# Finite Difference Methods for Financial Problems

# Part 4: ADI Schemes

Kwant Skool
Copenhagen
October 2010

Jesper Andreasen
Danske Markets
kwant.daddy@danskebank.com

## Outline

- ADI and split methods.
- The classic ADI scheme for zero correlation.
- Transforms to eliminate correlation.
- The Mitchell scheme.
- The Mitchell/Craig-Sneyd scheme for non-zero correlation.
- Predictor-corrector Craig-Sneyd scheme.
- 3-dimensional Craig-Sneyd for non-zero correlation.
- Higher dimensional Craig-Sneyd.
- Jump-diffusions.

- Forward schemes revisited.

- Conclusion.

## ADI and Split Schemes

- Consider the 2D PDE problem:

$$0=\frac{\partial V}{\partial t}+D_x V+D_y V \quad ,D_x=-\frac{r}{2}+\mu_x\frac{\partial}{\partial x}+\frac{1}{2}\sigma_x^2\frac{\partial^2}{\partial x^2} \quad ,D_y=-\frac{r}{2}+\mu_y\frac{\partial}{\partial y}+\frac{1}{2}\sigma_y^2\frac{\partial^2}{\partial y^2} \tag{1}$$

- For now we assume zero correlation (there is no $V_{xy}$) in the above PDE.

- Note, that we have done an equal split for the discounting into the $x$ and $y$ operators.

- A direct attack on (1) a la

$$[1-\frac{1}{2}\Delta t D]v(t)=[1+\frac{1}{2}\Delta t D]v(t+\Delta t) \quad ,D=D_x+D_y \tag{2}$$

…will work.

- *But:* the trouble is that direct solution of (1) is computationally very costly because the matrix $(1-\frac{1}{2}\Delta t D)$, though sparse, is very large.

- For 100x100 $(x,y)$ discretisation the matrix $(1-\frac{1}{2}\Delta t D)$ will be of size 10,000 x 10,000, i.e. 100,000,000 entries in total. A naive approach to this would carry a computational cost of $O(10000^3)=O(10^{12})$ operations.

- Instead, we will attempt to split the matrix solution in the two dimensions, which is what is normally termed Alternating Directions Implicit (ADI) methods.

- One of these methods is the Locally One Dimensional (LOD) scheme.

- Given a discretisation of the two spatial operators we have that the exact solution to (1) can be written

$$e^{-\Delta t D}V(t)=V(t+\Delta t) \qquad (3)$$

- Expanding the left hand side yields

$$
\begin{aligned}
&[1-\Delta t D_x - \Delta t D_y]V(t) = V(t+\Delta t) + O(\Delta t^2) \\
&\Downarrow \\
&[1-\Delta t D_y][1-\Delta t D_x]V(t) = V(t+\Delta t) + O(\Delta t^2)
\end{aligned}
\tag{4}
$$

- This leads to the $O(\Delta t)$ LOD scheme

$$
\begin{aligned}
&[1-\Delta t D_x]v(t+\Delta t/2) = v(t+\Delta t) \quad \forall y \\
&[1-\Delta t D_y]v(t) = v(t+\Delta t/2) \qquad \forall x
\end{aligned}
\tag{5}
$$

- The first equation has to be solved for all levels of $y$ and the second equation has to be solved for all levels of $x$.

- One can interpret the scheme as $x$ being diffused and $y$ standing still over the first half time step and vice-versa over the second half step.

- For a 100x100 discretisation in $(x, y)$ we have now broken down the computational task into

- 100 systems of $x$ problems of dimension 100.
- 100 systems of $y$ problems of dimension 100.

- As each of the matrix problems is linear, the computational cost of this scheme is $O(2\cdot 100\cdot 100)=O(10^4)$ which is much-much-much less than the cost of the direct naive attack $O(10^{12})$.

- So the LOD scheme eliminates the large computational cost.

- Further, the LOD scheme is unconditionally stable and it has the nice property that the transition can be guaranteed to be positive

$$(1-\Delta t D_y)^{-1}(1-\Delta t D_x)^{-1}\geq 0 \qquad (6)$$

- …under mild conditions on the parameters:

$$|\mu_x|\Delta x\leq\sigma_x^2 \quad ,|\mu_y|\Delta y\leq\sigma_y^2 \qquad \text{or upwind.} \qquad (7)$$

- The drawback of the LOD scheme is that it is only $O(\Delta t)$ accurate.

