The Smile in Stochastic Volatility Models

Julien Guyon

SOCIETE GENERALE
Corporate & Investment Banking
Global Markets Quantitative Research

Modeling and Managing Financial Risks Conference, Paris
January 13th, 2011

Joint work with Lorenzo Bergomi

julien.guyon@sgcib.com
lorenzo.bergomi@sgcib.com

Outline

- Motivation
- Expansion of the smile at order 2 in vol of vol
- First example: a family of Heston-like models
- Second example: the Bergomi model with 2 factors on the variance curve
- Numerical experiments
- Short-term and long-term asymptotics of the smile
- Rederiving the link between skew and skewness of log-returns
- Conclusion

Motivation

- Consider the following general dynamics for a diffusive stochastic volatility model:

$$\begin{aligned} dX_t &= -\frac{1}{2}\xi_t^t dt + \sqrt{\xi_t^t} dW_t^1, \qquad X_0 = x \\ d\xi_t^u &= \lambda(t, u, \xi_t^\cdot) \cdot dW_t, \qquad \xi_0^u = y^u \end{aligned} \tag{1}$$

- $X_t = \ln S_t$
- $\xi_t^\cdot \equiv (\xi_t^u, t \leq u)$: instantaneous forward variance curve from t onwards. For a given maturity u, ξ^u is a driftless process whose initial value is read on the market prices of variance swap contracts: $\xi_0^u = \frac{d}{du}\left(\hat{\sigma}_u^2 u\right)$, where $\hat{\sigma}_u$ is the implied variance swap volatility for maturity u.
- $\lambda = (\lambda_1, \ldots, \lambda_d)$: volatility of forward instantaneous variances.
- $W = \left(W^1, \ldots, W^d\right)$ is a d-dimensional Brownian motion. The first component of the Brownian motion, W^1, drives the spot dynamics.
- No dividend. Zero rates and repos (for the sake of simplicity)

- **No closed-form formula** is available **for the price of vanilla options** in Model (1).
- In a few particular cases of "first generation" stochastic volatility models, like the Heston model, some approximations of the price of vanilla options have been suggested in the literature.
- Here, we aim at finding a general approximation of the smile of implied volatility **which does not depend on a particular specification of the model**, i.e., on a particular choice of λ.

- $\Rightarrow$ We will derive **general asymptotic expansion of the smile, for small volatility of volatility, at second order**.
- We introduce a scaling factor ω for the volatility of instantaneous forward variance: $\lambda \to \omega\lambda$. Abusing language, we will speak of ω as the "vol of vol." X and ξ then depend on ω: $X \to X^\omega$ and $\xi \to \xi^\omega$.
- Our derivation relies on the fact that in Model (1), the volatility of the asset is an autonomous stochastic process, meaning that it incorporates **no local volatility component**, and that λ **does not depend on the asset value**.

Expansion of the price of a vanilla option

- Consider the vanilla option delivering $g(X_T^\omega)$ at time T.
- Its price is a function P^ω of $(t, X_t^\omega, \xi_t^{\cdot,\omega})$. We write $P^\omega(t,x,y)$: **the variable** $y \equiv (y^u, t \le u \le T)$ **is a curve**.
- P^ω solves the PDE $(\partial_t + L^\omega)\, P^\omega = 0$ with terminal condition $P^\omega(T,x,y) = g(x)$, where $L^\omega = L_0 + \omega L_1 + \omega^2 L_2$ with

$$
\begin{aligned}
L_0 &= -\frac{1}{2} y^t \partial_x + \frac{1}{2} y^t \partial_x^2 \\
L_1 &= \int_t^T du\ \mu(t,u,y)\ \partial^2_{xy^u} \\
L_2 &= \frac{1}{2} \int_t^T du \int_t^T du'\ \nu(t,u,u',y)\ \partial^2_{y^u y^{u'}} \\
\mu(t,u,y) &= \sqrt{y^t}\lambda_1(t,u,y) = \frac{\mathbb{E}\left[dX_t d\xi_t^u | \xi_t = y\right]}{dt} = \frac{\mathbb{E}\left[\frac{dS_t}{S_t} d\xi_t^u | \xi_t = y\right]}{dt} \\
\nu(t,u,u',y) &= \sum_{i=1}^d \lambda_i(t,u,y)\lambda_i(t,u',y) = \frac{\mathbb{E}\left[d\xi_t^u d\xi_t^{u'} | \xi_t = y\right]}{dt}
\end{aligned}
$$

The perturbation equations

- Assume that $P^\omega = P_0 + \omega P_1 + \omega^2 P_2 + \omega^3 P_3 + \cdots$

$$\begin{aligned} 0 &= \left(\partial_t + L_0 + \omega L_1 + \omega^2 L_2\right)\left(P_0 + \omega P_1 + \omega^2 P_2 + \omega^3 P_3 + \cdots\right) \\ &= \left(\partial_t + L_0\right) P_0 + \omega\left(\left(\partial_t + L_0\right) P_1 + L_1 P_0\right) \\ &\quad + \omega^2\left(\left(\partial_t + L_0\right) P_2 + L_1 P_1 + L_2 P_0\right) \\ &\quad + \omega^3\left(\left(\partial_t + L_0\right) P_3 + L_1 P_2 + L_2 P_1\right) + \cdots \end{aligned}$$

- $\Rightarrow$ We need to solve the following equations:

$$\begin{aligned} \left(\partial_t + L_0\right) P_0 &= 0, \quad P_0(T,x,y) = g(x) \\ \left(\partial_t + L_0\right) P_1 + L_1 P_0 &= 0, \quad P_1(T,x,y) = 0 \\ \left(\partial_t + L_0\right) P_n + L_1 P_{n-1} + L_2 P_{n-2} &= 0, \quad P_n(T,x,y) = 0, \qquad \forall n \geq 2 \end{aligned}$$

- L_0 is the infinitesimal generator associated to X^0, the unperturbed diffusion for which $\omega = 0$. It is the standard **one-dimensional Black-Scholes operator** with deterministic volatility $\sqrt{y^t}$ at time t.
- Each P_n is solution to the traditional **one-dimensional** diffusion equation **with a source term** $H_n = L_1 P_{n-1} + L_2 P_{n-2}$:

$$(\partial_t + L_0)\, P_n + H_n = 0$$

- Feynmann-Kac theorem $\Rightarrow$

$$\begin{aligned} P_0(t,x,y) &= \mathbb{E}\left[g\left(X_T^{0,t,x}\right)\right], \\ P_n(t,x,y) &= \mathbb{E}\left[\int_t^T H_n(s, X_s^{0,t,x}, y)ds\right], \qquad \forall n \geq 1 \end{aligned}$$

where $X^{0,t,x}$ is the unperturbed process where $\omega = 0$, starting at log-spot x at time t:

$$dX_s^{0,t,x} = -\frac{1}{2} y^s ds + \sqrt{y^s} dW_s^1, \qquad X_t^{0,t,x} = x$$

