Signature Methods in Variance Swap Pricing

Zuyuan Han, Haisu Ding
Mentors: Qi Feng, Bingyan Han

October 13, 2023

Abstract

This project addresses the intricate pricing challenge of variance swaps through the **application of signature methods** - a novel addition to quantitative finance area in recent years. We'll construct a **parameter-dense model** for the ultimate calibration process, a task characterized as a **high-dimensional regression**.

1 Introduction

Variance swap is a type of financial derivative that allows investors to trade or hedge against the volatility of an underlying asset. It is a contract in which two parties agree to exchange the realized variance of the underlying asset for a predetermined fixed payment.
The underlying asset of a variance swap is typically an equity index, such as the S&P 500, or a specific stock. The swap's payoff depends on the difference between the realized variance of the asset's returns over a specified period and a pre-agreed variance strike price. The realized variance represents the actual volatility observed in the market during the contract's duration. It is calculated based on the squared returns of the underlying asset over a specific time frame. The variance strike price K is a fixed level set at the inception of the swap, representing the expected or implied volatility of the underlying asset.

Theorem 1.1. *Mathematically, the payoff for a variance swap can be expressed as:*

$$Payoff = N \times (\sigma^2_{realized} - K)$$

Where

- N *is the notional amount of the swap*
- $\sigma^2_{realized}$ *is the realized variance of the underlying asset over the life of the swap.*
- K *is the variance strike, or the fixed variance level agreed upon at the initiation of the contract.*

Theorem 1.2. *The realized variance $\sigma^2_{realized}$ can be computed from the logarithmic returns of the underlying asset:*

$$\sigma^2_{realized} = \frac{1}{T}\sum_{i=1}^{T} r_i^2$$

Where

- r_i *is the logarithmic return at time i:* $r_i = ln(\frac{S_i}{S_{i-1}})$
- S_i *is the price of the underlying asset at time i.*
- T *is the total number of observations (e.g., trading days) over the life of the swap.*

Variance swaps are commonly used by investors and trades to hedge against or speculate on changes on volatility. By entering into a variance swap, market participants can effectively isolate and trade the volatility component of an asset's price movement, seperate from the asset's direction. This allows investors to manage their exposure to market volatility independently from their exposure to the asset's returns. Variance swaps provides a flexible and efficient means of gaining exposure to volatility, and they are often utilized by institutional investors, hedge funds and other sophisticated market participants. More information on variance swap we refer to *[JP Morgan] Variance Swap*[3].

In recent years, with the rise of big data and machine learning in finance, there have been efforts to employ neural networks and other machine learning models to predict the variance swap rates. *Ferguson & Green (2019)*[7] approximates pricing functions of derivatives using neural networks, whereas *Buehler et al. (2019)* focuses on optimal hedging using similar approaches. Although the learning methods have strong ability to adapt to complex patterns in data, model interpretability can be an issue. In contrast, model-based methods can capture volatility dynamics by using mathematical models such as Heston, Ornstein-Uhlenbeck process, etc., which describe the evolution of underlying assets.

With the inspiration of rough path theory work like *T. Lyons (1998)*[5], we find the recent development of signature method in quantitative finance. *C. Cuchiero, et al.(2022)*[1] formulates the time extended signature as the linear regression basis of continuous path functionals, aiming at applying data-driven, parameter-dense and tractable signature-based model in achieving outstanding calibration results in asset pricing problems.
Furthermore, *C. Cuchiero, et al.(2023)*[2] achieves highly accurate results in joint calibration to SPX and VIX options with signature-based models and the section 4 "*Expected signature of polynomial diffusion processes*" in *C. Cuchiero, et al.(2023)*[2] provides a direct inspiration of how signature approaches could be incorporated into describing polynomial processes and deriving close forms of expectation terms. Our work of taking signature model into computation starts from the pricing formula of variance swap and a small portion of our formula could have a similar form with VIX index presented as *Theorem 5.1* in *C. Cuchiero, et al.(2023)*[2]. Nevertheless, our model would be much more complex that VIX index and could involve higher dimensional calibration.

Theorem 1.3. *We consider asset price models whose dynamics are described by linear functions of the (time extended) signature of a primary underlying process, which is Brownian* *motion in our case. Throughout the paper we fix a filtered probability space* $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{Q})$ *on which we define the stochastic process*

$$dS_t = rS_t dt \ + \ \sigma_t S_t dB_t, \quad S_0 \in \mathbb{R}_+$$

where S_t *is the pricing process,* σ_t *is the volatility process that will be represented by signature model.* B_t *here is a Brownian motion which is correlated with* σ_t*.*

The variance swap contract is structured to ensure that losses are minimized when market volatility exceeds the agreed strike price. Thus in the following theorem, we formulate the closed form for our objective strike price K.

Theorem 1.4. *To price discretely sampled variance swaps, we first define pricing formula for vanilla variance swaps. Let the sampling dates be defined as*

$$0 = t_0 < t_1 < ... < t_N = T$$

where T *is the maturity date. The payoff of vanilla variance swap at* T *is defined as Theorem 1.1 and Theorem 1.2. where* K *is the strike price of the variance swap. To make the market free of arbitrage, the fair strike price* K *under risk-free measure* $\mathbb{Q}$ *is given by*

$$K = \mathbb{E}^{\mathbb{Q}}[\sum_{k=1}^{N}[\ln(\frac{S_{t_{k+1}}}{S_{t_k}})]^2|\mathcal{F}_0]$$

From Theorem 1.3, we get

$$S_{t_{k+1}} = S_{t_k} e^{r(t_{k+1}-t_k) - \int_{t_k}^{t_{k+1}} \frac{\sigma_s^2}{2} ds + \int_{t_k}^{t_{k+1}} \sigma_s dB_s}$$

