arXiv:2406.02319v2 [q-fin.CP] 24 Feb 2025

Pricing and calibration in the 4-factor path-dependent volatility model

Guido Gazzani* Julien Guyon†‡

February 25, 2025

Abstract

We consider the path-dependent volatility (PDV) model of Guyon and Lekeufack (2023), where the instantaneous volatility is a linear combination of a weighted sum of past returns and the square root of a weighted sum of past squared returns. We discuss the influence of an additional parameter that unlocks enough volatility on the upside to reproduce the implied volatility smiles of S&P 500 and VIX options. This PDV model, motivated by empirical studies, comes with computational challenges, especially in relation to VIX options pricing and calibration. We propose an accurate *pathwise* neural network approximation of the VIX which leverages on the Markovianity of the 4-factor version of the model. The VIX is learned pathwise as a function of the Markovian factors and the model parameters. We use this approximation to tackle the joint calibration of S&P 500 and VIX options, quickly sample VIX paths, and price derivatives that jointly depend on S&P 500 and VIX. As an interesting aside, we also show that this *time-homogeneous*, low-parametric, Markovian PDV model is able to fit the whole surface of S&P 500 implied volatilities remarkably well.

Keywords: path-dependent volatility, calibration of financial models, neural networks, S&P 500/VIX joint calibration
MSC (2020) Classification: 91B70, 91G20, 91G30, 91G60, 65C20.

1 Introduction

1.1 Path-dependent volatility models

In this article, we are interested in calibrating the path-dependent volatility (PDV) model of Guyon and Lekeufack (2023) to S&P 500 (SPX) and VIX futures and options, in quickly sampling VIX paths, and in pricing derivatives that may jointly depend on S&P 500 and

*University of Verona, Department of Economics, Via Cantarane 24, 37129 Verona, Italy guido.gazzani@univr.it.
†CERMICS, ENPC, Institut polytechnique de Paris, Marne-la-Vallée, France, julien.guyon@enpc.fr.
‡NYU Tandon School of Engineering, Department of Finance and Risk Engineering, One MetroTech Center, Brooklyn, NY 11201, USA, julien.guyon@nyu.edu.
The authors acknowledge financial support from the BNP Paribas Chair *Futures of Quantitative Finance*. The first author acknowledges financial support by the "PHC AMADEUS" program (project number: 47561RJ), funded by the French Ministry for Europe and Foreign Affairs, the French Ministry for Higher Education and Research, and the Austrian Ministry for Higher Education. The present work was initiated when the first author was affiliated to CERMICS, ENPC.

VIX. PDV models aim to provide an explicit and accurate description of the purely endogenous part of the joint dynamics of an asset price and its volatility ("spot-vol dynamics"): the volatility is described as a deterministic, explicit function of past asset returns. This feedback effect from past returns to volatility, which in turn impacts future returns, creates very rich, intricate, nonlinear dynamics of asset prices.

PDV models have recently received increased attention for the following reasons:

- Empirical data confirms the path-dependent nature of volatility, see for instance Zumbach (2009, 2010); Chicheportiche and Bouchaud (2014); Blanc et al. (2017); Guyon and Lekeufack (2023); Andrès et al. (2023), as well as the ARCH-GARCH literature.
- PDV models can naturally reproduce all the following important stylized facts about volatility: leverage effect, volatility clustering, heavy distribution tails, roughness at the daily scale, weak and strong Zumbach effects, jump-like behavior with very fast large volatility spikes followed by slower decays.

Like local volatility models, PDV models are complete models, but they can generate much more realistic joint spot-vol dynamics, as they capture the path-dependent nature of volatility. Of course, volatility is not purely path-dependent: some unanticipated, exogenous news have an impact on the volatility of financial markets. Stochastic volatility (SV) models start from the exogenous part, postulate a dynamics for the instantaneous volatility which typically depends on unobservable factors, such as Ornstein-Uhlenbeck factors, and can only generate some implicit, complicated path-dependency by correlating the Brownian motions that drive the dynamics of the asset price with those that drive the dynamics of the SV. Of course, one can build a fully PDV model from an SV model by fully correlating those Brownian motions; see for instance the quadratic rough Heston model by Gatheral et al. (2020) and the path-dependent two-factor Bergomi model in Guyon (2022a) where it is shown that the joint SPX/VIX smile calibration of the two-factor Bergomi model degenerates the model into its path-dependent version. However, PDV is not about fully correlating the Brownian motions of an underlying SV model. The PDV paradigm is different. It suggests another, more natural approach to volatility modeling:

(i) First, model the (large) purely endogenous part of volatility *explicitly* as well as possible, using PDV models that depend on *observable* factors, such as past asset returns.

(ii) Then, model the (smaller) exogenous part, for example by analyzing empirical residuals of the (pure) PDV model.

We call the resulting model a *path-dependent stochastic volatility* (PDSV) model. The class of explicit PDV models includes for instance: many contributions in the ARCH-GARCH literature, e.g., Sentana (1995), the complete model of Hobson and Rogers (1998) also discussed by Foschi and Pascucci (2005), the diagonal QARCH model of Chicheportiche and Bouchaud (2014), the ZHawkes process of Blanc et al. (2017), the PDV model of Guyon and Lekeufack (2023), as well as some version of the EWMA Heston of Parent (2023). In particular, the PDV model of Guyon and Lekeufack (2023) was shown to have a better predictive power of both implied volatility and future realized volatility, across equity indexes and train/test periods, than the best competing models available in the literature.

In the continuous-time version of the PDV model of Guyon and Lekeufack (2023), the instantaneous volatility σ_t takes a very simple form: it is an affine combination of a weighted

sum R_1 of past returns, and the square root of a weighted sum R_2 of past squared returns:

$$\sigma_t = \beta_0 + \beta_1 R_{1,t} + \beta_2 \sqrt{R_{2,t}}.$$

This particular, simple specification of a PDV came as the result of a thorough empirical study where more complicated features and functional forms (in particular, neural networks, including LSTM networks) were tested.

One distinctive feature of this PDV model is that it is homogeneous in volatility, in particular capturing the dependency on past realized volatility through the term $\sqrt{R_2}$. The other PDV models that have been studied in the literature are either homogeneous in variance, or mix terms homogeneous to a volatility with terms homogeneous to a variance, and/or lack the important R_2 feature.

This distinctive feature may explain the higher predictive power of the model. However, it also means that the model comes with computational challenges: unlike other models which are chosen because of their mathematical tractability, such as affine or polynomial models, it does not seem to have a nice mathematical structure allowing for fast pricing. In particular, in this PDV model, the forward instantaneous variances and the VIX are not directly accessible, since the instantaneous variance σ_t^2 contains $R_1\sqrt{R_2}$ and $\sqrt{R_2}$ terms whose conditional expectations cannot be exactly computed.

The main objective of this article is to address those computational challenges and compute the VIX in the model. We propose an accurate *pathwise* neural network approximation of the VIX which leverages on the Markovianity of the 4-factor version of the model. The VIX is learned pathwise as a function of the Markovian factors and the model parameters. This approximation is then used to quickly simulate VIX paths, jointly calibrate the model to SPX and VIX options, and quickly price derivatives that jointly depend on the VIX.

1.2 The joint calibration problem

Jointly calibrating to SPX and VIX futures and options is important to prevent arbitrage and ensure accurate pricing of liquid hedging instruments. Indeed, while calibrating to SPX options means incorporating market information on SPX spot implied volatilities, calibrating to VIX derivatives means incorporating market information on *future* SPX implied volatilities, an information that is not contained in SPX option prices. For instance, Bourgey and Guyon (2024) show the additional information contributed by VIX futures and VIX smiles, by quantifying how model-free bounds for SPX path-dependent payoffs tighten when the prices of VIX futures and VIX options are included. They also show that, to avoid mispricing some payoffs (particularly forward-starting payoffs, which are most sensitive to forward volatilities), it is important that the model not only fits SPX options but also fits VIX options, even when the payoff depends only on SPX prices.

In the rest of this section we review some of the literature on the joint calibration problem. We refer the reader to Di Nunno et al. (2023) for a recent survey on volatility modeling and advances on the joint calibration problem. We split the literature in two main streams of research: exact (nonparametric) and approximate (parametric) joint fits.

The first exact joint fit was obtained in Guyon (2020b), where a jointly calibrated nonparametric discrete-time model was built by minimum entropy, i.e., by solving a Schrödinger problem; see also Guyon (2024). Efficient algorithms for this problem (in particular, faster than the Sinkhorn algorithm) have recently been presented in Bourgey and Guyon (2024), who also extended the model to continuous time by martingale interpolation. A direct

continuous-time Schrödinger bridge approach, inspired by Henry-Labordère (2019), is presented in Guyon (2022b). A similar direct nonlinear optimal transport approach was suggested in Guo et al. (2022). The direct continuous-time approaches are much more computationally demanding than the novel discrete-time-continuous-time procedure of Bourgey and Guyon (2024). Note that an exact joint calibration ensures the absence of joint SPX/VIX arbitrages.

The other line of research concerns approximate fits of parametric models. The PDV model of Guyon and Lekeufack (2023) falls into this category. The benefit of low-parametric models is their interpretability. However, they cannot fit market smiles exactly, since the low-parametric structure induces rigidities. The main objective of a modeler is therefore to design parametric models that are able to accurately fit market SPX and VIX smiles. As explained in Guyon (2020a), the main challenge is to fit at the same time large short-term SPX at-the-money (ATM) skews with relatively low VIX implied volatilities. In a first attempt, Gatheral (2008) used a double constant elasticity of variance model, which, despite its quite large number of parameters, cannot jointly fit the implied volatilities of SPX and VIX options accurately enough for trading purposes. Later on, several authors have included jumps in the dynamics of the asset price and/or its volatility, e.g., the forward variance model of Cont and Kokholm (2013) described as the exponential of an affine process with Lévy jumps which allows for Fourier pricing, a regime-switching enhancement of the classical Heston model by Papanicolaou and Sircar (2014), the 3/2 model with jumps in the asset price of Baldeaux and Badran (2014), in the volatility (Kokholm and Stisen (2015)), or with co-jumps and idiosyncratic jumps in the volatility (Pacati et al. (2018)).

Continuous-paths models have also been employed to try to solve the joint calibration problem. For instance, in Fouque and Saporito (2018), a Heston model with stochastic vol-of-vol is calibrated, but only for maturities above 4 months when VIX options are less liquid. More recently, Rømer (2022) considered a model where the volatility is driven by two Ornstein-Uhlenbeck (OU) processes using a hyperbolic transformation function. Unfortunately, the study of Rømer (2022) does not address the calibration to VIX futures, which is crucial as VIX futures are the tradable assets on which VIX options are written. Guyon and Lekeufack (2023) show that the 4-factor PDV (4FPDV) model, a low-parametric Markovian model with only 10 parameters, is able to jointly fit SPX smiles, VIX futures, and VIX smiles with good accuracy. Some parametric models use a much large number of parameters, such as signature-based models or neural SDE models. The use of two OU processes inspired the first contribution with signature-based models, see Cuchiero et al. (2023, 2024). These models address the joint calibration problem by training only a subset of model parameters, namely the linear read out map, as is usually done in reservoir computing. The class of Gaussian polynomial processes introduced in Abi Jaber et al. (2022), which can be seen as a special case of signature-based models, lends itself to a quantization pricing technique for SPX and VIX options. Using neural SDEs, Guyon and Mustapha (2023) show that the joint calibration problem can be very accurately solved by a one-factor stochastic local volatility model, provided enough flexibility is allowed in the dynamics.

Rough volatility models, e.g., the quadratic rough Heston model (Gatheral et al. (2020); Rosenbaum and Zhang (2022)) have also been used to tackle the joint calibration problem with moderate success, due to their very parsimonious, low-parametric nature. Rømer (2022) indeed shows that the rough Bergomi model, the rough Heston model, and an extended rough Bergomi model are all outperformed in the joint calibration task by a hyperbolic 2-OU model; see also Abi Jaber and Li (2024) where different Bergomi-type models in this regard have been compared in terms of their calibration performances on SPX options

only. The rough Heston model belongs to the class of affine Volterra processes considered in Abi Jaber et al. (2019); Cuchiero and Teichmann (2019), and thus allows for Fourier pricing after solving the corresponding Riccati equations. This underlying structure is the building block of an extension with jumps (see Bondi et al. (2024a)) that is employed in the context of the joint calibration by Bondi et al. (2024b), where a rough Heston model with Hawkes-type jumps is shown, using Fourier pricing, to solve the joint calibration problem for short maturities with good accuracy.

1.3 Main contributions

The main contributions of this article are the following.

- The joint calibration procedure in Guyon and Lekeufack (2023) was manual, and the optimization depended strongly on an initial guess. In this work we present an automatic, more robust procedure.
- Moreover, the calibration in Guyon and Lekeufack (2023) was very slow, since the VIX was computed by nested Monte Carlo. In order to build a faster calibration procedure, we provide a novel pathwise deep learning approach to simulating the VIX under a given parametric Markov model, in our case the 4FPDV model, to efficiently price VIX futures and VIX options.
- With the addition of a single parameter we generate enough flexibility in the 4FPDV model to solve the joint SPX/VIX calibration problem for short maturities, leveraging on the neural approximation of the VIX.
- We also show that the 4FPDV model calibrates the SPX implied volatility surface very well, over a large range of maturities (up to at least 1 year).
- We provide an analysis of the stability over time of the calibrated parameters.
- Finally, we use our pathwise approximation of the VIX to quickly price light exotics whose payoffs depend both on SPX and VIX in the 4FPDV model.

