Numerical Methods for Stochastic Volatility: Fourier Methods, PDEs and Monte Carlo

Iain J. Clark

27 February, 2008

Imperial College Finance and Stochastics Seminar
Outline

- Introduction to FX options: vanillas and liquid exotics
- Heston’s stochastic volatility model in FX
- Developing market intuition about Heston
- Fast semianalytic techniques: characteristic functions
- Basic calibration of the model to the market smile
- Pricing using Monte Carlo
- Pricing using numerical finite differences in 2D
Introduction: liquid FX exotics and deviation from B-S prices

- Short dated FX options (out to 3Y or so):
  - ~90% vanillas
  - ~9% binaries/barriers [continuously monitored]
  - ~1% other complex exotics

- Vanillas
  - Almost solely OTC – not exchange traded
  - European, not American style. Value depends only on $S_T$

- Binaries/Barriers:
  - Common criticism of options is that they appear quite expensive to the buyer. Leads to demand for cheaper alternatives – e.g. knock-out options.
Vanillas in FX

- B-S inadequate – a single $\sigma$ will not match all vanillas in market
  - Structural deviation from B-S prices: volatility smile
- Benchmark FX instruments: 5 strikes per tenor
  - 10-delta put strike $K$ chosen so that $\Delta_p = -0.10$
  - 25-delta put strike $K$ chosen so that $\Delta_p = -0.25$
  - ATM option either* ATMF ($K=F$) or D-N ($\Delta_p + \Delta_c = 0$)
  - 25-delta call strike $K$ chosen so that $\Delta_c = +0.25$
  - 10 delta call strike $K$ chosen so that $\Delta_c = +0.10$
- Smiles are generally (JPY and EMs aside) reasonably symmetric

* depends on market convention.

Barriers in FX

- Introducing path dependency can make a vanilla option substantially cheaper

- European call: \( V_T = (S_T - K)^+ \)
  \[ = (S_T - K) \mathbf{1}_{\{S_T \geq K\}} \]

- \( m_T \) and \( M_T \) denote the minimum and maximum (resp.) of \( S_t \) over the time interval \([0,T]\)

- Cheaper alternatives:
  - Regular KO \( V_T = (S_T - K) \mathbf{1}_{\{S_T \geq K\}} \mathbf{1}_{\{m_T > L\}} \)
  - Reverse KO \( V_T = (S_T - K) \mathbf{1}_{\{S_T \geq K\}} \mathbf{1}_{\{M_T < U\}} \)
  - Double KO \( V_T = (S_T - K) \mathbf{1}_{\{S_T \geq K\}} \mathbf{1}_{\{m_T > L\}} \mathbf{1}_{\{M_T < U\}} \)
Binaries in FX

- Distant OTs (TV < 20%) typically trade above TV
- Nearer OTs typically trade below TV
  - Structural deviation from B-S prices: \textit{binary moustache}

\[
V_T = 1_{\{m_T \leq L\}} \quad \text{and} \quad V_T = 1_{\{M_T \geq U\}}
\]
Flow exotics – depend on which processes?

- All the flow exotics have value functions $V_T$ which depend at most on three of the following processes:

<table>
<thead>
<tr>
<th>Product</th>
<th>$S_T$</th>
<th>$m_T$</th>
<th>$M_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>European vanilla</td>
<td>YES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OT or NT [downside]</td>
<td></td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>OT or NT [upside]</td>
<td></td>
<td></td>
<td>YES</td>
</tr>
<tr>
<td>KI or KO [downside]</td>
<td>YES</td>
<td></td>
<td>YES</td>
</tr>
<tr>
<td>KI or KO [upside]</td>
<td>YES</td>
<td></td>
<td>YES</td>
</tr>
<tr>
<td>DNT or DT</td>
<td></td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>DKI or DKO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

- Ideally obtain market implied joint pdf of $\{S_T, m_T, M_T\}$
Heston’s Stochastic Volatility Model in FX

- In practice: seek a model that accurately describes (a) volatility smile, (b) binary moustaches [i.e. marginals for $S_T$, $m_T$ and $M_T$] and (c) DNT prices.
- The Heston model is a model for stochastic variance

\[
\begin{align*}
    dS_t &= \mu S_t \, dt + \sigma_t S_t \, dW^{(1)}_t, \quad \mu = r_d - r_f \\
    dV_t &= \kappa (m - V_t) \, dt + \alpha \sqrt{V_t} \, dW^{(2)}_t, \quad \sigma_t = \sqrt{V_t} \\
    \left\langle dW^{(1)}_t, dW^{(2)}_t \right\rangle &= \rho dt
\end{align*}
\]

- What intuition should we attach to the model parameters?

Stochastic volatility and vol convexity

- All stochastic volatility/variance models generate smiles, by correctly pricing in vol convexity
- Hull/White analysis: if processes driving spot and variance are uncorrelated

\[
PV = \int_0^\infty TV \bigg|_{\sigma=\sqrt{v}} \frac{f_V(v)}{\sqrt{v}} \, dv
\]

where \( f_V(v) \) is the pdf of average variance \( \bar{V} = 1/T \int_0^T \sigma_u^2 du \) over time interval \([0,T]\) and \(TV|_\sigma\) is the Black-Scholes price with constant volatility \(\sigma\).

