Zurich University
of Applied Sciences

School of
Management and Law

PDE Solvers for the Heston Model with Stochastic Correlation

Department of Banking, Finance and Insurance

Norbert Hilber

ZHAW School of Management and Law
St.-Georgen-Platz 2
P.O. Box
8401 Winterthur
Switzerland

Department of Banking, Finance and Insurance
https://www.zhaw.ch/en/sml/about-us/departments/department-of-banking-finance-insurance/

Author, Contact
Norbert Hilber
Norbert.Hilber@zhaw.ch

July 2019

PDE SOLVERS FOR THE HESTON MODEL WITH STOCHASTIC CORRELATION

NORBERT HILBER

Abstract. The Heston model with stochastic correlation is an extension to the stochastic volatility model of Heston. Modelling the dynamics of the correlation between the underlying and its instantaneous variance by an additional stochastic process renders the pricing of derivatives in this model trivariate. As a consequence, the pricing of European or weakly path-dependent options such as barrier options by means of partial differential equations (PDEs) leads to degenerate, parabolic PDEs in three space dimensions. We provide numerical schemes to approximate the solution of three-dimensional parabolic PDEs with variable, but factorising coefficients over a rectangular domain subject to an arbitrary combination of different conditions on the boundary of the domain. These schemes are based on a finite difference discretisation of 2nd or 4th order in space and a ADI time marching scheme supplemented with a Richardson extrapolation. We apply these PDEs solvers to price a European call option and to compare the results with those cited in the scientific literature. We are able to solve the pricing problem within a fraction of a second to an accuracy which is comparable to MC simulations and/or Fourier transform methods.

Contents

1. Introduction

Modeling volatility is a major challenge in quantitative finance, in discrete and continuous time alike. In financial time series modelling decades of research have

Date: July 10, 2019.

2010 *Mathematics Subject Classification.* 35K15, 65M06, 91G20, 91G30.

Key words and phrases. Finite differences, partial differential equations, options, stochastic correlation.

brought forth stochastic difference equations such as the GARCH model and its successors (see e.g. [Tsa10]), whereas in continuous time finance stochastic volatility models as solutions to stochastic differential equations were introduced to overcome the shortcomings of the standard Black-Scholes-Merton theory, see for example [Ber16] and the references therein. In the field of continuous time finance, the model of Heston [Hes93, Rou13] has become the benchmark stochastic volatility model. Here, the correlation of the Brownian motions that drive the underlying and its instantaneous variance are assumed to be constant (typically negative). One reason for the popularity of the Heston model is that its characteristic function (CF) is known in closed-form, which gives access to fast and accurate pricing and/or calibration. It is, however, well known that, besides other mis-modelling issues, the Heston model is not able to reproduce market volatility smiles for short-termed options. As a consequence, researchers relaxed the model by allowing its coefficients deterministically depend on time or by introducing additional sources of randomness. For example, Bates [Bat00] considers a two factor geometric jump-diffusion model with state dependent jump intensity. However, the correlation between the different sources of randomness is assumed - in the Bates model as well as in all the other attempts made to relax the Heston model - to be constant. Real data suggest that correlation between financial quantities is not constant, but stochastic as well. Recently, stochastic correlation models are used to price volatility and correlation derivatives [Bos14]. It comes therefore at no surprise that one tries to model the correlation between the underlying and its instantaneous variance in the model of Heston also stochastically [TEG16]. The price one has to pay for this increase of modelling flexibility is that the CF of the extended model is no longer known in closed form. Thus, to regain analytical tractability, one approximates the CF in such a way that the approximated CF admits a closed form solution. However, there is no error analysis available which quantifies the difference between option prices obtained by the exact and the approximate CF. Obviously, one could avoid this approximation problem by a Monte Carlo (MC) simulation, where one is confronted with the discretisation error only (with respect to the time steps and the number of simulated paths [TEG17]). Since the MC simulation delivers option prices randomly and somehow too slow, we suggest to obtain prices by solving the corresponding pricing partial differential equation (PDE). This is non-trivial as well, since the PDE is of degenerate, parabolic type and involves three space dimensions. However, combining a fourth-order finite difference scheme on a stretched grid with respect to space with an ADI scheme (with respect to time) supplemented with a Richardson extrapolation provides a simple and yet fast and accurate PDE solver. As we believe that the numerical solution of linear reaction-advection-diffusion equations is of its own interest, we provide a simple ADI finite difference scheme to solve such PDEs with respect to a variety of boundary conditions (admittedly only on rectangular domains).

The paper is organised as follows. In chapter 2, we describe the Heston model with stochastic correlation and state the pricing PDE in this model. In chapter 3, we develop a fully discrete approximation scheme to solve the (possibly transformed) pricing PDE of chapter 2 numerically. We apply a second and/or fourth-order finite difference discretisation with respect to the space variables and a ADI time-stepping scheme (followed by a Richardson extrapolation) to discretise with respect to time. Here, special attention is paid to the realisation of various boundary conditions. In the last chapter 4, we apply the approximation method suggested in chapter 3 to the particular pricing problem of the Heston model with stochastic correlation. Herein, we compare the results of different PDE solvers with results obtained by MC simulations and Fourier transform methods.