Then

$$\begin{aligned}
K \ &= \ \mathbb{E}\Big[\sum_{k=0}^{N}\Big(r(t_{k+1}-t_k) - \int_{t_k}^{t_{k+1}} \frac{\sigma_s^2}{2} ds + \int_{t_k}^{t_{k+1}} \sigma_s dB_s\Big)^2 \Big|\mathcal{F}_0\Big] \\
&= \ \mathbb{E}\Big[\sum_{k=0}^{N}\Big(r^2(t_{k+1}-t_k)^2 \ - \ r(t_{k+1}-t_k)\int_{t_k}^{t_{k+1}} \sigma_s^2 ds \ + \ 2r(t_{k+1}-t_k)\int_{t_k}^{t_{k+1}} \sigma_s dB_s \\
&\quad + \ \frac{1}{4}(\int_{t_k}^{t_{k+1}} \sigma_s^2 ds)^2 \ - \ \int_{t_k}^{t_{k+1}} \sigma_s^2 ds \int_{t_k}^{t_{k+1}} \sigma_s dB_s \ + \ (\int_{t_k}^{t_{k+1}} \sigma_s dB_s)^2\Big)\Big|\mathcal{F}_0\Big] \\
&= \ \mathbb{E}\Big[\sum_{k=0}^{N}\Big(r^2(t_{k+1}-t_k)^2 \ - \ r(t_{k+1}-t_k)\int_{t_k}^{t_{k+1}} \sigma_s^2 ds \\
&\quad + \ \frac{1}{4}(\int_{t_k}^{t_{k+1}} \sigma_s^2 ds)^2 \ - \ \int_{t_k}^{t_{k+1}} \sigma_s^2 ds \int_{t_k}^{t_{k+1}} \sigma_s dB_s \ + \ \int_{t_k}^{t_{k+1}} \sigma_s^2 ds\Big)\Big|\mathcal{F}_0\Big] \\
&= \ \mathbb{E}\Bigg[\sum_{k=0}^{N}\Bigg(\ r^2(t_{k+1}-t_k)^2 \ + \ \big(1 \ - \ r(t_{k+1}-t_k)\big)\int_{t_k}^{t_{k+1}} \sigma_s^2 ds + \ \frac{1}{4}(\int_{t_k}^{t_{k+1}} \sigma_s^2 ds)^2 \\
&\quad - \ \int_{t_k}^{t_{k+1}} \sigma_s^2 ds \int_{t_k}^{t_{k+1}} \sigma_s dB_s\Bigg)\Bigg|\mathcal{F}_0\Bigg]
\end{aligned}$$

2 Signature Model

In the section we'll be introducing the *signature* based model. The *signature* is an object associated with a path which captures many of the path's important analytic and geometric properties. The detailed theoretical properties and some numerical applications are discussed in *Chevyrev & Kormilitzin (2016)*[8].

2.1 Notions

- (Tensor Algebra) For $n, d \in \mathbb{N}$, the n-fold tensor product of $\mathbb{R}^d$ is given by

$$(\mathbb{R}^d)^{\otimes n} = \underbrace{(\mathbb{R}^d)\otimes(\mathbb{R}^d)\otimes ... \otimes (\mathbb{R}^d)}_{\text{n}}$$

  where we construct an **nd** -dimensional space out of n vectors from **d**-dimensional space.
  For $d \in \mathbb{N}$, the extended tensor algebra on $\mathbb{R}^d$ is given by

$$T(\mathbb{R}^d) = \{\mathbf{a} = (a_0, a_1, ..., a_n) : a_i \in (\mathbb{R}^d)^{\otimes i},\ i = 0, 1, 2, ..., n\}$$

  which could provide a structural understanding of term *signature*, which should be represented as a collection of iterated integrals of a path.

- (Operations) Suppose $\mathbf{a} = (a_i)_{i=0}^{\infty}$, $\mathbf{b} = (b_i)_{i=0}^{\infty} \in T((\mathbb{R}^d))$, define the sum $+$ and product $\otimes$ by

$$\mathbf{a}\ +\ \mathbf{b}\ := (a_i + b_i)_{i=0}^{\infty},$$
$$\mathbf{a}\ \otimes\ \mathbf{b}\ := (\sum_{k=0}^{i} a_k \otimes b_{i-k})_{i=0}^{\infty}$$

- (Multi-index) Suppose a multi-index $I\ :=\ (i_1, i_2, ..., i_n)$, then we set $|I| := n$. Remark that we also set $I' := (i_1, ..., i_{n-1})$ and $I'' := (i_1, ..., i_{n-2})$ whenever needed.
  We also have the notation $\{I : |I| = n\}\ :=\ \{1, 2, ..., d\}^n$.
  Now we combine the multi-index with the tensor basis:

$$e_I\ =\ e_{i_1} \otimes e_{i_2} \otimes\ ...\ e_{i_n}$$

  where $e_{i_1}, e_{i_2}, ..., e_{i_n}$ denotes the canonical basis vectors of $\mathbb{R}^d$. Denoting that $e_\emptyset$ is the basis element corresponding to $(\mathbb{R}^d)^{\otimes 0}$.