Intuition: implied ATM vol structure

- TV of the ATMF option is $TV \approx 0.4\sigma\sqrt{T} = 0.4\sqrt{VT}$
- Consider a driftless stochastic variance process.
- In that case the expectation of $V$ is just $V_0$
- Concavity of square root function means that PV decreases as the volatility becomes increasingly dispersed around $V_0$
- ATMF price under a driftless stochastic variance process decreases as volatility of variance increases.
- Implied ATMF vol is $\sqrt{V_0}$ adjusted downwards for this variance effect
Intuition: implied wing structure

- ATMF options are linear in volatility but wing options are convex.
  - Hence increasing volatility increases the implied smile.

![Graph showing the difference between ATMF and wing options in terms of volatility linearity and convexity.](image)
### Intuition: effects of Heston model parameters

- Five parameters have quite different effects on the shape of implied volatility surface generated

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial variance $V_0$</td>
<td>Fixes overall level of implied ATM vol</td>
</tr>
<tr>
<td>Vovariance $\alpha$</td>
<td>Generates volatility smile as $\alpha$ increases</td>
</tr>
<tr>
<td>Spot/Variance correlation $\rho$</td>
<td>Generates volatility skew for nonzero $\rho$</td>
</tr>
<tr>
<td>Mean reversion rate $\kappa$</td>
<td>Combined effect: increasing $\kappa$, term structure of implied ATM vol shifts in direction of $m^{1/2}$ &amp; smile flattens</td>
</tr>
<tr>
<td>Mean reversion level $m$</td>
<td></td>
</tr>
</tbody>
</table>
Risk neutral pricing

- Let asset have price process \( dS_t = (r_d - r_f)S_t dt + \sigma_t S_t dW_t \)
- Black-Scholes formula for European call option
  - take discounted expectation of payout under domestic RN measure
  - use Girsanov to change from domestic RN to foreign RN measure

\[
C(S, T) = e^{-r_d T} E^d [(S_T - K)^+] \\
= e^{-r_d T} E^d [(S_T - K)1_{\{S_T \geq K\}}] \\
= e^{-r_d T} E^d [S_T 1_{\{S_T \geq K\}}] - Ke^{-r_d T} E^d [1_{\{S_T \geq K\}}] \\
= S_0 e^{-r_f T} E^f [1_{\{S_T \geq K\}}] - Ke^{-r_d T} E^d [1_{\{S_T \geq K\}}] \\
= S_0 e^{-r_f T} P^f [S_T \geq K] - Ke^{-r_d T} P^d [S_T \geq K] \\
= S_0 e^{-r_f T} N(d_1) - Ke^{-r_d T} N(d_2)
\]
Risk neutral pricing

- Express in terms of asset log-returns $X_t = \ln(S_t)$
  - $S_t$ constrained to $[0, \infty)$ but $X_t$ defined on $(-\infty, \infty)$
  - $X_t$ in BS world follows an ABM and has normally distributed marginals (easier to compute characteristic functions)

- Call price is given by

$$C(S, T) = e^{-r_d T} \mathbb{E}^d [(\exp(X_T) - K) \mathbb{1}_{\{X_T \geq K\}}]$$

$$= S_0 e^{-r_f T} \mathbb{P}^f [X_T \geq \ln K] - Ke^{-r_d T} \mathbb{P}^d [X_T \geq \ln K]$$
Pricing in Fourier space

- Denote pdfs in foreign and domestic risk-neutral measures by \( f_{X_T}^d(x) \) and \( f_{X_T}^f(x) \) respectively. \( h(x) \) is the payout function.

\[
E^{d;f} [h(X_T)] = \int_{-\infty}^{\infty} h(x) f_{X_T}^{d;f}(x) \, dx
\]

an inner product in \( L^2(-\infty, \infty) \)

- Parseval’s theorem: inner products preserved under Fourier transforms

\[
\hat{h}(\phi) = \int_{-\infty}^{\infty} e^{i\phi x} h(x) \, dx
\]

characteristic function of \( X_T \)

\[
\hat{f}_{X_T}^{d;f}(\phi) = E^{d;f} [e^{i\phi X_T}]
\]

\[
= \int_{-\infty}^{\infty} e^{i\phi x} f_{X_T}^{d;f}(x) \, dx
\]
Pricing in Fourier space

- Clearly, expectations can be computed in $\phi$– space

\[ E^{d;f} [h(X_T)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{h}(\phi) \hat{f}_{X_T}^d (\phi) \, d\phi \]

- Now we can calculate the cdf’s – which take the form

\[ P^{d;f} [X_T \geq \ln K] = E^{d;f} [1_{\{X_T \geq \ln K\}}] \]

- The Fourier transform of $h(x) = 1_{\{x \geq \ln K\}}$ is given by

\[ \hat{h}(\phi) = \int_{-\infty}^{\infty} 1_{\{x \geq \ln K\}} e^{ix\phi} \, dx = \int_{\ln K}^{\infty} e^{ix\phi} \, dx = \left. \frac{e^{ix\phi}}{i\phi} \right|_{\ln K}^{\infty} \]

- Issues:
  - The limit as $x \to \infty$ of $e^{ix\phi}$ isn’t formally defined
  - Complex pole at the origin ($\phi = 0$)
  - No major impediment
Pricing in Fourier space

- Fourier inversion formula results (Heston; Bates; Bakshi et al.)

\[ P^{d:f}[X_T \geq \ln K] = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \text{Re}\left[ \hat{f}^{d:f}_{X_T}(\phi) \frac{\exp(-i\phi \ln K)}{i\phi} \right] d\phi \]

- Need to compute the c.f. \( \hat{f}^{d:f}_{X_T}(\phi) \) of the log-return asset process.
  - This is where things get interesting because it can be calculated analytically – e.g. for spot processes driven by Heston stochastic volatility process.