2. Pricing equation

Denote by $S(t) > 0$ the stock price at time t, and let $V(t) \geq 0$ be its instantaneous variance. Let $r \in \mathbb{R}$ be the (constant) continuously compounded risk free, and let $q \geq 0$ be the (constant) continuous dividend yield of the stock. The Heston model with stochastic correlation [TEG16] is as follows. For a stochastic correlation process $Z(t) \in [-1, 1]$ the vector process $\mathbf{X}(t) := (S(t), V(t), Z(t))^\top$ solves the system of stochastic differential equations (SDEs)

$$\mathrm{d}\mathbf{X}(t) = \boldsymbol{\mu}(\mathbf{X}(t), t)\mathrm{d}t + \boldsymbol{\sigma}(\mathbf{X}(t), t)\mathrm{d}\mathbf{W}(t)\ , \quad \mathbf{X}(0) = \mathbf{x}_0\ , \tag{2.1}$$

with $\mathbf{x}_0 = (s_0, v_0, z_0)^\top \in \mathbb{R}^+ \times \mathbb{R}^+ \times]-1, 1[$. The coefficients $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ of the SDE (2.1) are given by, with $\mathbf{x} = (s, v, z)^\top$ and $\delta, \kappa, m > 0$,

$$\boldsymbol{\mu}(\mathbf{x}, t) := \begin{pmatrix} (r-q)s \\ \kappa(m-v) \\ a(z) \end{pmatrix}\ , \quad \boldsymbol{\sigma}(\mathbf{x}, t) := \begin{pmatrix} \sqrt{v}s & 0 & 0 \\ 0 & \delta\sqrt{v} & 0 \\ 0 & 0 & b(z) \end{pmatrix} \mathbf{L}\ . \tag{2.2}$$

In (2.2), the 3×3-matrix $\mathbf{L}$ is the Cholesky decomposition, i.e., $\mathbf{L}\mathbf{L}^\top = \boldsymbol{\rho}$, of the correlation matrix

$$\boldsymbol{\rho} := \begin{pmatrix} 1 & z & \rho_1 \\ z & 1 & \rho_2 \\ \rho_1 & \rho_2 & 1 \end{pmatrix}\ , \quad \rho_i \in [-1; 1].$$

Furthermore, the functions a and b specify the correlation process. In particular, [TEG16] consider $Z(t)$ to be a OU process or a Jacobi process, compare with table 1. Therein, we also state conditions such that $Z(t) \in]-1, 1[$, $\forall t \geq 0$, with probability 1. Finally, remark that the standard Brownian motions $\mathbf{W}(t) := (W_1(t), W_2(t), W_3(t))^\top$ in (2.1) are independent.

model	$a(z)$	$b(z)$	condition for $Z(t) \in]-1, 1[$
OU	$\kappa_Z(m_Z - z)$	δ_Z	$\frac{\sqrt{\kappa_Z}(\pm 1 - m_Z)}{\delta_Z} \to \pm\infty$
Jacobi	$\kappa_Z(m_Z - z)$	$\delta_Z\sqrt{1 - z^2}$	$\kappa_Z > \frac{\delta_Z^2}{1 \pm m_Z}$

Table 1. The stochastic correlation process $Z(t)$ is specified as an OU process or a Jacobi process. $\kappa_Z, \delta_Z > 0$, $m_Z \in]-1, 1[$.

Remark 2.1. The Heston model [Hes93] without stochastic correlation can be obtained from (2.2) by omitting the third component, i.e., $\mathbf{X}(t) := (S(t), V(t))^\top$. That is, we delete the third row of $\boldsymbol{\mu}$ as well as the third row and third column of the matrix $\boldsymbol{\sigma}$ and replace the variable z in the matrix $\boldsymbol{\rho}$ by the constant ρ (which is the correlation of the Brownian motions that drive the stock and the instantaneous variance).

We consider a European style financial derivative with payoff function $s \mapsto g(s) \in \mathbb{R}$ and maturity $T > 0$ written on S in the model $\mathbf{X}$. By the general principles of derivatives pricing (see e.g. [DS06]), the value $V(\mathbf{x}, t) := V(x_1, x_2, x_3, t) := V(s, v, z, t)$ of the derivative is

$$V(\mathbf{x}, t) := \mathbb{E}\big[e^{-r(T-t)} g(S(T)) \mid \mathbf{X}(t) = \mathbf{x}\big] \ . \tag{2.3}$$

A formal application of the Feynman-Kac theorem (see e.g. [HS00]) yields the pricing partial differential equation (PDE) satisfied by V

$$\begin{cases} \partial_t V + \mathcal{A}V - rV &= 0 \quad \text{in} \quad G \times [0, T[\\ V(s, T) &= g(s) \quad \text{in} \quad G \end{cases} \tag{2.4}$$

where $\mathcal{A}$ is the infinitesimal generator of the process $\mathbf{X}$ and where $G := \mathbb{R}^+ \times \mathbb{R}^+ \times]-1, 1[$. The operator $\mathcal{A}$ acting on functions $f(\mathbf{x})$ is defined as

$$\mathcal{A}f := \frac{1}{2}\mathrm{tr}[\boldsymbol{\sigma}(\mathbf{x}, t)\boldsymbol{\sigma}(\mathbf{x}, t)^\top D^2 f] + \boldsymbol{\mu}(\mathbf{x}, t)^\top \nabla f \ , \tag{2.5}$$

