

Valuation of Volatility Derivatives as an Inverse Problem

Peter Friz, Jim Gatheral
Courant Institute and Merrill Lynch

May 24, 2004

Abstract

Ground-breaking recent work by Carr and Lee extends well-known results for variance swaps to arbitrary functions of realized variance, provided a zero-correlation assumption is made. We give a detailed mathematical analysis of some of their formal computations and work out the cases of volatility swaps and calls on variance. The latter leads to an ill-posed problem and we suggest a solution in a least-square sense.

The sum is divergent, that means we can do something.
-Heaviside¹

1 Introduction

Since the crash of 87, the study of volatility has become central to quantitative finance. In 1973, Black and Scholes introduced their lognormal model where both the actual and risk-neutral stock-price S_t are driven by the same volatility σ . With the usual notation we have the Black-Scholes SDE (assuming zero rates and dividends),

¹Quote suggested by Peter Carr. We take this opportunity to thank him, Jining Han, Bob Kohn and Roger Lee for related discussions.

$$\begin{aligned}dS_t &= \sigma S_t dW_t, \\d\langle S \rangle_t &= \sigma^2 S_t^2 dt.\end{aligned}$$

Pricing a financial contract by computing a risk-neutral expectation requires continuous-time hedging and the impact of doing this with a wrong σ is well studied, *e.g.* (Carr and Madan 1998) and the references therein. Market option prices are quoted as implied Black-Scholes volatilities and the resulting volatility surface is far from constant; the dependence of implied volatility on strike and expiration is referred to as the volatility smile. For background on how such implied volatility surfaces look and how they move around, see Cont and da Fonseca (2002).

Dupire suggested a one factor time-dependent Markov model in which

$$\sigma = \sigma(S_t, t)$$

is determined by the implied volatility surface such that all European option prices are recovered. It is a pure pricing process and σ represents an effective volatility for pricing processes rather than a true model of the dynamics of volatility.

More realistically, stochastic volatility models have been proposed in which

$$\sigma = \sigma(t, \omega)$$

itself follows a diffusion. These models are indeed able to generate a volatility smile. A particularly important measure of the smile, the at-the-money-skew, is known to be proportional to the correlation ρ . Exotic options can be very sensitive not only the skew but also its dynamics in time in which case a stochastic volatility model is a necessity for pricing and hedging, see Gatheral (2003).

If the stock price is assumed to follow a continuous-path diffusion, a semi-static *model-independent* hedge perfectly replicates the realized variance

$$\int_0^T v(s, \omega) ds \equiv \langle x \rangle_T$$

and instruments with this payoff, *variance swaps*, are actively traded.

In recent work (Carr and Lee 2003), Carr and Lee showed how to extend this to instruments with arbitrary payoff

$$f(\langle x \rangle_T)$$

provided that correlation is zero. For special f of the form $\exp[\lambda.]$ both prices and hedges are given explicitly in terms of associated European contracts on S_T . Consequently, the price of such *volatility derivatives* is determined by European puts and calls expiring at T or, equivalently, by the time slice T of the implied-volatility surface. Formally,

$$\mathbb{E}[f(\langle x \rangle_T)] = \int w_f(k) c(k, T) dk \tag{1}$$

where the $c(k, T)$ are European option prices and the $w_f(k)$ are weights to be found.

In their paper, Carr and Lee use Laplace techniques to decompose an arbitrary f in terms of payoffs of the form $\exp[\lambda.]$ and a formal computation suggests how to price (and hedge) a claim with payoff $f(\langle x \rangle_T)$. In effect, they show how to compute the weights for a replicating strip of European options.

Our main insight is that the weights w_f used for the replicating strip of European options do not depend nicely on f and in fact may not be well-defined for many functions f of interest (such as variance calls).

In our paper we implement a regularization scheme to permit computation of the weights and in so doing, demonstrate that inverting equation (1) directly is a more appealing approach.

We exhibit the problem as an ill-posed one and show how least-squares optimization can be used to construct solutions.

We focus on two payoffs of particular financial interest: the *volatility swap*

$$f(\langle x \rangle_T) = \sqrt{\langle x \rangle_T}$$

and the *call on variance*

$$f(\langle x \rangle_T) = (\langle x \rangle_T - K)^+ .$$

2 Laplace transform of quadratic variation

We assume no jumps and zero risk-free rate (for simplicity) so that the risk-neutral stock evolution is given by

$$dS_t = \sigma(t, \omega) S_t dW_t.$$

Set $x_T = \log[S_T/S_0]$ and observe that its quadratic variation $\langle x \rangle_T$ is exactly the total realized variance $\int_0^T \sigma^2(t, \omega) dt$.

Proposition 1 (Carr, Lee). *Assume independence between the Brownian motion W and volatility σ . With $p(\lambda) = 1/2 \pm \sqrt{1/4 + 2\lambda}$ we have*

$$\mathbb{E}[e^{\lambda \langle x \rangle_T}] = \mathbb{E}[e^{p(\lambda) x_T}]. \quad (2)$$

Proof. Note that x_t has a drift $-\frac{1}{2} \langle x \rangle_t$. Then

$$y_t = x_t + \frac{1}{2} \langle x \rangle_t$$

is a martingale with associated exponential martingale

$$\exp\left(py_t - \frac{1}{2}p^2 \langle y \rangle_t\right) = \exp\left(px_t - \frac{1}{2}(p^2 - p) \langle x \rangle_t\right).$$

First, assume $\langle x \rangle_t = \int_0^t \sigma^2(s) ds$ deterministic. Then taking expectations yields the required identity. In the stochastic case, we can condition on the volatility scenario which amounts to fixing $\langle x \rangle_t(\omega)$. Repeat the argument given before and then average over all possible volatility scenarios. The result follows. \square

Remark 2. *Either choice of the sign yields a correct expectation, interpreted as the price of an instrument. But the hedging strategy may be different. For instance, $\lambda = 0$ leads to $p = 0, 1$. The fair price 1 can be obtained by holding one dollar or as risk-neutral value of S_T/S_0 at time T .*

Remark 3. *Generalized European-style payoffs, as on the r.h.s. above, can be replicated by an infinite strip of (out of the money) calls/puts. Set $F = S_0$,*

$k = \log(K/F)$ and $c(k) = C(K, T; S_0)/K$, $p(k) = P(K, T; S_0)/K$. Note that put-call symmetry holds, $p(k) = e^{-k} c(-k)$.

$$\begin{aligned}
\mathbb{E}[e^{\lambda \langle x \rangle_T} - 1] &= \frac{1}{F^p} (\mathbb{E}[S_T^p] - F^p) \\
&= \frac{p(p-1)}{F^p} \left(\int_0^F dK P(K) K^{p-2} + \int_F^\infty dK C(K) K^{p-2} \right) \\
&= 2\lambda \left(\int_{-\infty}^0 dk \frac{P(K)}{K} e^{kp} + \int_0^\infty dk \frac{C(K)}{K} e^{kp} \right) \\
&= 2\lambda \left(\int_0^\infty dk c(k) e^k e^{-kp} + \int_0^\infty dk c(k) e^{kp} \right) \\
&= 4\lambda \int_0^\infty dk c(k) e^{k/2} \cosh[k\sqrt{1/4 + 2\lambda}] \tag{3}
\end{aligned}$$

where complex λ are not a problem².

