\prod_{i=1}^d \mathbf{J}_i^{\kappa_i \vartheta}\right)(\mathbf{j}) = \mathbf{T}_\vartheta^\kappa(\mathbf{j}), \quad \mathbf{j} \in \{\mathbf{0}, \mathbf{1}, \dots, \mathbf{d}\}.$$

Moreover, each $\Psi_{\mathbf{ii}}^{\mathbf{0}}$ satisfies the Leibniz rule,

$$\Psi_{\mathbf{ii}}^{\mathbf{0}}(\mathbf{u} \otimes \mathbf{v}) = \Psi_{\mathbf{ii}}^{\mathbf{0}}(\mathbf{u}) \otimes \mathbf{v} + \mathbf{u} \otimes \Psi_{\mathbf{ii}}^{\mathbf{0}}(\mathbf{v}),$$

and the same holds for $\sum_{i=1}^d \frac{1 - e^{-\kappa_i \vartheta}}{2} \Psi_{\mathbf{ii}}^{\mathbf{0}}$, so that its exponential is a tensor algebra homomorphism:

$$\exp\left(\sum_{i=1}^d \frac{1 - e^{-\kappa_i \vartheta}}{2} \Psi_{\mathbf{ii}}^{\mathbf{0}}\right)(\mathbf{u} \otimes \mathbf{v}) = \exp\left(\sum_{i=1}^d \frac{1 - e^{-\kappa_i \vartheta}}{2} \Psi_{\mathbf{ii}}^{\mathbf{0}}\right)(\mathbf{u}) \otimes \exp\left(\sum_{i=1}^d \frac{1 - e^{-\kappa_i \vartheta}}{2} \Psi_{\mathbf{ii}}^{\mathbf{0}}\right)(\mathbf{v}).$$

Moreover, for all $i = 1, \dots, d$, $\mathbf{J}_i^{\kappa_i \vartheta}$ is a homomorphism as well. Thus, the identity

$$\mathbf{T}_\vartheta^\kappa = \exp\left(\sum_{i=1}^d \frac{1 - e^{-\kappa_i \vartheta}}{2} \Psi_{\mathbf{ii}}^{\mathbf{0}}\right) \left(\prod_{i=1}^d \mathbf{J}_i^{\kappa_i \vartheta}\right)$$

is proved, yielding (5.14).

Step 2. Generator \mathbf{L}^κ . By definition,

$$\mathbf{L}^\kappa = \left[\frac{d}{d\vartheta} \left(\exp\left(\sum_{i=1}^d \frac{1 - e^{-2\kappa_i \vartheta}}{2} \Psi_{\mathbf{ii}}^{\mathbf{0}}\right) \right) \prod_{i=1}^d \mathbf{J}_i^{\kappa_i \vartheta} + \sum_{j=1}^d \exp\left(\sum_{i=1}^d \frac{1 - e^{-2\kappa_i \vartheta}}{2} \Psi_{\mathbf{ii}}^{\mathbf{0}}\right) \left(\frac{d}{d\vartheta} \mathbf{J}_j^{\kappa_j \vartheta} \right) \prod_{k \neq j} \mathbf{J}_k^{\kappa_k \vartheta} \right]_{\vartheta=0}.$$

A direct computation gives

$$\frac{d}{d\vartheta} \exp\left(\sum_{i=1}^d \frac{1 - e^{-2\kappa_i \vartheta}}{2} \Psi_{\mathbf{ii}}^{\mathbf{0}}\right) \Big|_{\vartheta=0} = \sum_{i=1}^d \kappa_i \Psi_{\mathbf{ii}}^{\mathbf{0}}, \quad \mathbf{J}_i^{\kappa_i \vartheta} \Big|_{\vartheta=0} = \exp\left(\sum_{i=1}^d \frac{1 - e^{-2\kappa_i \vartheta}}{2} \Psi_{\mathbf{ii}}^{\mathbf{0}}\right) \Big|_{\vartheta=0} = \text{Id}.$$

Recall that $\frac{d}{d\vartheta} \mathbf{J}_j^{\kappa_j \vartheta} \Big|_{\vartheta=0} = -\kappa_j \mathbf{\Lambda}_j$ so (5.14) follows.

Finally, the adjoint formulas (5.14) follow from Proposition 5.5 and the fact that $\mathbf{J}_i^{\kappa_i \vartheta}$ and $\mathbf{\Lambda}_i$ are diagonal operators. \square

Corollary 5.15. *The standard Ornstein–Uhlenbeck semigroup \mathbf{T}_ϑ and generator \mathbf{L} are given by*

$$\mathbf{T}_\vartheta \widehat{\mathbb{W}}_t = \exp\left(\frac{1-e^{-2\vartheta}}{2} \sum_{i=1}^d \Psi_{\mathbf{ii}}^{\mathbf{0}}\right) \left(\prod_{i=1}^d \mathbf{J}_i^\vartheta\right) \widehat{\mathbb{W}}_t, \quad \mathbf{L} \widehat{\mathbb{W}}_t = \sum_{i=1}^d (\Psi_{\mathbf{ii}}^{\mathbf{0}} - \mathbf{\Lambda}_i) \widehat{\mathbb{W}}_t.$$

The adjoint operators are given by

$$\mathbf{T}_\vartheta^* = \left(\prod_{i=1}^d \mathbf{J}_i^\vartheta\right) \exp\left(\frac{1-e^{-2\vartheta}}{2} \sum_{i=1}^d \Psi_{\mathbf{ii}}^{\mathbf{0}}\right) \quad \mathbf{L}^* = \sum_{i=1}^d (\Psi_{\mathbf{ii}}^{\mathbf{0}} - \mathbf{\Lambda}_i). \quad (5.14)$$

Remark 5.16. The semigroup \mathbf{T}_ϑ can equivalently be defined in terms of chaos expansions; see (Nualart, 2006, Definition 1.4.1). Namely, if $F \in L^2(\Omega, \mathcal{F}^W, \mathbb{P})$, then $F = \sum_{n \geq 0} J_n F$, where $J_n F$ denotes the projection of F onto the n -th Wiener chaos. The semigroup then acts as

$$\mathbf{T}_\vartheta F = \sum_{n \geq 0} e^{-n\vartheta} J_n F.$$

Using the recursive definition of the Wiener chaos and the Itô signature, a simple induction argument based on Fubini’s theorem shows that $\widehat{\mathbb{W}}_t^{\text{Itô}, \mathbf{v}}$ belongs to the Wiener chaos of order $\sum_{i=1}^d |\mathbf{v}|_i$.

$$\mathbf{T}_\vartheta \langle \ell, \widehat{\mathbb{W}}_t^{\text{Itô}} \rangle = \left\langle \prod_{i=1}^d \mathbf{J}_i^\vartheta \ell, \widehat{\mathbb{W}}_t^{\text{Itô}} \right\rangle.$$

Comparing this with (5.15), we observe that the operator $\exp\left(\frac{1-e^{-2\vartheta}}{2} \sum_{i=1}^d \Psi_{\mathbf{ii}}^{\mathbf{0}}\right)$ is related to the conversion of the Stratonovich signature into the Itô signature. Indeed, by writing the dynamics of $\widehat{\mathbb{W}}$ in terms of Itô integrals and proceeding as in the proof of Theorem 5.14, one can show that

$$\langle \ell, \widehat{\mathbb{W}}_t^{\text{Itô}} \rangle = \left\langle \exp\left(\frac{1}{2} \sum_{i=1}^d \Psi_{\mathbf{ii}}^{\mathbf{0}}\right) \ell, \widehat{\mathbb{W}}_t \right\rangle,$$

which provides a concise operator representation of the Itô–Stratonovich relationship for signatures similar to Ben Arous (1989, Proposition 1).

Corollary 5.17. *The “coordinate” Ornstein–Uhlenbeck semigroup \mathbf{T}_ϑ^i and generator \mathbf{L}^i are given by*

$$\mathbf{T}_\vartheta^i \widehat{\mathbb{W}}_t = \exp\left(\frac{1-e^{-2\vartheta}}{2} \Psi_{\mathbf{ii}}^{\mathbf{0}}\right) \mathbf{J}_i^\vartheta \widehat{\mathbb{W}}_t, \quad \mathbf{L}^i \widehat{\mathbb{W}}_t = (\Psi_{\mathbf{ii}}^{\mathbf{0}} - \mathbf{\Lambda}_i) \widehat{\mathbb{W}}_t.$$

The adjoint operators are given by

$$(\mathbf{T}_\vartheta^i)^* = \mathbf{J}_i^\vartheta \exp\left(\frac{1-e^{-2\vartheta}}{2} \Psi_{\mathbf{ii}}^{\mathbf{0}}\right) \quad (\mathbf{L}^i)^* = \Psi_{\mathbf{ii}}^{\mathbf{0}} - \mathbf{\Lambda}_i.$$

An alternative way to compute the generators \mathbf{L}^i consists in combining the formula $\mathbf{L}^i = -\delta^i \mathbf{D}^i$ and Theorem 5.6. We start by writing

$$\mathbf{L}^i \widehat{\mathbb{W}}_t = -\delta^i (\mathbf{D}^i \widehat{\mathbb{W}}_t) = -\delta^i \left(\widehat{\mathbb{W}} \cdot \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{\cdot, t} \right), \quad (5.15)$$

To proceed, one needs to express the Skorokhod integral in terms of the Stratonovich integral, as established in (Nualart, 2006, Theorem 3.1.1):

$$\delta^i(X) = \int_0^t X_u \circ dW_u^i - \frac{1}{2} \int_0^t ((\mathbf{D}^{+,i} X)_u + (\mathbf{D}^{-,i} X)_u) du, \quad (5.16)$$

where $\mathbf{D}^{+,i}X$ and $\mathbf{D}^{-,i}X$ denote the i -th components of the \mathbf{D}^+X and \mathbf{D}^-X defined in (3).

For the processes X appearing in the Skorokhod integral (5.3), $\mathbf{D}^{+,i}X$ and $\mathbf{D}^{-,i}X$ can be computed explicitly.

Lemma 5.18. *Let $\mathbf{v} \in V$ and consider the $T_2(\mathbb{R}^{d+1})$ -valued process*

$$\mathbb{X}_s = \widehat{\mathbb{W}}_s \otimes \mathbf{v} \otimes \widehat{\mathbb{W}}_{s,t}, \quad s \leq t.$$

Then $\mathbb{X} \in \mathbb{L}_2^{1,2}$ and, for $s \leq t$,

$$(\mathbf{D}^{+,i}\mathbb{X})_s = \widehat{\mathbb{W}}_s \otimes \mathbf{i}\mathbf{v} \otimes \widehat{\mathbb{W}}_{s,t}, \quad (\mathbf{D}^{-,i}\mathbb{X})_s = \widehat{\mathbb{W}}_s \otimes \mathbf{v}\mathbf{i} \otimes \widehat{\mathbb{W}}_{s,t}.$$

Proof. For $0 < s < u < t$, we have

$$\mathbf{D}_s^i(\widehat{\mathbb{W}}_u \otimes \mathbf{v} \otimes \widehat{\mathbb{W}}_{u,t}) = \widehat{\mathbb{W}}_s \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{s,u} \otimes \mathbf{v} \otimes \widehat{\mathbb{W}}_{u,t}.$$

We claim that for all $\mathbf{w} \in V$,

$$\mathbb{E} \left[\left(\mathbf{D}_s^i(\widehat{\mathbb{W}}_u \otimes \mathbf{v} \otimes \widehat{\mathbb{W}}_{u,t})^{\mathbf{w}} - (\widehat{\mathbb{W}}_s \otimes \mathbf{i}\mathbf{v} \otimes \widehat{\mathbb{W}}_{s,t})^{\mathbf{w}} \right)^2 \right] \rightarrow 0,$$

as $|u - s| \rightarrow 0$. Indeed, we have

$$\mathbf{D}_s^i(\widehat{\mathbb{W}}_u \otimes \mathbf{v} \otimes \widehat{\mathbb{W}}_{u,t}) - \widehat{\mathbb{W}}_s \otimes \mathbf{i}\mathbf{v} \otimes \widehat{\mathbb{W}}_{s,t} = \widehat{\mathbb{W}}_s \otimes \mathbf{i} \otimes (\widehat{\mathbb{W}}_{s,u} \otimes \mathbf{v} - \mathbf{v} \otimes \widehat{\mathbb{W}}_{s,u}) \otimes \widehat{\mathbb{W}}_{u,t},$$

so that

$$\begin{aligned} & \mathbb{E} \left[\left(\mathbf{D}_s^i(\widehat{\mathbb{W}}_u \otimes \mathbf{v} \otimes \widehat{\mathbb{W}}_{u,t})^{\mathbf{w}} - (\widehat{\mathbb{W}}_s \otimes \mathbf{i}\mathbf{v} \otimes \widehat{\mathbb{W}}_{s,t})^{\mathbf{w}} \right)^2 \right] \\ &= \mathbb{E} \left[\left(\sum_{\substack{\mathbf{w}_1, \mathbf{w}_2', \mathbf{w}_2'', \mathbf{w}_3 \\ \mathbf{w}_1 \mathbf{i} \mathbf{w}_2' \mathbf{v} \mathbf{w}_3 = \mathbf{w}_1 \mathbf{i} \mathbf{v} \mathbf{w}_2'' \mathbf{w}_3 = \mathbf{w}}} \widehat{\mathbb{W}}_s^{\mathbf{w}_1} (\widehat{\mathbb{W}}_{s,u}^{\mathbf{w}_2'} - \widehat{\mathbb{W}}_{s,u}^{\mathbf{w}_2''}) \widehat{\mathbb{W}}_{u,t}^{\mathbf{w}_3} \right)^2 \right] \\ &\leq \mathbb{E} \left[\left(\sum_{\substack{\mathbf{w}_1, \mathbf{w}_2', \mathbf{w}_2'', \mathbf{w}_3 \\ |\mathbf{w}_1|, |\mathbf{w}_3| < |\mathbf{w}| \\ 0 < |\mathbf{w}_2'|, |\mathbf{w}_2''| < |\mathbf{w}|}} |\widehat{\mathbb{W}}_s^{\mathbf{w}_1}| |\widehat{\mathbb{W}}_{s,u}^{\mathbf{w}_2'} - \widehat{\mathbb{W}}_{s,u}^{\mathbf{w}_2''}| |\widehat{\mathbb{W}}_{u,t}^{\mathbf{w}_3}| \right)^2 \right] \\ &\leq 4\mathbb{E} \left[\left(\sum_{|\mathbf{w}_1| < |\mathbf{w}|} |\widehat{\mathbb{W}}_s^{\mathbf{w}_1}| \right)^2 \right] \mathbb{E} \left[\left(\sum_{|\mathbf{w}_3| < |\mathbf{w}|} |\widehat{\mathbb{W}}_{u,t}^{\mathbf{w}_3}| \right)^2 \right] \mathbb{E} \left[\left(\sum_{0 < |\mathbf{w}_2| < |\mathbf{w}|} |\widehat{\mathbb{W}}_{s,u}^{\mathbf{w}_2}| \right)^2 \right] \\ &\leq 4\mathbb{E} \left[\left(\sum_{|\mathbf{w}_1| < |\mathbf{w}|} |\widehat{\mathbb{W}}_T^{\mathbf{w}_1}| \right)^2 \right]^2 \mathbb{E} \left[\left(\sum_{0 < |\mathbf{w}_2| < |\mathbf{w}|} |\widehat{\mathbb{W}}_{s,u}^{\mathbf{w}_2}| \right)^2 \right], \end{aligned}$$

where, in the last inequality, we used the diffusive scaling of the signature. The last expression depends only on $|u - s|$ and converges to 0 as $|s - u| \rightarrow 0$. Hence, the limit in (3) vanishes. The proof for $0 < u < s < t$ and $\mathbf{D}^{-,i}$ is analogous. \square

