Signatures methods in finance

Christa Cuchiero

University of Vienna

Mini course

Paris, Bachelier Seminar, March 2024

Part IV

Signature based asset price models for the joint calibration problem to SPX and VIX options

- based on joint work with Guido Gazzani, Janka Möller and Sara Svaluto-Ferro (https://arxiv.org/abs/2301.13235)

Outline

- Introduction
- VIX options with signature based models
- SPX options and hedging
- Joint calibration of SPX and VIX options

Signature based models

Highly parametric and overparametrized models gain in importance: instead of a few parameters, the goal is rather to learn the model's characteristics as a whole from data.

Relying on different universal approximation theorems yields different classes of models. We consider here ...

Signature based models

Highly parametric and overparametrized models gain in importance: instead of a few parameters, the goal is rather to learn the model's characteristics as a whole from data.

Relying on different universal approximation theorems yields different classes of models. We consider here ...

⇒ Signature based models:
the model itself or its characteristics are parameterized as linear functions of the signature of an underlying process, in the simplest case Brownian motion. Compare e.g. with

- I. Perez Arribas, C. Salvi, L. Szpruch ('20) "Sig-SDEs for quatitative finance"
- C.C., G. Gazzani, S.Svaluto-Ferro ('22) "Signature-based models: theory and calibration"
- E. Abi-Jaber and L.Gérard " Signature stochastic volatility models: pricing and hedging with Fourier"

Other recent applications of signature methods in finance: Akyildirim et al. ('22), Bayer et al. ('21), Bühler et al. ('20), Kalsi et al., Ni et al. ('20)), Salvi et al. ('21), etc.

Joint calibration problem and modeling framework

Goal: Find a model which calibrates jointly to options written on SPX and VIX, accurately and numerically efficiently.

- This is still regarded as a difficult problem in volatility modeling, even though significant progress has been made recently (see below in the literature overview).
- The challenge is, especially for short maturities, to reconcile the large negative skew of SPX options' implied volatilities with relatively lower implied volatilities arising from the VIX options.

Joint calibration problem and modeling framework

Goal: Find a model which calibrates jointly to options written on SPX and VIX, accurately and numerically efficiently.

- This is still regarded as a difficult problem in volatility modeling, even though significant progress has been made recently (see below in the literature overview).
- The challenge is, especially for short maturities, to reconcile the large negative skew of SPX options' implied volatilities with relatively lower implied volatilities arising from the VIX options.

Modeling framework:

- Stochastic volatility model under a risk-neutral probability measure $\mathbb{Q}$, to describe the discounted dynamics of the SPX for $t > 0$:

$$\mathrm{d}S_t(\ell) = S_t(\ell)\sigma_t^S(\ell)\mathrm{d}B_t, \quad \text{with} \quad \sigma_t^S(\ell) := \langle \ell, \widehat{\mathbb{X}}_t \rangle,$$

  i.e. the volatility σ_t^S is a linear map of the signature of a time-extended d-dimensional semimartingale X. The semimartingale X is fixed and its parameters are not trained, but can be considered as hyperparameters.

- The parameters ℓ have to be learned from option price data on SPX and VIX.

Implications of this modeling framework

- The modeling framework can be seen as universal in the class of continuous non-rough stochastic volatility models. It nests e.g. the recent work by Abi Jaber et al. (2022b) (apart from their additional input curve).

Implications of this modeling framework

- The modeling framework can be seen as universal in the class of continuous non-rough stochastic volatility models. It nests e.g. the recent work by Abi Jaber et al. (2022b) (apart from their additional input curve).

- If X is additionally a polynomial process, then VIX can be computed analytically via a matrix exponential.

Implications of this modeling framework

- The modeling framework can be seen as universal in the class of continuous non-rough stochastic volatility models. It nests e.g. the recent work by Abi Jaber et al. (2022b) (apart from their additional input curve).

- If X is additionally a polynomial process, then VIX can be computed analytically via a matrix exponential.

- Defining $\widehat{Z} := (t, X, B)$, then not only $\sigma^S(\ell)$ but also $\log(S(\ell))$ can be expressed as a linear function of the signature of $\widehat{Z}$.

Implications of this modeling framework

- The modeling framework can be seen as universal in the class of continuous non-rough stochastic volatility models. It nests e.g. the recent work by Abi Jaber et al. (2022b) (apart from their additional input curve).
- If X is additionally a polynomial process, then VIX can be computed analytically via a matrix exponential.
- Defining $\widehat{Z} := (t, X, B)$, then not only $\sigma^S(\ell)$ but also $\log(S(\ell))$ can be expressed as a linear function of the signature of $\widehat{Z}$.
- A Monte Carlo approach for option pricing, potentially with variance reduction, is tractable since we can compute offline the signature samples of $\widehat{Z}$.
  $\Rightarrow$ The calibration task can be split into an offline sampling and a standard optimization.

Implications of this modeling framework

- The modeling framework can be seen as universal in the class of continuous non-rough stochastic volatility models. It nests e.g. the recent work by Abi Jaber et al. (2022b) (apart from their additional input curve).

- If X is additionally a polynomial process, then VIX can be computed analytically via a matrix exponential.

- Defining $\widehat{Z} := (t, X, B)$, then not only $\sigma^S(\ell)$ but also $\log(S(\ell))$ can be expressed as a linear function of the signature of $\widehat{Z}$.

- A Monte Carlo approach for option pricing, potentially with variance reduction, is tractable since we can compute offline the signature samples of $\widehat{Z}$.
$\Rightarrow$ The calibration task can be split into an offline sampling and a standard optimization.

- We illustrate that the joint calibration problem can be solved in this framework without jumps and rough volatility (compare also Guyon and Lekeufack (2022); Abi Jaber et al. (2022b)).

VIX options with signature based models

The VIX Index

- The VIX is a popular measure of the market's expected volatility of the SPX, calculated and published by the Chicago Board Options Exchange (CBOE). It is given by

$$\mathrm{VIX}_T := \sqrt{\mathbb{E}\left[-\frac{2}{\Delta}\log\left(\frac{S_{T+\Delta}}{S_T}\right)|\mathcal{F}_T\right]},$$

where $\Delta = 30$ days and $S = (S_t)_{t\geq 0}$ denotes the SPX.

