Ultra-Sparse Finite-Differencing For Arbitrage-Free Volatility Surfaces From Your Favourite Stochastic Volatility Model



25th of September 2014

	Sent 27 H Harpenta State State
2	Which stochastic volatility model?
3	The DHI volatility framework
4	Application examples
5	Spot Shock Sensitivities
6	Normal Extensions
7	The generator
8	Spatial discretisation
9	Continuous-time perfect martingale
10	Finite-differencing stencils
11	Boundary Conditions
12	Ensuring continuous-time stability
13	Numerical integation in time
14	Box transition probability translation
15	Cash dividends
16	Conclusion

(VTB Capital)

Ultra-Sparse Finite-Differencing

25th of September 2014

2 / 151

Implied volatilities have long been managed via *parametric formulations*.

Parametric volatility formulations are used for various purposes:-

- Convenient management of implied volatility surfaces and option books.
- Intuition about the shape of the volatility surface.
- Parametrization of the response of volatility to spot moves.
- Control of skew/smile-adjusted deltas of vanilla options to the typical behaviour of the market.



• Constant Elasticity of Variance (CEV) [CR76]

$$\mathrm{d}F = \sigma_{\mathsf{CEV}} \cdot F^{\beta} \cdot \mathrm{d}W \qquad \beta \leq 1$$

This gives only a skew, not a smile.

• Displaced Diffusion [Rub83]

$$dF = \sigma_{ extsf{DD}} \cdot [\beta F + (1 - \beta)F(0)] \cdot dW \qquad eta \leq 1$$

For $\beta > 0$, this gives implied volatility shapes very similar to CEV.

Note that Displaced Diffusion converges to the Bachelier model for $\beta \rightarrow 0$, but <u>CEV doesn't</u>.

• Mixture of log-normals

$$w(F,K,T) = w \cdot B(F,K,\sigma_1,T) + (1-w) \cdot B(F,K,\sigma_2,T)$$

This is a very old practitioner trick. If you cannot be sure about the volatility, take the average of (Black prices from) your two best guesses.

This gives only a smile with no skew at the money.

It can be extended to include a skew by allowing for two different forwards F_1 and F_2 subject to

$$w_1F_1 + w_2F_2 = F$$
.

Aka "Log-normal mixture dynamics" when mapped via Gyöngyi's theorem

$$\sigma^2_{_{ ext{effective local volatility}}}(K) = \mathsf{E} \left[\left. \sigma(F)^2 \right| F = K
ight]$$

to a continuous-time (Dupire-style) local volatility model.



Heston [Hes93]

$$\mathsf{d} F = \sqrt{
u} \cdot F \cdot \mathsf{d} W$$
, $\mathsf{d}
u = \kappa(heta -
u) \cdot \mathsf{d} t + \xi \sqrt{
u} \cdot \mathsf{d} Z$, $\mathsf{d} W \mathsf{d} Z =
ho \cdot \mathsf{d} t$

- Generates skews and smiles.
- Has a genuine second driver of risk, unlike local volatility models.
- The skew it can attain is often not enough in equities.
- When calibrated, the variance process typically has a significant positive probability (not just density!) of being at 0.
- Its forward volatility distribution is economically and financially doubtful.

• Numerically troubled: $\lim_{\nu \to 0} \frac{d}{d\nu} \sqrt{\nu} = \infty$ (infinite slope at 0) !

Dozens of articles on Monte Carlo or Finite Differencing for it.

Parametric volatility types

Stochastic volatility models

Schöbel-Zhu-Stein-Stein [SZ99, SS91]

 $\mathrm{d} F = \sigma \cdot F \cdot \mathrm{d} W, \quad \mathrm{d} \sigma = \kappa (\theta - \sigma) \cdot \mathrm{d} t + \xi \cdot \mathrm{d} Z, \quad \mathrm{d} W \mathrm{d} Z = \rho \cdot \mathrm{d} t$

- Normal volatility process with mean reversion.
- Vanilla options via characteristic functions, as in Heston, but without the analytical trap of the multi-valued logarithm.
- Numerically tractable.
- Good smiles and skews, but usually not enough for some equity markets.
- There is no lump of probability for volatility to be at zero, but there is positive density (and a lot of it).
- A good model, superior to Heston. Sadly much less commonly known.
- If only it had a local volatility component...

and x(z) is defined by

$$x(z) = \log\left\{\frac{\sqrt{1 - 2\rho z + z^2} + z - \rho}{1 - \rho}\right\}.$$
 (2.17c)

SABR is:-

• based on an asymptotic expansion,

Introduction

- a parametric formula for implied volatility,
- not a dynamic model,
- likely to give rise to arbitrage (negative densities, digitals > 1 or < 0),

Parametric volatility types

- fraught with danger for strikes near zero,
- prone to give infinite Libor-in-arrears prices,
- ... but it does make for nice smiles and skews, and is intuitive in its parameters.

For some time, the SABR formula has been considered the pinnacle of implied volatility parametrisation.





(VTB Capital)

Ultra-Sparse Finite-Differencing 25th of September 2014 11 / 151





```
(VTB Capital) Ultra-Sparse Finite-Differencing 25th of September 2014 13 / 151
```



Also, it was noticed that the dynamics of the underlying SDE are unsuitable for numerical evaluation:

- Moment explosions
- Singular local volatility slope at F = 0 (when $\beta < 1$). The same issue was the curse of the Heston model in numerical implementations.

However, and this is very important, SABR does give:-

- a dynamic response of the smile to spot movements which enables us to compute smile-adjusted deltas of vanillas, "*Managing Smile Risk*" (!),
- a sensible volatility surface for constant parameters,
- and thus allows for parametric interpolation.

(VTB Capital)	Ultra-Sparse Finite-Differencing	25th of September 2014	15 / 151
No. 17" MENERSE		6	
1 Introduction	Parametric volatility types	Stochastic volatility models	5
last but not least			

- J. Gatheral [Gat04] introduced the Stochastic Volatility Inspired (SVI) form.
- It can also create arbitrage.
- Its parameters are not as intuitive as SABRs.
- It is not quite as flexible as SABR.
- It gives no parametric generation of a volatility surface over time.
- It gives no parametric response of the volatility surface to spot moves.
- It is difficult to compute smile-adjusted deltas of vanillas.
- An alternative formulation for SVI with restricted parameters was published in [GJ13]. This, however, is so limited that many find it not useful.

Many people tried to remedy SABR.

- There are at least half a dozen of alternative asymptotic formulas similar to (2.17a), all of which can still have negative densities and exploding second moments (⇒ Libor in arrears).
- A range of researchers attempted to "correct the wings" of SABR. This does indeed remove the arbitrage. It does not, however, fix the infinite Libor-in-arrears case. Piecemeal and unsatisfactory, imho.
- Other authors started looking into finite-differencing approximations for the dynamic equations behind SABR.



Dropping the aim to have the "correct" solution of the dynamic equations...

- Andreasen and Huge use a finite differencing solver of a local volatility projection of the CEV/log-normal volatility process behind SABR [AH11].
- Lipton and Sepp [LS11] calibrate implied volatilities generated by a local volatility finite differencing solver to a given market smile.
- P. Hagan [Hag13] also came to the fore with a finite differencing approach to generating smiles from the CEV/log-normal stochastic volatility model he proposed more than ten years earlier.
- All of the above are for one-dimensional local volatility equations.
- None of the above make statements about interpolation and extrapolation after the finite-differencing stage.



Ok, so we are happy to use finite-differencing to create volatility smiles.

Based on our wishlist, we like the Hyperbolic-Hyperbolic model [JK07] :-

$$dx = \sigma_0 f(x)g(y)dW, \quad dy = -\kappa y dt + \alpha \sqrt{2\kappa} dZ, \quad dW dZ = \rho dt$$

$$x(t) = \frac{F_t(t)}{F_t(0)} \qquad \qquad y(0) = 0$$

with the two hyperbolic forms

$$f(x) = \frac{1}{\beta} \left[(1 - \beta + \beta^2) \cdot x + (\beta - 1) \cdot \left(\sqrt{x^2 + \beta^2 (1 - x)^2} - \beta \right) \right], \text{ and } g(y) = y + \sqrt{y^2 + 1}$$

(when $\beta < 0$ and x > 1, we actually use $f(x) = x^{\beta}$ which is well behaved under these circumstances).

The Hyp-Hyp model:-

- matches the CEV local volatility form based on the parameter β up to second order at the money via its hyperbolic local volatility f(x);
- does not have an infinite slope in its local volatility function in F = 0,
- and thus does not permit the underlying to become exactly 0;
- does not to converge to zero in its (relative) local volatility function for very high spot values,
- and thus does not imply zero implied volatility for very high strikes from its local volatility part alone;





(VTB Capital)

Ultra-Sparse Finite-Differencing

25th of September 2014 23 / 151

151



(VTB Capital)

Why would we have $\beta < 0$?

- Some markets, e.g., the S&P, have such a strong skew that it can only be approximated with $\beta < -3$ (or even lower).
- Econometric analyses have previously suggested negative β .
- E.g., in 1979, Macbeth and Merville [MM80] compute for $\theta \equiv 2\beta$:

<u>Stock</u>	Confidence Region	<u>Point Estimates</u>
ATT	$-2 \le \theta \le 6$	3.84
AVON	$-8 \le heta \le -2$	-3.63
ETKD	$-1 \leq heta \leq 5$	3.04
EXXN	$-1 \leq heta \leq 5$	1.62
IBM	$-8 \le \theta \le 2$	-4.16
XERX	$-4 \le heta \le 2$	-1.69

(VTB Capital)	Ultra-Sparse Finite-Differencing	25th of September 2014	25 / 151
17" H Harperster		6	
2 Which stochastic volatility model?	Hyp-Hyp	Negative β ?	

The Hyp-Hyp model:-

- matches the log-normal distribution of volatility of the SABR model up to second order near the most likely point of the distribution;
- does not have a log-normal tail for volatility at the upper end,
- and thus doesn't easily result in moment explosions;
- has economically realistic mean-reversion for volatility, typically not less than 50% ($\tau_{\rm memory} \approx 2$ Y) but readily as high as 1200% ($\tau_{\rm memory} \approx 1$ M);
- is benign with respect to aspects of numerical analysis.