- First, we work through Black-Scholes case, then look at Heston.


Pricing in Fourier space – Black-Scholes case

- Assume volatility constant. Spot follows 
\[ dS_t = (r_d - r_f)S_t dt + \sigma S_t dW_t \]
and the log-returns therefore follow 
\[ dX_t = (r_d - r_f - \frac{1}{2} \sigma^2) dt + \sigma dW_t \]
with solution 
\[ X_t = X_0 + (r_d - r_f - \frac{1}{2} \sigma^2) t + \sigma W_t \]
- This is normal with mean \( X_0 + (r_d - r_f - \frac{1}{2} \sigma^2) t \) and variance \( \sigma \sqrt{t} \)
- For a \( N(\mu, \sigma^2) \) r.v., with pdf 
\[ f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \]
the c.f. is given by (complete the square) 
\[ \hat{f}_X(\phi) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{i\phi x} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) dx = \exp(i\mu\phi - \frac{1}{2} \sigma^2 \phi^2) \]
Pricing in Fourier space – Black-Scholes case

- We now have all we need to price any European option using Fourier integration in Black-Scholes. For example, a call is priced at

\[ C(S, T) = S_0 e^{-r_f T} \mathbb{P}^f [X_T \geq \ln K] - K e^{-r_d T} \mathbb{P}^d [X_T \geq \ln K] \]

- where

\[ \mathbb{P}^{d:f} [X_T \geq \ln K] = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \text{Re} \left[ \hat{f}_{X_T}^{d:f}(\phi) \frac{\exp(-i\ln K)}{i\phi} \right] d\phi \]

- For Black-Scholes:

\[ \hat{f}_{X_T}^{d:f}(\phi) = \exp(i[(r_d - r_f \pm \frac{1}{2} \sigma^2)T + X_0]) \phi - \frac{1}{2} \sigma^2 T \phi^2 \]

- Questions: What do the integrands look like? For stochastic volatility models, how might we obtain the required characteristic functions?
Examination of integrands – Black-Scholes case

- **ATMF**: $S_0=1$, $\sigma = 10\%$, $r_d=0$, $r_f=0$, $K=1$, $T=1$

  ![Graph showing $d_1$ and $d_2$ values]

  - $d_1 > 0$, so $N(d_1) > \frac{1}{2}$
  - $d_2 < 0$, so $N(d_2) < \frac{1}{2}$

- The symmetry reminds us of $d_1 = -d_2$ for the ATMF case
Examination of integrands – Black-Scholes case

- **Market:** $S_0=1$, $\sigma=10\%$, $r_d=0$, $r_f=0$
- **ITM:** $K=0.8$, $T=1$
- **OTM:** $K=1.2$, $T=1$

- RN probabilities of exceeding 0.8 at expiry > 0.5: integrals are positive
- RN probabilities of exceeding 1.2 at expiry < 0.5: integrals are negative
Examination of integrands – Black-Scholes case

- **ATMS:** \( K=1, \ T=1, \ S_0=1 \)
- **Market:** \( \sigma=10\%, \ r_d=8\%, \ r_f=8\% \)

- **Market:** \( \sigma=10\%, \ r_d=8\%, \ r_f=0\% \)

- Unaffected by changes in rates which maintain same IR differential

- Affected by changes in IR differential
  - forward rate moves, affects RN probs.
Examination of integrands – Black-Scholes case

- Following is a sensible choice: $\phi_{\text{max}} = Q/(\sigma\sqrt{T})$
  Makes sense as $\sigma\sqrt{T}$ is dimensionless.
- In fact you can see this analytically from
  $$
P[d,f][X_T \geq \ln K] = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \text{Re} \left[ \hat{f}_{d,f} \left( \phi \right) \frac{\exp(-i\phi \ln K)}{i\phi} \right] d\phi$$
  where
  $$\hat{f}_{d,f} \left( \phi \right) = \exp(i[(r_d - r_f) \pm \frac{1}{2} \sigma^2 T + X_0] \phi - \frac{1}{2} \sigma^2 T \phi^2)$$
  $$\text{Re} \left[ \hat{f}_{d,f} \left( \phi \right) \frac{\exp(-i\phi \ln K)}{i\phi} \right] = \frac{\exp(-\frac{1}{2} \sigma^2 T \phi^2)}{\phi} \times \sin\left(\left[(r_d - r_f) \pm \frac{1}{2} \sigma^2 T + X_0 - \ln K\right] \phi\right)$$
  - Choice of $Q$ – somewhere between 2 and 5 is generally sufficient
  - Simple trapezoidal integration on $[0, \phi_{\text{max}}]$ is OK in practice.
  - Backtest against exact Black-Scholes price to make sure integration is OK
Examination of integrands – Black-Scholes case