where $D^2 f = (\partial_{x_i x_j} f)_{1 \le i,j \le 3}$ and $\nabla f = (\partial_{x_1} f, \partial_{x_2} f, \partial_{x_3} f)^\top$ denotes the Hessian and the gradient, respectively, of f. Furthermore, $\mathrm{tr}[\mathbf{M}] = \sum_{i=1}^d m_{i,i}$ denotes the trace of a $d \times d$-matrix $\mathbf{M}$. Calculating $\mathcal{A}$ for the Heston model with stochastic correlation (2.1)-(2.2) yields

$$\begin{aligned} \mathcal{A} \ = \ & \frac{1}{2}s^2 v \partial_{ss} + \frac{1}{2}\delta^2 v \partial_{vv} + \frac{1}{2} b(z)^2 \partial_{zz} \\ & + \delta s v z \partial_{sv} + \rho_1 s \sqrt{v} b(z) \partial_{sz} + \rho_2 \delta \sqrt{v} b(z) \partial_{vz} \\ & + (r - q) s \partial_s + \kappa(m - v)\partial_v + a(z)\partial_z \ . \end{aligned} \tag{2.6}$$

Remark 2.2. The described European style setting can be easily relaxed to barrier options. For example, a down-and-out barrier option with barrier $0 < B < S(0)$ written on the stock in the Heston model with stochastic correlation admits the price

$$V(\mathbf{x}, t) := \mathbb{E}\big[e^{-r(T-t)} g(S(T)) 1_{\{\min_{t \in]0,T]} S(t) > B\}} \mid \mathbf{X}(t) = \mathbf{x}\big] \ .$$

This value function still solves the PDE (2.4), but on the domain $G =]B, \infty[\times \mathbb{R}^+ \times]-1, 1[$ and subject to the boundary condition $V(B, v, z, t) = 0$, $\forall (v, z, t) \in \mathbb{R}^+ \times]-1, 1[\times \mathbb{R}^+$.

To solve (2.4) with a finite difference scheme, we switch to time-to-maturity $t \mapsto T - t$, restrict G to the bounded domain $[0, s_r[\times [0, v_r[\times]-1, 1[$ (which we again call G) and set boundary conditions on (parts) of ∂G. Note carefully that we do not specify boundary conditions on the faces $\{s = 0\}$ and $\{v = 0\}$, i.e., we solve the PDE also on these, see below. Thus, the pricing problem becomes a special case of the following parabolic partial differential equation: Find $w = w(\mathbf{x}, t)$, $\mathbf{x} := (x_1, x_2, x_3)$, such that

$$\left\{ \begin{array}{rcll} \partial_t w + \mathcal{B}w & = & 0 & \text{in} \quad G\times]0,T] \\ w(\mathbf{x},0) & = & g(\mathbf{x}) & \text{in} \quad G \end{array} \right. , \tag{2.7}$$

subject to homogeneous boundary conditions. The operator $\mathcal{B}$ is given by

$$\begin{aligned} \mathcal{B} \;:=\; & a_1(\mathbf{x})\partial_{x_1x_1} + a_2(\mathbf{x})\partial_{x_2x_2} + a_3(\mathbf{x})\partial_{x_3x_3} \\ & + a_4(\mathbf{x})\partial_{x_1x_2} + a_5(\mathbf{x})\partial_{x_1x_3} + a_6(\mathbf{x})\partial_{x_2x_3} \\ & + b_1(\mathbf{x})\partial_{x_1} + b_2(\mathbf{x})\partial_{x_2} + b_3(\mathbf{x})\partial_{x_3} + c(\mathbf{x}) \\ =\; & \sum_{\substack{i,j=1 \\ j>i}}^{3} a_{i+j+1}(\mathbf{x})\partial_{x_ix_j} + \sum_{i=1}^{3} \Big[a_i(\mathbf{x})\partial_{x_ix_i} + b_i(\mathbf{x})\partial_{x_i} + \frac{1}{3}c(\mathbf{x}) \Big] \end{aligned} \tag{2.8}$$

for trivariate functions a_j, b_j and c, and the domain is

$$G :=]x_1^l, x_1^r[\times]x_2^l, x_2^r[\times]x_3^l, x_3^r[$$

for some $-\infty < x_i^l < x_i^r < \infty$, $i = 1, 2, 3$.

Example 2.1. If we set $(x_1, x_2, x_3) = (s, v, z)$, then it follows from (2.6) that the coefficients a_j, b_j, c appearing in the operator $\mathcal{B}$ (2.8) of the Heston model with stochastic correlation are given by

$$\begin{aligned} & a_1(\mathbf{x}) = -\frac{1}{2}x_1^2x_2, \;\; a_2(\mathbf{x}) = -\frac{1}{2}\delta^2 x_2, \;\; a_3(\mathbf{x}) = -\frac{1}{2}b(x_3)^2, \\ & a_4(\mathbf{x}) = -\delta x_1 x_2 x_3, \;\; a_5(\mathbf{x}) = -\rho_1 x_1 \sqrt{x_2}\, b(x_3), \;\; a_6(\mathbf{x}) = -\rho_2 \delta \sqrt{x_2}\, b(x_3), \\ & b_1(\mathbf{x}) = -(r-q)x_1, \;\; b_2(\mathbf{x}) = -\kappa(m - x_2), \;\; b_3(\mathbf{x}) = -a(x_3), \;\; c(\mathbf{x}) = r\ . \end{aligned}$$