- The challenge for the ADI technique is to extend it to obtain $O(\Delta t^2)$ accuracy and to include correlation while retaining stability.

- This is what most of the material presented today is going to focus on.

## The Classic ADI Split

- Given a discretisation of the two spatial operators we can write the solution to (1) as

$$
\begin{aligned}
&V(t+\Delta t/2)=e^{\Delta t/2D_x+\Delta t/2D_y}V(t+\Delta t)\\
&V(t)=e^{\Delta t/2D_x+\Delta t/2D_y}V(t+\Delta t/2)
\end{aligned}
\tag{8}
$$

- If the two operators *commute* ($D_xD_y=D_yD_x$) then we can write (8) as

$$
\begin{aligned}
&e^{-\Delta t/2D_x}V(t+\Delta t/2)=e^{\Delta t/2D_y}V(t+\Delta t)\\
&e^{-\Delta t/2D_y}V(t)=e^{\Delta t/2D_x}V(t+\Delta t/2)
\end{aligned}
\tag{9}
$$

- Equation (9) leads to the classic ADI split scheme

$$
\begin{aligned}
&(1-\frac{1}{2}\Delta tD_x)v(t+\Delta t/2)=(1+\frac{1}{2}\Delta tD_y)v(t+\Delta t)\\
&(1-\frac{1}{2}\Delta tD_y)v(t)=(1+\frac{1}{2}\Delta tD_x)v(t+\Delta t/2)
\end{aligned}
\tag{10}
$$

- In (10) the first equation has to be solved for each level of $y$, whereas the second has to be solved for each level of $x$.

- The classic ADI split scheme has a very nice symmetry to it: first you go explicit in one direction and implicit in other. Then over the next step you reverse the order.

- Exercise: Find sufficient conditions for $D_x$ and $D_y$ to commute.

- Exercise: Show that the scheme (10) has an accuracy of $O(\Delta t^2 + \Delta x^2 + \Delta y^2)$.

- Exercise: Show that the scheme (10) is von Neumann stable.

**Transformation and Orthogonalisation**

- It is often possible to transform away the correlation terms in models.

- As an example consider the model

$$
\begin{aligned}
dx &= \mu_x(t,x,y)dt + \sigma_x(t,x)dW^x \\
dy &= \mu_y(t,x,y)dt + \sigma_y(t,y)dW^y \\
dW^x \cdot dW^y &= \rho(t)dt
\end{aligned}
\tag{11}
$$

- Define the *increasing* and thereby, *invertible*, functions

$$
\begin{aligned}
X(t,x) &= \frac{1}{\sigma_x(t,x_0)} + \int_{x_0}^{x} \frac{1}{\sigma_x(t,z)}dz \\
Y(t,y) &= \frac{1}{\sigma_y(t,y_0)} + \int_{y_0}^{y} \frac{1}{\sigma_y(t,z)}dz
\end{aligned}
\tag{12}
$$

- Then

$$
\begin{aligned}
dX &= [\frac{\partial}{\partial t} + D_x]Xdt + dW^x \\
dY &= [\frac{\partial}{\partial t} + D_y]Ydt + dW^y
\end{aligned}
\tag{13}
$$

- Now consider $Y' = -\rho X + Y$. We have

$$
\begin{aligned}
& dY' = (-\rho[\frac{\partial}{\partial t} + D_x]X - \rho[\frac{\partial}{\partial t} + D_y]Y)dt + \sqrt{1-\rho^2}dW' \\
& dW^x \cdot dW' = 0
\end{aligned}
\tag{14}
$$

- This indicates that many problems can be transformed into problems that can be handled by the classic ADI (10).

- It is, however, not for all problems that this technique can be applied.

- When we apply this technique it can be necessary to do it so that the transform preserves one of the coordinates.

- This is for example the case if we wish to solve barrier option problems.

- Exercise: Perform the transform a 2F Black-Scholes case.

- Exercise: Perform the transform for the Heston model case.

