The Step Stochastic Volatility Model (SSVM)

Peter K. Friz[*], Paolo Pigato[†], Jonathan Seibel[‡]

May 7, 2020

Abstract

Stochastic Volatility Models (SVMs) are ubiquitous in quantitative finance. But is there a Markovian SVM capable of producing extreme ($T^{-1/2}$) short-dated implied volatility skew?

We here propose a modification of a given SVM "backbone", Heston for instance, to achieve just this - without adding jumps or non-Markovian "rough" fractional volatility dynamics. This is achieved via non-smooth leverage function, such as a step function. The resulting Step Stochastic Volatility Model (SSVM) is thus a parametric example of local stochastic volatility model (LSVM). From an IT perspective, SSVM amounts to trivial modifications in the code of existing SVM implementations. From a QF perspective, SSVM offers new flexibility in smile modelling and towards assessing model risk. For comparison, we then exhibit the market-induced leverage function for LSVM, calibrated with the particle method.

1 Overview

SVMs are widely used for pricing and risk management of financial products. As is well documented, e.g. [12, 3], SVMs offer reasonable and parsimonious dynamics, at the price of imperfect market fits, in contrast to local volatility models (LVMs) that provide essentially perfect fits to European market data [5], at the price of unreasonable (forward) dynamics and infinitely many parameters. Local stochastic volatility models (LSVMs) have been successful in achieving essentially perfect fits with reasonable dynamics, largely due to the particle (calibration) method of Guyon and Henry-Labordère [15, 16], which makes the leverage function (by which we mean the local vol type decoration of a stochastic vol backbone) effectively non-parametric. In a different development, rough SVMs [1], in which volatility follows an anomalous diffusion, have shown remarkable ability to combine reasonable and parsimonious dynamics with (very) good market fits, including the ability to produce extreme short-dated skews as seen in market data. Rough SVMs come at price of non-Markovian dynamics, with much recent progress on efficient numerical schemes [23]. We shall not comment here on the vast literature of SVM with jumps but mention [25, 13, 6] for a discussion of pro and cons with regard to smile and skew modelling.

In this note we propose a parametric LSVM which decorates a given SVM with a leverage function, which has a discontinuity at-the-money (ATM). An immediate way to achieve this amounts

[*]TU and WIAS Berlin
[†]University of Rome Tor Vergata
[‡]Munich Re

to multiply the (backbone) stochastic volatility with distinct factors, say σ_-, σ_+ depending on when the considered option is out-of-the-money or in-the-money. This step function form led us to the proposed name *Step Stochastic Volatility Model* (SSVM). Econometric intuition suggests higher-vol in the OTM region, so that one may want to take $\sigma_- > \sigma_+$. If we assume a suitable average to be of order one (which makes sense if the SVM backbone was previously calibrated to market data), we are left with one single extra parameter[1]. The implied skew generated by such model explodes as $T^{-1/2}$, as was mathematically demonstrated in [27] in the flattest of all SVMs: the Black-Scholes model with constant volatility (in this case our SSV model degenerates to a singular LV.)

To summarize, our SSV model is parsimonious (one effective additional parameter compared to the backbone SVM), able to produce extreme short-dated skews, still Markovian (hence PDE methods are available, at least in low dimensions), trivially implemented (modify one line in the code of an existing Monte Carlo SVM implementation).

We would like to point out that much of the recent success of rough SVMs [1, 9] is due to the (desired) implied skew blow up at rate $T^{H-1/2}$, where the Hurst parameter $H \in (0, 1/2]$ quantifies the roughness of the volatility process. We recall in this context the blowup can be *at most* of order $T^{-1/2}$; this is a model free consequence of no-arbitrage (see the ATM skew short-dated bounds in [19] and [8, Remark 2.4]). Rough SVMs thus cover the entire range of possible skew blowups, however become "super-critical" at $H = 0$. (We refer to [26] for a first study of the super-critical limit $H \downarrow 0$, relying on Kahane's theory of Gaussian multiplicative chaos.) It is remarkable that SSVM achieves the limiting rate $T^{-1/2}$ without any mathematical complication such as jumps or giving up Markovianity. In particular, PDE based pricing methods remain possible.

The question remains to what extent SSVMs, or variations thereof (see below), are actually practically useful. Trivial things first, SSVMs will perform no worse than the backbone SVM by itself (take $\sigma_- = \sigma_+ = 1$, i.e. a unit leverage function), while it can be expected - by design - to perform better when dealing with steep short-dated skews. We will see this numerically in Section 2 below. At the same time, SSVM amounts to trivial modifications in the code of existing SVM implementations, an appealing feature to practitioners in need for additional flexibility, such as to assess model risk for exotic products. (It is common to compare exotic prices under stochastic vs local volatility; see e.g. Reghai [29] where this is quantified as "toxicity index".) Parametric ramifications of SSVM are easy: it suffices to replace the step function by other (low dimensional parametric) leverage functions with well-designed singularity. In Section 4, we see that a leverage function of the form $|x|^{H(x)}$, were $H(x)$ is a step function, provides a very good fit to the short-term leverage function calibrated to the Euro Stoxx 50, with Heston backbone, for multiple initial dates. That said, "perfect calibration" is really the domain of LSVM and in the final part of the paper we address the looming question of how a parametric guess compares to the "market implied" leverage functions, as computed by the Guyon–Henry-Labordère' particle method (see Section 4).

Acknowledgements: This work was presented (PKF) at *Quant Minds Europe*. Participants are thanked for their feedback, especially Julien Guyon and Lorenzo Bergomi for repeated discussions and valuable comments. JS is grateful to Jan-Hiroshi Schwarz for discussion and helpful remarks. PKF and PP acknowledge support from DFG Excellence Cluster MATH+ (Projekt AA4-2).

[1]As is typical for LSVMs (resp. LVMs), the leverage function (resp. local volatility function) is chosen as function of today's volatility surface and is not consistent over time. We discuss a remedy in Remark 1, following Guyon [14].

1.1 Square-root law for ATM implied volatility skew term structure

Conventional one-factor SVMs may fail to reproduce the steep volatility market smiles observed for short dates, since they predict, in short time, a finite limit for the ATM implied volatility skew [12, 9]. Such skew has a slope, which can be strong when mean-reversion is strong, but the skew of these models always converges to a finite limit. Multi-factor SVM, for example the 2-factor Bergomi model, with one fast mean-reverting factor and one slow mean-reverting factor, can well approximate the term-strucure of ATM market skew, even for very short maturities, even though the short time limit of their implied skew is finite (see Figure 2).

One (main) motivation of the recent success of rough volatility models is the fact that the empirical ATM implied volatility skew, defined as

$$\psi(T) = \left| \frac{\partial}{\partial k} \sigma_{BS}(T,k)\,|_{k=0} \right|, \quad k = \log(K/F_T)$$

is consistently found [1] to be proportional to a power law

$$\psi(T) \sim \frac{1}{T^{\alpha}}, \quad \text{with } \alpha \in (0, 0.5)$$

for a wide range of expirations T. Calibrations of various authors give $\alpha \in (0.3, 0.5)$, consistent with a practitioners rule-of-thumb [19] that states $\alpha = 0.5$, see also [10, Remark 4.1] for a recent explanation why this should be so.