- Suppose we have $e_I : |I| = N$ to be an orthonormal basis of $(\mathbb{R}^d)^{\otimes N}$. Then for any $\mathbf{a} \in T((\mathbb{R}^d))$, it can be written as

$$\mathbf{a} = \sum_{|I|>=0} a_I e_I$$

  where $a_I \in \mathbb{R}$ is a coefficient for a vector basis on multi-index I. Equivalently,

$$a_I = \langle e_I,\ \mathbf{a}\rangle$$

2.2 Definitions

Theorem 2.1 (Definition 2.1 [1]). *Let $(X_t)_{t\in[0,T]}$ be a continuous $\mathbb{R}^d$-valued semimartingale. The signature of X is the $T((\mathbb{R}^d))$-valued process $(s,t)\mapsto \mathbb{X}_{s,t}$ whose components are recursively defined as*

$$\langle e_\emptyset, \mathbb{X}_{s,t}\rangle = 1$$

$$\langle e_I, \mathbb{X}_{s,t}\rangle = \int_s^t \langle e_{I'}, \mathbb{X}_{s,t}\rangle \circ dX_r^{i_n}$$

for each $I=(i_1,i_2,...,i_n)$, $I'=(i_1,i_2,...,I_{n-1})$ and $0\leqslant s\leqslant t\leqslant T$, where $\circ$ denotes the Stratonovich integral. Its projection $\mathbb{X}^N$ on $T^{(N)}(\mathbb{R}^d)$ is given by

$$\mathbb{X}_{s,t}^N = \sum_{|I|\leqslant N} \langle e_I, \mathbb{X}_{s,t}\rangle e_I$$

and is called signature of X truncated at level N.
The equivalent notation:

$$\mathbb{X}_t = (1,\ \int_0^t 1\circ dX_s^1, ...,\ \int_0^t 1\circ dX_s^d,\ \int_0^t(\int_0^t 1\circ dX_s^1)\circ dX_s^1,\ ...\ ,\ \int_0^t(\int_0^t 1\circ dX_s^d)\circ dX_s^d,\ ...\ )$$

Signature is a collection of iterated integrals of the given multidimensional path. The integrals are listed in the collection in a strict order. Similar to 1.2.1 presented in *Chevyrev & Kormilitzin (2016)*[8], here is an example for a single signature term.

Example.

$$\langle e_1, \mathbb{X}_{s,t}\rangle = \int_0^t 1\circ dX_s^1 = X_t^1 - X_s^1$$

Note that $\langle e_1\otimes e_2, \mathbb{X}_{s,t}\rangle \neq \langle e_2\otimes e_1, \mathbb{X}_{s,t}\rangle$ as $\langle e_1\otimes e_2, \mathbb{X}_{s,t}\rangle = \int_0^t(\int_0^t 1\circ dX_s^1)\circ dX_s^2$ while $\langle e_2\otimes e_1, \mathbb{X}_{s,t}\rangle = \int_0^t(\int_0^t 1\circ dX_s^2)\circ dX_s^1$

Theorem 2.2 (Definition 2.2 [1]). *(Shuffle Product) Given two multi-indices $I=(i_1,i_2,...,i_n)$ and $J=(j_1,j_2,...,j_m)$, the **shuffle product** is defined recursively as*

$$e_I \sqcup\!\sqcup\, e_J = (e_{I'} \sqcup\!\sqcup\, e_J)\otimes e_{i_n} + (e_I \sqcup\!\sqcup\, e_{J'})\otimes e_{j_m}$$

where $e_I \sqcup\!\sqcup\, e_\emptyset \;=\; e_\emptyset \sqcup\!\sqcup\, e_I \;=\; e_I$.

Proposition 2.1. *Shuffle Property[Proposition 2.3 [1]] Let $((X_t))t\in[0,T]$ be a continuous $\mathbb{R}^d$-valued semimartingale and I, J be two multi-indices. Then*

$$\langle e_I, \mathbb{X}_{s,t}\rangle\langle e_J, \mathbb{X}_{s,t}\rangle \;=\; \langle e_I \sqcup\!\sqcup\, e_J, \mathbb{X}_{s,t}\rangle$$

Proof. *We prove this proposition by induction with Stratonovich integrals.*
Induction Basis. $\langle e_\emptyset, \mathbb{X}_{s,t}\rangle\langle e_J, \mathbb{X}_{s,t}\rangle \;=\; 1\ \langle e_J, \mathbb{X}\rangle = \langle e_\emptyset \sqcup\!\sqcup\, e_J, \mathbb{X}_{s,t}\rangle$, which agrees with the propostion.
Induction Steps. Assumed that for arbitrary subsets S_1 and S_2 of multi indices I and J, we have

$$\langle e_{S_1}, \mathbb{X}_{s,t}\rangle\langle e_{S_2}, \mathbb{X}_{s,t}\rangle \;=\; \langle e_{S_1} \sqcup\!\sqcup\, e_{S_2}, \mathbb{X}_{s,t}\rangle$$

After that,

$$
\begin{aligned}
\langle e_I, \mathbb{X}_{s,t}\rangle\langle e_J, \mathbb{X}_{s,t}\rangle &= \int_s^t \langle e_J, \mathbb{X}_{s,u}\rangle \circ d\langle e_I, \mathbb{X}_{s,u}\rangle + \int_s^t \langle e_I, \mathbb{X}_{s,u}\rangle \circ d\langle e_J, \mathbb{X}_{s,u}\rangle \\
&= \int_s^t \langle e_{I'}, \mathbb{X}_{s,r}\rangle\langle e_J, \mathbb{X}_{s,r}\rangle \circ d\mathbb{X}_r^{i_n} + \int_s^t \langle e_I, \mathbb{X}_{s,r}\rangle\langle e_{J'}, \mathbb{X}_{s,r}\rangle \circ d\mathbb{X}_r^{j_m} \\
&= \int_s^t \langle e_{I'} \sqcup\!\sqcup\, e_J, \mathbb{X}_{s,r}\rangle \circ d\mathbb{X}_r^{i_n} + \int_s^t \langle e_I \sqcup\!\sqcup\, e_{J'}, \mathbb{X}_{s,r}\rangle \circ d\mathbb{X}_r^{j_m} \\
&= \langle e_{I'} \sqcup\!\sqcup\, e_J \otimes e_{i_n}, \mathbb{X}_{s,t}\rangle + \langle e_I \sqcup\!\sqcup\, e_{J'} \otimes e_{j_m}, \mathbb{X}_{s,t}\rangle \\
&= \langle e_{I'} \sqcup\!\sqcup\, e_J \otimes e_{i_n} + e_I \sqcup\!\sqcup\, e_{J'} \otimes e_{j_m}, \mathbb{X}_{s,t}\rangle \\
&= \langle e_I \sqcup\!\sqcup\, e_J, \mathbb{X}_{s,t}\rangle
\end{aligned}
$$