The price at order 0

- P_0 is just the Black-Scholes price with time-dependent volatility $\sqrt{y^t}$:

$$P_0(t,x,y) = \mathbb{E}\left[g\left(x + \int_t^T \sqrt{y^s}dW_s^1 - \frac{1}{2}\int_t^T y^s ds\right)\right] = P_{BS}\left(x, \int_t^T y^s ds\right)$$

where

$$P_{BS}(x,v) = \mathbb{E}\left[g\left(x + \sqrt{v}G - \frac{1}{2}v\right)\right], \qquad G \sim \mathcal{N}(0,1) \tag{2}$$

- $v = \int_t^T y^s ds$ is the total variance of X^0 integrated from t to T.
- $P_0(t,x,y)$ depends on the curve $y \equiv (y^s, t \leq s \leq T)$ only through v.
- P_{BS} is solution to the PDE

$$\partial_v P_{BS} = \frac{1}{2}\left(\partial_x^2 - \partial_x\right) P_{BS}, \qquad P_{BS}(x,0) = g(x) \tag{3}$$

Links the vega and gamma of a vanilla option in the unperturbed state.

The price at order 0

An important observation:

- Because L_0 incorporates no local volatility, L_0 and ∂_x commute so $(\partial_t + L_0)\, \partial_x^p P_0 = \partial_x^p \, (\partial_t + L_0)\, P_0 = 0$.
- $\Rightarrow \partial_x^p P_{BS} \left(X_t^0, \int_t^T y^s ds \right) \equiv \partial_x^p P_0(t, X_t^0, y)$ is a martingale for all integer p.
- Equation (3) then shows that **for all integers** m, n, $\partial_v^m \partial_x^n P_{BS} \left(X_t^0, \int_t^T y^s ds \right)$ **is a martingale**.
- This is crucial in the computations of P_1 and P_2.

The price at order 1

- Let us define the **integrated spot-variance covariance** function $C_t^{X\xi}(y)$:

$$C_t^{X\xi}(y) = \int_t^T ds \int_s^T du \;\mu(s,u,y) = \int_t^T ds \int_s^T du \;\frac{\mathbb{E}\left[\frac{dS_s}{S_s} d\xi_s^u | \xi_s = y\right]}{ds}$$

- We then have

$$\begin{aligned}
P_1(t,x,y) &= \mathbb{E}\left[\int_t^T L_1 P_0(s, X_s^{0,t,x}, y) ds\right] \\
&= \mathbb{E}\left[\int_t^T ds \int_s^T du \;\mu(s,u,y)\; \partial_{y^u}\left(\partial_x P_{BS}\left(X_s^{0,t,x}, \int_s^T y^r\, dr\right)\right)\right] \\
&= \mathbb{E}\left[\int_t^T ds \int_s^T du \;\mu(s,u,y)\; \partial^2_{xv} P_{BS}\left(X_s^{0,t,x}, \int_s^T y^r\, dr\right)\right] \\
&= \int_t^T ds \int_s^T du \;\mu(s,u,y) \mathbb{E}\left[\;\partial^2_{xv} P_{BS}\left(X_s^{0,t,x}, \int_s^T y^r\, dr\right)\right] \\
&= C_t^{X\xi}(y)\; \partial^2_{xv} P_{BS}\left(x, \int_t^T y^r\, dr\right)
\end{aligned}$$

The price at order 2

A similar result holds for the second order correction:

$$
\begin{aligned}
P_2 &= P_2^{L_2 P_0} + P_2^{L_1 P_1} \\
P_2^{L_2 P_0}(t,x,y) &= \frac{1}{2} C_t^{\xi\xi}(y)\ \partial_v^2 P_{BS}\left(x, \int_t^T y^r\,dr\right) \\
P_2^{L_1 P_1} &= P_{2,0}^{L_1 P_1} + P_{2,1}^{L_1 P_1} \\
P_{2,0}^{L_1 P_1}(t,x,y) &= \frac{1}{2} C_t^{X\xi}(y)^2\ \partial_x^2 \partial_v^2 P_{BS}\left(x, \int_t^T y^r\,dr\right) \\
P_{2,0}^{L_1 P_1}(t,x,y) &= C_t^{\mu}(y)\ \partial_x^2 \partial_v P_{BS}\left(x, \int_t^T y^r\,dr\right)
\end{aligned}
$$

$$
\begin{aligned}
C_t^{\xi\xi}(y) &= \int_t^T ds \int_s^T du \int_s^T du'\ \nu(s,u,u',y) = \int_t^T ds \int_s^T du \int_s^T du'\ \frac{\mathbb{E}\left[d\xi_s^u\, d\xi_t^{u'} | \xi_s = y\right]}{ds} \\
C_t^{\mu}(y) &= \int_t^T ds \int_s^T du\ \mu(s,u,y)\ \partial_{y^u}\left(C_s^{X\xi}(y)\right)
\end{aligned}
$$

$C_t^{\xi\xi}(y)$: **integrated variance-variance covariance** function

Expansion of the implied volatility

- We write $C^{X\xi} = C_0^{X\xi}(y)$, $C^{\xi\xi} = C_0^{\xi\xi}(y)$ and $C^{\mu} = C_0^{\mu}(y)$.
- **In the general diffusive stochastic volatility model (1), at second order in the vol of vol ω, the implied volatility for maturity T and strike K is quadratic in $L = \ln\left(\frac{K}{S_0}\right)$:**

$$I^{\omega}(T,K) = I_T^{ATM} + \mathcal{S}_T \ln\left(\frac{K}{S_0}\right) + \kappa_T \ln^2\left(\frac{K}{S_0}\right) + O(\omega^3) \qquad (4)$$

- Coefficients are

$$I_T^{ATM} = \sqrt{\frac{v}{T}} + \frac{C^{X\xi}}{4\sqrt{vT}}\omega + \frac{1}{32v^{5/2}\sqrt{T}}\left(12C^{X\xi\,2} - C^{\xi\xi}v\,(v+4) + 4C^{\mu}v\,(v-4)\right)\omega^2$$

$$\mathcal{S}_T = \frac{C^{X\xi}}{2v^{3/2}\sqrt{T}}\omega + \frac{1}{8v^{5/2}\sqrt{T}}\left(4C^{\mu}v - 3C^{X\xi\,2}\right)\omega^2$$

$$\kappa_T = \frac{1}{8v^{7/2}\sqrt{T}}\left(4C^{\mu}v + C^{\xi\xi}v - 6C^{X\xi\,2}\right)\omega^2$$

- $v = \int_0^T y^r\,dr$ = integrated variance.