The remainder of this article is structured as follows. Section 2 is a reminder about the 4FPDV model. The novel pathwise deep learning estimation of the VIX is presented in Section 3. Calibration results are presented in Section 4 for the calibration to the SPX implied volatility surface only and in Section 5 for the joint SPX/VIX calibration. The stability analysis is carried out in Section 6. Finally, pricing examples are reported in Section 7. The market data used in our numerical experiments was provided by BNP Paribas via the Chair *Futures of Quantitative Finance*. The codes for the present work are available at GuidoGazzani-ai/pdv_nn.

2 The model

Throughout this article, $(r_t)_{t\geq 0}$ and $(q_t)_{t\geq 0}$ denote the interest rate and the dividend yield curves, respectively, and are assumed deterministic. The dividend yield includes the repo rate. We recall the 4-factor path-dependent volatility (4FPDV) model of Guyon and Lekeufack (2023). The price $(S_t)_{t\geq 0}$ of the asset is modeled under the unique risk-neutral

probability measure $\mathbb{Q}$ by

$$\mathrm{d}S_t = (r_t - q_t)S_t\,\mathrm{d}t + S_t\sigma_t\,\mathrm{d}W_t, \tag{2.1}$$

where $S_0 > 0$, $\sigma = (\sigma_t)_{t\geq 0}$ is the instantaneous volatility process to be specified, and $W = (W_t)_{t\geq 0}$ is a standard one-dimensional $\mathbb{Q}$-Brownian motion. The volatility is path-dependent and takes the following simple explicit form as a function of past returns:

$$\begin{aligned}
\sigma_t &:= \sigma(R_{1,t}, R_{2,t}),\\
\sigma(R_1, R_2) &:= \beta_0 + \beta_1 R_1 + \beta_2\sqrt{R_2} + \beta_{1,2}R_1^2\mathbf{1}_{\{R_1>0\}},\\
R_{1,t} &:= (1-\theta_1)R_{1,0,t} + \theta_1 R_{1,1,t},\\
R_{2,t} &:= (1-\theta_2)R_{2,0,t} + \theta_2 R_{2,1,t},
\end{aligned} \tag{2.2}$$

where $R_{1,0,t}$, $R_{1,1,t}$, $R_{2,0,t}$, $R_{2,1,t}$ are four path-dependent factors defined by[1]

$$R_{n,p,t} = \int_{-\infty}^{t} \lambda_{n,p}e^{-\lambda_{n,p}(t-u)}\left(\frac{\mathrm{d}S_u}{S_u}\right)^n, \qquad n\in\{1,2\}, \quad p\in\{0,1\}.$$

The two factors $R_{1,p,t}$ are exponentially weighted moving averages of past returns, measuring the recent trend in the asset price, while the two factors $R_{2,p,t}$ are exponentially weighted moving averages of past squared returns, measuring the recent volatility in the asset price, with exponential weights $\lambda_{n,p} > 0$. Note that

$$R_{n,t} = \int_{-\infty}^{t} K_n(t-u)\left(\frac{\mathrm{d}S_u}{S_u}\right)^n, \qquad n\in\{1,2\},$$

is a weighted average of past returns (when $n = 1$) or past squared returns (when $n = 2$), where the convolution kernels

$$K_n(t) := (1-\theta_n)\lambda_{n,0}e^{-\lambda_{n,0}t} + \theta_n\lambda_{n,1}e^{-\lambda_{n,1}t} \tag{2.3}$$

are convex combinations of two exponential kernels. We use the convention that $\lambda_{n,0} \geq \lambda_{n,1}$, only for interpretation purposes, so $\lambda_{n,0}$ codes for the short memory and $\lambda_{n,1}$ for the long memory. This parametrization of the kernels allows us to

(i) mix short memory (large $\lambda_{n,0}$) and long memory (smaller $\lambda_{n,1}$)—in particular, such kernels are well approximated by power laws over a large range of maturities;

(ii) build a Markovian model, since

$$\begin{aligned}
\mathrm{d}R_{1,p,t} &= -\lambda_{1,p}R_{1,p,t}\,\mathrm{d}t + \lambda_{1,p}\sigma(R_{1,t}, R_{2,t})\,\mathrm{d}W_t,\\
\mathrm{d}R_{2,p,t} &= \lambda_{2,p}(\sigma(R_{1,t}, R_{2,t})^2 - R_{2,p,t})\,\mathrm{d}t.
\end{aligned}$$

Nutz and Riveros Valdevenito (2024) have established the strong existence and uniqueness of a solution $(R_{1,0,t}, R_{1,1,t}, R_{2,0,t}, R_{2,1,t})$ of the above SDE, hence the wellposedness of the 4FPDV model, when $\beta_{1,2} = 0$. Andrès and Jourdain (2024) have extended this result to the case of generic kernels. In particular, the instantaneous volatility cannot blow up in finite time.

The parameters have the following interpretation:

[1] By $(\frac{\mathrm{d}S_u}{S_u})^2$ we mean $\sigma_u^2\,\mathrm{d}u$.

- $\lambda_{1,0}$ captures the dependence of R_1 on recent returns; the larger $\lambda_{1,0}$, the more weight is given to recent returns;
- $\lambda_{1,1} < \lambda_{1,0}$ captures the dependence of R_1 on older returns; the smaller $\lambda_{1,1}$, the more weight is given to old returns;
- $\lambda_{2,0}$ captures the dependence of R_2 on recent squared returns; the larger $\lambda_{2,0}$, the more weight is given to recent squared returns;
- $\lambda_{2,1} < \lambda_{2,0}$ captures the dependence of R_2 on older squared returns; the smaller $\lambda_{2,1}$, the more weight is given to old squared returns;
- $\theta_1 \in [0,1]$ (resp., $\theta_2 \in [0,1]$) mixes the dependence on recent and older returns (resp., squared returns) to form the summary variable R_1 (resp., R_2);
- $\beta_0 > 0$ is the baseline instantaneous volatility;
- $\beta_1 < 0$ and $\beta_2 > 0$ are the sensitivities of the volatility to the trend R_1 and the historical volatility $\sqrt{R_2}$, respectively;
- $\beta_{1,2} \geq 0$ is an additional parameter which produces volatility when the asset price trends up—see below.

The role of $\beta_{1,2}$. The original 4FPDV model, which aimed to explain a single global level of implied volatility (e.g., the VIX) using past daily returns, has $\beta_{1,2} = 0$. As explained in Guyon and Lekeufack (2023), allowing $\beta_{1,2}$ to vary did not significantly increase the predictive power of the model. However, the continuous-time limit of the original 4FPDV model, that is, Model (2.2) with $\beta_{1,2} = 0$, generates out-the-money (OTM) call implied volatilities that are too small and do not increase with strike for large strikes. In order for the continuous-time model to fit not only one single global level of implied volatility but full smiles of implied volatility, Guyon and Lekeufack (2023) added the term $\beta_{1,2} R_1^2 \mathbf{1}_{\{R_1>0\}}$, with $\beta_{1,2} \geq 0$. Indeed, compared to the case where $\beta_{1,2} = 0$, this term generates added volatility when the asset price experiences a positive trend. This allows the asset price to take large values with higher probability, and this results in increased risk-neutral density for large asset price values, therefore increased implied volatility for large strikes.

Table 1

$\lambda_{1,0} = 62.11$	$\lambda_{1,1} = 32.25$	$\theta_1 = 0.23$	$\lambda_{2,0} = 9.57$	$\lambda_{2,1} = 3.51$	$\theta_2 = 0.99$
$\beta_0 = 0.026$	$\beta_1 = -0.138$	$\beta_2 = 0.69$			

This is illustrated in Figure 1 (top), where we plot the model SPX smile and risk-neutral distribution at maturity 16 days. We use the model parameters from Table 8 in Guyon and Lekeufack (2023), calibrated to market smiles, except that we vary $\beta_{1,2} \in \{0, 0.05, 0.1, 0.15\}$; the parameters are reported in Table 1. Figure 1 (top left) shows that the extra term $\beta_{1,2} R_1^2 \mathbf{1}_{\{R_1>0\}}$ allows us to reproduce the hockey-stick shape that is typical of market SPX smiles; Figure 1 (top right) illustrates the increased risk-neutral density on the large SPX side. Figure 1 (bottom) shows the model VIX smile with maturity 14 days. The increased VIX future value, denoted by a dashed line, results from the additional volatility.

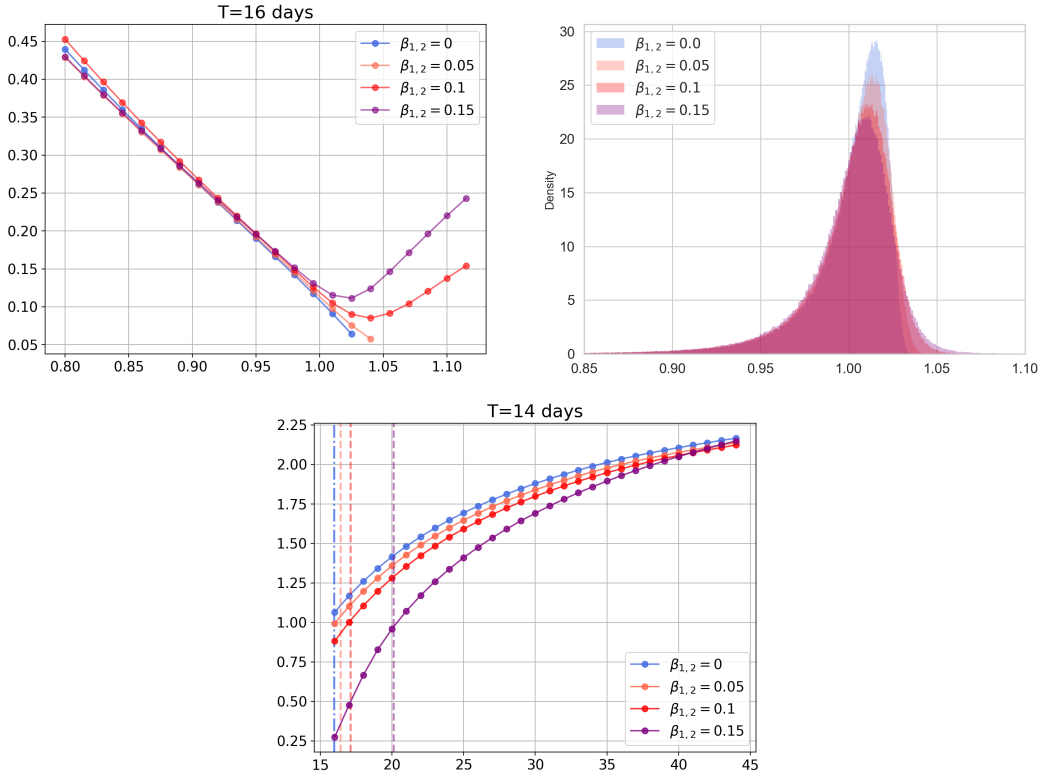


Figure 1: Impact of $\beta_{1,2}$ on model SPX implied volatilities at maturity 16 days (top left), the corresponding model SPX density (top right), and model VIX futures and VIX implied volatilities (bottom). Model parameters are reported in Table 1, except $\beta_{1,2}$ which varies in $\{0, 0.05, 0.1, 0.15\}$. Prices of call options on S are computed by Monte Carlo using $N_{\text{MC}} = 10^5$ trajectories and a discretization step $\Delta t = \frac{1}{2520}$. VIX call option prices are computed using $3 \cdot 10^3$ nested Monte Carlo paths. Initial values of the factors: $R_{1,0,0} = 0.2988$, $R_{1,1,0} = 0.2397$, $R_{2,0,0} = 0.016$, $R_{2,1,0} = 0.02$.

The dynamics of the instantaneous volatility in model (2.2) is given by the following lemma. For ease of notation, we denote $R_t := (R_{1,0,t}, R_{1,1,t}, R_{2,0,t}, R_{2,1,t})$ and introduce the following quantities:

$$\bar{\lambda}_n := (1-\theta_n)\lambda_{n,0} + \theta_n\lambda_{n,1}, \quad \bar{R}_{n,t} := \frac{(1-\theta_n)\lambda_{n,0}R_{n,0,t} + \theta_n\lambda_{n,1}R_{n,1,t}}{\bar{\lambda}_n}, \quad n \in \{1,2\}.$$

Lemma 2.1. Let $\sigma = (\sigma_t)_{t\geq 0}$ satisfy (2.2). Then

$$\mathrm{d}\sigma_t = \mu(R_t)\,\mathrm{d}t + \nu(R_t)\,\mathrm{d}W_t,$$

where

$$\mu(R_t) = -(\beta_1 + 2\beta_{1,2}R_{1,t}\mathbf{1}_{\{R_{1,t}>0\}})\bar{\lambda}_1\bar{R}_{1,t} + \beta_2\bar{\lambda}_2\frac{(\sigma_t^2 - \bar{R}_{2,t})}{2\sqrt{R_{2,t}}} + \beta_{1,2}\bar{\lambda}_1^2\sigma_t^2,$$
$$\nu(R_t) = (\beta_1 + 2\beta_{1,2}R_{1,t}\mathbf{1}_{\{R_{1,t}>0\}})\bar{\lambda}_1\sigma_t.$$

Proof. The proof is a straightforward application of Itô's formula. □

When $\beta_{1,2} = 0$, the instantaneous (lognormal) volatility of the instantaneous volatility is constant (equal to $|\beta_1|\bar{\lambda}_1$), like in Bergomi models. When $\beta_{1,2} > 0$, it is equal to $|(\beta_1 + 2\beta_{1,2}R_{1,t}\mathbf{1}_{\{R_{1,t}>0\}})|\bar{\lambda}_1$ and thus depends on the trend $R_{1,t}$. It vanishes when $R_{1,t} = -\frac{\beta_1}{2\beta_{1,2}} > 0$, and can get large when the asset price trends very positively. In the numerical experiments, to avoid a possible explosion of the instantaneous volatility caused by the superlinear term R_1^2, the instantaneous volatility σ is capped at 1.5.

