ADI FD Schemes for the Numerical Solution of the Three-dimensional Heston–Cox–Ingersoll–Ross PDE

Tinne Haentjens

Department of Mathematics and Computer Science, University of Antwerp, Middelheimlaan 1, 2020 Antwerp, Belgium

Abstract. This paper deals with the numerical solution of the time-dependent, three-dimensional Heston–Cox–Ingersoll–Ross PDE, with all correlations nonzero, for the fair pricing of European call options. We apply a finite difference discretization on non-uniform spatial grids and then numerically solve the semi-discrete system in time by using an Alternating Direction Implicit scheme. We show that this leads to a highly efficient and stable numerical solution method.

Keywords: financial option pricing, Heston model, Cox–Ingersoll–Ross model, finite difference methods, ADI schemes.
PACS: 02.60.Lj, 02.70.Bf, 89.65.Gh

HESTON–COX–INGERSOLL–ROSS MODEL AND THE OPTION PRICING PDE

We consider a general asset price model given by the system of stochastic differential equations (SDEs)

$$dS_\tau = R_\tau S_\tau \, d\tau + S_\tau \sqrt{V_\tau}\, dW_\tau^1, \quad dV_\tau = \kappa(\eta - V_\tau)\, d\tau + \sigma_1 \sqrt{V_\tau}\, dW_\tau^2, \quad dR_\tau = a(b(\tau) - R_\tau)\, d\tau + \sigma_2 \sqrt{R_\tau}\, dW_\tau^3 \quad (1)$$

for $0 < \tau \leq T$ with $T > 0$ the maturity time of the option. Here S_τ, V_τ, R_τ denote random variables that represent the asset price, its variance, and the interest rate, resp., at time τ. The asset price model (1) constitutes a natural extension of the well-known Black–Scholes model where the volatility $\sqrt{V_\tau}$ and the interest rate R_τ both evolve randomly over time. The stochastic one-factor models for V_τ and R_τ were proposed by Heston and Cox, Ingersoll & Ross, resp., cf. [1, 9, 10, 11]. The full Heston–Cox–Ingersoll–Ross model has been considered, for example, in [2, 6].

In (1), the quantities κ, η, σ_1 and a, σ_2 are given positive real constants and W_τ^1, W_τ^2, W_τ^3 are Brownian motions, with (arbitrary) given correlation factors ρ_{12}, ρ_{13}, $\rho_{23} \in [-1,1]$. The function b is a given, deterministic, positive function of time, which renders a slightly more general model than the original Cox–Ingersoll–Ross model as described in [3]. In contrast to the Hull–White interest rate model, the interest rate in the Cox–Ingersoll–Ross model cannot become negative, which forms an advantage of the latter model. We note that in the numerical literature it is often assumed that the so-called *Feller condition* is fulfilled for both the volatility and the interest rate models, i.e., $2\kappa\eta > \sigma_1^2$ and $2ab(\tau) > \sigma_2^2$ resp. In the present paper we make no assumptions concerning this condition.

Let $u(s,v,r,t)$ denote the fair price of a European call option if at time $\tau = T - t$ the asset price equals s, its variance equals v and the interest rate equals r. The payoff yields the initial condition $u(s,v,r,0) = \max(0, s-K)$, for a given strike price $K > 0$. It can be shown from financial option pricing theory that $u(s,v,r,t)$ satisfies the partial differential equation (PDE)

$$\begin{aligned}\frac{\partial u}{\partial t} &= \tfrac{1}{2}s^2 v \frac{\partial^2 u}{\partial s^2} + \tfrac{1}{2}\sigma_1^2 v \frac{\partial^2 u}{\partial v^2} + \tfrac{1}{2}\sigma_2^2 r \frac{\partial^2 u}{\partial r^2} + \rho_{12}\sigma_1 s v \frac{\partial^2 u}{\partial s \partial v} + \rho_{13}\sigma_2 s \sqrt{vr} \frac{\partial^2 u}{\partial s \partial r} \\ &\quad + \rho_{23}\sigma_1\sigma_2\sqrt{vr}\frac{\partial^2 u}{\partial v \partial r} + rs\frac{\partial u}{\partial s} + \kappa(\eta - v)\frac{\partial u}{\partial v} + a(b(T-t) - r)\frac{\partial u}{\partial r} - ru \qquad (2)\end{aligned}$$

for $s > 0$, $v > 0$, $r > 0$, $0 < t \leq T$. We refer to (2) as the *Heston–Cox–Ingersoll–Ross (HCIR) PDE*. This is a time-dependent convection-diffusion-reaction equation, with mixed derivative terms, on an unbounded, three-dimensional spatial domain. In case of all correlations nonzero, an exact solution of the HCIR PDE in (semi) closed-form is not known in the literature. We numerically solve (2) via the method-of-lines approach. First we use finite difference (FD) schemes on non-uniform Cartesian grids in the (s,v,r)-domain for the space discretization of the PDE. Then we apply an Alternating Direction Implicit (ADI) method. This type of time discretization methods is tailored to semi-discrete systems stemming from multi-dimensional PDEs with mixed spatial-derivative terms.