From the second line above observe the ATM-weight

$$\frac{p(p-1)}{F^p} K^{p-2} |_{K=F} = \frac{p(p-1)}{F^2} = \frac{2\lambda}{S_0^2}$$

Remark 4. For the well-known variance swap,

$$\mathbb{E}[\langle x \rangle_T] = 2 \left(\int_0^F dK \frac{P(K)}{K^2} + \int_F^\infty dk \frac{C(K)}{K^2} \right) \tag{4}$$

$$= 2 \left(\int_{-\infty}^0 dk p(k) + \int_0^\infty dk c(k) \right) \tag{5}$$

$$= 2 \left(\int_0^\infty dk c(k) e^k + \int_0^\infty dk c(k) \right) \quad \text{using put-call symmetry}$$

$$= 4 \int_0^\infty dk c(k) e^{k/2} \cosh[k/2] \tag{6}$$

As a consistency check, the last expression can also be obtained from (3) by differentiating with respect to λ at $\lambda = 0$.

For later reference we note the ATM-weight with respect to actual strikes $2/F^2 = 2/S_0^2$ from (4) and log-strikes $4 e^{k/2} \cosh[k/2] |_{k=0} = 4$ from (6).

²a) Note that the l.h.s. can become infinity for $\text{Re}[\lambda]$ too large, this will only depend on the pdf of volatility in the considered model. Any stochastic volatility model in which volatility stay below some σ_{\max} implies finite exponential moments.

b) Note that $\cosh[k\sqrt{1/4 + 2\lambda}]$ is an entire function in λ .

3 Arbitrary functions of realized variance

Assume we want to replicate $f(\langle x \rangle_T)$ and manage to find a way to represent (or at least approximate) f by a *linear* combination of Laplace functionals $y \mapsto \exp[\lambda y]$, $\lambda \in \mathbb{C}$.

Example 5. $f(y) = \sqrt{y}$. By a well-known formula,

$$\sqrt{y} = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - e^{-\lambda y}}{\lambda^{3/2}} d\lambda. \quad (7)$$

Example 6. Whenever $f(y)$ has a Laplace-transform $F(\lambda)$, for $a \in \mathbb{R}$ on the right of all singularities of F ,

$$f(y) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} F(\lambda) e^{\lambda y} d\lambda \quad (8)$$

.

In general, an infinite number of Laplace functionals has to be combined to recreate a given payoff. However, a finite number will suffice to approximate f . We quote a classical result due to L. Schwartz, (Schwartz 1959)

Theorem 7. *The linear span of $\{e^{-\lambda y} : \lambda \in \mathbb{R}, \lambda > 0\}$ is dense in $C([0, \infty), \mathbb{R})$.*

In practical terms, it will suffice to find an approximation for say $y = \langle x \rangle_T \in [0, T]$ since we are almost sure that, say, S&P 500 volatility remains below 100%.

3.1 Volatility swaps

Using (7) and taking expectations

$$\mathbb{E} \left[\sqrt{\langle x \rangle_T} \right] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - \mathbb{E} \left[e^{-\lambda \langle x \rangle_T} \right]}{\lambda^{3/2}} d\lambda. \quad (9)$$

Substituting the expression for $\mathbb{E} \left[e^{-\lambda \langle x \rangle_T} \right]$ from equation (3), we obtain

$$\begin{aligned} \mathbb{E} \left[\sqrt{\langle x \rangle_T} \right] &= \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{d\lambda}{\sqrt{\lambda}} \int_0^\infty dk e^{k/2} c(k) \cosh \left[k \left(\sqrt{1/4 - 2\lambda} \right) \right] \\ &= \int_0^\infty dk c(k) w_{vol}(k) \end{aligned}$$

with

$$w_{vol}(k) \equiv \frac{2}{\sqrt{\pi}} e^{k/2} \int_0^\infty \frac{d\lambda}{\sqrt{\lambda}} \cosh \left[k \left(\sqrt{1/4 - 2\lambda} \right) \right]. \quad (10)$$

Theorem 8. *We have*

$$w_{vol}(k) = \sqrt{\frac{\pi}{2}} e^{k/2} I_1(k/2) + \sqrt{2\pi} \delta(k) \quad (11)$$

To a very good approximation, only the delta-function term counts.

Proof. From (10), the weight of the option with log-strike k in the replicating strip is given by

$$\frac{2}{\sqrt{\pi}} e^{k/2} \int_0^\infty \frac{d\lambda}{\sqrt{\lambda}} \cosh k \left(\sqrt{1/4 - 2\lambda} \right) = \frac{2}{\sqrt{\pi}} e^{k/2} \{J + K\}$$

with

$$J = \int_0^{1/8} \frac{d\lambda}{\sqrt{\lambda}} \cosh \left[k \left(\sqrt{1/4 - 2\lambda} \right) \right]$$

and

$$K = \int_{1/8}^\infty \frac{d\lambda}{\sqrt{\lambda}} \cosh \left[ik \left(\sqrt{2\lambda - 1/4} \right) \right] = \int_{1/8}^\infty \frac{d\lambda}{\sqrt{\lambda}} \cos \left[k \left(\sqrt{2\lambda - 1/4} \right) \right]$$

With the changes of variable $u = \sqrt{1/4 - 2\lambda}$ and $v = \sqrt{2\lambda - 1/4}$, we obtain

$$J = \sqrt{2} \int_0^{1/2} \frac{u du}{\sqrt{1/4 - u^2}} \cosh(ku) = \frac{\pi}{2\sqrt{2}} L_{-1}(k/2)$$

where $L_n(\cdot)$ denotes the Struve L-function and

$$K = \sqrt{2} \int_0^\infty \frac{v dv}{\sqrt{1/4 + v^2}} \cos(kv)$$

The K integrand clearly diverges as $v \rightarrow \infty$. We may eliminate this singularity by defining

$$\begin{aligned} K' &= \sqrt{2} \int_0^\infty dv \left(\frac{v}{\sqrt{1/4 + v^2}} - 1 \right) \cos(kv) \\ &= \frac{\pi}{2\sqrt{2}} \{I_1(k/2) - L_{-1}(k/2)\} \end{aligned}$$

where $I_n(\cdot)$ represents a modified Bessel function of the first kind. Adding together J and K , we obtain the following analytical expression for the weights

$$w_{vol}(k) = \sqrt{\frac{\pi}{2}} e^{k/2} I_1(k/2) + \frac{2\sqrt{2}}{\sqrt{\pi}} \int_0^\infty dv \cos k v$$