- How oscillatory can these integrals get?
- Difficult cases: \( K = \{0.5, 0.66, 0.75\} \), \( S_0 = 1 \), \( \sigma = 3\% \), \( r_d = 0 \), \( r_f = 0 \), \( T = 1/12 \)

Deeply OTM options with small \( \sigma \sqrt{T} \) are hard to handle – but they’re not worth much
Computing price with just one Fourier integral

- Wanted to show how the domestic and foreign risk-neutral probabilities can be calculated using Fourier methods and related back to $N(d_1)$ and $N(d_2)$. Easy to visualise.
- In fact, for European calls and puts, the computation can be performed using a single Fourier integral along a contour in the complex plane – see Lewis, p.37 for details
- This is more efficient as only one integral to compute.
- Need to use inversion formula (2.5) in Lewis.
- Recommend starting with the 2 integral technique, then implement the single integral technique as a companion scheme. Ensure results agree.


Out of print but Chapters 1 and 2 available at www.optioncity.net
Computing price with just one Fourier integral

- Integrals are less singular for Europeans using the 1-integral technique
- Difficult cases again: $S_0=1$, $\sigma=3\%$, $r_d=0$, $r_f=0$, $T=1/12$.
  Horizontal axis is $\log(\phi)$ – hence integral on positive half-line OK

Single integration should recover Black-Scholes price accurately
Pricing in Fourier space – Heston stochastic volatility model

- This seems like a lot of extra work when we can just go directly to Black-Scholes closed form formulae. What’s the point?
- **Key point:** this method *extends* to stochastic volatility models.
- How? Go back to the definition of characteristic function

\[
\hat{f}_{X_T}^d (\phi) = \mathbb{E}^d \left[ e^{i\phi X_T} \right] = \int_{-\infty}^{\infty} e^{i\phi x} f_{X_T}^d (x) dx
\]

- Following Section 2.2.2 of Zhu (2000), compute in risk neutral measures

\[
\hat{f}_{X_T}^d (\phi) = \mathbb{E}^d \left[ e^{i\phi X_T} \right]
\]

\[
\hat{f}_{X_T} f (\phi) = e^{(r_f-r_d)T} \mathbb{E}^d \left[ \frac{S_T}{S_0} e^{i\phi X_T} \right] = e^{(r_f-r_d)T} e^{-X_0} \mathbb{E}^d \left[ e^{(i\phi+1)X_T} \right]
\]


Pricing in Fourier space – Heston stochastic volatility model

- Clearly we need to compute $E^d[e^{(i\phi+j)X_T}], \ j \in \{0,1\}$

- Consider by way of example the Heston model.

$$dS_t = \mu S_t dt + \sqrt{V_t} S_t dW^{(1)}_t, \quad \langle dW^{(1)}_t, dW^{(2)}_t \rangle = \rho dt$$

$$dV_t = \kappa (m - V_t) dt + \alpha \sqrt{V_t} dW^{(2)}_t$$

- Log-returns:

$$dX_t = (\mu - \frac{1}{2} V_t) dt + \sqrt{V_t} dW^{(1)}_t$$

- Integrate the log-return process to get

$$X_T = X_0 + \mu T - \frac{1}{2} \int_0^T V_t dt + \int_0^T \sqrt{V_t} dW^{(1)}_t$$
Pricing in Fourier space – Heston stochastic volatility model

A few pages to show how Girsanov works here. Substituting $X_T$ in:

$E^d [e^{(i\phi + j)X_T}] = e^{(i\phi + j)(X_0 + \mu T)} E^d [\exp(-\frac{1}{2} \int_0^T V_t dt + \int_0^T \sqrt{V_t} dW_t^{(1)})]$ 

Several terms cancel out, leaving the c.f.s ($j=0$ for $d$, $j=1$ for $f$)

$\hat{f}^{d,f}_{X_T}(\phi) = e^{i\phi[X_0 + \mu T]} E^d \left[ \exp\{(i\phi + j)(-\frac{1}{2} \int_0^T V_t dt + \int_0^T \sqrt{V_t} dW_t^{(1)})\} \right]$ 

Apply Cholesky decomposition

$dW_t^{(1)} = \rho dW_t^{(2)} + \overline{\rho} dW_t^{(-2)}$ 

$\overline{\rho} = \sqrt{1 - \rho^2}$

to obtain

$E^d \left[ \exp\{(i\phi + j)(-\frac{1}{2} \int_0^T V_t dt + \rho \int_0^T \sqrt{V_t} dW_t^{(2)} + \overline{\rho} \int_0^T \sqrt{V_t} dW_t^{(-2)})\} \right]$
Pricing in Fourier space – Heston stochastic volatility model