Here F_T represents the forward $F_T = S_0 \exp(\int_0^T (r_s - q_s) ds)$, as it is common in empirical studies and in general between practitioners. In theoretical papers, the definition of k is usually slightly different, as $k = \log(K/S_0)$, where S_0 is the spot price. Anyway, many works assume zero rates and dividends ($r \equiv 0$ and $q \equiv 0$), so that $F_T = S_0$.

In Figure 1, we see an implied volatility surface of the Euro Stoxx 50.

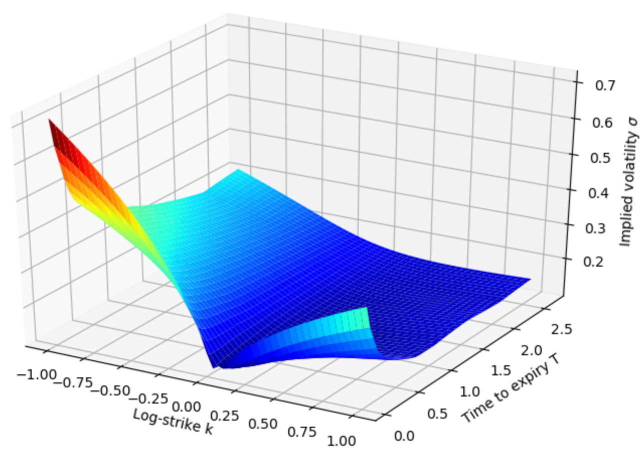

Figure 1: Implied Volatility Surface Euro Stoxx 50 as of 04-June-2018.

In Figure 2, we see the implied skew of the surface in Figure 1 and the corresponding 2-factor Bergomi vs power-law fit.

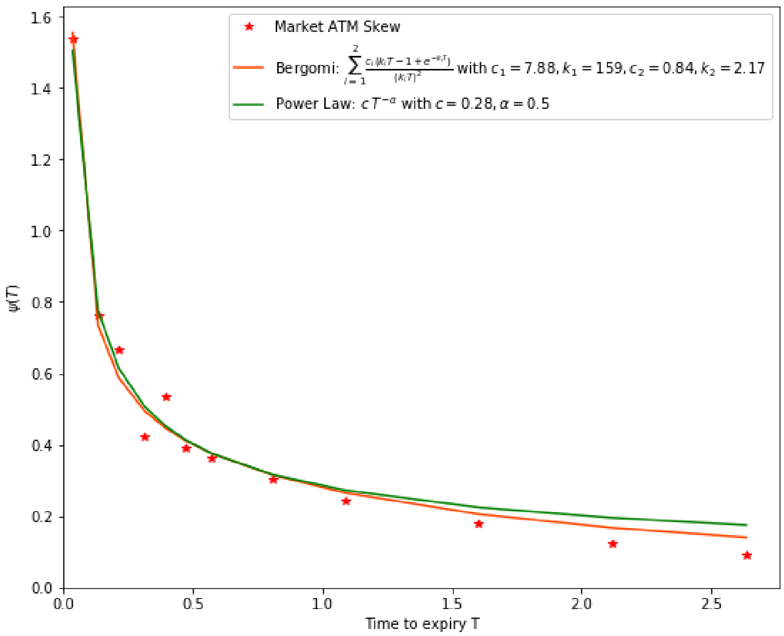

Figure 2: Euro Stoxx 50 ATM skew as of 04-June-2018 versus time to expiry. The green dashed line is a power law fit. The orange line is a 2-factor Bergomi fit. Time is measured in years.

In this context, a key quantity is the Skew Stickyness Ratio $\mathcal{R}_T$ (SSR): Bergomi [3] shows that for smooth LV and standard SV (even multi-factor) models, the SSR tends to 2 as time-to-maturity goes to 0. In particular, for smooth LV models, [3, Equation 2.64] relates Skew Stickyness Ratio $\mathcal{R}_T$ and implied skew and shows that if $\lim_{T\downarrow 0} \psi_T$ is finite, then $\lim_{T\downarrow 0} \mathcal{R}_T = 2$. On the other hand, it is observed in [32] that the estimated SSR becomes larger than 2 for small maturities[2]. This result can be seen as an ulterior motivation to explore non-smooth local volatilities and look for simple models with exploding implied skew.

2 Stepping StochVol

StochVol backbone. In LSV models, the SV backbone is fixed a priori and follows the dynamics of a classical SV model, often taken as Heston which allows for efficient calibration of the backbone[3]. Our implementation in Section 4 is hence based on a Heston backbone. To set notation, recall [17]

[2]This conclusion seems to be the opposite of [3, Section 9.9], but this may be due to the fact that in [3] the short-time limit of the SSR is not extrapolated, but taken at a fixed, small positive time (one month).

[3]Heston dynamics are "reasonable" compared to local volatility but have their own deficiencies, such as unrealistic joint dynamics of spot and implied vol , see e.g. [3, Chapter 6].

$\sqrt{v_0}$	κ	η	ρ	$\sqrt{\bar{v}}$
0.13	3.83	0.76	-0.69	0.18

Table 1: Calibrated parameters for Heston model on Euro Stoxx 50 as of 04-June-2018.

that the stock price S_t and its variance v_t evolve as

$$\begin{aligned} \frac{dS_t}{S_t} &= \sqrt{v_t}(\rho dW_t + \sqrt{1-\rho^2}dZ_t) \\ dv_t &= \kappa(\bar{v} - v_t)dt + \eta\sqrt{v_t}dW_t \end{aligned} \tag{2.1}$$

where W_t and Z_t are two uncorrelated Brownian motions, $\rho \in (-1, 1)$, $\bar{v}, \kappa, \eta, r, q \geq 0$ are numerical parameters and $S_0, v_0 > 0$ are the initial values. (There is no problem to adapt to non-zero deterministic rates and dividends, it suffices to include a drift term of the form $(r - q)dt$ in the first line.) The parameters have the following interpretation: v_0 is the initial variance, $\bar{v}$ the long term variance, η the volatility-of-volatility, κ the speed of mean reversion, ρ is the correlation.

The following short-time limit holds for the implied skew $\psi(T) = |\partial_k \sigma_{BS}(T, k)|$ under Heston dynamics with correlation ρ and volatility-of-volatility η (see [12]):

$$\frac{\partial}{\partial k}\sigma_{BS}(T, 0) \xrightarrow{T\to 0} \frac{\rho\eta}{4\sqrt{v_0}} \tag{2.2}$$

(equivalently $\frac{\partial}{\partial k}\sigma^2_{BS}(T, 0) \xrightarrow{T\to 0} \rho\eta/2$). As an example, consider now the empirical volatility surface of the Euro Stoxx 50 as of 04-June-2018. We have the calibrated parameters as in Table 1 (see Section 4 for details). The corresponding short time skew would be

$$\psi(T) \approx \frac{|\rho|\eta}{4\sqrt{v_0}} \approx 1.01 \tag{2.3}$$

In the short end, the empirical skew blows up and this value will be far from data, cf. Figure 2, where the shortest maturity skew is $\psi(T) \approx 1.6$. (As we have seen in Section 1.1, 2-factors SV or rough stochastic volatility will do better.)