We recognize the second term in this expression as the Dirac-delta function. Integrating the second term gives

$$w_{vol}(k) = \sqrt{\frac{\pi}{2}} e^{k/2} I_1(k/2) + \sqrt{2\pi} \delta(k)$$

□

Remark 9. *It is worth emphasizing here what Carr and Lee have achieved: to compute the fair value of volatility under our standing zero-correlation assumption we only need a knowledge of the volatility smile at expiration. Given that we already know how to compute the expected variance the so-called convexity adjustment follows immediately. This suggests that the dynamics of volatility are highly constrained by the current implied volatility surface.*

Remark 10. *We give a financial interpretation of formula (11). To this end, recall that the weights w_{vol} are universal for all zero-correlation models.*

$$E \left[\sqrt{\langle x \rangle_T} \right] = \int_0^\infty dk c(k) w_{vol}(k).$$

In particular, we may restrict attention to a standard Black-Scholes model where (with obvious notation) the normalized calls $c(k) = c_{BS}(k, y)$ only depend on log-strike and $y = \langle x \rangle_T = \sigma^2 T$. Thus

$$\sigma \sqrt{T} = \int_0^\infty dk c(k) w_{vol}(k).$$

However, by a well-known approximation for ATM Black-Scholes calls,

$$c_{BS}(0, \sigma^2 T) \approx \frac{\sigma \sqrt{T}}{\sqrt{2\pi}}$$

and we recover $w_{vol}(k) \approx \sqrt{2\pi} \delta(k)$.

3.2 Generalized payoffs

Using (8) and taking expectations, a formal computation yields

$$\begin{aligned}
\mathbb{E}[f(\langle x \rangle_T)] &= \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} F(\lambda) \mathbb{E}[e^{\lambda \langle x \rangle_T}] d\lambda \\
&= \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} F(\lambda) \left[1 + 4\lambda \int_0^\infty dk c(k) e^{k/2} \cosh[k\sqrt{1/4 + 2\lambda}] \right] d\lambda \\
&= f(0) + \int_0^\infty dk c(k) 4 e^{k/2} \left(\frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} F(\lambda) \lambda \cosh[k\sqrt{1/4 + 2\lambda}] d\lambda \right).
\end{aligned} \tag{12}$$

Elegant though the complex integral in (12) might look, its convergence is not guaranteed; indeed for typical non-smooth payoffs such as calls, it diverges. The rest of our paper concerns itself with how to deal with such payoffs.

Example 11 (Variance call). $f(y) = (y - K)^+$ with $K \geq 0$ has Laplace-transform $F(\lambda) = e^{-\lambda K}/\lambda^2$. Formally

$$\mathbb{E}[(\langle x \rangle_T - K)^+] = \int_0^\infty dk c(k) \underbrace{\left(\frac{4 e^{k/2}}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{e^{-\lambda K}}{\lambda} \cosh[k\sqrt{1/4 + 2\lambda}] d\lambda \right)}_{=: w_{call}(k) = w_{call}(k, K)}. \tag{13}$$

Note that $e^{-\lambda K}$ is a purely oscillating term as $\lambda \rightarrow a \pm i\infty$. For $k > 0$ we have the exponentially divergent term

$$|\cosh[k\sqrt{1/4 + 2\lambda}]| \sim \exp\left[k\sqrt{\text{Im}[\lambda]}\right] \text{ as } \text{Im}[\lambda] \rightarrow +\infty.$$

Later we will introduce the mollification technique which will provide a strong enough decay to ensure existence of all integrals.

In anticipation of our discussion of ATM weights in Section 4, note that for $k = 0$ we observe

$$w_{call}(k = 0, K) = 4 \left(\frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{e^{\lambda(-K)}}{\lambda} d\lambda \right) = 4\theta(-K)$$

using the fact that the Heaviside-function $\theta(\cdot)$ has Laplace-transform $1/\lambda$. In particular,

$$w_{call}(k = 0, K) = 0 \text{ for all } K > 0$$

where as from the discussion of the variance swap

$$\mathbb{E}[(\langle x \rangle_T - 0)^+] = \mathbb{E}[\langle x \rangle_T] = 4 \int_0^\infty dk c(k) e^{k/2} \cosh[k/2]$$

and we observe the discontinuity

$$w_{call}(k = 0, K = 0) = 4.$$

4 The structure of ATM weights

In this section only, we revert to dollar call and put prices (C, P) and dollar strikes K .

The following result does not require a zero-correlation assumption. In particular, local volatility and arbitrary stochastic volatility underlying dynamics are allowed.

Theorem 12. *Assume zero risk-free rate, no dividends and a decomposition*

$$\mathbb{E}[f(\langle x \rangle_T)] = \int_0^\infty dK w(K; S_0) \begin{cases} P(T, S_0; K) & \text{if } K < S_0 \\ C(T, S_0; K) & \text{if } K \geq S_0 \end{cases} .$$

Note that for this decomposition in terms of out-of-the-money options to hold we necessarily have $f(0) = 0$.

Then the ATM-weight $w(S_0, S_0)$ is given by

$$\frac{2f'(0)}{S_0^2} .$$

Proof. We only discuss two special cases, generalizations are obvious.

Case 1 (Local Volatility)

$$\begin{aligned} dS_t &= \sigma_{loc}(S_t, t) S_t dW_t \\ dy_t &= \sigma_{loc}^2(S_t, t) dt \end{aligned}$$

Note $\langle x \rangle_T \equiv y_T$. As a 2-dimensional time-inhomogeneous Markov process, its generator is written as

$$\mathcal{L} = \frac{1}{2} \sigma_{loc}^2(S_t, t) S^2 \partial_{SS} + \sigma_{loc}^2(S_t, t) \partial_y$$

Observe that

$$\begin{aligned} \mathbb{E}[f(\langle x \rangle_T)] &= \mathbb{E}[f(y_T)] \\ &= \mathbb{E}[f(\langle x \rangle_T) | S_0, y_0 = 0] \\ &= u(T, S_0, 0) \end{aligned}$$

where $u(T, S, y)$ is a solution of the PDE

$$\partial_T u = \mathcal{L}[u]$$

with initial condition

$$u(0, S, y) = f(y).$$

Pricing standard (out-of-the-money) Europeans calls and puts is based on the same PDE but uses different initial conditions,

$$g_K(S) := (K - S)^+ \text{ when } K < S_0 \text{ and } g_K(S) := (S - K)^+ \text{ otherwise.}$$

