Consistent modeling and efficient pricing of VIX options

A. Barletta
Department of Economics and Business Economics
Aarhus University

London, 14 March 2016

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n$^\circ$289032.

HPCFinance

AARHUS UNIVERSITY

Introduction

- The Volatility Index (VIX) is a measure of market volatility solely based on OTM option prices on the S&P 500.
- It is determined as the square root of the value of a portfolio that replicates the "fair" price of a 30-days ($=\tau$) variance swap.
- The continuous time version of the VIX is the square root of the log-price conditional annualized quadratic variation

$$\begin{aligned}\mathrm{VIX}_t^2 &= \frac{1}{\tau} E_t \left[[X]_t^{t+\tau} \right] \\ &:= \frac{1}{\tau} \lim_{N \to +\infty} E_t \left[\sum_{n=0}^{N} \left(\log \left(S_{t+n\frac{\tau}{N}} / S_{t+(n-1)\frac{\tau}{N}} \right) \right)^2 \right].\end{aligned}$$

Modeling VIX under stochastic volatility

In our setting today's price of a VIX call option with strike K and maturity T is always

$$C(K, T) = e^{-rT} E\left[(\mathrm{VIX}_T - K)^+\right],$$

therefore we model the risk-neutral dynamics of the VIX.

- The VIX has no explicit dependence on the stock price, but they can be linked through the instantaneous variance v.
- A starting point is the Heston model

$$\begin{aligned} d\log(S_t) &= \left(r - \frac{v_t}{2}\right) dt + \sqrt{v_t}\, dZ_t \\ dv_t &= \kappa(\theta - v_t)\, dt + \epsilon\sqrt{v_t}\, dW_t, \\ d\langle Z, W\rangle_t &= \rho\, dt. \end{aligned}$$

Advantages of the Heston model

- We can express VIX_t^2 as an affine function of v_t

$$\mathrm{VIX}_t = 100 \cdot \sqrt{a v_t + b}.$$

- The density of v_t is available in closed-form.

A pitfall of the Heston model

It is known that the Heston model fails in matching the upward sloping skew of VIX implied volatilities.

Possible generalizations

- Adding jumps in stock and volatility (AJD framework).
  - The affine relation between VIX^2 and v is preserved.
  - The Laplace transform of v is known in semi-closed form.
  - The moments of v can be determined very efficiently by the matrix-based technique of Cuchiero et al. (2010).
- Introducing additional volatility factors (double Heston, double mean reverting, etc.).
  - The affine relation between VIX^2 and v is preserved as long as all the volatility factors have linear drift.
  - Remaining in pure-diffusion allows for passing through PDEs.

SUBPROJECT 1 - PRICING VIX UNDER AJD

THE AJD FRAMEWORK

Under the affine jump-diffusion (AJD) framework we have

$$
\begin{aligned}
d\log(S_t) &= \left(r - \lambda\bar{\mu} - \frac{1}{2}v_t\right)dt + \sqrt{v_t}\,dZ_t + dJ_t^X, \\
dv_t &= \kappa(\theta - v_t)dt + \epsilon\sqrt{v_t}\,dW_t + dJ_t^v, \\
d\langle Z, W\rangle_t &= \rho\, dt,
\end{aligned}
$$

where $J = (J^X, J^v)$ is a pure jump process taking values in $\mathbb{R} \times \mathbb{R}^+$ and $\bar{\mu}$ is the compensator of J^X.

Orthogonal polynomial expansions

Denoting by f the risk-neutral density related to v we approximate f as follows

$$f(x) \approx \phi(x)\left(1 + \sum_{k=0}^{n} c_k h_k(x)\right)$$

where ϕ (kernel) represents a "tractable initial guess" for f and

- h_k are polynomials that only depend on ϕ,
- c_k are corrective factors embedding all the information on f.

The polynomials $(h_k)_{n\in\mathbb{N}}$ are chosen to be orthogonal with respect to ϕ, i.e.

(i) $\mu_k := \int_{-\infty}^{+\infty} x^k \phi(x)\,dx < +\infty, \quad \forall k \in \mathbb{N}$

(ii) $\int_{-\infty}^{+\infty} h_k(x) h_\ell(x) \phi(x)\,dx = 0, \quad \forall k \neq \ell$

Why using orthogonal polynomials?

If $(h_k)_{k\in\mathbb{N}}$ are orthogonal to ϕ then c_k are linear combinations of moments of f

$$c_k = \frac{1}{C_k} \sum_{i=0}^{k} w_{ki} \int_{-\infty}^{+\infty} x^i f(x)\,dx.$$

with $h_k(x) = \sum_{i=0}^{k} w_{ki} x^i$ and $C_k = \sqrt{\int_{-\infty}^{+\infty} h_k^2(x)\phi(x)\,dx}$

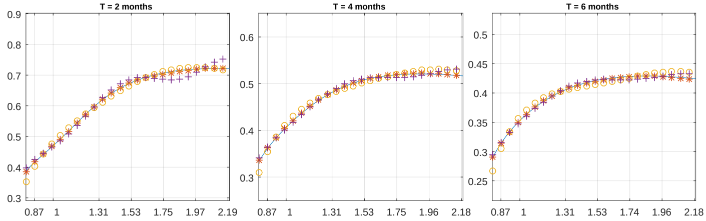
The eGIG kernel

We consider an enriched generalized inverse Gaussian (eGIG) kernel

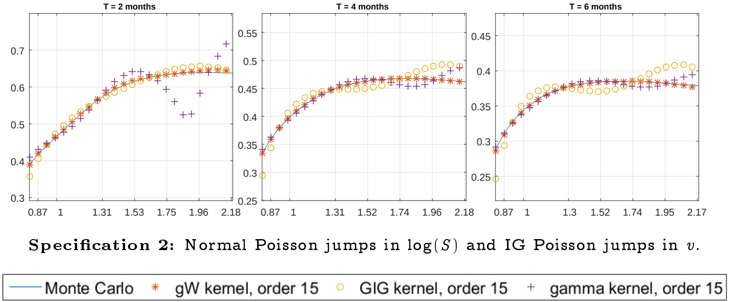
$$\phi(x) = x^{\alpha-1} e^{-\left(\beta x^{p} + \gamma x^{-1}\right)} 1_{[0,+\infty)}(x), \quad \alpha, \beta, \gamma > 0,\ p \in (\frac{1}{2}, 1].$$