(VTB Capital)

25th of September 2014 28 / 151





(VTB Capital)

25th of September 2014 30 / 151

to define

$$\operatorname{zabr}(z;\gamma) := \left[1 + (1 - \gamma) \cdot z\right]^{\frac{1}{(1 - \gamma)}} . \tag{4}$$

What does this look like for $\gamma=$ 80% in comparison to the SABR (exponential) case, and how does $g(\cdot)$ compare?



To compare this with the exponential form of SABR, and the hyperbolic form g(y) of Hyp-Hyp, we take, assuming for simplicity that $\tilde{\alpha}_0 = 1$,

Hyp-Hyp

P. Hagan mentioned that the response of volatility to its factor should, on a

The suggestion was to use the ZABR [AH11] form for the volatility process:

 $d\tilde{\alpha} = \tilde{\alpha}^{\gamma} \cdot dZ$

$$\int \frac{\mathrm{d}\tilde{\alpha}}{\mathrm{d}z} = \ln \tilde{\alpha} - \ln \tilde{\alpha}_0 \qquad \Rightarrow \quad \tilde{\alpha}(z) \sim \mathrm{e}^z \tag{2}$$

$$\int_{\tilde{\alpha}_0} \frac{\mathrm{d}\tilde{\alpha}}{\tilde{\alpha}} = \ln \tilde{\alpha} - \ln \tilde{\alpha}_0 \qquad \Rightarrow \quad \tilde{\alpha}(z) \sim \mathrm{e}^z \tag{2}$$

$$= \ln \tilde{\alpha} - \ln \tilde{\alpha}_0 \qquad \Rightarrow \quad \tilde{\alpha}(z) \sim e^z \qquad (2)$$

logarithmic scale, only be about 80%.

(1)

(3)



(VTB Capital)

Ultra-Sparse Finite-Differencing 25th of September 2014

4 33 / 151



The scope of the DHI volatility framework

is to:-

- take in parameters defined in terms of the HypHyp local-stochastic volatility process,
- output implied volatilities at a computational performance that is not significantly different from the use of an actual analytical formula,
- be *completely free of arbitrage*, without any exceptions,
- provide a wide range of smile and skew shapes,
- be amenable to specifying term structures of (piecewise constant) parameter values to create a complete *volatility surface* that is by construction free of arbitrage.

- σ₀: initial level of instantaneous volatility. Approximately equal to short term ATF implied volatility. LEVEL.
- *β*: local volatility skew coefficients. Same purpose and behaviour as in CEV or SABR. SKEW (via local volatility).
- α: uncertainty of volatility in the sense of *relative standard deviation*. This is *not* volatility of volatility¹. Recall that volatility scales by e^y and that α is the width of the stationary distribution of y.
 SMILE (via stochastic volatility).

¹Volatility of volatility is $\alpha\sqrt{2\kappa}$ but this quantity is often misleading in a mean-reverting context.

- *ρ*: correlation of the spot process driver diffusion and the volatilit process driver diffusion. Same purpose and behaviour as in SABR or Heston, etc. SKEW (via stochastic volatility).
- *k*: mean reversion of volatility. May need to be > 100% to straighten smiles or to accommodate long term volatility being not significantly higher than short term, as well as reduction of long term smile (which is something that SABR cannot match).

Larger κ makes y converge to the stationary distribution more rapidly.

Volatility memory time: $\left| au_{ ext{memory}} \sim rac{1}{\kappa}
ight|$

ĸ

STRAIGHTENING (both in the strike direction and in time).



We define the DHI framework

on a spatial discretisation for the process of the underlying, and of volatility.

To be crystal clear: our design is *not* an approximation to any process that is defined on a (spatially) continuous domain.

We define a process with stochastic volatility on a discrete set of spatial levels, both for the financial underlying, and for volatility.

This means that the chosen number of discrete levels is technically part of the set of parameters.

In practice, we use 25×11 nodes throughout.

DHI calculations are mathematically equivalent to a <u>continuous-time-finite-state Markov chain</u>.



3 The DHI volatility framework

Spatial discretisation

6

For fixed aspect ratios n_x/n_y , the discretisation dependence is small.





Implementation aspects

Implementation aspects

- Forward propagation to generate discrete Green's function for all T.
- A Markov-chain-consistent finite-differencing stencil for $\frac{d^2}{dzdv}$.
- Algebraic separation of the spatially discrete infinitesimal generator \tilde{L} into a horizontal, vertical, and diagonal component: $\tilde{L} = \tilde{L}_x + \tilde{L}_y + \tilde{L}_{xy}$.
- Absorbing boundary conditions at either end of each of the separate layers of \tilde{L}_x and \tilde{L}_{xy} in aid of making x a martingale.
- Reflecting boundary conditions for \tilde{L}_y .



- Analytical Martingale enforcement on all nodes of the full discrete generator \tilde{L} in the x-direction.
- Upwind evaluation of convection terms where necessary to remain Markovchain-consistent.
- A nested application of the second order Strang scheme [Str68] to split the continuous-time exponential propagator into five sequential stages.
- The second order fully implicit Padé(0,2) scheme for numerical integration in time.



A PARTY AND A PARTY AND A	and the second second	6
4 Application examples	Equity	RDXUSD

.RDXUSD 2013-09-30 market data





(VTB Capital) Ultra-Sparse Finite-Differencing 25th of September 2014 45 / 151

4 Application examples Pegged FX USDHKD



Market smiles



(VTB Capital) Ultra-Sparse Finite-Differencing 25th of September 2014 47 / 151

4 Application examplesPegged FXUSDHKDMarket smile + scenario"6M volatilities down by 1%"





Spot movement responses

- Since the dynamics are written on $x = \frac{S}{S_0}$, the initial value of x is always 1.
- ullet This applies even when the spot is shocked $S_0 o ilde S_0$.
- We can, however, derive a spot shock response by going back to S/S_0 and computing new parameters $\tilde{\sigma}'_0$ and $\tilde{\beta}$ that are (to first order) consistent with the model's dynamics.

The absolute volatility σ_{abs} of the spot $S \equiv F_t(t)$ is given by:

$$\sigma_{\mathsf{abs}}(S, F_t(0), \sigma_0, \beta) := \sigma_0 \cdot F_t(0) \cdot f(\frac{S}{F_t(0)}; \beta) \cdot g(y).$$
(5)

If the forward curve moves from $F_t(0)$ to $\tilde{F}_t(0)$, we look for $\tilde{\sigma}_0$ and $\tilde{\beta}$ such that:

$$\sigma_{\mathsf{abs}}(S, F_t(0), \sigma_0, \beta) \Big|_{S = \tilde{F}_t(0)} = \sigma_{\mathsf{abs}}(S, \tilde{F}_t(0), \tilde{\sigma}_0, \tilde{\beta}) \Big|_{S = \tilde{F}_t(0)}$$
(6)

$$\partial_{S}\sigma_{\mathsf{abs}}(S,F_{t}(0),\sigma_{0},\beta)\Big|_{S=\tilde{F}_{t}(0)} = \partial_{S}\sigma_{\mathsf{abs}}(S,\tilde{F}_{t}(0),\tilde{\sigma}_{0},\tilde{\beta})\Big|_{S=\tilde{F}_{t}(0)}$$
(7)

Denoting $\tilde{x} := \frac{\tilde{F}_t(0)}{F_t(0)}$, we obtain the analytical solution

$$\widetilde{\sigma}_0 = \sigma_0 \cdot f(\widetilde{x};\beta)/\widetilde{x}$$
(8)

$$\tilde{\beta} = \tilde{x} \cdot \partial_x \ln f(x;\beta) \Big|_{x=\tilde{x}}.$$
(9)



The flexibility of the DHI framework allows us to match virtually the same smile for different values of β .

Different choices of parameters give rise to different vega-adjusted Deltas

$$\Delta = \frac{\partial \mathsf{B}}{\partial S_0} + \frac{\partial \hat{\sigma}}{\partial S_0} \frac{\partial \mathsf{B}}{\partial \hat{\sigma}} \tag{10}$$

where B is the Black or Bachelier formula, and $\hat{\sigma}$ is implied volatility.

P. Hagan [Hag13] observed that, with a model-consistent response $\frac{\partial \hat{\sigma}}{\partial S_0}$, we are largely immunized against different parametrisations of the same smile.

The key point is that *pure spot movements* are fundamentally inconsistent with the model, especially when $|\rho| \approx 1$.

Given a shock for S, we therefore compute a model-consistent response of y, move our initial Dirac mass of probability to that location in the (x, y) plane, and build the shocked DHI volatility surface from that.

We call this response of the volatility surface *Dynamic Delta*².







(VTB Capital) Ultra-Sparse Finite-Differencing 25th of September 2014 55 / 151

5 Spot Shock Sensitivities Vanilla Δs for different parametrisations





(VTB Capital) Ultra-Sparse Finite-Differencing 25th of September 2014 57 / 151

5 Spot Shock Sensitivities Vanilla Δs for different parametrisations

Vanilla call Δ





(VTB Capital) Ultra-Sparse Finite-Differencing 25th of September 2014 59 / 151



Vanilla call Δ



Deltas from four different response types.

For hedging purposes, we'd like to have the model mimic the response of market volatilities to spot movements. Take RDXUSD on 2014-01-07.



Undoubtedly, the 3M smile is an excellent fit. But do we like the parametric spot-vol response, compared to the recent trend of the market?



Surely, though, given a choice, we'd prefer a better match of model and market spot-vol response.

The local-stochastic volatility framework enables us to match the current options smile **and** the dynamic spot-volatility response of the market!



Caution is required as to the choice of the target spot-vol slope. This, however, now becomes a trading choice, instead of just model output!

Ultra-Sparse Finite-Differencing (VTB Capital) 25th of September 2014 63 / 151 5 Spot Shock Sensitivities Calibration to spot-vol slope

Alas, when we use the Dynamic Delta response logic, we give up the control over the spot-vol relationship:



The (95%-105%) skew is pprox 2.6%, the response skew is pprox 5.2%. Ratio: pprox 2.

The Dynamic Delta response is *the local volatility projection* !



6 Normal Extensions

When S can go negative, we can define

$$x(t) := F_t(t) - F_t(0)$$
 (11)

as the model process. In addition, we need a different local volatility function f.

We could use

$$f_{normal}(x;\beta) = 1 + \frac{|\beta|}{2} \left(\sqrt{1+x^2} + sign(\beta) \cdot x - 1 \right)$$
 (12)

which satisfies f(0) = 1.