The resulting PDE solutions (*e.g.* the prices) at time 0 are

$$P(T, S_0; K) \text{ resp. } C(T, S_0; K).$$

By put-call parity,

$$P(T, S_0; S_0) = C(T, S_0; S_0).$$

By assumption,

$$u(T; S_0, 0) = \int_0^\infty dK w(K; S_0) \begin{cases} P(T, S_0; K) & \text{if } K < S_0 \\ C(T, S_0; K) & \text{if } K \geq S_0 \end{cases}.$$

Now differentiate w.r.t. T and evaluate at $T = 0$. From the PDE

$$[\mathcal{L}f](S_0, 0) = \int_0^\infty dK w(K; S_0) [\mathcal{L}g_K](S_0, 0).$$

Note that $[\mathcal{L}g_K](S, 0) = \frac{1}{2} \sigma_{loc}^2(S, 0) S^2 \delta_K(S)$ and $[\mathcal{L}f](S, y) = \sigma_{loc}^2(S_0, 0) f'(y)$. Hence

$$\begin{aligned} \sigma_{loc}^2(S_0, 0) f'(0) &= \frac{1}{2} \int_0^\infty dK w(K; S_0) \sigma_{loc}^2(S_0, 0) S_0^2 \delta_K(S_0) \\ &= \frac{1}{2} \sigma_{loc}^2(S_0, 0) S_0^2 w(S_0; S_0). \end{aligned}$$

Case 2 (Stochastic Volatility)

The modifications are minor. The underlying Markovian diffusion now contains three components: the stock price process S_t , its instantaneous variance process v_t and the accumulated total variance y_t . As before

$$dy_t = v_t dt$$

and

$$\mathcal{L} = \frac{1}{2} v S^2 \partial_{SS} + (\text{terms involving } \partial_v, \partial_{vv}, \partial_{Sv}) + v \partial_y.$$

The derivation goes through line by line as in the local volatility case. It suffices to note that all the payoffs under consideration, namely $f(y)$, $(K - S)^+$ and $(S - K)^+$ are insensitive to v so differentiating w.r.t v yields a zero contribution. \square

Example 13 (Variance Swap). $f(z) = z$. We find ATM-weight $2/S_0^2$ in agreement with the well-known result (see Remark 4).

Example 14 (Volatility Swap). $f(z) = \sqrt{z}$ leads to $f'(0) = \infty$. This implies mass-concentration of $k \mapsto w(k, S_0)$ at the money ($k = S_0$) in agreement with Section 3.1.

Example 15 (Laplace payoff). $f(z) = \exp[\lambda z] - 1$. The ATM-weight equals $2\lambda/S_0^2$ in agreement with equation (3).

Example 16 (Variance Call). $f(z) = (z - K)^+$. We have $f'(0) = 0$ implying zero ATM weight in agreement with Example 11.

The last example provides a different proof of the discontinuity of ATM-weights for variance calls as $K \rightarrow 0$ which we found earlier.

5 Pdf of variance

Fix some strike $K > 0$. Assume $\mathbb{P}[\langle x \rangle_T \in dy] = g(y)dy$. Crystallizing the essence of the Carr-Lee result, we propose to recover the pdf of realized quadratic variation from a strip of European calls.

Let $\delta_K^{(h)}$ denote the Gaussian density with mean K and standard deviation $\sqrt{h} \ll 1$. Then, using (3) we obtain

$$\begin{aligned}
g(K) &= \mathbb{E}[\delta_K(\langle x \rangle_T)] \approx \mathbb{E}[\delta_K^{(h)}(\langle x \rangle_T)] \\
&= \mathbb{E} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} dp e^{-ip\langle x \rangle_T} e^{ipK} e^{-hp^2/2} \right] \quad (\text{Fourier-inversion}) \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} dp e^{ipK} e^{-hp^2/2} \mathbb{E} [e^{-ip\langle x \rangle_T}] \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} dp e^{ipK} e^{-hp^2/2} \left(1 - 4ip \int_0^{\infty} dk c(k) e^{k/2} \cosh \left[k \sqrt{1/4 - 2ip} \right] \right) \\
&= \underbrace{\delta_K^{(h)}(0)}_{\rightarrow 0 \text{ as } h \rightarrow 0} + \int_0^{\infty} c(k) w_{pdf}(k, K; h) dk
\end{aligned}$$

by application of Fubini's Theorem³ and with call option weights

$$w_{pdf}(k, K; h) = 4 e^{k/2} \frac{1}{2\pi} \int_{-\infty}^{\infty} dp e^{-hp^2/2} (-ip) e^{ipK} \cosh \left[k \sqrt{1/4 - 2ip} \right].$$

³We sketch how to justify the use of Fubini's theorem above as we don't believe it is trivial. It suffices to check that the iterated integral

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} dp e^{ipK} e^{-hp^2/2} \left(1 - 4ip \int_0^{\infty} dk c(k) e^{k/2} \cosh \left[k \sqrt{1/4 - 2ip} \right] \right)$$

(which we know to be finite) exists as absolutely convergent iterated integral. Noting that

$$\cosh \left[k \sqrt{1/4 - 2ip} \right] \sim \exp \left[k |p|^{1/2} \right]$$

the essence is

$$\int_{-\infty}^{\infty} dp e^{-hp^2/2} p \int_0^{\infty} dk c(k) \exp \left[k(1/2 + |p|^{1/2}) \right].$$

W.l.o.g. we may consider BS-calls where $c(k) \sim \exp(-\eta k^2)$ and we estimate

$$\begin{aligned}
&\int_{-\infty}^{\infty} dp e^{-hp^2/2} p \int_0^{\infty} dk \exp(-\eta k^2) \exp \left[k(1/2 + |p|^{1/2}) \right] \\
&\sim \int_{-\infty}^{\infty} dp e^{-hp^2/2} \int_0^{\infty} dk \exp(-\eta k^2 + k |p|^{1/2}) \\
&\leq \int_{-\infty}^{\infty} dp e^{-hp^2/2} e^{p/4} \sqrt{\pi} \\
&< \infty.
\end{aligned}$$

Now all integrals exist (and converge absolutely) thanks to the exponential damping factor $e^{-hp^2/2}$. Note also that w_{pdf} is a real-valued function as the integrand is an analytic function of ip .