We focus on three subcases:

- the gamma ($p = 1, \gamma = 0$) kernel
- the GIG ($p = 1$) kernel
- the gW ($\gamma = 0$) kernel



Specification 1: Normal Poisson jumps in $\log(S)$ and exponential Poisson jumps in v.



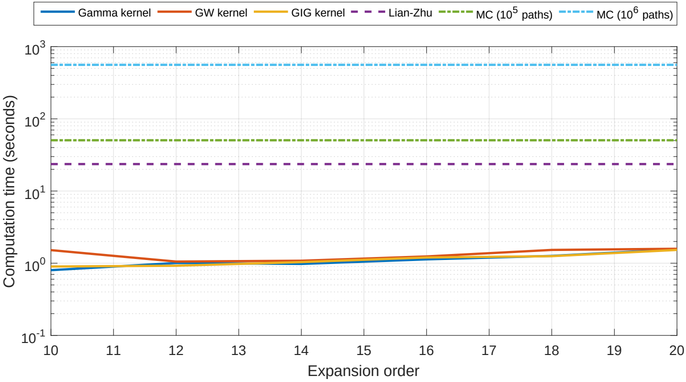
Specification 2: Normal Poisson jumps in $\log(S)$ and IG Poisson jumps in v.

Figure: VIX implied volatilities/moneyness under AJD with Poisson jumps.

Efficiency

Under Specification 2 we compare

- Our technique (orders ranging from 10 to 20).
- Laplace integration optimized formula of Lian and Zhu (2013).
- Monte Carlo with Euler scheme on a time grid of 10^3 points/month.



Subproject 2 - Pricing VIX under multi-factor pure diffusion

We consider the following multi-factor SV model

$$d\log(S_t) = \left(r - \frac{1}{2}\sum_{i=1}^{q} v_t^i\right) dt + \sum_{i=1}^{q} \sqrt{v_t^i}\, dZ_t^i,$$
$$dv_t = (\alpha v_t + \beta)\, dt + \eta(t, v_t)\, dW_t, \quad v_t \in \mathbb{R}^m, \ m \geq q,$$

where $\alpha \in \mathbb{R}^{q\times q}$, $\beta \in \mathbb{R}_+^q$, and $\eta = \mathrm{diag}(\eta_1, \ldots, \eta_q) \in \mathbb{R}^{q\times q}$ is a diagonal matrix of functions.

If Z have no correlation we have

$$\mathrm{VIX}_t = 100 \cdot \sqrt{\langle a, v_t\rangle + b}$$

for some $a \in \mathbb{R}^m$, $b \in \mathbb{R}$.

AARHUS UNIVERSITY

Some history:

- Hagan and Woodward (1999) found asymptotics for the implied volatilities under a CEV model based on a PDE-technique.
- Lorig et al. (2015) have recently extended this technique to a broader pure-diffusion setting.
- Based on this, Pagliarani and Pascucci (2016) have extended these results to include the implied volatility sensitivities.
- We extend these results to a context where the underlying (the VIX futures) dynamics are not explicit.
- All these results apply to short-time and near ATM options.

Examples

Introducing an elasticity coefficient in the vol-of-vol

$$d\log(S_t) = -\frac{1}{2}v_t\,dt + \sqrt{v_t}\,dW_t^X$$
$$dv_t = k\left(v_t - \theta^h\right)dt + \epsilon^h v_t^\delta\,dW_t^Y$$

We have

$$\lim_{\substack{K\to \mathrm{VIX}_0\\ T\to 0}} \frac{\partial}{\partial K}\sigma^{\mathrm{imp}}(T,K) = \frac{\epsilon v_0^{\delta-\frac{1}{2}}\left(2\delta \mathrm{VIX}_0^2 - 2av_0\right)}{4\mathrm{VIX}_0^2\sqrt{v_0}}.$$

The slope is controlled by

$$2\delta\mathrm{VIX}_0^2 - 2av_0 = 2(\delta-1)av_0 + 2\delta b$$

Randomizing the vol-of-vol size (SVV model)

$$
\begin{aligned}
d\log(S_t) &= -\frac{1}{2} v_t^1\, dt + \sqrt{v_t^1}\, dZ_t \\
dv_t^1 &= \kappa_1 \left(v_t^1 - \theta_1\right) dt + v_t^2 \sqrt{v_t^1}\, dW_t^1 \\
dv_t^2 &= \kappa_2 \left(v_t^2 - \theta_2\right) dt + \eta\left(v_t^2\right) dW_t^2 \\
d\left\langle W^1, W^2 \right\rangle_t &= \rho\, dt
\end{aligned}
$$

We have

$$
\lim_{\substack{K \to \mathrm{VIX}_0 \\ T \to 0}} \frac{\partial}{\partial K} \sigma^{\mathrm{imp}}(T, K) = \frac{v_0^2 \left(\mathrm{VIX}_0^2 - 2 a_1 v_0^1\right)}{4\sqrt{v_0^1}\,\mathrm{VIX}_0^2} + \rho \frac{\eta(v_0^2)}{2 v_0^2}.
$$

Thank you for your attention.