Implied skew and Local Volatility: the one half-rule. It was shown in [9] that for general LV models

$$\frac{dS_t}{S_t} = \sigma_{loc}(t, S_t)dW_t, \tag{2.4}$$

under minimal regularity conditions, the implied skew $\partial_k \sigma_{BS}(T, 0)$ converges for short expiries to a constant proportional to the LV skew $\partial_S \sigma_{loc}(0, S_0)$, and therefore there is no blow-up of the implied skew. More precisely, this happens for the approximation of the derivative given by the finite difference at the central limit regime, meaning that for fixed $z \neq \zeta$ we have

$$\frac{\sigma_{BS}(T, \sqrt{T}z) - \sigma_{BS}(T, \sqrt{T}\zeta)}{\sqrt{T}(z - \zeta)} \sim \frac{S_0}{2}\frac{\partial}{\partial S}\sigma_{loc}(0, S_0) \quad \text{for } T \downarrow 0. \tag{2.5}$$

Heuristically, this follows from the out-of-the money (OTM) short time asymptotic results proved in [2], holding for very general local volatilities. For $T \downarrow 0$, the implied volatility generated by a LV model is approximated by the harmonic average of the LV function

$$\lim_{T\downarrow 0} \frac{1}{\sigma_{BS}(T,x)} = \int_0^1 \frac{ds}{\sigma_{loc}(0,sx)}, \tag{2.6}$$

with a slight abuse of notation since here σ_{loc} depends on time and log-price $X = \log S$. Formally differentiating such relation one can recover (2.5). Hence, exploding skew of implied volatility can be produced by a local vol model only if $\sigma_{loc}(0,\cdot)$ is not nice.

Stepping local volatility. The simplest way to introduce a singularity ATM in the LV function, which we expect to produce a negative skew, is to consider a Brownian motion run at volatility σ_+ on $B \in R^+$, at volatility σ_- on $B \in R^-$. This is formally defined as $(B_t)_{t\geq 0}$ solution to the SDE

$$dB_t = H(B_t)dW_t \tag{2.7}$$

with

$$H(x) = \sigma_+ 1_{x\geq 0} + \sigma_- 1_{x<0}, \tag{2.8}$$

so that H is a Heaviside step function. This process, introduced by Keilson-Wellner in [18] is referred to as oscillating Brownian motion [18, 21]. It has a strict relation with the skew Brownian motion (SBM) in the sense that the process $(B_t/H(B_t))_{t\geq 0}$ is a SBM. For pricing, we consider its geometric variant

$$\frac{dS_t}{S_t} = H\left(\log\left(S_t/S_0\right)\right)dW_t. \tag{2.9}$$

This model was recently studied as theoretical model by Gairat, Shcherbakov [11] and Lipton [21] (see also the previous work on the multi-tile case [22]). The SBM itself had been previously considered for pricing and arbitrage in [4, 30]. In [27], it is shown that the implied volatility of such model, in short time, has the following ATM expansion, for $k \downarrow 0$

$$\sigma_{BS}(T,k) = \frac{2\sigma_+\sigma_-}{\sigma_+ + \sigma_-} + \frac{\sqrt{\pi}}{\sqrt{2}}\frac{\sigma_+ - \sigma_-}{\sigma_+ + \sigma_-}\frac{k}{\sqrt{T}} + O\left(\frac{k^2}{T}\right) \tag{2.10}$$

This means that the implied vol level (effective volatility) is the harmonic average $\frac{2\sigma_+\sigma_-}{\sigma_+ + \sigma_-}$ and, most interestingly for us, the skew explodes in short time as

$$\psi(T) \sim 1/\sqrt{T} \text{ as } T \downarrow 0.$$

Qualitatively, such behavior looks similar to the empirical one (Section 1.1). It is quite extreme, as it almost attains a model-free bound (cf. [27, Remark 8], [8, Remark 2.4] and [19]). Also notice that the skew is negative when $\sigma_+ < \sigma_-$.

Stepping Stochastic Volatility. We propose to change the SVM backbone, Heston for the sake of presentation, (2.1) to a LSV dynamics

$$\begin{aligned} \frac{dS_t}{S_t} &= H\left(\log\left(\frac{S_t}{S_0}\right)\right)\sqrt{v_t}\,(\rho dW_t + \sqrt{1-\rho^2}dZ_t) \\ dv_t &= \kappa(\bar{v} - v_t)dt + \eta\sqrt{v_t}dW_t \end{aligned} \tag{2.11}$$

with the "step" leverage function H defined (2.8). (Most works on local volatility we are aware of, assume regularity of the leverage functions, but see [7].) We can consider $\sigma_\pm$ as free parameters, or impose that $\sigma_\pm$ have unit harmonic average (since the harmonic average is the implied vol level of H, cf. (2.10)). In this case, the model has only one additional parameter with respect to simple Heston, but we expect exploding skew in short time.

This approach has the following advantages:

- The model is parsimonious (e. g., with Heston backbone, unit harmonic average, we have Heston parameters + 1)
- It is at least as "good" as SV (e.g. Heston) with regard to fitting, dynamics but allows for extreme skews in the short end
- It is at least as "good" as skew or oscillating Brownian motion in (2.7)
- It displays exploding skew with no need for jumps
- It is still a Markov model, with the possibility (in principle) to use PDE methods for pricing
- The implementation is extremely easy, via a trivial modification of existing (Monte Carlo) Heston code.

As one can see in Figure 3, the additional step parameter can indeed be used to adjust the Heston model for extreme short-term skews. Since Heston fits are already good for long maturities, it may be useful to introduce a time-dependence in order to turn off the skew correction for long expiries. To fit specific datasets, we may also work with simple modifications of this leverage function. The point in any case is to impose a priory a singularity ATM in the leverage function.. Concerning the skew explosion, one finds $T^{-1/2}$ term-structure of the ATM implied vol skew [27], to be compared with $T^{H-1/2}$ in rough volatility [9, 1].

Remark 1 (Guyon's Stepping path-dependent volatility)**.** *A feature of financial markets which is desirable, but hard, to reproduce via simple volatility models is the presence of very negative short term forward skews. In [14], Guyon studies path-dependent local volatility models of the form*

$$\frac{dS_t}{S_t} = \sigma(t, S_t, X_t) l(t, S_t) dW_t \tag{2.12}$$

where X_t can be a past value of the stock ($S_{t-\Delta}$, with fixed $\Delta > 0$), or for example a moving average or minimum/maximum of past values of S. The structure of the path dependent function is imposed a priori, and when it involves a step function (e.g. $\sigma(S, X) = \sigma_+ 1_{S/X \leq 0} + \sigma_- 1_{S/X > 0}$) such model generates very negative short term forward skews (as the ones observed on data). In [14], the local volatility part $l(t, S_t)$ is then calibrated using the particle method, so that a perfect fit of the volatility surface is obtained. As in (2.11), the step is used to "enforce" the extremely asymmetric behavior of the (forward) skew. As mentioned in the introduction, such type of time dependent structure could also be a remedy for the time-inconsistency of the LSVM (2.11): taking X as a moving average of past values of S, time 0 plays no particular role and the model is time-consistent.