Proposition 2.2 (Chen's identity)**.** *[Lemma 2.10 [1]] Let $(X_t)_{t\in[0,T]}$ be an $\mathbb{R}^d$-valued semi-martingale. Then*

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

for each $s \leqslant u \leqslant t \leqslant T$. This can equivalently be written as

$$\langle e_I, \mathbb{X}_{s,t}\rangle = \sum_{e_{I_1}\otimes e_{I_2} = e_I} \langle e_{I_1}, \mathbb{X}_{s,u}\rangle\langle e_{I_2}, \mathbb{X}_{u,t}\rangle$$

where I represents an arbitrary multi-index.

Proof. We prove this by induction on the length of the multi-index I.
Induction Basis. When $|I| = 0$, obviously the statement holds.
Induction Steps. Assumed that the statement holds for all multi-indices $J \subseteq I$, where $|J| < n$ and $I = (i_1, i_2, ..., i_n)$. Then,

$$
\begin{aligned}
\langle e_I, \mathbb{X}_{s,t}\rangle &= \int_s^t \langle e_{I'}, \mathbb{X}_{s,r}\rangle \circ d\mathbb{X}_r^{i_n} \\
&= \int_s^u \langle e_{I'}, \mathbb{X}_{s,r}\rangle \circ d\mathbb{X}_r^{i_n} + \int_u^t \langle e_{I'}, \mathbb{X}_{s,r}\rangle \circ d\mathbb{X}_r^{i_n} \\
&= \langle e_I, \mathbb{X}_{s,u}\rangle + \int_u^t \sum_{e_{I_1'}\otimes e_{I_2'} = e_{I'}} \langle e_{I_1'}, \mathbb{X}_{s,u}\rangle\langle e_{I_2'}, \mathbb{X}_{u,r}\rangle \circ d\mathbb{X}_r^{i_n} \\
&= \langle e_I, \mathbb{X}_{s,u}\rangle + \sum_{e_{I_1'}\otimes e_{I_2'} = e_{I'}} \langle e_{I_1'}, \mathbb{X}_{s,u}\rangle\langle e_{I_2'} \otimes e_{i_n}, \mathbb{X}_{u,t}\rangle \\
&= \sum_{e_{I_1}\otimes e_{I_2} = e_I} \langle e_{I_1}, \mathbb{X}_{s,u}\rangle\langle e_{I_2}, \mathbb{X}_{u,t}\rangle
\end{aligned}
$$

Note: paying attention to the difference between shuffle $\sqcup\!\sqcup$ and tensor product $\otimes$.

Theorem 2.3. *Our goal is to parameterize the volatility process σ_t as a signature model (Sig-SDE), which is, applying signature model to describe the dynamics of S_t as*

$$\sigma_t(\ell) = \ell_\emptyset + \sum_{|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{W}}_t\rangle$$

where W is the underlying d-dimensional Brownian motion and $\widehat{W}_t = (t, W_t)$.

Remark. With the same motivation,

$$(\sigma_t(\ell))^2 = \ell_\emptyset + \sum_{|I|,|J|\leq n} \ell_I \ell_J \langle e_I \sqcup\!\sqcup\, e_J, \widehat{\mathbb{W}}_t\rangle$$

3 Expected Signature

Theorem 3.1. *Since every polynomial function in the signature has a linear representation, suppose the d-dimensional real-valued process X is a weak solution of*

$$dX_t = b(\widehat{\mathbb{X}}_t)dt + \sqrt{a(\widehat{\mathbb{X}}_t)}dW_t$$

where W is a d-dimensional Brownian motion and a and b are linear maps. Thus we can define X as a polynomial process:

$$dX_t = b(X_t)dt + \sqrt{a(X_t)}dW_t, \quad X_0 = x_0$$

where maps $a : \mathbb{R}^d \mapsto \mathbb{S}^d_+$ and $b : \mathbb{R}^d \mapsto \mathbb{R}^d$ such that each a_{ij} is a polynomial of degree at most 2 and b_j is a polynomial of degree at most 1 for each $i,\ j = 1, ..., d+1$. Furthermore, to define the correlation between processes X and B, we introduce process $Z = (X_t, B_t)$ and denote its time extension by $\widehat{Z} = (t, X, B)$ with its signature $(\widehat{\mathbb{Z}}_t)_{t\geq 0}$. Thus, the correlation matrix process between X and B can be defined as

$$\rho_{ij} = \frac{[Z^i, Z^j]}{\sqrt{[Z^i]}\sqrt{[Z^j]}} \in [-1, 1]$$

for all $i, j = 1, ..., d+1$, where $[\ ,\ ]$ denotes the quadratic variation. In this case, the volatility process σ_t can be defined equivalently as a signature model (Sig-SDE) of the primary process Z:

$$\sigma_t(\ell) = \ell_\emptyset + \sum_{|I|\leq n, e_i\neq d+1} \ell_I \langle e_I, \widehat{\mathbb{Z}}_t\rangle$$