Applying Lemma 5.18 to $\widehat{\mathbb{W}} \cdot \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{\cdot,t}$ yields

$$\left(\mathbf{D}^{+,i}(\widehat{\mathbb{W}} \cdot \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{\cdot,t}) \right)_u + \left(\mathbf{D}^{-,i}(\widehat{\mathbb{W}} \cdot \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{\cdot,t}) \right)_u = 2(\widehat{\mathbb{W}}_u \otimes \mathbf{i}\mathbf{i} \otimes \widehat{\mathbb{W}}_{u,t}).$$

The Skorokhod integral (5.3) can therefore be expressed in terms of the Stratonovich integral using (5.3):

$$\begin{aligned}\delta^i(\widehat{\mathbb{W}} \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{\cdot,t}) &= \int_0^t (\widehat{\mathbb{W}}_u \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{u,t}) \circ dW_u^i - \int_0^t (\widehat{\mathbb{W}}_u \otimes \mathbf{ii} \otimes \widehat{\mathbb{W}}_{u,t}) du \\ &= \Psi_{\mathbf{i}}^i(\widehat{\mathbb{W}}_t) - \Psi_{\mathbf{ii}}^0(\widehat{\mathbb{W}}_t) \\ &= \Lambda_{\mathbf{i}}(\widehat{\mathbb{W}}_t) - \Psi_{\mathbf{ii}}^0(\widehat{\mathbb{W}}_t),\end{aligned}$$

where we used Theorem 5.6 and observed that $\Psi_{\mathbf{i}}^i = \Lambda_{\mathbf{i}}$. This recovers the formula obtained in Corollary 5.17.

6 Integration by parts

For conciseness of notation, we will denote by $\langle \cdot, \cdot \rangle_{L^2}$ the scalar product in the space $L^2(\Omega \times [0, T], \mathbb{R}^d)$. That is, for $h = (h^1, \dots, h^d)$ and $g = (g^1, \dots, g^d)$ in $L^2(\Omega \times [0, T], \mathbb{R}^d)$, we set

$$\langle h, g \rangle_{L^2} = \int_0^T \langle h_t, g_t \rangle_{\mathbb{R}^d} dt.$$

The following proposition (a multidimensional version of Nualart (2006, Proposition 6.2.1)) allows for computation of Malliavin weights in the integration-by-parts formula (2.2).

Proposition 6.1. *Let G and F be random variables such that $G \in \mathbb{D}^{1,2}(\mathbb{R})$, and let $h \in L^2(\Omega \times [0, T], \mathbb{R}^d)$ be such that $\langle \mathbf{D}G, h \rangle_{L^2} \neq 0$ a.s. and $\frac{Fh}{\langle \mathbf{D}G, h \rangle_{L^2}} \in \text{Dom}(\delta)$. Then, for any $f \in C_b^1(\mathbb{R}, \mathbb{R})$ we have*

$$\mathbb{E}[f'(G)F] = \mathbb{E}\left[f(G)\delta\left(\frac{Fh}{\langle \mathbf{D}G, h \rangle_{L^2}}\right)\right].$$

Moreover, if $\frac{F}{\langle \mathbf{D}G, h \rangle_{L^2}} \in \mathbb{D}^{1,2}$ and $h \in \text{Dom}(\delta)$, then

$$\mathbb{E}[f'(G)F] = \mathbb{E}\left[f(G)\left(\frac{F\delta(h)}{\langle \mathbf{D}G, h \rangle_{L^2}} - \frac{\langle \mathbf{D}F, h \rangle_{L^2}}{\langle \mathbf{D}G, h \rangle_{L^2}} + \frac{F\langle \mathbf{D}\langle \mathbf{D}G, h \rangle_{L^2}, h \rangle_{L^2}}{\langle \mathbf{D}G, h \rangle_{L^2}^2}\right)\right]. \quad (6.1)$$

Proof. By the chain rule, $\langle \mathbf{D}f(G), h \rangle_{L^2} = f'(G)\langle \mathbf{D}G, h \rangle_{L^2}$, so that

$$f'(G) = \frac{\langle \mathbf{D}f(G), h \rangle_{L^2}}{\langle \mathbf{D}G, h \rangle_{L^2}},$$

and

$$\mathbb{E}[f'(G)F] = \mathbb{E}\left[\frac{\langle \mathbf{D}f(G), h \rangle_{L^2}}{\langle \mathbf{D}G, h \rangle_{L^2}}F\right] = \mathbb{E}\left[f(G)\delta\left(\frac{Fh}{\langle \mathbf{D}G, h \rangle_{L^2}}\right)\right],$$

where we used the definition of the Skorokhod integral (5.2). The formula (5.2) yields

$$\mathbb{E}[f'(G)F] = \mathbb{E}\left[f(G)\left(\frac{F\delta(h)}{\langle \mathbf{D}G, h \rangle_{L^2}} - \left\langle \mathbf{D}\left(\frac{F}{\langle \mathbf{D}G, h \rangle_{L^2}}\right), h \right\rangle_{L^2}\right)\right],$$

and (6.1) follows from a direct computation of the Malliavin derivative inside the integral. \square

6.1 Linearization of the L^2 scalar product

Typically, computing $\langle \mathbf{D}G, h \rangle_{L^2}$ and $\langle \mathbf{D}\langle \mathbf{D}G, h \rangle_{L^2}, h \rangle_{L^2}$, which appear in (6.1), is non-trivial, and their numerical evaluation may present a significant challenge. However, this computation is significantly simplified if the process h is chosen in such a way that it allows one to linearize $\langle \mathbf{D}G, h \rangle_{L^2}$, i.e., to express it as a linear functional of $\widehat{\mathbb{W}}_T$. Assuming additionally that both G and F are linear or rational (i.e., ratios of two linear

functionals) functions of $\widehat{\mathbb{W}}_T$ makes it possible to rewrite the weight given by (6.1) as a rational function of $\widehat{\mathbb{W}}_T$ as well.

We start by considering two random variables $F, F' \in \mathcal{P}(\widehat{\mathbb{W}}_T)$ with corresponding coefficients $\ell, \ell' \in T(\mathbb{R}^{d+1})$, and we introduce an operation $\diamond: T(\mathbb{R}^d) \times T(\mathbb{R}^d) \rightarrow T(\mathbb{R}^d)$ that allows us to linearize the scalar product $\langle \mathbf{D}F, \mathbf{D}F' \rangle_{L^2}$, i.e., to express it as

$$\langle \mathbf{D}F, \mathbf{D}F' \rangle_{L^2} = \left\langle \mathbf{D}\langle \ell, \widehat{\mathbb{W}}_T \rangle, \mathbf{D}\langle \ell', \widehat{\mathbb{W}}_T \rangle \right\rangle_{L^2} = \left\langle \ell \diamond \ell', \widehat{\mathbb{W}}_T \right\rangle.$$

Definition 6.2. Fix three words $\mathbf{u}_1, \mathbf{u}_2, \mathbf{w} \in V$. For two words \mathbf{v}, \mathbf{v}' , define $\diamond_{\mathbf{w}}^{\mathbf{u}_1, \mathbf{u}_2}: T(\mathbb{R}^d) \times T(\mathbb{R}^d) \rightarrow T(\mathbb{R}^d)$ by

$$\mathbf{v} \diamond_{\mathbf{w}}^{\mathbf{u}_1, \mathbf{u}_2} \mathbf{v}' = \sum_{\substack{\mathbf{v}_1, \mathbf{v}_2: \\ \mathbf{v}_1 \mathbf{u}_1 \mathbf{v}_2 = \mathbf{v}}} \sum_{\substack{\mathbf{v}'_1, \mathbf{v}'_2: \\ \mathbf{v}'_1 \mathbf{u}_2 \mathbf{v}'_2 = \mathbf{v}'}} (\mathbf{v}_1 \sqcup \mathbf{v}'_1) \otimes \mathbf{w} \otimes (\mathbf{v}_2 \sqcup \mathbf{v}'_2). \quad (6.2)$$

For two tensor sequences $\ell, \ell' \in T(\mathbb{R}^d)$, the operator is extended by bilinearity. We also define $\diamond := \sum_{i=1}^d \diamond_{\mathbf{0}}^{\mathbf{i}, \mathbf{i}}$.

The proposition below clarifies the link between this algebraic operation and the integration of the product of linear functionals of two ‘‘pierced’’ signatures $\mathbb{X}_t \otimes \mathbf{u} \otimes \mathbb{X}_{t,T}$.

Proposition 6.3. Let X be an \mathbb{R}^d -valued continuous semimartingale. Then, for all words $\mathbf{u}_1, \mathbf{u}_2 \in V$, letter $\mathbf{k} \in A$, and $\ell, \ell' \in T(\mathbb{R}^d)$,

$$\int_0^T \langle \ell, \mathbb{X}_t \otimes \mathbf{u}_1 \otimes \mathbb{X}_{t,T} \rangle \langle \ell', \mathbb{X}_t \otimes \mathbf{u}_2 \otimes \mathbb{X}_{t,T} \rangle \circ dX_t^{\mathbf{k}} = \langle \ell \diamond_{\mathbf{k}}^{\mathbf{u}_1, \mathbf{u}_2} \ell', \mathbb{X}_T \rangle. \quad (6.3)$$

Proof. Since the integral in (6.3) is a bilinear functional of ℓ and ℓ' , it suffices to consider the case $\ell = \mathbf{v} \in V$ and $\ell' = \mathbf{v}' \in V$. Then the integrand expands as

$$\begin{aligned} & \int_0^T \langle \mathbf{v}, \mathbb{X}_t \otimes \mathbf{u}_1 \otimes \mathbb{X}_{t,T} \rangle \langle \mathbf{v}', \mathbb{X}_t \otimes \mathbf{u}_2 \otimes \mathbb{X}_{t,T} \rangle \circ dX_t^{\mathbf{k}} \\ &= \sum_{\substack{\mathbf{v}_1, \mathbf{v}_2: \\ \mathbf{v}_1 \mathbf{u}_1 \mathbf{v}_2 = \mathbf{v}}} \sum_{\substack{\mathbf{v}'_1, \mathbf{v}'_2: \\ \mathbf{v}'_1 \mathbf{u}_2 \mathbf{v}'_2 = \mathbf{v}'}} \int_0^T \mathbb{X}_t^{\mathbf{v}_1} \mathbb{X}_{t,T}^{\mathbf{v}'_1} \mathbb{X}_{t,T}^{\mathbf{v}_2} \mathbb{X}_{t,T}^{\mathbf{v}'_2} \circ dX_t^{\mathbf{k}} \\ &= \sum_{\substack{\mathbf{v}_1, \mathbf{v}_2: \\ \mathbf{v}_1 \mathbf{u}_1 \mathbf{v}_2 = \mathbf{v}}} \sum_{\substack{\mathbf{v}'_1, \mathbf{v}'_2: \\ \mathbf{v}'_1 \mathbf{u}_2 \mathbf{v}'_2 = \mathbf{v}'}} \int_0^T \langle \mathbf{v}_1 \sqcup \mathbf{v}'_1, \mathbb{X}_t \rangle \langle \mathbf{v}_2 \sqcup \mathbf{v}'_2, \mathbb{X}_{t,T} \rangle \circ dX_t^{\mathbf{k}} \\ &= \sum_{\substack{\mathbf{v}_1, \mathbf{v}_2: \\ \mathbf{v}_1 \mathbf{u}_1 \mathbf{v}_2 = \mathbf{v}}} \sum_{\substack{\mathbf{v}'_1, \mathbf{v}'_2: \\ \mathbf{v}'_1 \mathbf{u}_2 \mathbf{v}'_2 = \mathbf{v}'}} \langle (\mathbf{v}_1 \sqcup \mathbf{v}'_1) \mathbf{k} (\mathbf{v}_2 \sqcup \mathbf{v}'_2), \mathbb{X}_T \rangle \\ &= \langle \mathbf{v} \diamond_{\mathbf{k}}^{\mathbf{u}_1, \mathbf{u}_2} \mathbf{v}', \mathbb{X}_T \rangle, \end{aligned}$$

where Lemma 5.1 is used in the third equality. This proves (6.3). \square

Corollary 6.4. Let $F, F' \in \mathcal{P}(\widehat{\mathbb{W}}_T)$ with the coefficients $\ell, \ell' \in T(\mathbb{R}^{d+1})$. Then

$$\langle \mathbf{D}F, \mathbf{D}F' \rangle_{L^2} = \left\langle \ell \diamond \ell', \widehat{\mathbb{W}}_T \right\rangle. \quad (6.4)$$

Proof. We have

$$\begin{aligned}
\langle \mathbf{D}F, \mathbf{D}F' \rangle_{L^2} &= \sum_{i=1}^d \int_0^T \langle \ell, \mathbf{D}_t^i \widehat{\mathbb{W}}_T \rangle \langle \ell', \mathbf{D}_t^i \widehat{\mathbb{W}}_T \rangle dt \\
&= \sum_{i=1}^d \int_0^T \langle \ell, \widehat{\mathbb{W}}_t \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{t,T} \rangle \langle \ell', \widehat{\mathbb{W}}_t \otimes \mathbf{i} \otimes \widehat{\mathbb{W}}_{t,T} \rangle dt \\
&= \sum_{i=1}^d \langle \ell \diamond_{\mathbf{0}}^{\mathbf{i}, \mathbf{i}} \ell', \widehat{\mathbb{W}}_T \rangle = \langle \ell \diamond \ell', \widehat{\mathbb{W}}_T \rangle,
\end{aligned}$$

where we used Proposition 6.3 and the definition of \diamond . \square

Corollary 6.5. Fix $\mathbf{i}, \mathbf{j}, \mathbf{k} \in A$ and $\ell \in T(\mathbb{R}^{d+1})$. Then $\ell \diamond_{\mathbf{k}}^{\mathbf{i}, \mathbf{j}} \mathbf{j} = \Psi_{\mathbf{k}}^{\mathbf{i}}(\ell)$.