**The Mitchell Scheme**

- The Mitchell scheme looks less symmetrical than the classic ADI split:

$$\begin{aligned} &[1-\theta_x \Delta t D_x]u = [1+(1-\theta_x)\Delta t D_x + \Delta t D_y]v(t+\Delta t) \\ &[1-\theta_y \Delta t D_y]v(t) = u - \theta_y \Delta t D_y v(t+\Delta t) \end{aligned} \tag{15}$$

- Again, in (15) the first equation has to be solved for each level of $y$, whereas the second has to be solved for each level of $x$.

- Here $u$ does *not* have the interpretation as a half step in time.

- The scheme can be shown to be generally $O(\Delta t)$ accurate and stable for $\theta_x, \theta_y \in [1/2, 1]$, and $O(\Delta t^2)$ accurate for the case $\theta_x = \theta_y = 1/2$.

- To prove the accuracy, we note that (15) can be re-written as

$$[1-\theta_x \Delta t D_x][1-\theta_y \Delta t D_y]v(t) = [1+(1-\theta_x)\Delta t D_x + (1-\theta_y)\Delta t D_y]v(t+\Delta t) + \theta_x \theta_y \Delta t^2 D_x D_y v(t+\Delta t)$$

(16)

- Expanding the left-hand side and rearranging yields

$$[1-\theta_x \Delta t D_x - \theta_y \Delta t D_y]v(t) = [1+(1-\theta_x)\Delta t D_x + (1-\theta_y)\Delta t D_y]v(t+\Delta t) + \theta_x \theta_y \Delta t^2 D_x D_y (v(t+\Delta t) - v(t))$$
$$\Downarrow$$
$$[1-\theta_x \Delta t D_x - \theta_y \Delta t D_y]v(t) = [1+(1-\theta_x)\Delta t D_x + (1-\theta_y)\Delta t D_y]v(t+\Delta t) + O(\Delta t^3) \quad , v(t+\Delta t) = v(t) + O(\Delta t)$$

(17)

- On the other hand the exact solution (1) satisfies

$$e^{-\Delta t/2D} v(t) = e^{\Delta t/2D} v(t+\Delta t) \quad , D = D_x + D_y$$
$$\Downarrow \qquad (18)$$
$$[1-\Delta t/2D]v(t) = [1+\Delta t/2D]v(t+\Delta t) + O(\Delta t^3)$$

- So (18) implies that for the case $\theta_x = \theta_y = 1/2$, (16) is valid to order $O(\Delta t^3)$. Hence, for the centred case the total scheme is $O(\Delta t^2)$ order accurate.

- Similar arguments can be used to show $O(\Delta t)$ accuracy of the scheme for the $\theta_x, \theta_y \neq 1/2$ cases.

- We note that, surprisingly, the accuracy of the scheme does *not* depend on the operators commuting.

- With respect to stability, we note that insertion of an eigensolution of the form

$$e(t,x,y) = g^{-t} e^{ikx} e^{ily} \tag{19}$$

- …into (16) leads to

$$g = \frac{[1+(1-\theta_x)\Delta t D_x + (1-\theta_y)\Delta t D_y + \theta_x\theta_y \Delta t^2 D_x D_y](e^{ikx+ily})}{[1-\theta_x \Delta t D_x - \theta_y \Delta t D_y + \theta_x\theta_y \Delta t^2 D_x D_y](e^{ikx+ily})}$$

$$= \frac{1+\theta_x\theta_y\Delta t^2 a_x a_y - (1-\theta_x)\Delta t a_x - (1-\theta_y)\Delta t a_y}{1+\theta_x\theta_y\Delta t^2 a_x a_y + \theta_x \Delta t a_x + \theta_y \Delta t a_y} \tag{20}$$

$$D_x e^{ikx} = -[\frac{r}{2} - i\frac{\mu_x}{\Delta x}\sin(k\Delta x) + \frac{\sigma_x^2}{\Delta x^2}(1-\cos(k\Delta x))]e^{ikx} \equiv -a_x e^{ikx} \quad ,a_x \in \mathbb{C} \quad ,\mathrm{Re}\, a_x \geq 0$$

$$D_y e^{ily} = -[\frac{r}{2} - i\frac{\mu_y}{\Delta y}\sin(l\Delta y) + \frac{\sigma_y^2}{\Delta y^2}(1-\cos(l\Delta y))]e^{ilxy} \equiv -a_y e^{ily} \quad ,a_y \in \mathbb{C} \quad ,\mathrm{Re}\, a_y \geq 0$$

- From this we have that $|g| \leq 1$ for $\theta_x, \theta_y \in [1/2, 1]$ if

$$|\mu_x|\Delta x \leq \sigma_x^2 \quad ,|\mu_y|\Delta y \leq \sigma_y^2 \tag{21}$$

- Condition (21) can be replaced by upwinding and is probably not strictly necessary.