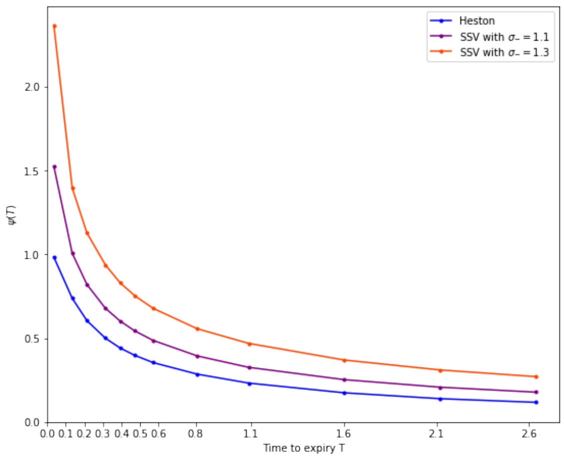

Figure 3: Term-structure of ATM implied volatility skews produced by SSV model for different choices of σ_-. Used Heston parameters: $\sqrt{v_0} = 0.13, \kappa = 3.83, \eta = 0.76, \rho = -0.69, \sqrt{\bar{v}} = 0.18$. The parameters were taken from a calibration of the Heston model to SX5E implied volatility data as of 04-June-2018 (cf. Table 1).

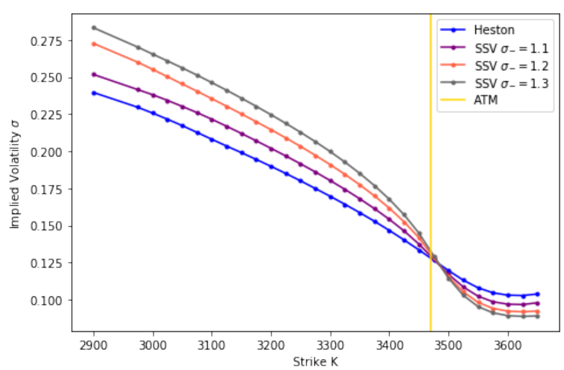

Figure 4: Implied volatility smiles produced by SSV model for different choices of σ_- for time to expiry $T = 0.04$ years. Used Heston parameters: $\sqrt{v_0} = 0.13, \kappa = 3.83, \eta = 0.76, \rho = -0.69, \sqrt{\bar{v}} = 0.18$ (as in Figure 3). Note how higher steps produce steeper skews.

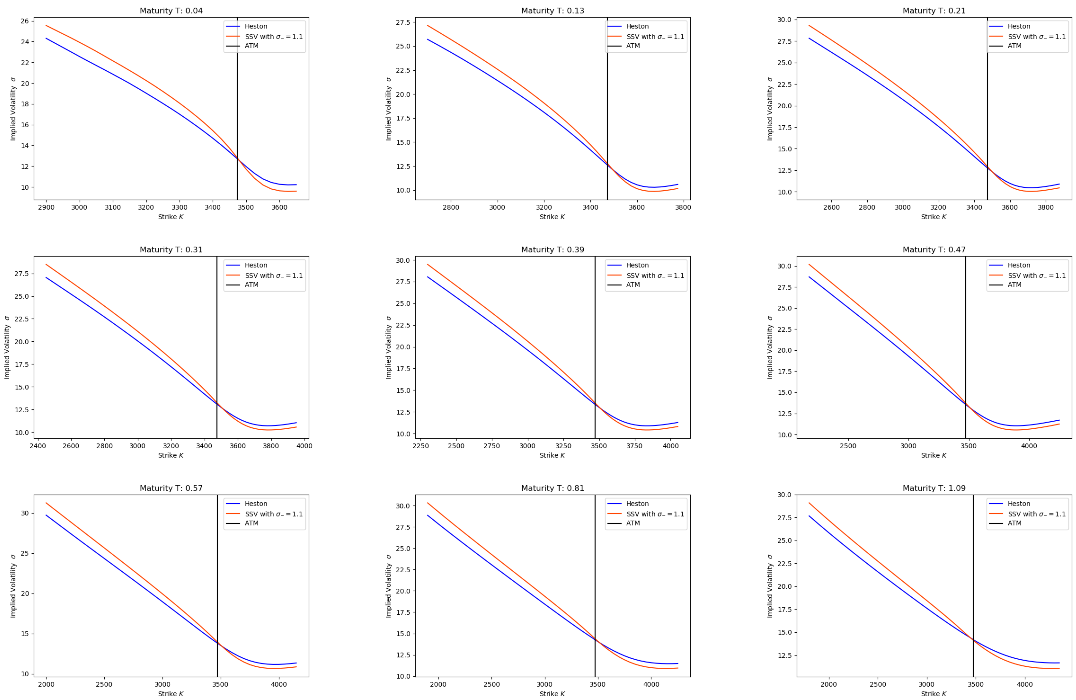

Figure 5: Heston model parameters: $\sqrt{v_0} = 0.13, \kappa = 3.83, \eta = 0.76, \rho = -0.69, \sqrt{\bar{v}} = 0.18$. The blue lines are implied volatilities computed on a Monte Carlo simulation of the Heston model. The orange lines represent implied volatilities computed on a Monte Carlo simulation of the SSV model with $\sigma_- = 1.1$.

3 Local stochastic volatility and the Guyon and Henry-Labordère particle method

Local Stochastic Volatility (LSV) models were introduced to combine the best features of SV and LV models, namely realistic dynamics and perfect fits to the vanilla market [28, 20]. This is accomplished by embedding a deterministic leverage function $l(t, x)$ into a SV base model (e.g. Heston), assuming instantaneous volatility to be modeled through

$$\sigma_t = l(t, S_t)\sqrt{v_t} \tag{3.1}$$

where v_t denotes the variance process of the base model. Consequently, we assume that in a LSV model the spot S_t and its variance v_t evolve as

$$\begin{aligned} \frac{dS_t}{S_t} &= l(t, S_t)\sqrt{v_t}\,(\rho dW_t + \sqrt{1-\rho^2} dZ_t) \\ dv_t &= \alpha(t, v_t)dt + \beta(t, v_t)dW_t, \end{aligned} \tag{3.2}$$

where W_t, Z_t are two uncorrelated Brownian motions. For simplicity, we have taken zero interest rates, repo, and dividends. Note that the leverage function is usually meant to serve as a correction of the SV model. One calibrates the naked SV model first, in order to have a model that generates realistic base dynamics, and then calibrates the leverage function afterwards to eliminate deviations between the SV base model and the market.

However, the particle method can also serve different purposes. For example, in order to obtain a precise, specific behavior, one can first choose specific model parameters, instead of calibrating the SV model to the market. Then, one can add on top a leverage function generated by the particle method, so that the resulting model still produces the desired behavior, and it is also perfectly calibrated to the market smiles (compare also with [14], where such procedure is applied to PDV models).

Applying Itô-Tanaka's formula one can show that model 3.2 is perfectly calibrated to the market smile if and only if the leverage function $l(t, x)$ satisfies the following functional equation

$$l(t, x) = \frac{\sigma_{\text{loc}}(t, x)}{\sqrt{\mathbb{E}_{\mathbb{Q}}\left[v_t | S_t = x\right]}}. \tag{3.3}$$

where $\sigma_{\text{loc}}(t, x)$ is the Dupire local volatility [15]. In fact, as the conditional expectation on the right-hand side of 3.3 depends on the joint probability distribution $\mathbb{Q}_t$ of $X_t := (S_t, v_t)$, the calibrated LSV model is an example of a non-linear McKean SDE, in which the drift and the volatility depend not only on the current value X_t of the process, but also on the probability distribution $\mathbb{Q}_t$ of X_t, i.e.

$$dX_t = b(t, X_t, \mathbb{Q}_t)dt + \sigma(t, X_t, \mathbb{Q}_t)dW_t, \quad \mathbb{Q}_t = \text{Law}(X_t) \tag{3.4}$$

where W_t is a d-dimensional Brownian motion [24].