Remark 2.2. Let us stress that in classical stochastic volatility models, the initial values of the instantaneous volatility or of the latent volatility factors, e.g., Ornstein-Uhlenbeck factors, are not observable and thus have to be calibrated, like model parameters. The same holds for Markovian approximations of rough volatility models, e.g., Rosenbaum and Zhang (2022). By contrast, in pure PDV models, the factors driving the instantaneous volatility are *observable* in the market—in the case of the 4FPDV model, weighted averages of past returns and past squared returns. The model parameters thus directly yield the initial values of the factors, which do not need to be calibrated.

3 Learning the VIX with neural networks

We aim to fastly compute the VIX in the 4FPDV model, pathwise, at any date in the future. This will allow us to quickly price derivatives depending on the VIX, generate VIX paths, and build a computer procedure to jointly calibrate the 4FPDV model to SPX smiles, VIX futures, and VIX smiles. In view of calibrating the model to VIX futures and VIX smiles, we will not only learn VIX_T as a function of the Markovian factors $(R_{1,0,T}, R_{1,1,T}, R_{2,0,T}, R_{2,1,T})$, but also as a function of the model parameters.

3.1 A brief reminder on the VIX

The CBOE Volatility Index, also known as the VIX, is a popular measure of the market's expected volatility of the SPX index over the next 30 days. It is calculated and published by the Chicago Board Options Exchange (CBOE). The VIX is meant to represent the 30-day

implied volatility of the log-contract on the SPX index. We thus take as stylized definition of the VIX

$$\mathrm{VIX}_T^2 := \mathrm{Price}_T\left[-\frac{2}{\Delta}\log\left(\frac{S_{T+\Delta}}{F_T^{T+\Delta}}\right)\right], \tag{3.1}$$

where $\Delta = 30$ days and $F_t^{T+\Delta}$ denotes the price at t of the SPX future with maturity $T+\Delta$. Within a pricing model, we therefore define the VIX by

$$\mathrm{VIX}_T^2 := \mathbb{E}\left[-\frac{2}{\Delta}\log\left(\frac{S_{T+\Delta}}{F_T^{T+\Delta}}\right)\middle|\mathcal{F}_T\right], \tag{3.2}$$

where $\mathbb{E}[\cdot]$ denotes the expectation under the pricing measure $\mathbb{Q}$ and $(\mathcal{F}_t)$ denotes the filtration representing the flow of available information in the market. Note that the nonnegativity of the r.h.s. of (3.1) results from absence of arbitrage, while that of the r.h.s. of (3.2) results from the martingale property of $(F_t^{T+\Delta})_{0\le t\le T+\Delta}$, the concavity of the logarithm, and Jensen's inequality.

In order to compute the VIX in model (2.2), we thus need to compute the conditional expectation (3.2), where $(\mathcal{F}_t)$ is the filtration generated by the Brownian motion W. From Neuberger (1994); Dupire (1994), since the SPX model price has continuous paths,

$$\mathrm{VIX}_T^2 = \mathbb{E}\left[\frac{1}{\Delta}\int_T^{T+\Delta}\sigma_t^2\,\mathrm{d}t\middle|\mathcal{F}_T\right]. \tag{3.3}$$

This equality, which immediately follows from (3.2) by applying the Itô formula to $\log(F_t^{T+\Delta})$ between T and $T+\Delta$, explains the choice (3.1) to represent the expected volatility of the SPX index over the next 30 days. We observe that when σ_t satisfies (2.2), the expression (3.3) is not known in closed form, i.e., one cannot explicitly compute

$$\mathrm{VIX}_T^2 = \frac{1}{\Delta}\mathbb{E}\left[\int_T^{T+\Delta}\left(\beta_0+\beta_1 R_{1,t}+\beta_2\sqrt{R_{2,t}}+\beta_{1,2}R_{1,t}^2\mathbf{1}_{\{R_{1,t}>0\}}\right)^2\mathrm{d}t\middle|\mathcal{F}_T\right]. \tag{3.4}$$

Nevertheless, due to the Markovianity of $R_t = (R_{1,0,t}, R_{1,1,t}, R_{2,0,t}, R_{2,1,t})$ and the time-homogeneity of its dynamics, the VIX squared at time $T > 0$ is a measurable function $f(\Theta, R_T)$, where Θ denotes the collection of the model parameters,

$$\Theta = (\lambda_{1,0}, \lambda_{1,1}, \theta_1, \lambda_{2,0}, \lambda_{2,1}, \theta_2, \beta_0, \beta_1, \beta_2, \beta_{1,2}) \in \mathbb{R}^{10}.$$

Our key idea is to learn the function f by parameterizing it as a feed-forward neural network.

Remark 3.1. An important remark is that we use neural networks to directly learn the VIX at future dates, *pathwise*, as a function of the model parameters and the Markovian factors—not just the prices at time 0 of VIX futures and VIX options as a function of the model parameters. Our use of neural networks thus differs from the so-called "deep pricing" and "deep calibration" approaches, which we briefly recall. Let $\mathcal{T}$ and $\mathcal{K}$ be collections of maturities and strikes, respectively. Using the notation $\phi_{T,K}$ to represent the price of a European vanilla option or its implied volatility with maturity $T \in \mathcal{T}$ and strike $K \in \mathcal{K}$, the "deep pricing" and "deep calibration" methodologies can be summarized in the three following network parametrizations:

- Global deep pricing: $\mathcal{NN} : \Theta \mapsto (\phi_{T,K}^{\mathrm{model}})_{T\in\mathcal{T},K\in\mathcal{K}}$ as in Horvath et al. (2021); Rosenbaum and Zhang (2022); Rømer (2022), for fixed $\mathcal{T}$ and $\mathcal{K}$.

- Pointwise deep pricing: $\mathcal{NN} : (\Theta, T, K) \mapsto \phi^{\text{model}}_{T,K}$ as in Bayer and Stemper (2018); Baschetti et al. (2024).
- Deep calibration: $\mathcal{NN} : (\phi^{\text{model}}_{T,K})_{T\in\mathcal{T},K\in\mathcal{K}} \mapsto \Theta$ as for instance in Hernandez (2016); Gambara and Teichmann (2020) to directly solve the calibration problem.

Those methodologies are most relevant for non-Markovian models, for which Monte Carlo simulations can be very costly. In this article, we learn pathwise quantities (here, the VIX), i.e., we learn at the finest granular level, and our calibration procedure will involve Monte Carlo sampling. One could directly use neural networks to learn the prices at time 0 of VIX futures and VIX options, like in the "deep pricing" and "deep calibration" approaches, without learning the VIX pathwise. However, learning the VIX pathwise as a function of the factors and model parameters is necessary to (1) quickly price derivatives involving the VIX (see Section 7) and (2) qualitatively and quantitatively study the model spot-volatility dynamics, by analyzing the joint behavior of SPX and VIX. Note, moreover, that since the 4FPDV model is Markovian, simulating sample paths in this model is easier and faster than in non-Markovian models, so our Monte Carlo calibration is relatively fast (see Sections 4.2 and 5).

3.2 The learning procedure

The steps of the learning procedure are outlined as follows:

(i) For each of the 10 model parameters Θ_i, sample $N > 0$ independent configurations $\Theta_i(\omega_j)$ accounting for the natural constraint $\Theta_i \in [a_i, b_i]$:

$$\forall\, i \in \{1,\dots,10\},\ \forall\, j \in \{1,\dots,N\}, \qquad \Theta_i(\omega_j) \sim \mathcal{U}[a_i, b_i],$$

where $\mathcal{U}[a,b]$ denotes the uniform distribution over the interval $[a,b]$.

(ii) For each parameter combination $j \in \{1,\dots,N\}$, simulate one random realization of the vector $R_t(\omega_j) := (R^j_{1,0,t}, R^j_{1,1,t}, R^j_{2,0,t}, R^j_{2,1,t})$ in the 4FPDV model (2.2) for a grid of future dates $t \in \{t_1(\omega_j),\dots,t_M(\omega_j)\}$ to be specified later.

(iii) Also store the corresponding $\widehat{\text{VIX}}_t(\omega_j)$ for $t \in \{t_1(\omega_j),\dots,t_M(\omega_j)\}$ and $j \in \{1,\dots,N\}$, where $\widehat{\text{VIX}}$ denotes the square root of the random variable (3.4) estimated using nested Monte Carlo.

(iv) Randomly split the total number of parameter configurations into $N_1 + N_2 = N$. The first N_1 samples consist of the training set, while the next N_2 samples form the validation set.

(v) For a limited number of epochs, train a neural network $\mathcal{NN} : (\Theta, R) \in \mathbb{R}^{14} \mapsto \text{VIX} \in \mathbb{R}^+$ on N_1 while tuning the architecture based on the validation set N_2 using a cross-validation approach, e.g., Random Search, Bayesian Optimization, or Hyperband; see Li et al. (2017). We use as loss function the root mean squared error (RMSE), i.e., the square root of

$$\frac{1}{NM}\sum_{j=1}^{N}\sum_{k=1}^{M}\left(\widehat{\text{VIX}}_{t_k(\omega_j)}(\omega_j) - \mathcal{NN}(\Theta(\omega_j), R_{t_k(\omega_j)}(\omega_j))\right)^2.$$

The inputs of the network are standardized to help the training process.

Remark 3.2. We did not consider the loss function

$$\frac{1}{NM}\sum_{j=1}^{N}\sum_{k=1}^{M}\left(\frac{1}{\Delta}\int_{t_k(\omega_j)}^{t_k(\omega_j)+\Delta}\sigma_t^2(\omega_j)\,\mathrm{d}t-\mathcal{NN}(\Theta(\omega_j),R_{t_k(\omega_j)}(\omega_j))\right)^2$$

(which does not require simulating nested paths) to learn the VIX squared due to the very large conditional variance of the integrated instantaneous variance. Using averages $\widehat{\mathrm{VIX}}_t(\omega_j)$ improves the learning, as analyzed in Alfonsi et al. (2023).

Note that this methodology naturally extends to any parametric Markovian volatility model. The above routine is first used to select the optimal architecture for the neural network. The training further continues to achieve higher accuracy. In the next sections, we report our numerical results.

3.3 Training and testing of the network

We consider $N_1 = 6.8 \cdot 10^5$, $N_2 = 1.2 \cdot 10^5$ configurations of the model parameters to train and validate the network, respectively. Natural constraints yielding enough flexibility for the volatility smiles generated by the 4FPDV model (2.2) are given by

$$\begin{aligned}&\beta_0 \in [0,0.2], \quad \beta_1 \in [-0.25,0), \quad \beta_2 \in [0,1), \quad \beta_{1,2} \in [0,0.3],\\ &\lambda_{n,p} \in [1,100], \quad \theta_n \in [0,1], \qquad n \in \{1,2\}, \quad p \in \{0,1\}.\end{aligned}$$

We require that $\lambda_{n,0} > \lambda_{n,1}$ so that $\lambda_{n,0}$ encodes the short memory and $\lambda_{n,1}$ the long memory. Notice that possible combinations of the above may give rise to extremely large values of the instantaneous volatility of the instantaneous volatility, hence in the simulation of the training set we remove the combinations for which $|\beta_1|\bar{\lambda}_1 > 10$. We observe the four factors $R_t(\omega_j)$ and the VIX, namely $\widehat{\mathrm{VIX}}_t(\omega_j)$, at $M = 200$ evenly spaced time points $t_1(\omega_j) < \cdots < t_M(\omega_j)$, with $t_1(\omega_j) = \max\{1/\lambda_{1,0}(\omega_j), 1/\lambda_{2,0}(\omega_j)\}$ and $t_M(\omega_j) = 1$. With this choice of $t_1(\omega_j)$, the four factors are often observed close to or in their stationary states, as we numerically checked.[2] We compute $\widehat{\mathrm{VIX}}_t$ by nested Monte Carlo: 10^4 nested paths are used to simulate the conditional expectation in (3.4), where the integral is approximated with Riemann sums. Both processes σ and R are sampled with discretization step $\Delta t = \frac{1}{2520}$.

To assess the goodness of the training, we consider two tests. First, we build the histogram of the average absolute error over 10,000 different Monte Carlo samples between the optimal neural network approximation $\mathcal{NN}^\star$ and $\widehat{\mathrm{VIX}}$ for 1,000 random configurations of the model parameters. In Figure 2, we observe that for 99% of the tested configurations, the mean absolute error is smaller than 0.55 (around half a volatility point).

[2] For a stronger guarantee, one could choose $t_1(\omega_j) = \max\{1/\lambda_{1,1}(\omega_j), 1/\lambda_{2,1}(\omega_j)\}$ and use a larger value of $t_M(\omega_j)$. The results of our learning procedure show that our choice of $t_1(\omega_j)$ and $t_M(\omega_j)$ is good enough to sample a large variety of values of R.