We describe the boundary conditions in (2.7). The (spatial) cube G has six faces, which we define by the short hand notation

$$\mathcal{F}_{x_i^l} := \{x_i = x_i^l\}, \quad \mathcal{F}_{x_i^r} := \{x_i = x_i^r\}, \quad i = 1, 2, 3\ .$$

To be more precise, by $\mathcal{F}_{x_1^l}$ for example we understand the set

$$\mathcal{F}_{x_1^l} := \{x_1 = x_1^l\} := \big\{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 = x_1^l,\ x_2 \in]x_2^l, x_2^r[,\ x_3 \in]x_3^l, x_3^r[\big\}\ .$$

The other faces are defined analogously. On each of the six faces we assume that the k-th partial derivative with respect to x_i of the unknown function w is zero. For example, we might assume that, for some $k \in \{0, 1, 2\}$,

$$\partial_{x_1}^{(k)} w(x_1^l, x_2, x_3, t) = 0 \;\text{ on }\; \mathcal{F}_{x_1^l}\times]0,T]\ .$$

Thus, in particular, $k = 0$ corresponds to a Dirichlet condition and $k = 1$ to a Neumann condition. We may also solve the PDE on some of the faces, i.e., we assume that $\partial_t w + \mathcal{B}w = 0$ not only holds in $G\times]0,T]$, but additionally on $\mathcal{F}_{x_i^{l,r}}\times]0,T]$. In such a case, we do not need to specify boundary conditions on the corresponding faces. For the Heston model with (and without) stochastic correlation it is indeed not necessary to specify boundary conditions on $\mathcal{F}_{x_1^l}$ and $\mathcal{F}_{x_2^l}$, see for example [ET10].

Later on, it will become necessary to define grid stretching functions. By this we mean bijective transformations ϕ_i in each coordinate direction as follows (where $\mathbf{z} := (z_1, z_2, z_3)$)

$$\boldsymbol{\phi} : \bigtimes_{i=1}^{3} [z_i^l, z_i^r] \to \bigtimes_{i=1}^{3} [x_i^l, x_i^r], \;\; \mathbf{z} \mapsto \mathbf{x} = \boldsymbol{\phi}(\mathbf{z}) := (\phi_1(z_1), \phi_2(z_2), \phi_3(z_3)) \; .$$

We then define

$$u(\mathbf{z}, t) := w\big(\boldsymbol{\phi}(\mathbf{z}), t\big)$$

and equivalently solve, instead of the PDE (2.7) for w, the transformed PDE

$$\left\{ \begin{array}{rcll} \partial_t u + \widehat{\mathcal{B}} u &=& 0 & \text{in } \; \widehat{G} \times]0, T] \\ u(\mathbf{z}, 0) &=& \widehat{g}(\mathbf{z}) & \text{in } \; \widehat{G} \end{array} \right. , \tag{2.9}$$

for u. Herein, we define $\widehat{g}(\mathbf{z}) := g(\boldsymbol{\phi}(\mathbf{z}))$, $\widehat{G} := \bigtimes_{i=1}^{3}]z_i^l, z_i^r[$, and the differential operator $\widehat{\mathcal{B}}$ is given by

$$\widehat{\mathcal{B}} := \sum_{\substack{i,j=1 \\ j>i}}^{3} \widehat{a}_{i+j+1}(\mathbf{z}) \partial_{z_i z_j} + \sum_{i=1}^{3} \Big[\widehat{a}_i(\mathbf{z}) \partial_{z_i z_i} + \widehat{b}_i(\mathbf{z}) \partial_{z_i} + \frac{1}{3} \widehat{c}(\mathbf{z}) \Big] \tag{2.10}$$

The coefficients of $\widehat{\mathcal{B}}$ follow from the coefficients of $\mathcal{B}$ in (2.8) and are given by

$$\begin{array}{rcl} \widehat{a}_i(\mathbf{z}) &:=& \dfrac{a_i(\boldsymbol{\phi}(\mathbf{z}))}{(\partial_{z_i} \phi_i)^2}, \;\; i = 1, 2, 3 \\ \widehat{a}_{i+j+1}(\mathbf{z}) &:=& \dfrac{a_{i+j+1}(\boldsymbol{\phi}(\mathbf{z}))}{\partial_{z_i} \phi_i \partial_{z_j} \phi_j}, \;\; i = 1, 2, \;\; j = 2, 3 \, ; \; i \neq j \\ \widehat{b}_i(\mathbf{z}) &:=& \dfrac{b_i(\boldsymbol{\phi}(\mathbf{z}))}{\partial_{x_i} \phi_i} - \dfrac{a_i(\boldsymbol{\phi}(\mathbf{z})) \partial_{z_i z_i} \phi_i}{(\partial_{z_i} \phi_i)^3}, \;\; i = 1, 2, 3 \\ \widehat{c}(\mathbf{z}) &:=& c(\boldsymbol{\phi}(\mathbf{z})) \; . \end{array}$$

For $\xi \in [x^l, x^r]$ and $\gamma \in \mathbb{R}^+$, a common used grid stretching function is

$$[0, 1] \ni z \mapsto x = \phi(z) := \xi + \gamma \sinh(\alpha z + \beta(1 - z)) \in [x^l, x^r] \; , \tag{2.11}$$

where $\alpha := \sinh^{-1}((x^r - \xi)/\gamma)$ and $\beta := \sinh^{-1}((x^l - \xi)/\gamma)$. The smaller the grid stretching parameter γ, the more grid points are concentrated around the point ξ. See figure 1 for an example.