Proof. It is sufficient to observe that for all continuous semimartingales X , we have

$$\langle \ell \diamond_{\mathbf{k}}^{\mathbf{i}, \mathbf{j}} \mathbf{j}, \mathbb{X}_T \rangle = \int_0^T \langle \ell, \mathbb{X}_t \otimes \mathbf{i} \otimes \mathbb{X}_{t,T} \rangle \underbrace{\langle \mathbf{j}, \mathbb{X}_t \otimes \mathbf{j} \otimes \mathbb{X}_{t,T} \rangle}_{=1} \circ dX_t^k = \langle \Psi_{\mathbf{k}}^{\mathbf{i}}(\ell), \mathbb{X}_T \rangle,$$

where the last equality follows from Theorem 5.3. \square

Although it enjoys the desired linearization property (6.4), the direct computation of \diamond via its definition (6.2) is rather cumbersome and numerically intricate. The following proposition shows that \diamond can be expressed in terms of the switching operators $\Psi_{\mathbf{0}}^{\mathbf{i}, \mathbf{i}}$, enabling a significantly more efficient numerical implementation.

Proposition 6.6. If $\mathbf{i}, \mathbf{k} \in A$ and $\ell, \ell' \in T(\mathbb{R}^{d+1})$, then the diamond product $\diamond_{\mathbf{k}}^{\mathbf{i}, \mathbf{i}}$ can be written in carré du champ form:

$$\ell \diamond_{\mathbf{k}}^{\mathbf{i}, \mathbf{i}} \ell' = \frac{1}{2} (\Psi_{\mathbf{k}}^{\mathbf{i}, \mathbf{i}}(\ell \sqcup \ell') - \Psi_{\mathbf{k}}^{\mathbf{i}, \mathbf{i}}(\ell) \sqcup \ell' - \ell \sqcup \Psi_{\mathbf{k}}^{\mathbf{i}, \mathbf{i}}(\ell')). \quad (6.5)$$

In particular, we have

$$\ell \diamond \ell' = \frac{1}{2} \sum_{i=1}^d (\Psi_{\mathbf{0}}^{\mathbf{i}, \mathbf{i}}(\ell \sqcup \ell') - \Psi_{\mathbf{0}}^{\mathbf{i}, \mathbf{i}}(\ell) \sqcup \ell' - \ell \sqcup \Psi_{\mathbf{0}}^{\mathbf{i}, \mathbf{i}}(\ell')).$$

Proof. For $t \in [0, T]$ and $\varepsilon > 0$, take an arbitrary continuous semimartingale X and define

$$h_t(\varepsilon) := \langle \ell \sqcup \ell', \mathbb{X}_t \otimes \exp^{\otimes}(\varepsilon \mathbf{i}) \otimes \mathbb{X}_{t,T} \rangle.$$

Since $\mathbb{X}_t \otimes \exp^{\otimes}(\varepsilon \mathbf{i}) \otimes \mathbb{X}_{t,T}$ is group-like (i.e., corresponds to the signature of a path as a product of signatures), it satisfies the shuffle property:

$$h_t(\varepsilon) = \langle \ell \sqcup \ell', \mathbb{X}_t \otimes \exp^{\otimes}(\varepsilon \mathbf{i}) \otimes \mathbb{X}_{t,T} \rangle = \langle \ell, \mathbb{X}_t \otimes \exp^{\otimes}(\varepsilon \mathbf{i}) \otimes \mathbb{X}_{t,T} \rangle \langle \ell', \mathbb{X}_t \otimes \exp^{\otimes}(\varepsilon \mathbf{i}) \otimes \mathbb{X}_{t,T} \rangle.$$

Differentiating both sides twice with respect to ε at $\varepsilon = 0$ gives

$$\begin{aligned}
\left. \frac{d^2}{d\varepsilon^2} \right|_{\varepsilon=0} h_t(\varepsilon) &= \langle \ell \sqcup \ell', \mathbb{X}_t \otimes \mathbf{i} \otimes \mathbb{X}_{t,T} \rangle \\
&= \langle \ell, \mathbb{X}_t \otimes \mathbf{i} \otimes \mathbb{X}_{t,T} \rangle \langle \ell', \mathbb{X}_T \rangle + \langle \ell, \mathbb{X}_T \rangle \langle \ell', \mathbb{X}_t \otimes \mathbf{i} \otimes \mathbb{X}_{t,T} \rangle \\
&\quad + 2 \langle \ell, \mathbb{X}_t \otimes \mathbf{i} \otimes \mathbb{X}_{t,T} \rangle \langle \ell', \mathbb{X}_t \otimes \mathbf{i} \otimes \mathbb{X}_{t,T} \rangle.
\end{aligned}$$

Integrating this identity with respect to $\circ dX_t^k$ over $[0, T]$ and applying Theorem 5.3, Proposition 5.5, and (6.3) yields

$$\langle \Psi_{\mathbf{k}}^{\mathbf{i}, \mathbf{i}}(\ell \sqcup \ell'), \mathbb{X}_T \rangle = \langle \Psi_{\mathbf{k}}^{\mathbf{i}, \mathbf{i}}(\ell), \mathbb{X}_T \rangle \langle \ell', \mathbb{X}_T \rangle + \langle \ell, \mathbb{X}_T \rangle \langle \Psi_{\mathbf{k}}^{\mathbf{i}, \mathbf{i}}(\ell'), \mathbb{X}_T \rangle + 2 \langle \ell \diamond_{\mathbf{k}}^{\mathbf{i}, \mathbf{i}} \ell', \mathbb{X}_T \rangle,$$

so that

$$\langle \ell \diamond_{\mathbf{k}}^{\mathbf{i}, \mathbf{i}} \ell', \mathbb{X}_T \rangle = \frac{1}{2} \langle \Psi_{\mathbf{k}}^{\mathbf{i}, \mathbf{i}}(\ell \sqcup \ell') - \Psi_{\mathbf{k}}^{\mathbf{i}, \mathbf{i}}(\ell) \sqcup \ell' - \ell \sqcup \Psi_{\mathbf{k}}^{\mathbf{i}, \mathbf{i}}(\ell'), \mathbb{X}_T \rangle$$

Finally, noting that this equality remains true for all possible X , we conclude that (6.6) holds. \square

6.2 Integration by parts formula

The linearization property (6.4), proved in the previous subsection, suggests that taking $h = \mathbf{D}H$ for some $H \in \mathcal{P}(\widehat{\mathbb{W}}_T)$ is a natural choice for the function h in Proposition 6.1. However, one can obtain a much larger class of admissible functions h by taking

$$h = \left(\mathbf{D}^i \left\langle \ell_i^h, \widehat{\mathbb{W}}_T \right\rangle \right)_{i=1, \dots, d}, \quad \ell_i^h \in T(\mathbb{R}^{n+1}), \quad i = 1, \dots, d, \quad (6.6)$$

so that we associate with each function h a collection of coefficients

$$\vec{\ell}^h = (\ell_1^h, \dots, \ell_d^h) \in T(\mathbb{R}^{d+1})^d.$$

A natural extension of the ‘‘diamond’’ product \diamond to vectors of coefficients is

$$\vec{\ell} \diamond \vec{\ell}' := \sum_{i=1}^d \ell_i \diamond_{\mathbf{0}}^{i,i} \ell'_i, \quad \vec{\ell}, \vec{\ell}' \in T(\mathbb{R}^{d+1})^d.$$

Hence, for $G = \langle \ell^G, \widehat{\mathbb{W}}_T \rangle$ and h of the form (6.2), we obtain

$$\langle \mathbf{D}G, h \rangle_{L^2} = \langle \ell^G \diamond \vec{\ell}^h, \widehat{\mathbb{W}}_T \rangle, \quad (6.7)$$

where we identify ℓ^G with the vector $\vec{\ell}^G = (\ell^G, \dots, \ell^G)$.

Before stating the main theorem of this section, we introduce the class $\mathcal{R}(\widehat{\mathbb{W}}_T)$ of random variables that can be written as a ratio of two linear functionals of the Brownian signature.

Definition 6.7. *We say that a random variable F belongs to $\mathcal{R}(\widehat{\mathbb{W}}_T)$ if there exist $\ell_1, \ell_2 \in T(\mathbb{R}^d)$ such that $\langle \ell_2, \widehat{\mathbb{W}}_T \rangle \neq 0$ a.s. and*

$$F = \frac{\langle \ell_1, \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2, \widehat{\mathbb{W}}_T \rangle} \quad \text{a.s.}$$

Note that $\mathcal{R}(\widehat{\mathbb{W}}_T)$ forms an algebra.

Theorem 6.8. *Under the assumptions of Proposition 6.1, suppose that*

1. $G \in \mathcal{P}(\widehat{\mathbb{W}}_T)$ with the coefficient $\ell^G \in T(\mathbb{R}^d)$.

2. $F \in \mathcal{R}(\widehat{\mathbb{W}}_T)$, i.e., $F = \frac{\langle \ell_1^F, \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle}$ for some $\ell_1^F, \ell_2^F \in T(\mathbb{R}^d)$.

3. h is a process of the form (6.2).

4. The coefficients $\ell^G, \ell^{F_1}, \ell^{F_2}, \vec{\ell}^h$ are such that $\frac{F}{\langle \mathbf{D}G, h \rangle_{L^2}} = \frac{\langle \ell_1^F, \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle \langle \ell^G \diamond \vec{\ell}^h, \widehat{\mathbb{W}}_T \rangle} \in \mathbb{D}^{1,2}$.

Then,

$$\mathbb{E}[f'(G)F] = \mathbb{E}[f(G)\pi_T], \quad (6.8)$$

where $\pi_T \in \mathcal{R}(\widehat{\mathbb{W}}_T)$ is given by

$$\begin{aligned} \pi_T = & \frac{\langle \ell_1^F, \widehat{\mathbb{W}}_T \rangle \langle \sum_{i=1}^d (\Lambda_i - \Psi_0^{ii})(\ell_i^h), \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle \langle \ell^G \diamond \vec{\ell}^h, \widehat{\mathbb{W}}_T \rangle} - \frac{\langle \ell_1^F \diamond \vec{\ell}^h, \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle \langle \ell^G \diamond \vec{\ell}^h, \widehat{\mathbb{W}}_T \rangle} \\ & + \frac{\langle \ell_1^F, \widehat{\mathbb{W}}_T \rangle \langle \ell_2^F \diamond \vec{\ell}^h, \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle^2 \langle \ell^G \diamond \vec{\ell}^h, \widehat{\mathbb{W}}_T \rangle} + \frac{\langle \ell_1^F, \widehat{\mathbb{W}}_T \rangle \langle \ell^G \diamond \vec{\ell}^h \diamond \vec{\ell}^h, \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle \langle \ell^G \diamond \vec{\ell}^h, \widehat{\mathbb{W}}_T \rangle^2}. \end{aligned} \quad (6.9)$$

Proof. The proof relies on formula (6.1), obtained in Proposition 6.1. Our goal is to show that each term in the Malliavin weight given by (6.1) can be written as a fraction of two random variables in $\mathcal{P}(\widehat{\mathbb{W}}_T)$.

First, all the terms in (6.1) containing $\langle \mathbf{D}G, h \rangle_{L^2}$ can be computed via (6.2).

To compute $\langle \mathbf{D}F, h \rangle$, we apply the chain rule (2.1):

$$\mathbf{D}F = \mathbf{D} \left(\frac{\langle \ell_1^F, \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle} \right) = \frac{\mathbf{D} \langle \ell_1^F, \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle} - \frac{\langle \ell_1^F, \widehat{\mathbb{W}}_T \rangle \mathbf{D} \langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle^2},$$

so that, using (6.2),

$$\langle \mathbf{D}F, h \rangle_{L^2} = \frac{\langle \ell_1^F \diamond \bar{\ell}^h, \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle} - \frac{\langle \ell_1^F, \widehat{\mathbb{W}}_T \rangle \langle \ell_2^F \diamond \bar{\ell}^h, \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle^2}. \quad (6.10)$$

As for the term $\delta(h)$, we observe that

$$\delta^i(h^i) = \delta^i \left(\mathbf{D}^i \langle \ell_i^h, \widehat{\mathbb{W}}_T \rangle \right) = -\mathbf{L}^i \langle \ell_i^h, \widehat{\mathbb{W}}_T \rangle = \langle (\Lambda_i - \Psi_0^{\text{ii}})(\ell_i^h), \widehat{\mathbb{W}}_T \rangle, \quad (6.11)$$

as follows from Corollary 5.17.

Combining (6.2), (6.2), and (6.2) and substituting the results into (6.1) yields (6.8).