Comments

ATM implied volatility:

$$I_T^{ATM} = \sqrt{\frac{v}{T}} + \frac{C^{X\xi}}{4\sqrt{vT}}\omega + \frac{1}{32v^{5/2}\sqrt{T}}\left(12C^{X\xi 2} - C^{\xi\xi}v\,(v+4) + 4C^{\mu}v\,(v-4)\right)\omega^2$$

- ATM implied volatility = variance swap volatility + spread. At first order, spread $= \frac{C^{X\xi}}{4\sqrt{vT}}\omega$.
- Typically, on the equity market, $C^{X\xi} < 0$: the ATM implied volatility lies below the variance swap volatility.
- When spot returns and forward variances are uncorrelated, $C^{X\xi} = C^{\mu} = 0$ so that

  $$I_T^{ATM} = \sqrt{\frac{v}{T}} - \frac{C^{\xi\xi}\,(v+4)}{32v^{3/2}\sqrt{T}}\omega^2$$

  Because $C^{\xi\xi} \geq 0$, the ATM implied volatility lies again below the variance swap volatility. The higher the volatility of variances, the smaller the ATM implied volatility.

Comments (continued)

ATM skew: $\mathcal{S}_T = \frac{C^{X\xi}}{2v^{3/2}\sqrt{T}}\omega + \frac{1}{8v^{5/2}\sqrt{T}}\left(4C^{\mu}v - 3C^{X\xi 2}\right)\omega^2$

- The ATM skew $\mathcal{S}_T$ is of order ω. It has the sign of $C^{X\xi}$. $\mathcal{S}_T$ vanishes when spot returns and forward variances are uncorrelated, even at second order. The skew is produced by the spot-variance correlation.
- If $\mu \leq 0$, $\mathcal{S}_T \leq 0$ at first and at second order.
- At first order in ω, the ATM skew has the same sign as the difference between the ATM implied volatility and the variance swap volatility.

ATM convexity: $\kappa_T = \frac{1}{8v^{7/2}\sqrt{T}}\left(4C^{\mu}v + C^{\xi\xi}v - 6C^{X\xi 2}\right)\omega^2$

- The convexity κ_T is of order ω^2. It is an increasing function of $C^{\xi\xi}$ and of C^{μ}, and a decreasing function of the absolute value of the spot-variance covariance $\left|C^{X\xi}\right|$.
- If spot and variances are uncorrelated,

$$\kappa_T = \frac{C^{\xi\xi}}{8v^{5/2}\sqrt{T}}\omega^2 \geq 0$$

The convexity is positive, and produced by variance-variance correlation.

Another derivation which stays at the level of operators

- Recall that the price P^ω of the vanilla option is solution to $\left(\partial_t + L_t^\omega\right) P^\omega = 0$ with $L_t^\omega = L_{0,t} + \omega L_{1,t} + \omega^2 L_{2,t}$, and terminal condition $P^\omega\left(T,x,y\right) = g(x)$.
- The price can be expressed in terms of the semigroup $(U_{st}^\omega, 0 \leq s \leq t \leq T)$ attached to the family of differential operators L_t^ω: $P^\omega(t,\cdot) = U_{tT}^\omega g$.
- The semigroup is defined by

$$U_{st}^\omega = \lim_{n\to\infty} \left(1 - \delta t L_{t_0}^\omega\right)\left(1 - \delta t L_{t_1}^\omega\right)\cdots\left(1 - \delta t L_{t_{n-1}}^\omega\right), \quad \delta t = \frac{t-s}{n}, \quad t_i = s+i\delta t$$

- It satisfies $U_{rt}^\omega = U_{rs}^\omega U_{st}^\omega$ for $0 \leq r \leq s \leq t \leq T$, hence the notation $:\exp\left(\int_s^t L_\tau^\omega d\tau\right):$, where :: denotes time ordering.
- **We can directly expand U_{st}^ω in powers of ω.** This is the usual time-dependent perturbation technique in quantum mechanics. U_{st}^0 is called the free propagator.

- Consider the general situation where a differential operator L_t is perturbed by another operator H_t: $L_t^\varepsilon = L_t + \varepsilon H_t$
- From the definition of the semigroup, $U_{st}^\varepsilon = U_{st}^{(0)} + \varepsilon U_{st}^{(1)} + \varepsilon^2 U_{st}^{(2)} + \cdots$ with

$$
\begin{aligned}
U_{st}^{(1)} &= \int_s^t d\tau \; U_{s\tau}^0 H_\tau U_{\tau t}^0 \\
U_{st}^{(2)} &= \int_s^t d\tau_1 \int_{\tau_1}^t d\tau_2 \; U_{s\tau_1}^0 H_{\tau_1} U_{\tau_1\tau_2}^0 H_{t_2} U_{\tau_2 t}^0
\end{aligned}
$$

- $\Rightarrow P^\omega = P_0 + \omega P_1 + \omega^2 P_2 + \cdots$, with

$$
\begin{aligned}
P_1 &= \int_t^T d\tau \; U_{t\tau}^0 L_{1,\tau} U_{\tau T}^0 g \\
P_2 &= \int_t^T d\tau \; U_{t\tau}^0 L_{2,\tau} U_{\tau T}^0 g + \int_t^T d\tau_1 \int_{\tau_1}^T d\tau_2 \; U_{t\tau_1}^0 L_{1,\tau_1} U_{\tau_1\tau_2}^0 L_{1,\tau_2} U_{\tau_2 T}^0 g
\end{aligned}
$$

- We recover the expressions of P_1 and P_2. Quicker.

First example: a Heston-like model

$$
\begin{aligned}
dX_t^\omega &= -\frac{1}{2}V_t^\omega dt + \sqrt{V_t^\omega}dW_t^1, \qquad X_0^\omega = x \\
dV_t^\omega &= -k\left(V_t^\omega - v_\infty\right)dt + \omega\left(V_t^\omega\right)^\varphi\left(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2\right), \qquad V_0^\omega = v
\end{aligned} \tag{5}
$$

- The instantaneous forward variance reads

$$\xi_t^{u,\omega} = \mathbb{E}\left[V_u^\omega \middle| V_t^\omega\right] = v_\infty + \left(V_t^\omega - v_\infty\right)e^{-k(u-t)}$$

  and has dynamics

$$d\xi_t^{u,\omega} = \omega e^{-k(u-t)}\left(\xi_t^{t,\omega}\right)^\varphi\left(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2\right)$$

- The initial term-structure of instantaneous forward variances is

$$y^u \equiv \xi_0^u = v_\infty + \left(v - v_\infty\right)e^{-ku}$$

- Like **in all classic “first generation” stochastic volatility models, this term-structure is determined by the model parameters**, and the current value of the instantaneous volatility.