The simulation of a McKean SDE can be accomplished using the particle method. The idea is to replace $\mathbb{Q}_t$ by the empirical distribution

$$\mathbb{Q}_t^N = \frac{1}{N}\sum_{i=1}^{N} \delta_{X_t^{i,N}} \tag{3.5}$$

with $(X_t^{i,N})_{1\le i\le N}$ being a solution of the N-dimensional system of SDEs

$$dX_t^{i,N} = b(t, X_t^{i,N}, \mathbb{Q}_t^N)dt + \sigma(t, X_t^{i,N}, \mathbb{Q}_t^N)dW_t^i, \quad \mathrm{Law}(X_0^{i,N}) = \mathbb{Q}_0 \tag{3.6}$$

driven by N independent Brownian motions $(W_t^i)_{1\le i\le N}$ [16]. Note that the N particles $(X^{i,N})_{1\le i\le N}$ interact with each other in the sense that one needs to know the positions of all N particles at time t in order to simulate the position of the i-th particle $X^{i,N}$ at time $t+\Delta t$.

In their book [16], Guyon and Henry-Labordère show that given that the SDE flow (3.6) fulfills the so-called propagation of chaos property, the random measure $\mathbb{Q}_t^N$ converges in distribution towards the true, deterministic measure $\mathbb{Q}_t$, i.e.

$$\mathbb{Q}_t^N \xrightarrow{d} \mathbb{Q}_t,$$

which implies

$$\frac{1}{N}\sum_{i=1}^{N} \phi(X_t^{i,N}) \xrightarrow[N\to\infty]{L^1} \int \phi(x)\, d\mathbb{Q}_t(dx) \quad \forall \phi \in C_b. \tag{3.7}$$

This means, for N big enough, we can expect to obtain the fair value for a payoff $\phi(X_T)$, when approximating $\mathbb{E}_{\mathbb{Q}}\left[\phi(X_T)\right]$ with the average payoff associated with the simulated particles, i.e.

$$\frac{1}{N}\sum_{i=1}^{N} \phi(X_T^{i,N}).$$

However, since for model 3.2 the empirical measure $\mathbb{Q}_t^N$ obtained from the positions of the particles is atomic [16], and thus admits no density $p_N(t,S,v)$, the conditional expectation

$$\mathbb{E}_{\mathbb{Q}_t^N}\left[v_t | S_t = x\right] = \frac{\int v\, p_N(t,x,v)dv}{\int p_N(t,x,v)dv} = \frac{\sum_{i=1}^N v_t^{i,N}\delta(S_t^{i,N}-x)}{\sum_{i=1}^N \delta(S_t^{i,N}-x)}$$

is not properly defined, and thus the N interacting particles cannot evolve as

$$\frac{dS_t^{i,N}}{S_t^{i,N}} = l(t, S_t^{i,N}, \mathbb{Q}_t^N)\sqrt{v_t^{i,N}}\,dW_t^i$$

with

$$l(t, S_t, \mathbb{Q}_t) = \frac{\sigma_{\mathrm{loc}}(t, S_t)}{\sqrt{\mathbb{E}_{\mathbb{Q}_t}\left[v_t | S_t\right]}}.$$

Therefore, Guyon and Henry-Labordère suggest to replace the Dirac peaks $\delta(\cdot)$ by regularizing kernels $\delta_{t,N}(\cdot)$ as in the Nadaraya-Watson regression [16]. In this sense, they define

$$l_N(t,x) := \sigma_{\mathrm{loc}}(t,x)\sqrt{\frac{\sum_{i=1}^N \delta_{t,N}(S_t^{i,N}-x)}{\sum_{i=1}^N v_t^{i,N}\delta_{t,N}(S_t^{i,N}-x)}} \tag{3.8}$$

and simulate

$$\frac{dS_t^{i,N}}{S_t^{i,N}} = l_N(t,S_t^{i,N})\sqrt{v_t^{i,N}}dW_t^i. \tag{3.9}$$

The choice of a proper bandwidth of the kernel, is of utmost importance for the efficiency and convergence of the algorithm. Assuming the kernels to be Gaussian

$$\delta_{t,N}(x) := \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2\,h_{t,N}}}$$

we can use Silverman's rule of thumb [31] for the estimation of the optimal bandwidth, i.e.

$$h_{t,N}^{opt} = \left(\frac{4}{3}\right)^{1/5}\sigma_{t,N}\,N^{-1/5} \tag{3.10}$$

where $\sigma_{t,N}$ denotes the standard deviation of the simulated paths $(S_t^{i,N})_{1\le i\le N}$ at time t. In practice the shape of the kernel does not matter much - choosing a quartic kernel can be more efficient as it is faster to compute than an exponential. However, using another kernel, one would have to adjust the estimator for the optimal bandwidth as in [16, Section 11.6.2].

4 A case study: parametric fitting the particle and implied volatility surfaces

We have seen that we are able to obtain exploding ATM skew in short time by decorating the Heston model with a stepping local volatility function. In order to examine if this feature of the SSV model is already sufficient in order to improve the Heston model with regard to fitting short-dated vanilla market, we study short-time shapes of leverage surfaces computed with the particle method. In fact, if there exists a parametric leverage function $H(t,x)$ such that the vanilla market can be fully recovered by a local stochastic volatility model of the form

$$\frac{dS_t}{S_t} = H\left(t,\log\left(\frac{S_t}{S_0}\right)\right)\sqrt{v_t}\,dW_t,$$

where v_t is a pre-calibrated Heston variance process, the particle method would recover this function. In particular: If the proposed SSV dynamics were consistent with the market, the particle method would recover a step function or something similar to it.

As we know that the ATM skew produced by the Heston model tends towards a constant, it is reasonable to expect a singularity ATM in the true leverage function, which causes the skew explosion in the corresponding LSV model, regardless whether the specific shape is a step function or not. However, since our numerical procedure involves several steps that should smooth down such a singularity (e.g. Gaussian kernels, parameterization of the implied volatility surface), we cannot expect to fully recover the singularity numerically, but one can hope to find persistent short-time shapes that indicate such an ATM irregularity.

Figure 6 shows the short-end of the particle approximation of the leverage surface corresponding to the Euro Stoxx 50 as of 04-June 2018. As one can see in Figure 10, we are able to perfectly reproduce the vanilla market using the computed particle approximation of the leverage surface. In order to get a better idea of the recovered shapes for small t, seen as a function of the asset price, we plot two cross-sections of the computed leverage surface in Figure 7 .