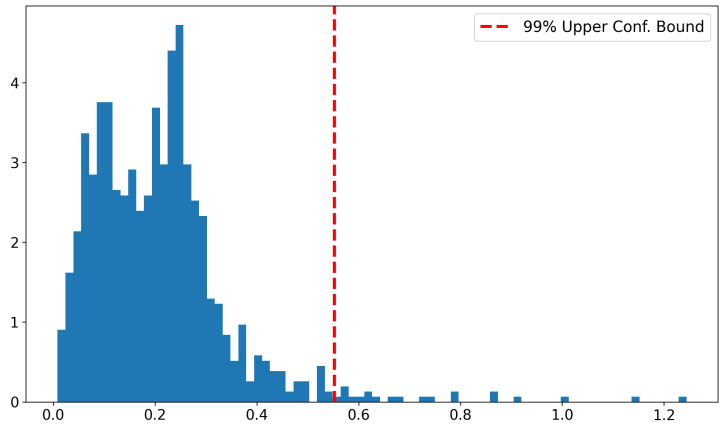


Figure 2: Histogram of the average absolute error between $\mathcal{NN}^\star$ and $\widehat{\text{VIX}}$ for 10^3 configurations of parameters of the model. The average is taken over 10^4 Monte Carlo samples.

Some of the 1,000 configurations of model parameters tested in Figure 2 generate unrealistic VIX smiles and may produce the largest average absolute errors. Therefore, in our second test, we compare the VIX implied volatility smile computed using the neural approximation $\mathcal{NN}^\star$ of the VIX with the one computed using $\widehat{\text{VIX}}$ for a few sets of market-calibrated parameters. Figure 3 shows an excellent agreement between the two VIX smiles and the two VIX futures for one set of realistic parameters.

In Figure 3 we additionally plot the predictions of the neural network against the nested Monte Carlo benchmark as well as the histogram of the absolute error between the two. For this particular set of realistic parameters, the average absolute error is 0.2, which is in the bulk of the distribution of average absolute errors plotted in Figure 2. For 99% of the 10^6 Monte Carlo samples, the (pathwise) absolute error between $\mathcal{NN}^\star$ and $\widehat{\text{VIX}}$ is smaller than 0.65.

In Appendix A, we report similar results for parameters jointly calibrated to the SPX and VIX smiles on June 2, 2021, and June 3, 2021. This informs us that for market-calibrated parameters the optimal neural network accurately learns the VIX computed with nested Monte Carlo and can be used to price VIX futures and VIX options. In any case, when we calibrate model parameters to VIX market data using the $\mathcal{NN}^\star$ approximation of the VIX, we systematically check that the VIX smile computed using $\mathcal{NN}^\star$ and the one computed using the nested estimator $\widehat{\text{VIX}}$ match each other closely (see Figures 10 and 11).

3.4 The architecture of the neural network

Here, we report the architecture of the network obtained after the hyperparameter tuning via Bayesian optimization. We asked the KerasTuner to select from 1 to 5 dense hidden layers for a feed-forward neural network, with a number of nodes ranging from 64 to 512 with 32 as step size, and two possible activation functions, namely

$$\text{ReLU}(x) := \max\{x, 0\}, \quad \tanh(x) := \frac{e^x - e^{-x}}{e^x + e^{-x}}, \qquad \forall x \in \mathbb{R}.$$

We report the optimal values in Table 2. Note that the optimal architecture uses the full 5 hidden layers, with varying numbers of nodes per layer, and that it mixes both activation

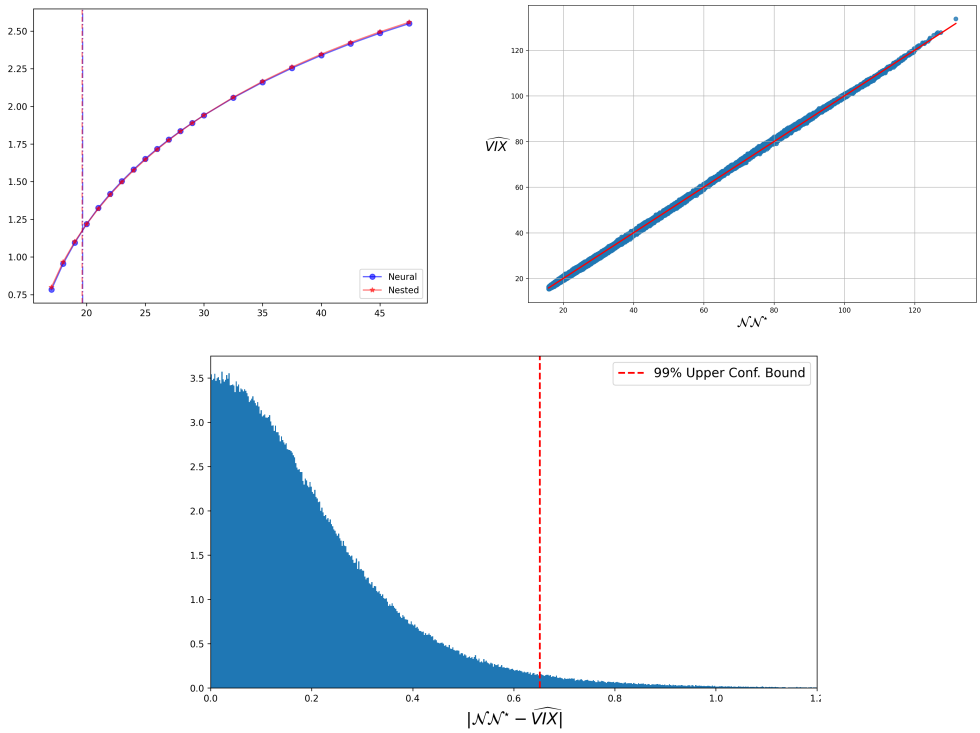


Figure 3: Model parameters: $\lambda_{1,0} = 29$, $\lambda_{1,1} = 20$, $\theta_1 = 0.69$, $\lambda_{2,0} = 81$, $\lambda_{2,1} = 66$, $\theta_2 = 0.25$, $\beta_0 = 0.11$, $\beta_1 = -0.057$, $\beta_2 = 0.1$, $\beta_{1,2} = 0.256$. Top left: VIX smiles and VIX future (dashed) with maturity 16 days. In blue: the VIX is computed via the pre-trained neural network $\mathcal{NN}^\star$. In red: the VIX is computed via nested Monte Carlo, $\widehat{\text{VIX}}$. Top right: A total of 10^6 predictions of $\mathcal{NN}^\star$ vs. $\widehat{\text{VIX}}$. Bottom: histogram of the absolute error between $\mathcal{NN}^\star$ and $\widehat{\text{VIX}}$.

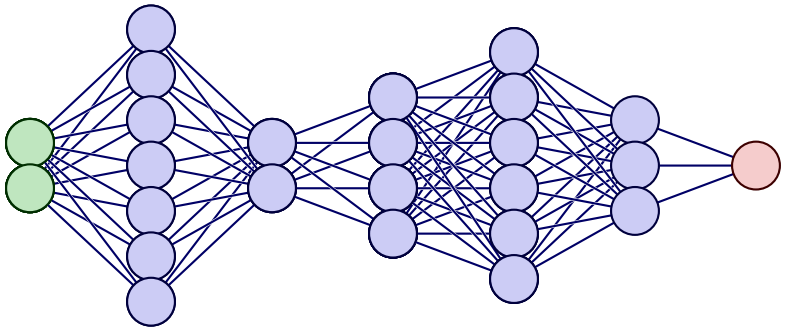


Figure 4: Example of architecture for a feed-forward neural network.

functions. Moreover, we tuned the learning rate of the Adam optimizer between 10^{-5} and 10^{-2} with log-sampling. The optimal learning rate is $4.2 \cdot 10^{-5}$.

Table 2: Architecture of our neural network

Layer type	Activation	No. Nodes
Input	ReLU	14
Hidden 1	tanh	448
Hidden 2	tanh	64
Hidden 3	ReLU	224
Hidden 4	tanh	416
Hidden 5	ReLU	128
Output	linear	1

3.5 Joint SPX/VIX sample paths

Since we learn the VIX in the 4FPDV model *pathwise*, we can very quickly simulate VIX paths in the model along with SPX paths using $\mathcal{NN}^\star$. This is very useful to investigate the dynamical pathwise properties of the model. For instance, in Figure 5, we consider parameters used in Guyon and Lekeufack (2023) to produce realistic paths. We check that VIX spikes correspond to recent drops in the SPX, as observed in the market. Thanks to the R_2 factors (the memory of volatility), the model also reproduces the time-asymmetry of large VIX spikes observed in the market, i.e., the fact that for large VIX spikes, the VIX tend to spike very quickly and decrease more slowly when the market recovers from a large drop. This is due to the fact that both the R_1 and R_2 factors push the volatility up when the SPX is down (a series of negative Brownian increments dW_t is drawn), while they have opposite effects on volatility when the market recovers. In particular, the long memory of volatility factor $R_{2,1}$ tends to keep volatility large for a while, as it remembers large square returns that have happened in the past weeks or months. In Figure 5, we also display a realization of the path of the VIX computed via the nested Monte Carlo, i.e., $\widehat{\text{VIX}}$. As we already knew from the tests reported in Section 3.3, sampling via $\mathcal{NN}^\star$ is very accurate. Moreover, sampling via $\mathcal{NN}^\star$ is of course much faster: it is more than 2,400 times faster than via the nested Monte Carlo with 10^4 inner simulations: approximately 0.09 seconds vs. 4.3 minutes (on a local computer with Core i7).

4 Calibration to SPX options

We now turn to the issue of calibrating the model to option prices. We first consider SPX options only; in Section 5 we will add VIX options, using the neural network approximation built in Section 3.

4.1 Pricing and calibration of SPX options

Suppose we are given a set $\mathcal{T}^{\text{SPX}}$ of SPX option maturities. Then, for each $T \in \mathcal{T}^{\text{SPX}}$, we consider a collection $\mathcal{K}_T^{\text{SPX}}$ of strike prices. Under model (2.1) with parameters Θ, the price

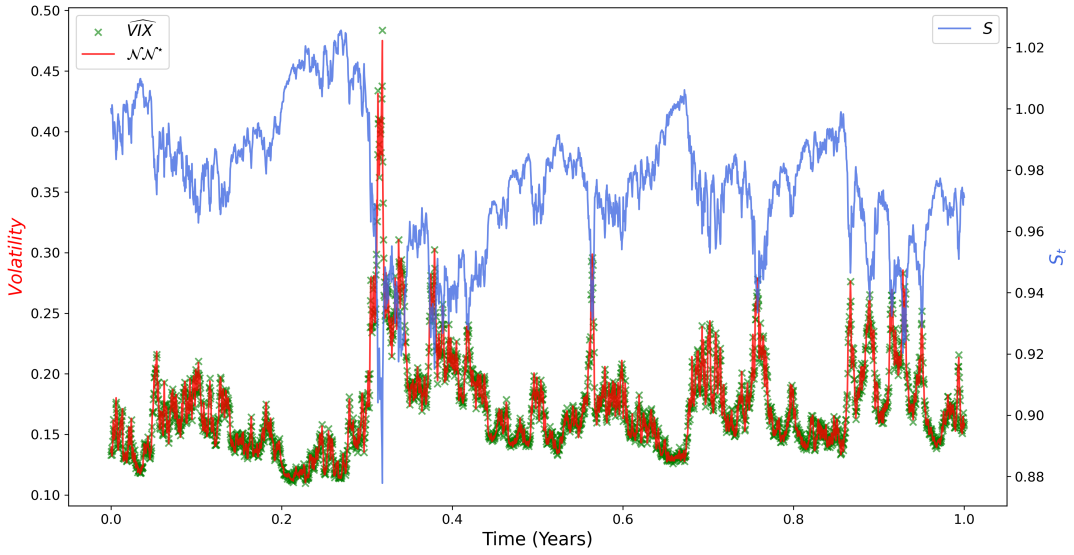


Figure 5: Sample paths of SPX and VIX in the 4FPDV model. In red: approximation of the pre-trained neural network $\mathcal{NN}^\star$; green markers: nested Monte Carlo estimator $\widehat{\text{VIX}}$; in blue: SPX. Model parameters: $\lambda_1 = (55, 10)^\top$, $\lambda_2 = (20, 3)^\top$, $\beta = (0.04, -0.13, 0.65)^\top$, $\beta_{1,2} = 0$, $\theta_1 = 0.25$, $\theta_2 = 0.5$, $S_0 = 1$. Initial values of the Markovian factors: $R_{1,0,0} = 0.078$; $R_{1,1,0} = 0.16$ $R_{2,0,0} = 0.074$, $R_{2,1,0} = 0.016$. Discretization step: $\Delta t = \frac{1}{2520}$.