It remains to notice that $\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle \neq 0$ a.s. since $F \in \mathcal{R}(\widehat{\mathbb{W}}_T)$, and that

$$\langle \ell^G \diamond \bar{\ell}^h, \widehat{\mathbb{W}}_T \rangle = \langle \mathbf{D}G, h \rangle_{L^2} \neq 0 \quad (a.s.),$$

by the assumption of Proposition 6.1, so that π_T belongs to $\mathcal{R}(\widehat{\mathbb{W}}_T)$. \square

Remark 6.9. Theorem 6.8 states that the integration by parts formula (6.8) preserves the class $\mathcal{R}(\widehat{\mathbb{W}}_T)$. This implies that if $f \in C^k$ for some $k \geq 2$, then one can iterate the formula to obtain

$$\mathbb{E} \left[f^{(k)}(G)F \right] = \mathbb{E} \left[f(G)\pi_T^k \right], \quad (6.12)$$

where $\pi_T^k \in \mathcal{R}(\widehat{\mathbb{W}}_T)$ is given recursively by (6.8), replacing ℓ_m^F by $\ell_m^{\pi_T^{k-1}}$ for $m \in \{1, 2\}$ and $k \geq 2$. \blacksquare

Remark 6.10. The formula (6.8) can be written explicitly in the compact form

$$\pi_T = \frac{\langle \ell_1^{\pi_T}, \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^{\pi_T}, \widehat{\mathbb{W}}_T \rangle}, \quad \langle \ell_2^{\pi_T}, \widehat{\mathbb{W}}_T \rangle \neq 0,$$

with

$$\begin{aligned} \ell_1^{\pi_T} &= \left(\sum_{i=1}^d (\Lambda_i - \Psi_0^{\text{ii}})(\ell_i^{\pi_T}) \right) \sqcup \ell_1^F \sqcup \ell_2^F \sqcup (\ell^G \diamond \bar{\ell}^h) - (\ell_1^F \diamond \bar{\ell}^h) \sqcup \ell_2^F \sqcup (\ell^G \diamond \bar{\ell}^h) \\ &\quad + \ell_1^F \sqcup (\ell_1^F \diamond \bar{\ell}^h) \sqcup (\ell^G \diamond \bar{\ell}^h) + \ell_1^F \sqcup \ell_2^F \sqcup (\ell^G \diamond \bar{\ell}^h \diamond \bar{\ell}^h), \\ \ell_2^{\pi_T} &= (\ell_2^F)^{\sqcup 2} \sqcup (\ell^G \diamond \bar{\ell}^h)^{\sqcup 2}. \end{aligned}$$

Although this makes the implementation of the recursive formula (6.9) straightforward, it has an important practical disadvantage: computing $\langle \ell_1^{\pi_T}, \widehat{\mathbb{W}}_T \rangle$ and $\langle \ell_2^{\pi_T}, \widehat{\mathbb{W}}_T \rangle$ requires evaluating $\widehat{\mathbb{W}}_t^{\leq N}$ for a much higher order N than would be necessary when using the ‘‘decomposed’’ formula (6.8). \blacksquare

Example 6.11. The following examples show that two widely used choices of h are covered by Theorem 6.8.

1. Consider first the simplest choice $h^i \equiv \mathbb{1}_{[0,T]}$, $i \in I$, for some set of indices $I \subset \{1, \dots, d\}$, as considered in (Al-Foraih et al., 2023). Note that $\mathbb{1}_{[0,T]} \equiv \mathbf{D}^i \langle \mathbf{i}, \widehat{\mathbb{W}}_T \rangle$, so that $\ell_i^h = \mathbb{1}_{\{i \in I\}} \mathbf{i}$. Since, by Corollary 6.5, $\ell \diamond_{\mathbf{0}}^{\mathbf{i}} \mathbf{i} = \Psi_{\mathbf{0}}^{\mathbf{i}}(\ell)$ for all $\ell \in T(\mathbb{R}^d)$, and since $(\Lambda_i - \Psi_{\mathbf{0}}^{\mathbf{i}})(\mathbf{i}) = \mathbf{i}$ by direct computation, the expression (6.8) simplifies to

$$\begin{aligned} \pi_T = & \frac{\langle \sum_{i \in I} \mathbf{i}, \widehat{\mathbb{W}}_T \rangle}{\langle \sum_{i \in I} \Psi_{\mathbf{0}}^{\mathbf{i}}(\ell^G), \widehat{\mathbb{W}}_T \rangle} F - \frac{\langle \sum_{i \in I} \Psi_{\mathbf{0}}^{\mathbf{i}}(\ell_1^F), \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle \langle \sum_{i \in I} \Psi_{\mathbf{0}}^{\mathbf{i}}(\ell^G), \widehat{\mathbb{W}}_T \rangle} \\ & + \frac{\langle \ell_1^F, \widehat{\mathbb{W}}_T \rangle \langle \sum_{i \in I} \Psi_{\mathbf{0}}^{\mathbf{i}}(\ell_2^F), \widehat{\mathbb{W}}_T \rangle}{\langle \ell_2^F, \widehat{\mathbb{W}}_T \rangle^2 \langle \sum_{i \in I} \Psi_{\mathbf{0}}^{\mathbf{i}}(\ell^G), \widehat{\mathbb{W}}_T \rangle} + \frac{\langle \sum_{i,j \in I} \Psi_{\mathbf{0}}^{\mathbf{i}}(\Psi_{\mathbf{0}}^{\mathbf{j}}(\ell^G)), \widehat{\mathbb{W}}_T \rangle}{\langle \sum_{i \in I} \Psi_{\mathbf{0}}^{\mathbf{i}}(\ell^G), \widehat{\mathbb{W}}_T \rangle^2} F. \end{aligned}$$

Despite the relative simplicity of this formula, it is necessary to justify that $\langle \sum_{i \in I} \Psi_{\mathbf{0}}^{\mathbf{i}}(\ell^G), \widehat{\mathbb{W}}_T \rangle \neq 0$ a.s. If this cannot be guaranteed, one might prefer the second option.

2. The second natural choice is $h^i = \mathbf{D}^i G$ (i.e. $\ell_i^h = \ell^G$) for $i \in I$. It is often used to establish the regularity of densities (see, for instance, Nualart (2006, Proposition 2.1.1)). This choice ensures that

$$\langle \mathbf{D}G, h \rangle_{L^2} = \sum_{i \in I} \|\mathbf{D}^i G\|^2 \geq 0. \quad (6.13)$$

The expression (2) is strictly positive if a.s. there exist $i \in I$ and $t \in [0, T]$ such that $\mathbf{D}_t^i G \neq 0$, which is much easier to verify than the corresponding condition in the first case. Note that, when $I = \{1, \dots, d\}$, the condition $\langle \mathbf{D}G, h \rangle_{L^2} \neq 0$ represents the standard non-degeneracy condition for the Malliavin covariance matrix, which reduces to a scalar in this setting. ■

Remark 6.12. The assumptions on the process h in Theorem 6.8 allow us not only to encompass the most common choices of h , as shown in the example above, but also to express the Malliavin weight using the unified formula (6.8). However, in general, other classes of processes h may be considered, and the corresponding integration by parts formulae can be derived using similar arguments. ■

7 Greeks in the signature volatility model

We illustrate our results by computing the Greeks within the Signature Volatility model. Throughout this section, we assume that $d = 2$. The asset price under the risk-neutral measure in the Signature Volatility model is given by

$$\begin{aligned} \frac{dS_t}{S_t} &= \sigma_t dB_t, \quad t \in [0, T], \\ \sigma_t &= \langle \boldsymbol{\sigma}, \widehat{\mathbb{W}}_t \rangle. \end{aligned} \quad (7.1)$$

where $\boldsymbol{\sigma} \in T(\mathbb{R}^3)$, $\rho \in [-1, 1]$, $\widehat{\mathbb{W}}_t$ denotes the signature of $\widehat{W}_t = (t, W_t^1, W_t^2)$, and $B_t = \rho W_t^1 + \bar{\rho} W_t^2$ with $\bar{\rho} = \sqrt{1 - \rho^2}$. We assume the interest rate is $r = 0$ and that the volatility process is independent of W^2 , meaning no letter **2** appears in $\boldsymbol{\sigma}$. To ensure the martingale property of the price process S established in (Abi Jaber et al., 2025a, Theorem 3.1), we take odd truncation orders N (that is, $\boldsymbol{\sigma}^{\mathbf{v}} = 0$ for all $|\mathbf{v}| > N$) and impose the condition $\rho \boldsymbol{\sigma}^{\mathbf{1}^{\otimes N}} < 0$, unless the volatility process is linear in W .

This model was shown to be universal in the sense that a large class of Markovian and non-Markovian stochastic volatility models can be exactly represented in the form (7) with an infinite collection of coefficients $\boldsymbol{\sigma}$, and can be well approximated by such models with finitely many coefficients $\boldsymbol{\sigma}$; see (Abi Jaber and Gérard, 2025a). This highlights the universality of the approach proposed in this paper, which can be applied to any model that can be written in the form (7) without any modification of the pricing or Greeks formulae.

We are interested in general path-dependent options with maturity $T > 0$ that depend on the signature $\widehat{\mathbb{X}}_T$ of the time-augmented log-price $\widehat{X} : t \mapsto (t, \log S_t)$ for $t \in [0, T]$. We show in Proposition 7.2 below that X_t

can be represented as a linear functional $\langle \boldsymbol{\ell}^X, \widehat{\mathbb{W}}_t \rangle$ of the Brownian signature. Hence, each element of $\widehat{\mathbb{X}}_T$ can be expressed as a linear functional of $\widehat{\mathbb{W}}_T$ as well; see (Cuchiero et al., 2023, Theorem 3.16) for an explicit formula. Therefore, we can assume the payoff is written in the form $f(G)$, $G = \langle \boldsymbol{\ell}^G, \widehat{\mathbb{W}}_T \rangle$, where $\boldsymbol{\ell}^G \in T(\mathbb{R}^3)$. The option price is given by $\mathbb{E}[f(G)]$. Our goal is to calculate the sensitivities

$$\frac{\partial}{\partial \theta} \mathbb{E}[f(G)], \quad (7.2)$$

with respect to the model parameters θ , where $\theta \in \{S_0, \rho, \boldsymbol{\sigma}\}$.

Assuming for a moment that $f \in C_b^1$ and that we can exchange the expectation and the differentiation, the Greek (7) is given by

$$\mathbb{E} \left[f' \left(\langle \boldsymbol{\ell}^G, \widehat{\mathbb{W}}_T \rangle \right) \left\langle \frac{\partial}{\partial \theta} \boldsymbol{\ell}^G, \widehat{\mathbb{W}}_T \right\rangle \right] = \mathbb{E} \left[f \left(\langle \boldsymbol{\ell}^G, \widehat{\mathbb{W}}_T \rangle \right) \pi_T \right]. \quad (7.3)$$

Note that $\left\langle \frac{\partial}{\partial \theta} \boldsymbol{\ell}^G, \widehat{\mathbb{W}}_T \right\rangle \in \mathcal{P}(\widehat{\mathbb{W}}_T) \subset \mathcal{R}(\widehat{\mathbb{W}}_T)$, so that Theorem 6.8 applies and the weight π_T is given by (6.8). This weight is independent of the function f , but it does depend on the option payoff through the coefficient $\boldsymbol{\ell}^G$.

Remark 7.1. The integration-by-parts result given by Theorem 6.8 requires the function f to be differentiable with a bounded derivative. However, functions of practical interest are typically non-differentiable, such as call option payoffs, or even discontinuous, such as digital option payoffs. In these cases, equation (7) no longer makes sense. Nevertheless, for a square-integrable function f , in the sense that

$$\mathbb{E} \left[f \left(\langle \boldsymbol{\ell}^G, \widehat{\mathbb{W}}_T \rangle \right)^2 \right] < \infty,$$

assuming that the maps $\theta \mapsto \mathbb{E} \left[f \left(\langle \boldsymbol{\ell}^G, \widehat{\mathbb{W}}_T \rangle \right)^2 \right]$ and $\theta \mapsto \mathbb{E}[\pi_T^2]$ are continuous, one can prove that

$$\frac{\partial}{\partial \theta} \mathbb{E} \left[f \left(\langle \boldsymbol{\ell}^G, \widehat{\mathbb{W}}_T \rangle \right) \right] = \mathbb{E} \left[f \left(\langle \boldsymbol{\ell}^G, \widehat{\mathbb{W}}_T \rangle \right) \pi_T \right]. \quad (7.4)$$

This is achieved by approximating f in L^2 by a sequence $(f_n)_{n \geq 0}$ of compactly supported, infinitely differentiable functions for which (7.1) holds, and then passing to the limit as $n \rightarrow \infty$. A rigorous proof of this approximation argument can be obtained by a straightforward adaptation of the proof of (Fournié et al., 1999, Proposition 3.2(ii)). Note that, for the examples considered below, the continuity assumptions are trivially satisfied.

Before considering concrete examples, we rewrite $\log S_T$ as a linear functional of the Brownian signature.

Proposition 7.2. *The log-price $X_t = \log S_t$ for $t \in [0, T]$ is given by*

$$X_t = \langle \boldsymbol{\ell}^X, \widehat{\mathbb{W}}_t \rangle, \quad \boldsymbol{\ell}^X := \log(S_0) \boldsymbol{\varnothing} - \frac{1}{2} \boldsymbol{\sigma} \uplus \boldsymbol{\sigma} \otimes \mathbf{0} + \rho \left(\boldsymbol{\sigma} \otimes \mathbf{1} - \frac{1}{2} \boldsymbol{\sigma}|_{\mathbf{1}} \otimes \mathbf{0} \right) + \bar{\rho} \boldsymbol{\sigma} \otimes \mathbf{2} \in T(\mathbb{R}^d), \quad (7.5)$$

where $\boldsymbol{\sigma}|_{\mathbf{i}} = \sum_{n \geq 0} \sum_{|\mathbf{v}|=v} \boldsymbol{\sigma}^{\mathbf{v}|\mathbf{i}}$.