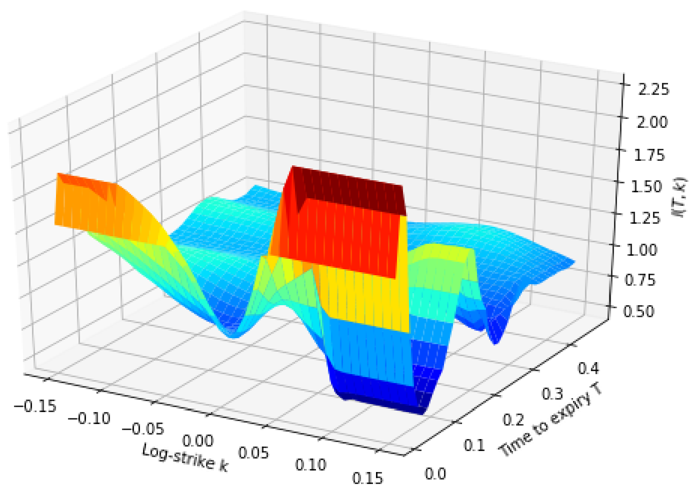

Figure 6: Particle approximation ($N = 2 \cdot 10^5$ particles) of the leverage function in short-end: Euro Stoxx 50 as of 04-June-2018. Heston parameters: $\sqrt{v_0} = 0.13, \kappa = 3.83, \eta = 0.76, \rho = -0.69, \sqrt{\bar{v}} = 0.18$

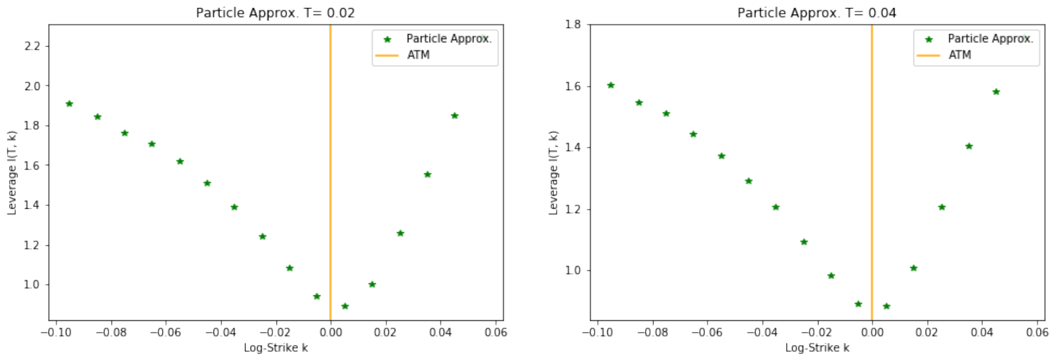

Figure 7: Cross-sections of the particle approximation ($N = 2 \cdot 10^5$ particles) of the Euro Stoxx 50 leverage surface as of 04-June-2018.

As one can see, the short-term shapes clearly differ from a step function. Instead, a power law like function of the form

$$H(x) := \sigma_0 + \gamma |x|^{\sigma_+ 1_{x \geq 0} + \sigma_- 1_{x<0}}$$

with $\sigma_0, \gamma > 0$ and $\sigma_+, \sigma_- \in (0, 1)$ seems to be a much better choice to capture the recovered short-term shapes (see figure 8).

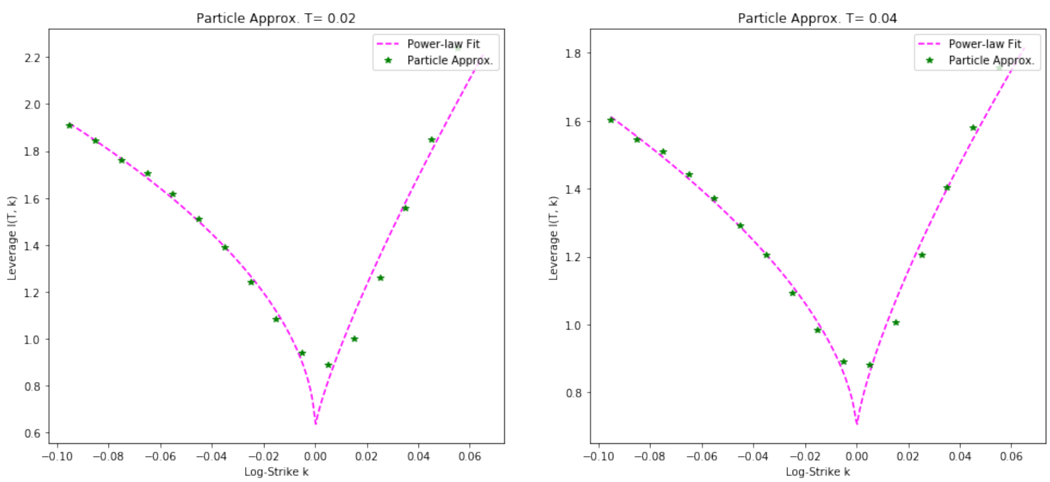

Figure 8: Cross-sections of the particle approximation ($N = 2 \cdot 10^5$ particles) of the Euro Stoxx 50 leverage surface as of 04-June-2018.

In order to investigate the ATM skew behavior generated by a leverage function of the discovered power law like form, we consider a parametric local volatility model of the following form

$$\frac{dS_t}{S_t} = H\left(\log\left(\frac{S_t}{S_0}\right)\right) dW_t \tag{4.1}$$

with

$$H(x) := \sigma_0 + \gamma |x|^{\sigma_+ 1_{x\geq 0} + \sigma_- 1_{x<0}}.$$

We find that the step function in the exponent does not only add the ability to capture non-symmetric shapes, but also produces and controls the short-term ATM skew explosion of the implied volatility corresponding to model 4.1. In fact, given that $\sigma_+ \neq \sigma_-$, the ATM skew generated by model 4.1 explodes like a power-law, which is in line with the skew explosion we observe in the market. Moreover, while for $\sigma_+ > \sigma_-$ the model generates negative skew, we obtain positive skew for $\sigma_+ < \sigma_-$.

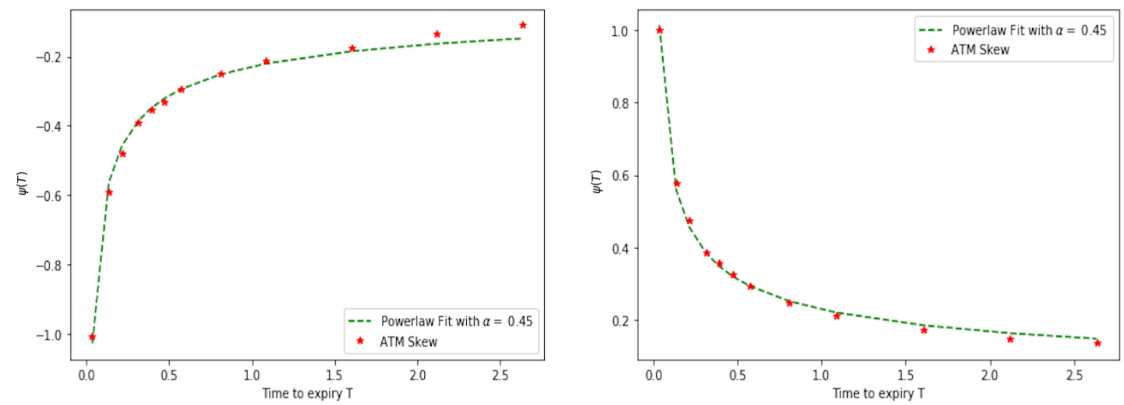

Figure 9: Implied volatility skew computed on Monte Carlo simulations of model 4.1. Left: $\sigma_+ = 0.7$, $\sigma_- = 0.5$, Right: $\sigma_+ = 0.5$, $\sigma_- = 0.7$. Moreover: $\sigma_0 = 0.1$ and $\gamma = 1$. The green dashed line is a power-law fit ($CT^{-\alpha}$).