- This can be simplified by some Girsanov sleight of hand

\[
E^d \left[ \exp \left\{ (i\phi + j) \left( -\frac{1}{2} \int_0^T V_t dt + \rho \int_0^T \sqrt{V_t} dW_t^{(2)} + \bar{\rho} \int_0^T \sqrt{V_t} dW_t^{(-2)} \right) \right\} \right] \\
= E^d \left[ e^{\int_0^T (i\phi + j)(-\frac{1}{2} \int_0^T V_t dt + \rho \int_0^T \sqrt{V_t} dW_t^{(2)})} \right] \\
= E^d \left[ \frac{dQ^d}{dQ^d^*} = e^{\int_0^T \lambda_t dW_t^{(-2)} - \frac{1}{2} \int_0^T \lambda_t^2 dt} \right] \\
= E^d * \left[ e^{\frac{1}{2}(i\phi + j)^2 \bar{\rho}^2 \int_0^T V_t dt} e^{(i\phi + j)(-\frac{1}{2} \int_0^T V_t dt + \rho \int_0^T \sqrt{V_t} dW_t^{(2)})} \right] \\
= E^d \left[ e^{\left[ \frac{1}{2}(i\phi + j)^2 \bar{\rho}^2 - \frac{1}{2} (i\phi + j) \right] \int_0^T V_t dt} e^{\rho (i\phi + j) \int_0^T \sqrt{V_t} dW_t^{(2)}} \right]
\]
Pricing in Fourier space – Heston stochastic volatility model

- Integrating the Heston process we obtain

\[ V_T = V_0 + \kappa n T - \kappa \int_0^T V_t \, dt + \alpha \int_0^T \sqrt{V_t} \, dW_t^{(2)} \]

- So obviously

\[ \int_0^T \sqrt{V_t} \, dW_t^{(2)} = \frac{V_T - V_0 - \kappa n T}{\alpha} + \frac{\kappa}{\alpha} \int_0^T V_t \, dt \]

- We obtain

\[
E^d \left[ e^{\frac{1}{2} \left(i \phi + j\right)^2 \bar{\rho}^2 - \frac{1}{2} \left(i \phi + j\right)}} \int_0^T V_t \, dt \, e^{\rho \left(i \phi + j\right)} \int_0^T \sqrt{V_t} \, dW_t^{(2)} \right]
\]

\[ = E^d \left[ e^{\frac{1}{2} \left(i \phi + j\right)^2 \bar{\rho}^2 - \frac{1}{2} \left(i \phi + j\right) + \rho \kappa / \alpha} \int_0^T V_t \, dt \, e^{\rho \left(i \phi + j\right)(V_T - V_0 - \kappa n T) / \alpha} \right] \]

and by suitably defining terms \( s_1^{(j)} \), \( s_2^{(j)} \) (given on next slide) we obtain

\[ \hat{f}_{X_T}^d (\phi) = e^{i \phi \left[X_0 + \mu T\right]} \exp \left( - (V_0 + \kappa n T) s_2^{(j)} \right) E^d \left[ \exp \left( - s_1^{(j)} \int_0^T V_t \, dt + s_2^{(j)} V_T \right) \right] \]
Zhu (2000) computes this expectation, obtaining the following characteristic functions (\(j=0\) for \(d\), \(j=1\) for \(f\)) – note I use \(\mu = r_d - r_f\)

\[
\hat{f}_{X_T}^{d,f} (\phi) = \exp(i\phi[ X_0 + \mu T] - (V_0 + \kappa m T) s_2^{(j)}) \exp \left(A^{(j)} V_0 + C^{(j)} \right)
\]

where

\[
s_1^{(j)} = -(i\phi + j)(-\frac{1}{2} + \frac{\rho \kappa}{\alpha} + \frac{1}{2} (i\phi + j)(1 - \rho^2))
\]

\[
s_2^{(j)} = (i\phi + j) \rho / \alpha
\]

\[
\gamma_1^{(j)} = \sqrt{\kappa^2 + 2\alpha^2 s_1^{(j)}}
\]

\[
\gamma_2^{(j)} = 2 \gamma_1^{(j)} \exp(-\gamma_1^{(j)} T) + (\kappa + \gamma_1^{(j)} - \alpha^2 s_2^{(j)}) (1 - \exp(-\gamma_1^{(j)} T))
\]

\[
A^{(j)} = \left[\gamma_1^{(j)} s_2^{(j)} (1 + \exp(-\gamma_1^{(j)} T)) - (1 - \exp(-\gamma_1^{(j)} T))(2 s_1^{(j)} + \kappa s_2^{(j)})\right] / \gamma_2^{(j)}
\]

\[
C^{(j)} = 2 \kappa m \alpha^{-2} \ln \left[2 \gamma_1^{(j)} \exp \left(\frac{1}{2} (\kappa - \gamma_1^{(j)}) T\right) / \gamma_2^{(j)} \right]
\]
Examination of integrands – stochastic volatility models

- ATMF: $K=1.0$, $S_0=1$, $\sigma = 10\%$, $r_d=0$, $r_f=0$, $T=1$
- Heston with vovol $\alpha = 5\%$, 10\% or 20\%
Examination of integrands – stochastic volatility models

- Wings: $K=0.94$, $S_0=1$, $\sigma = 10\%$, $r_d=0$, $r_f=0$, $T=1$
  Heston with vovol $\alpha=5\%$, $10\%$ or $20\%$

RN prob of exceeding 0.94 at $T$ is greater as vovariance increases. Narrow shoulders.
Examination of integrands – stochastic volatility models