- The volatility $\lambda(t,u,y)$ of instantaneous forward variances depends on the instantaneous forward variance curve $y=(y^s, t\le s\le T)$ only through the instantaneous spot variance y^t:

$$\begin{aligned}\lambda_1(t,u,y) &= \rho\left(y^t\right)^{\varphi} e^{-k(u-t)}\\ \lambda_2(t,u,y) &= \sqrt{1-\rho^2}\left(y^t\right)^{\varphi} e^{-k(u-t)}\end{aligned}$$

- As a consequence,

$$\begin{aligned}C^{X\xi} &= \frac{\rho}{k}\int_0^T ds\ \left(y^s\right)^{\varphi+\frac{1}{2}}\left(1-e^{-k(T-s)}\right)\\ C^{\xi\xi} &= \sum_{i=1}^{2}\int_0^T ds\left(\int_s^T du\,\lambda_i(s,u,y)\right)^2=\frac{1}{k^2}\int_0^T ds\ \left(y^s\right)^{2\varphi}\left(1-e^{-k(T-s)}\right)^2\\ C^{\mu} &= \left(\varphi+\frac{1}{2}\right)\frac{\rho^2}{k}\int_0^T ds\ \left(y^s\right)^{\varphi+\frac{1}{2}}\int_s^T du\ \left(y^u\right)^{\varphi-\frac{1}{2}} e^{-k(u-s)}\left(1-e^{-k(T-u)}\right)\end{aligned}$$

- This coincides with Equations (3.7) to (3.10) in Lewis [5], where $J^{(1)}=C^{X\xi}$, $J^{(3)}=\frac{1}{2}C^{\xi\xi}$, and $J^{(4)}=C^{\mu}$

Second example: the Bergomi model

$$
\begin{aligned}
dX_t^\omega &= -\frac{1}{2}\xi_t^{t,\omega}dt + \sqrt{\xi_t^{t,\omega}}dW_t^S \\
d\xi_t^{u,\omega} &= \omega\xi_t^{u,\omega}\alpha_\theta\left((1-\theta)\,e^{-k_X(u-t)}dW_t^X + \theta e^{-k_Y(u-t)}dW_t^Y\right) \\
&= \omega\lambda(t,u,\xi_t^{\cdot,\omega})\cdot dW_t
\end{aligned}
$$

$$d\langle W^S, W^X\rangle_t = \rho_{SX}dt,\ d\langle W^S, W^Y\rangle_t = \rho_{SY}dt,\ d\langle W^X, W^Y\rangle_t = \rho_{XY}dt.$$

- The normalizing factor

$$\alpha_\theta = \left((1-\theta)^2 + 2\rho_{XY}\theta\,(1-\theta) + \theta^2\right)^{-1/2}$$

  is such that the very-short term variance $\xi_t^{t,\omega}$ has log-normal volatility ω.
- We pick $k_X > k_Y$, θ is a parameter which mixes the short-term factor W^X and the long-term factor W^Y.

- After a Cholesky transform, this can be restated using independent Brownian motions W^1, W^2 and W^3 as follows:

$$
\begin{aligned}
W^S &= W^1 \\
W^X &= \rho_{SX} W^1 + \sqrt{1-\rho_{SX}^2} W^2 \\
W^Y &= \rho_{SY} W^1 + \chi_{XY} \sqrt{1-\rho_{SY}^2} W^2 + \sqrt{\left(1-\chi_{XY}^2\right)\left(1-\rho_{SY}^2\right)} W^3
\end{aligned}
$$

where $\chi_{XY} = \frac{\rho_{XY} - \rho_{SX}\rho_{SY}}{\sqrt{1-\rho_{SX}^2}\sqrt{1-\rho_{SY}^2}}$

- ρ_{SX}, ρ_{SY} and ρ_{XY} define a correlation matrix $\iff \chi_{XY} \in [-1, 1]$.
- The volatility of variance $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ reads

$$
\begin{aligned}
\lambda_1(t,u,y) &= y^u \alpha_\theta \left((1-\theta) \rho_{SX} e^{-k_X(u-t)} + \theta \rho_{SY} e^{-k_Y(u-t)} \right) \\
\lambda_2(t,u,y) &= y^u \alpha_\theta \left((1-\theta) \sqrt{1-\rho_{SX}^2} e^{-k_X(u-t)} + \theta \chi_{XY} \sqrt{1-\rho_{SY}^2} e^{-k_Y(u-t)} \right) \\
\lambda_3(t,u,y) &= y^u \alpha_\theta \theta \sqrt{\left(1-\chi_{XY}^2\right)\left(1-\rho_{SY}^2\right)} e^{-k_Y(u-t)}
\end{aligned}
$$

- We write $\lambda_i(t,u,y) = y^u \alpha_\theta \left(w_{iX} e^{-k_X(u-t)} + w_{iY} e^{-k_Y(u-t)} \right)$

The covariance functions read

$$
\begin{aligned}
C^{X\xi} &= \int_0^T du \int_0^u dt \sqrt{y^t} \lambda_1(t,u,y) \\
&= \alpha_\theta \left(1-\theta\right) \rho_{SX} \int_0^T du\ y^u \int_0^u dt \sqrt{y^t} e^{-k_X(u-t)} \\
&\quad + \alpha_\theta \theta \rho_{SY} \int_0^T du\ y^u \int_0^u dt \sqrt{y^t} e^{-k_Y(u-t)} \\
C^{\xi\xi} &= \sum_{i=1}^{3} \int_0^T ds \left(\int_s^T du\, \lambda_i(s,u,y) \right)^2 \\
&= \alpha_\theta^2 \sum_{i=1}^{3} \int_0^T ds \left(w_{iX} \int_s^T du\ y^u e^{-k_X(u-s)} + w_{iY} \int_s^T du\ y^u e^{-k_Y(u-s)} \right)^2 \\
C^{\mu} &= \int_0^T ds \int_s^T du \sqrt{y^s} \lambda_1(s,u,y^u) \left(\frac{1}{2\sqrt{y^u}} \int_u^T dt \lambda_1(u,t,y^t) \right. \\
&\quad \left. + \int_s^u dr \sqrt{y^r} \frac{\partial \lambda_1}{\partial z}(r,u,z)|_{z=y^u} \right)
\end{aligned}
$$