The VIX Index

- The VIX is a popular measure of the market's expected volatility of the SPX, calculated and published by the Chicago Board Options Exchange (CBOE). It is given by

$$\mathrm{VIX}_T := \sqrt{\mathbb{E}\left[-\frac{2}{\Delta}\log\left(\frac{S_{T+\Delta}}{S_T}\right)|\mathcal{F}_T\right]},$$

  where $\Delta = 30$ days and $S = (S_t)_{t\geq 0}$ denotes the SPX.

- Recall that we consider the following model

$$\mathrm{d}S_t = S_t\sigma_t^S\mathrm{d}B_t, \tag{SPX}$$

  where $B = (B_t)_{t\geq 0}$ is a one-dimensional Brownian motion and $\sigma^S = (\sigma_t^S)_{t\geq 0}$ the volatility process. Define additionally the instantanous variance process $V = (\sigma^S)^2$ and suppose that $\mathbb{E}[\int_0^T V_t dt] < \infty$.

As well known, under the above model choice, VIX_T can be expressed as follows.

Lemma

Let $S = (S_t)_{t\geq 0}$ be a price process described by Equation (SPX). Then,

$$VIX_T = \sqrt{\frac{1}{\Delta}\mathbb{E}\left[\int_T^{T+\Delta} V_t\mathrm{d}t|\mathcal{F}_T\right]}.$$

Modeling assumptions

Our modeling choices are as follows:

Assumption (on σ^S)

Fix $n > 0$ and consider

$$\sigma_t^S := \ell_\emptyset + \sum_{0<|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{X}}_t \rangle, \tag{vol}$$

where

- $X = (X^1, \ldots, X^d)$ is a d-dimensional continuous semimartingale and $\widehat{X}$ its time-extension. Denoting $Z = (X, B)$, then the correlation matrix between X and B is given by

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

  for all $i, j = 1, \ldots, d+1$, where $[\cdot, \cdot]$ denotes the quadratic variation.
- $\ell := \{\ell_I \in \mathbb{R} : |I| \leq n\}$ the collection of parameters of the model, i.e., $\ell \in \mathbb{R}^{(d+1)_n}$.

Modeling assumptions

Our modeling choices are as follows:

Assumption (on σ^S)

Fix $n > 0$ and consider

$$\sigma_t^S := \ell_\emptyset + \sum_{0<|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{X}}_t \rangle, \qquad \text{(vol)}$$

where

- $X = (X^1, \dots, X^d)$ is a d-dimensional continuous semimartingale and $\widehat{X}$ its time-extension. Denoting $Z = (X, B)$, then the correlation matrix between X and B is given by

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

  for all $i, j = 1, \dots, d+1$, where $[\cdot, \cdot]$ denotes the quadratic variation.
- $\ell := \{\ell_I \in \mathbb{R} : |I| \leq n\}$ the collection of parameters of the model, i.e., $\ell \in \mathbb{R}^{(d+1)_n}$.

Alternative model formulation

Alternatively the volatility process can also be specified a

$$\sigma_t^S := \ell_\emptyset + \sum_{0<|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{X}}_{t-\delta,t} \rangle,$$

where $\widehat{\mathbb{X}}_{t-\delta,t}$ denotes the signature of $\widehat{X}$ between $t-\delta$ and t for some lag $\delta > 0$. This yields similar tractability features, e.g. in view of explicit formulas of the VIX.

Modeling X as a polynomial diffusion

Assumption (on the underlying semimartingale X)

We assume that X is a d-dimensional polynomial diffusion process, i.e.

$$dX_t = b(X_t)dt + \sqrt{a(X_t)}dW_t, \qquad \text{(poly-process)}$$

where

- b, a are polynomials of order one and two, respectively.
- $W = (W_t)_{t \geq 0}$ is a d-dimensional Brownian motion.

Modeling X as a polynomial diffusion

Assumption (on the underlying semimartingale X)

We assume that X is a d-dimensional polynomial diffusion process, i.e.

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sqrt{a(X_t)}\mathrm{d}W_t, \qquad \text{(poly-process)}$$

where

- b, a are polynomials of order one and two, respectively.
- $W = (W_t)_{t\geq 0}$ is a d-dimensional Brownian motion.

Theorem (C.C., G. Gazzani, J. Möller, S. Svaluto-Ferro)

Let $(X_t)_{t\geq 0}$ be a polynomial diffusion process of form (poly-process). Then the truncated signature $(\widehat{\mathbb{X}}_t^n)_{t\geq 0}$ is also polynomial diffusion process. Moreover, let $G \in \mathbb{R}^{(d+1)_n \times (d+1)_n}$ be the matrix representative of the infinitesimal generator of $(\widehat{\mathbb{X}}_t^n)_{t\geq 0}$. Then for each $t, x \geq 0$ it holds

$$\mathbb{E}[\mathrm{vec}(\widehat{\mathbb{X}}_{t+x}^n)|\mathcal{F}_t] = \exp(xG^\top)\mathrm{vec}(\widehat{\mathbb{X}}_t^n),$$

where $\exp(\cdot)$ *denotes the matrix exponential.*

Modeling X as a polynomial diffusion

Assumption (on the underlying semimartingale X)

We assume that X is a d-dimensional polynomial diffusion process, i.e.

$$dX_t = b(X_t)dt + \sqrt{a(X_t)}dW_t, \qquad \text{(poly-process)}$$

where

- b, a are polynomials of order one and two, respectively.
- $W = (W_t)_{t\geq 0}$ is a d-dimensional Brownian motion.