If the recovered power law like short-term shapes turn out to be stylized facts of leverage surfaces corresponding to LSV models with Heston SV backbone, a model of the form

$$\begin{aligned}\frac{dS_t}{S_t} &= H\left(\log\left(\frac{S_t}{S_0}\right)\right) \sqrt{v_t}\,(\rho dW_t + \sqrt{1-\rho^2} dZ_t) \\ dv_t &= \kappa(\bar{v} - v_t)dt + \eta\sqrt{v_t} dW_t\end{aligned} \tag{4.2}$$

with

$$H(x) := \sigma_0 + \gamma |x|^{\sigma_+ 1_{x\geq 0} + \sigma_- 1_{x<0}}$$

could be a well-suited candidate for a parsimonious extension of the Heston model which is capable of improving the Heston model with regard to fitting the short dated vanilla market. For the pricing of contracts with longer maturities, however, one would have to introduce some additional time-dependency that smoothes down the parametric leverage function as time increases, in order to account for the fact that the leverage surface is usually flattening out over time (see e.g. Figure 6).

In order to examine the persistency of the discovered power law-like shapes, we run our numerics for another Euro Stoxx 50 data set as of 26-July-2018, where we find the power law shapes to be persistent. The computed particle approximation of the leverage surface corresponding to the Euro Stoxx 50 data as of 26-July-2018, as well as cross-sections of the short-end of this surface, can be found in the appendix.

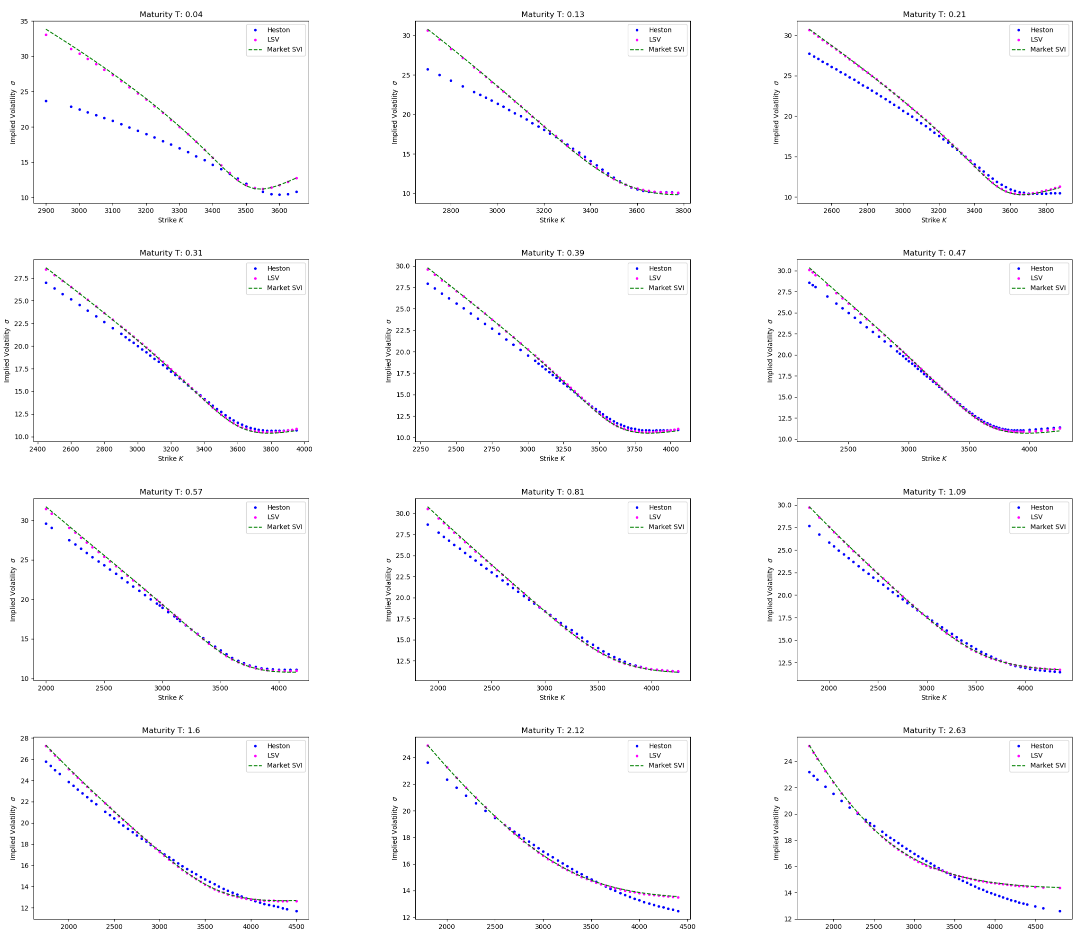

Figure 10: Implied volatilities of the Euro Stoxx 50 as of 04-June-2018 for different maturities. Heston model parameters: $\sqrt{v_0} = 0.13, \kappa = 3.83, \eta = 0.76, \rho = -0.69, \sqrt{\bar{v}} = 0.18$. The blue dots are implied volatilities computed on a Monte Carlo simulation of the Heston model. The magenta dots represent implied volatilities computed on a Monte Carlo simulation of the LSV model. The green dashed line represents the arbitrage-free SVI parameterization of market implied volatilities.

A Appendix

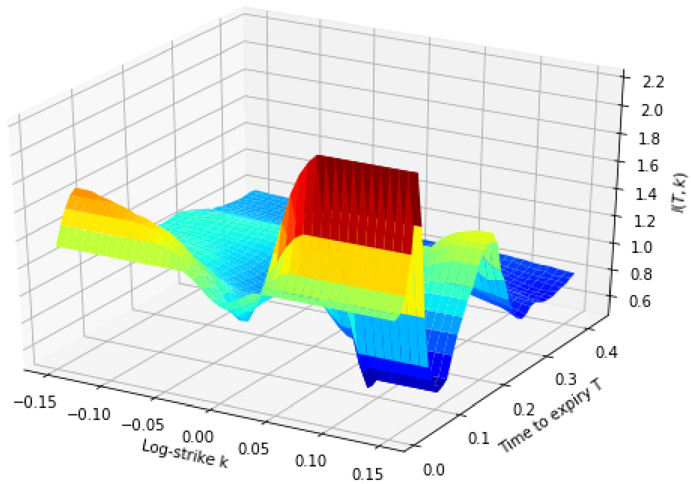

Figure 11: Particle approximation ($N = 2\cdot 10^5$ particles) of the leverage function in short-end: Euro Stoxx 50 as of 26-July-2018. Heston parameters: $\sqrt{v_0} = 0.08, \kappa = 3.99, \eta = 0.81, \rho = -0.68, \sqrt{\bar{v}} = 0.19$

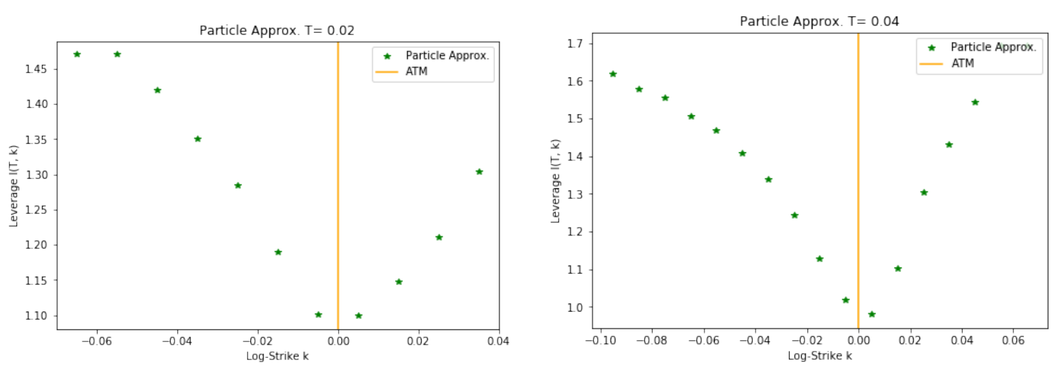

Figure 12: Cross-sections of the particle approximation ($N = 2 \cdot 10^5$ particles) of the Euro Stoxx 50 leverage surface as of 26-July-2018.