Proof. We have

$$\begin{aligned} S_T &= S_0 \exp \left(-\frac{1}{2} \int_0^T \sigma_u^2 du + \int_0^T \sigma_u \rho dW_u^1 + \int_0^T \sigma_u \bar{\rho} dW_u^2 \right) \\ &= S_0 \exp \left(\left\langle -\frac{1}{2} \boldsymbol{\sigma} \uplus \boldsymbol{\sigma} \otimes \mathbf{0} + \rho \left(\boldsymbol{\sigma} \otimes \mathbf{1} - \frac{1}{2} \boldsymbol{\sigma}|_{\mathbf{1}} \otimes \mathbf{0} \right) + \bar{\rho} \boldsymbol{\sigma} \otimes \mathbf{2}, \widehat{\mathbb{W}}_T \right\rangle \right). \end{aligned}$$

where we used the identity $\int_0^T \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \rangle dW_t^i = \langle \boldsymbol{\ell} \otimes \mathbf{i} - \frac{1}{2} \boldsymbol{\ell}|_{\mathbf{i}} \otimes \mathbf{0}, \widehat{\mathbb{W}}_T \rangle$, which follows from the relationship between the Itô and Stratonovich integrals, as well as the dynamics of $\langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \rangle$:

$$\begin{aligned} d \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \rangle &= \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \otimes \mathbf{0} \rangle dt + \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \otimes \mathbf{1} \rangle \circ dW_t^1 + \langle \boldsymbol{\ell}, \widehat{\mathbb{W}}_t \otimes \mathbf{2} \rangle \circ dW_t^2, \\ &= \langle \boldsymbol{\ell}|_{\mathbf{0}}, \widehat{\mathbb{W}}_t \rangle dt + \langle \boldsymbol{\ell}|_{\mathbf{1}}, \widehat{\mathbb{W}}_t \rangle \circ dW_t^1 + \langle \boldsymbol{\ell}|_{\mathbf{2}}, \widehat{\mathbb{W}}_t \rangle \circ dW_t^2. \end{aligned}$$

It remains to notice that $\sigma|_2 = 0$ by assumption on σ . Taking the logarithm then concludes the proof. \square

7.1 European options

In this subsection, we consider a general European option with the payoff $f(\log S_T) = f(\langle \ell^X, \widehat{\mathbb{W}}_T \rangle)$ at maturity T , so that $\ell^G = \ell^X$.

It is straightforward to compute $F := \frac{\partial}{\partial \theta} \langle \ell^X, \widehat{\mathbb{W}}_T \rangle$ for $\theta \in \{S_0, \rho, \sigma\}$ using (7.2). Indeed,

$$\frac{\partial}{\partial S_0} \langle \ell^X, \widehat{\mathbb{W}}_T \rangle = \frac{1}{S_0}, \quad \frac{\partial}{\partial \rho} \langle \ell^X, \widehat{\mathbb{W}}_T \rangle = \left\langle \sigma \mathbf{1} - \frac{1}{2} \sigma|_1 \mathbf{0} - \frac{\rho}{\bar{\rho}} \sigma \mathbf{2}, \widehat{\mathbb{W}}_T \right\rangle,$$

and the derivatives $\frac{\partial}{\partial \sigma^{\mathbf{v}}} \langle \ell^X, \widehat{\mathbb{W}}_T \rangle$, $\mathbf{v} \in V$, are given by

$$\frac{\partial}{\partial \sigma^{\mathbf{v}}} \langle \ell^X, \widehat{\mathbb{W}}_T \rangle = \left\langle -(\mathbf{v} \sqcup \sigma) \mathbf{0} + \rho \left(\mathbf{v} \mathbf{1} - \frac{1}{2} \mathbf{v}|_1 \mathbf{0} \right) + \bar{\rho} \left(\mathbf{v} \mathbf{2} - \frac{1}{2} \mathbf{v}|_2 \mathbf{0} \right), \widehat{\mathbb{W}}_T \right\rangle.$$

We will compare the numerical computation of the option's delta (i.e., $\theta = S_0$, which implies $\ell_1^F = \emptyset$, $\ell_2^F = S_0 \emptyset$) for the following choices of the function h :

$$h_1 = (1, 0), \quad h_2 = (\mathbf{D}^1 X_T, 0), \quad h_3 = (0, 1), \quad h_4 = (0, \mathbf{D}^2 X_T).$$

In this case, we observe that since $\ell_1^F = \emptyset$, $\ell_2^F = S_0 \emptyset$, we have

$$\ell_1^F \diamond_0^{\mathbf{i}, \mathbf{i}} \ell^{h^i} = \ell_2^F \diamond_0^{\mathbf{i}, \mathbf{i}} \ell^{h^i} = 0, \quad \Psi_0^{\mathbf{i}}(\ell_1^F) = \Psi_0^{\mathbf{i}}(\ell_2^F) = 0, \quad i = 1, 2.$$

Furthermore, since all the non-zero coefficients of ℓ^X correspond to the words with at most one letter $\mathbf{2}$, we have

$$\ell^X \diamond_0^{\mathbf{2}, \mathbf{2}} \ell^X \diamond_0^{\mathbf{2}, \mathbf{2}} \ell^X = 0, \quad \Psi_0^{\mathbf{2}}(\Psi_0^{\mathbf{2}}(\ell^X)) = 0, \quad \Psi_0^{\mathbf{2}\mathbf{2}}(\ell^X) = 0.$$

Moreover, a straightforward computation shows that

$$\widehat{\Psi}_0^{\mathbf{2}}(\ell^X) = \bar{\rho} \sigma \mathbf{0}, \quad \ell^X \diamond_0^{\mathbf{2}, \mathbf{2}} \ell^X = \bar{\rho}^2 \sigma \sqcup \mathbf{2} \mathbf{0}, \quad \Lambda_2(\ell^X) = \bar{\rho} \sigma \mathbf{2}.$$

This leads to a significant simplification of the Malliavin weights. Indeed, applying Theorem 6.8 and the formulae derived in Example 6.11 yields the Malliavin weights given in the second column of Table 1.

Note that the formulas corresponding to h^3 and h^4 can be rewritten simply as

$$\frac{W_T^2}{\bar{\rho} S_0 \int_0^T \sigma_t dt} \quad \text{and} \quad \frac{\int_0^T \sigma_t dW_t^2}{\bar{\rho} S_0 \int_0^T \sigma_t^2 dt},$$

respectively. These weights can be obtained via standard integration by parts with respect to W^2 , without invoking Theorem 6.8. Hence, they do not depend on the choice of the volatility process as long as $\mathbf{D}^2 \sigma_t \equiv 0$. We include these weights as a sanity check, observing that our general result, Theorem 6.8, indeed covers this particular case and leads to the same expressions. Moreover, we observe that the weights corresponding to h_3 and h_4 diverge when $\rho \in \{-1, 1\}$.

Another possible choice of h is $h_t = \frac{1}{\mathbf{D}_t^2 X_T}$, as considered by Alós et al. (2021, Section 4.2.4). While this choice simplifies the computation of the weight, it introduces the integral $\int_0^T \frac{1}{\sigma_t} dW_t^2$, which is not well defined if the volatility process σ touches zero. Furthermore, this integral cannot be expressed as a linear functional of the signature and therefore requires discretization. This, in turn, results in a loss of numerical stability and tractability. For this reason, we do not include such estimators in our comparison.

h	Malliavin weight π_T	Signature order N ($\deg(\sigma) = M$)	Numerical stability
$h_1 = (1, 0)$	$\frac{\langle \mathbf{1}, \widehat{\mathbb{W}}_T \rangle}{S_0 \langle \Psi_0^1(\ell^X), \widehat{\mathbb{W}}_T \rangle} + \frac{\langle \Psi_0^1(\Psi_0^1(\ell^X)), \widehat{\mathbb{W}}_T \rangle}{S_0 \langle \Psi_0^1(\ell^X), \widehat{\mathbb{W}}_T \rangle^2}$	$N = 2M + 1$	✗
$h_2 = (\mathbf{D}^1 X_T, 0)$	$\frac{\langle (\Lambda_1 - \Psi_0^{11})(\ell^X), \widehat{\mathbb{W}}_T \rangle}{S_0 \langle \ell^X \diamond_0^{1,1} \ell^X, \widehat{\mathbb{W}}_T \rangle} + \frac{\langle \ell^X \diamond_0^{1,1} \ell^X \diamond_0^{1,1} \ell^X, \widehat{\mathbb{W}}_T \rangle}{S_0 \langle \ell^X \diamond_0^{1,1} \ell^X, \widehat{\mathbb{W}}_T \rangle^2}$	$N = 6M + 1$	✓
$h_3 = (0, 1)$	$\frac{\langle \mathbf{2}, \widehat{\mathbb{W}}_T \rangle}{\bar{\rho} S_0 \langle \sigma \mathbf{0}, \widehat{\mathbb{W}}_T \rangle}$	$N = M + 1$	✗
$h_4 = (0, \mathbf{D}^2 X_T)$	$\frac{\langle \sigma \mathbf{2}, \widehat{\mathbb{W}}_T \rangle}{\bar{\rho} S_0 \langle \sigma^{\perp \perp 2} \mathbf{0}, \widehat{\mathbb{W}}_T \rangle}$	$N = 2M + 1$	✓ ($ \rho \neq 1$)

Table 1: Comparison of Malliavin weights, signature order, and numerical stability for different h

Remark 7.3. Different choices of h generally require computing different maximal levels of the signature in order to obtain the exact expressions. For instance, if one takes the coefficient σ of degree M (that is, $\sigma^{\mathbf{v}} \neq 0$ only for \mathbf{v} such that $|\mathbf{v}| \leq M$), then the degree of the coefficients appearing in π_T is $6M + 1$ for h_2 , only $2M + 1$ for h_1 and h_4 , and $M + 1$ for h_3 . However, there is no theoretical guarantee that the denominators of the weights corresponding to h_1 and h_3 do not vanish. In Appendix C, we illustrate that it is indeed possible to find model parameters such that π_T blows up. We summarize the discussion above in Table 1. ■

We illustrate the convergence for standard vanilla and digital at-the-money (ATM) options

$$f_{\text{vanilla}}(x) = (e^x - K)^+, \quad f_{\text{digital}}(x) = \mathbf{1}_{\{e^x \geq K\}},$$

with $K = S_0$.

We provide in Figure 2 the convergence diagrams for (7) under the signature volatility model with parameters

$$\sigma = 0.2 \cdot \emptyset + 0.1 \cdot (\mathbf{1} + \mathbf{10} + \mathbf{110} + \mathbf{111}), \quad \rho = -1. \quad (7.6)$$

The signature truncation order was set to $N = 8$. Note that in this particular case of perfect correlation $\rho = -1$, the formulas corresponding to the functions h_3 and h_4 are not applicable, so we plot the results only for the first two functions h_1 and h_2 .

To reduce the variance of the Malliavin delta estimator, we apply the localization procedure proposed by Fournié et al. (1999) with $\delta = 10$, which consists in localizing the singularity of the payoff. More precisely, the payoff function f is decomposed as

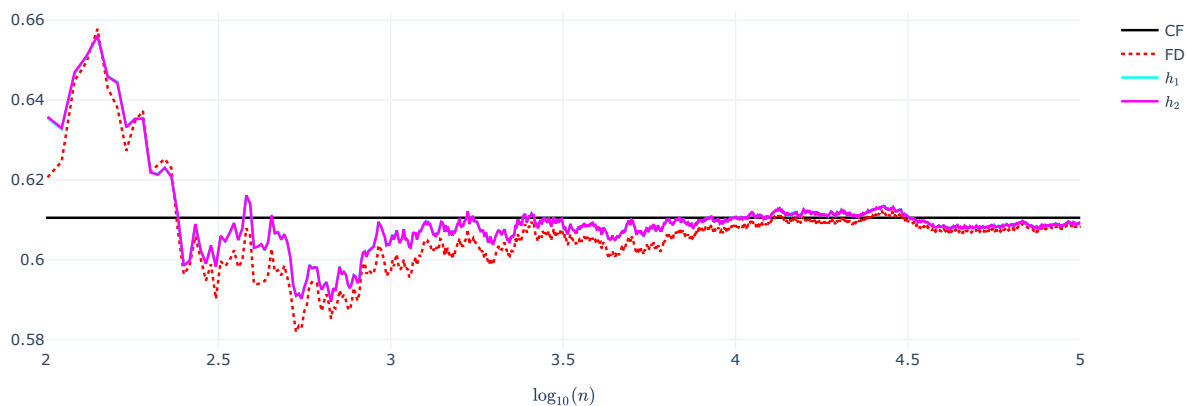
$$f(x) = G_\delta(x) + F_\delta(x), \quad x \in \mathbb{R},$$

where $G_\delta \in C_b^1$ is a regular function, while F_δ has a singularity at $x = \log K$ but vanishes whenever $|e^x - K| > \delta$. One then applies the Malliavin method only to the function F_δ in order to reduce the Monte Carlo variance induced by the weight π_T , while using pathwise differentiation (i.e., differentiation under the expectation) for the function G_δ . For the explicit construction of the functions G_δ and F_δ in the case of a call option payoff, we refer to (Fournié et al., 1999, Section 4).

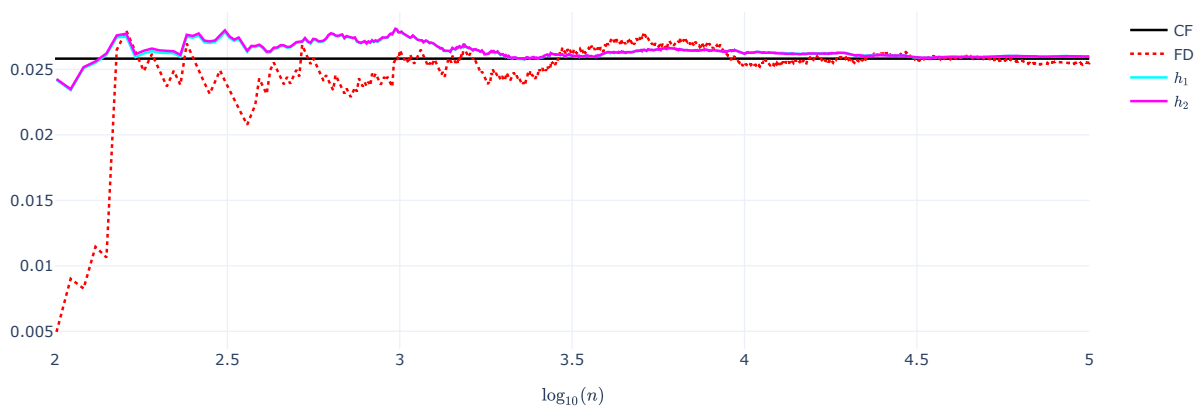
We compare the Malliavin-based Monte Carlo approach using $n_{\max} = 10^5$ trajectories with the option's delta obtained from the characteristic function of the log-price (see Appendix B), and with the standard finite-difference approach. The latter consists of approximating the option's delta by

$$\frac{\partial}{\partial s} \mathbb{E}[f(\log S_T) | S_0 = s] \approx \frac{\mathbb{E}[f(\log S_T) | S_0 = (1 + \varepsilon)s] - \mathbb{E}[f(\log S_T) | S_0 = (1 - \varepsilon)s]}{2\varepsilon s},$$

where the expectations are estimated via the Monte Carlo method using common random numbers. Throughout the experiments, we fix $\varepsilon = 0.01$.