Numerical Analysis and Applied Mathematics ICNAAM 2012
AIP Conf. Proc. 1479, 2195-2199 (2012); doi: 10.1063/1.4756628

The aim of this paper is to show that using an ADI scheme for the time discretization yields an efficient and stable numerical solution method for the three-dimensional HCIR PDE.

NUMERICAL SOLUTION OF THE HESTON–COX–INGERSOLL–ROSS PDE

To numerically solve the HCIR PDE (2), the spatial domain is restricted to a bounded set $[0,S_{\max}]\times[0,V_{\max}]\times[0,R_{\max}]$ with fixed values $S_{\max},V_{\max},R_{\max}$ chosen sufficiently large. The following Dirichlet and Neumann boundary conditions are applied for a European call option

$$u(0,v,r,t)=0, \qquad \frac{\partial u}{\partial s}(S_{\max},v,r,t)=1, \qquad u(s,V_{\max},r,t)=s \qquad \frac{\partial u}{\partial r}(s,v,R_{\max},t)=0. \tag{3}$$

At the boundaries $v=0$ and $r=0$, we consider inserting these values into the HCIR PDE. A mathematical foundation for this is given in [4, 5]. Note that this is done irrespective of whether or not the Feller condition is satisfied for v and/or r.

First, a Cartesian spatial grid is chosen, analogously as in [7]. A difference here is that $r\in[0,R_{\max}]$ instead of $r\in[-R_{\max},R_{\max}]$. We have smooth non-uniform meshes $0=s_0<s_1<\ldots<s_{m_1}=S_{\max}$, $0=v_0<v_1<\ldots<v_{m_2}=V_{\max}$ and $0=r_0<r_1<\ldots<r_{m_3}=R_{\max}$. The mesh in the s-direction has relatively many points throughout a given interval $[S_{\text{left}},S_{\text{right}}]\subset[0,S_{\max}]$ containing the strike K. This is natural, because it alleviates numerical difficulties due to the initial (payoff) function that has a discontinuous derivative at $s=K$. It is also natural to place relatively many mesh points around $v=0$, resp. $r=0$, for numerical reasons, as the HCIR PDE is convection-dominated in the v- resp. r-direction for $v\approx 0$, resp. $r\approx 0$ and the initial function is nonsmooth. A second motivation for these choices is that the region of interest in the (s,v,r) domain lies around $(K,0,0)$.

For the FD discretization of the HCIR PDE, we also follow the ideas in [7], where a combination of central and upwind second-order schemes was used to obtain an effective spatial discretization. However, we now have a different model for the interest rate, which is similar to the model for the volatility. The derivatives $\partial u/\partial r$, $\partial^2 u/\partial r^2$ and $\partial^2 u/\partial s\partial r$, $\partial^2 u/\partial v\partial r$ are discretized in the same manner as the corresponding derivatives for v in the Heston–Hull–White PDE [7]. For example, as for $\partial u/\partial v$ in the region $v>\eta$, we apply an upwind scheme for $\partial u/\partial r$ in the region $r>\max\{b(T-t)|0\le t\le T\}$. Complete details will be given in a future paper.

The FD discretization of the initial-boundary value problem for the HCIR PDE leads to an initial value problem for a large system of stiff ordinary differential equations (ODEs),

$$U'(t)=A(t)U(t)+g(t) \quad (0\le t\le T), \quad U(0)=U_0. \tag{4}$$

For any $t\ge 0$, $A(t)$ is a given real matrix and $g(t)$, U_0 are given real vectors, of order M, with $M=m_1m_2(m_3+1)$. The vector g depends on the boundary conditions (3) and the vector U_0 is obtained from the initial condition. The entries of the solution vector $U(t)$ to (4) form approximations to the exact option values $u(s,v,r,t)$ at the spatial grid points.