We conclude that under mild technical assumptions on g (continuity and boundedness), the pdf is given by

$$g(y) = \lim_{h \rightarrow 0} \int_0^\infty c(k) w_{pdf}(k, y; h) dk.$$

Remark 17. *The last relation may be written as a linear integral transform*

$$g(y) = \lim_{h \rightarrow 0} [\mathfrak{W}_h c(\cdot)](y).$$

Conversely, if we know $g(y)$ for all y , we know the pdf of volatility. Under our standing zero-correlation assumption, calls are priced by averaging BS-prices (Hull and White 1987),

$$c(k) = \int_0^\infty c_{BS}(k, y) g(y) dy =: [\mathfrak{L}g(\cdot)](k).$$

We observe that $\lim_{h \rightarrow 0} \mathfrak{W}_h = \mathfrak{L}^{-1}$. In the theory of (linear) ill-posed equations, the family $\{\mathfrak{W}_h : h > 0\}$ is called a regularization scheme, see Monk (2003).

6 Digital variance payoff

Fix some strike $K > 0$. We consider a claim that pays one dollar if $\langle x \rangle_T \geq K$ and zero otherwise; we denote the payoff function by $\mathbb{I}_{[K, \infty)}$.

As before, we compute $\mathbb{E}[\mathbb{I}_{[K, \infty)}(\langle x \rangle_T)]$ as a weighted integral of European option prices. We proceed by representing the digital payoff as an integral of delta-function payoffs. Formally,

$$\mathbb{I}_{[K, \infty)}(z) = \int_K^\infty \delta_y(z) dy.$$

In order to ensure convergence, we again mollify the payoffs. Recall that $\delta_0^{(h)}$ is a Gaussian with peak at 0 and variance h . Set

$$\mathbb{I}_{[K, \infty)}^{(h)}(z) := (\mathbb{I}_{[K, \infty)} * \delta_0^{(h)})(z) = \int_K^\infty \delta_y^{(h)}(z) dy.$$

Then

$$\begin{aligned}
& \mathbb{E} [\mathbb{I}_{[K,\infty)}(\langle x \rangle_T)] \\
& \approx E \left[\mathbb{I}_{[K,\infty)}^{(h)}(\langle x \rangle_T) \right] \\
& = \int_K^\infty \mathbb{E}[\delta_y^{(h)}(\langle x \rangle_T)] dy.
\end{aligned}$$

From Section 5, this equals

$$\begin{aligned}
& \int_K^\infty dy \left(\delta_y^{(h)}(0) + \int_0^\infty c(k) 4 e^{k/2} \frac{1}{2\pi} \int_{-\infty}^\infty dp e^{-hp^2/2} (-ip) e^{ipy} \cosh \left[k\sqrt{1/4 - 2ip} \right] dk \right) \\
& = \underbrace{\mathbb{I}_{[K,\infty)}^{(h)}(0)}_{\rightarrow 0 \text{ as } h \rightarrow 0} + \underbrace{\int_0^\infty c(k) 4 e^{k/2} \frac{1}{2\pi} \int_{-\infty}^\infty dp e^{-hp^2/2} e^{ipK} \cosh \left[k\sqrt{1/4 - 2ip} \right] dk}_{\text{weights } w_{dig}=w_{dig}(k,K;h)}.
\end{aligned}$$

The last equality is based on the fact that $\int_K^M dy (-ip) e^{ipy} = e^{ipK} - e^{ipM}$ and that e^{ipM} tends weakly to zero as $M \rightarrow \infty^4$.

We note the ATM-weight ($k = 0$) of

$$4 \frac{1}{2\pi} \int_{-\infty}^\infty dp e^{-hp^2/2} e^{ipK} = 4 \delta_K^{(h)}(0) \approx 0 \quad (14)$$

for h small enough relative to variance-strike K (a harmless and realistic assumption). Note also that with

$$f(z) := \mathbb{I}_{[K,\infty)}^{(h)}(z) - \mathbb{I}_{[K,\infty)}^{(h)}(0) \approx \mathbb{I}_{[K,\infty)}^{(h)}(z) \approx \mathbb{I}_{[K,\infty)}(z)$$

we have a decomposition

$$\mathbb{E}[f(\langle x \rangle_T)] = \int_0^\infty c(k) w_{dig}(k, K; h) dk$$

to which our Theorem 12 applies:

$$w_{dig}(0, K; h) \sim f'(0) \approx 0.$$

which confirms (14).

⁴Formally set $M = \infty$. The oscillations of $e^{ip\infty}$ will completely cancel the remaining smooth integrand w.r.t. p

7 Variance calls

Fix a strike $K > 0$. Once again, we propose to compute $\mathbb{E}[(\langle x \rangle_T - K)^+]$ from a strip of European calls. We proceed by piling up digital payoffs,

$$g_K(z) := (z - K)^+ = \int_K^\infty \mathbb{I}_{[y, \infty)}(z) dy.$$

It is easy to check that

$$g_K^{(h)}(z) := (g_K * \delta_0^{(h)})(z) = \int_K^\infty \mathbb{I}_{[y, \infty)}^{(h)}(z) dy.$$

As before, $g_K^{(h)}(0) \approx 0$ for $K > 0$ and h small enough. Hence

$$\begin{aligned} \mathbb{E}[(\langle x \rangle_T - K)^+] &\approx \mathbb{E}\left[g_K^{(h)}(\langle x \rangle_T) - g_K^{(h)}(0)\right] \\ &= \mathbb{E}\left[\int_K^\infty \left(\mathbb{I}_{[y, \infty)}^{(h)}(\langle x \rangle_T) - \mathbb{I}_{[y, \infty)}^{(h)}(0)\right) dy\right] \\ &= \int_K^\infty dy \int_0^\infty c(k) w_{dig}(k, y; h) dk \\ &= \int_0^\infty dk c(k) \left(\int_K^\infty w_{dig}(k, y; h) dy\right) \\ &=: \int_0^\infty dk c(k) w_{call}(k, K; h) \end{aligned}$$

Proposition 18. *Let $a > 0$. We have*

$$w_{call}(k, K; h) = \frac{4 e^{k/2}}{2\pi i} \int_{a-i\infty}^{a+i\infty} e^{h\lambda^2/2} \frac{e^{-\lambda K}}{\lambda} \cosh\left[k \sqrt{1/4 + 2\lambda}\right] d\lambda \quad (15)$$

and this integral converges absolutely due to the exponential damping factor $e^{h\lambda^2/2}$. Setting $h = 0$ we recover the (divergent) weight $w_{call}(k, K)$ which was obtained by a formal computation earlier (Example 11).