In the case of a flat initial term structure of variance swaps ($y_0^t \equiv \xi$), this reads

$$\begin{aligned}
C^{X\xi} &= \alpha_\theta \xi^{3/2} \left(w_{1X} \mathcal{J}\left(k_X, T\right) + w_{1Y} \mathcal{J}\left(k_Y, T\right)\right) \\
C^{\xi\xi} &= \alpha_\theta^2 \xi^2 (w_0 T + w_X \mathcal{I}(k_X, T) + w_Y \mathcal{I}(k_Y, T) + w_{XX} \mathcal{I}(2k_X, T) \\
&\quad + w_{YY} \mathcal{I}(2k_Y, T) + w_{XY} \mathcal{I}(k_X + k_Y, T)) \\
C^{\mu} &= \alpha_\theta^2 \xi^2 \left(\frac{1}{2} A_1 + A_2 \right)
\end{aligned}$$

with

$$\begin{aligned}
\mathcal{I}(k, u) &= \int_0^u dt\, e^{-k(u-t)} = \frac{1 - e^{-ku}}{k} \\
\mathcal{J}(k, T) &= \int_0^T du \int_0^u dt\, e^{-k(u-t)} = \frac{kT - 1 + e^{-kT}}{k^2} \\
\mathcal{K}(k, T) &= \int_0^T ds\ (T - s) e^{-k(T-s)} = \int_0^T ds\ s e^{-ks} = \frac{1 - (1 + kT) e^{-kT}}{k^2} \\
w_0 &= \sum_{i=1}^3 \left(\frac{w_{iX}}{k_X} + \frac{w_{iY}}{k_Y} \right)^2, \qquad w_X = -2 \sum_{i=1}^3 \frac{w_{iX}}{k_X} \left(\frac{w_{iX}}{k_X} + \frac{w_{iY}}{k_Y} \right) \\
w_Y &= -2 \sum_{i=1}^3 \frac{w_{iY}}{k_Y} \left(\frac{w_{iX}}{k_X} + \frac{w_{iY}}{k_Y} \right), \qquad w_{XX} = \sum_{i=1}^3 \frac{w_{iX}^2}{k_X^2} \\
w_{YY} &= \sum_{i=1}^3 \frac{w_{iY}^2}{k_Y^2}, \qquad w_{XY} = 2 \sum_{i=1}^3 \frac{w_{iX} w_{iY}}{k_X k_Y}
\end{aligned}$$

and

$$
\begin{aligned}
A_1 &= -\frac{w_{1X}^2}{k_X}\mathcal{K}(k_X,T) - \frac{w_{1Y}^2}{k_Y}\mathcal{K}(k_Y,T) + w'_X\mathcal{J}(k_X,T) + w'_Y\mathcal{J}(k_Y,T) \\
&\quad - \frac{w_{1X}w_{1Y}}{k_X+k_Y}\left(e^{-k_XT}\frac{1-e^{-k_XT}}{k_X^2} + e^{-k_YT}\frac{1-e^{-k_YT}}{k_Y^2}\right. \\
&\quad \left. - \frac{e^{-2k_XT}\left(1-e^{-k_YT}\right) + e^{-2k_YT}\left(1-e^{-k_XT}\right)}{k_Xk_Y}\right) \\
A_2 &= w''_X\mathcal{J}(k_X,T) + w''_Y\mathcal{J}(k_Y,T) + w''_{XX}\mathcal{J}(2k_X,T) + w''_{YY}\mathcal{J}(2k_Y,T) + w''_{XY}\mathcal{J}(k_X+k_Y,T)
\end{aligned}
$$

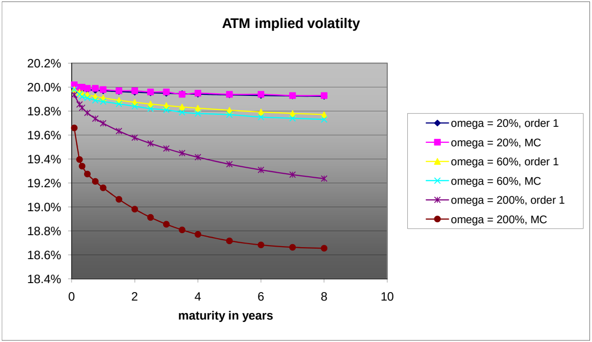
with

$$
\begin{aligned}
w'_X &= \frac{w_{1X}^2}{k_X} + \frac{w_{1X}w_{1Y}}{k_Y}, \qquad w'_Y = \frac{w_{1Y}^2}{k_Y} + \frac{w_{1X}w_{1Y}}{k_X} \\
w''_X &= \frac{w_{1X}^2}{k_X} + \frac{w_{1X}w_{1Y}}{k_Y}, \qquad w''_Y = \frac{w_{1Y}^2}{k_Y} + \frac{w_{1X}w_{1Y}}{k_X} \\
w''_{XX} &= -\frac{w_{1X}^2}{k_X}, \qquad w''_{YY} = -\frac{w_{1Y}^2}{k_Y}, \qquad w''_{XY} = -\frac{w_{1X}w_{1Y}}{k_X} - \frac{w_{1X}w_{1Y}}{k_Y}
\end{aligned}
$$

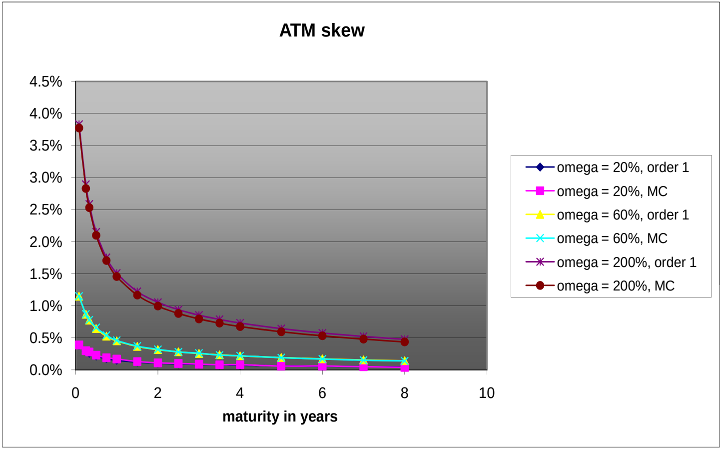
Numerical experiments

We pick the Bergomi model with a flat initial term structure of variance swap prices and

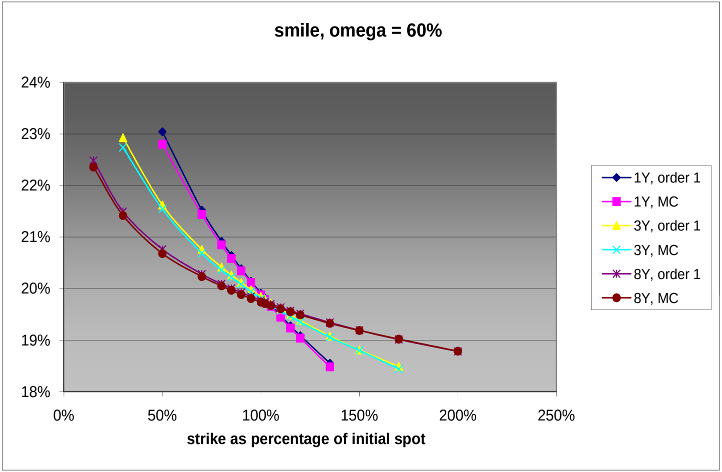
θ	k_X	k_Y	ρ_{SX}	ρ_{SY}	ρ_{XY}	χ_{XY}	ξ
0.25	8	0.35	-0.8	-0.48	0	-0.73	$(0.2)^2$