- Hence, the Mitchell scheme (15) is von Neumann stable and the time centred version is second order accurate in time.

- So the Mitchell scheme is the equivalent of the CN scheme for one dimension.

**Mitchell/Craig-Sneyd with Correlation**

- Though we can often find transformations that eliminate the correlation/cross derivative in the PDE, it is not always the case, and in all circumstances it would be nice with a method that directly handles correlation/cross terms.

- Consider the PDE

$$0 = \frac{\partial V}{\partial t} + D_x V + D_y V + D_{xy} V \quad , D_{xy} = \sigma_x \sigma_y \rho \frac{\partial^2}{\partial x \partial y} \tag{22}$$

- …here the cross derivative can be approximated by

$$\begin{aligned}
&\frac{\partial^2 f(x,y)}{\partial x \partial y} = \delta_{xy} f(x,y) + O(\Delta x \cdot \Delta y) \\
&\delta_{xy} f(x,y) = \delta_x \delta_y f(x,y) \\
&\qquad = \frac{1}{4\Delta x \Delta y}[f(x+\Delta x, y+\Delta y) - f(x+\Delta x, y-\Delta y) - f(x-\Delta x, y+\Delta y) + f(x-\Delta x, y-\Delta y)]
\end{aligned} \tag{23}$$

- The Craig-Sneyd scheme is

$$\begin{aligned}
&[1-\theta_x \Delta t D_x]u = [1+(1-\theta_x)\Delta t D_x + \Delta t D_y + \Delta t D_{xy}]v(t+\Delta t) \\
&[1-\theta_y \Delta t D_y]v(t) = u - \theta_y \Delta t D_y v(t+\Delta t)
\end{aligned} \tag{24}$$

- Applying the cross derivative to the von Neumann eigen solution yields

$$\begin{aligned}
\delta_{xy}[e^{ikx+ily}] &= \frac{e^{ikx+ily}}{4\Delta x \Delta y}(e^{ik\Delta x} - e^{-ik\Delta x})(e^{il\Delta y} - e^{-il\Delta y}) \\
&= -\frac{e^{ikx+ily}}{\Delta x \Delta y}\sin(k\Delta x)\sin(l\Delta y)
\end{aligned} \tag{25}$$

- So the von Neumann growth factor for this scheme is

$$g=\frac{[1+(1-\theta_x)\Delta tD_x+(1-\theta_y)\Delta tD_y+\theta_x\theta_y\Delta t^2D_xD_y+\Delta tD_{xy}](e^{ikx+ily})}{[1-\theta_x\Delta tD_x-\theta_y\Delta tD_y+\theta_x\theta_y\Delta t^2D_xD_y](e^{ikx+ily})}$$

$$=\frac{1+\theta_x\theta_y\Delta t^2a_xa_y-(1-\theta_x)\Delta ta_x-(1-\theta_y)\Delta ta_y-\Delta ta_{xy}}{1+\theta_x\theta_y\Delta t^2a_xa_y+\theta_x\Delta ta_x+\theta_y\Delta ta_y}$$

$$D_xe^{ikx}=-[\frac{r}{2}-i\frac{\mu_x}{\Delta x}\sin(k\Delta x)+\frac{\sigma_x^2}{\Delta x^2}(1-\cos(k\Delta x))]e^{ikx}\equiv-a_xe^{ikx}\quad,a_x\in\mathbb{C}\quad,\mathrm{Re}\,a_x\geq 0 \tag{26}$$

$$D_ye^{ily}=-[\frac{r}{2}-i\frac{\mu_y}{\Delta y}\sin(l\Delta y)+\frac{\sigma_y^2}{\Delta y^2}(1-\cos(l\Delta y))]e^{ily}\equiv-a_ye^{ily}\quad,a_y\in\mathbb{C}\quad,\mathrm{Re}\,a_y\geq 0$$

$$D_{xy}e^{ikx+ily}=-[\frac{\sigma_x\sigma_y\rho}{\Delta x\Delta y}\sin(k\Delta x)\sin(l\Delta y)]e^{ikx+ily}\equiv-a_{xy}e^{ikx+ily}\quad,a_{xy}\in\mathbb{R}$$