(a) Delta of the vanilla ATM call option.



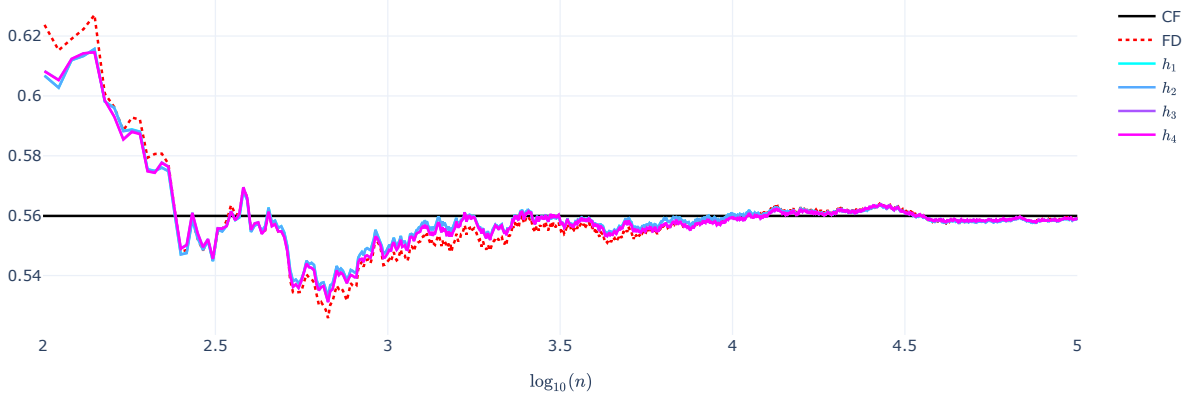
(b) Delta of the digital ATM call option.

Figure 2: Convergence diagram for vanilla and digital ATM options in the perfect-correlation model (7.1). The black horizontal line represents the delta computed using the characteristic function. The dotted red curve shows the results of the finite-difference method, while the two solid curves correspond to the Malliavin method with functions h_1 and h_2 , respectively.

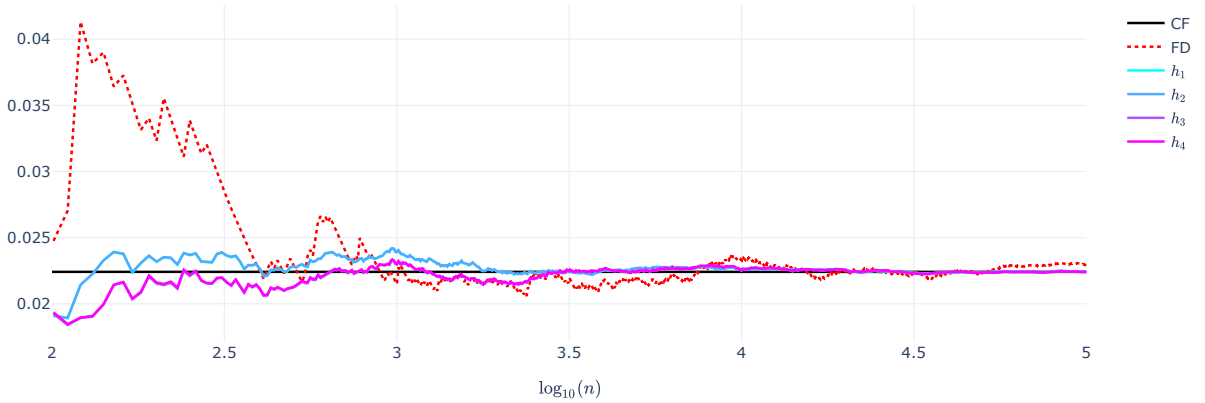
In Figure 3, we show the convergence diagram corresponding to the model

$$\sigma = 0.25 \cdot \emptyset + 0.04 \cdot (\mathbf{1} + \mathbf{01} + \mathbf{110} + \mathbf{111}), \quad \rho = -0.9, \quad (7.7)$$

with the signature truncated at order $N = 7$, as the number of simulations n varies from 0 to $n_{\max} = 10^5$. As expected, in both cases the Malliavin method exhibits faster convergence than the finite-difference method when the payoff function f is less regular (discontinuous in our case).



(a) Delta of the vanilla ATM call option.



(b) Delta of the digital ATM call option.

Figure 3: Convergence diagram for vanilla and digital ATM options in the stochastic volatility model (7.1). The dotted red curve shows the results of the finite-difference method, while the four solid curves correspond to the Malliavin method with functions h_1, h_2, h_3 , and h_4 , respectively.

7.2 Asian options

We now consider a genuinely path-dependent Asian option with payoff function $(S_t)_{t \in [0, T]} \mapsto f\left(\frac{1}{T} \int_0^T S_t dt\right)$, which depends on the average underlying price over $[0, T]$. Note that since $\log S_T = \langle \boldsymbol{\ell}^X, \widehat{\mathbb{W}}_T \rangle$, we have $S_T = e^{\langle \boldsymbol{\ell}^X, \widehat{\mathbb{W}}_T \rangle} = \langle \exp^{\sqcup}(\boldsymbol{\ell}^X), \widehat{\mathbb{W}}_T \rangle$ and

$$Y_T = \frac{1}{T} \int_0^T S_t dt = \left\langle \frac{1}{T} \exp^{\sqcup}(\boldsymbol{\ell}^X) \otimes \mathbf{0}, \widehat{\mathbb{W}}_T \right\rangle,$$

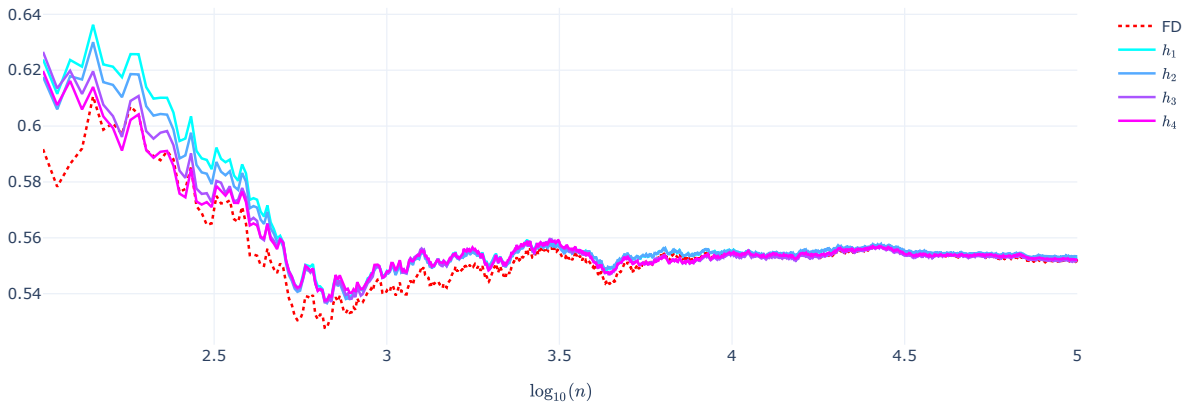
recall the definition of the shuffle exponential (3). We apply Theorem 6.8 to standard vanilla and digital payoff functions

$$f_{\text{vanilla}}(x) = (x - K)^+, \quad f_{\text{digital}}(x) = \mathbb{1}_{\{x \geq K\}},$$

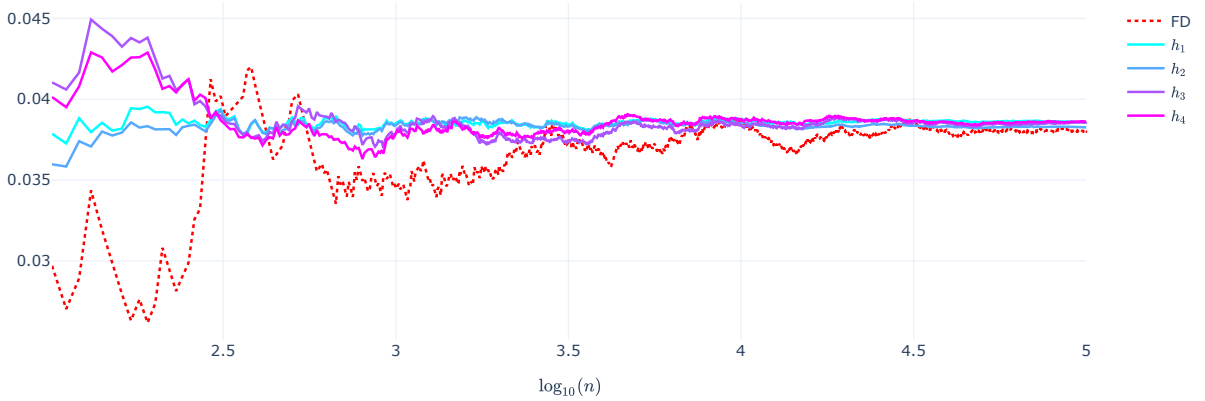
with $K = S_0$ and the coefficients

$$\ell^{G, \text{Asian}} = \ell_1^F = \frac{1}{T} \exp^{\sqcup}(\ell^X) \otimes \mathbf{0}, \quad \ell_2^F = S_0 \emptyset.$$

Unlike the European option case, where a significant simplification of the Malliavin weight was possible for the functions h_k , $k = 1, 2, 3, 4$, defined in the previous subsection, we now use the generic formula (6.8) for all four functions.



(a) Delta of the Asian ATM call option.



(b) Delta of the digital Asian ATM call option.

Figure 4: Convergence diagram for vanilla and digital Asian ATM options in the stochastic volatility model (7.2). The dotted red curve shows the results of the finite-difference method, while the four solid curves correspond to the Malliavin method with functions h_1, h_2, h_3 , and h_4 , respectively.

We consider the signature volatility model with parameters

$$\sigma = 0.25 \cdot \varnothing + 0.1 \cdot \mathbf{1} + 0.05 \cdot \mathbf{10}, \quad \rho = -0.9, \quad (7.8)$$

truncated at order $N = 8$.

The convergence diagrams are shown in Figure 4. Again, we observe faster convergence for the discontinuous payoff, while for the vanilla Asian option the estimator remains close to the one given by the finite-difference approach. We also note that $\ell^{G, \text{Asian}}$ is an infinite sequence, so truncation error is inevitable, especially when performing two “diamond” products in the weight formula (6.8). Despite this, we observe that the Malliavin method provides sufficiently accurate results even with a relatively small truncation order, namely $N = 8$.

A Malliavin differentiability of Stratonovich integrals

For simplicity of notation, we present the proof of Theorem 4.3 for the process

$$X_t = \int_0^t a_s \circ dW_s,$$

where a satisfies (2)–(4) and W is a one-dimensional Brownian motion. The general case can be treated analogously.

The proof is an adaptation of the results in (Nualart, 2006, Section 3.1.1) to the case of Stratonovich integrals considered as processes $\int_0^\cdot a_s \circ dW_s$, rather than random variables $\int_0^T a_s \circ dW_s$, and also establishes the existence of the derivative $(r, t) \mapsto \mathbf{D}_r \int_0^t a_s \circ dW_s$ as an element of $L^2(\Omega \times [0, T]^2)$. Our goal is therefore to show that

$$X \in \mathbb{L}^{1,2} \quad \text{and} \quad \mathbf{D}_r X_t = a_r \mathbf{1}_{[0,t]}(r) + \int_0^t \mathbf{D}_r a_s \circ dW_s.$$

Recall that for a square-integrable process h , the Stratonovich integral $\int_0^T h_s \circ dW_s$, when it exists, is defined as the limit in probability of the Riemann sums

$$S^\pi(h_\cdot) = \sum_{j=0}^{n-1} \frac{1}{t_{j+1} - t_j} \left(\int_{t_j}^{t_{j+1}} h_s ds \right) (W_{t_{j+1}} - W_{t_j}),$$

along partitions $\pi = \{0 = t_0 < t_1 < \dots < t_n = T\}$ as their diameter $|\pi| := \sup_{j=0, \dots, n-1} |t_{j+1} - t_j| \rightarrow 0$; see, e.g., Nualart (2006, Definition 3.1.1).

Let (Y, \mathcal{Y}, ν) be a measurable space with a σ -finite atomless measure ν , and let $h \in L^2(\Omega \times Y \times [0, T])$. In what follows, we will need to define the Stratonovich integrals $\int_0^T h(y, s) \circ dW_s$, depending on a parameter $y \in Y$, as elements of $L^2(\Omega \times Y)$.

We denote by $P^\pi h$ the stepwise approximation of h corresponding to the partition π :

$$P^\pi h(y, \cdot) = \sum_{j=0}^{n-1} \frac{1}{t_{j+1} - t_j} \int_{t_j}^{t_{j+1}} h(y, s) ds \mathbf{1}_{(t_j, t_{j+1}]}(\cdot),$$

which, for almost every $(\omega, y) \in \Omega \times Y$, is the orthogonal projection of $h(y, \cdot)$ onto the space of step functions generated by $\{\mathbf{1}_{(t_j, t_{j+1}]}\}_{j=0}^{n-1}$ associated with the partition π .

The following lemma extends (Nualart, 2006, Lemma 3.1.2).