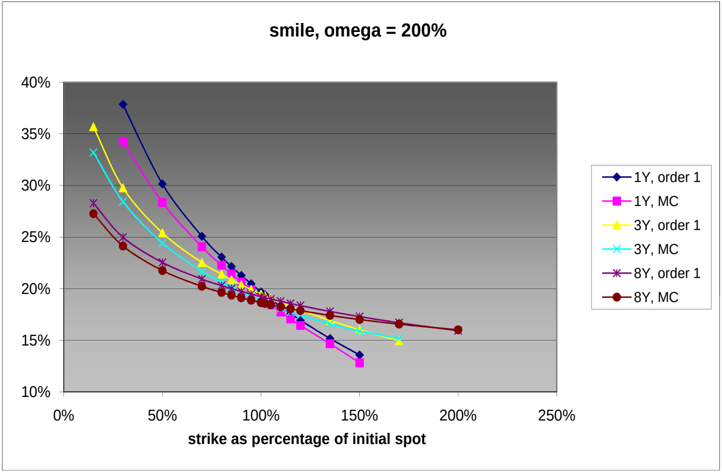
First order


ATM skew
4.5%
4.0%
3.5%
3.0%
2.5%
2.0%
1.5%
1.0%
0.5%
0.0%
0
2
4
6
8
10
maturity in years
omega = 20%, order 1
omega = 20%, MC
omega = 60%, order 1
omega = 60%, MC
omega = 200%, order 1
omega = 200%, MC

First order


smile, omega = 60%
24%
23%
22%
21%
20%
19%
18%
0%
50%
100%
150%
200%
250%
strike as percentage of initial spot
1Y, order 1
1Y, MC
3Y, order 1
3Y, MC
8Y, order 1
8Y, MC

First order


smile, omega = 200%
40%
35%
30%
25%
20%
15%
10%
0%
50%
100%
150%
200%
250%
strike as percentage of initial spot
1Y, order 1
1Y, MC
3Y, order 1
3Y, MC
8Y, order 1
8Y, MC

First order

- The **ATM skew is very sharply estimated by the first order expansion**, even for large values of the volatility of variance ω.
- The **ATM volatility is well captured by the expansion at first order in ω only for small values of ω** (say, up to 60%).
- True ATM implied volatilities are below their first order approximates $\Rightarrow$ the ATM volatility is a very concave function of ω, around $\omega = 0$. In view of the expression for I_T^{ATM}, this means that, for the set of parameters picked,

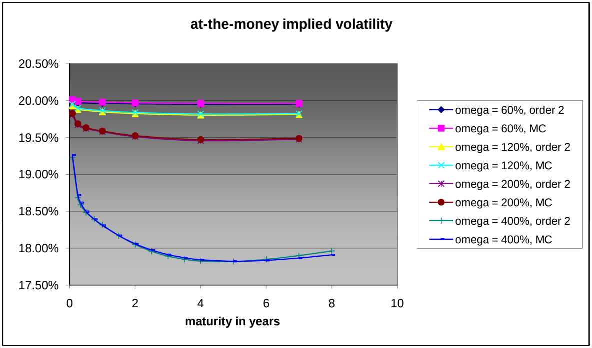
$$12C^{X\xi 2} - C^{\xi\xi}v\,(v+4) + 4C^{\mu}v\,(v-4) \leq 0$$

- **The global shape of the smile is well captured by the first order expansion**: the true implied volatility for strike K is indeed approximately affine in $\ln(K/S_0)$,
- But the level of the smile is well captured only for small valus of ω.

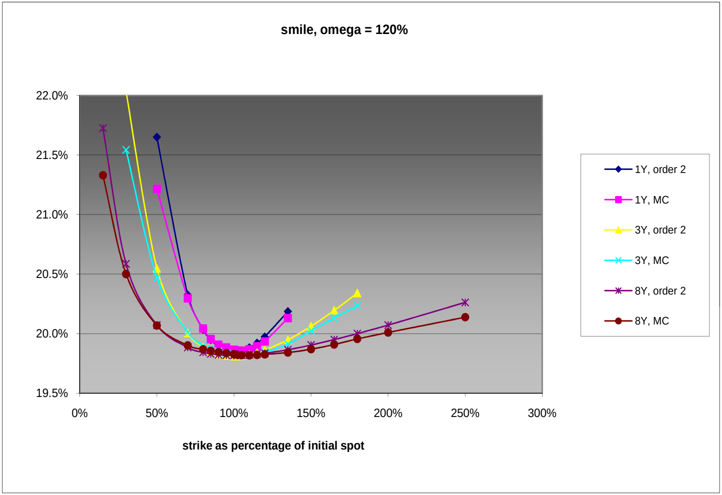
Second order

We first consider the situation when spot returns and forward variances are uncorrelated. In this case, the ATM skew vanishes, and so does its expansion at second order in ω. We pick

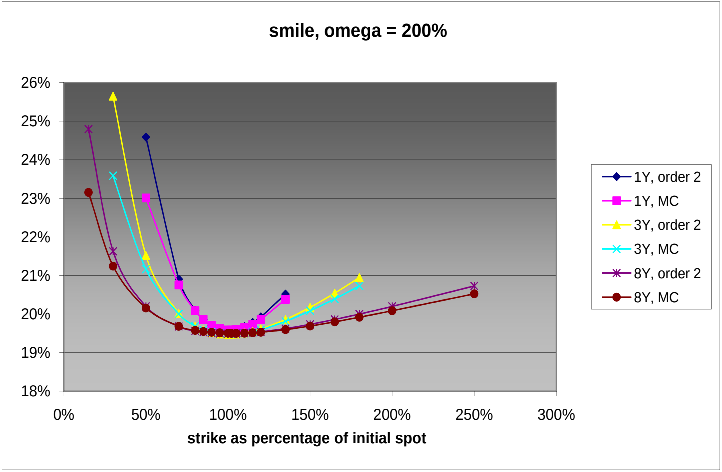
θ	k_X	k_Y	ρ_{SX}	ρ_{SY}	ρ_{XY}	ξ
0.25	8	0.35	0	0	0.6	$(0.2)^2$