- Provided that $\mathrm{Re}(a_x+a_y)/2+a_{xy}\geq 0$ (which can be shown using hi skool trigonometrix), the scheme will be von Neumann stable for $\theta_x,\theta_y\in[1/2,1]$ under the same drift condition as for the uncorrelated case:

$$|\mu_x|\Delta x\leq\sigma_x^2\quad,|\mu_y|\Delta y\leq\sigma_y^2 \tag{27}$$

- Again, the drift conditions are probably not strictly necessary.

## Craig-Sneyd Predictor-Corrector

- The scheme of the previous section has the drawback that directly applied it will only be $O(\Delta t)$ accurate.

- This can be improved to $O(\Delta t^2)$ if the scheme is combined with a predictor-corrector methodology.

- The predictor step is to solve for $w(t)$ in:

$$\begin{aligned} &[1-\theta_x \Delta t D_x]u = [1+(1-\theta_x)\Delta t D_x + \Delta t D_y + \Delta t D_{xy}]v(t+\Delta t) \\ &[1-\theta_y \Delta t D_y]w(t) = u - \theta_y \Delta t D_y v(t+\Delta t) \end{aligned} \tag{28}$$

- The corrector step is to solve for $v(t)$ in:

$$\begin{aligned} &[1-\theta_x \Delta t D_x]u = [1+(1-\theta_x)\Delta t D_x + \Delta t D_y + (1-\theta_{xy})\Delta t D_{xy}]v(t+\Delta t) + \theta_{xy}\Delta t D_{xy} w(t) \\ &[1-\theta_y \Delta t D_y]v(t) = u - \theta_y \Delta t D_y v(t+\Delta t) \end{aligned} \tag{29}$$

- For $\theta_x = \theta_y = \theta_{xy} = 1/2$ second order accuracy $O(\Delta t^2)$ is obtained.

- To prove this, we note that (29) can be rearranged as

$$
\begin{aligned}
&[1-\theta_x \Delta t D_x - \theta_y \Delta t D_y] v(t) = [1+(1-\theta_x)\Delta t D_x + (1-\theta_y)\Delta t D_y + (1-\theta_{xy})\Delta t D_{xy}] v(t+\Delta t) \\
&\qquad + \theta_x \theta_y \Delta t^2 D_x D_y (v(t+\Delta t) - v(t)) + \theta_{xy} \Delta t D_{xy} w \\
&\qquad \Downarrow \\
&[1-\theta_x \Delta t D_x - \theta_y \Delta t D_y] v(t) = [1+(1-\theta_x)\Delta t D_x + (1-\theta_y)\Delta t D_y + (1-\theta_{xy})\Delta t D_{xy}] v(t+\Delta t) \\
&\qquad + \theta_{xy} \Delta t D_{xy} v(t) + O(\Delta t^3)
\end{aligned}
\tag{30}
$$

- Here we have used that

$$
\begin{aligned}
&v(t) = v(t+\Delta t) + O(\Delta t) \\
&w(t) = v(t) + O(\Delta t^2)
\end{aligned}
\tag{31}
$$

- We again note that the exact solution satisfies

$$
\begin{aligned}
&e^{-\Delta t/2D}V(t)=e^{\Delta t/2D}V(t+\Delta t) \quad ,D=D_x+D_y+D_{xy} \\
&\Downarrow \\
&[1-\Delta t/2D]V(t)=[1+\Delta t/2D]V(t+\Delta t)+O(\Delta t^3)
\end{aligned}
\tag{32}
$$

- Combining (30) and (32) leads to the conclusion the that the predictor-corrector scheme is $O(\Delta t^2)$ accurate for $\theta_x=\theta_y=\theta_{xy}=1/2$.

- Note, that the Craig-Sneyd predictor-corrector approach is roughly 2.5 times more computational intensive than the Mitchell scheme for zero correlation.