Lemma A.1. *If $h \in L^2(\Omega \times Y \times [0, T])$, then*

$$\|P^\pi h - h\|_{L^2(\Omega \times Y \times [0, T])} \xrightarrow{|\pi| \rightarrow 0} 0.$$

Proof. For almost every $(\omega, y) \in \Omega \times Y$, we have $h(y, \cdot) \in L^2([0, T])$, and by the density of step functions,

$$\int_0^T |P^\pi h(y, s) - h(y, s)|^2 ds \xrightarrow{|\pi| \rightarrow 0} 0.$$

Moreover,

$$\int_0^T |P^\pi h(y, s) - h(y, s)|^2 ds \leq 2 \int_0^T |P^\pi h(y, s)|^2 ds + 2 \int_0^T |h(y, s)|^2 ds \leq 4 \int_0^T |h(y, s)|^2 ds,$$

where we used the projection property $\|P^\pi h(y, \cdot)\|_{L^2([0, T])} \leq \|h(y, \cdot)\|_{L^2([0, T])}$. Integrating this bound with respect to $\mathbb{P} \otimes \nu(dy)$ yields

$$\|P^\pi h - h\|_{L^2(\Omega \times Y \times [0, T])}^2 = \mathbb{E} \left[\int_Y \int_0^T |P^\pi h(y, s) - h(y, s)|^2 ds \nu(dy) \right] \leq 4 \|h\|_{L^2(\Omega \times Y \times [0, T])}^2.$$

The dominated convergence theorem then implies

$$\|P^\pi h - h\|_{L^2(\Omega \times Y \times [0, T])}^2 \xrightarrow{|\pi| \rightarrow 0} 0.$$

□

Lemma A.2. *If $h_n \rightarrow 0$ in $L^2(\Omega \times Y \times [0, T])$ and $\mathbf{D}h_n \rightarrow 0$ in $L^2(\Omega \times Y \times [0, T]^2)$ as $n \rightarrow \infty$, then*

$$\|\delta(s \mapsto h_n(\cdot, s))\|_{L^2(\Omega \times Y)} \xrightarrow{n \rightarrow \infty} 0,$$

where we recall that δ is the Skorokhod integral defined in Subsection 5.2.

Proof. The proof relies on the standard L^2 -bound for $\delta(u)$, where $u \in \mathbb{L}^{1,2}$, that is,

$$\mathbb{E}[\|u\|_{L^2([0, T])}^2] + \mathbb{E}[\|\mathbf{D}u\|_{L^2([0, T]^2)}^2] < \infty.$$

In this case, the following bound holds

$$\mathbb{E}[|\delta(u)|^2] \leq \mathbb{E}[\|u\|_{L^2([0, T])}^2] + \mathbb{E}[\|\mathbf{D}u\|_{L^2([0, T]^2)}^2]. \quad (\text{A.1})$$

For the proof, we refer to (Nualart, 2006, Proposition 1.3.1).

Applying (A) to $\delta(h_n(y, \cdot))$, integrating with respect to $\nu(dy)$ and applying Fubini's theorem, we obtain

$$\|\delta(s \mapsto h_n(\cdot, s))\|_{L^2(\Omega \times Y)}^2 \leq \|h_n\|_{L^2(\Omega \times Y \times [0, T])}^2 + \|\mathbf{D}h_n\|_{L^2(\Omega \times Y \times [0, T]^2)}^2 \xrightarrow{n \rightarrow \infty} 0.$$

□

Lemma A.3. *Suppose that $h \in L^2(\Omega \times Y \times [0, T])$ and $\mathbf{D}h \in L^2(\Omega \times Y \times [0, T]^2)$, and that there exist processes $\mathbf{D}^+h, \mathbf{D}^-h \in L^2(\Omega \times Y \times [0, T])$ satisfying*

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_Y \int_0^T \sup_{\substack{s < u < T \\ |u-s| \leq 1/n}} \mathbb{E} \left[|\mathbf{D}_s h(y, u) - (\mathbf{D}^+h)(y, s)|^2 \right] ds \nu(dy) &= 0, \\ \lim_{n \rightarrow \infty} \int_Y \int_0^T \sup_{\substack{0 < u < s \\ |s-u| \leq 1/n}} \mathbb{E} \left[|\mathbf{D}_s h(y, u) - (\mathbf{D}^-h)(y, s)|^2 \right] ds \nu(dy) &= 0. \end{aligned} \quad (\text{A.2})$$

Then, the Stratonovich integral $\int_0^T h(y, s) \circ dW_s$ is well-defined in $L^2(\Omega \times Y)$ as the limit of the Riemann sums $S^\pi(s \mapsto h(\cdot, s))$ as $|\pi| \rightarrow 0$.

Proof. We first note that the Riemann sums can be rewritten as

$$\begin{aligned} S^\pi(s \mapsto h(y, s)) &= \sum_{j=0}^{n-1} \frac{1}{t_{j+1} - t_j} \left(\int_{t_j}^{t_{j+1}} h(y, s) ds \right) (W_{t_{j+1}} - W_{t_j}) \\ &= \delta(P^\pi h(y, \cdot)) + \sum_{j=0}^{n-1} \frac{1}{t_{j+1} - t_j} \int_{t_j}^{t_{j+1}} \int_{t_j}^{t_{j+1}} \mathbf{D}_t h(y, s) ds dt. \end{aligned} \tag{A.3}$$

The final equality follows from the fact that $P^\pi h(y, \cdot)$ is a step function with random coefficients, in conjunction with the formula (5.2).

By Lemma A.1, the process $s \mapsto P^\pi h(\cdot, s)$ converges to h in $L^2(\Omega \times Y \times [0, T])$, and $\mathbf{D}P^\pi h$ converges to $\mathbf{D}h$ in $L^2(\Omega \times Y \times [0, T]^2)$. It then follows from Lemma A.2 that the Skorokhod integral term in (A) converges to $\delta(h(y, \cdot))$ in $L^2(\Omega \times Y)$.

The convergence of the second term in (A) to

$$\frac{1}{2} \int_0^T (\mathbf{D}^+ h)(y, s) ds + \frac{1}{2} \int_0^T (\mathbf{D}^- h)(y, s) ds$$

in $L^2(\Omega \times Y)$ is obtained by repeating the proof of (Nualart, 2006, Theorem 3.1.1), replacing the $L^1(\Omega)$ norms with $L^2(\Omega \times Y)$ norms. \square

A.1 Proof of Theorem 4.3

Step 1. Well-posedness of the Stratonovich integral. To show that the Stratonovich integral

$$t \mapsto \int_0^t a_s \circ dW_s$$

is well-defined as an element of $L^2(\Omega \times [0, T])$, we define the process $h(t, s) = a_s \mathbf{1}_{[0, t]}(s)$ and verify the assumptions of Lemma A.3 with $Y = [0, T]$ and $\nu(dy) = dy$. Clearly,

$$\|h\|_{L^2(\Omega \times [0, T]^2)}^2 = \mathbb{E} \left[\int_0^T \int_0^T |a_s \mathbf{1}_{[0, t]}(s)|^2 ds dt \right] \leq T \mathbb{E} \left[\int_0^T |a_s|^2 ds \right] < \infty,$$

and

$$\|\mathbf{D}h\|_{L^2(\Omega \times [0, T]^3)}^2 = \mathbb{E} \left[\int_0^T \int_0^T \int_0^T |\mathbf{D}_r a_s \mathbf{1}_{[0, t]}(s)|^2 ds dt dr \right] \leq T \mathbb{E} \left[\int_0^T \int_0^T |\mathbf{D}_r a_s|^2 ds dr \right] < \infty,$$

by (2).

We set $(\mathbf{D}^+ h)(t, s) := (\mathbf{D}^+ a)_s \mathbf{1}_{[0, t]}(s)$. Condition (A.3) becomes

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_0^T \int_0^T \sup_{\substack{s < u < T \\ |u-s| \leq 1/n}} \mathbb{E} \left[|\mathbf{D}_s h(t, u) - (\mathbf{D}^+ h)(t, s)|^2 \right] ds dt \\ &= \lim_{n \rightarrow \infty} \int_0^T \int_0^T \sup_{\substack{s < u < T \\ |u-s| \leq 1/n}} \mathbb{E} \left[|\mathbf{D}_s a_u \mathbf{1}_{[0, t]}(u) - (\mathbf{D}^+ a)_s \mathbf{1}_{[0, t]}(s)|^2 \right] ds dt \\ &\leq \lim_{n \rightarrow \infty} \left(T \int_0^T \sup_{\substack{s < u < T \\ |u-s| \leq 1/n}} \mathbb{E} \left[|\mathbf{D}_s a_u - (\mathbf{D}^+ a)_s|^2 \right] ds + \frac{1}{n} \int_0^T \mathbb{E} [|(\mathbf{D}^+ a)_s|^2] ds \right) = 0, \end{aligned}$$

by (3). The condition for $(\mathbf{D}^- h)(t, s) := (\mathbf{D}^- a)_s \mathbf{1}_{[0, t]}(s)$ is verified similarly. Lemma A.3 therefore yields convergence of the Riemann sums $S^\pi(s \mapsto h(\cdot, s))$ to the Stratonovich integral in $L^2(\Omega \times [0, T])$.

Step 2. Malliavin derivative of the Stratonovich integral. We next show that the $L^2(\Omega \times [0, T]^2)$ limit of $\mathbf{D}S^\pi(s \mapsto h(\cdot, s))$ exists and is equal to

$$\lim_{|\pi| \rightarrow 0} \mathbf{D}S^\pi(s \mapsto h(\cdot, s)) = a \mathbb{1}_{[0, \cdot]} + \int_0^T \mathbf{D}a_s \mathbb{1}_{[0, \cdot]}(s) \circ dW_s. \quad (\text{A.4})$$

By the closability of the operator $\mathbf{D}: L^2(\Omega \times [0, T]) \rightarrow L^2(\Omega \times [0, T]^2)$, this implies

$$\mathbf{D}_r \int_0^t a_s \circ dW_s = a_r \mathbb{1}_{[0, t]}(r) + \int_0^t \mathbf{D}_r a_s \circ dW_s,$$

for a.e. $(r, t) \in [0, T]^2$, and consequently $\int_0^\cdot a_s \circ dW_s \in \mathbb{L}^{1,2}$.

To prove (A.1), we compute

$$\begin{aligned} \mathbf{D}_r S_t^\pi(s \mapsto h(t, s)) &= \sum_{j=0}^{n-1} \frac{1}{t_{j+1} - t_j} \left(\int_{t_j}^{t_{j+1}} a_s \mathbb{1}_{[0, t]}(s) ds \right) \mathbb{1}_{(t_j, t_{j+1}]}(r) \\ &\quad + \sum_{j=0}^{n-1} \frac{1}{t_{j+1} - t_j} \left(\int_{t_j}^{t_{j+1}} \mathbf{D}_r a_s \mathbb{1}_{[0, t]}(s) ds \right) (W_{t_{j+1}} - W_{t_j}). \end{aligned} \quad (\text{A.5})$$

By Lemma A.1, the first term converges to $a_r \mathbb{1}_{[0, t]}(r)$ in $L^2(\Omega \times [0, T]^2)$ as $|\pi| \rightarrow 0$.

The second term in (A.5) is a Riemann sum corresponding to the process $\tilde{h} \in L^2(\Omega \times [0, T]^3)$ defined by $\tilde{h}(r, t, s) = \mathbf{D}_r a_s \mathbb{1}_{[0, t]}(s)$. It follows from (2) that

$$\|\tilde{h}\|_{L^2(\Omega \times [0, T]^3)}^2 < \infty, \quad \|\mathbf{D}\tilde{h}\|_{L^2(\Omega \times [0, T]^4)}^2 < \infty.$$

Define $(\mathbf{D}^+ \tilde{h})(r, t, s) := \mathbf{D}_r(\mathbf{D}^+ a)_s \mathbb{1}_{[0, t]}(s)$ and $(\mathbf{D}^- \tilde{h})(r, t, s) := \mathbf{D}_r(\mathbf{D}^- a)_s \mathbb{1}_{[0, t]}(s)$. Using the same argument as in Step 1 together with (4), we verify condition (A.3) for \tilde{h} . Applying Lemma A.3 then yields convergence in $L^2(\Omega \times [0, T]^2)$ of the second term in (A.5) to $\int_0^t \mathbf{D}_r a_s \circ dW_s$. This completes the proof.

B Greeks via the characteristic function

If the following two conditions are satisfied:

1. the model admits a computable characteristic function $\phi(u) := \mathbb{E} \left[e^{iu \log \frac{S_T}{S_0}} \right]$, $u \in \mathbb{R}$;
2. for a given payoff f , a Fourier inversion formula is available,

then one can obtain a quasi-analytic Fourier representation for Greeks with respect to S_0 , computable via numerical integration without the need for Monte Carlo simulation. The characteristic function $\phi(u)$ in the signature volatility model can be computed by solving an infinite-dimensional Riccati equation; see (Abi Jaber and Gérard, 2025a, Section 5).

For example, for a vanilla call option, Lewis' formula yields

$$C(T, K) = \mathbb{E}[(S_T - K)^+] = S_0 - \frac{K}{\pi} \int_0^\infty \Re \left(e^{i(u - \frac{i}{2}) \log \frac{S_0}{K}} \phi \left(u - \frac{i}{2} \right) \right) \frac{du}{u^2 + \frac{1}{4}}, \quad (\text{B.1})$$

so that the option Delta is given by formal differentiation with respect to S_0 :

$$\Delta(T, K) = 1 - \frac{1}{\pi} \frac{K}{S_0} \int_0^\infty \Re \left(i \left(u - \frac{i}{2} \right) e^{i(u - \frac{i}{2}) \log \frac{S_0}{K}} \phi \left(u - \frac{i}{2} \right) \right) \frac{du}{u^2 + \frac{1}{4}},$$

which can be computed much faster than Monte Carlo methods.

Similarly, using the relationship $C_{\text{digital}}(T, K) = -\partial_K C(T, K)$ between digital and vanilla call options, one obtains a Fourier representation for the digital option price by differentiating (B) with respect to K :

$$C_{\text{digital}}(T, K) = \mathbb{E}[\mathbf{1}_{\{S_T \geq K\}}] = \frac{1}{\pi} \int_0^\infty \Re \left(\left(\frac{1}{2} - iu \right) e^{i(u - \frac{i}{2}) \log \frac{S_0}{K}} \phi \left(u - \frac{i}{2} \right) \right) \frac{du}{u^2 + \frac{1}{4}}.$$

Taking the derivative with respect to S_0 , we obtain the Delta of the digital option:

$$\Delta_{\text{digital}}(T, K) = \frac{1}{\pi S_0} \int_0^\infty \Re \left(e^{i(u - \frac{i}{2}) \log \frac{S_0}{K}} \phi \left(u - \frac{i}{2} \right) \right) du.$$

This approach is rather formal, since we do not justify the differentiation of the integrals with respect to parameters, and no rigorous result concerning the existence of solutions to the infinite-dimensional Riccati equations is currently available. Hence, we emphasize that the Greeks obtained by this method are used only as sanity checks to demonstrate consistency with other methods.