Proof. Up to a factor $4 e^{k/2}$ the weights $w_{dig}(k, y; h)$ for the (mollified) digital payoff are

$$\begin{aligned} &\frac{1}{2\pi} \int_{-\infty}^\infty dp e^{-hp^2/2} e^{ipy} \cosh\left[k \sqrt{1/4 - 2ip}\right] \\ &= \frac{1}{2\pi i} \int_{0-i\infty}^{0+i\infty} d\lambda e^{h\lambda^2/2} e^{-\lambda y} \cosh\left[k \sqrt{1/4 + 2\lambda}\right] \quad (\text{set } \lambda = -ip) \\ &= \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} d\lambda e^{h\lambda^2/2} e^{-\lambda y} \cosh\left[k \sqrt{1/4 + 2\lambda}\right] \quad \text{for any } a \in \mathbb{R}. \end{aligned}$$

The last equality follows by change of contour noting that the integrand decays uniformly with $\mathbf{Im}(z) \rightarrow \pm\infty$.

In order to synthesize a variance call payoff we integrate as before

$$\int_K^M e^{-\lambda y} dy = \frac{1}{\lambda}(e^{-\lambda K} - e^{-\lambda M})$$

and get

$$\frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} d\lambda e^{\lambda^2/2} \frac{1}{\lambda} (e^{-\lambda K} - e^{-\lambda M}) \cosh \left[k \sqrt{1/4 + 2\lambda} \right] =: F(K, a) - F(M, a)$$

Then $\lim_{M \rightarrow \infty} F(M, a) = 0$ since $a > 0$ and

$$\left| e^{h\lambda^2/2} \frac{1}{\lambda} e^{-\lambda M} \cosh \left[k \sqrt{1/4 + 2\lambda} \right] \right| \leq e^{-aM} \underbrace{\left| e^{h\lambda^2/2} \frac{1}{\lambda} \cosh \left[k \sqrt{1/4 + 2\lambda} \right] \right|}_{\text{integrable over } a \pm i\infty}$$

□

7.1 Visualization of the weights

Note first that as $h \rightarrow 0$, the integrand in (15) becomes highly oscillatory and so direct integration is not obviously the right way to compute the weights. To see what this means in practice, consider the examples presented in Figures 1 and 2.

It is clear from the form of the regularization scheme that increasing h in $g_K^{(h)}(z)$ effectively smoothes the payoff $g_K(z)$ of a variance call. For reference, in Figure 3 we graph $g_K^{(h)}(z)$ for $h = .0001$ and $h = 0.00001$ respectively with $K = 0.04$ in both cases.

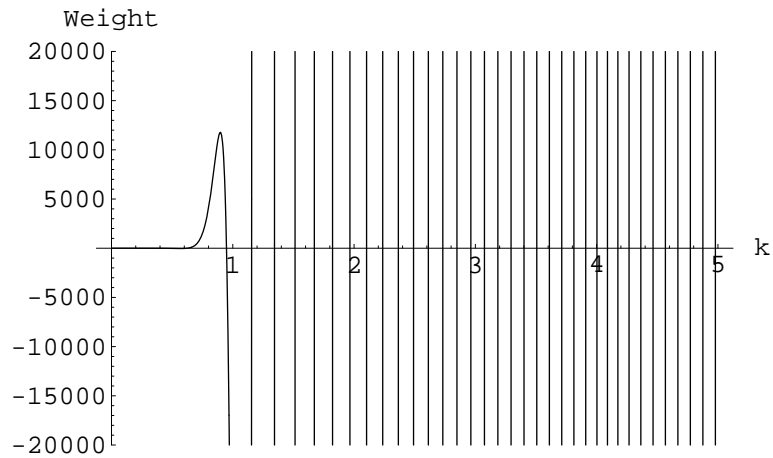


Figure 1: Here we see a plot of the weights $w_{call}(k, K; h)$ as a function of log-strike k with $K = 0.04$ and $h = 0.0001$. Note that the scale has been carefully chosen so the first peak (at +10,000 or so) can be seen. It would be an understatement to say that the weights are oscillatory.

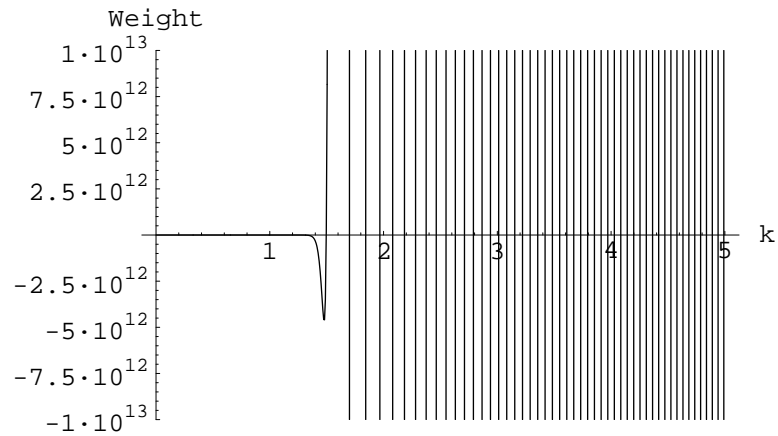


Figure 2: Now we see a plot of the weights $w_{call}(k, K; h)$ as a function of log-strike k with $K = 0.04$ and $h = 0.00001$ (ten times smaller than in the figure above). The weights are even more oscillatory with this smaller value of h .

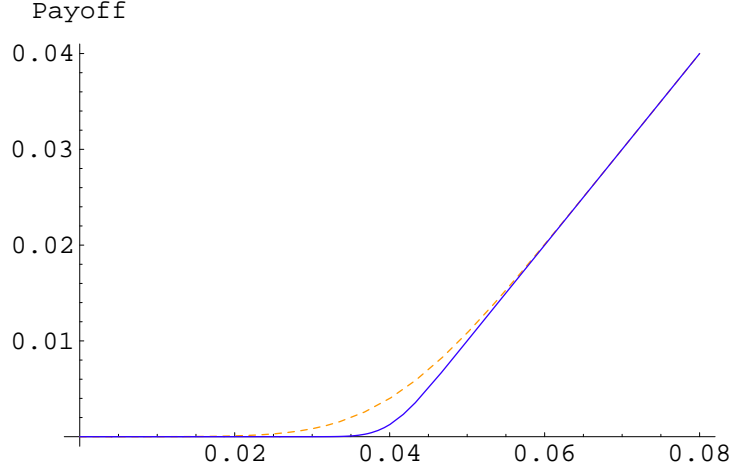


Figure 3: The solid blue line is the regularized payoff $g_K^{(h)}(z)$ with $h = 0.00001$ plotted as a function of total realized variance z and the dashed red line, the payoff with $h = 0.0001$ both with $K = 0.04 = 20\%^2$.