- This means that there might be a significant computational gain from performing the orthogonalisation when technically feasible.

- We have successfully used the CS predictor-approach for a variety of models: Heston, Heston with local volatility, 2-dimensional Gaussian model, Cheyette, etc.

- Exercise: Why is the difference between the CS and the Mitchell zero correlation scheme approximately a factor 2.5 in computational effort?

## Craig-Sneyd in Dimension 3

- For dimension 3 the Craig-Sneyd predictor step is

$$
\begin{aligned}
&[1-\theta_x \Delta t D_x]u_x = [1+(1-\theta_x)\Delta t D_x + \Delta t D_y + \Delta t D_y + \Delta t D_z + \Delta t D_c]v(t+\Delta t)\\
&[1-\theta_y \Delta t D_y]u_y = u_x - \theta_y \Delta t D_y v(t+\Delta t)\\
&[1-\theta_z \Delta t D_z]w(t) = u_y - \theta_z \Delta t D_z v(t+\Delta t)\\
&D_c = \sigma_{xy}\delta_{xy} + \sigma_{xz}\delta_{xz} + \sigma_{yz}\delta_{yz}
\end{aligned}
\qquad (33)
$$

- The corrector step is to solve for $v(t)$ in:

$$
\begin{aligned}
&[1-\theta_x \Delta t D_x]u_x = [1+(1-\theta_x)\Delta t D_x + \Delta t D_y + \Delta t D_y + \Delta t D_z + (1-\theta_c)\Delta t D_c]v(t+\Delta t) + \theta_c \Delta t D_c w(t)\\
&[1-\theta_y \Delta t D_y]u_y = u_x - \theta_y \Delta t D_y v(t+\Delta t)\\
&[1-\theta_z \Delta t D_z]v(t) = u_y - \theta_z \Delta t D_z v(t+\Delta t)
\end{aligned}
$$

(34)

- The Craig-Sneyd (1988) paper proves stability and $O(\Delta t^2)$ convergence for the $\theta_x = \theta_y = \theta_c = 1/2$ case.

- We have used this scheme extensively for PRDC trades with the model

$$\begin{aligned} \frac{dS}{S} &= (r_0 - r_1)dt + \sigma dW_S \quad [FX] \\ dr_i &= \kappa_i(a_i - r_i)dt + \eta_i dW_i \quad [Dom\,\&\,For\ Rates] \end{aligned} \tag{35}$$

- …with predictor-corrector or rotation to handle/eliminate the correlation terms.

- Computational times in the order of 5-30 seconds for combat accuracy for 30 year trades.

## Craig-Sneyd Schemes in n Dimensions

- The Craig-Sneyd Scheme can be extended to dimension $n$:

$$\begin{aligned}
&u_0 = [1 + \Delta t \sum_i D_i + \Delta t D_c] v(t+\Delta t) \quad , D_c = \sum_{i \neq j} \sigma_{ij} \delta_{ij} \\
&[1 - \theta_i \Delta t D_i] u_i = u_{i-1} - \theta_i \Delta t D_i v(t+\Delta t) \\
&v(t) = u_n
\end{aligned} \tag{36}$$

- …but it will only be stable for $\theta_i = 1/2$ for non-zero correlation in up to three dimensions.

- In that case $O(\Delta t^2)$ convergence can be obtained if predictor-corrector stepping is applied.

- Word on the street is that a modification of this type of scheme successfully has been applied to a four factor model of the type

$$
\begin{aligned}
&\frac{dS}{S} = (r_0 - r_1)dt + \sqrt{z}\sigma(t,S)dW_S \quad [FX] \\
&dr_i = \kappa_i(a_i - r_i)dt + \eta_i dW_i \quad [Dom\,\&\,For\ Rates] \\
&dz = \theta(1-z)dt + \varepsilon\sqrt{z}dZ \quad [SV] \\
&dZ \cdot dW_i = dZ \cdot dW_S = 0
\end{aligned}
\tag{37}
$$

- …at a well renowned place.

- Exercise: For the zero-correlation case prove that the scheme (36) is von Neumann stable.