Theorem 3.2. *[Lemma 4.1 in [2]] Let $(X_t)_{t\geq 0}$ be the polynomial process given by (4) and b, a be the corresponding drift and diffusion coefficients. Then*

$$b_j(x) = b^c_j \ + \ \sum_{k=1}^{d} b^k_j x_k$$

$$a_{ij}(x) = a^c_{ij} + \sum_{k=1}^{d} a^k_{ij} x_k + \sum_{k,h=1}^{d} a^{kh}_{ij} x_k x_h$$

for $b^c_j, b^k_j, a^c_{ij}, a^k_{ij}, a^{kh}_{ij} \in \mathbb{R}$
Moreover we have,

$$b_j(Y_t) = \langle \boldsymbol{b}_j, \mathbb{Y}^1_t\rangle$$

and

$$a_{ij}(Y_t) = \langle \boldsymbol{a}_{ij}, \mathbb{Y}_t^2 \rangle$$

where

$$\boldsymbol{b}_j = b_j^c e_\emptyset + \sum_{k=1}^d (b_j^k X_0^k e_\emptyset \ + \ b_j^k e_k) = (b_j^c + \sum_{k=1}^d b_j^k X_0^k) e_\emptyset \ + \ \sum_{k=1}^d b_j^k e_k$$

$$\boldsymbol{a}_{ij} = (a_{ij}^c + \sum_{k=1}^d a_{ij}^k X_0^k + \sum_{k,h=1}^d a_{ij}^{kh} X_0^k X_0^h) e_\emptyset \ + \ \sum_{k=1}^d (a_{ij}^k \ + \ 2 \sum_{h=1}^d a_{ij}^{kh} X_0^h) e_k \ + \ \sum_{k,h=1}^d a_{ij}^{kh} e_k \sqcup e_h$$

Theorem 3.3 (Lemma 4.2 [2].)**.** *Let $(X_t)_{t \geq 0}$ be the polynomial process given by (4) and take b, a from Theorem 3.2. The truncated signature $(\mathbb{X}_t^n)_{t \geq 0}$ is a polynomial process and for each $|I| \leq n$ it holds that*

$$\langle e_I, \mathbb{X}_t^n \rangle = \int_0^t \langle Le_I, \mathbb{X}_s^n \rangle ds \ + \ \int_0^t \langle e_{I'}, \mathbb{X}_s^n \rangle \sigma_{i_{|I|}}(X_s) dW_s$$

where the operator $L : T(\mathbb{R}^d) \to T(\mathbb{R}^d)$ satisfies $L(T^{(n)}(\mathbb{R}^d)) \subseteq T^{(n)}(\mathbb{R}^d)$ and is given by

$$Le_I = e_{I'} \sqcup \boldsymbol{b}_{i_{|I|}} \ + \ \frac{1}{2} e_{I''} \sqcup \boldsymbol{a}_{i_{|I|-1} i_{|I|}}$$

Definition 3.1. [Definition 4.3 in [2].] For each $|I| \leq n$ set then $\eta_{IJ} \in \mathbb{R}$ such that

$$Le_I = \sum_{|J| \leq n} \eta_{IJ} e_J$$

Then we fix a labelling injective function $\mathcal{L} : \{I : |I| \leq n\} \to \{1, 2, ..., d_n\}$. We then call the matrix $G \in \mathbb{R}^{d_n \times d_n}$ where

$$G_{\mathcal{L}(I)\mathcal{L}(J)} = \eta_{IJ}$$

the d_n dimensional matrix representative of L.

Theorem 3.4 (Theorem 4.4 in [2].)**.** *Let $(X_t)_{t \geq 0}$ be the polynomial process given by (4). $((F)_t)_{t \geq 0}$ be the filtration generated by $(Y_t)_{t \geq 0}$ and let G be the d_n-dimensional matrix representative of the dual operator corresponding to $\mathbb{X}$. Then for each $T, t \geq 0, |I| \leq n$ we have*

$$\mathbb{E}[\langle e_I, \mathbb{X}_{T+t} \rangle | \mathcal{F}_T] = \sum_{|J| \leq n} (e^{tG^\intercal})_{\mathcal{L}(I)\mathcal{L}(J)} \langle e_J, \mathbb{X}_T \rangle$$

Here we introduce a lifting operator $\mathbf{P}$ which is defined as:

$$\mathbf{P}_t^I(\mathbb{X}_T) \ = \ \mathbb{E}[\langle e_I, \mathbb{X}_{T+t} \rangle | \mathcal{F}_T] = \sum_{|J| \leq n} (e^{tG^\intercal})_{\mathcal{L}(I)\mathcal{L}(J)} \langle e_J, \mathbb{X}_T \rangle$$

Remark. Here's an another case when $0 \leq t_0 < t_1 < t_2$, by Tower's property and Theorem 3.4

$$\begin{aligned}
\mathbb{E}\Big[\langle e_I, \mathbb{X}_{t_2}\rangle\langle e_H, \mathbb{X}_{t_1}\rangle \,|\mathcal{F}_{t_0}\Big] &= \mathbb{E}\,\Big[\, \mathbb{E}[\langle e_I, \mathbb{X}_{t_2}\rangle|\mathcal{F}_{t_1}]\, \langle e_H, \mathbb{X}_{t_1}\rangle \mid \mathcal{F}_{t_0}\Big] \\
&= \mathbb{E}\Big[\sum_{|J|\leq n} (e^{(t_2-t_1)G^\intercal})_{\mathcal{L}(I)\mathcal{L}(J)} \langle e_J, \mathbb{X}_{t_1}\rangle\langle e_H, \mathbb{X}_{t_1}\rangle \mid \mathcal{F}_{t_0}\Big] \\
&= \mathbb{E}\Big[\sum_{|J|\leq n} (e^{(t_2-t_1)G^\intercal})_{\mathcal{L}(I)\mathcal{L}(J)} \langle e_J \sqcup e_H, \mathbb{X}_{t_1}\rangle \mid \mathcal{F}_{t_0}\Big] \\
&= \sum_{|J|\leq n} (e^{(t_2-t_1)G^\intercal})_{\mathcal{L}(I)\mathcal{L}(J)} \mathbb{E}\Big[\langle e_J \sqcup e_H, \mathbb{X}_{t_1}\rangle \mid \mathcal{F}_{t_0}\Big] \\
&= \sum_{|J|\leq n} (e^{(t_2-t_1)G^\intercal})_{\mathcal{L}(I)\mathcal{L}(J)} \mathbf{P}^{J\sqcup H}_{t_1-t_0}(\mathbb{X}_{t_0}) \\
&= \Phi^{(I,t_2),(H,t_1)}(\mathbb{X}_{t_0})
\end{aligned}$$