For the time discretization ADI schemes are considered. Standard implicit schemes such as Crank–Nicolson are in general computationally too demanding. Splitting schemes of the ADI type are particularly useful for semi-discrete systems stemming from multi-dimensional PDEs. The matrix $A(t)$ and the vector $g(t)$ are split into four simpler matrices, resp. vectors,

$$A(t)=A_0+A_1+A_2+A_3(t), \quad g(t)=g_0+g_1+g_2+g_3(t). \tag{5}$$

Here A_0 represents the part of $A(t)$ that stems from the FD discretization of all mixed derivative terms in the HCIR PDE. Note that A_0 is nonzero whenever one of the correlation factors ρ_{12}, ρ_{13}, ρ_{23} is nonzero. The matrices A_1, A_2, $A_3(t)$ represent the parts of $A(t)$ that stem from the FD discretization of all spatial derivatives in the s-, v- and r-directions, resp., and the ru term is distributed evenly over $A_1,A_2,A_3(t)$. We decompose $g(t)$ analogous to that of $A(t)$.

We study two ADI schemes, namely the Douglas (Do) scheme and the Modified Craig–Sneyd (MCS) scheme. Let $\theta>0$ be a given real parameter, let $\Delta t=T/N$ with integer $N\ge 1$, and $t_n=n\cdot\Delta t$. Set $\Delta g_n=g_3(t_n)-g_3(t_{n-1})=g(t_n)-g(t_{n-1})$. The following schemes generate, in a one-step manner, approximations U_n to the exact solution values $U(t_n)$ of (4) for $n=1,2,\ldots,N$:

Do scheme:

$$\begin{cases} Y_0 = U_{n-1} + \Delta t \left(A\left(t_{n-1}\right) U_{n-1} + g\left(t_{n-1}\right)\right), \\ Y_j = Y_{j-1} + \theta \, \Delta t \, A_j \left(Y_j - U_{n-1}\right) \quad (j = 1, 2), \\ Y_3 = Y_2 + \theta \, \Delta t \left(A_3\left(t_n\right) Y_3 - A_3\left(t_{n-1}\right) U_{n-1} + \Delta g_n\right), \\ U_n = Y_3. \end{cases}$$

MCS scheme:

$$\begin{cases} Y_0 = U_{n-1} + \Delta t \left(A\left(t_{n-1}\right) U_{n-1} + g\left(t_{n-1}\right)\right), \\ Y_j = Y_{j-1} + \theta \, \Delta t \, A_j \left(Y_j - U_{n-1}\right) \quad (j = 1, 2), \\ Y_3 = Y_2 + \theta \, \Delta t \left(A_3\left(t_n\right) Y_3 - A_3\left(t_{n-1}\right) U_{n-1} + \Delta g_n\right), \\ \widehat{Y}_0 = Y_0 + \theta \, \Delta t \, A_0 \left(Y_3 - U_{n-1}\right), \\ \widetilde{Y}_0 = \widehat{Y}_0 + \left(\tfrac{1}{2} - \theta\right) \Delta t \left(A\left(t_n\right) Y_3 - A\left(t_{n-1}\right) U_{n-1} + \Delta g_n\right), \\ \widetilde{Y}_j = \widetilde{Y}_{j-1} + \theta \, \Delta t \, A_j \left(\widetilde{Y}_j - U_{n-1}\right) \quad (j = 1, 2), \\ \widetilde{Y}_3 = \widetilde{Y}_2 + \theta \, \Delta t \left(A_3\left(t_n\right) \widetilde{Y}_3 - A_3\left(t_{n-1}\right) U_{n-1} + \Delta g_n\right), \\ U_n = \widetilde{Y}_3. \end{cases}$$

The MCS scheme can be seen as an extension to the Do scheme. The Do scheme has a classical order of consistency one and the MCS scheme is of order two for any given θ. The A_0 part, representing all mixed derivative terms, is always treated in an *explicit* fashion. The A_1, A_2, $A_3(t)$ parts are successively treated in an *implicit* fashion.

The key advantage of ADI schemes is that the arising linear systems are efficiently solved by *LU* factorization, as the pertinent matrices are essentially tridiagonal or pentadiagonal. For standard implicit methods, the matrices that arise have a (very) large bandwidth and solving by *LU* factorization is in general computationally too demanding. Moreover, with the splitting (5), the time-dependency of A is only passed onto the simpler matrix A_3, which yields an additional advantage of ADI schemes over standard implicit methods.

Under the objective of retaining unconditional stability of these ADI schemes, we take the same values of θ as in [7], namely, for Do $\theta = \frac{2}{3}$ and for MCS $\theta = \max\{\frac{1}{3}, \frac{2}{13}(2\gamma+1)\}$ with $\gamma = \max\{|\rho_{12}|, |\rho_{13}|, |\rho_{23}|\}$. Although there are no theoretical results known in literature at present on the stability of ADI schemes for general three-dimensional convection-diffusion-reaction equations with mixed derivative terms , these values for θ were carefully selected from the theoretical stability results that are available for two-dimensional convection-diffusion-reaction equations with mixed derivative terms and three-dimensional pure diffusion problems with mixed derivatives. Here stability is always understood in the von Neumann sense. It is our aim to rigorously investigate the performance of these ADI schemes in the solution of the HCIR PDE. In the following section we provide a numerical illustration.