Theorem (C.C., G. Gazzani, J. Möller, S. Svaluto-Ferro)

Let $(X_t)_{t\geq 0}$ be a polynomial diffusion process of form (poly-process). Then the truncated signature $(\widehat{\mathbb{X}}_t^n)_{t\geq 0}$ is also polynomial diffusion process. Moreover, let $G \in \mathbb{R}^{(d+1)_n \times (d+1)_n}$ be the matrix representative of the infinitesimal generator of $(\widehat{\mathbb{X}}_t^n)_{t\geq 0}$. Then for each $t, x \geq 0$ it holds

$$\mathbb{E}[\mathrm{vec}(\widehat{\mathbb{X}}_{t+x}^n)|\mathcal{F}_t] = \exp(xG^\top)\mathrm{vec}(\widehat{\mathbb{X}}_t^n),$$

where $\exp(\cdot)$ denotes the matrix exponential.

The pool of processes X satisfying (poly-process) is rather large: correlated Brownian motions, geometric ones, Ornstein-Uhlenbeck processes, CIR processes, Jacobi processes, all affine processes,...

Analytic expression for the VIX

Theorem (C.C., G. Gazzani, J. Möller, S. Svaluto-Ferro ('23))

Let S and σ^S satisfy (SPX) *and* (vol) *with X being a polynomial diffusion process. Then,*

$$VIX_T(\ell) = \sqrt{\frac{1}{\Delta}\ell^\top Q(T)\ell}, \qquad \text{(VIX-model)}$$

where, for an injective labeling function $\mathscr{L} : \mathcal{I}^n_{d+1} \to \{1, \dots, (d+1)_n\}$,

$$Q_{\mathscr{L}(I),\mathscr{L}(J)}(T) = \mathbb{E}\big[\langle (e_I \sqcup\!\sqcup e_J) \otimes e_0, \widehat{\mathbb{X}}_{T+\Delta}\rangle | \mathcal{F}_T\big] - \langle (e_I \sqcup\!\sqcup e_J) \otimes e_0, \widehat{\mathbb{X}}_T\rangle$$
$$= \mathrm{vec}((e_I \sqcup\!\sqcup e_J) \otimes e_0)(e^{\Delta G^\top} - \mathrm{Id}) \cdot \mathrm{vec}(\widehat{\mathbb{X}}^{2n+1}_T),$$

with $G \in \mathbb{R}^{(d+1)_{2n+1} \times (d+1)_{2n+1}}$ being the matrix representative of the infinitesimal generator of $\widehat{\mathbb{X}}^{2n+1}_t$.

The matrix $Q(T)$ admits a Cholesky decomposition such that $Q(T) = U(T)U(T)^\top$, whence

$$\mathrm{VIX}_T(\ell) = \frac{1}{\sqrt{\Delta}}\|U_T\ell\|$$

Pricing of VIX options

- Note that VIX options are written on future contracts whose price at time at time $t \in [0, T]$ is given by $F_t^T := \mathbb{E}[\mathrm{VIX}_T | \mathcal{F}_t]$.

Pricing of VIX options

- Note that VIX options are written on future contracts whose price at time at time $t \in [0, T]$ is given by $F_t^T := \mathbb{E}[\mathrm{VIX}_T | \mathcal{F}_t]$.

- Hence, in a calibration, also the futures' prices should be calibrated (see Pacati et al. (2018); Guyon (2020)), as model implied volatilites should be computed with model future prices.

Pricing of VIX options

- Note that VIX options are written on future contracts whose price at time at time $t \in [0, T]$ is given by $F_t^T := \mathbb{E}[\mathrm{VIX}_T | \mathcal{F}_t]$.

- Hence, in a calibration, also the futures' prices should be calibrated (see Pacati et al. (2018); Guyon (2020)), as model implied volatilites should be computed with model future prices.

- Let $\mathcal{T}$ be a set of maturities and $\mathcal{K}$ a set of strikes. We then compute the model prices via

$$\pi_{\mathrm{VIX}}^{\text{model}}(\ell, T, K) := \frac{e^{-rT}}{N_{MC}} \sum_{i=1}^{N_{MC}} (\mathrm{VIX}_T(\ell, \omega_i) - K)^+,$$

$$F_{\mathrm{VIX}}^{\text{model}}(\ell, T) := \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} \mathrm{VIX}_T(\ell, \omega_i),$$

  with $\mathrm{VIX}_T(\ell, \omega) = \frac{1}{\sqrt{\Delta}} \| U_T(\omega) \ell \|$ and where N_{MC} denotes the number of Monte Carlo samples.

- Note that for every ℓ the same samples of $\mathbb{X}_T^{2n+1}(\omega)$ to determine $U_T(\omega)$ can be used.
  $\Rightarrow$ No simulation during the optimization task.
  $\Rightarrow$ Variance reduction by using polynomials of VIX squared.

Calibration task for VIX options

The calibration functional reads as follows,

$$L_{\mathrm{VIX}}(\ell) := \sum_{T,K} \mathcal{L}\left(\pi_{\mathrm{VIX}}^{\text{model}}(\ell, T, K), F_{\mathrm{VIX}}^{\text{model}}(\ell, T), \pi_{\mathrm{VIX}}^{mkt,b,a}(T, K), F_{\mathrm{VIX}}^{mkt}(T)\right)$$

where $\mathcal{L}$ denotes some loss function and $\pi_{\mathrm{VIX}}^{mkt,b,a}(T, K), F_{\mathrm{VIX}}^{mkt}(T)$ are the market's option bid/ask prices and market's futures' prices, respectively.

Calibration task for VIX options

The calibration functional reads as follows,

$$L_{\mathrm{VIX}}(\ell) := \sum_{T,K} \mathcal{L}\left(\pi^{\text{model}}_{\mathrm{VIX}}(\ell, T, K), F^{\text{model}}_{\mathrm{VIX}}(\ell, T), \pi^{mkt,b,a}_{\mathrm{VIX}}(T, K), F^{mkt}_{\mathrm{VIX}}(T)\right)$$

where $\mathcal{L}$ denotes some loss function and $\pi^{mkt,b,a}_{\mathrm{VIX}}(T, K), F^{mkt}_{\mathrm{VIX}}(T)$ are the market's option bid/ask prices and market's futures' prices, respectively.