This function has the reflection symmetry

$$f_{\text{normal}}(-x;-\beta) = f_{\text{normal}}(x;\beta) .$$
(13)



Normal model absolute local volatility function $f_{normal}(x;\beta)$ given in (12).

• The system of stochastic differential equations

$$dx = \sigma_0 f(x) g(y) dW$$
(14)

$$dy = -\kappa y dt + \alpha \sqrt{2\kappa} dZ$$
 (15)

• gives us the spatially continuous instantaneous generator

$$L = -\kappa y \frac{\partial}{\partial y} + \frac{1}{2} \sigma_0^2 f(x)^2 g(y)^2 \frac{\partial^2}{\partial x^2}$$
(16)
+ $\rho \sigma_0 \alpha \sqrt{2\kappa} f(x) g(y) \frac{\partial^2}{\partial x \partial y} + \kappa \alpha^2 \frac{\partial^2}{\partial y^2}$

Pricing derivatives *backward Kolmogorov* equation:

The generator

$$\frac{\partial}{\partial T}v = -L \cdot v . \tag{17}$$

Transition probabilities from t = 0 to T forward Kolmogorov or Fokker-Planck equation:

$$\frac{\partial}{\partial t} p = L^* \cdot p . \tag{18}$$

 \implies L* is the adjoint of the operator L.

Starting from a Dirac spike at t = 0, this gives us

Green's functions.



Transform to the logarithmic coordinate

$$z := \ln(x) . \tag{19}$$

The generator L becomes

$$L = -\frac{1}{2}\varsigma(z,y)^2 \frac{\partial}{\partial z} - \kappa y \frac{\partial}{\partial y}$$
(20)

$$+ \frac{1}{2}\varsigma(z,y)^2 \frac{\partial^2}{\partial z^2} + \rho\varsigma(z,y)\alpha\sqrt{2\kappa}\frac{\partial^2}{\partial z\partial y} + \kappa\alpha^2\frac{\partial^2}{\partial y^2}$$

with the time-homogenous separable local (in z and y) volatility

$$\varsigma(z, y) := \sigma_0 e^{-z} f(e^z) g(y) .$$
(21)

We form a lattice of (typically) $n_z = 25 \times n_y = 11$ spatial nodes.

The dynamics are confined to these spatial levels.

The operator L is replaced by a spatially discretised generator \tilde{L} by the aid of finite differencing stencils.

The probability function p(t, z, y) becomes a vector-valued function of time alone

$$\tilde{p}(t) \in \mathbb{R}^{n_z \times n_y}$$
.

 $\implies \tilde{p}(t)$ is still <u>continuous</u> in time.



The (forward) dynamics are now governed by the

system of ordinary differential equations

$$\frac{\partial}{\partial t}\tilde{p} = \tilde{L}^* \cdot \tilde{p} \tag{22}$$

with solution

$$\tilde{p}(t+\tau) = e^{\tau \cdot \tilde{L}^*} \cdot \tilde{p}(t)$$
 (23)

The dynamics are now *exactly* in the form of a

continuous-time finite-state Markov chain.
8 Spatial discretisation Continuous-time finite-state Markov The Metzler property

• Since \tilde{L} is a real-valued matrix,

$$\tilde{L}^* = \tilde{L}^\top$$
 .

• Choose boundary conditions for the generator \tilde{L} , and \tilde{L}^* follows trivially!

(This is a common problem when implementing the forward-Kolmogorov system: what boundary conditions to use in a forward scheme? As we show here, it shouldn't be an issue at all.)

- From $e^{\tau \cdot \tilde{L}^*} \cdot \tilde{p}$, we see that \tilde{L} must not have any eigenvalues with positive real parts: sup $\left\{ \Re(\lambda) : \lambda \in \Sigma(\tilde{L}) \right\} \leq 0$ (24)
- Since the rows of *L̃* always sum up to zero (Markov chain property), for (24) to hold, via the Gershgorin circle theorem,

it is sufficient that all off-diagonal elements of \tilde{L} are ≥ 0 .

This is known as the *Metzler* property.

 \Rightarrow The spatial discretisation and boundary conditions must result in \tilde{L} being a Metzler matrix.

8 Spatial discretisation Continuous-time finite-state Markov The Metzler property

- The generator \tilde{L} is also know under the names *transition rate matrix* or *intensity matrix*.
- The dynamics are a *pure jump process* with nearest-neighbour-transitions.
- We can compute the (residual) drift in the x-direction from Itô's lemma for pure jump processes.
- We can use this knowledge to make \tilde{L} an analytically exact x-martingale.
- This COMPLETELY removes any residual drift error as is commonly observed in finite-differencing schemes.

This residual drift is the dominant reason why most finite-differencing implementations need significant numbers (at least hundreds, but even thousands) of discretisation levels in the spot (x) direction.

 We can have *L̃* as a perfect x-martingale with any number of nodes, be it 100, 10, or even just 3!

Example: spatially discretised local volatility generator.

Consider the transport out of the node at x_0 to its nearest neighbours:

$$x_{-1} = e^{-\Delta z + z_0}$$
 $x_0 = e^{z_0}$ $x_1 = e^{\Delta z + z_0}$ (25)

In spot-logarithmic coordinates:

$$dz = -\frac{\sigma^2(z)}{2} dt + \sigma(z) dW$$
(26)

The spatially discrete generator with standard centre differencing for both diffusion and advection:

$$\zeta(z) = \frac{\sigma^2(z)}{2\Delta z^2} \; .$$

(VTB Capital) Ultra-Sparse Finite-Differencing 25th of September 2014 75 / 151

9 Continuous-time perfect martingale Local volatility example

This makes the *actual* drift in x out of the node at $x_0 = e^{z_0}$:

$$\mu(z_0) := \frac{\mathsf{E}[\mathsf{d}_x|_{x=x_0}]}{x_0 \mathsf{d}_t}$$
(28)

$$= \underbrace{\frac{\sigma^2(z_0)}{2\Delta z^2} \left[(e^{-\Delta z} - 1)(1 + \frac{\Delta z}{2}) + (e^{\Delta z} - 1)(1 - \frac{\Delta z}{2}) \right]}_{z = 0}$$

This is Itô's lemma for our pure jump process to neighbouring nodes. (29)

$$= -\frac{\sigma^2(z_0)}{24}\Delta z^2 + \mathcal{O}(\Delta z^4)$$
(30)

$$\neq \mathbf{0}$$
 . (31)

However, using $\mu(z_0)$ as given by (29), and changing the advection coefficient in \tilde{L} to

$$\tilde{L}' = \begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \zeta(z_0)(1 + \frac{\Delta z}{2} \cdot [1 + \varepsilon]) & -2\zeta(z_0) & \zeta(z_0)(1 - \frac{\Delta z}{2} \cdot [1 + \varepsilon]) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$
(32)

with

$$\boldsymbol{\varepsilon} = \frac{\mu(z_0)}{\Delta z \cdot \sinh(\Delta z)} \tag{33}$$

makes \tilde{L}' an *exact* x-martingale out of each and every level z_0 .

Mutatis mutandis, the same can be done for our original (z, y) process to make its generator \tilde{L} an exact continuous-time x-martingale in each node.

To form \tilde{L} from L, we use

• for diffusive terms, centre-differencing:

$$\frac{\partial^2}{\partial z^2} p(\cdot, z, y_j) \bigg|_{z=z_i} \implies \frac{1}{\Delta z^2} \Big(\tilde{p}_{i-1,j} - 2\tilde{p}_{i,j} + \tilde{p}_{i+1,j} \Big)$$
(34)

$$\frac{\partial^2}{\partial y^2} p(\cdot, z_i, y) \bigg|_{y=y_j} \implies \frac{1}{\Delta y^2} \Big(\tilde{p}_{i,j-1} - 2\tilde{p}_{i,j} + \tilde{p}_{i,j+1} \Big)$$
(35)

• for advective terms, *preferably*, centre-differencing

$$\frac{\partial}{\partial z} p(\cdot, z, y_j) \Big|_{z=z_i} \implies \frac{1}{2\Delta z} \Big(\tilde{p}_{i+1,j} - \tilde{p}_{i-1,j} \Big)$$
(36)

$$\left. \frac{\partial}{\partial y} p(\cdot, z_i, y) \right|_{y=y_j} \implies \frac{1}{2\Delta y} \left(\tilde{p}_{i,j+1} - \tilde{p}_{i,j-1} \right)$$
(37)

though we switch to the (partial) upwind, aka forward-differencing, stencil when the above would violate the Metzler property of \tilde{L} .

Regarding the mixed diffusion term, conventional implementations tend to use the four-point stencil

$$\frac{\partial^2}{\partial z \partial y} \tilde{p}_{i,j} \approx \frac{1}{4\Delta z \Delta y} \Big(\tilde{p}_{i+1,j+1} + \tilde{p}_{i-1,j-1} - \tilde{p}_{i+1,j-1} - \tilde{p}_{i-1,j+1} \Big).$$
(38)

An alternative is to use another second order discretisation for mixed diffusion terms, the little known *seven-point stencil*

$$\frac{\partial^{2}}{\partial z \partial y} \tilde{p}_{i,j} \approx \frac{1}{2\Delta z \Delta y} \cdot \begin{cases} \tilde{p}_{i+1,j+1} + 2\tilde{p}_{i,j} + \tilde{p}_{i-1,j-1} \\ -\tilde{p}_{i+1,j} - \tilde{p}_{i-1,j} - \tilde{p}_{i,j+1} - \tilde{p}_{i,j-1} \\ 0 \\ -\tilde{p}_{i-1,j+1} - 2\tilde{p}_{i,j} - \tilde{p}_{i+1,j-1} \\ + \tilde{p}_{i+1,j} + \tilde{p}_{i-1,j} + \tilde{p}_{i,j-1} + \tilde{p}_{i,j+1} \\ 0 \\ 0 \end{cases}$$
(39)



The four-point (38) and the seven-point mixed diffusion stencil (39).

Cross-diffusion

Example:

2D standard homogenous diffusion on 3×3 lattice with diffusion only on centre node.

The four-point stencil only permits the Metzler property when $\rho = 0$, which defeats the point of having a mixed-diffusion stencil in the first place.