of an SPX call or put option is approximated by

$$C_\Theta^{\text{SPX}}(T,K) = e^{-\int_0^T r_s\,\mathrm{d}s}\frac{1}{N_{MC}}\sum_{j=1}^{N_{MC}}(S_T(\omega_j)-K)^+, \tag{4.1}$$

$$P_\Theta^{\text{SPX}}(T,K) = e^{-\int_0^T r_s\,\mathrm{d}s}\frac{1}{N_{MC}}\sum_{j=1}^{N_{MC}}(K-S_T(\omega_j))^+, \tag{4.2}$$

for all $T \in \mathcal{T}^{\text{SPX}}$ and for all $K \in \mathcal{K}_T^{\text{SPX}}$ respectively. Denote by $\sigma_{\text{IV},\Theta}^{\text{SPX}}(T,K)$ the model implied volatility under the parameters Θ and for a fixed maturity and strike price, $T, K > 0$ computed for OTM options. Since by construction the model and market SPX futures curve agree we can consider a loss function which aims to measure the discrepancy between either option prices (see, e.g., Cont and Ben Hamida (2005); Papanicolaou and Sircar (2014); Cuchiero et al. (2024)) or implied volatilities (as, e.g., in Pacati et al. (2018); Abi Jaber et al. (2022); Bondi et al. (2024b); Guyon and Mustapha (2023)). In the following we consider the latter approach and introduce the loss function

$$L_{\text{SPX}}(\Theta) = \omega_{\text{SPX}}\frac{1}{\#\mathcal{T}^{\text{SPX}}}\sum_{T\in\mathcal{T}^{\text{SPX}}}\frac{1}{\#\mathcal{K}_T^{\text{SPX}}}\sum_{K\in\mathcal{K}_T^{\text{SPX}}}\ell\left(\sigma_{\text{IV},\Theta}^{\text{SPX}}(T,K),\sigma_{\text{IV}}^{\text{SPX}}(T,K)\right), \tag{4.3}$$

where $\omega_{\text{SPX}} > 0$ is a hyperparameter weight that will be useful when we address the joint SPX/VIX calibration problem and $\ell : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is the score function

$$\ell(x,y) = \left(\frac{x}{y} - 1\right)^2. \tag{4.4}$$

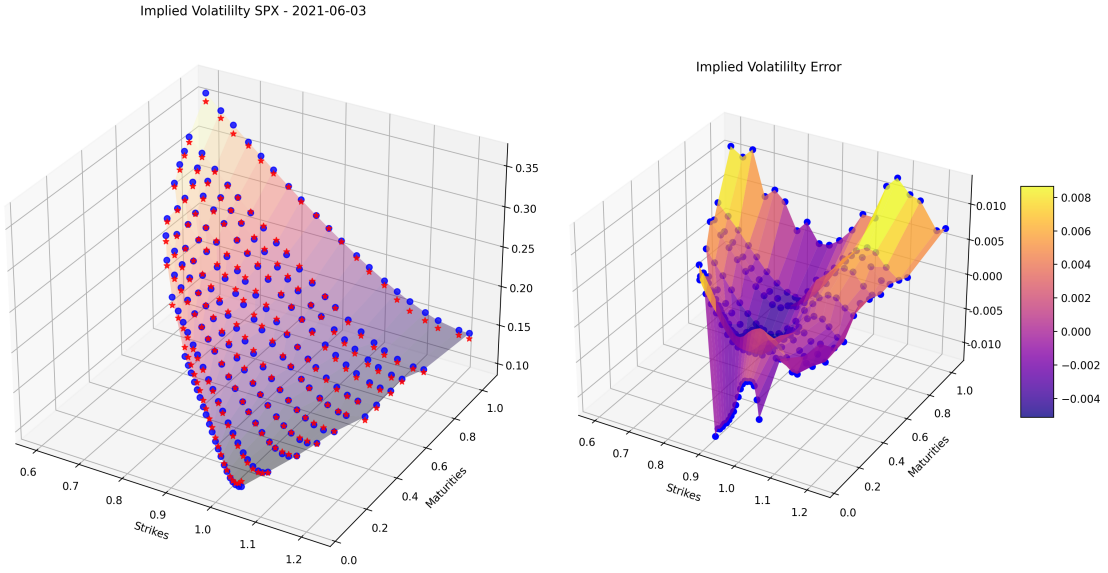


Figure 6: Calibration as of June 3, 2021. Left: in blue dots, the calibrated implied volatility smiles under the PDV model; in red stars, the market mid implied volatility smiles. Right: the interpolated relative error between the calibrated implied volatility smiles and the market mid ones displayed on the left. We report the initial values of the factors given the calibrated parameters in the Table 3. Recall that these are uniquely determined by the calibrated λs and the observed past daily returns for June 3, 2021 with a cut-off at 1000 days: $R_{1,0} = 0.0894$, $R_{1,1} = -0.1602$, $R_{2,0} = 0.0031$, $R_{2,1} = 0.0476$.

4.2 Calibration to the SPX surface

In Figures 6 and 7 we report the results of two calibration exercises: as of June 3, 2021, we calibrate to monthly SPX options with maturities between 15 and 351 days; and as of October 25, 2023, to monthly SPX options with maturities from 23 days to 296 days. At each iteration of the optimizer we simulate $N_{MC} = 2 \cdot 10^5$ trajectories with discretization time step $\Delta t = \frac{1}{504}$. We have carefully checked that this time step is small enough to produce accurate prices and implied volatilities, even in the case of very large mean reversion ($\lambda_{1,0} = \lambda_{1,1} = 100$) and vol-of-vol ($|\bar{\lambda}_1 \beta_1| = 10$). Note that a mean reversion speed of 100 corresponds to a characteristic time of mean reversion of $\frac{1}{100}$, which is five times larger than our time step. The calibrated parameters of the path-dependent volatility model are reported in Tables 3 and 4, respectively. It is remarkable that the model, which is low-parametric (10 parameters) and time-homogeneous (no deterministic time-dependent input curve is used to calibrate to a term-structure of volatilities) is able to fit SPX smiles for a wide range of maturities so accurately. In particular, the two-exponential kernels allow us to accurately capture the market term-structure of the ATM skew (see Figure 8). We observe that two different mean reversion speeds need to be coupled, see the values of the λs and θs in Table 3.

The calibration to the SPX surface up to the one-year maturity takes approximately 30 minutes using Cuda 11.4 with PyTorch on a GPU-NVIDIA-SMI. The optimization was

Table 3: Calibrated parameters of the 4FPDV model. Calibration to the SPX surface as of June 3, 2021.

$\lambda_{1,0} = 34.39$	$\lambda_{1,1} = 13.26$	$\theta_1 = 0.501$	$\lambda_{2,0} = 95.63$	$\lambda_{2,1} = 1.428$	$\theta_2 = 0.448$
$\beta_0 = 0.0493$	$\beta_1 = -0.1999$	$\beta_2 = 0.5479$	$\beta_{1,2} = 0.2285$		

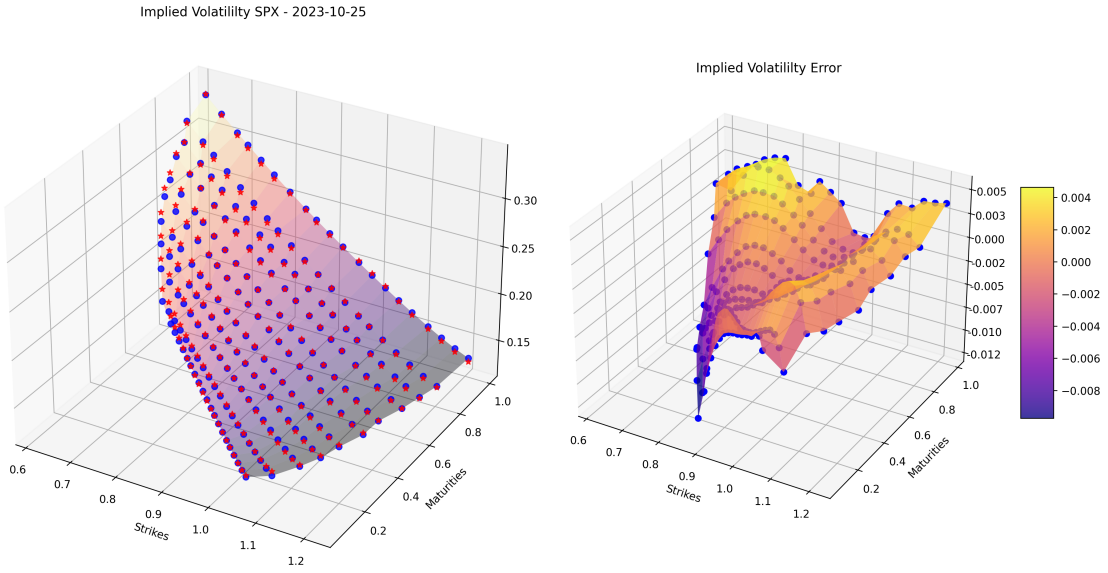
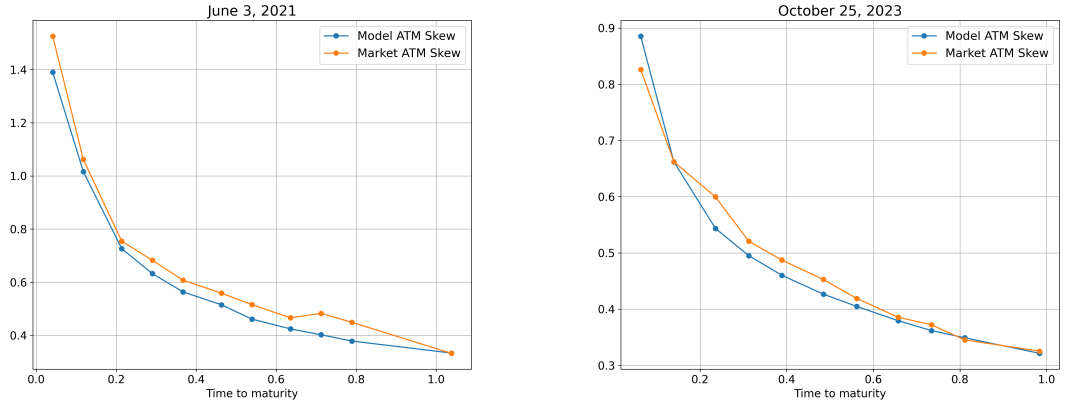


Figure 7: Calibration as of October 25, 2023. See caption of Figure 6 for color reference. The initial values of the factors with a cut-off at 1000 days are given by: $R_{1,0} = -1.2054$, $R_{1,1} = -0.2428$, $R_{2,0} = 0.0178$, $R_{2,1} = 0.0166$.

Table 4: Calibrated parameters of the 4FPDV model. Calibration to the SPX surface as of October 25, 2023.

$\lambda_{1,0} = 53.03$	$\lambda_{1,1} = 6.031$	$\theta_1 = 0.685$	$\lambda_{2,0} = 12.03$	$\lambda_{2,1} = 8.325$	$\theta_2 = 0.2876$
$\beta_0 = 0.0381$	$\beta_1 = -0.1483$	$\beta_2 = 0.7097$	$\beta_{1,2} = 0.1671$		



(a) (b)

Figure 8: Comparison of the market and model term-structures of ATM skew.

carried out, starting from a randomized offline initial configuration, using the derivative-free optimizer Py-BOBYQA (see Cartis et al. (2019, 2022)) relying on a trust-region method which proved particularly efficient with respect to other optimizers. We enforced natural bounds on the parameters during optimization and then scaled model parameters to help the learning.

5 Joint calibration of SPX and VIX options

5.1 Pricing and calibration of VIX futures and VIX options

Pricing. Suppose we are given a set $\mathcal{T}^{\mathrm{VIX}}$ of VIX option maturities. For each $T \in \mathcal{T}^{\mathrm{VIX}}$, we consider a collection $\mathcal{K}_T^{\mathrm{VIX}}$ of strike prices. We use the classical Monte Carlo method together with the neural approximation $\mathcal{NN}^\star$ of the VIX under model (2.2) with parameters Θ built in Section 3. The price of the VIX future and the price of a VIX call are respectively approximated by

$$F_\Theta^{\mathrm{VIX}}(T) = \frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} \mathcal{NN}^\star(\Theta, R_T(\omega_j)),$$

$$C_\Theta^{\mathrm{VIX}}(T,K) = e^{-\int_0^T r_s \,\mathrm{d}s} \frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} (\mathcal{NN}^\star(\Theta, R_T(\omega_j)) - K)^+,$$

for all $T \in \mathcal{T}^{\mathrm{VIX}}$ and for all $K \in \mathcal{K}_T^{\mathrm{VIX}}$.

Calibration. We now turn to the problem of calibrating the model parameters to VIX futures and VIX options. Notice that since the actual tradable underlying of VIX options are the VIX futures with the same expiry as the option, it is crucial that the VIX futures be very well calibrated, as already argued in, e.g., Pacati et al. (2018); Guyon (2020a); Cuchiero et al. (2024). To be consistent, the model VIX implied volatilities are then of course computed by inverting the Black formula using the corresponding model VIX future.

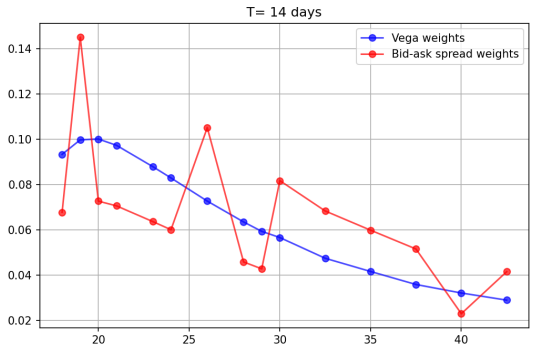


Figure 9: In blue: normalized vega weights; in red: normalized bid-ask spread weights. Data as of June 2, 2021.