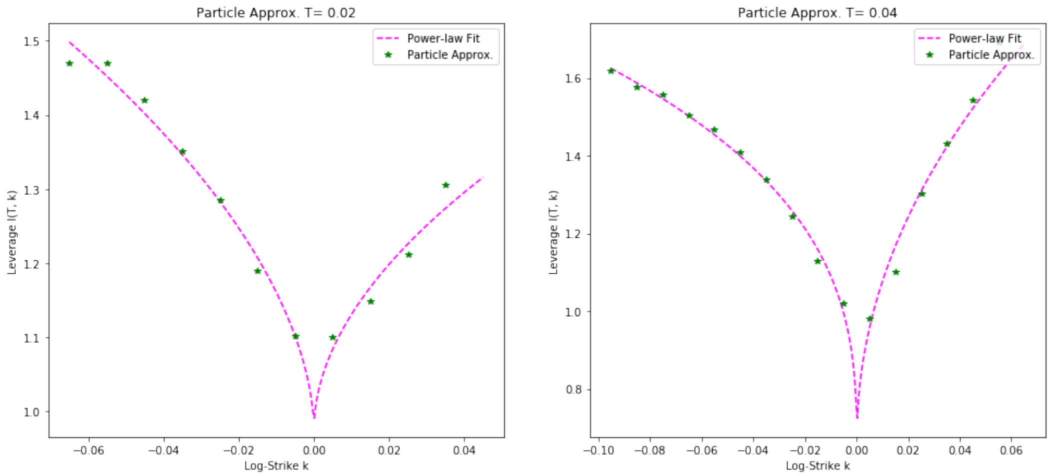

Figure 13: Cross-sections of the particle approximation ($N = 2 \cdot 10^5$ particles) of the Euro Stoxx 50 leverage surface as of 26-July-2018.

References

[1] C. Bayer, P. K. Friz, and J. Gatheral. Pricing under rough volatility. *Quantitative Finance*, 16(6):887–904, 2016.

[2] H. Berestycki, J. Busca, and I. Florent. Asymptotics and calibration of local volatility models. *Quantitative Finance*, 2(1):61–69, 2002.

[3] L. Bergomi. *Stochastic Volatility Modeling*. Financial Mathematics, Chapman and Hall/CRC, 2016.

[4] T. R. A. Corns and S. E. Satchell. Skew Brownian Motion and Pricing European Options. *The European Journal of Finance*, 13(6):523–544, 2007.

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

[6] J. E. Figueroa-López and S. Olafsson. Short-time asymptotics for the implied volatility skew under a stochastic volatility model with Lévy jumps. *Finance and Stochastics*, 20(4):973–1020, 2016.

[7] P. K. Friz, S. Gerhold, and M. Yor. How to make Dupire's local volatility work with jumps. *Quantitative Finance, 14(8):1327–1331, 2014. Special Issue: Themed Issue on Financial Models with Jumps. ISSN: 1469-7688, Publisher: Taylor & Francis*, 2014.

[8] M. Fukasawa. Normalization for Implied Volatility. *arXiv e-prints*, page arXiv:1008.5055, Aug 2010.

[9] M. Fukasawa. Short-time at-the-money skew and rough fractional volatility. *Quantitative Finance*, 17(2):189–198, 2017.

[10] M. Fukasawa. Volatility has to be rough. *arXiv e-prints*, page arXiv:2002.09215, February 2020.

[11] A. Gairat and V. Shcherbakov. Density of Skew Brownian motion and its functionals with application in finance. *Mathematical Finance*, 27(4):1069–1088, Oct 2017.

[12] J. Gatheral. *The Volatility Surface, A Practitioner's Guide*. Wiley, 2006.

[13] S. Gerhold, I. C. Gülüm, and A. Pinter. Small-maturity asymptotics for the at-the-money implied volatility slope in lévy models. *Applied Mathematical Finance*, 23(2):135–157, 2016.

[14] J. Guyon. Path-dependent volatility. *Risk Technical Paper, Longer version available at ssrn.com/abstract=2425048*, 2014.

[15] J. Guyon and P. Henry-Labordère. Being particular about calibration. *Risk Magazine*, pages 88–93, 1 2012.

[16] J. Guyon and P. Henry-Labordère. *Nonlinear Option Pricing*. Chapman and Hall/CRC Financial Mathematics Series, 2013.

[17] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.

[18] J. Keilson and J. A. Wellner. Oscillating Brownian motion. *J. Appl. Probability*, 15(2):300–310, 1978.

[19] R. W. Lee. Implied volatility: statics, dynamics, and probabilistic interpretation. In *Recent advances in applied probability*, pages 241–268. Springer, New York, 2005.

[20] A. Lipton. The vol smile problem. *Risk Magazine*, pages 61–65, 2002.

[21] A. Lipton. Oscillating Bachelier and Black-Scholes Formulas. In *Financial Engineering*. World Scientific, 2018.

[22] A. Lipton and A. Sepp. Filling the gaps. *Risk Magazine*, pages 66–71, 10 2011.

[23] R. McCrickerd and M. S. Pakkanen. Turbocharging Monte Carlo pricing for the rough Bergomi model. *Quantitative Finance*, 18(11):1877–1886, 2018.

[24] H. P. McKean. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci. USA*, 56(6):1907–1911, 1966.

[25] A. Mijatović and P. Tankov. A new look at short-term implied volatility in asset price models with jumps. *Mathematical Finance*, 26(1):149–183, 2016.

[26] E. Neuman and M. Rosenbaum. Fractional Brownian motion with zero Hurst parameter: a rough volatility viewpoint. *Electron. Commun. Probab.*, 23:12 pp., 2018.

[27] P. Pigato. Extreme at-the-money skew in a local volatility model. *Finance and Stochastics*, 23:827–859, 2019.

[28] V. Piterbarg. Markovian projection method for volatility calibration. *Risk Magazine*, pages 84–89, 2007.

[29] A. Reghai. Model evolution. Presentation at the Parisian Model Validation seminar, available at https://sites.google.com/site/projeteuclide/les-seminaires-vmf/archives-vmf, February, 2011.

[30] D. Rossello. Arbitrage in skew Brownian motion models. *Insurance Math. Econom.*, 50(1):50–56, 2012.

[31] B.W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, New York, 1986.

[32] V. Vargas, T.-L. Dao, and J.-P. Bouchaud. Skew and implied leverage effect: smile dynamics revisited. *International Journal of Theoretical and Applied Finance*, 18(04):1550022, 2015.