(VTB Capital)	Ultra-Sparse Finite-Differencing	25th of September 2014	81 / 151
		6	
10 Finite-differencing stencils	Cross-diffusion	The Seven Point Stencil	

- Worse even: in the limit of $|\rho| \rightarrow 1$, the generator \tilde{L}_{4P} doesn't preserve the stratification invariance of the two-dimensional heat equation.
- In that limit, the diffusion becomes one-dimensional along the diagonal (or anti-diagonal), and no flow should occur between different diagonal (or anti-diagonal) layers.
- Ultimately, for |
 ho|
 ightarrow 1, the four-point stencil becomes *inconsistent*.
- In addition, the four-point stencil, not having the Metzler property, makes it impossible to *preserve positivity* (as we will show later).

Note: the only non-zero eigenvalue of \tilde{L}_{4P} in (40) is actually -2, and thus there is no "instability" at stake here.

However, in addition to stability, for the later purpose of numerical integration in time to preserve non-negative probabilities, ideally,

we would like at least the fully implicit Padé(0,1) scheme propagator

$$A^{(0,1)}(\tilde{L}) = \left[1 - \tau \cdot \tilde{L}^*\right]^{-1}$$
(41)

over a time step of length au > 0 to have no negative entries.

(VTB Capital)	U	ltra-Sparse	Finite-D	ifferen	cing	25th o	83 / 151		
271 M-166	and the second	i l'ante				(5		
10 Finite-differencing	stencils T	ne importan	ce of be	ing Me	etzler				
We can compute	$A^{(0,1)}$ and	alyticall	y for	a g	eneric	Ĩ w	hose	sole no	n-zero
entries are in its centre row:									
${\mathcal{A}^{(0,1)}}^{ op}(ilde{L})\ =\ c\ \cdot$	$ \begin{pmatrix} \frac{1}{c} & 0 \\ 0 & \frac{1}{c} \\ 0 & 0 \\ 0 & 0 \\ \tilde{\ell}_{5,1} & \tilde{\ell}_{5,2} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} $	$ \begin{array}{c} 0\\ 0\\ \frac{1}{c}\\ 0\\ \tilde{\ell}_{5,3}\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0\\$	$0 \\ 0 \\ \frac{1}{c} \\ \tilde{\ell}_{5,4} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ $	$ \begin{array}{c} 0 \\ 0 \\ 0 \\ \frac{1}{\tau} \\ 0 $	$\begin{array}{c}0\\0\\0\\\tilde{\ell}_{5,6}\\\frac{1}{c}\\0\\0\end{array}$	$ \begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ \tilde{\ell}_{5,7} \\ 0 \\ \frac{1}{c} \\ 0 \\ 0 \\ 0 \end{array} $	$\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ \tilde{\ell}_{5,8} \\ 0 \\ 0 \\ \frac{1}{c} \\ 0 \end{array}$	$ \begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ \tilde{\ell}_{5,9} \\ 0 \\ 0 \\ 0 \\ 1 \end{array} $	(42)
		0	0	0	U	0	0	$\frac{1}{c}$ /	
with $c:=rac{ au}{1- au\cdot\ell_{{\tt 5},{\tt 5}}}$	$> 0. \Rightarrow V$	Ve need							
	$ ilde{\ell}_{5,5} \leq 0$	and	$\widetilde{\ell}_5$,i ≥	0 ∀	i ≠	5		
which makes $ ilde{\mathcal{L}}$ Metzler. This generalizes:									
if $ ilde{L}$ is Metzler, $A^{(0,1)}(ilde{L})_{i,j} \geq 0 orall \ \ i,j.$									

```
(VTB Capital)
```

In continuous time, we can invoke the generic theorem that

the exponential of a Metzler matrix is a non-negative matrix. So, \tilde{L} being Metzler is sufficient for $e^{\tau \cdot \tilde{L}}$ to be non-negative. However, it is straightforward to see from the Taylor expansion

$$e^{\tau \cdot \tilde{L}} = \mathbf{1} + \tau \cdot \tilde{L} + \frac{1}{2}\tau^2 \cdot \tilde{L}^2 + \frac{1}{6}\tau^3 \cdot \tilde{L}^3 + \cdots$$
(43)

that for any non-Metzler \tilde{L} , there is some value for $\tau > 0$ for which

(any truncated Taylor expansion of) $e^{\tau \cdot \tilde{L}}$

has negative off-diagonal elements!

The Metzler property for \tilde{L} is both sufficient and necessary !

Formally, we have $(e^{\tau \cdot \tilde{L}})_{ij} \ge 0$ $(\forall i, j)$ for all $\tau \ge 0$ if and only if \tilde{L} is Metzler [Min88].

10 Finite-differencing stencils The importance of being Metzler Metzler is sufficient and necessary

To summarize, we want \tilde{L}^* to be Metzler because:-

- Ideally, we want whichever numerical scheme we use later for integration in time to preserve positivity of probabilities.
- The only (simple) scheme that (we know of that) can preserve positivity of *p* from t to t + τ (for all τ > 0) is the first order fully implicit scheme

$$(1 - \tau \cdot \tilde{L}^*) \cdot \boldsymbol{\rho}(t + \tau) = \boldsymbol{\rho}(t)$$
(44)

• For (44) to preserve positivity (in general, for all $\tau > 0$),

\tilde{L}^* must be Metzler.

- Note that we may actually soften up our aim to have strictly positive transitions later on, and not actually use (44)
- but we still aim to preserve the Metzler property with other schemes, too,
- since we find (heuristically, but definitely not suprisingly) that non-positivity is much less of an issue when \tilde{L}^* is Metzler.
- An example is the first order fully explicit scheme which preserves positivity if and only if \tilde{L}^* is Metzler (and τ is small enough).

In contrast to the four-point stencil, consider the net generator with the seven-point stencil for the same example (with positive correlation):

The seven-point stencil preserves the Metzler property for all $|
ho| \leq 1$!



The good news is, the seven-point stencil

separates into three uni-directional diffusion components.

For
$$\rho \ge 0$$
:

$$\frac{\partial^2}{\partial z \partial y} \tilde{p}_{i,j} \approx \frac{1}{\Delta z \Delta y} \left[\underbrace{p_{i+1,j+1} - 2p_{i,j} + p_{i-1,j-1}}_{\text{zy-component}} - \underbrace{(p_{i+1,j} - 2p_{i,j} + p_{i-1,j})}_{\text{z-component}} - \underbrace{(p_{i,j+1} - 2p_{i,j} + p_{i,j-1})}_{\text{y-component}} \right]$$
By combining the

- z-component of the mixed-diffusion (seven-point) stencil with the z-diffusion generator terms,
- and the y-component of the mixed-diffusion (seven-point) stencil with the y-diffusion generator terms,

we can write the spatially discrete generator as

$$\tilde{L} = \tilde{L}_{zy} + \tilde{L}_{z} + \tilde{L}_{y} .$$
(47)

25th of September 2014

```
(VTB Capital)
```

 10 Finite-differencing stencils
 Cross-diffusion
 The Seven Point Stencil

Ultra-Sparse Finite-Differencing

Notes:

- The net operators \tilde{L}_z and \tilde{L}_y can have, at this stage, negative local diffusivity.
- This will be dealt with later with *anti-diffusive flux limiting*.
- Equation (47) is in spirit akin to a *differential* operator split.
- In practice, we employ a full algebraic split of \tilde{L} .
- In essence, this means that we really only split the matrix \tilde{L} into three parts after the boundary conditions have been fully taken into account.
- This enables us to minimize the amount of anti-diffusive flux limiting we have to impose as part of the definition of our spatially discretised model.³

89 / 151

³For more details on the difference between differential and algebraic operatorsplitting, *Algebraic Flux Corrections* (AFC), and *Flux Corrected Transport* (FCT) see, e.g., [Kuz07, Kuz10, BLOG93].

Common "wisdom" has it that:-

- boundary conditions matter little,
- the boundary always has to be moved out a long way,
- one always needs a lot of nodes,
- "if the boundary conditions seem to affect your result, the boundary needs to be moved out further, and you need to use more nodes".



(VTB Capital)

Ultra-Sparse Finite-Differencing

25th of September 2014

91 / 151

11 Boundary Conditions

- There is a large volume of literature in engineering and Computational Fluid Dynamics dedicated to boundary conditions. *They <u>do</u> matter.*
- Our design here is to be *ultra-sparse*, e.g., typically, 25×11 nodes.
- With very few nodes, the boundaries (by mere node-count-proximity) always have a significant influence on the central region.
- Real physical spatially discrete (model) systems with few nodes behave perfectly reasonably, so why shouldn't our model equations? (recall that our equations are discrete in space but *continuous in time*, and thus can still be represented by an actual physical experiment)
- On sparse lattices, the boundary nodes make up a significant percentage:

Discretisation	Total nodes	Boundary nodes	Percentage
15 imes 5	75	36	$48\% \approx 1/_2$
21 imes 7	147	52	$35.4\% pprox 1/_3$
25 imes 11	275	68	24.7% ≈ ¹/₄

• We should make full use of the boundary nodes for the sake of efficiency!

In support of the intended *algebraic* split

$$\tilde{L} = \tilde{L}_{z} + \tilde{L}_{y} + \tilde{L}_{zy}$$
(47)

we choose boundary conditions directly for \tilde{L}_z , \tilde{L}_y , and \tilde{L}_{zy} according to:-

- The generator must be an *x*-martingale.
- This means that x- (and thus z-) direction boundary nodes must have no (outoging) transport in the x- (and thus z-) direction. Absorption.
- By the same token, zy-direction boundary nodes must have no (outoging) transport in the zy-direction.
 Absorption.
- In contrast, y-direction boundary nodes have no y-martingale requirements: transport is permissible in and out in the y-direction. *Reflection*.