- Distant wings: $K=0.80, S_0=1, \sigma = 10\%, r_d=0, r_f=0, T=1$
  Heston with vovol $\alpha = 5\%, 10\%$ or $20\%$

RN prob of exceeding 0.8 at $T$ is less as vovariance increases. RN prob of being below 0.8 at expiry greater as vovariance gets higher. Fat tails.
Examination of implied pdfs – stochastic volatility models

- Implied pdfs behave as expected – chart generated by pricing up a strip of Arrow-Debrue securities – see Lewis, p. 37 for details

- $S_0=1$, $\sigma = 10\%$, $r_d=0$, $r_f=0$, $T=1$
Examination of integrands – stochastic volatility models

- Heston implied pdfs [vvar=10%] become skewed with correlation
- $S_0=1$, $\sigma=10\%$, $r_d=0$, $r_f=0$, $T=1$
Examination of implied smiles – stochastic volatility models

- It is quite clear that smiles are generated by increasing volatility.
- $S_0=1$, $\sigma=10\%$, $r_d=0$, $r_f=0$, $T=1$
Examination of implied smiles – stochastic volatility models

- Skews are generated by nonzero values for the correlation between spot and variance

![Graph showing the relationship between implied smiles and correlation]

- BS
- Heston, vovar=10%, corr=-0.25
- Heston, vovar=10%, corr=+0.25
Basic calibration of the model to market smile

- Heston model has no problem generating smiles and skews
- SV calibration is a fairly simple optimisation exercise using semianalytic methods discussed in this talk.
- **Terminal calibration:** take as inputs the volatilities at three strikes (25-d-P, ATM, 25-d-C), at one expiry time $T$. Lock down $\kappa$ and $m$. Attempt to minimise objective function which measures the sum of squares of the errors in the vol by varying $V_0$, $\rho$, $\alpha$. The objective function calculates Heston prices using the characteristic function method and backs out implied volatilities.
- **Term structure calibration:** With suitably chosen mean reversion parameters $\kappa$ and $m$, possible to generate upward sloping or downward sloping ATM volatility surfaces. Increasing mean reversion causes smiles to flatten and diminish as the mean reversion of variance takes effect.
Pricing in Heston using Monte Carlo

- Monte Carlo is always a useful check for testing other algorithms against.
  - Draw samples $\Delta W$ from a standardised 2D bivariate normal distribution at timepoints $\{0, \Delta t, 2\Delta t, \ldots T-\Delta t\}$
  - Compute drift vector $\mu_t$ and volatility vector $\Sigma_t$ at time $t_i$ (see below)
  - Integrate the factor from its initial value $X_0=(\log(S_0), V_0)$ out to time $T$

$$X_{t+\Delta t} = X_t + \mu_t \Delta t + (\Sigma_t \Delta W) \sqrt{\Delta t}$$

$$\mu_t = \left(r_d - r_f - \frac{1}{2} \kappa \left[m - V_t\right]\right)$$

$$\Sigma_t = \begin{bmatrix}
V_t^{1/2} & 0 \\
\alpha \rho V_t^{1/2} & \alpha \sqrt{1-\rho^2} V_t^{1/2}
\end{bmatrix}$$

- Evaluate payoff (function of $X_T$) at time $T$. Integrate over all simulations.
Numerical solution of the Heston PDE by finite differences

- Characteristic function technique can be used for any option that has value at $T$ as a function of $S_T$. Europeans, digitals, etc…
- Rules out all path dependent options (barriers and binaries) and early exercisable options.
- Heston model can be solved using 2D PDE for these products with suitable boundary conditions
  - Approximate spatial & temporal differences with mesh differences

$$\frac{1}{2} V S^2 \frac{\partial^2 U}{\partial S^2} + \rho \sigma V S \frac{\partial^2 U}{\partial S \partial V} + \frac{1}{2} \alpha^2 V \frac{\partial^2 U}{\partial V^2}$$

$$+ (r_d - r_f) S \frac{\partial U}{\partial S} + \{\kappa[m - V] - \lambda\} \frac{\partial U}{\partial V} - r_d U + \frac{\partial U}{\partial t} = 0$$

* $\lambda$ denotes the market price of volatility risk
Black-Scholes volatility (constant)

- Imagine a solution diffusing through the **gaps** in the following uniform mesh.
  Note: *not* a representation of a finite difference mesh. *Schematic* illustration of diffusion.
- B-S solution obtained by diffusing the source solution backwards on the mesh.
- Analogous to tree methods.
Stochastic volatility (or variance)

- Stochastic volatility: extend from one “spatial” to two “spatial” dimensions
- Source solution diffuses more rapidly where volatility/variance is larger

More diffusion through gaps in mesh where \( V \) is large

Source solution e.g. \((S_T - K)^+\) independent of \( V \)

Less diffusion through gaps in mesh where \( V \) is small

Diffusion independent of \( S \)
Solution of 2D Heston PDE using finite differences

- Easiest to start with a 2D explicit PDE scheme. Simple to code up.
  - However this will be far too slow for anything except development
  - 2D EFD prices should converge (slowly) to Fourier & MC prices
- The standard method for these problems is the ADI [alternating direction implicit] scheme. References given below.
  - Quite useful to set up PDE engines so that the mesh can be output to files – makes it quite easy to see when there are problems with boundary conditions, or stability.
- Also helps to compare with output from 1D PDE engines (B-S)