In the next section, we develop a finite-difference-method to solve (2.9) numerically under the assumption that the coefficient functions a_i, b_i and c of the operator $\mathcal{B}$ (2.8) factorise with respect to the coordinates x_i. To be more precise, we assume that these functions can be written as products

$$a_i(\mathbf{x}) = a_i^1(x_1) a_i^2(x_2) a_i^3(x_3), \;\; b_i(\mathbf{x}) = b_i^1(x_1) b_i^2(x_2) b_i^3(x_3), \;\; c(\mathbf{x}) = c^1(x_1) c^2(x_2) c^3(x_3)$$

for univariate functions a_i^j, b_i^j and c^j. Since the grid stretching functions ϕ_i act on z_i only, the coefficients $\widehat{a}_i$, $\widehat{b}_i$ and $\widehat{c}$ of the operator $\widehat{\mathcal{B}}$ (2.10) can also be written as (sums of) products. In particular, we have with the decomposition $\widehat{b}_i =: \widehat{b}_{i,1} + \widehat{b}_{i,2}$

$$\begin{array}{rcl}
\widehat{a}_i(\mathbf{z}) & = & \widehat{a}_i^1(z_1)\widehat{a}_i^2(z_2)\widehat{a}_i^3(z_3) \\
\widehat{b}_i(\mathbf{z}) & = & \widehat{b}_{i,1}^1(z_1)\widehat{b}_{i,1}^2(z_2)\widehat{b}_{i,1}^3(z_3) + \widehat{b}_{i,2}^1(z_1)\widehat{b}_{i,2}^2(z_2)\widehat{b}_{i,2}^3(z_3) \\
\widehat{c}(\mathbf{z}) & = & \widehat{c}^1(z_1)\widehat{c}^2(z_2)\widehat{c}^3(z_3)
\end{array}$$

where the factors become

$$\begin{aligned}
&\widehat{a}_j^j(z_j) := \frac{a_j^j(\phi_j(z_j))}{(\partial_{z_j}\phi_j)^2}, \quad j = 1,2,3 \\
&\widehat{a}_i^j(z_j) := a_i^j(\phi_j(z_j)), \quad i,j = 1,2,3, \quad i \neq j \\
&\widehat{b}_{j,1}^j(z_j) := \frac{b_j^j(\phi_j(z_j))}{\partial_{z_j}\phi_j}, \quad \widehat{b}_{j,2}^j(z_j) := -\frac{a_j^j(\phi_j(z_j))\partial_{z_j z_j}\phi_j}{(\partial_{z_j}\phi_j)^3}, \quad j = 1,2,3 \\
&\widehat{b}_{i,1}^j(z_j) := b_i^j(\phi_j(z_j)), \quad \widehat{b}_{i,2}^j(z_j) := a_i^j(\phi_j(z_j)), \quad i,j = 1,2,3, \quad i \neq j \\
&\widehat{c}^j(z_j) := c^j(\phi_j(z_j)), \quad j = 1,2,3 \; .
\end{aligned}$$

3. Finite difference method

We first discretise the PDE (2.9) in space; the resulting linear system of ODEs (with respect to time) is then approximatively solved by a time-marching scheme. The idea of the finite-difference-method is to consider the PDE not for all $\mathbf{z} \in \widehat{G}$ but only for a finite number of (equidistant) grid points $\mathbf{z_i} := (z_{1,i_1}, z_{2,i_2}, z_{3,i_3})$

$$\begin{aligned}
&\mathbf{z_i} \in \mathcal{G} := \\
&\quad \{(z_1^l + i_1 h_1, z_2^l + i_2 h_2, z_3^l + i_3 h_3) \mid i_j = 0, \ldots, N_j + 1, \; j = 1,2,3\}
\end{aligned}$$

where $h_j := \frac{x_j^r - x_j^l}{N_j + 1}$, $N_j \in \mathbb{N}^\times$, and to replace at these grid points the partial derivatives by their corresponding finite difference quotients.

3.1. **Finite difference discretisation with respect to space.**

For $n \in \{2, 4\}$ we introduce the following centred difference-operators in the i-th coordinate direction

$$\begin{array}{rcl}
\delta_{z_i,h_i}^{(1),n} f(\mathbf{z}) & := & h_i^{-1} \displaystyle\sum_{j=-2}^{2} m_j^{(1),n} f(\ldots, z_i + j h_i, \ldots) \; , \\
\delta_{z_i,h_i}^{(2),n} f(\mathbf{z}) & := & h_i^{-2} \displaystyle\sum_{j=-2}^{2} m_j^{(2),n} f(\ldots, z_i + j h_i, \ldots) \; .
\end{array}$$

where the masks $\mathbf{m}^{(k),n} := (m_{-2}^{(k),n}, m_{-1}^{(k),n}, m_0^{(k),n}, m_1^{(k),n}, m_2^{(k),n})$ are given by

$$\begin{aligned}
\mathbf{m}^{(1),2} &:= \frac{1}{2}(0,-1,0,1,0) \\
\mathbf{m}^{(2),2} &:= (0,1,-2,1,0) \\
\mathbf{m}^{(1),4} &:= \frac{1}{12}(1,-8,0,8,-1) \\
\mathbf{m}^{(2),4} &:= \frac{1}{12}(-1,16,-30,16,-1)
\end{aligned}$$