4 Calibration of variance swap price with signature model

Theorem 4.1. *Assume Theorem 1.4 and the linear signature presentation for dynamic process, the risk-neutral strike K for variance swap is given by*

$$K = \sum_{k=0}^{N} \Big(r^2 (t_{k+1} - t_k)^2 + \big(1 - r(t_{k+1} - t_k)\big) \mathcal{K}^1_{t_k, t_{k+1}} + \frac{1}{4} \mathcal{K}^2_{t_k, t_{k+1}} - \mathcal{K}^3_{t_k, t_{k+1}} \Big)$$

where

- *(1)* $\mathcal{K}^1_{t_k,t_{k+1}} = \sum_{|I|,|J| \le n} \ell_I \ell_J \left(\mathbf{P}^{I \sqcup J \otimes \mathbf{0}}_{t_{k+1}}(\widehat{\mathbb{W}}_0) - \mathbf{P}^{I \sqcup J \otimes \mathbf{0}}_{t_k}(\widehat{\mathbb{W}}_0) \right)$

- *(2)* $\mathcal{K}^2_{t_k,t_{k+1}} = \sum_{|I|,|J|,|M|,|N| \le n} \ell_I \ell_J \ell_M \ell_N \Bigg(\mathbf{P}^{(I \sqcup J \otimes \mathbf{0}) \sqcup (M \sqcup N \otimes \mathbf{0})}_{t_{k+1}}(\widehat{\mathbb{W}}_0) + \mathbf{P}^{(I \sqcup J \otimes \mathbf{0}) \sqcup (M \sqcup N \otimes \mathbf{0})}_{t_k}(\widehat{\mathbb{W}}_0)$
  $- 2\Phi^{(I \sqcup J \otimes \mathbf{0}, t_{k+1}),(M \sqcup N \otimes \mathbf{0}, t_k)}(\widehat{\mathbb{W}}_0) \Bigg)$

- *(3)* $\mathcal{K}^3_{t_k,t_{k+1}} = \sum_{|I|,|J|,|M| \le n} \ell_I \ell_J \ell_M \Bigg(\mathbf{P}^{(I \sqcup J \otimes \mathbf{0}) \sqcup \tilde{M}^{d+1}}_{t_{k+1}}(\widehat{\mathbb{Z}}_0) + \mathbf{P}^{(I \sqcup J \otimes \mathbf{0}) \sqcup \tilde{M}^{d+1}}_{t_k}(\widehat{\mathbb{Z}}_0)$
  $- \Phi^{(I \sqcup J \otimes \mathbf{0}, t_{k+1}),(\tilde{M}^{d+1}, t_k)}(\widehat{\mathbb{Z}}_0) - \Phi^{(\tilde{M}^{d+1}, t_{k+1}),(I \sqcup J \otimes \mathbf{0}, t_k)}(\widehat{\mathbb{Z}}_0) \Bigg)$

Proof. See appendix A.

5 Model Modification

In section 4, our general process is given by $\widehat{Z}_t = (t, W_t, B_t)$ where t is the time process, W is the underlying d-dimensinal brownian motion and B is the one dimensional brownian motion from the pricing formula.

According Theorem 3.1, we introduced the general diffusion process X_t into our study. Note that the process X_t is also driven by underlying d-dimensinal brownian motion W_t. That provides an inspiration to us to apply high-level process into calibration instead of calibrating directly on the underlying brownian motions.

The only thing that needs to be modified here is the G matrix. Our future work would be incorporating OU process and Heston model into our project.

A Appendix A

The full proof for Theorem 4.1.

- *step* **1**. We take $\mathcal{K}^1_{t_k,t_{k+1}} = \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} \sigma_s^2 ds\Big]$. With Remark 2.2 where we apply signature formula in term σ and Theorem 3.4

$$\begin{aligned}
\mathcal{K}^1_{t_k,t_{k+1}} &= \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} \sigma_s^2 ds\Big] \\
&= \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} \sum_{|I|,|J|\le n} \ell_I \ell_J \langle e_I \sqcup e_J, \widehat{\mathbb{W}}_s\rangle ds\Big] \\
&= \mathbb{E}\Big[\int_{0}^{t_{k+1}} \sum_{|I|,|J|\le n} \ell_I \ell_J \langle e_I \sqcup e_J, \widehat{\mathbb{W}}_s\rangle ds\Big] - \mathbb{E}\Big[\int_{0}^{t_{k}} \sum_{|I|,|J|\le n} \ell_I \ell_J \langle e_I \sqcup e_J, \widehat{\mathbb{W}}_s\rangle ds\Big] \\
&= \mathbb{E}\Big[\sum_{|I|,|J|\le n} \ell_I \ell_J \langle e_I \sqcup e_J \otimes e_0, \widehat{\mathbb{W}}_{t_{k+1}}\rangle\Big] - \mathbb{E}\Big[\sum_{|I|,|J|\le n} \ell_I \ell_J \langle e_I \sqcup e_J \otimes e_0, \widehat{\mathbb{W}}_{t_{k}}\rangle\Big] \\
&= \sum_{|I|,|J|\le n} \ell_I \ell_J \mathbb{E}\Big[\langle e_I \sqcup e_J \otimes e_0, \widehat{\mathbb{W}}_{t_{k+1}}\rangle\Big] - \sum_{|I|,|J|\le n} \ell_I \ell_J \mathbb{E}\Big[\langle e_I \sqcup e_J \otimes e_0, \widehat{\mathbb{W}}_{t_{k}}\rangle\Big] \\
&= \sum_{|I|,|J|\le n} \ell_I \ell_J \mathbf{P}^{I\sqcup J\otimes \mathbf{0}}_{t_{k+1}}(\widehat{\mathbb{W}}_0) - \sum_{|I|,|J|\le n} \ell_I \ell_J \mathbf{P}^{I\sqcup J\otimes \mathbf{0}}_{t_{k}}(\widehat{\mathbb{W}}_0) \\
&= \sum_{|I|,|J|\le n} \ell_I \ell_J \Big(\mathbf{P}^{I\sqcup J\otimes \mathbf{0}}_{t_{k+1}}(\widehat{\mathbb{W}}_0) - \mathbf{P}^{I\sqcup J\otimes \mathbf{0}}_{t_{k}}(\widehat{\mathbb{W}}_0)\Big)
\end{aligned}$$