C Numerical instabilities in the Malliavin weights

We consider the simplest possible stochastic volatility model,

$$\sigma = 0.02 \cdot \varnothing + 0.05 \cdot \mathbf{2}, \quad \rho = -0.95, \quad (\text{C.1})$$

with signature truncation order $N = 7$. Equivalently, the volatility is an affine function of Brownian motion, $\sigma_t = 0.02 + 0.05 W_t$. Even in this elementary setting, the denominators $\langle \mathbf{D}G, h \rangle$ appearing in the Malliavin weights of Theorem 6.8 may vanish, causing the method to become unstable for $h = h_1$ and $h = h_3$ from Table 1. Figure 5 shows that the support of $\langle \mathbf{D}G, h \rangle$ is contained in \mathbb{R}^+ for h_2 and h_4 , whereas for h_1 and h_3 , zero lies in the interior of the support. As a consequence, the corresponding Greek estimators exhibit large jumps, as seen in Figure 6, ultimately leading to divergence of the Monte Carlo estimator.

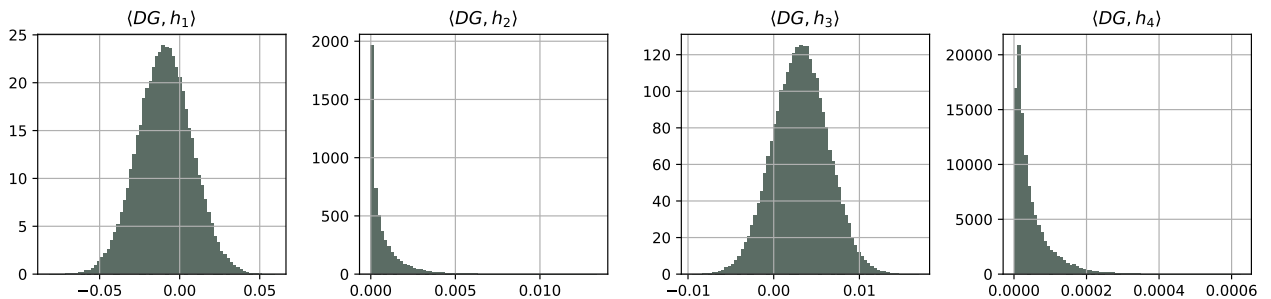


Figure 5: Empirical distribution of $\langle \mathbf{D}G, h \rangle$ for the functions h_k , $k = 1, 2, 3, 4$, listed in Table 1.

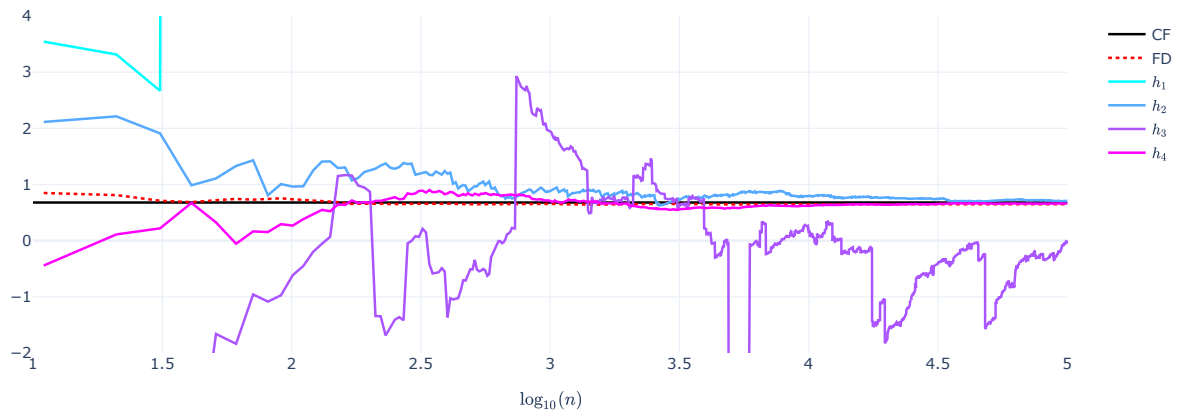


Figure 6: “Convergence” diagram for vanilla ATM options in the stochastic volatility model (C)

References

- Eduardo Abi Jaber and Louis-Amand Gérard. Signature volatility models: Pricing and hedging with Fourier. *Siam Journal on Financial Mathematics*, to appear, 2025a.
- Eduardo Abi Jaber and Louis-Amand Gérard. Hedging with memory: shallow and deep learning with signatures. [arXiv:2508.02759](https://arxiv.org/abs/2508.02759), 2025b.
- Eduardo Abi Jaber, Louis-Amand Gérard, and Yuxing Huang. Path-dependent processes from signatures. [arXiv:2407.04956](https://arxiv.org/abs/2407.04956), 2024.
- Eduardo Abi Jaber, Paul Gassiat, and Dimitri Sotnikov. Martingale property and moment explosions in signature volatility models. [arXiv:2503.17103](https://arxiv.org/abs/2503.17103), 2025a.
- Eduardo Abi Jaber, Donatien Hainaut, and Edouard Motte. Signature approach for pricing and hedging path-dependent options with frictions. [arXiv:2511.23295](https://arxiv.org/abs/2511.23295), 2025b.
- Mishari Al-Foraih, Jan Pospíšil, and Josep Vives. Computation of Greeks under rough Volterra stochastic volatility models using the Malliavin calculus approach. [arXiv:2312.00405](https://arxiv.org/abs/2312.00405), 2023.
- Elisa Alós, David Lorite, and Dariusz Gatarek. *Malliavin Calculus in Finance: Theory and Practice*. CRC Press, 05 2021. ISBN 9781003018681. doi: 10.1201/9781003018681.
- Imanol Perez Arribas, Cristopher Salvi, and Lukasz Szpruch. Sig-SDEs model for quantitative finance. In *Proceedings of the First ACM International Conference on AI in Finance*, pages 1–8, 2020.
- Vlad Bally. Lower bounds for the density of locally elliptic Itô processes. *The Annals of Probability*, 34(6): 2406 – 2440, 2006.
- Vlad Bally, Lucia Caramellino, and Antonino Zanette. Pricing and hedging american options by monte carlo methods using a malliavin calculus approach. *Monte Carlo Methods and Applications*, 11, 01 2003.
- Fabrice Baudoin. *An Introduction to the Geometry of Stochastic Flows*. Imperial College Press, 11 2004. ISBN 978-1-86094-481-9.
- Fabrice Baudoin, Qi Feng, and Cheng Ouyang. Density of the signature process of fbm. *Transactions of the American Mathematical Society*, 373(12):8583–8610, 2020.
- Christian Bayer, Paul P Hager, Sebastian Riedel, and John Schoenmakers. Optimal stopping with signatures. *The Annals of Applied Probability*, 33(1):238–273, 2023.

- Christian Bayer, Luca Pelizzari, and John Schoenmakers. Primal and dual optimal stopping with signatures. Finance and Stochastics, pages 1–34, 2025.
- G erard Ben Arous. Flots et s eries de Taylor stochastiques. Probability Theory and Related Fields, 81:29–77, 1989.
- Franziska Bielert. Rough functional Itô formula. arXiv:2409.02532, 2025.
- Jean-Michel Bismut. Martingales, the malliavin calculus and hypoellipticity under general h ormander’s conditions. Zeitschrift f ur Wahrscheinlichkeitstheorie und verwandte Gebiete, 56(4):469–505, 1981.
- Horatio Boedihardjo, Xi Geng, Terry Lyons, and Danyu Yang. The signature of a rough path: uniqueness. Advances in Mathematics, 293:720–737, 2016.
- Fabio Bugini, Michele Coghi, and Torstein Nilssen. Malliavin calculus for rough stochastic differential equations. arXiv:2402.12056, 2024.
-  lvaro Cartea, Imanol P erez Arribas, and Leandro S anchez-Betancourt. Double-execution strategies using path signatures. SIAM Journal on Financial Mathematics, 13(4):1379–1417, 2022.
- Thomas Cass and Emilio Ferrucci. On the Wiener chaos expansion of the signature of a Gaussian process. Probability Theory and Related Fields, 189(3):909–947, 2024.
- Kuo-Tsai Chen. Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. Annals of Mathematics, 65(1):163–178, 1957.
- Ilya Chevyrev and Andrey Kormilitzin. A primer on the signature method in machine learning. In Signature Methods in Finance: An Introduction with Computational Applications, pages 3–64. Springer Nature Switzerland Cham, 2025.
- Rama Cont and David-Antoine Fourni e. Functional Itô calculus and stochastic integral representation of martingales. The Annals of Probability, 41(1):109 – 133, 2013.
- Christa Cuchiero, Guido Gazzani, and Sara Svaluto-Ferro. Signature-based models: Theory and calibration. SIAM Journal on Financial Mathematics, 14(3):910–957, 2023.
- Christa Cuchiero, Guido Gazzani, Janka M oller, and Sara Svaluto-Ferro. Joint calibration to SPX and VIX options with signature-based models. Mathematical Finance, 35(1):161–213, 2025a.
- Christa Cuchiero, Xin Guo, and Francesca Primavera. Functional Itô-formula and Taylor expansions for non-anticipative maps of c adl ag rough paths. arXiv:2504.06164, 2025b.
- Mark Davis and Martin Johansson. Malliavin monte carlo Greeks for jump diffusions. Stochastic Processes and their Applications, 116:101–129, 01 2006.
- Nina Drobac, Margaux Br eg ere, Joseph de Vilmar est, and Olivier Wintenberger. Sliding-window signatures for time series: Application to electricity demand forecasting. arXiv:2510.12337, 2025.
- Bruno Dupire. Functional itô calculus. Quantitative Finance, 19(5):721–729, 2019.
- Bruno Dupire and Valentin Tissot-Daguette. Functional expansions. arXiv:2212.13628, 2022.
- Youssef El-Khatib. Computations of Greeks in stochastic volatility models via the Malliavin calculus. arXiv:0904.3247, 2009.
- Youssef El-Khatib and Nicolas Privault. Computations of Greeks in a market with jumps via the Malliavin calculus. Finance and Stochastics, 8, 09 2003.
- Thomas Fawcett. Problems in stochastic analysis. Connections between rough paths and non-commutative harmonic analysis. PhD thesis, University of Oxford, 2003.
- 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.
- Eric Fourni e, Jean-Michel Lasry, Pierre-Louis Lions, J er ome Lebuchoux, and Nizar Touzi. An application of Malliavin calculus to Monte Carlo methods in finance. Finance and Stochastics, 3:391–412, 09 1999.

- 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.
- Emmanuel Gobet and Arturo Kohatsu-Higa. Computation of greeks for barrier and look-back options using malliavin calculus. Electronic Communications in Probability, 8:51–62, 2003.
- Emmanuel Gobet and Remi Munos. Sensitivity analysis using itô–malliavin calculus and martingales, and application to stochastic optimal control. SIAM Journal on Control and Optimization, 43(5):1676–1713, 2005.
- Yuzuru Inahama. Malliavin differentiability of solutions of rough differential equations. Journal of Functional Analysis, 267(5):1566–1584, 2014.
- Samy Jazaerli and Yuri F. Saporito. Functional Itô calculus, path-dependence and the computation of Greeks. Stochastic Processes and their Applications, 127(12):3997–4028, 2017. ISSN 0304-4149.
- Peter E Kloeden and Eckhard Platen. Stochastic differential equations. Springer, 1992.
- Arturo Kohatsu-Higa. Lower bounds for densities of uniformly elliptic random variables on wiener space. Probability Theory and Related Fields, 126, 04 2003.
- Shigeo Kusuoka and Daniel Stroock. Applications of the Malliavin calculus, part III. J. Fac. Sci. Univ. Tokyo Sect IA Math, 34:391–442, 1987.
- Shigeo Kusuoka and Daniel W. Stroock. Applications of the Malliavin calculus. II. Journal of the Faculty of Science, the University of Tokyo. Sect. 1 A, Mathematics, 32:1–76, 1985.
- Daniel Levin, Terry Lyons, and Hao Ni. Learning from the past, predicting the statistics for the future, learning an evolving system. arXiv:1309.0260, 2016.
- Terry Lyons and Nicolas Victoir. Cubature on Wiener space. Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences, 460:169–198, 01 2004.
- Terry Lyons, Michael Caruana, and Thierry Lévy. Differential equations driven by rough paths. Springer, 2007.
- Terry Lyons, Sina Nejad, and Imanol Perez Arribas. Non-parametric pricing and hedging of exotic derivatives. Applied Mathematical Finance, 27(6):457–494, 2020.
- Paul Malliavin. Stochastic calculus of variation and hypoelliptic operators. In Proc. Intern. Symp. SDE Kyoto 1976, pages 195–263. Kinokuniya, 1978.
- Nicolas Marie. Sensitivities via rough paths. ESAIM: Probability and Statistics, 19:515–543, 2015.
- Ivan Nourdin and Frederi Viens. Density formula and concentration inequalities with Malliavin calculus. Electronic Journal of Probability, 14:2287 – 2309, 2009.
- David Nualart. The Malliavin calculus and related topics. Springer, 2006.
- Kevin D. Oden. Malliavin calculus as stochastic backpropogation. arXiv:2602.17013, 2025.
- Evangelia Petrou. Malliavin calculus in Lévy spaces and applications to finance. Electronic Journal of Probability, 13:852–879, 2008.
- Jakiw Pidstrigach, Elizabeth Louise Baker, Carles Domingo-Enrich, George Deligiannidis, and Nikolas Nüsken. Conditioning diffusions using malliavin calculus. In Forty-second International Conference on Machine Learning, 2025.
- Josef Teichmann. Calculating the Greeks by cubature formulae. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 462(2066):647–670, 2006.
- Bilgi Yilmaz. Computation of option Greeks under hybrid stochastic volatility models via Malliavin calculus. Modern Stochastics: Theory and Applications, 5(2):145–165, 2018.