Our choice:

$$\mathcal{L}\left(\pi^{\text{model}}_{\mathrm{VIX}}(\ell, T, K), F^{\text{model}}_{\mathrm{VIX}}(\ell, T), \pi^{mkt,b,a}_{\mathrm{VIX}}(T, K), F^{mkt}_{\mathrm{VIX}}(T)\right) =$$

$$\frac{1}{(\upsilon^{mkt}(\sigma^{mkt,a} - \sigma^{mkt,b}))^2}\Big(\big(\beta\tilde{1}_{\{\pi\notin[\pi^{mkt,b},\pi^{mkt,a}]\}} + (1-\beta)\big)\left|\pi - (\pi^{mkt,a} + \pi^{mkt,b})/2\right|$$

$$+ \left|\delta^{mkt}(e^{-rT} F^{model}_{\mathrm{VIX}}(\ell, T) - e^{-rT} F^{mkt}_{\mathrm{VIX}}(T))\right|\Big)^2,$$

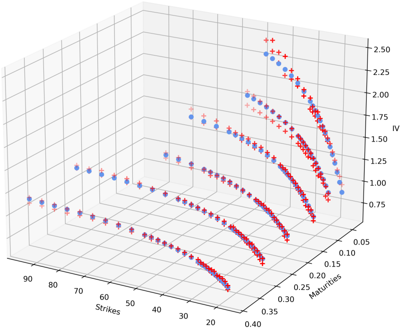
where $\beta \in \{0, 1\}$, $\tilde{1}$ is a smoothed version of the indicator function, δ^{mkt} and υ^{mkt} are the market delta and vega and $\sigma^{mkt,a,b}$ the market implied bid/ask volatilities.

Numerical results I

As a model for X, we use here a two-dimensional OU-process with the following parameters and take the truncation in its signature $n = 3$.

$$\kappa = (0.1, 25)^\top, \qquad \theta = (0.1, 4)^\top, \qquad \sigma = (0.7, 10)^\top, \qquad \rho = \begin{pmatrix} 1 & -0.577 & 0.3 \\ \cdot & 1 & -0.6 \\ \cdot & \cdot & 1 \end{pmatrix}$$

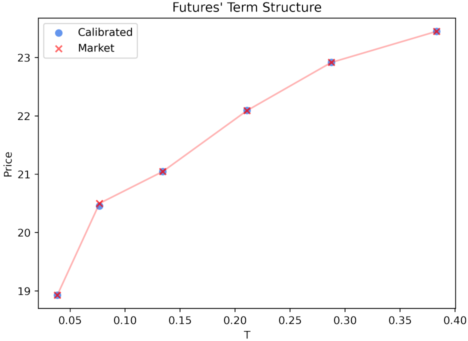
Implied Volatilities VIX 02-06-2021



$T_1 = 0.0383$	$T_2 = 0.0767$	$T_3 = 0.1342$	$T_4 = 0.2108$	$T_5 = 0.2875$	$T_6 = 0.3833$
(90%,250%)	(90%,250%)	(80%,310%)	(80%,300%)	(75%,395%)	(80%,405%)

$T_1 = 0.0383$	$T_2 = 0.0767$	$T_3 = 0.1342$
$\varepsilon_{T_1} = 7.0 \times 10^{-6}$	$\varepsilon_{T_2} = 2.1 \times 10^{-3}$	$\varepsilon_{T_3} = 1.3 \times 10^{-5}$

$T_4 = 0.2108$	$T_5 = 0.2875$	$T_6 = 0.3833$
$\varepsilon_{T_4} = 1.5 \times 10^{-4}$	$\varepsilon_{T_5} = 1.9 \times 10^{-6}$	$\varepsilon_{T_6} = 1.3 \times 10^{-6}$



SPX options and hedging

SPX options with signatures

Signature-based models for the SPX of the form

$$S_n(\ell)_t = \ell_\emptyset + \sum_{0<|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{X}}_t \rangle,$$

can be used to calibrate efficiently to SPX options (see e.g. Perez Arribas et al. (2020), Cuchiero et al. (2022)).

SPX options with signatures

Signature-based models for the SPX of the form

$$S_n(\ell)_t = \ell_\emptyset + \sum_{0<|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{X}}_t \rangle,$$

can be used to calibrate efficiently to SPX options (see e.g. Perez Arribas et al. (2020), Cuchiero et al. (2022)).

Goal: Obtain a tractable (in terms of sampling) form for S satisfying

$$\mathrm{d}S_t(\ell) = S_t(\ell)\sigma_t^S(\ell)\mathrm{d}B_t,$$

where $\sigma_t^S(\ell) = \sum_{|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{X}}_t \rangle$, such that $S(\ell)$ can be expressed in terms of the signature of $\widehat{Z} = (\widehat{X}, B)$, allowing again to precompute all samples and use them for every ℓ.

A signature based model for SPX

Theorem (C.C., G. Gazzani, J. Möller, S. Svaluto-Ferro ('23))

Let $S = (S_t)_{t \geq 0}$ and $\sigma^S = (\sigma_t^S)_{t \geq 0}$ satisfy (SPX) and (vol) with X being a polynomial diffusion process. Then, with $Z = (X, B)$

$$S_t(\ell) = S_0 \exp\left\{-\frac{1}{2}\ell^\top Q^0(t)\ell + \sum_{|I| \leq n} \ell_I \langle \tilde{e}_I^B, \widehat{\mathbb{Z}}_t \rangle\right\}, \qquad \text{(SPX-sig-model)}$$

where $\tilde{e}_\emptyset^B := e_{d+1}$ and $\tilde{e}_I^B$ a transformation of e_I (involving the coefficients of the quadratic variation of X).
The components of the matrix $Q^0(t) \in \mathbb{R}^{(d+1)_n \times (d+1)_n}$ are given, for a labeling function $\mathscr{L} : \mathcal{I} \to \{1, \dots, (d+1)_n\}$, by

$$Q^0_{\mathscr{L}(I), \mathscr{L}(J)}(t) = \langle (e_I \sqcup\!\sqcup e_J) \otimes e_0, \widehat{\mathbb{X}}_t \rangle.$$