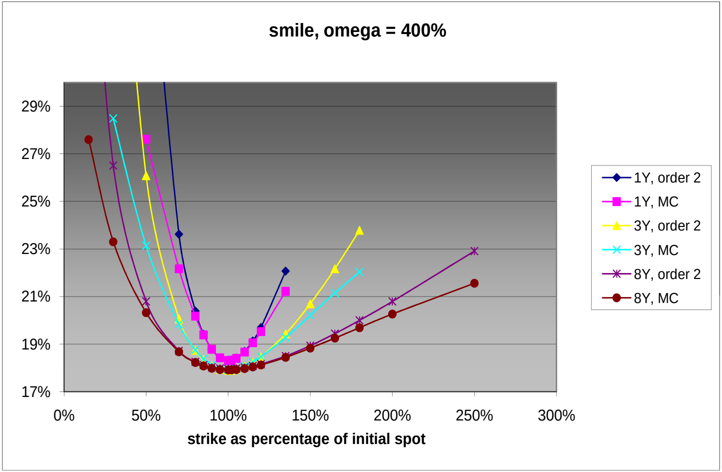
Second order



Second order



Second order

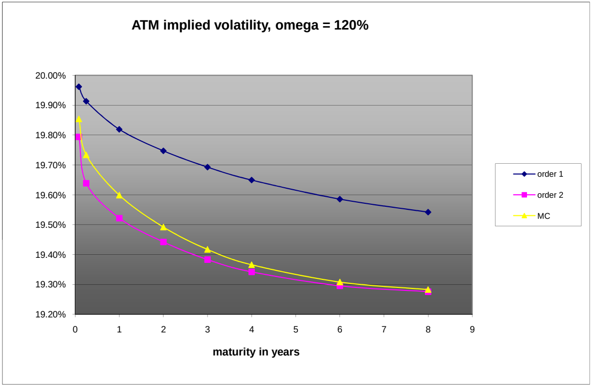


Second order

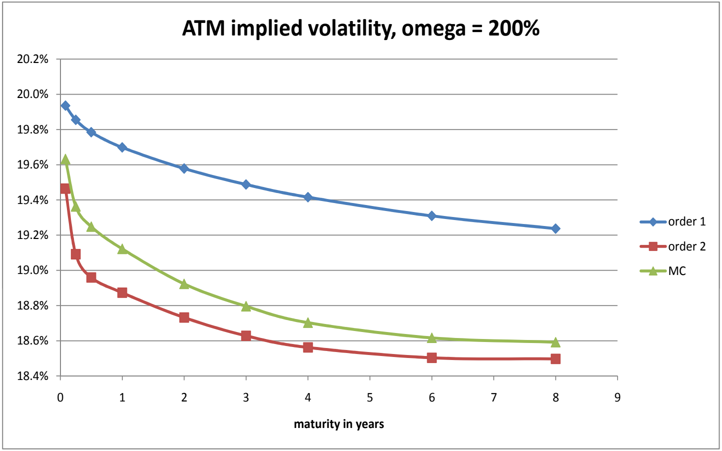
- **The ATM implied volatility is very sharply estimated by the second order expansion**, even up to $\omega = 400\%$ and to long maturities. For $T = 15$ years, the estimate is less than 15 bps above the true ATM volatility.
- Looking at the whole smile: the second order expansion of the implied volatility is excellent around the money, but becomes too large for strikes far from the money.
- Not surprising, because no arbitrage requires that for very small and very large strikes, the implied volatility squared $I(T,K)^2$ grows at most linearly with $\ln(K/S_0)$ (see Lee [4]), whereas the second order estimate for $I(T,K)^2$ grows like $\ln^4(K/S_0)$, see (4). This means that the remainder $O(\omega^3) = R(\omega, T, K)$ is large for large K, for finite ω.
- Nevertheless, even for $\omega = 400\%$, a maturity of 8 years and a deep out-the-money strike of 250%, the error is only 1.5 point of volatility.

Second order

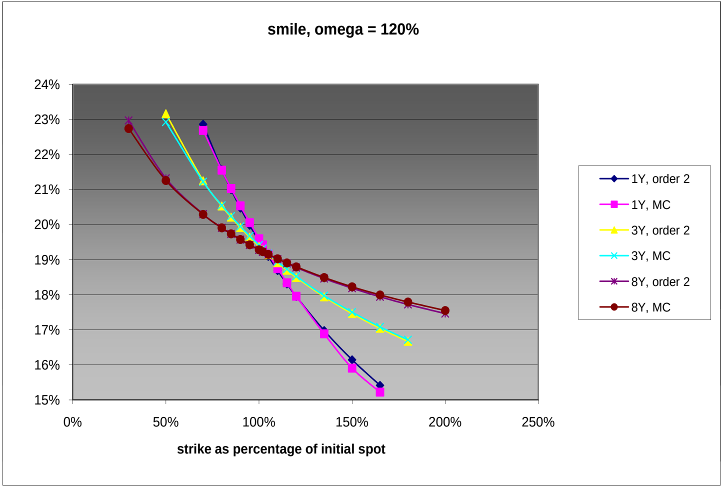
We now check numerically the accuracy of the second order expansion of the smile in the general case of correlated spot returns and variances.



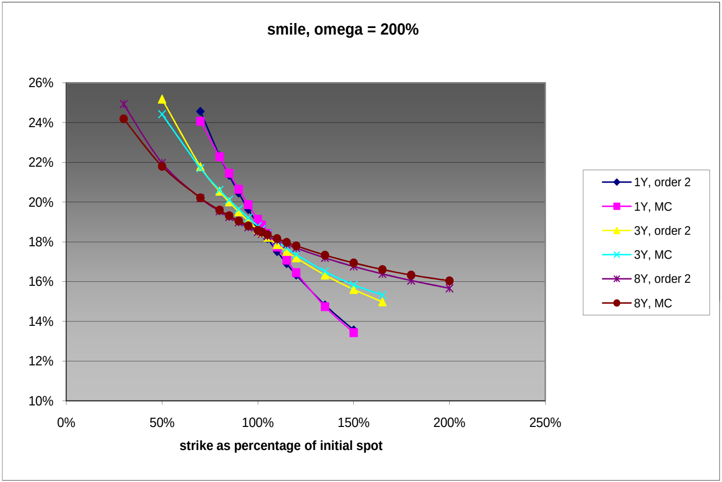
Second order



Second order



Second order



Short-term asymptotics of implied volatility

- Assume $d\xi_t^t = \cdots dt + \omega(\xi_t^t)^\varphi dB_t$
- Let ρ_{SV} be the correlation between S_t and instantaneous variance $V_t = \xi_t^t$
- Heston: $\varphi = \frac{1}{2}$, $\rho_{SV} = \rho$;
  Bergomi: $\varphi = 1$, $\rho_{SV} = \alpha_\theta \left((1-\theta)\rho_{SX} + \theta\rho_{SY}\right)$
- Then for short maturities