(VTB Capital)

Ultra-Sparse Finite-Differencing 25th of September 2014 95 / 151





			6	
11 Boundary Conditions Permissible transport An extended alternative	11 Boundary Conditions	Permissible transport	An extended alternative	

An optional extended alternative version (not used in the following):



11 Boundary Conditions

Absortion and Reflection

In double-indexed notation, the element $\tilde{L}(i, j, k, l)$ with $(i, j) \neq (k, l)$ is the *instantaneous transition rate of transfer* from the node at (i, j) to the node at (k, l) in our spatially discrete stochastic process, and

$$\tilde{L}(i,j,i,j) = -\sum_{k\neq i,j\neq l} \tilde{L}(i,j,k,l) .$$

In this notation, we have the absorbing boundary conditions for \tilde{L}_z :

$$\tilde{L}_{z}(1,\cdot,\cdot,\cdot) = \tilde{L}_{z}(n_{z},\cdot,\cdot,\cdot) = 0$$
(48)

For \tilde{L}_{zy} , we have the absorbing conditions:

$$\tilde{L}_{zy}(1,\cdot,\cdot,\cdot) = \tilde{L}_{zy}(n_z,\cdot,\cdot,\cdot) = \tilde{L}_{zy}(\cdot,1,\cdot,\cdot) = \tilde{L}_{zy}(\cdot,n_y,\cdot,\cdot) = 0.$$
(49)

The reflecting boundary conditions for L_y are:

$$\widetilde{L}_{y}(i,1,i,1) = -D_{y}(z_{i},y_{1}) \qquad \widetilde{L}_{y}(i,n_{y},i,n_{y}) = -D_{y}(z_{i},y_{n_{y}})
\widetilde{L}_{y}(i,1,i,2) = D_{y}(z_{i},y_{1}) \qquad \widetilde{L}_{y}(i,n_{y},i,n_{y}-1) = D_{y}(z_{i},y_{n_{y}})
(50)$$

The local diffusivity coefficient $D_y(z, y)$ will be explained later.

(VTB Capital)	Ultra-Sparse Finite-Differencing	25th of September 2014	99 / 151

11 Boundary Conditions Deriving the reflecting conditions The method of images

To derive the reflecting conditions, we can use *method of images*.

- For a left-most boundary node value v_0 at location y_0 , pose an imaginary node \breve{v}_{-1} to the left of it, at y_{-1} .
- This imaginary node at y_{-1} mirrors what is to the right of y_0 .
- Hence, it has the same value as the second (i.e., the first interior) node v₁, and y₁ y₀ = y₀ y₋₁.



100

This gives us for the first derivative with standard centre-differencing

$$\frac{\partial v}{\partial y}\Big|_{y=y_0} \approx \frac{v_1 - \breve{v}_{-1}}{2\Delta y}$$

$$= \frac{v_1 - v_1}{2\Delta y}$$

$$= 0$$
(51)

 \implies The reflecting boundary condition means NO ADVECTION!



For the diffusion term, we obtain from the method of images that the total local diffusion flux is proportional to

$$\frac{\partial^2 v}{\partial y^2}\Big|_{y=y_0} \approx \frac{v_1 - 2v_0 + \breve{v}_{-1}}{\Delta y^2}$$

$$= \frac{v_1 - 2v_0 + v_1}{\Delta y^2}$$

$$= \frac{2 \cdot \frac{v_1 - v_0}{\Delta y^2}}{\Delta y^2} .$$
(52)

This, however, includes the flux to and from the imaginary node!

By symmetry, only half of this is *actual* flux to and from the interior node, and hence no factor 2 in (50).

⇒The reflecting boundary condition makes a diffusion term look like a one-sided, i.e., forward-differencing, advection term.

11 Boundary Conditions Deriving the reflecting conditions

From continuous equations

We note that conventional (spatially continuous) notation for a reflecting boundary condition is

$$\frac{\partial v}{\partial y} = 0.$$
 (53)

In a spatially discretised form, this forces the advection term to be 0.

The diffusion term can be realized by *extending* the domain by an additional (non-imaginary) node at y_{-1} (though this can in principle be asymmetric, i.e., closer to y_0):



11 Boundary Conditions Deriving the reflecting conditions From continuous equations

Using forward differencing for advection terms in this extra node, the condition

$$0 = \left. \frac{\partial v}{\partial y} \right|_{y=y_{-1}} \approx \frac{v_0 - v_{-1}}{\Delta y}$$
(54)

enforces the identity

$$v_{-1} \equiv v_0 . \tag{55}$$

This gives us for the diffusion term in y_0

$$\frac{\partial^2 v}{\partial y^2}\Big|_{y=y_0} \approx \frac{v_1 - 2v_0 + v_{-1}}{\Delta y^2} = \frac{v_1 - 2v_0 + v_0}{\Delta y^2} = \frac{v_1 - v_0}{\Delta y^2}$$
(56)

which is exactly what we had from the spatially discretised method of images.

This derivation here is based on the spatially continuous reflection condition (53) results in the same dynamics on the interior *and* the boundary node in y_0 . The only difference is the extra auxiliary node at y_{-1} which is of no dynamic value or modelling contribution. It is literally a waste of space. The three components of $\tilde{L} = \tilde{L}_z + \tilde{L}_y + \tilde{L}_{zy}$ are:-

• \tilde{L}_z : advection-diffusion generator with *local diffusivity*

$$D_{z}(z,y) = \frac{1}{2} \left(\frac{\varsigma(z,y)^{2}}{\Delta z^{2}} - \frac{|\rho|\varsigma(z,y)\alpha\sqrt{2\kappa}}{\Delta z\Delta y} \right)$$
(57)

• \tilde{L}_y : advection-diffusion generator with local diffusivity

$$D_{y}(z,y) = \frac{1}{2} \left(\frac{2\kappa\alpha^{2}}{\Delta y^{2}} - \frac{|\rho|\varsigma(z,y)\alpha\sqrt{2\kappa}}{\Delta z\Delta y} \right)$$
(58)

• \tilde{L}_{zy} : pure diffusion generator with local diffusivity

$$D_{zy}(z,y) = \frac{1}{2} \frac{|\rho|\varsigma(z,y)\alpha\sqrt{2\kappa}}{\Delta z \Delta y}$$
(59)

with

$$\varsigma(z, y) := \sigma_0 e^{-z} f(e^z) g(y) .$$
(21)

(VTB Capital) Ultra-Sparse Finite-Differencing 25th of September 2014 105 / 151

12 Ensuring continuous-time stability Local diffusivity in continuous time

Why the special attention on the diffusion part?

• A pure advection term can give rise to oscillatory modes, but

not to growing modes.

- The sign of the advection coefficient can be either way it has no impact on mode stability.
- A *physical* diffusion term has eigenvalues with non-positive real parts. Real world diffusion does not concentrate distributions, it *diffuses* them. The clue is in the name.
- A diffusion term is physical if its local diffusivity is non-negative.
- A diffusion term with negative local diffusivity causes exponentially growing modes. It violates the *Maximum Principle*.
- A growing (localized) mode can be caused by any local diffusivity of any of the three of \tilde{L}_z , \tilde{L}_y , and \tilde{L}_{zy} to be negative anywhere.

In order to ensure continuous-time stability, we demand that all local diffusivities are non-negative everywhere.

12 Ensuring continuous-time stability Local diffusivity in continuous time

For any given value of $|\rho|$, either D_z or D_y can be negative, but not both.

To see this, define

$$R(z,y) := \frac{\varsigma(z,y)}{\alpha\sqrt{2\kappa}} \frac{\Delta y}{\Delta z} , \qquad (60)$$

then

$$D_{z}(z,y) = \frac{1}{2} \frac{\varsigma(z,y) \alpha \sqrt{2\kappa}}{\Delta z \Delta y} \Big(R(z,y) - |\rho| \Big)$$
(61)

$$D_{y}(z,y) = \frac{1}{2} \frac{\varsigma(z,y)\alpha\sqrt{2\kappa}}{\Delta z \Delta y} \left(\frac{1}{R(z,y)} - |\rho|\right).$$
 (62)

Since $R(z, y) \ge 0$, both of $D_z(z, y)$ and $D_y(z, y)$ are non-negative when

$$|\rho| \leq \min\left(R(z,y), \frac{1}{R(z,y)}\right)$$
 (63)

(VTB Capital)	Ultra-Sparse Finite-Differencing	25th of September 2014	107 / 151
A PARTY AND A PART		6	
12 Ensuring continuous-time stability	Local diffusivity in continuous time	Anti-diffusive flux limiting	

However, when

$$\rho| > |\rho|_{\max} \tag{64}$$

with

$$|\rho|_{\max} := \min_{z,y} \left(R(z,y), \frac{1}{R(z,y)} \right) , \qquad (65)$$

then some of the local diffusivities $D_z(z, y)$ and $D_y(z, y)$ on the lattice are negative, which is not admissible.

When this happens, we floor the respective local diffusivity at zero.

This is called Anti-Diffusive Flux Limiting.

Note that only one of $D_z(z, y)$ and $D_y(z, y)$ can be negative at any one node.

We emphasize that this flooring of coefficients does not constitute an inconsistency with our model choice because *our model choice is the net result of discretisation*, after all adjustments, in the form of the final (spatially discrete) generator \tilde{L} .