Since we have learned the VIX not only as a function of the four Markovian factors but also as a function of the 10 model parameters, Θ, we can use the loss function

$$L_{\mathrm{VIX}}(\Theta) = \omega_F \frac{1}{\#\mathcal{T}^{\mathrm{VIX}}} \sum_{T\in\mathcal{T}^{\mathrm{VIX}}} \ell\left(F_\Theta^{\mathrm{VIX}}(T), F^{\mathrm{VIX}}(T)\right) \tag{5.1}$$
$$+ \omega_{\mathrm{VIX}} \frac{1}{\#\mathcal{T}^{\mathrm{VIX}}} \sum_{T\in\mathcal{T}^{\mathrm{VIX}}} \frac{1}{\#\mathcal{K}_T^{\mathrm{VIX}}} \sum_{K\in\mathcal{K}_T^{\mathrm{VIX}}} \gamma_{T,K}^{\mathrm{VIX}} \ell\left(C_\Theta^{\mathrm{VIX}}(T,K), C^{\mathrm{VIX}}(T,K)\right)$$

to calibrate the model to VIX futures and VIX options, where

- $\omega_F,\ \omega_{\mathrm{VIX}} > 0$ are hyperparameters weights;
- $\#$ denotes the cardinality of a set;
- $(\gamma_{T,K}^{\mathrm{VIX}})_{T,K}$ are normalized vega weights; given $\mathcal{K}(T)$ the strikes available for maturity $T > 0$,
$$\gamma_{T,K} := \frac{\mathcal{V}_{T,K}}{\sum_{K'\in\mathcal{K}(T)} \mathcal{V}_{T,K'}},$$
where $\mathcal{V}_{T,K}$ denotes the Black vega of the call option with maturity T and strike K, computed using the corresponding market mid implied volatility.
- $F^{\mathrm{VIX}}, C^{\mathrm{VIX}}$ denote the market futures and the market call option prices, respectively;
- ℓ is the score function (4.4).

Note that we calibrate the model to VIX option prices, not VIX implied volatilities, as model VIX implied volatilities depend on model VIX futures; this prevents a mismatch between model and market VIX futures to cascade into a mismatch between model and market VIX implied volatilities. The weights $(\gamma_{T,K}^{\mathrm{VIX}})_{T,K}$ are often taken to be the inverse of the market bid-ask spreads, which gives more weight to those options where the bid-ask spread is tighter (see, e.g., Cont and Ben Hamida (2005) for a discussion on this topic). In Figure 9 we compare the normalized vega weights with the inverse of the market bid-ask spreads as of June 2, 2021, for VIX options with 33 days to expiration. The vega weights are a smoothed version of the inverse of the bid-ask spread.

5.2 Joint calibration results

We now report the results of the joint calibration of SPX and VIX monthly options as of June 2, 2021 (Figure 10 and Table 5) and June 3, 2021 (Figure 11 and Table 6). As loss function we take the sum of the two loss functions (4.3) and (5.1). Recall that SPX monthly options expire the third Friday of the expiration month, while the monthly VIX futures and VIX options mature 30 days before the third Friday of the following month.

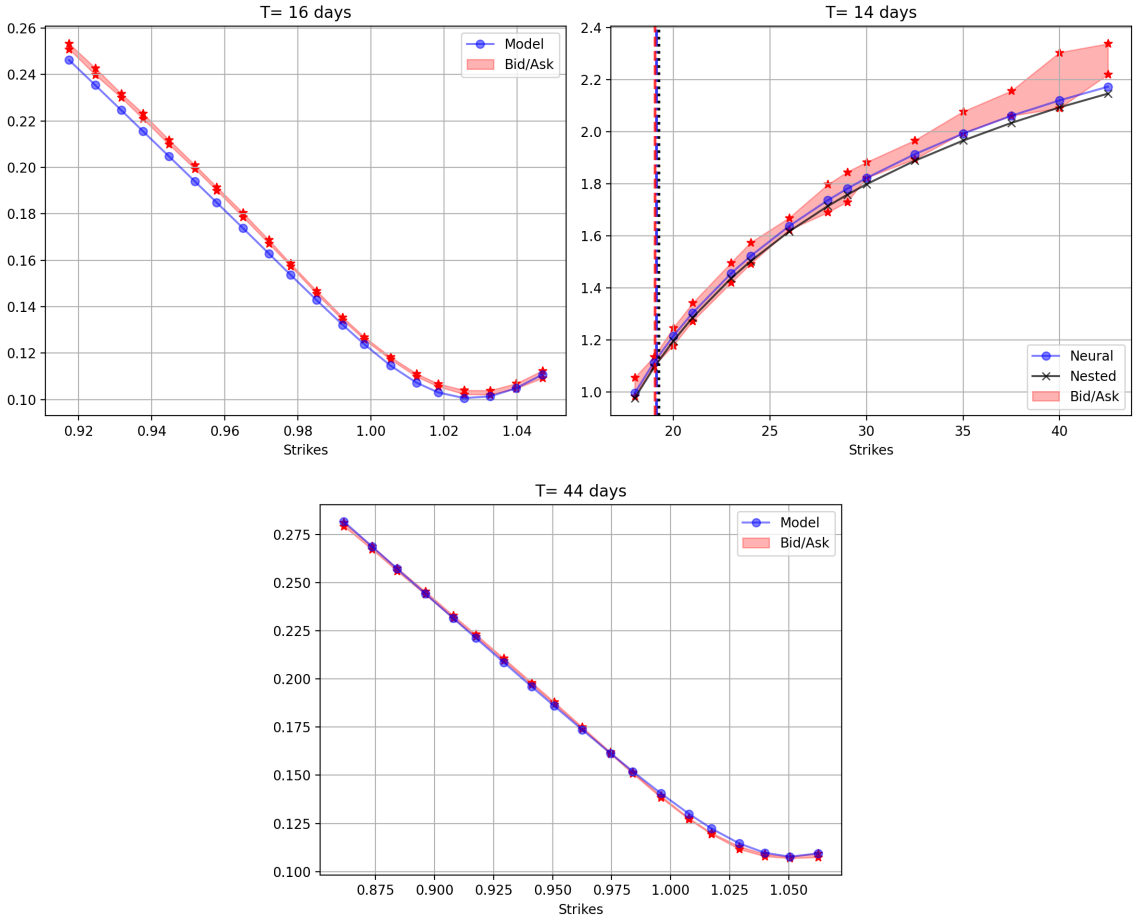


Figure 10: Joint calibration as of June 2, 2021. Comparison of calibrated and market SPX/VIX implied volatility smiles. Top left and bottom: SPX smiles. Top right: the market VIX future is displayed with dotted red lines while the model VIX future is indicated by the dotted blue line; in black (resp., dotted black) the VIX smile (resp., VIX future) computed with nested Monte Carlo. The initial values of the factors given the calibrated parameters in Table 5 are $R_{1,0} = 0.2689$, $R_{1,1} = 0.2375$, $R_{2,0} = 0.0249$, $R_{2,1} = 0.02491$.

Table 5: Calibrated parameters of the 4FPDV model. Joint calibration to the SPX and VIX smiles as of June 2, 2021.

$\lambda_{1,0} = 44.42$	$\lambda_{1,1} = 33.19$	$\theta_1 = 0.398$	$\lambda_{2,0} = 4.311$	$\lambda_{2,1} = 3.254$	$\theta_2 = 0.72$
$\beta_0 = 0.0254$	$\beta_1 = -0.1602$	$\beta_2 = 0.6922$	$\beta_{1,2} = 0.1639$		

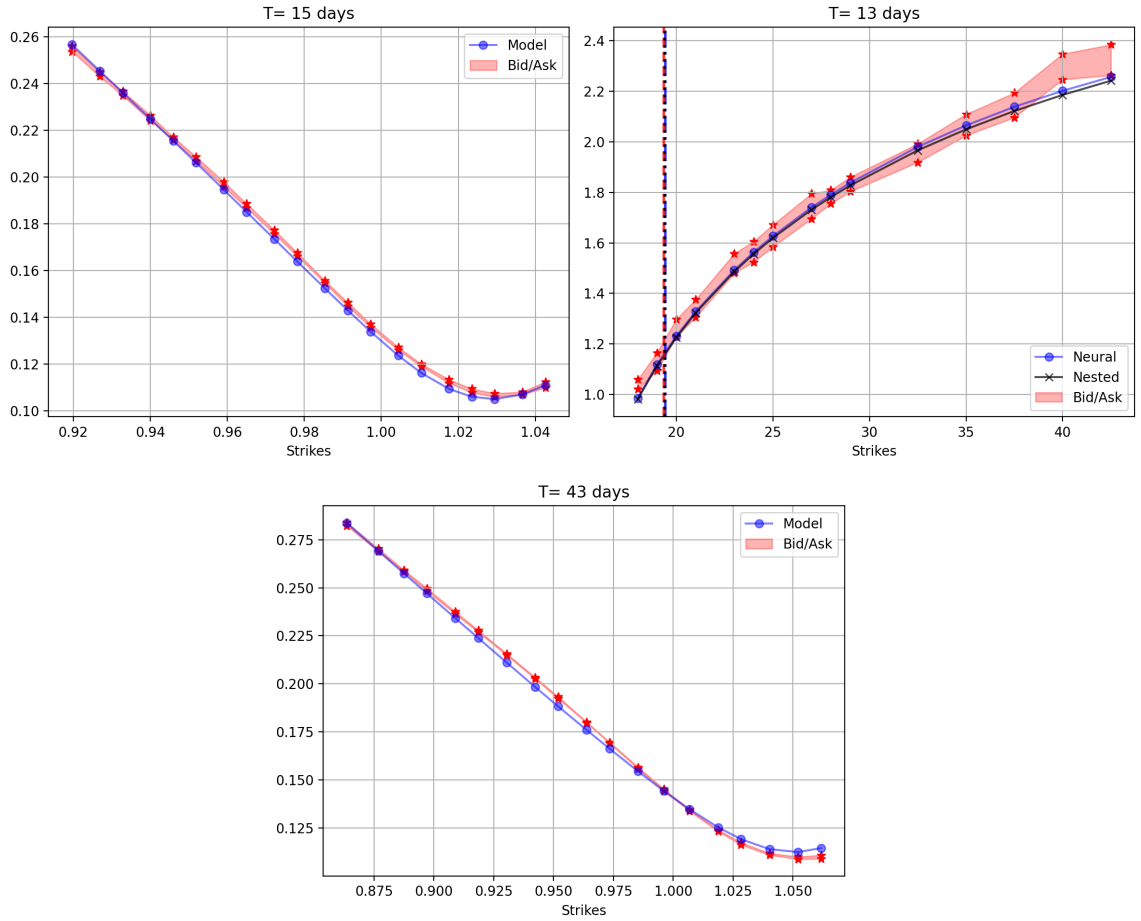


Figure 11: Joint calibration as of June 3, 2021. See Figure 10. The initial values of the factors given the calibrated parameters in Table 6 are $R_{1,0} = 0.0669$, $R_{1,1} = 0.0916$, $R_{2,0} = 0.02197$, $R_{2,1} = 0.02725$.

Table 6: Calibrated parameters of the 4FPDV model. Joint calibration to the SPX and VIX smiles as of June 3, 2021.

$\lambda_{1,0} = 42.78$	$\lambda_{1,1} = 31.51$	$\theta_1 = 0.389$	$\lambda_{2,0} = 3.694$	$\lambda_{2,1} = 3.693$	$\theta_2 = 0.698$

$\beta_0 = 0.0264$	$\beta_1 = -0.1665$	$\beta_2 = 0.6829$	$\beta_{1,2} = 0.1628$

In Figures 10 and 11 we observe that we are able to calibrate jointly the first two monthly smiles of the SPX (top left and bottom) and the first monthly VIX future and VIX smile (top right) with a very good accuracy. Note that for both calibration dates, $\lambda_{1,0} \approx \lambda_{1,1}$ and $\lambda_{2,0} \approx \lambda_{2,1}$, which means that one-exponential kernels seem enough for this fit. This is due to the fact that we only calibrate to short-dated options here. Recall that, by contrast, in Section 4.2, we have seen that two-exponential kernels are needed to jointly fit short-dated and long-dated SPX implied volatilities. For each iteration of the optimizer, we simulated $N_{MC} = 2 \cdot 10^5$ trajectories with discretization step $\Delta t = \frac{1}{504}$. The hyperparameters of the loss functions for SPX and VIX options are fixed to be $(\omega_{\text{SPX}}, \omega_{\text{VIX}}, \omega_F) = (10, 5, 20)$. Like for the calibration to the SPX surface, we use the Py-BOBYQA optimizer and randomize

MAE	
October 23, 2023	$2.23 \cdot 10^{-3}$
October 24, 2023	$2.26 \cdot 10^{-3}$
October 25, 2023	$2.12 \cdot 10^{-3}$
October 26, 2023	$2.63 \cdot 10^{-3}$
October 27, 2023	$2.32 \cdot 10^{-3}$

Table 7: Mean absolute error of the SPX implied volatility fit. The range of moneyness considered for each maturity $T > 0$, is given by $[K_{\min}, K_{\max}]$ where $K_{\min} = 1 - 0.4\sqrt{T}$ and $K_{\max} = 1 + 0.25\sqrt{T}$.

the initial guess within the natural bounds of the parameters. The joint calibration takes around 8 minutes to complete.

In Figures 10 and 11, we do not only report the VIX future and VIX implied volatilities in the calibrated model when we use our neural network approximation $\mathcal{NN}^\star$ of the VIX. We also report those quantities when we use nested Monte Carlo paths to estimate the VIX. Figures 10 and 11 show that for the jointly calibrated parameters our neural network approximation of the VIX is accurate enough for trading purposes.