These operators are approximations to the partial derivatives $\partial_{z_i} f =: \partial_{z_i}^{(1)} f$ and $\partial_{z_i z_i} f =: \partial_{z_i}^{(2)} f$ of 2nd or 4th order in the sense that

$$\partial_{z_i}^{(k)} f(\mathbf{z}) = \delta_{z_i,h_i}^{(k),n} f(\mathbf{z}) + \mathcal{O}(h_i^n) \ .$$

With $\delta_{z_i,h_i}^{(k),n}$, we define a discrete version of the differential operator $\widehat{\mathcal{B}}$ as follows

$$\widehat{\mathcal{B}}_{\mathbf{h}}^n := \sum_{\substack{i,j=1 \\ j>i}}^{3} \widehat{a}_{i+j+1}(\mathbf{z_i}) \delta_{z_i,h_i}^{(1),n} \delta_{z_j,h_j}^{(2),n} + \sum_{i=1}^{3} \left[\widehat{a}_i(\mathbf{z_i}) \delta_{z_i,h_i}^{(2),n} + \widehat{b}_i(\mathbf{z_i}) \delta_{z_1,h_1}^{(1),n} + \frac{1}{3}\widehat{c}(\mathbf{z_i}) \right]$$

and consider, for each $\mathbf{z_i} \in \mathcal{G}$, the approximation

$$\partial_t u(\mathbf{z_i},t) + \widehat{\mathcal{B}}_{\mathbf{h}}^n u(\mathbf{z_i},t) = \sum_{i=1}^{3} \mathcal{O}(h_i^n)$$

to the PDE (2.9). By replacing $u(\mathbf{z_i},t)$ with $u_\ell(t)$ for some mapping

$$\{1,\dots,N_1\} \times \{1,\dots,N_2\} \times \{1,\dots,N_3\} \ni \mathbf{i} := (i_1,i_2,i_3) \mapsto \ell := \ell(\mathbf{i}) \in \{1,\dots,N\} \ ,$$

the above is equivalent to the linear system of $N := N_1 N_2 N_3$ ordinary differential equations

$$\left\{ \begin{aligned} \mathbf{u}'(t) + \mathbf{A}\mathbf{u}(t) &= \mathbf{0} \\ \mathbf{u}(0) &= \mathbf{g} \end{aligned} \right. \ , \tag{3.1}$$

where we denote by $\mathbf{u}(t) = (u_1(t),\dots,u_N(t))^\top$ the vector of unknowns and by $\mathbf{u}'(t) = (\partial_t u_1(t),\dots,\partial_t u_N(t))^\top$ its derivative with respect to time. The components of the initial condition $\mathbf{u}(0)$ are given by $g(\mathbf{z_i})$. If we set Dirichlet conditions on all of $\partial\widehat{G}$, then the $N \times N$-matrix $\mathbf{A}$ in (3.1) is a sum of Kronecker products

$$\begin{aligned}
\mathbf{A} \;:=\; & \mathbf{M}_{\widehat{a}_1^3}^{(0)} \otimes \mathbf{M}_{\widehat{a}_1^2}^{(0)} \otimes \mathbf{M}_{\widehat{a}_1^1}^{(2),n} + \mathbf{M}_{\widehat{a}_2^3}^{(0)} \otimes \mathbf{M}_{\widehat{a}_2^2}^{(2),n} \otimes \mathbf{M}_{\widehat{a}_2^1}^{(0)} + \mathbf{M}_{\widehat{a}_3^3}^{(2),n} \otimes \mathbf{M}_{\widehat{a}_3^2}^{(0)} \otimes \mathbf{M}_{\widehat{a}_3^1}^{(0)} \\
& + \mathbf{M}_{\widehat{a}_4^3}^{(0),n} \otimes \mathbf{M}_{\widehat{a}_4^2}^{(1),n} \otimes \mathbf{M}_{\widehat{a}_4^1}^{(1),n} + \mathbf{M}_{\widehat{a}_5^3}^{(1),n} \otimes \mathbf{M}_{\widehat{a}_5^2}^{(0)} \otimes \mathbf{M}_{\widehat{a}_5^1}^{(1),n} + \mathbf{M}_{\widehat{a}_6^3}^{(1),n} \otimes \mathbf{M}_{\widehat{a}_6^2}^{(1),n} \otimes \mathbf{M}_{\widehat{a}_6^1}^{(0)} \\
& + \mathbf{M}_{\widehat{b}_{1,1}^3}^{(0)} \otimes \mathbf{M}_{\widehat{b}_{1,1}^2}^{(0)} \otimes \mathbf{M}_{\widehat{b}_{1,1}^1}^{(1),n} + \mathbf{M}_{\widehat{b}_{2,1}^3}^{(0)} \otimes \mathbf{M}_{\widehat{b}_{2,1}^2}^{(1),n} \otimes \mathbf{M}_{\widehat{b}_{2,1}^1}^{(0)} + \mathbf{M}_{\widehat{b}_{3,1}^3}^{(1),n} \otimes \mathbf{M}_{\widehat{b}_{3,1}^2}^{(0)} \otimes \mathbf{M}_{\widehat{b}_{3,1}^1}^{(0)} \\
& + \mathbf{M}_{\widehat{b}_{1,2}^3}^{(0)} \otimes \mathbf{M}_{\widehat{b}_{1,2}^2}^{(0)} \otimes \mathbf{M}_{\widehat{b}_{1,2}^1}^{(1),n} + \mathbf{M}_{\widehat{b}_{2,2}^3}^{(0)} \otimes \mathbf{M}_{\widehat{b}_{2,2}^2}^{(1),n} \otimes \mathbf{M}_{\widehat{b}_{2,2}^1}^{(0)} + \mathbf{M}_{\widehat{b}_{3,2}^3}^{(1),n} \otimes \mathbf{M}_{\widehat{b}_{3,2}^2}^{(0)} \otimes \mathbf{M}_{\widehat{b}_{3,2}^1}^{(0)} \\
& + \mathbf{M}_{\widehat{c}^3}^{(0)} \otimes \mathbf{M}_{\widehat{c}^2}^{(0)} \otimes \mathbf{M}_{\widehat{c}^1}^{(0)} \ .
\end{aligned}$$