$$I_T^{ATM} \simeq \sqrt{\xi_0^0} + \frac{\rho_{SV}\left(I_T^{ATM}\right)^{2\varphi} T}{8}\omega + \frac{T}{96}\left(I_T^{ATM}\right)^{4\varphi-3}\left(\left(5-8\varphi\right)\rho_{SV}^2 - 4\right)\omega^2 + O\left(\omega^3\right)$$

$$\mathcal{S}_T \simeq \frac{\rho_{SV}\left(I_T^{ATM}\right)^{2\varphi-2}}{4}\omega + \frac{T}{8}\left(I_T^{ATM}\right)^{4\varphi-3}\rho_{SV}^2\left(\frac{2}{3}\varphi - \frac{5}{12}\right)\omega^2 + O\left(\omega^3\right)$$

$$\kappa_T \simeq \frac{\left(I_T^{ATM}\right)^{4\varphi-5}}{8}\left(\left(\frac{2}{3}\varphi - \frac{7}{6}\right)\rho_{SV}^2 + \frac{1}{3}\right)\omega^2 + O\left(\omega^3\right)$$

- $\Rightarrow$ Short-term ATM skew does not depend on short-term ATM vol iff $\varphi = 1$ (observed in equity markets)
- $\Rightarrow$ Short-term ATM convexity does not depend on short-term ATM vol iff $\varphi = \frac{5}{4}$. And $(\forall \rho_{SV},\ \kappa_T \geq 0) \iff \varphi \geq \frac{5}{4}$

Long-term asymptotics of implied volatility

- Assume the term-structure of variance swaps volatilities is flat: $\xi_0^t \equiv \xi$.
- Assume that for large $u-t$, $\mu(t,u,y) \propto (u-t)^{-\alpha}$, $\alpha > 0$.
  Then at higher order in ω, for long maturities,

$$\begin{aligned} \mathcal{S}_T &\propto T^{-\alpha} && \text{if } \alpha < 1 \\ \mathcal{S}_T &\propto T^{-1} && \text{if } \alpha > 1 \end{aligned}$$

  Cf Bergomi, *Smile Dynamics 4* [2], for the link with the skew-stickiness ratio.
- Assume that for large $u-t$ and $u'-t$,
  $\nu(t,u,u',y) \propto (u-t)^{-\alpha}(u'-t)^{-\alpha}$, $\alpha > 0$.
  Also assume that spots and volatilities are uncorrelated ($\mu \equiv 0$). Then at higher order in ω, for long maturities,

$$\begin{aligned} \kappa_T &\propto T^{-2\alpha} && \text{if } \alpha < 1 \\ \kappa_T &\propto T^{-2} && \text{if } \alpha > 1 \end{aligned}$$

- Exponential decay $\leftrightarrow \alpha > 1$.

- Remember $\mathcal{S}_T = \frac{C^{X\xi}}{2v^{3/2}\sqrt{T}}\omega + O\left(\omega^2\right)$
- Let us now compute the skewness s_T of log-returns:

$$s_T = \frac{\mathbb{E}\left[\mathcal{X}_T^3\right]}{\mathbb{E}\left[\mathcal{X}_T^2\right]^{3/2}}, \qquad \mathcal{X}_T = X_T - \mathbb{E}\left[X_T\right] = \int_0^T \sqrt{\xi_t^{t,\omega}} dW_t^1$$

- We have $\mathbb{E}\left[\mathcal{X}_T^2\right] = \int_0^T \mathbb{E}\left[\xi_t^{t,\omega}\right] dt = \int_0^T \xi_0^t dt$ and

$$\mathbb{E}\left[\mathcal{X}_T^3\right] = 3\omega C^{X\xi} + O\left(\omega^2\right)$$

- At first order in the vol of vol, the skewness of (the distribution of) $\ln\left(S_T/S_0\right)$ is thus

$$s_T = \frac{3\omega C^{X\xi}}{\left(\int_0^T \xi_0^t dt\right)^{3/2}}$$

- The ATM skew $\mathcal{S}_T$ simply reads

$$\mathcal{S}_T = \frac{s_T}{6\sqrt{T}} + O(\omega^2)$$

Conclusion

- We have considered **general "second generation" stochastic volatility models** and derived an expansion of the smile of implied volatility at second order in the volatility of variance.
- At this order, **the smile is quadratic in** $L = \ln(K/S_0)$.
- This expansion shows that **the smile is driven by three model-dependent quantities**:
  - $C^{X\xi}$, the integrated spot-variance covariance function,
  - $C^{\xi\xi}$, the integrated variance-variance covariance function,
  - C^{μ}, which, like $C^{X\xi}$, depends only on the instantaneous spot-variance covariance, but in a more complex way.
- $C^{X\xi}$ drives the ATM implied volatility and ATM skew, at first order. When spot returns are uncorrelated with variances, the smile is U-shaped, $C^{X\xi} = C^{\mu} = 0$ and $C^{\xi\xi}$ drives both the ATM implied volatility and ATM convexity.
- In the general case where spot returns are correlated with variances, the second order correction for ATM implied volatility, the second order correction for the ATM skew, and the ATM convexity are all driven by a linear combination of $\left(C^{X\xi}\right)^2$, $C^{\xi\xi}$ and C^{μ}.

- Our derivation relies on the fact that in Model (1), the volatility of the asset is an autonomous stochastic process, meaning that it incorporates **no local volatility component**, and that λ **does not depend on the asset value**.
- These three fundamental covariance functions are computed for Heston-like traditional stochastic volatility models (In particular we give a new derivation of A. Lewis' results [5]).
- They are also computed for the Bergomi model with two factors on the variance curve. In this case, numerical experiments show an excellent agreement between the estimate and the true quantity **at first order for the ATM skew**, and **at second order for the ATM implied volatility**, hence for the whole smile, up to typical values of the volatility of variance (say, 400% in the equity market).
- We have given short-term and long-term behaviour of the smile.
- We have also rederived that the ATM skew is just the skewness of the distribution of the log-spot, divided by 6 times the square root of maturity, at first order in the volatility of variance.

- Bergomi L., *Smile Dynamics 2*, Risk Magazine, pages 67-73, October 2005.
- Bergomi L., *Smile Dynamics 4*, Risk Magazine, December 2009.
- Backus D., Foresi S., Li K. and Wu L., *Accounting for Biases in Black-Scholes*, unpublished.
- Lee R., *The moment formula for implied volatility at extreme strikes*, Stanford University and Courant Institute, 2002.
- Lewis A., *Option valuation under stochastic volatility*, Finance Press, 2000.