## Jump Diffusion Problems

- Merton's jump diffusion model

$$
\begin{aligned}
& dx = (r - q - \frac{1}{2}\sigma^2 - \lambda m)dt + \sigma dW + \ln(1+J)dN \quad, S = e^x \\
& \Pr(dN = 1) = \lambda dt \quad, 1 + J \sim e^{\mu - \frac{1}{2}\delta^2 + \delta N(0,1)} \quad, m = E[J] = e^{\mu} - 1
\end{aligned}
\tag{38}
$$

- …leads to the Partial *Integro* Differential Equation (PIDE):

$$
\begin{aligned}
& 0 = \frac{\partial V}{\partial t} + DV + \lambda E[\Delta V] \\
& D = -r + (r - q - \lambda m)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(\frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x}) \\
& E[\Delta V] = \int (V(t,y) - V(t,x))\varphi(y - x)dy \\
& \varphi(\Delta x) = \phi(\Delta x; \mu - \frac{1}{2}\delta^2, \delta)
\end{aligned}
\tag{39}
$$

- The point is that expected jump in price can be seen as a *convolution*

$$I \equiv E[\Delta V] = \varphi * V - V = (\varphi * -1)V \tag{40}$$

- Convolutions can be handled via Fourier transforms, or more specifically FFT (Nrc::four1(), Nrc::realft()), so

$$\begin{aligned} & F[I] = F[\varphi] \cdot F[V] \quad , F[g](\omega) = \int e^{i2\pi\omega x} g(x) dx \quad , F^{-1}[h](x) = \int e^{-i2\pi\omega x} h(\omega) d\omega \quad , i = \sqrt{-1} \\ & \Downarrow \\ & I(x) = F^{-1}[F[\varphi] \cdot F[V]](x) - V(x) \end{aligned} \tag{41}$$

- Given a function observed at the discrete points

$$g(x_1), \ldots, g(x_n) \quad , x_i = x_0 + i\Delta x \quad , n = 2^k , \tag{42}$$

- …the Fast Fourier Transform (FFT) algorithm can compute all values

$$h(\omega_i) = F[g(\omega_i)] \quad , \omega_i = i / \Delta x \tag{43}$$

- …in one go, at the cost of $O(n\ln n)$. Equivalently for the inverse Fourier transform.

- Using the FFT the values of $I(x_1),\ldots,I(x_n)$ can be obtained in $O(3n\ln n)=O(n\ln n)$ time.

- Inspired by the ADI/split schemes one could be inspired to use the scheme

$$[1-\frac{1}{2}\Delta tD]v(t_h)=[1+\frac{1}{2}\Delta tD]v(t_{h+1})+\Delta t\lambda(\varphi^*-1)v(t_{h+1}) \qquad (44)$$

- …where the convolution part is handled by FFT.

- The scheme (44) is stable but only $O(\Delta t)$ accurate. Besides that, schemes of this type appear to create a bias in the options prices because the Brownian motion part is handled CN style whereas the jump part is handled explicitly.

- A better approach is to use a predictor-corrector scheme where

$$
\begin{aligned}
&[1-\frac{1}{2}\Delta t D]w(t_h) = [1+\frac{1}{2}\Delta t D]v(t_{h+1}) + \Delta t\lambda(\varphi^* - 1)v(t_{h+1}) \\
&[1-\frac{1}{2}\Delta t D]v(t_h) = [1+\frac{1}{2}\Delta t D]v(t_{h+1}) + \Delta t\lambda(\varphi^* - 1)\frac{1}{2}[w(t_h) + v(t_{h+1})]
\end{aligned}
\tag{45}
$$

- …again with the application of FFT for the convolution part. Using this we obtain a $O(\Delta t^2)$ accurate and stable scheme.

- A more funky approach (Andersen and Andreasen (1999)) is

$$
\begin{aligned}
&[1-\frac{1}{2}\Delta t D]v(t_{h+1/2}) = [1+\frac{1}{2}\Delta t\lambda(\varphi^* - 1)]v(t_{h+1}) \\
&[1+\frac{1}{2}\Delta t\lambda(\varphi^* - 1)]v(t_h) = [1+\frac{1}{2}\Delta t D]v(t_{h+1/2})
\end{aligned}
\tag{46}
$$