PDE implementation

- Standard PDE schemes:
  - 1D: (i) fully explicit, (ii) fully implicit, (iii) Crank-Nicolson
  - 2D: (i) fully explicit, (ii) ADI

- Consider dimensionless pde:
  \[
  \frac{\partial U}{\partial \tau} = \sum_{i,j} A_{ij} \frac{\partial^2 U}{\partial x_i \partial x_j} + \sum_i b_i \frac{\partial U}{\partial x_i} + fU
  \]
  - transform Heston pde (slide 43) to log-spot \( x = \log(S) \) and read off convection & diffusion coefficients

- In 1D:
  \[
  U_\tau = AU_{xx} + bU_x + fU
  \]

- Apply \( x \) and \( \tau \) discretisation
  \[
  U_i^j = U(x_i, \tau_j)
  \]
PDE implementation

- Time derivative is given by:
  \[ \frac{\partial U}{\partial \tau} = \frac{U_{i}^{j+1} - U_{i}^{j}}{\Delta \tau} \]

- (Central) spatial derivatives can either be taken at \( \tau_{j} \)
  \[ \frac{\partial U}{\partial x} = \frac{U_{i}^{j+1} - U_{i-1}^{j}}{2\Delta x} \quad \frac{\partial^{2} U}{\partial x^{2}} = \frac{U_{i+1}^{j} - 2U_{i}^{j} + U_{i-1}^{j}}{\Delta x^{2}} \]

- ...or at \( \tau_{j+1} \)
  \[ \frac{\partial U}{\partial x} = \frac{U_{i}^{j+1} - U_{i+1}^{j+1}}{2\Delta x} \quad \frac{\partial^{2} U}{\partial x^{2}} = \frac{U_{i+1}^{j+1} - 2U_{i}^{j+1} + U_{i-1}^{j+1}}{\Delta x^{2}} \]

- ...or in between
  \[ \frac{\partial U}{\partial x} = \frac{\theta[U_{i+1}^{j+1} - U_{i-1}^{j+1}] + (1 - \theta)[U_{i+1}^{j} - U_{i-1}^{j}]}{2\Delta x} \]
  \[ \frac{\partial^{2} U}{\partial x^{2}} = \frac{\theta[U_{i+1}^{j+1} - 2U_{i}^{j+1} + U_{i-1}^{j+1}] + (1 - \theta)[U_{i+1}^{j} - 2U_{i}^{j} + U_{i-1}^{j}]}{\Delta x^{2}} \]
PDE implementation

- ...leading to the fully explicit scheme

\[ U_{i}^{j+1} = U_{i}^{j} + \frac{A\Delta\tau}{\Delta x^2} [U_{i+1}^{j} - 2U_{i}^{j} + U_{i-1}^{j}] + \frac{b\Delta\tau}{2\Delta x} [U_{i+1}^{j} - U_{i-1}^{j}] + f\Delta \tau U_{i}^{j} \]

- ... the fully implicit scheme

\[ U_{i}^{j+1} - \frac{A\Delta\tau}{\Delta x^2} [U_{i+1}^{j+1} - 2U_{i}^{j+1} + U_{i-1}^{j+1}] - \frac{b\Delta\tau}{2\Delta x} [U_{i+1}^{j+1} - U_{i-1}^{j+1}] - f\Delta \tau U_{i}^{j+1} = U_{i}^{j} \]

- ...or the Crank-Nicolson scheme (\( \theta = 1/2 \))

\[ U_{i}^{j+1} - \frac{A\Delta\tau}{2\Delta x^2} [U_{i+1}^{j+1} - 2U_{i}^{j+1} + U_{i-1}^{j+1}] - \frac{b\Delta\tau}{4\Delta x} [U_{i+1}^{j+1} - U_{i-1}^{j+1}] - \frac{f}{2} \Delta \tau U_{i}^{j+1} \]

\[ = U_{i}^{j} + \frac{A\Delta\tau}{2\Delta x^2} [U_{i+1}^{j} - 2U_{i}^{j} + U_{i-1}^{j}] + \frac{b\Delta\tau}{4\Delta x} [U_{i+1}^{j} - U_{i-1}^{j}] + \frac{f}{2} \Delta \tau U_{i}^{j} \]
Handling boundary conditions

- Extinguishing options (NT, DNT, KO, DKO) are easily handled by placing a Dirichlet boundary condition at the barrier level.
- Without KO barriers (e.g. Europeans without barriers), common technique is to assume 2\textsuperscript{nd} derivative vanishes on the boundaries. Hence solution is linear.
- Suffices therefore to use one-sided differences (neglect diffusion terms), to time-step the solution on the boundary
  - This is for explicit finite differences; use \( j+1 \) for IFD

\[
\frac{\partial U}{\partial x} \bigg|_{x=x_0, t=\tau_j} = \frac{U_j^1 - U_0^j}{2\Delta x} \quad \frac{\partial U}{\partial x} \bigg|_{x=x_N, t=\tau_j} = \frac{U_N^j - U_{N-1}^j}{2\Delta x}
\]