7.2 Computing the weights: an ill-posed PDE

Adopting the same notation as in the last section, consider the reduced weight

$$w_{red}(k, K) = \frac{w_{call}}{4e^{k/2}} = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} e^{h\lambda^2/2} \frac{e^{-\lambda K}}{\lambda} \cosh \left[k \sqrt{1/4 + 2\lambda} \right] d\lambda.$$

Then

$$\begin{aligned} \frac{\partial^2}{\partial k^2} w_{red} &= \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} e^{h\lambda^2/2} \frac{e^{-\lambda K}}{\lambda} (1/4 + 2\lambda) \cosh \left[k \sqrt{1/4 + 2\lambda} \right] d\lambda \\ &= \frac{1}{4} w_{red} - 2 \frac{\partial}{\partial K} w_{red}. \end{aligned} \quad (16)$$

This PDE lives on the domain $(k, K) \in [0, \infty) \times [0, \infty)$ with boundary conditions (as $h \rightarrow 0$)

$$w_{red}(k = 0, K > 0) = 0, \quad w_{red}(k > 0, K = 0) = \cosh[k/2].$$

The first boundary condition follows from Example 16 and the second one from the known variance swap weights, see (4). Note however the discontinuity at $k = 0, K = 0$.

At first sight, equation (16) is a standard parabolic PDE with time variable K and space variable k . It is the negative (wrong) sign of the K -derivative that leads us to identify it as ill-posed. To solve it, we would effectively have to solve a backward equation forward in time given an initial condition⁵.

8 Dealing with ill-posedness

We have accumulated evidence that the map

$$f(\cdot) \mapsto \text{“weights”}$$

is ill-posed in the sense that small perturbations of f in sup-topology may result in dramatic changes in the weights. The most that we can hope for in that case is a solution in some least-squares sense.

To this end, recall Remark 17 that given the law of realized variance $g(y)$ and under the standing zero-correlation assumption, calls are priced by averaging Black and Scholes prices,

$$c(k) = \int_0^\infty c_{BS}(k, y) g(y) dy =: [\mathfrak{L}g(\cdot)](k).$$

Formally then, we may obtain the pdf of variance by inverting the linear integral operator \mathfrak{L} . In terms of the regularization scheme $\{\mathfrak{W}_h : h > 0\}$ of Section 17,

$$\mathfrak{L}^{-1} = \lim_{h \rightarrow 0} \mathfrak{W}_h.$$

8.1 Matrix implementation

We pick a number of (log) strikes (*e.g.* from the market)

$$\{k_i\}, \quad i = 1, \dots, n$$

and a number of variance levels

$$\{v_j\}, \quad j = 1, \dots, m.$$

⁵In Fourier space, a symptom of this problem would be exponential blowup of the modes.

To capture the shape of total variance, $\{v_j\}$ should be fine enough and cover a realistically large range of variances. Setting $c_i = c(k_i)$ and $A_{ij} = c_{BS}(k_i, v_j)$ we obtain

$$c_i = \sum_{j=1}^m A_{ij} g_j$$

where $\mathbf{g} = \{g_j\}$ is a probability vector. From the injectivity of the BS-integral operator, we see that $\mathbf{A} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is injective.

If $m > n$, we have an under-determined linear system, whose minimal solution ($\|\mathbf{g}\|^2 = \sum g_j^2 \rightarrow \min!$) is computed using the Moore-Penrose pseudo-inverse,

$$\mathbf{g} = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{c} =: \mathbf{M} \mathbf{c}.$$

The numerical quality of this procedure is determined by the conditioning number γ , the ratio of the largest to the smallest eigenvalue of $\mathbf{A} \mathbf{A}^T$ (these are the singular values of \mathbf{A}). Observe that at no point we did we impose on \mathbf{g} that it should be an actual probability vector, that is

$$g_j \geq 0, \quad \sum_j g_j = 1.$$

In fact, from the consistency of the European call option prices used as input (assuming they come from a zero-correlation world so that Hull-White applies), these conditions are automatically satisfied. In practice, although in principle we can find a discretization with sufficiently many points sufficiently finely spaced that these conditions are approximately satisfied, it may make sense to add these conditions as explicit constraints.

Remark 19. *One might wonder whether or not the Moore-Penrose pseudo inverse procedure generates a result that is stable with respect to small perturbations in the input option prices. In fact, finiteness of the matrix norm of \mathbf{M} guarantees stability of the solution with respect to such small perturbations.*

8.2 Variance call options revisited

With the same notation as before, pricing a variance-call struck at K now amounts to computing

$$\sum_i (v_i - K)^+ g_i =: C_{var}(K)$$

On the other hand, in the context of a stochastic volatility model, variance-calls may be priced by Monte Carlo simulation of

$$\mathbb{E} \left[\left(\left(\int_0^T \sigma^2(s, \omega) ds \right) - K \right)^+ \right] =: \tilde{C}_{var}(K)$$

or by solving a PDE using the ideas of Section 4.

Alternatively, if the characteristic function of quadratic variation (total realized variance) is known in closed-form (as it is for many commonly-used models), we may apply the Fourier transform techniques of Carr and Madan (1999).

Example 20. *Consider the Heston model with the well-known Bakshi, Cao and Chen parameters (see Bakshi, Cao, and Chen (1997) and the references therein). Heston volatility of volatility η is 0.39. We price n one-year European calls with reasonably spread out (log-)strikes in increments of Δk . We also pick m levels of total variance in increments of Δv . We choose*

$$n = 5, \Delta k = 0.14, m = 45, \Delta v = 0.005$$

(this choice actually minimizes the conditioning number γ from the last section). This allows us to compute $C_{var}(K)$ as explained above.

On the other hand, by applying the bond pricing formula of Cox, Ingersoll, and Ross (1985), we obtain the characteristic function of quadratic variation in the Heston model.

$$\phi_T(u) = \mathbb{E}[e^{iu(x)_T}] = \mathbb{E}[e^{iu \int_0^T v_t dt}]$$

As noted earlier, we can then apply the Fourier transform methods of Carr and Madan (1999) to compute the fair value of a variance call. With the same choice of parameters, we obtain the nearly perfect fit shown in Figure 4. The maximum numerical error in our regularization procedure in this case was 0.00043 which is around 1% of the variance swap value.