Note that the three sums involving the matrices $\mathbf{M}^{(1),n}_{\widetilde{b}^j_{j,2}}$, $j = 1, 2, 3$, are zero if we do not apply grid stretching, since these matrices are zero itself ($\partial_{z_j z_j}\phi_j$ is zero in this case). For $n = 2$ and a function f the matrices $\mathbf{M}^{(0)}_f$ and $\mathbf{M}^{(k),n}_f$, $k \in \{1, 2\}$ are defined in [Hil19]. For $n = 4$, these matrices are as follows.

Definition 3.1. Let $f : [z^l, z^r] \to \mathbb{R}$ be a continuous function. For $N \in \mathbb{N}^\times$ define $h := \frac{z^r - z^l}{N+1}$ and let $f_j := f(z^l + jh)$, $j = 1, \ldots, N$. The $N \times N$-matrices $\mathbf{M}^{(k),4}_f$, $k = 1, 2$ corresponding to the operators $f(x)\partial^{(k)}_x = f(x)\delta^{(k),4}_{x,h} + \mathcal{O}(h^4)$ are defined as

$$\mathbf{M}^{(1),4}_f := \frac{1}{12h}\begin{pmatrix} -10f_1 & 18f_1 & -6f_1 & f_1 & & & \\ -8f_2 & 0 & 8f_2 & -f_2 & & & \\ f_3 & -8f_3 & 0 & 8f_3 & -f_3 & & \\ & & & \ddots & & & \\ & & f_{N-2} & -8f_{N-2} & 0 & 8f_{N-2} & -f_{N-2} \\ & & & f_{N-1} & -8f_{N-1} & 0 & 8f_{N-1} \\ & & & -f_N & 6f_N & -18f_N & 10f_N \end{pmatrix}$$

$$\mathbf{M}^{(2),4}_f := \frac{1}{12h^2}\begin{pmatrix} -15f_1 & -4f_1 & 14f_1 & -6f_1 & f_1 & & \\ 16f_2 & -30f_2 & 16f_2 & -f_2 & & & \\ -f_3 & 16f_3 & -30f_3 & 16f_3 & -f_3 & & \\ & & & \ddots & & & \\ & & -f_{N-2} & 16f_{N-2} & -30f_{N-2} & 16f_{N-2} & -f_{N-2} \\ & & & -f_{N-1} & 16f_{N-1} & -30f_{N-1} & 16f_{N-1} \\ & & f_N & -6f_N & 14f_N & -4f_N & -15f_N \end{pmatrix}$$

For non-Dirichlet boundary conditions, we have to use non-centred finite difference quotients and the matrices $\mathbf{M}^{(k),n}_f$ have to be changed to matrices ${}^r_l\mathbf{M}^{(k),n}_f$. Here, the lower prescript $l \in \{n, s, i\}$ indicates the typ of the boundary condition on the left boundary of the corresponding interval and the upper prescript $r \in \{n, s, i\}$ specifies the condition on the right boundary of the interval. The letter n stands for a Neumann condition, the letter s means that the second derivative is specified, and i indicates that we solve the PDE also on the boundary (no boundary condition is needed in this case). Again, for the second order finite difference discretisation $n = 2$, the matrices ${}^r_l\mathbf{M}^{(k),2}_f$ are defined in [Hil19], whereas for the fourth-order difference quotients $n = 4$, we specify the matrices ${}^r_l\mathbf{M}^{(k),4}_f$ in table 2 and table 3. Herein, we only need to specify the first two and the last two rows, since the remaining rows do not change and are therefore equal to rows of the corresponding Dirichlet matrices $\mathbf{M}^{(k),4}_f$.