- Set

$$
\begin{aligned}
&g = [1+\frac{1}{2}\Delta t D]v(t_{h+1/2}) \\
&h = [1+\frac{1}{2}\Delta t\lambda(\varphi^* - 1)]
\end{aligned}
\tag{47}
$$

- …then the second step in (46) can be handled by the FFT using

$$
\begin{aligned}
&F[h]\cdot F[v(t_h)] = F[g] \\
&\Downarrow \\
&v(t_h) = F^{-1}\left[\frac{F[g]}{F[h]}\right] = F^{-1}\left[\frac{F[g]}{1+\Delta t\lambda/2(F[\varphi]-1)}\right]
\end{aligned}
\tag{48}
$$

- Both methods (45) and (46) are stable and $O(\Delta t^2)$ accurate.

- Exercise: Prove accuracy and stability of the scheme (45).

## Forward Equations -- Again

- As for the 1-dimensional case it is possible to find corresponding forward equations.

- For the continuous time matrix problem, $D = D_x + D_y + D_{xy}$, we have that

$$
\begin{aligned}
& 0 = p(t)'\frac{\partial v(t)}{\partial t} + p(t)'Dv(t) \\
& \Rightarrow \\
& 0 = \int_0^T p(t)'\frac{\partial v(t)}{\partial t}dt + \int_0^T p(t)'Dv(t)dt \\
& = [p(t)'v(t)]_{t=0}^{t=T} - \int_0^T \frac{\partial p(t)'}{\partial t}v(t)dt + \int_0^T p(t)'Dv(t)dt \\
& \Rightarrow \\
& v(t, x_k, y_l) = p(T)'v(T) \\
& 0 = -\frac{\partial p}{\partial t} + D'p \quad , p(0) = 1_{x=x_k}1_{y=y_l}
\end{aligned}
\tag{49}
$$

- This can be used for direct discretisation using the methods we have presented in this section.

- It still avoids any extra considerations for boundary conditions at the edges, but the approach is not fully consistent with the backward schemes.

- A direct approach for the LOD scheme yields

$$
\begin{aligned}
&p(t_h) = 1_{x=x_k} 1_{y=y_l} \\
&[1-\Delta t D_y]' \, p(t_{h+1/2}) = p(t_h) \\
&[1-\Delta t D_x]' \, p(t_{h+1}) = p(t_{h+1/2})
\end{aligned}
\tag{50}
$$

- For the CS scheme we get the predictor step

$$
\begin{aligned}
&p(t_h) = 1_{x=x_k} \cdot 1_{y=y_l} \\
&[1-\theta_y \Delta t D_y]' q_0 = p(t_h) \\
&[1-\theta_x \Delta t D_x]' r_0 = q_0 \\
&p(t_{h+1}) = [1+(1-\theta_x)\Delta t D_x + \Delta t D_y + \Delta t D_{xy}]' r_0 - \theta_y \Delta t D_y{}' q_0
\end{aligned}
\tag{51}
$$

- …and the corrector step

$$
\begin{aligned}
&p(t_h) = 1_{x=x_k} \cdot 1_{y=y_l} \\
&[1-\theta_y \Delta t D_y]' q = p(t_h) \\
&[1-\theta_x \Delta t D_x]' r = q \\
&p(t_{h+1}) = [1+(1-\theta_x)\Delta t D_x + \Delta t D_y + (1-\theta_{xy})\Delta t D_{xy}]' r + \theta_{xy} \Delta t D_{xy}' r_0 - \theta_y \Delta t D_y' q
\end{aligned}
\tag{52}
$$

- Morten will show the use of this.

- Exercise: Show (50-52).

## Conclusion

- The ADI methods are powerful FD methods for solving multidimensional problems and eliminates the need for solving very large matrix systems..

- The challenge is to retain second order accuracy in time and to include correlation without sacrificing the stability. But it can be done, at least up to dimension 3.

- The Craig-Sneyd scheme is the equivalent of the Crank-Nicolson method for one dimensional problems.

- Correlation can be handled by rotation or by predictor-corrector technique in the CS scheme.

- The ADI schemes have successfully been applied to many practical problems in finance.

- The ADI split technique can be applied to other problems such as for example the jump-diffusion problem.

- The forward equation idea for one dimension also applies to ADI schemes.

- ADI schemes naturally lend themselves to GPU style parallelisation.
- Questions?