Apart from not preserving positivity, a violation of the Metzler property can have another, <i>much much worse</i> , consequence For instance, take these parameters in a province of the second	12 Ensuring continuous-time stability Local diffusivit	v in conti	nuous	time	How	S bad cou	uld it no	ssibly b	e?
For instance, take these parameters which lead to this generator \tilde{L} : T σ β κ α ρ n x n_{ν} 1 1 1 1 1 1 1 0.07 7 5 T σ β κ α ρ n_{ν} n_{ν} T σ γ σ	Apart from not preserving positivity, have another, <i>much much worse</i> , co	a viol nseque	atio ence	n of 	the N	letzl	er pro	opert	y can
which lead to this generator \tilde{L} : 1	For instance, take these parameters	Т	σ	β	К	α	ρ	n_x	n_y
(VTB Capital) Utra-Sparse Finite-Differencing 2th of September 2014 109 / 151 12 Ensuring continuous-time stability Local diffusivity in continuous time Yew bad could it possibly be? This (continuous-time stability Local diffusivity in continuous time Yew bad could it possibly be? 14 Finite and the state and	which lead to this generator $\widetilde{\mathcal{L}}$:	1	1	1	1	1	-0.7	7	5
(VTB Capital) Ultra-Sparse Finite-Differencing 25th of September 2014 109 / 151 12 Ensuring continuous-time stability Local diffusivity in continuous time How bad could it possibly be? This (continuous-time) generator has this spectrum:	$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	$\begin{array}{cccccc} 16 & 17 & 18 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 &$	$\begin{array}{c cccc} 19 & 20 \\ \hline 0 & 0 \\ 1 & 1.366 \\ 0 & 0 \\ 1 & 1.366 \\ 0 & 0 \\ 1 & 1.366 \\ 0 & 0 \\ 0 $	21 22 0 0	23 24 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.566 0.0 0 0 0 0 0.566 0.0 0 0 0 0.202 0 0 0 2.895 -6.83 0.202 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$\begin{array}{cccccccccccccccccccccccccccccccccccc$
This (continuous-time) generator has this spectrum:	12 Ensuring continuous-time stability Local diffusivit	y in conti	nuous	time	How	S bad cou	ıld it po	ssibly b	e?
Image: Non-Section of the section o	I his (continuous-time) generator ha	is this	spec	ctrun	า:				-1
							×		0.0
Image: Second									0.8
							X		0.6
a X 0.2 a X X 0.2 x X X X 0.2 y -25 -20 -15 -10 -5 0 y X 0.4 X 0.4 X 0.4 x 0.4 X 0.6 0.6 0.6 All real parts are non-positive. All is well. 0.8 0.8 0.1	m								0.4
x x <td>a g</td> <td></td> <td></td> <td></td> <td>×</td> <td></td> <td></td> <td></td> <td>0.2</td>	a g				×				0.2
a X X X X X X X X X X X X X X X X X X X 0.2 0.2 0.4 X 0.4 X 0.4 X 0.4 X 0.6 X 0.6 X 0.6 X 0.6 X 0.8 X 0.8 X 0.8 X 0.8 X 0.8 X 0.8 X 0.1 X 0.8 X 0.1 X 0.1 X 0.8 X 0.1 X 0.8 X 0.1 X 0.8 X 0.1 X 0.8 X 0.1 X 0.1 <t< td=""><td>n han han han han han han han han han ha</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></t<>	n han han han han han han han han han ha								
y 10 10 -0.2 P a r t All real parts are non-positive. All is well. 0.8 -0.6 -0.4	a X X 15	×	_1(XXX		```		
P a r t All real parts are non-positive. All is well. 0.8 -1	Y 45 40 -15)±-		• •				-0.2
r t All real parts are non-positive. All is well. 0.4 0.6 0.8 -1	p a				X				-0.4
All real parts are non-positive. All is well.	r t							•	
	All roal parts are no		sitiv	ο Λ	II ie		X		-0.6
		-ho:		e. A	11 13	vvCII.	• •		-0.8
±							X		-1

(VTB Capital)

25th of September 2014 110 / 151

Ultra-Sparse Finite-Differencing

For the same parameters, without anti-diffusive flux limiting, we have for the generator \tilde{L} :

	-																																		
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
	0 -0.87	0	0	0	0	0	0	0.866	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1 -0.11	-0.72	-0.04	0	0	0	0	0	0.866	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2 0	-0.11	-0.72	-0.04	0	0	0	0	0	0.866	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	3 0	0	-0.11	-0.72	-0.04	0	0	0	0	0	0.866	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4 0	0	0	-0.11	-0.72	-0.04	0	0	0	0	0	0.866	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	5 0	0	0	0	-0.11	-0.72	-0.04	0	0	0	0	0	0.866	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	6 0	0	0	0	0	0	-0.87	0	0	0	0	0	0	0.866	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	7 0.266	0	0	0	0	0	0	-1.53	0	0	0	0	0	0	1.266	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	8 0	0.266	0.234	0	0	0	0	0	-1.85	-0.15	0	0	0	0	0.234	1.266	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	9 0	0	0.266	0.234	0	0	0	0	0	-1.85	-0.15	0	0	0	0	0.234	1.266	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0 0	0	0	0.266	0.234	0	0	0	0	0	-1.85	-0.15	0	0	0	0	0.234	1.266	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1 0	0	0	0	0.266	0.234	0	0	0	0	0	-1.85	-0.15	0	0	0	0	0.234	1.266	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	2 0	0	0	0	0	0.266	0.234	0	0	0	0	0	-1.85	-0.15	0	0	0	0	0.234	1.266	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	3 0	0	0	0	0	0	0.266	0	0	0	0	0	0	-1.53	0	0	0	0	0	0	1.266	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	4 0	0	0	0	0	0	0	0.434	0	0	0	0	0	0	-0.87	0	0	0	0	0	0	0.434	0	0	0	0	0	0	0	0	0	0	0	0	0
1	5 0	0	0	0	0	0	0	0	0.434	0.566	0	0	0	0	0.793	-2.73	-0.07	0	0	0	0	0.566	0.434	0	0	0	0	0	0	0	0	0	0	0	0
1	6 0	0	0	0	0	0	0	0	0	0.434	0.566	0	0	0	0	0.793	-2.73	-0.07	0	0	0	0	0.566	0.434	0	0	0	0	0	0	0	0	0	0	0
1	7 0	0	0	0	0	0	0	0	0	0	0.434	0.566	0	0	0	0	0.793	-2.73	-0.07	0	0	0	0	0.566	0.434	0	0	0	0	0	0	0	0	0	0
1	8 0	0	0	0	0	0	0	0	0	0	0	0.434	0.566	0	0	0	0	0.793	-2.73	-0.07	0	0	0	0	0.566	0.434	0	0	0	0	0	0	0	0	0
1	.9 0	0	0	0	0	0	0	0	0	0	0	0	0.434	0.566	0	0	0	0	0.793	-2.73	-0.07	0	0	0	0	0.566	0.434	0	0	0	0	0	0	0	0
2	0 0	0	0	0	0	0	0	0	0	0	0	0	0	0.434	0	0	0	0	0	0	-0.87	0	0	0	0	0	0	0.434	0	0	0	0	0	0	0
2	1 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.634	0	0	0	0	0	0	-0.27	0	0	0	0	0	0	-0.37	0	0	0	0	0	0
2	2 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.634	1.366	0	0	0	0	2.895	-6.1	0.202	0	0	0	0	1.366	-0.37	0	0	0	0	0
2	23 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.634	1.366	0	0	0	0	2.895	-6.1	0.202	0	0	0	0	1.366	-0.37	0	0	0	0
2	4 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.634	1.366	0	0	0	0	2.895	-6.1	0.202	0	0	0	0	1.366	-0.37	0	0	0
2	15 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.634	1.366	0	0	0	0	2.895	-6.1	0.202	0	0	0	0	1.366	-0.37	0	0
2	6 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.634	1.366	0	0	0	0	2.895	-6.1	0.202	0	0	0	0	1.366	-0.37	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.634	0	0	0	0	0	0	-0.27	0	0	0	0	0	0	-0.37
2	8 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1.4	0	0	0	0	0	0	1.396	0	0	0	0	0	0
2	9 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1.4	0	0	0	0	0	9.615	-11.8	3.537	0	0	0	0
1	0 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1.4	0	0	0	0	0	9.615	-11.8	3.537	0	0	0
	0 10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1.4	0	0	0	0	U	9.615	-11.8	3.537	2 5 2 7	0
	2 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1.4	14	0	0	0	0	9.015	-11.8	3.33/	2 5 2 7
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1.4	1.4	0	0	0	0	9.015	-11.8	3.33/
1	4 U	0	0	0	0	U	0	0	0	0	0	0	0	0	0	0	0	0	0	U	0	0	U	U	0	0	0	-1.4	0	U	U	U	U	U	1.396



Ultra-Sparse Finite-Differencing

rencing 25th of September 2014

10 nber 2014 111 / 151



(VTB Capital)

How

The first growing mode with real part of eigenvalue = 1.676:



12 Ensuring continuous-time stability Local diffusivity in continuous time How bad could it possibly be?

The second growing mode, also with real part of eigenvalue = 1.676:



The instability arises (mainly) in the corner of strongest anti-diffusive flux!

Note that these instabilities are not to be confused with those arising from schemes for numerical integration in time.

Those, typically, (as for the first order explicit scheme), are associated with the mode of the most negative eigenvalue (here, about -22):



13 Numerical integation in time Actual continuous time

Recall that the spatially discrete dynamic equations

$$\frac{\partial}{\partial t}\tilde{p} = \tilde{L}^* \cdot \tilde{p} \tag{22}$$

have the continuous-time solution

$$\tilde{p}(t+\tau) = e^{\tau \cdot \tilde{L}^*} \cdot \tilde{p}(t) . \qquad (23)$$

In principle, this can be evaluated directly by the aid of a numerical routine for the computation of the exponential of $\tau \cdot \tilde{L}^*$.

Much more efficiently, one would use a method that gives the result of the action of $e^{\tau \cdot \tilde{L}^*}$ on a given vector \tilde{p} .

E.g., R. Sidje's ExpoKit [Sid98] which combines a Krylov-Arnoldi projection to reduce the dimensionality combined with a 12-th order (!) Padé(6,6) approximation to the exponential of the reduced linear system (which is much smaller than the original linear system). ExpoKit also automatically inserts sub-steps based on a formal error analysis.

For production purposes, we can do something faster.

With respect to the exponential solution, we notice

$$e^{\tau \cdot \left(\tilde{\boldsymbol{L}}_{\boldsymbol{z}}^{*} + \tilde{\boldsymbol{L}}_{\boldsymbol{y}}^{*} + \tilde{\boldsymbol{L}}_{\boldsymbol{z}\boldsymbol{y}}^{*}\right)} \neq e^{\tau \cdot \tilde{\boldsymbol{L}}_{\boldsymbol{z}}^{*}} \cdot e^{\tau \cdot \tilde{\boldsymbol{L}}_{\boldsymbol{y}}^{*}} \cdot e^{\tau \cdot \tilde{\boldsymbol{L}}_{\boldsymbol{z}\boldsymbol{y}}^{*}}$$
(66)

Strang [Str68]:

$$e^{\tau \cdot (A+B)} = e^{\frac{\tau}{2} \cdot A} \cdot e^{\tau \cdot B} \cdot e^{\frac{\tau}{2} \cdot A} + \mathcal{O}(\tau^3) .$$
(67)

We nest this:

$$\mathrm{e}^{\tau \cdot \left(\tilde{\boldsymbol{L}}_{\boldsymbol{z}}^{*} + \tilde{\boldsymbol{L}}_{\boldsymbol{y}}^{*} + \tilde{\boldsymbol{L}}_{\boldsymbol{z}\boldsymbol{y}}^{*}\right)} = \mathrm{e}^{\frac{\tau}{2} \cdot \tilde{\boldsymbol{L}}_{\boldsymbol{z}}^{*}} \cdot \mathrm{e}^{\frac{\tau}{2} \cdot \tilde{\boldsymbol{L}}_{\boldsymbol{y}}^{*}} \cdot \mathrm{e}^{\tau \cdot \tilde{\boldsymbol{L}}_{\boldsymbol{z}}^{*}} \cdot \mathrm{e}^{\frac{\tau}{2} \cdot \tilde{\boldsymbol{L}}_{\boldsymbol{y}}^{*}} \cdot \mathrm{e}^{\frac{\tau}{2} \cdot \tilde{\boldsymbol{L}}_{\boldsymbol{y}}^{*}} + \mathcal{O}(\tau^{3}) .$$
(68)

- We now have a sequence of one-dimensional propagations to perform.
- Even the mixed diffusion terms have been transformed to genuinely onedimensional propagations.
- This is as yet agnostic of any choice of *explicit* versus *implicit*! Most numerical integration schemes are a rigid combination of splitting choices and explicit or implicit steps.