A signature based model for SPX

Theorem (C.C., G. Gazzani, J. Möller, S. Svaluto-Ferro ('23))

Let $S = (S_t)_{t\geq 0}$ and $\sigma^S = (\sigma^S_t)_{t\geq 0}$ satisfy (SPX) and (vol) with X being a polynomial diffusion process. Then, with $Z = (X, B)$

$$S_t(\ell) = S_0 \exp\left\{ -\frac{1}{2}\ell^\top Q^0(t)\ell + \sum_{|I|\leq n} \ell_I \langle \tilde{e}_I^B, \widehat{\mathbb{Z}}_t\rangle \right\}, \qquad \text{(SPX-sig-model)}$$

where $\tilde{e}_\emptyset^B := e_{d+1}$ and $\tilde{e}_I^B$ a transformation of e_I (involving the coefficients of the quadratic variation of X).
The components of the matrix $Q^0(t) \in \mathbb{R}^{(d+1)_n \times (d+1)_n}$ are given, for a labeling function $\mathscr{L} : \mathcal{I} \to \{1, \dots, (d+1)_n\}$, by

$$Q^0_{\mathscr{L}(I),\mathscr{L}(J)}(t) = \langle (e_I \sqcup\!\sqcup\, e_J) \otimes e_0, \widehat{\mathbb{X}}_t\rangle.$$

Remark: Note that $\log(S_t)$ can also be rewritten as

$$\mathrm{d}\log(S_t) = -\frac{1}{2}\ell^\top \tilde{Q}(t)\ell \mathrm{d}t + \ell^\top \mathbf{vec}(\widehat{\mathbb{X}}^n_t)\mathrm{d}B_t,$$

where $\tilde{Q}$ is given by $\tilde{Q}_{\mathscr{L}(I)\mathscr{L}(J)}(t) := \langle e_I \sqcup\!\sqcup\, e_J, \widehat{\mathbb{X}}_t\rangle$.

$\Rightarrow$ $(\log S, \widehat{\mathbb{X}}^{2n})$ is a $1 + (d+1)_{2n}$-dimensional polynomial Markov process.

$\Rightarrow$ Path-dependent factor model in spirit of Guyon and Lekeufack (2022).

An affine process point of view

Applying the results from 'Signature SDEs from an affine and polynomial perspective' (see C.C., S.Svaluto-Ferro and J.Teichmann ('23)), here in the case when X is a polynomial diffusion, we obtain:

Theorem (C.C., G. Gazzani, J. Möller, S. Svaluto-Ferro ('23))

The process $(\log(S(\ell)), \widehat{\mathbb{X}})$ *is an* $\mathbb{R} \times T((\mathbb{R}^d))$*-valued affine process.*

An affine process point of view

Applying the results from 'Signature SDEs from an affine and polynomial perspective' (see C.C., S.Svaluto-Ferro and J.Teichmann ('23)), here in the case when X is a polynomial diffusion, we obtain:

Theorem (C.C., G. Gazzani, J. Möller, S. Svaluto-Ferro ('23))

The process $(\log(S(\ell)), \widehat{\mathbb{X}})$ is an $\mathbb{R} \times T((\mathbb{R}^d))$-valued affine process. Suppose that the corresponding $T((\mathbb{R}^d))$-Riccati ODE $\partial_t \psi(t, u, \mathrm{v}) = \mathcal{R}(\psi(t, u, \mathrm{v}))$ has a solution for appropriate $u \in \mathbb{C}$ and $\mathrm{v} \in T((\mathbb{R}^d))$. Then under certain moment conditions it holds

$$\mathbb{E}[\exp(u \log S_T + \langle \mathrm{v}, \widehat{\mathbb{X}}_T \rangle) | \mathcal{F}_t] = \exp(u \log S_t + \langle \psi(T - t, u, \mathrm{v}), \widehat{\mathbb{X}}_t \rangle).$$

An affine process point of view

Applying the results from 'Signature SDEs from an affine and polynomial perspective' (see C.C., S.Svaluto-Ferro and J.Teichmann ('23)), here in the case when X is a polynomial diffusion, we obtain:

Theorem (C.C., G. Gazzani, J. Möller, S. Svaluto-Ferro ('23))

The process $(\log(S(\ell)), \widehat{\mathbb{X}})$ is an $\mathbb{R} \times T((\mathbb{R}^d))$-valued affine process. Suppose that the corresponding $T((\mathbb{R}^d))$-Riccati ODE $\partial_t \psi(t,u,\mathrm{v}) = \mathcal{R}(\psi(t,u,\mathrm{v}))$ has a solution for appropriate $u \in \mathbb{C}$ and $\mathrm{v} \in T((\mathbb{R}^d))$. Then under certain moment conditions it holds

$$\mathbb{E}[\exp(u \log S_T + \langle \mathrm{v}, \widehat{\mathbb{X}}_T \rangle) | \mathcal{F}_t] = \exp(u \log S_t + \langle \psi(T-t, u, \mathrm{v}), \widehat{\mathbb{X}}_t \rangle).$$

In the case (for simplicity) of X being an d-dimensional OU-process of the form

$$dX_t^j = \kappa^j(\theta^j - X_t^j)dt + \sigma^j dW_t^j, \qquad X_0^j = 0, \quad j = 1, \dots, d,$$

with W a d-dimensional Brownian motion independent of B, the function $\mathcal{R}$ reads as

$$\mathcal{R}(u,\mathrm{v}) = \sum_{0 \le |I|,|J| \le n} \frac{1}{2}(u^2 - u)\ell_I \ell_J e_J \sqcup\!\sqcup e_I$$

$$+ \sum_{j=0}^{d} \sum_{|I| \ge 0} \left(\kappa^j \theta^j \mathrm{v}_{(Ij)} e_I + \kappa^j \mathrm{v}_{(Ij)} e_j \sqcup\!\sqcup e_I + \frac{1}{2}(\sigma^j)^2 (\mathrm{v}_{(Ijj)} e_I + \mathrm{v}_{(Ij)}^2 e_I \sqcup\!\sqcup e_I) \right).$$

Towards a hedging formula - expressing $\widehat{\mathbb{X}}$ via the forward variance curve

- Therefore Fourier pricing can be applied and it implies that for any sufficiently integrable payoff of the form $h(S_T, \mathrm{VIX}_T)$ its price at time t is given by some function $H(T-t, S_t, \widehat{\mathbb{X}}_t)$. The price of the SPX call at time t is denoted by $C(T-t, S_t, \widehat{\mathbb{X}}_t)$.
- In order to use this for hedging we express now $\widehat{\mathbb{X}}$ via the forward variance.