- *step* **2**. $\mathcal{K}^2_{t_k,t_{k+1}} = \mathbb{E}\Big[(\int_{t_k}^{t_{k+1}} \sigma_s^2 ds)^2\Big]$. With Theorem 3.4 and Remark 3

$$\begin{aligned}
\mathcal{K}^2_{t_k,t_{k+1}} &= \mathbb{E}\Big[(\int_{t_k}^{t_{k+1}} \sigma_s^2 ds)^2\Big] \\
&= \mathbb{E}\Big[(\int_{0}^{t_{k+1}} \sigma_s^2 ds - \int_{0}^{t_{k}} \sigma_s^2 ds)^2\Big|\mathcal{F}_0\Big] \\
&= \mathbb{E}\Big[(\int_{0}^{t_{k+1}} \sigma_s^2 ds)^2 + (\int_{0}^{t_{k}} \sigma_s^2 ds)^2 - 2(\int_{0}^{t_{k+1}} \sigma_s^2 ds \int_{0}^{t_{k}} \sigma_s^2 ds)\Big|\mathcal{F}_0\Big] \\
&= \sum_{|I|,|J|,|M|,|N|\le n} \ell_I \ell_J \ell_M \ell_N \Bigg(\mathbf{P}^{(I\sqcup J\otimes \mathbf{0})\sqcup(M\sqcup N\otimes \mathbf{0})}_{t_{k+1}}(\widehat{\mathbb{W}}_0) + \mathbf{P}^{(I\sqcup J\otimes \mathbf{0})\sqcup(M\sqcup N\otimes \mathbf{0})}_{t_{k}}(\widehat{\mathbb{W}}_0) \\
&\quad - 2\Phi^{(I\sqcup J\otimes \mathbf{0},t_{k+1}),(M\sqcup N\otimes \mathbf{0},t_k)}(\widehat{\mathbb{W}}_0)\Bigg)
\end{aligned}$$

- *step* **3**. Before we dive into the actual computation of the cross term, we introduce the special case when we take the extra correlated Brownian process B into the picture. If we take W as our multidimensional brownian process and the underlying primary process under the signature model of $\sigma = \sum_{|I|\le n} \ell_I \langle e_I, \widehat{\mathbb{W}}\rangle$, here we need to define the process $Z = (W, B)$ and the time-extended version $\widehat{Z} = (t, W, B)$ as well as its signature $\widehat{\mathbb{Z}}$.

Assumption A.1. *C. Cuchiero, G. Gazzani, S. Svaluto-Ferro(2022) [1] For all $i \in 1, 2, ..., d$, we have*

$$d[W^i, Z^{d+1}]_t = \sum_{|J|\leq m} a^J_{i(d+1)} \langle e_J, \widehat{\mathbb{Z}}_t \rangle dt$$

for some $m \in \mathbb{N}$ where $\widehat{Z} = (t, W, B)$.

Now we compute the special case correlated with B.

$$\begin{aligned}
\int_0^T \sigma_s dZ^{d+1}_s &= \int_0^T \sigma_s \circ dZ^{d+1}_s - \frac{1}{2}[\sigma, Z^{d+1}]_T \\
&= \int_0^T \sum_{|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{Z}}_s \rangle \circ dZ^{d+1}_s - \frac{1}{2}[\sum_{|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{Z}} \rangle, Z^{d+1}]_T \\
&= \sum_{|I|\leq n} \ell_I \langle e_I \otimes e_{d+1}, \widehat{\mathbb{Z}}_T \rangle - \frac{1}{2} \int_0^T \sum_{|I|\leq n} \ell_I \langle e_{I'}, \widehat{\mathbb{Z}}_s \rangle d[Z^{i_{|I|}}, Z^{d+1}]_s \\
&= \sum_{|I|\leq n} \ell_I (\ \langle e_I \otimes e_{d+1}, \widehat{\mathbb{Z}}_T \rangle - \frac{1}{2} \int_0^T \langle e_{I'}, \widehat{\mathbb{Z}}_s \rangle d[Z^{i_{|I|}}, Z^{d+1}]_s \ ) &\quad (1) \\
&= \sum_{|I|\leq n} \ell_I (\ \langle e_I \otimes e_{d+1}, \widehat{\mathbb{Z}}_T \rangle - \frac{1}{2} \sum_{|J|\leq m} a^J_{i_{|I|}(d+1)} \ \langle e_{I'} \sqcup\!\sqcup\, e_J \otimes e_0, \widehat{\mathbb{Z}}_T \rangle \ ) &\quad (2) \\
&= \sum_{|I|\leq n} \ell_I (\ \langle e_I \otimes e_{d+1} - \frac{1}{2} \sum_{|J|\leq m} a^J_{i_{|I|}d+1} \ e_{I'} \sqcup\!\sqcup\, e_J \otimes e_0, \widehat{\mathbb{Z}}_T \rangle \ ) \\
&= \sum_{|I|\leq n} \ell_I \ \langle \tilde{e}^{d+1}_I, \widehat{\mathbb{Z}}_T \rangle
\end{aligned}$$