We use the fully implicit second order Padé(0,2) scheme, e.g., [Jäc13b]

$$e^{\tau \tilde{\mathcal{L}}_{i}^{*}} = \left[1 - \tau \cdot \tilde{\mathcal{L}}_{i}^{*} + \frac{\tau^{2}}{2} \cdot \tilde{\mathcal{L}}_{i}^{*2}\right]^{-1} + \mathcal{O}(\tau^{3}).$$
(69)

- One (set of) pentadiagonal linear systems for each of the five steps
- Solving a pentadiagonal linear system is efficient with direct methods.
- The equations here are benign.
- The linear systems are small. Maximum size = 25.
- No pivoting required in the pentadiagonal solver.

- 6
- Other schemes are available, e.g., Craig-Sneyd [CS88], Modified Craig-Sneyd [itHM10], Hundsdorfer-Verwer [itHW09, HV03],
- or, when combined with the Strang splitting method, Crank-Nicolson scheme [CN47] Lawson-Morris [LM78], etc.
- All of these schemes have *explicit* components.
- Explicit components forego *positivity preservation*.
- We propagate Green's functions, i.e., probabilities: we cannot afford having any negative p
 _z marginals since that would result in arbitrageable vanilla option prices.
- With all of the \tilde{L}_i being Metzler, the first order fully implicit scheme perfectly preserves positivity.
- The second order fully implicity scheme is not perfect, but much better than any of the alternatives, *and* of second order (!)
- As long as we take 5 steps to the first observation horizon, we found we always have non-negative probabilities with Padé(0,2).

• In continuous time, our generator \tilde{L}^* was a perfect x-martingale operator.

- With a numerical integration scheme, we incur a small error of order τ^3 .
- At every target horizon, all we want is an implied volatility interpolator from the option prices computed from the discrete probabilities.
- We correct for the (very small) drift error by a numerical rescaling of the interpolator strike levels by the effective numerical forward.
- We interpolate implied volatilities with the arbitrage-free interpolation method discussed in [Jäc13a].
- The net result is a *spatially continuous* implied volatility profile at the target horizon that has *no drift error whatsoever*.
- No arbitrage, no drift error on the forward, everybody wins.



14 Box transition probability translation Transfer via Implied Volatility

- At *box transition points*, we need to redistribute the discrete probability masses from the previous lattice to the new lattice.
- This is not the same task as interpolating a continuous value function such as a contract price or a payoff!
- The conditions for a meaningful transfer are subtle.
- Conventional concepts of "interpolation" are simply not applicable to the probability translation problem.
- The most important requirement is: continuity of option prices!
- We use an arbitrage-free implied volatility interpolator [Jäc13a] constructed from the previous lattice nodes to infer the distribution on the new lattice.

- First, we compute out-of-the money option prices $\{v_i\}$, $i = 1...n_z$, struck at the new lattice's spot levels.
- Dropping the leftmost and rightmost nodes, we bootstrap a set of discrete z-marginal probabilities $\tilde{p}'(z'_i)$ such that

$$v'_{k} = \sum_{i=1}^{i_{\min}(e^{z'_{k}})} \tilde{p}'(z'_{i}) \cdot (e^{z'_{k}} - e^{z'_{i}}) \quad \text{for } 1 < k < \frac{n_{z}+1}{2}$$

$$v'_{k} = \sum_{i=i_{\min}(e^{z'_{k}})}^{n_{z}} \tilde{p}'(z'_{i}) \cdot (e^{z'_{i}} - e^{z'_{k}}) \quad \text{for } \frac{n_{z}+1}{2} < k < n_{z}$$
(70)
$$\tilde{p}'(z'_{\frac{n_{z}+1}{2}}) = 1 - \sum_{i \neq \frac{n_{z}+1}{2}} \tilde{p}'(z'_{i})$$

where we have assumed that n_z is odd and greater than four.

- We then redistribute p'_i in the y direction by building a two-dimensional interpolator Q(z, y) of discrete probabilities from $\tilde{p}(z_i, y_j)$, i.e., from the data of the earlier box's lattice.
- We use this two-dimensional interpolator for the purpose of interpolation in the y-direction, *conditional on a given z-level*, and so to distribute the z-marginal probability mass at some level z_i in the y-direction.
- After flooring and conditioning, the discrete Green's function on the new lattice is given by

$$\tilde{p}(z'_i, y'_j) = \tilde{p}'(z_i) \cdot \frac{\left(Q(z'_i, y'_j)\right)_+}{\sum_l \left(Q(z'_i, y'_l)\right)_+} .$$

$$(71)$$

This procedure is a generic methodology to redistribute discrete probabilities from one set of discrete nodes to another, whilst preserving the quantities that are of most importance in our context, namely:-

- the sum of all probabilities
- the expectation of the underlying, i.e., the forward,
- the prices of vanilla options on the new lattice's nodes as implied by the earlier lattice's probability distribution.

This method of translating a set of discrete probabilities from one lattice discretisation to another is in its own right a subject that is little documented in the literature and can be of use in other contexts.

(VTB Capital)	Ultra-Spa	rse Finite-	Differenc	ing	25th of September 2014	125 / 151
17 H Harres					6	
14 Box transition probability translation	on Iransfer	via Implie	d Volatılı	ty	What does this look like	
For these parameters	β κ	σ	α	ρ	, at <i>T</i> = 1,	
·	1 1	25%	50%	-50%	J.	
the implied volatility t	ransition no	des are	9:			
		0.6				
_				→ Im	pliedVolOnOldLattice	
		0.55				
		0.5		 Im	pliedVolOnNewLattice	
		0.5				
		0.45				
					_	
		0.4				
		0.25				
		0.55				
		0.3				
		0.25				
		0.2				
0.1		<u>0.2</u>				10
0.1		K, F				10
(VTB Capital)	Ultra-Spa	, Irse Fini <u>te-</u>	Differe <u>nc</u>	ing	25th of September 2014	126 / 151



e 14 Box transition probability translation Transfer via Implied Volatility What does this look like?



These translate into:

The marginal probabilities in the z-direction $(z = \ln(S))$ over z:



(VTB Capital) Ultra-Sparse Finite-Differencing 25th of September 2014 129 / 151

14 Box transition probability translation Transfer via Implied Volatility What does this look like?

To understand the apparent "enveloping", we need to view the discrete probabilities over their node count position:





(VTB Capital) Ultra-Sparse Finite-Differencing 25th of September 2014 131 / 151

15 Cash dividends

Cash dividends

- Absolute cash dividends have long been the bane of equity modelling.
- The problem is that the simple model assumption of an absolute fixed amount *irrespective of the attained spot level* is fundamentally flawed and generates hard arbitrage.
- There is a range of modelling approximations for the incorporation of cash dividends into equity modelling, and its impact on the volatility smile.

All of them generate hard arbitrage in the limit of strikes going to zero.

- In essence: when the spot has dropped to levels comparable to the dividend, the dividend *will* be cut, not least because of regulatory restrictions.
- Empirical analysis also shows that when the stock S becomes similar in size or even smaller than the forecast dividend D,

dividends become proportional.

We use the following simple dividend *process* model:-

• Define the cash dividend forecast *D* as the drop in the forward curve across the ex-dividend date *T* as seen out of today.

$$D := F(T_{-}) - F(T_{+})$$
(72)

Define the jump of the spot across the ex-dividend date from S₋ := S(T₋) to S₊ := S(T₊) given by a chosen function

$$S_+ = f(S_-)$$
. (73)

- As a balanced choice between matching the reality of dividend cuts for collapsed spots, and simplicity, we choose a continuous piecewise linear function f(·) that comprises:
 - an outright downwards jump by $D^* \approx D$, i.e., an actual cash dividend,
 - *unless* the spot S is below some threshold $\theta \approx 2D$,
 - with f(0) = 0.



The actual difference between D and D^* is *tiny for real data*.

For T = 1, $\sigma_{ATF} = 15\%$, $F_{-} = 100$, D = 5, we have (smile data to follow):



Assuming θ is given, D^* is uniquely determined by (72). Rewriting (74):

$$f(S_{-}) = S_{-} - D^{*} + \frac{D^{*}}{\theta} (\theta - S_{-})_{+}$$
 (75)

gives us

$$F_{+} - F_{-} = -D = -D_{*} + \frac{D^{*}}{\theta} \cdot p_{-}(\theta)$$
, (76)

and thus

$$D^* = \frac{D}{1 - \frac{p_-(\theta)}{\theta}}$$
(77)

where $p_{-}(K)$ is the price of a T_{-} -put struck at K.

The option price transition rules for puts and calls pre- $[p_{-}(K) \text{ and } c_{-}(K)]$ and post-dividend $[p_{+}(K) \text{ and } c_{+}(K)]$ can be readily derived:-

$$p_{+}(S_{+}) = \begin{cases} \left(1 - \frac{D^{*}}{\theta}\right) \cdot p_{-}(S_{-}) & \text{if } S_{-} \leq \theta \\ \\ p_{-}(S_{-}) - \frac{D^{*}}{\theta} \cdot p_{-}(\theta) & \text{if } S_{-} > \theta \end{cases}$$
(78)

$$c_{+}(S_{+}) = \begin{cases} \left(1 - \frac{D^{*}}{\theta}\right) \cdot c_{-}(S_{-}) + \frac{D^{*}}{\theta} \cdot c_{-}(\theta) & \text{if } S_{-} \leq \theta \\ c_{-}(S_{-}) & \text{if } S_{-} > \theta \end{cases}$$
(79)

where $c_{-}(K)$ is the price of a T_{-} -call struck at K.