Towards a hedging formula - expressing $\widehat{\mathbb{X}}$ via the forward variance curve

- Therefore Fourier pricing can be applied and it implies that for any sufficiently integrable payoff of the form $h(S_T, \mathrm{VIX}_T)$ its price at time t is given by some function $H(T-t, S_t, \widehat{\mathbb{X}}_t)$. The price of the SPX call at time t is denoted by $C(T-t, S_t, \widehat{\mathbb{X}}_t)$.
- In order to use this for hedging we express now $\widehat{\mathbb{X}}$ via the forward variance.
- Note that the forward variance is given by

$$\mathbb{E}[V_u|\mathcal{F}_t] = \ell^\top \widehat{Q}(u-t, \widehat{\mathbb{X}}_t)\ell, \quad u \geq t$$

where $\widehat{Q}_{\mathscr{L}(I),\mathscr{L}(J)}(u-t, \widehat{\mathbb{X}}_t) = \mathbf{vec}(e_I \sqcup\!\sqcup e_J)e^{(u-t)G^\top}\mathbf{vec}(\widehat{\mathbb{X}}_t^{2n})$, whence we have a linear dependence of the forward variance on $\mathbf{vec}(\widehat{\mathbb{X}}^{2n})$.

- From the forward variance curve $(\mathbb{E}[V_u|\mathcal{F}_t])_{u\geq t}$ we can thus get – via linear regression – estimates of $\widehat{\mathbb{X}}^{2n}$. Using e.g. an OLS estimator we get

$$\widehat{\mathbb{X}}_t^{2n} \approx g_t((\mathbb{E}[V_{u_j}|\mathcal{F}_t])_{j\in J}),$$

where for each $t \geq 0$, g_t is a linear function of $(\mathbb{E}[V_{u_j}|\mathcal{F}_t])_{j\in J}$ for some index set J.

A hedging formula

Inspired by Rosenbaum and El Euch (2018) and by expressing $\widehat{\mathbb{X}}$ as above by the forward variance, the following approximate hedging formula becomes feasible by hedging with the asset S and the forward variances $(\mathbb{E}[V_{r+x_j}|\mathcal{F}_r])_{j\in J}$.

Corollary

Under the above assumptions, the following approximate hedging formula for the SPX call holds:

$$C(T-t, S_t, \widehat{\mathbb{X}}_t) \approx \int_0^t \partial_S C^{2n}(T-r, S_r, g_r((\mathbb{E}[V_{r+x_j}|\mathcal{F}_r])_{j\in J}))dS_r + \int_0^t \sum_{j\in J} \partial_j C^{2n}(T-r, S_r, g_r((\mathbb{E}[V_{r+x_j}|\mathcal{F}_r])_{j\in J}))d\mathbb{E}[V_{r+x_j}|\mathcal{F}_r],$$

where

- *C^{2n} denotes the function C obtained via Fourier pricing when $\widehat{\mathbb{X}}$ is restricted to $\widehat{\mathbb{X}}^{2n}$ (which means to truncate the characteristics exponent ψ accordingly);*
- *the corresponding derivatives can be explicitly computed from the Fourier prices C^{2n};*
- *$d\mathbb{E}[V_s|\mathcal{F}_r]$ denotes the Itô differential at time r of the martingale $M_r = \mathbb{E}[V_s|\mathcal{F}_r]$, $r \leq s$.*

Joint calibration of SPX and VIX options

A non exhaustive literature review

- One of the first approaches was via a double CEV model, in Gatheral (2008).
- Jumps literature: Sepp (2012); Papanicolaou and Sircar (2014); Baldeaux and Badran (2014); Pacati et al. (2018).
- Optimal transport: Guo et al. (2020); Guyon (2020, 2021); Guyon and Bourgey (2022).
- Within rough volatility models Gatheral et al. (2020); Rømer (2022); Jacquier et al. (2021); Abi Jaber et al. (2022a) and more recently Bondi et al. (2022) with additional jumps.
- Machine learning techniques for rough volatility and neural SDEs: Rosenbaum and Zhang (2021); Rømer (2022); Guyon and Mustapha (2022).
- Gaussian polynomial volatility models: Abi Jaber et al. (2022a,b).
- Multi-factor (rough) models: Rømer (2022).
- Path-dependent models: Guyon and Lekeufack (2022).

Joint calibration of SPX and VIX options

- Denote by $L_{\mathrm{SPX}}(\ell)$ the SPX calibration functional, where

$$L_{\mathrm{SPX}}(\ell) := \sum_{T \in \mathcal{T}, K \in \mathcal{K}} \mathcal{L}(\pi^{\mathrm{model}}_{\mathrm{SPX}}(\ell, T, K), \pi^{mkt,b}_{\mathrm{SPX}}(T, K), \pi^{mkt,a}_{\mathrm{SPX}}(T, K)).$$

- Then, in order to achieve a joint calibration of the SPX/VIX options and VIX futures, we have to minimize for $\lambda \in (0, 1)$

$$L_{\mathrm{joint}}(\ell, \lambda) := \lambda L_{\mathrm{SPX}}(\ell) + (1 - \lambda) L_{\mathrm{VIX}}(\ell).$$

- Maturity and moneyness specifications:

$T_1^{\mathrm{VIX}} = 0.0383$	$T_2^{\mathrm{VIX}} = 0.0767$
(90%,220%)	(90%,220%)