6 On the stability of the calibrated parameters

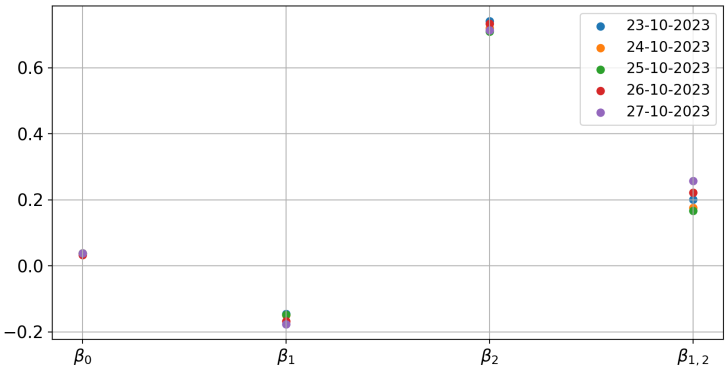
In this section we address the stability of the calibrated parameters for the 4FPDV model. Stability of the calibrated parameters is desirable as, usually, models are periodically re-calibrated for hedging purposes; parameter stability prevents oscillating hedge ratios and higher transaction costs. A similar analysis has been carried out in Cuchiero et al. (2024).

We split our analysis in two parts. In the first part (Section 6.1), we calibrate OTM SPX monthly options every day of the fourth trading week of October 2023. In the second part (Section 6.2), we compare the parameters found for the two joint calibration exercises in Section 5, i.e., for the two consecutive days June 2, 2021 and June 3, 2021. In both cases we display the corresponding kernels of the 4FPDV model as defined in (2.3).

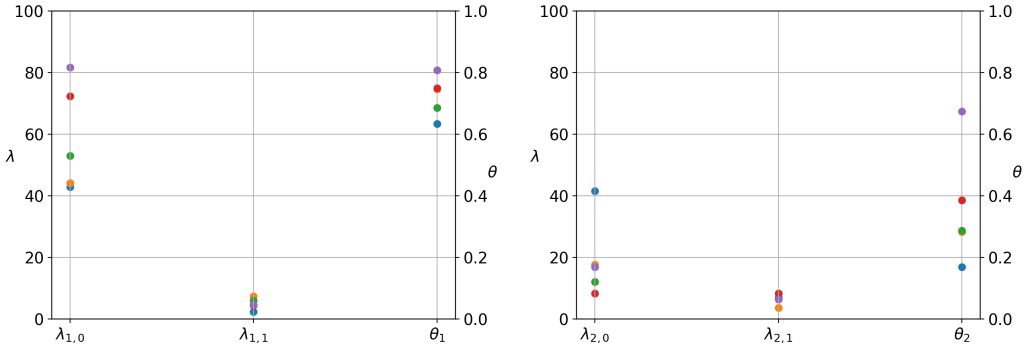
6.1 Calibration to SPX options only

In this section, we calibrate OTM SPX monthly options every day of the fourth trading week of October 2023, that is, October 23 to October 27, 2023. We calibrate to monthly maturities up to one year. Figure 7 shows the result of the calibration as of October 25, 2023. Table 7, in which we report the mean absolute error (MAE) of the implied volatility fit, shows that the fits as of October 23, 24, 26, and 27, not reported here, achieve the same degree of accuracy.

The values of the calibrated parameters on the five consecutive trading days are reported in Figure 12. The β parameters $(\beta_0, \beta_1, \beta_2, \beta_{1,2})$ (Figure 12a) are remarkably stable over time. The kernel parameters $(\lambda_{1,0}, \lambda_{1,1}, \theta_1, \lambda_{2,0}, \lambda_{2,1}, \theta_2)$ (Figure 12b) appear less stable, in particular the K_2 parameters. However, note that different values of $(\lambda_{1,0}, \lambda_{1,1}, \theta_1)$ (resp., $(\lambda_{2,0}, \lambda_{2,1}, \theta_2)$) produce similar two-exponential kernels. This is illustrated in Figure 13, and is mostly observed for K_1.



(a) Sensitivities $(\beta_0, \beta_1, \beta_2, \beta_{1,2})$ of the volatility for five consecutive calibration dates.



(b) Parameters of the kernels K_1 and K_2 for five consecutive calibration dates. L.h.s.: parameters $(\lambda_{1,0}, \lambda_{1,1}, \theta_1)$ of the kernel K_1. R.h.s.: parameters $(\lambda_{2,0}, \lambda_{2,1}, \theta_2)$ of the kernel K_2. For each plot, the y-axis on the l.h.s. refers to the natural scale of the $\lambda_{n,p}$s while the y-axis on the r.h.s. refers to the values of the θ_ps.

Figure 12

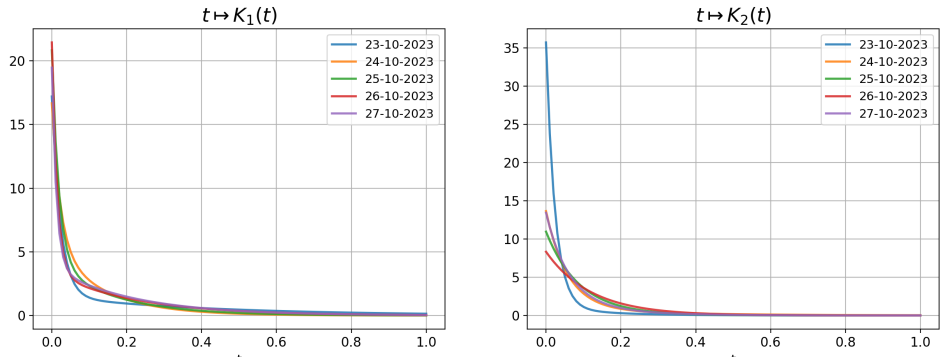
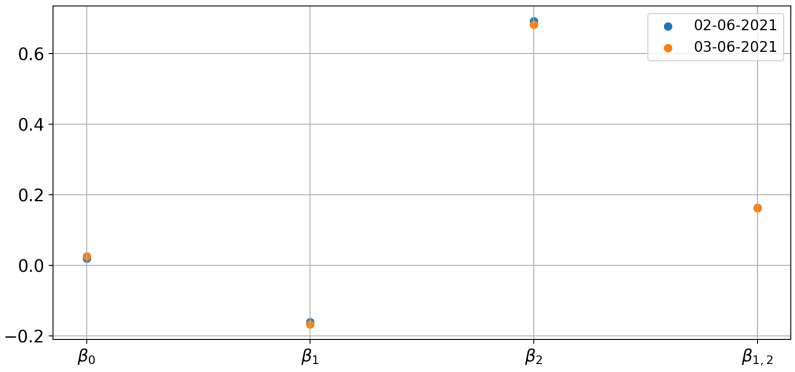


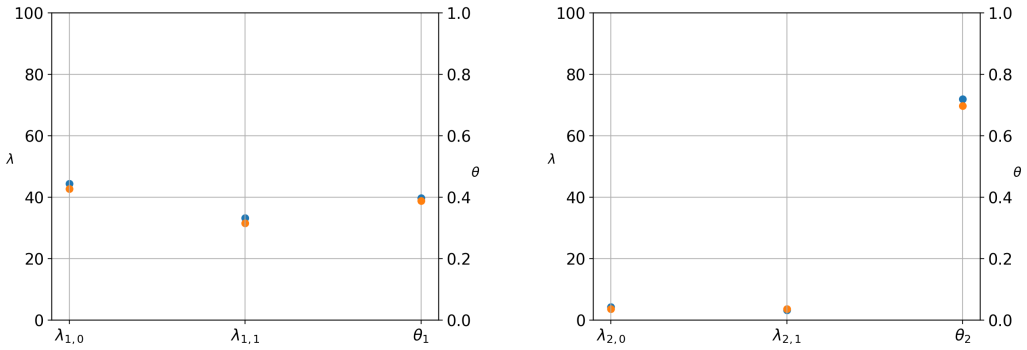
Figure 13: Comparison of the kernels on five consecutive trading days. L.h.s.: K_1. R.h.s.: K_2

6.2 Joint SPX/VIX calibration

In Figure 14, we plot the jointly calibrated parameters of Tables 5 (June 2, 2021) and 6 (June 3, 2021). We observe that the calibrated parameters are very stable from one calibration date to the next, including the kernel parameters. In particular Figure 15 illustrates that both kernels K_1 and K_2 do not change drastically from one calibration date to the next.



(a) Sensitivities $(\beta_0, \beta_1, \beta_2, \beta_{1,2})$ of the volatility for two consecutive calibration dates.



(b) Parameters of the kernels K_1 and K_2 for two consecutive calibration dates. L.h.s.: parameters $(\lambda_{1,0}, \lambda_{1,1}, \theta_1)$ of the kernel K_1. R.h.s.: parameters $(\lambda_{2,0}, \lambda_{2,1}, \theta_2)$ of the kernel K_2. For each plot, the y-axis on the l.h.s. refers to the natural scale of the $\lambda_{n,p}$s while the y-axis on the r.h.s. refers to the values of the θ_ps.

Figure 14

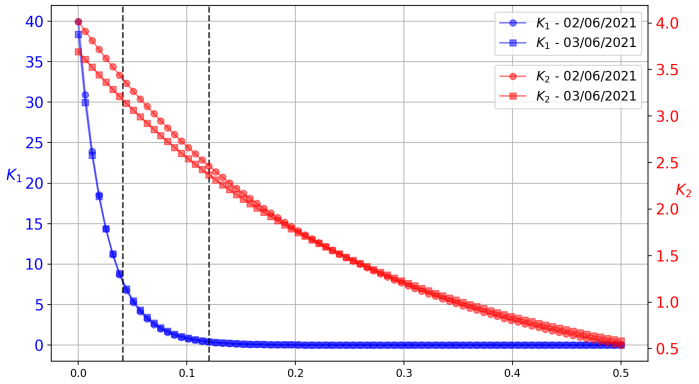


Figure 15: Comparison of the calibrated exponential kernels jointly to SPX and VIX. The two vertical dotted lines denote the smallest maturity considered for calibration (13 days) and the largest one (44 days).

7 Pricing light exotics

One of the benefits of learning the VIX pathwise, and not only the prices at time 0 of VIX vanillas, is that in addition to a faster joint calibration discussed in the previous section, we are also able to quickly price derivatives involving the VIX. For instance, we consider the following light exotics, typically sold by banks to hedge funds, a VIX call with a lower barrier on the SPX (Payoff 1), and a VIX put with an upper barrier on the SPX (Payoff 2):

$$\text{Payoff 1}: \quad (\text{VIX}_{T_1} - K)_+ \mathbf{1}_{\{\inf_{t\in[0,T_1]} \frac{S_t}{S_0} \geq B\}},$$
$$\text{Payoff 2}: \quad (K - \text{VIX}_{T_1})_+ \mathbf{1}_{\{\sup_{t\in[0,T_1]} \frac{S_t}{S_0} \leq B\}}.$$

The barriers are introduced to make the derivative much cheaper, as often the SPX has hit the low (resp., high) barrier when the VIX call (resp., put) is in the money. We can very quickly price these payoffs in the jointly calibrated 4FPDV model, using our fast pathwise approximation $\mathcal{NN}^\star$ for the VIX. We use a discretization step of $\Delta t = \frac{1}{2520}$, $N_{MC} = 2 \times 10^6$ Monte Carlo samples, the model parameters reported in Table 6, and $T_1 = 14$ days. Figures 16 and 17 show the impact of the strike and barrier levels.

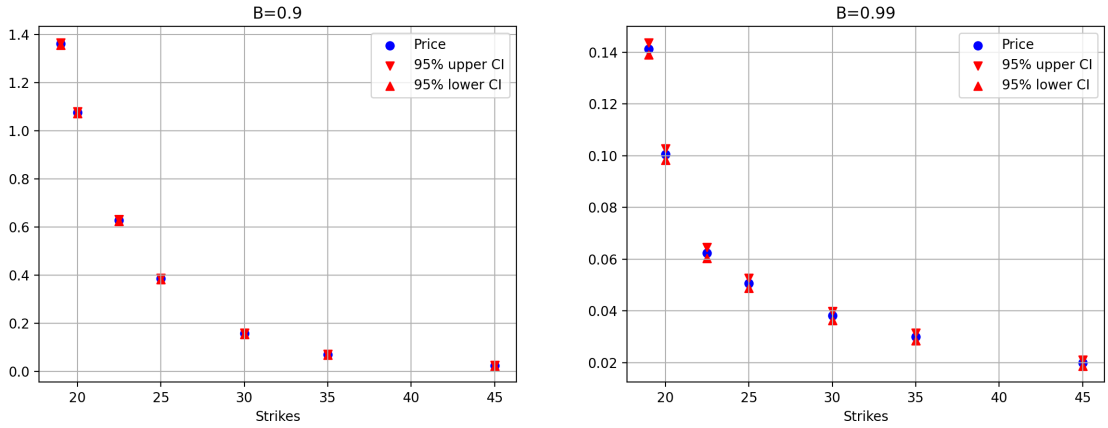


Figure 16: Payoff 1

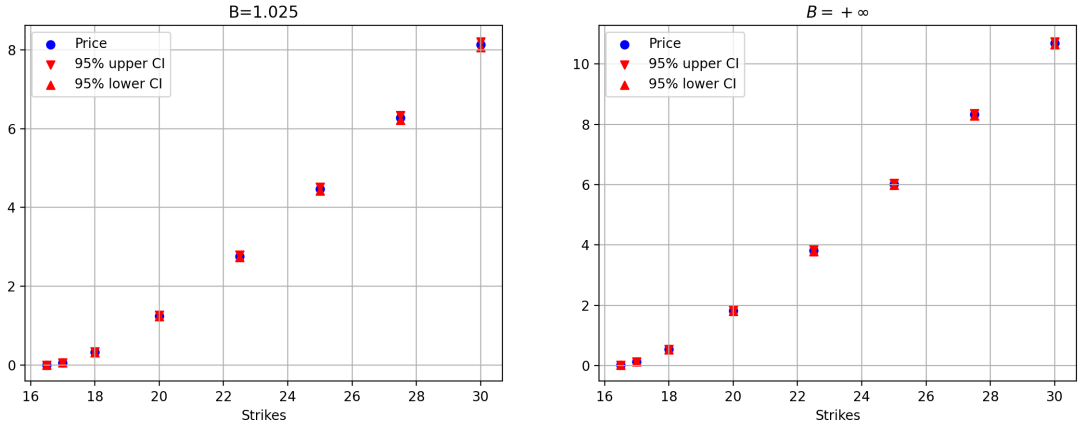


Figure 17: Payoff 2

8 Conclusion

This paper investigates the VIX in path-dependent volatility (PDV) models, focusing on the 4FPDV model by Guyon and Lekeufack (2023). It introduces a *pathwise* neural network approximation of the VIX as a function of both the Markovian factors driving the volatility dynamics and the model parameters. The VIX is indeed not known in closed form in the model. For given model parameters, the dependence on the factors allows us to quickly sample VIX paths and price derivatives involving the VIX. The dependence of our neural net approximation on the model parameters allows us to quickly calibrate the model jointly to SPX and VIX options. Compared to the existing *deep pricing* and *deep calibration* techniques, our methodology allows us to have access to the VIX pathwise in the model, which would be lost if we learned directly the map to model option prices or model implied volatilities.