(VTB Capital) Ultra-Sparse Finite-Differencing 25th of September 2014 137 / 151 15 Cash dividends Including a proportional part

If we include a proportional dividend component δ such that

$$F_{+} = (1 - \delta) \cdot F_{-} - D$$
, (80)

we have the spot transition rule

$$S_{+} = f(S_{-}) = \begin{cases} \left((1-\delta) - \frac{D^{*}}{\theta}\right) \cdot S_{-} & \text{if } S_{-} \leq \theta \\ (1-\delta) \cdot S^{-} - D^{*} & \text{if } S^{-} > \theta \end{cases}$$
(81)

and equation (77) for D^* still holds. The price transitions become:-

$$p_{+}(S_{+}) = \begin{cases} \left((1-\delta) - \frac{D^{*}}{\theta}\right) \cdot p_{-}(S_{-}) & \text{if } S_{-} \leq \theta\\ (1-\delta) \cdot p_{-}(S_{-}) - \frac{D^{*}}{\theta} \cdot p_{-}(\theta) & \text{if } S_{-} > \theta \end{cases}$$
(82)

$$c_{+}(S_{+}) = \begin{cases} \left((1-\delta) - \frac{D^{*}}{\theta}\right) \cdot c_{-}(S_{-}) + \frac{D^{*}}{\theta} \cdot c_{-}(\theta) & \text{if } S_{-} \leq \theta\\ (1-\delta) \cdot c_{-}(S_{-}) & \text{if } S_{-} > \theta \end{cases}$$
(83)



(VTB Capital)

Ultra-Sparse Finite-Differencing 25th of September 2014 139 / 151

Mary Mary Parts		6
15 Cash dividends	The smile before and after	Actual implied volatility smile data



Near the money:-





(VTB Capital)

Ultra-Sparse Finite-Differencing 25th of September 2014 141 / 151

		6
15 Cash dividends	The smile before and after	Actual implied volatility smile data



Near D = 5, it's a *bend* not a *kink*:-



15 Cash dividends The smile before and after The risk-neutral density


Within the DHI framework, we can include this piecewise affine dividend model with great ease using the techniques developed for <u>box transitions</u>!

- All we need, is a transition at which we *do not change the lattice layout*, i.e., a box transition without change in lattice dimensions.
- Instead, we populate the probability levels at T_+ from option prices struck at the lattice node levels, following the price transitions rules (82) and (83).
- This of course requires option prices at T_{-} at strikes at which there are no lattice nodes. For this, we have the implied volatility interpolator at T_{-} as at any other box transition.
- All else is already in place! And not a shred of arbitrage!
- This is another example where the probability transfer method via transformation to an implied volatility smile, and back, is of great use.

- We presented a framework to effectively convert your favourite (smooth) stochastic volatility equations into a (quasi-) parametric implied volatility surface.
- We used the Hyp-Hyp local-stochastic volatility model.
- The purpose is volatility parametrisation and Δ -hedging of vanillas.
- The framework is *defined* on a spatial discretisation of an idealised model process.
- It is numerically implemented via ultra-sparse finite differencing.
- *Dynamic Delta* largely immunizes against different parameter combinations for the same smile. It turns out that this is the same as the Local Volatility Delta.
- Extensions to normal, as opposed to log-normal, volatilities are possible.
- Spatial discretisation gives us a continuous-time-finite-state Markov chain.

- A continuous-time perfect martingale from *Itô's lemma for pure jump processes*.
- The importance of being Metzler.

This ensures continuous-time stability and is necessary for positivity preservation.

- The *seven point stencil for cross-derivatives* allows us to remain consistent and stable for all levels of correlation.
- Absorbing boundary conditions in all spot-like directions, reflecting boundary conditions for volatility to maximize the use of the grid. *No redundancies.*
- Algebraic operator splitting and anti-diffusive flux limiting. *Keep it physical*.
- The second order Strang/Padé(0,2) scheme.
 Fully implicit and fully Locally-One-Dimensional.
- Box transition probability transfer via mapping to implied volatilities, *z*-marginal bootstrapping from prices, and redistribution in the volatility direction from a conditional probability interpolator.

Also useful for *genuinely* arbitrage-free cash dividend modelling.

((VTB Capital)	Ultra-Sparse Finite-Differencing	25th of September 2014	147 / 151
	Peferences		6	
	References			
[AH11]	J. Andreasen and B. Huge Volatility Interpolation. <i>Risk</i> , March 2011. ssrn.com/abstract=169497	2.		
[BLOG93]	J.P. Boris, A.M. Landsber, LCPFCT - A Flux-Correct Technical report, Naval Re http://www.dtic.mil/get-	g, A.S. Oran, and J.H. Gardner. ed Transport Algorithm for Solving Gen esearch Laboratory, April 1993. tr-doc/pdf?AD=ADA265011.	eralized Continuity Equation	S.
[CN47]	J. Crank and P. Nicolson. A practical method for n conduction type. <i>Proc. Camb. Philos. Soc.</i> ,	umerical evaluation of solutions of part 43:50–67, 1947.	tial differential equations of	the heat-
[C R 76]	J. C. Cox and S. A. Ross. The valuation of options f Journal of Financial Econo	or alternative stochastic processes. omics, 3:145–166, March 1976.		
[C S88]	I.J.D. Craig and A.D. Sne An alternating-direction in Comput. Math. Applic., 10 www.sciencedirect.com/sc	yd. nplicit scheme for parabolic equations w 6:341–350, 1988. ience/article/pii/0898122188901502.	ith mixed derivatives.	
[Gat04]	J. Gatheral. A parsimonious arbitrage- volatility derivatives. In <i>ICBI Global Derivatives</i> faculty.baruch.cuny.edu/	free implied volatility parameterization <i>Conference Madrid</i> , 2004. jgatheral/madrid2004.pdf.	with application to the va	aluation of
[GJ13]	J. Gatheral and A. Jacquie Arbitrage-free SVI volatilit arxiv.org/pdf/1204.0646.	er. ty surfaces, March 2013. pdf.		
	(VTB Capital)	Ultra-Sparse Finite-Differencing	25th of September 2014	148 / 151

_____ References

[Hag13]	P. Hagan.
[0]	Arbitrage-free SABR. WBS 9th Fixed Income Conference, Munich, October 2013.
[Hes93]	S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options.
	The Review of Financial Studies, 6:327–343, 1993.
[HKL02]	P. Hagan, D. Kumar, and A. S. Lesniewski. Managing Smile Risk. <i>Wilmott Magazine</i> , pages 84–108, September 2002.
[HV03]	W. Hundsdorfer and J.G. Verwer. Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations. Springer, 2003. ISBN 3-540-03440-4.
[itHM10]	K.J. in 't Hout and C. Mishra. Stability of the Modified Craig-Sneyd scheme for two-dimensional convection-diffusion equations with mixed derivative term. Working paper, University of Antwerp, April 2010. arxiv.org/pdf/1011.6528.pdf.
[itHW09]	K.J. in 't Hout and B.D. Welfert. Unconditional stability of second-order ADI schemes applied to multi-dimensional diffusion equations with mixed derivative terms. <i>Appl. Num. Math.</i> , 59:677–692, 2009.
[Jäc13a]	P. Jäckel. Clamping Down on Arbitrage, December 2013. www.jaeckel.org/ClampingDownOnArbitrage.pdf.

(VTB Capital)

Ultra-Sparse Finite-Differencing 25th of September 2014 149 / 151

6

A STAN H. H. P. P. Stand Stand Stand	6
References	

[Jäc13b]	P. Jäckel. Finite differencing schemes as Padé approximants, April 2013. www.jaeckel.org/FiniteDifferencingSchemesAsPad%C3%A9Approximants.pdf.
[Jäc13c]	P. Jäckel. Let's be rational. www.jaeckel.org/LetsBeRational.pdf, November 2013.
[JK07]	P. Jäckel and C. Kahl. Hyp Hyp Hooray, June 2007. www.jaeckel.org/HypHypHooray.pdf.
[Kuz07]	D. Kuzmin. Operator splitting techniques. Technical report, Universität Dortmund, July 2007. www.mathematik.uni-dortmund.de/~kuzmin/cfdintro/lecture11.pdf.
[Kuz10]	D. Kuzmin. A Guide to Numerical Methods for Transport Equations. self-published, 2010. www.mathematik.uni-dortmund.de/~kuzmin/Transport.pdf.
[LM78]	J.D. Lawson and J.LI. Morris. The Extrapolation of First Order Methods for Parabolic Partial Differential Equations I. SIAM Journal of Numerical Analysis, 15:1212–1224, 1978. www.cs.uwaterloo.ca/research/tr/1977/CS-77-20.pdf.
[LS11]	A. Lipton and A. Sepp. Filling the Gaps. Risk, pages 66-71, October 2011. ssrn.com/abstract=2150646.

[Min88]	H. Minc. <i>Nonnegative Matrices.</i> John Wiley and Sons, 1988.
[MM80]	J.D. Macbeth and L.J. Merville. Tests of the Black-Scholes and Cox Call Option Valuation models. Journal of Finance, 35(2):285-301, 1980. efinance.org.cn/cn/fe/19800505Tests%20of%20the%20Black-Scholes%20and%20Cox%20Call%20Option% 20Valuation%20Models,%20pppdf.
[Rub83]	M. Rubinstein. Displaced diffusion option pricing. Journal of Finance, 38:213–217, March 1983.
[Sid98]	R. Sidje. Expokit: A Software Package for Computing Matrix Exponentials. ACM Trans. Math. Softw., 24:130–156, 1998. www.maths.uq.edu.au/expokit.
[SS91]	E. M. Stein and J. C. Stein. Stock Price Distribution with Stochastic Volatility : An Analytic Approach. <i>Review of Financial Studies</i> , 4:727–752, 1991.
[Str68]	G. Strang. On the construction and comparison of difference schemes. SIAM J. Numerical Analysis, 5:509–517, 1968.
[SZ99]	R. Schöbel and J. Zhu. Stochastic Volatility With an Ornstein Uhlenbeck Process: An Extension. <i>European Finance Review</i> , 3:23–46, 1999. ssrn.com/abstract=100831.

(VTB Capital)

Ultra-Sparse Finite-Differencing 25th of September 2014 151 / 151

6