NUMERICAL EXPERIMENTS

In the special case where $\rho_{13} = \rho_{23} = 0$, a semi closed-form analytic formula for the exact European call option price function u is known [2]. We employ this formula to gain insight into the *global spatial error* of the FD discretization, defined by $e(m_1, m_2, m_3) = \max\{|u(s_i, v_j, r_k, T) - U_l(T)|\}$. For the ADI schemes we study the *global temporal error*, defined by $\widehat{e}(\Delta t; m_1, m_2, m_3) = \max\{|U_l(T) - U_{N,l}|\}$. Here $U(T)$ denotes the exact solution vector to the semidiscrete HCIR PDE (4) at time T. The maximum is taken over all points $(s_i, v_j, r_k) \in ROI$, the index $l = l(i, j, k)$ corresponds to the spatial grid point (s_i, v_j, r_k) and *ROI* is a natural region of interest, chosen here as $\frac{1}{2}K < s_i < \frac{3}{2}K,\ 0 < v_j < 1,\ 0 < r_k < \frac{1}{4}$. The temporal and spatial discretization errors are both measured in the maximum norm, which is highly relevant to financial applications. In order to compute the temporal error for a given spatial grid, we use a sufficiently accurate reference value for $U(T)$, obtained by applying the MCS scheme to (4) with $N = 10000$ time steps.

For our numerical experiments we consider two sets for the Heston and Cox–Ingersoll–Ross parameter values

	κ	η	σ_1	a	b	σ_2	ρ_{12}	ρ_{13}	ρ_{23}	T
set 1	3.0	0.12	0.04	0.20	0.05 $(-0.01e^{-(T-t)} + 0.05)$	0.03	0.6	0 (0.2)	0 (0.4)	1
set 2	1.0	0.09	1.00	0.22	0.034 $(-0.014e^{-2.1(T-t)} + 0.034)$	0.11	-0.3	0 (-0.5)	0 (-0.2)	5

The strike price and spatial domain are taken as $K = 100$ and $[0, 14K] \times [0, 10] \times [0, 1]$. Note that the Feller condition is not satisfied for the volatility model parameters in set 2.

Figure 1 displays in the left column the estimated global spatial errors vs. $1/m$ for $m = 10, 15, \ldots, 75$ where $m_1 = 2m$ and $m_2 = m_3 = m$. Here b is constant and $\rho_{13} = \rho_{23} = 0$. The observed order of convergence for the spatial error is appr. equal to two, namely 1.94 for set 1 and 1.96 for set 2, which is as desired. The middle column shows