where $\tilde{e}^{d+1}_I = e_I \otimes e_{d+1} - \frac{1}{2} \sum_{|J|\leq m} a^J_{i_{|I|}d+1} \ e_{I'} \sqcup\!\sqcup\, e_J \otimes e_0$.
(Note that the highest order of $\tilde{e}^{d+1}_I$ is $n + m$)
From (1) to (2) with assumption A.1 :

$$\begin{aligned}
\int_0^T \langle e_{I'}, \widehat{\mathbb{Z}}_s \rangle d[Z^{i_{|I|}}, Z^{d+1}]_s &= \int_0^T \langle e_{I'}, \widehat{\mathbb{Z}}_s \rangle \sum_{|J|\leq m} a^J_{i_{|I|}d+1} \langle e_J, \widehat{\mathbb{Z}}_s \rangle ds \\
&= \int_0^T \sum_{|J|\leq m} a^J_{i_{|I|}d+1} \langle e_{I'}, \widehat{\mathbb{Z}}_s \rangle \langle e_J, \widehat{\mathbb{Z}}_s \rangle ds \\
&= \sum_{|J|\leq m} a^J_{i_{|I|}d+1} \langle e_{I'} \sqcup\!\sqcup\, e_J \otimes e_0, \widehat{\mathbb{Z}}_T \rangle
\end{aligned}$$

Remark. When we are not incorporate any process form (like OU, Heston) into our model and we are calibrating directly on the underlying brownian motions, the term could be written as

$$\tilde{e}_I^{d+1} = e_I \otimes e_{d+1} - \frac{1}{2}\sigma^{i_{|I|}}\rho_{i_{|I|}d+1}\ e_{I'} \otimes e_0$$

Now we compute the explicit formula of the cross term.

$$\mathbb{E}[\int_{t_k}^{t_{k+1}} \sigma_s^2 ds \int_{t_k}^{t_{k+1}} \sigma_s dZ_s^{d+1}]$$

$$\begin{aligned}
&= \mathbb{E}\big[\int_{t_k}^{t_{k+1}} \sigma_s^2 ds \int_{t_k}^{t_{k+1}} \sigma_s dZ_s^{d+1} \big| \mathcal{F}_0\big] \\
&= \mathbb{E}\big[(\int_0^{t_{k+1}} \sigma_s^2 ds - \int_0^{t_k} \sigma_s^2 ds)(\int_0^{t_{k+1}} \sigma_s dZ^{d+1}s - \int_0^{t_k} \sigma_s dZ^{d+1}s) \big| \mathcal{F}_0\big] \\
&= \mathbb{E}\big[\int_0^{t_{k+1}} \sigma_s^2 ds \int_0^{t_{k+1}} \sigma_s dZ^{d+1}s + \int_0^{t_k} \sigma_s^2 ds \int_0^{t_k} \sigma_s dZ^{d+1}s \\
&- \int_0^{t_{k+1}} \sigma_s^2 ds \int_0^{t_k} \sigma_s dZ^{d+1}s - \int_0^{t_{k+1}} \sigma_s dZ^{d+1}s \int_0^{t_k} \sigma_s^2 ds \big| \mathcal{F}_0\big] \\
&= \sum_{|I|,|J|,|M| \le n} \ell_I \ell_J \ell_M \Big(\mathbf{P}_{t_{k+1}}^{(I \sqcup J \otimes \mathbf{0}) \sqcup \tilde{M}^{d+1}}(\widehat{\mathbb{Z}}_0) \ + \ \mathbf{P}_{t_k}^{(I \sqcup J \otimes \mathbf{0}) \sqcup \tilde{M}^{d+1}}(\widehat{\mathbb{Z}}_0) \\
&- \ \Phi^{(I \sqcup J \otimes \mathbf{0}, t_{k+1}), (\tilde{M}^{d+1}, t_k)}(\widehat{\mathbb{Z}}_0) \ - \ \Phi^{(\tilde{M}^{d+1}, t_{k+1}), (I \sqcup J \otimes \mathbf{0}, t_k)}(\widehat{\mathbb{Z}}_0) \Big)
\end{aligned}$$

References

[1] C. Cuchiero, G. Gazzani, S. Svaluto-Ferro(2022) *Signature-based models: theory and calibration*. From: arXiv:2207.13136v1 [q-fin.MF] 26 Jul 2022

[2] C.Cuchiero, G.Gazzani, J.Moller, S.S.Ferro(2023) ·*Joint calibration to SPX and VIX options with signature-based models*. From: arXiv:2301.13235v1 [q-fin.MF] 30 Jan 2023

[3] P.Allen, S.Einchcomb, N. Granger, (2006) ·*[JP Morgan] Variance Swap*. From: European Equity Derivatives Research, J.P. Morgan Securities Ltd. London, 17 November, 2006

[4] "Yahoo Finance" *CBOE Volatility Index (VIX) historical data*. 2023. *Website link*.

[5] Lyons, Terry J., (1998) *Differential Equations driven by Rough Signals*. From: Revista Mathemática Iberoamericana, vol.14, No.2, 1998

[6] Buehler et al. (2019) *Deep Hedging* From: Quantitative Finance, Vol 19, 2019- Issue 8

[7] Ferguson & Green (2019) *Deep Learning Derivatives*

[8] Chevyrev & Kormilitzin (2016) *A Primer on the Signature Method in Machine Learning*