$T_1^{\mathrm{SPX}} = 0.0383$	$T_2^{\mathrm{SPX}} = 0.1205$	$T_3^{\mathrm{SPX}} = 0.1588$
(92%,105%)	(70%,105%)	(80%,120%)

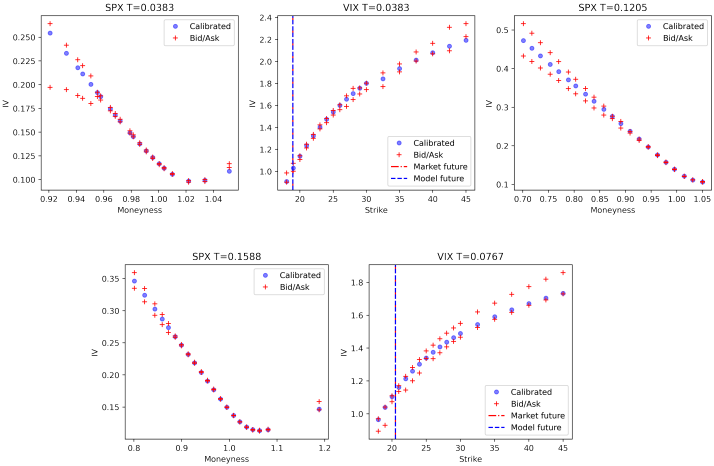
- As a model for X, we here us a three dimensional OU-process with the following parameters and take the truncation in its signature $n = 3$.

$$\kappa = (0.1, 25, 10)^{\top}, \qquad \theta = (0.1, 4, 0.08)^{\top}, \qquad \sigma = (0.7, 10, 5)^{\top},$$

$$\rho = \begin{pmatrix} 1 & 0.213 & -0.576 & 0.329 \\ \cdot & 1 & -0.044 & -0.549 \\ \cdot & \cdot & 1 & -0.539 \\ \cdot & \cdot & \cdot & 1 \end{pmatrix}, \qquad X_0 = (1, 0.08, 2)^{\top},$$

Numerical results II

We calibrate to the same day as in Guyon and Lekeufack (2022).



In blue the calibrated implied volatilities smiles from top-left at maturities T_1^{SPX}, T_1^{VIX}, T_2^{SPX}, T_3^{SPX}, T_2^{VIX}. In red the corresponding bid-ask spreads.

Simulation of time-series of SPX and VIX

- Let $\ell^\star \in \mathbb{R}^{85}$ be the parameters calibrated to SPX and VIX options.
- We sample trajectories for $(V_t(\ell^\star))_{t\in[0,T]}$, $(\mathrm{VIX}_t(\ell^\star))_{t\in[0,T]}$, $(S_t(\ell^\star))_{t\in[0,T]}$.

Simulation of time-series of SPX and VIX

- Let $\ell^\star \in \mathbb{R}^{85}$ be the parameters calibrated to SPX and VIX options.
- We sample trajectories for $(V_t(\ell^\star))_{t\in[0,T]}$, $(\mathrm{VIX}_t(\ell^\star))_{t\in[0,T]}$, $(S_t(\ell^\star))_{t\in[0,T]}$.
- Even though ℓ^* was only calibrated to option prices, the trajectories are economically reasonable and also in line with several stylized facts, such as negative correlation between SPX and VIX or volatility clustering.

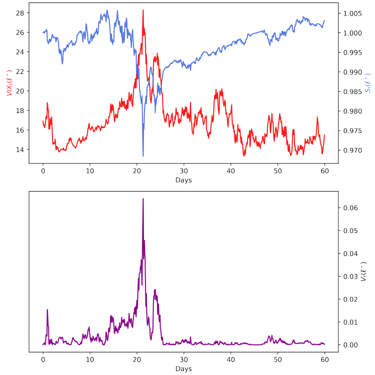


Figure: On the top: one realization of the calibrated model $S(\ell^\star)$ for the SPX (in blue) and the corresponding calibrated VIX (in red). On the bottom: the corresponding realization of the calibrated variance process $V(\ell^\star)$.

Numerical results III

Using time varying parameters we can first jointly calibrate the maturities T_1^{SPX}, T_1^{VIX}, T_2^{SPX} and then add the following new maturities and strikes:

$T_2^{\text{VIX}} = 0.1342$	$T_3^{\text{VIX}} = 0.2875$	$T_4^{\text{VIX}} = 0.3833$
(90%,330%)	(78%,395%)	(80%,405%)

$T_3^{\text{SPX}} = 0.2163$	$T_4^{\text{SPX}} = 0.3696$	$T_5^{\text{SPX}} = 0.4654$
(75%,125%)	(60%,135%)	(50%,145%)

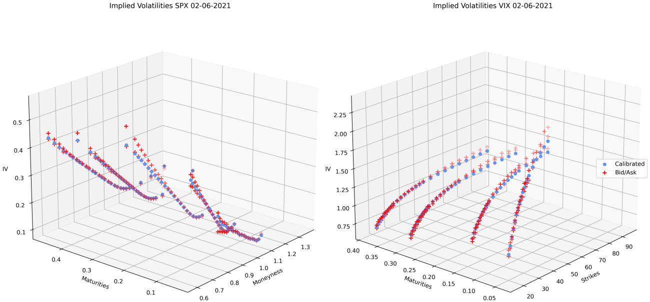


Figure: On the left hand side: SPX smiles, in blue the calibrated implied volatilities and in red the bid-ask spreads. On the right hand side: VIX smiles, in blue the calibrated implied volatilities and in red the bid-ask spreads.

Conclusion

- Universal modeling framework in the class of continuous non-rough stochastic volatility models.

Conclusion

- Universal modeling framework in the class of continuous non-rough stochastic volatility models.
- $\sigma^S(\ell)$ and $\log(S(\ell))$ are expressed as a linear function of the signature of some underlying process $\widehat{Z}$.

Conclusion

- Universal modeling framework in the class of continuous non-rough stochastic volatility models.
- $\sigma^S(\ell)$ and $\log(S(\ell))$ are expressed as a linear function of the signature of some underlying process $\widehat{Z}$.
- Within the polynomial specification VIX can be computed analytically via a matrix exponential.