Example: 1D PDE scheme

- Black-Scholes (Crank-Nicolson)
PDE implementation

- Algebra of Crank-Nicolson can be simplified introducing a node at half-time

\[ U_i^{j+1} - \frac{A \Delta \tau}{2 \Delta x^2} \left[ U_i^{j+1} - 2U_i^j + U_i^{j-1} \right] - \frac{b \Delta \tau}{4 \Delta x} \left[ U_i^{j+1} - U_i^{j-1} \right] - \frac{f}{2} \Delta \tau U_i^j = U_i^{j+1/2} \]

\[ U_i^{j+1/2} = U_i^j + \frac{A \Delta \tau}{2 \Delta x^2} \left[ U_i^{j+1} - 2U_i^j + U_i^{j-1} \right] + \frac{b \Delta \tau}{4 \Delta x} \left[ U_i^{j+1} - U_i^{j-1} \right] + \frac{f}{2} \Delta \tau U_i^j \]

- Can be seen as equivalent to an explicit step over time interval \((\tau_j, \tau_{j+1/2})\) followed by an implicit step over time interval \((\tau_{j+1/2}, \tau_j)\).

- The ADI [alternating direction implicit] scheme, which we use for problems with \textbf{two} spatial variables, works similarly by applying an explicit step in one spatial direction, followed by an implicit step in the \textbf{other} spatial direction.

- Since each diffusion & convection term is only applied over half of the time stepping, we have to double the effective contribution of these terms when they are in fact applied. Correlation handled in the explicit steps.

- Boundary conditions handled similarly to 1D PDEs (no variance barriers).
PDE implementation

- In 2D (correlation neglected) with discretisation $U_{i,j}^k = U(x_i, y_j, \tau_k)$

$$U_{\tau} = A_{11} U_{xx} + A_{22} U_{yy} + b_1 U_x + b_2 U_y + fU$$

- Explicit in X, then implicit in Y
  $$U_{i,j}^{k+1} = U_{i,j}^k + \frac{A_{22} \Delta \tau}{2 \Delta y} \left[U_{i+1,j}^{k+1} - 2U_{i,j}^{k+1} + U_{i-1,j}^{k+2}\right] - \frac{b_2 \Delta \tau}{2 \Delta y} \left[U_{i+1,j}^{k+1} - U_{i,j}^{k+1}\right] - \frac{f}{2} \Delta \tau U_{i,j}^{k+1} = U_{i,j}^{k+1/2}$$

- $U_{i,j}^{k+1/2} = U_{i,j}^k + \frac{A_{11} \Delta \tau}{\Delta x^2} \left[U_{i+1,j}^k - 2U_{i,j}^k + U_{i-1,j}^k\right] + \frac{b_1 \Delta \tau}{2 \Delta x} \left[U_{i+1,j}^k - U_{i-1,j}^k\right] + \frac{f}{2} \Delta \tau U_{i,j}^k$

- ...then explicit in Y, and implicit in X
  $$U_{i,j}^{k+2} = U_{i,j}^{k+1} + \frac{A_{11} \Delta \tau}{\Delta x^2} \left[U_{i+1,j}^{k+1} - 2U_{i,j}^{k+2} + U_{i-1,j}^{k+2}\right] - \frac{b_1 \Delta \tau}{2 \Delta x} \left[U_{i+1,j}^{k+2} - U_{i-1,j}^{k+2}\right] - \frac{f}{2} \Delta \tau U_{i,j}^{k+2} = U_{i,j}^{k+3/2}$$

- $U_{i,j}^{k+3/2} = U_{i,j}^{k+1} + \frac{A_{22} \Delta \tau}{\Delta y^2} \left[U_{i,j+1}^{k+1} + 2U_{i,j}^{k+1} + U_{i,j-1}^{k+1}\right] + \frac{b_2 \Delta \tau}{2 \Delta y} \left[U_{i,j+1}^{k+1} - U_{i,j-1}^{k+1}\right] + \frac{f}{2} \Delta \tau U_{i,j}^{k+1}$
2D algorithm: best-of-call & BlackScholes

- Best-of 2 call option

ADI cycles between $x$ and $y$ directions

Explicit $x$  
Implicit $y$  
Implicit $x$  
Explicit $y$
2D algorithm: European call under Heston dynamics

- $x$ direction is log-spot
- $y$ direction is variance
- Solution has diffused out most where variance is large
The binary moustache generated by Heston model broadly exhibits correct qualitative features (priced using 2D ADI)

Model: $S_0=1$, $V_0=0.01$, $r_d=0$, $r_f=0$, strip of binaries with $T=1$

vovol $\alpha=5\%$
Summary

- Heston stochastic volatility model:
  - capable of generating realistic smiles and skews for vanillas
  - generates sensible deviations from B-S prices for binaries
  - able to admit very fast calibration scheme via semianalytic pricing
- Monte Carlo is easy to implement and provides useful “reality check” for other pricing algorithms
- When 2D finite difference engine such as ADI implemented, fast pricing of flow exotics in FX is quite straightforward
  - pricing requires solution of 2D convection-diffusion problem where diffusion is anisotropic in variance direction
  - compare with 2-factor Black-Scholes: isotropic in both log-spots