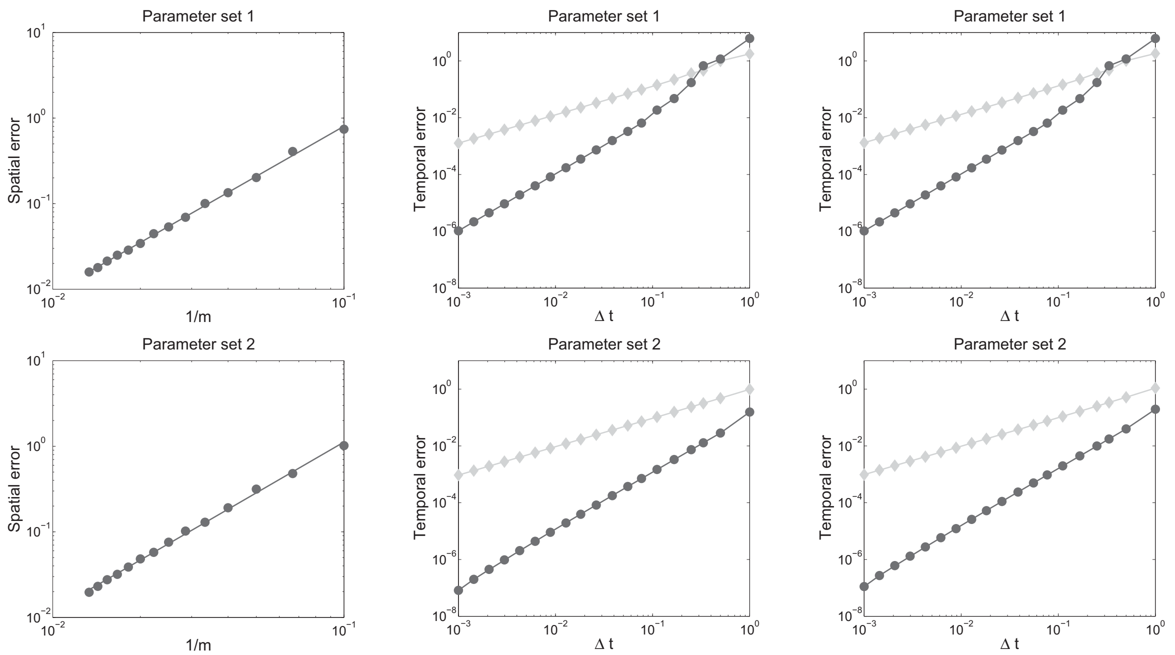

FIGURE 1. Left: Spatial error $e(2m,m,m)$ vs $1/m$ for $m = 10, 15, \ldots, 75$ with b constant, $\rho_{13} = \rho_{23} = 0$. Middle: Temporal error $\widehat{e}(\Delta t; 100, 50, 50)$ vs Δt with b constant, $\rho_{13} = \rho_{23} = 0$. Right: Temporal error $\widehat{e}(\Delta t; 100, 50, 50)$ vs Δt with b time-dependent and nonzero correlations. ADI schemes: Do scheme (diamond) and MCS scheme (circle).

the temporal errors of the Do and MCS schemes for a range of step sizes $10^{-3} \leq \Delta t \leq 10^{0}$ when $m = 50$. Here, b is again constant and $\rho_{13} = \rho_{23} = 0$. The temporal errors are uniformly bounded from above by a moderate value and decay monotonically as Δt decreases. Further experiments show that the temporal errors are only weakly affected by the chosen number of spatial grid points. These observations are compatible with an unconditionally stable behavior of the Do and MCS schemes. A convergence behavior is observed for the Do and MCS schemes with an order appr. equal to one, resp. two, which agrees with their classical orders of consistency. Note also that temporal errors are in general much smaller for MCS compared to Do. The right column displays the temporal errors for the parameter sets with nonzero correlations and time-dependent mean-reversion level b. In this case there is no longer an analytical price formula available. The observations concerning the temporal errors remain the same as for $\rho_{13} = \rho_{23} = 0$ and b constant. Note that for set 2, now, the Feller condition is also not satisfied for the interest rate model parameters whenever $t \geq 4.6381$. We conclude that the MCS scheme is preferred over the Do scheme. In the future, more exotic and American-style options will be studied in detail, where we will consider the extension of the pertinent research in [7, 8].

ACKNOWLEDGEMENTS

The author is grateful to Karel in ’t Hout for sharing his knowledge and for his valuable remarks. This work has been supported financially by the Research Foundation – Flanders, FWO contract no. G.0125.08.

REFERENCES

1. L. B. G. Andersen & V. V. Piterbarg, Interest Rate Models, Volume II: Term Structure Models, Atlantic Financial Press, 2010.
2. G. Bakshi, C. Cao & Z. Chen, *Empirical performance of alternative option pricing models*, J. Fin. **52** (1997) 2003–2049
3. J. C. Cox, J. E. Ingersoll & S. A. Ross, *A theory of the term structure of interest rates*, Econometrica **53**2 (1985) 385–407
4. E. Ekström & J. Tysk, *The Black–Scholes equation in stochastic volatility models*, J. Math. Anal. Appl. **368** (2010) 498–507
5. E. Ekström & J. Tysk, *Boundary conditions for the single-factor term structure equation*, Ann. Appl. Prob. **21** (2011) 332–350
6. L. A. Grzelak & C. W. Oosterlee, *On the Heston model with stochastic interest rates*, SIAM J. Fin. Math. **2** (2011) 255–286
7. T. Haentjens & K. J. in 't Hout, *ADI finite difference schemes for the Heston–Hull–White PDE*, J. Comp. Finan. (Fall 2012).
8. T. Haentjens, K. J. in 't Hout & K. Volders, *ADI schemes with Ikonen–Toivanen splitting for pricing American put options in the Heston model*, Numer. Anal. Appl. Math., eds. T. E. Simos et. al., AIP Conf. Proc. **1281** (2010) 231–234 .
9. J. C. Hull, Options, Futures and Other Derivatives, 6th ed., Prentice Hall, New Jersey, 2006.
10. S. E. Shreve, Stochastic Calculus for Finance II, Springer, New York, 2004.
11. D. Tavella & C. Randall, Pricing Financial Instruments, Wiley, New York, 2000.