Conclusion

- Universal modeling framework in the class of continuous non-rough stochastic volatility models.
- $\sigma^S(\ell)$ and $\log(S(\ell))$ are expressed as a linear function of the signature of some underlying process $\widehat{Z}$.
- Within the polynomial specification VIX can be computed analytically via a matrix exponential.
- Viewing $(\log S(\ell), \widehat{\mathbb{X}})$ as affine process allows to obtain approximate hedging formulas.

Conclusion

- Universal modeling framework in the class of continuous non-rough stochastic volatility models.
- $\sigma^S(\ell)$ and $\log(S(\ell))$ are expressed as a linear function of the signature of some underlying process $\widehat{Z}$.
- Within the polynomial specification VIX can be computed analytically via a matrix exponential.
- Viewing $(\log S(\ell), \widehat{\mathbb{X}})$ as affine process allows to obtain approximate hedging formulas.
- A Monte Carlo approach for option pricing, potentially with variance reduction, is tractable since we can compute offline the signature samples of $\widehat{Z}$.

Conclusion

- Universal modeling framework in the class of continuous non-rough stochastic volatility models.
- $\sigma^S(\ell)$ and $\log(S(\ell))$ are expressed as a linear function of the signature of some underlying process $\widehat{Z}$.
- Within the polynomial specification VIX can be computed analytically via a matrix exponential.
- Viewing $(\log S(\ell), \widehat{\mathbb{X}})$ as affine process allows to obtain approximate hedging formulas.
- A Monte Carlo approach for option pricing, potentially with variance reduction, is tractable since we can compute offline the signature samples of $\widehat{Z}$.
- For several short maturities the joint calibration problem can be solved with non-time dependent parameters. By using time-varying parameters we can also fit longer ones.

Conclusion

- Universal modeling framework in the class of continuous non-rough stochastic volatility models.
- $\sigma^S(\ell)$ and $\log(S(\ell))$ are expressed as a linear function of the signature of some underlying process $\widehat{Z}$.
- Within the polynomial specification VIX can be computed analytically via a matrix exponential.
- Viewing $(\log S(\ell), \widehat{\mathbb{X}})$ as affine process allows to obtain approximate hedging formulas.
- A Monte Carlo approach for option pricing, potentially with variance reduction, is tractable since we can compute offline the signature samples of $\widehat{Z}$.
- For several short maturities the joint calibration problem can be solved with non-time dependent parameters. By using time-varying parameters we can also fit longer ones.
- Outlook:
  - Verification of the assumptions to get the affine transform formula and in turn the hedging formula for specific model choices
  - Identification (canoncial form)/ economic meaning of primary process

References I

E. Abi Jaber, C. Illand, and S. Li. Joint SPX-VIX calibration with Gaussian polynomial volatility models: deep pricing with quantization hints. *Available at SSRN 4292544*, 2022a.

E. Abi Jaber, C. Illand, and S. Li. The quintic Ornstein-Uhlenbeck volatility model that jointly calibrates SPX & VIX smiles. *Available at SSRN*, 2022b.

J. Baldeaux and A. Badran. Consistent modelling of VIX and equity derivatives using a 3/2 plus jumps model. *Applied Mathematical Finance*, 21(4):299–312, 2014.

A. Bondi, S. Pulido, and S. Scotti. The rough Hawkes Heston stochastic volatility model. *Preprint arXiv:2210.12393*, 2022.

C. Cuchiero, G. Gazzani, and S. Svaluto-Ferro. Signature-based models: theory and calibration. *Preprint arXiv:2207.13136*, 2022.

J. Gatheral. Consistent modeling of SPX and VIX options. *Bachelier Congress*, 2008.

J. Gatheral, P. Jusselin, and M. Rosenbaum. The quadratic rough Heston model and the joint S&P 500/VIX smile calibration problem. *Preprint arXiv:2001.01789*, 2020.

I. Guo, G. Loeper, J. Obloj, and S. Wang. Joint modelling and calibration of SPX and VIX by optimal transport. *Preprint arXiv:2004.02198*, 2020.

J. Guyon. The joint S&P 500/VIX smile calibration puzzle solved. *Risk, April*, 2020.

J. Guyon. Dispersion-constrained martingale Schrödinger problems and the exact joint S&P 500/VIX smile calibration puzzle. *Available at SSRN 3853237*, 2021.

J. Guyon and F. Bourgey. Fast Exact Joint S&P 500/VIX Smile Calibration in Discrete and Continuous Time. *Available at SSRN 4315084*, 2022.

J. Guyon and J. Lekeufack. Volatility Is (Mostly) Path-Dependent. *Available at SSRN 4174589*, 2022.

References II

J. Guyon and S. Mustapha. Neural Joint S&P 500/VIX Smile Calibration. *Available at SSRN 4309576*, 2022.

A. Jacquier, A. Muguruza, and A. Pannier. Rough multifactor volatility for spx and vix options. *arXiv preprint arXiv:2112.14310*, 2021.

C. Pacati, G. Pompa, and R. Renò. Smiling twice: the Heston++ model. *Journal of Banking & Finance*, 96:185–206, 2018.

A. Papanicolaou and R. Sircar. A regime-switching Heston model for VIX and S&P 500 implied volatilities. *Quantitative Finance*, 14(10):1811–1827, 2014.

I. Perez Arribas, C. Salvi, and L. Szpruch. Sig-SDEs model for quantitative finance. In *Proceedings of the First ACM International Conference on AI in Finance*, pages 1–8, 2020.

S. Rømer. Empirical analysis of rough and classical stochastic volatility models to the SPX and VIX markets. *Quantitative Finance*, pages 1–34, 2022.

M. Rosenbaum and J. Zhang. Deep calibration of the quadratic rough Heston model. *Preprint arXiv:2107.01611*, 2021.

A. Sepp. Achieving consistent modeling of VIX and equities derivatives. *Global Derivatives Conf. Barcelona*, 2012.