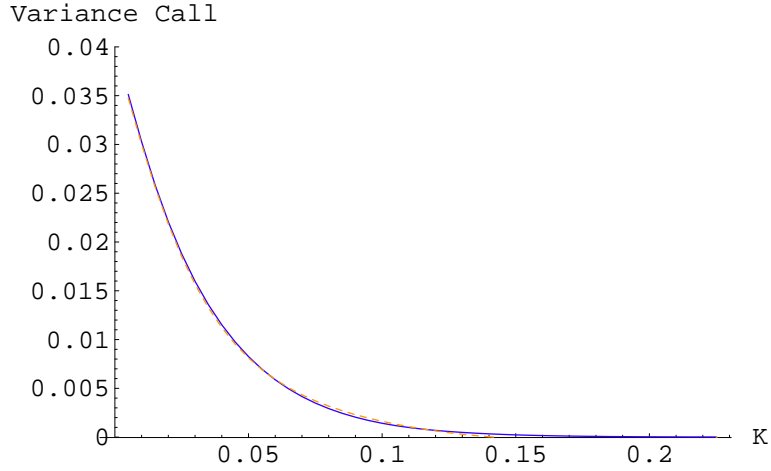


Figure 4: Value of one-year variance call vs variance strike K with Bakshi, Cao and Chen Heston model parameters. The solid blue line is the Fourier transform computation. The dashed red line comes from our regularization procedure.

8.3 Consistency with variance swaps

Finally, we connect our regularization scheme results to the variance call “weights” studied earlier. Specifically,

$$\begin{aligned}
 C^{var}(K) &= \sum_{i=1}^m (v_i - K)^+ g_i. \\
 &= \sum_{i=1}^m \left(\sum_{j=1}^n (v_i - K)^+ M_{ij} \right) c_j \\
 &=: \sum_{j=1}^n c_j w_j[K]
 \end{aligned}$$

yields a discretization $w_j[K]$ of the continuous (but not defined) weights $w_{call}(k, K)$.

Setting $K = 0$, we retrieve the (well-defined) variance swap weights $w_j[0]$.

Then we have

$$\begin{aligned}
C^{var}(0) &= \mathbb{E}[\langle x \rangle_T] \\
&= 4 \int_0^\infty c(k) e^{k/2} \cosh[k/2] dk \\
&\simeq \sum_{j=1}^n c_j 4 e^{k_j/2} \cosh[k_j/2] (\Delta k)_j.
\end{aligned}$$

Now we can enforce consistency between calls on variance and variance swaps by requiring that

$$4 e^{k_j/2} \cosh[k_j/2] (\Delta k)_j = \sum_{i=1}^m v_i M_{ij} \quad \text{for all } j = 1, \dots, n$$

which generates n constraints on our choice of $\{v_i, k_j\}$ pairs. Of course, the set of log-strikes k_j will be dictated in practice by what is available in the (zero-correlation) market.

9 Relaxing the zero-correlation assumption

Throughout this paper, we have depended on the assumption of zero-correlation between quadratic variation and underlying returns. However, in the context of stochastic volatility for example, the prices of volatility derivatives clearly do not depend on the correlation assumption.

To make this observation concrete, suppose we were to fit a stochastic volatility model such as the Heston model to European option prices. We would obtain values for all the parameters of the model including the correlation ρ . Suppose we were to take that same model and regenerate option prices with a different value of ρ (*e.g.* $\rho = 0$): clearly the fair values of volatility derivatives would not change.

We further observe that as noted in Lee (2001) for example the volatility smile must be symmetric around the at-the-money forward strike ($k = 0$). This leads us to propose a recipe⁶ for computing the fair value of volatility derivatives under the assumption of continuous paths of the underlying but with non-zero correlation between quadratic variation and returns:

⁶The idea of rotating the smile is due to Jining Han.

- Fit a model to European option prices of a given expiration. This model could either be a proper model or merely a functional form for implied volatility as a function of strike.
- Compute the fair value of a variance swap from the fit.
- Change the parameters of the model in such a way that the volatility smile retains its shape and the fair value of the variance swap is preserved but the at-the-money volatility skew is zero⁷. Effectively, rotate the smile while retaining its overall level. For example, in a stochastic volatility model, set the correlation ρ to zero.
- Apply the results of this paper to compute the values of other volatility derivatives under the zero-correlation assumption.

10 Concluding remarks

In this paper, we have studied in detail the formal expression for the value of volatility derivatives in terms of the prices of standard European options proposed by Carr and Lee under the zero-correlation assumption. After showing that the formal expression for the weights that they derive diverges in many cases of interest, we showed how to determine the value of volatility derivatives as the solution of a linear system using the Moore-Penrose pseudo-inverse.

Taken together with our observations on the extension to non-zero correlation, our results strongly suggest that under diffusion assumptions, and given the prices of European options of all strikes and expirations, the values of volatility derivatives are uniquely determined.

We emphasize here that if indeed we had a proper model with dynamics that we really believed in, neither the Carr-Lee results nor our proposed methods for their implementation would be of great interest; we could value volatility derivatives directly (by Monte Carlo for example). On the contrary, the value of this work lies in its applicability to the case where all we have is a parameterization of the implied volatility surface.

Finally, it should be clear from the graphs of the weights that we presented that this work does not concern itself with how to hedge volatility

⁷We will show how to make this more precise in a forthcoming paper.

derivatives. All we have done is to indicate how they should be priced relative to the prices of European options on the underlying.

References

- Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997, Empirical performance of alternative option pricing models, *The Journal of Finance* 52, 2003–2049.
- Carr, Peter, and Roger W. Lee, 2003, Robust replication of volatility derivatives, Discussion paper Courant Institute, NYU and Stanford University, Department of Mathematics.
- Carr, Peter, and Dilip Madan, 1998, Towards a theory of volatility trading, in Robert A. Jarrow, ed.: *Volatility: New Estimation Techniques for Pricing Derivatives*. chap. 29, pp. 417–427 (Risk Books: London).
- Carr, Peter, and Dilip B. Madan, 1999, Option valuation using the fast fourier transform, *The Journal of Computational Finance* 2, 61–73.
- Cont, Rama, and José da Fonseca, 2002, Dynamics of implied volatility surfaces, *Quantitative Finance* 2, 45–60.
- Cox, John C., Jonathan E. Ingersoll, and Steven A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–407.
- Gatheral, Jim, 2003, Case studies in financial modeling lecture notes, http://www.math.nyu.edu/fellows_fin_math/gatheral/case_studies.html.
- Gradshteyn, I. S., and I. M. Ryzhik, 1980, *Table of Integrals, Series and Products* (Academic Press: New York, NY, USA).
- Hull, John, and Alan White, 1987, The pricing of options with stochastic volatilities, *The Journal of Finance* 19, 281–300.
- Lee, Roger W., 2001, Implied and local volatilities under stochastic volatility, *International Journal of Theoretical and Applied Finance* 4, 45–89.
- Monk, P., 2003, Basic theory of linear ill-posed problems, <http://www.inria.fr/actualites/colloques/2003/pnla/monk1.pdf>.

Schwartz, Laurent, 1959, *Étude des sommes d'exponentielles* (Hermann: Paris).