Also importantly, the paper shows, using a few random calibration dates, that the *time-homogeneous*, low-parametric, Markovian 4FPDV model is able to fit the surface of SPX

implied volatilities remarkably well.

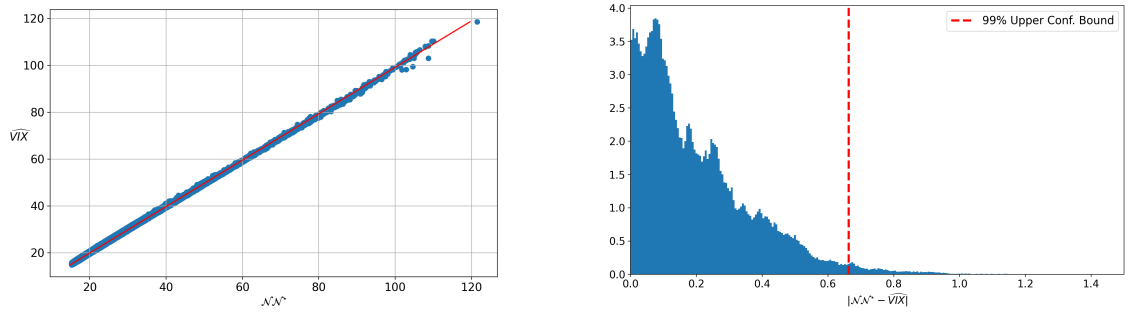
Perspectives of future research include performing a large-scale empirical study to measure how well the 4FPDV model fits the SPX implied volatility surface over several years of calibration dates; jointly calibrating the model to more, longer maturities; measuring the impact of introducing a stochastic volatility component and/or discrete time; and investigating the tractability properties of affine or polynomial versions, directly written on the variance, of this PDV model, in which the VIX is known in analytical form.

Acknowledgements

The authors thank the two anonymous referees for useful comments which helped improve this article. The first author would also like to thank Fabio Baschetti for fruitful discussions.

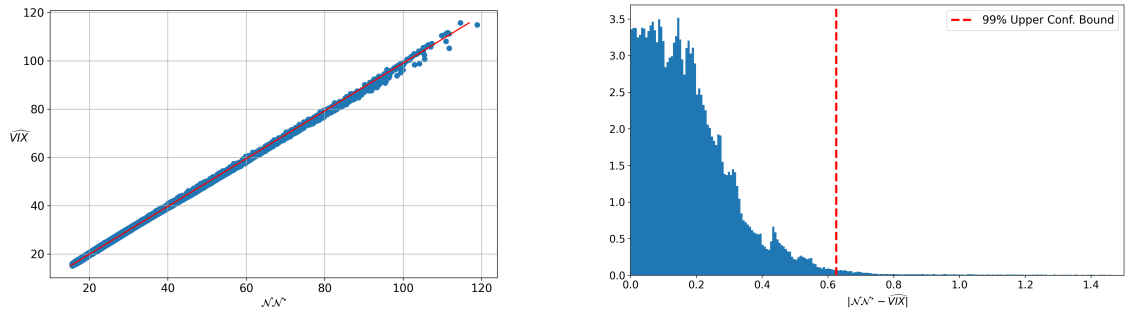
A Appendix: Additional tests of the neural network

In this appendix, we report the error between the neural network predictions $\mathcal{NN}^\star$ and the nested Monte Carlo estimator $\widehat{\text{VIX}}$ for market-calibrated parameters, see Figures 18 and 19.



(a) (b)

Figure 18: Model parameters are jointly calibrated on June 2, 2021. Left: comparison between $N_{MC} = 3 \cdot 10^5$ realizations of the nested Monte Carlo estimator $\widehat{\text{VIX}}$ and the neural approximation $\mathcal{NN}^\star$. Here the $\widehat{\text{VIX}}$ is computed with $\Delta t = \frac{1}{2520}$, $3 \cdot 10^5$ outer paths and 10^4 nested paths. Right: histogram of the absolute error between the neural network predictions $\mathcal{NN}^\star$ and $\widehat{\text{VIX}}$. The average absolute error is 0.202%.



(a) (b)

Figure 19: Same as Figure 18, with model parameters jointly calibrated on June 3, 2021. The average absolute error between the neural network predictions $\mathcal{NN}^\star$ and $\widehat{\text{VIX}}$ is 0.185%.

Disclosure of interest

The authors declare that there are no relevant financial or non-financial competing interests to report.

References

E. Abi Jaber and S. Li. Volatility models in practice: Rough, Path-dependent or Markovian? *Preprint arXiv:2401.03345*, 2024.

E. Abi Jaber, M. Larsson, and S. Pulido. Affine Volterra processes. *Annals of Applied Probability*, 29(5):3155–3200, 2019.

E. Abi Jaber, C. Illand, and S. Li. Joint SPX-VIX calibration with Gaussian polynomial volatility models: deep pricing with quantization hints. *Preprint arXiv:2212.08297*, 2022.

A. Alfonsi, B. Lapeyre, and J. Lelong. How many inner simulations to compute conditional expectations with least-square Monte Carlo? *Methodology and Computing in Applied Probability*, 25(3):71, 2023.

H. Andrès and B. Jourdain. Existence, uniqueness and positivity of solutions to the Guyon-Lekeufack path-dependent volatility model with general kernels. *Preprint arXiv:2408.02477*, 2024.

H. Andrès, A. Boumezoued, and B. Jourdain. Implied volatility (also) is path-dependent. *Preprint arXiv:2312.15950*, 2023.

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.

F. Baschetti, G. Bormetti, and P. Rossi. Deep calibration with random grids. *Quantitative Finance*, 1(1):1–23, 2024.

C. Bayer and B. Stemper. Deep calibration of rough stochastic volatility models. *Preprint arXiv:1810.03399*, 2018.

P. Blanc, J. Donier, and J.-P. Bouchaud. Quadratic Hawkes processes for financial prices. *Quantitative Finance*, 17(2):171–188, 2017.

A. Bondi, G. Livieri, and S. Pulido. Affine volterra processes with jumps. *Stochastic Processes and their Applications*, 168:104264, 2024a.

A. Bondi, S. Pulido, and S. Scotti. The rough Hawkes Heston stochastic volatility model. *Mathematical Finance*, pages 1–45, 2024b.

F. Bourgey and J. Guyon. Fast exact joint S&P 500/VIX smile calibration in discrete and continuous time. *Risk*, February, 2024.

C. Cartis, J. Fiala, B. Marteau, and L. Roberts. Improving the flexibility and robustness of model-based derivative-free optimization solvers. *ACM Transactions on Mathematical Software (TOMS)*, 45(3):1–41, 2019.

C. Cartis, L. Roberts, and O. Sheridan-Methven. Escaping local minima with local derivative-free methods: a numerical investigation. *Optimization*, 71(8):2343–2373, 2022.

R. Chicheportiche and J. Bouchaud. The fine-structure of volatility feedback I: Multi-scale self-reflexivity. *Physica A: Statistical Mechanics and its Applications*, 410:174–195, 2014.

R. Cont and S. Ben Hamida. Recovering volatility from option prices by evolutionary optimization. *Journal of Computational Finance*, 8(4):43–76, 2005.

R. Cont and T. Kokholm. A consistent pricing model for index options and volatility derivatives. *Mathematical Finance*, 23(2):248–274, 2013.

C. Cuchiero and J. Teichmann. Markovian lifts of positive semidefinite affine Volterra type processes. *Decisions in Economics and Finance*, 42(2):407–448, 2019.

C. Cuchiero, G. Gazzani, and S. Svaluto-Ferro. Signature-based models: Theory and calibration. *SIAM Journal on Financial Mathematics*, 14(3):910–957, 2023.

C. Cuchiero, G. Gazzani, J. Möller, and S. Svaluto-Ferro. Joint calibration to SPX and VIX options with signature-based model. *Mathematical Finance*, page 1–53, 2024.

G. Di Nunno, K. Kubilius, Y. Mishura, and A. Yurchenko-Tytarenko. From constant to rough: A survey of continuous volatility modeling. *Mathematics*, 11(19):4201, 2023.

B. Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

P. Foschi and A. Pascucci. Calibration of the Hobson and Rogers model: empirical tests. *Computational Statiscs & Data Analysis*, 53:2219–2235, 2005.

J.-P. Fouque and Y. Saporito. Heston stochastic vol-of-vol model for joint calibration of VIX and S&P 500 options. *Quantitative Finance*, 18(6):1003–1016, 2018.

M. Gambara and J. Teichmann. Consistent recalibration models and deep calibration. *Preprint arXiv:2006.09455*, 2020.

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. *Risk*, May, 2020.

I. Guo, G. Loeper, J. Obloj, and S. Wang. Optimal transport for model calibration. *Risk*, January, 2022.

J. Guyon. Inversion of convex ordering in the VIX market. *Quantitative Finance*, 20(10): 1597–1623, 2020a.

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

J. Guyon. The VIX future in Bergomi models: Fast approximation formulas and joint calibration with S&P 500 skew. *SIAM Journal on Financial Mathematics*, 13(4):1418–1485, 2022a.

J. Guyon. Dispersion-constrained martingale Schrödinger bridges: Joint entropic calibration of stochastic volatility models to S&S 500 and VIX smiles. *Available at SSRN 4165057*, 2022b.

J. Guyon. Dispersion-constrained martingale Schrödinger problems and the exact joint S&P 500/VIX smile calibration puzzle. *Finance and Stochastics*, 28(1):27–79, 2024.

J. Guyon and J. Lekeufack. Volatility is (mostly) path-dependent. *Quantitative Finance*, 23(9):1221–1258, 2023.

J. Guyon and S. Mustapha. Neural joint S&P 500/VIX smile calibration. *Risk*, December, 2023.

P. Henry-Labordère. From (martingale) Schrödinger bridges to a new class of stochastic volatility models. *Preprint arXiv:1904.04554*, 2019.

A. Hernandez. Model calibration with neural networks. *Available at SSRN 2812140*, 2016.

D. Hobson and L. Rogers. Complete models with stochastic volatility. *Mathematical Finance*, 8(1):27–48, 1998.

B. Horvath, A. Muguruza, and M. Tomas. Deep learning volatility: a deep neural network perspective on pricing and calibration in (rough) volatility models. *Quantitative Finance*, 21(1):11–27, 2021.

T. Kokholm and M. Stisen. Joint pricing of VIX and SPX options with stochastic volatility and jump models. *The Journal of Risk Finance*, 16(1):27–48, 2015.

L. Li, K. Jamieson, G. De Salvo, A. Rostamizadeh, and A. Talwalkar. Hyperband: A novel bandit-based approach to hyperparameter optimization. *The journal of machine learning research*, 18(1):6765–6816, 2017.

A. Neuberger. The log contract. *Journal of portfolio management*, 20:74–74, 1994.

M. Nutz and A. Riveros Valdevenito. On the Guyon-Lekeufack volatility model. *Finance and Stochastics*, 4:1203–1223, 2024.

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.

L. Parent. The EWMA Heston model. *Quantitative Finance*, 23(1):71–93, 2023.

S. E. Rømer. Empirical analysis of rough and classical stochastic volatility models to the SPX and VIX markets. *Quantitative Finance*, 22(10):1805–1838, 2022.

M. Rosenbaum and J. Zhang. Deep calibration of the quadratic rough Heston model. *Risk*, September, 2022.

E. Sentana. Quadratic ARCH models. *The Review of Economic Studies*, 62(4):639–661, 1995.

G. Zumbach. Time reversal invariance in finance. *Quantitative Finance*, 9(5):505–515, 2009.

G. Zumbach. Volatility conditional on price trends. *Quantitative Finance*, 10(4):431–442, 2010.