If we solve the PDE also on the boundary ($r = l = i$) of the interval $[z^l_j, z^r_j]$, then the matrices ${}^i_i\mathbf{M}^{(0)}_f$ and ${}^i_i\mathbf{M}^{(k),4}_f$ have dimension $(N+2) \times (N+2)$. In table 4 we find the corresponding first two and last two rows, which follow again by an application of non-centred finite difference quotients. Note that we may combine the different

matrix	entries of	row ℓ	Neumann, $j = n$
${}^{n}_{n}\mathbf{M}_f^{(1),4}$	first five	1	$\frac{f_1}{12h}(-\frac{394}{25}, \frac{558}{25}, \frac{-198}{25}, \frac{34}{25}, 0)$
	first five	2	$\frac{f_2}{12h}(-3, -10, 18, -6, 1)$
	last five	$N-1$	$\frac{f_{N-1}}{12h}(-1, 6, -18, 10, 3)$
	last five	N	$\frac{f_N}{12h}(-\frac{34}{25}, \frac{198}{25}, -\frac{558}{25}, \frac{394}{25})$
${}^{n}_{n}\mathbf{M}_f^{(2),4}$	frist six	1	$\frac{f_1}{12h^2}(\frac{21}{5}, -\frac{92}{5}, \frac{102}{5}, -\frac{36}{5}, 1, 0)$
	first six	2	$\frac{f_2}{12h^2}(10, -15, -4, 14, -6, 1)$
	last six	$N-1$	$\frac{f_{N-1}}{12h^2}(1, -6, 14, -4, -15, 10)$
	last six	N	$\frac{f_N}{12h^2}(0, 1, -\frac{36}{5}, \frac{102}{5}, -\frac{92}{5}, \frac{21}{5})$

Table 2. Neumann boundary condition. The rows $\ell = 1, 2, N-1, N$ of the matrices ${}^{n}_{n}\mathbf{M}_f^{(2),4}$. The remaining rows $\ell = 3, \ldots, N-2$ are equal to the rows of $\mathbf{M}_f^{(k),4}$ in definition 3.1.

matrix	entries of	row ℓ	second derivative, $j = s$
${}^{s}_{s}\mathbf{M}_f^{(1),4}$	first five	1	$\frac{f_1}{12h}(-\frac{304}{15}, \frac{484}{15}, -\frac{82}{5}, \frac{76}{15}, -\frac{2}{3})$
	first five	2	$\frac{f_2}{12h}(-3, -10, 18, -6, 1)$
	last five	$N-1$	$\frac{f_{N-1}}{12h}(-1, 6, -18, 10, 3)$
	last five	N	$\frac{f_N}{12h}(\frac{2}{3}, -\frac{76}{15}, \frac{82}{5}, -\frac{484}{15}, \frac{304}{15})$
${}^{s}_{s}\mathbf{M}_f^{(2),4}$	first six	1	$\frac{f_1}{12h^2}(\frac{173}{9}, -\frac{464}{9}, \frac{146}{3}, -\frac{176}{9}, \frac{29}{9}, 0)$
	first six	2	$\frac{f_2}{12h^2}(10, -15, -4, 14, -6, 1)$
	last six	$N-1$	$\frac{f_{N-1}}{12h^2}(1, -6, 14, -4, -15, 10)$
	last six	N	$\frac{f_N}{12h^2}(0, \frac{29}{9}, -\frac{176}{9}, \frac{146}{3}, -\frac{464}{4}, \frac{173}{9})$

Table 3. Second derivative boundary condition. The rows $\ell = 1, 2, N-1, N$ of the matrices ${}^{s}_{s}\mathbf{M}_f^{(2),4}$. The remaining rows $\ell = 3, \ldots, N-2$ are equal to the rows of $\mathbf{M}_f^{(k),4}$ in definition 3.1.

boundary conditions arbitrarily. For example, we may impose no condition on the left boundary and a Neumann condition on the right boundary of the interval $[z_j^l, z_j^r[$. The resulting matrices $\mathbf{M}_f^{(0)}$ and ${}^{n}_{i}\mathbf{M}_f^{(k),4}$ have then dimension $(N+1) \times (N+1)$.

matrix	entries of	row ℓ	no condition $j = i$
${}^i_i\mathbf{M}^{(1),4}_f$	first five	0	$\frac{f_0}{12h}(-25, 48, -36, 16, -3)$
	first five	1	$\frac{f_1}{12h}(-3, -10, 18, -6, 1)$
	last five	N	$\frac{f_N}{12h}(-1, 6, -18, 10, 3)$
	last five	$N+1$	$\frac{f_{N+1}}{12h}(3, -16, 36, -48, 25)$
${}^i_i\mathbf{M}^{(2),4}_f$	frist six	0	$\frac{f_0}{12h^2}(45, -154, 214, -156, 61, -10)$
	first six	1	$\frac{f_1}{12h^2}(10, -15, -4, 14, -6, 1)$
	last six	N	$\frac{f_N}{12h^2}(1, -6, 14, -4, -15, 10)$
	last six	$N+1$	$\frac{f_{N+1}}{12h^2}(-10, 61, -156, 214, -154, 45)$

Table 4. No boundary condition. The rows $\ell = 0, 1, N, N+1$ of the matrices ${}^i_i\mathbf{M}^{(2),4}_f$. The remaining rows $\ell = 2, \ldots, N-1$ are equal to the rows of $\mathbf{M}^{(k),4}_f$ in definition 3.1..

We leave it to the reader to specify the entries of these matrices by combining the definition 3.1 with tables 2, 3 and 4.

3.2. **Time stepping.** Let $M \in \mathbb{N}^\times$ be the number of time steps, let $k := T/M$ be the step size and let be n the number of rows and columns of the matrix $\mathbf{A}$ in (3.1). A time stepping scheme is a mapping $\Phi : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{u}_0 \mapsto \mathbf{u}_M = \Phi(\mathbf{u}_0)$ that approximates the solution $\mathbf{u}(T) = e^{-\mathbf{A}T}\mathbf{g}$ of the system of ODEs (3.1) by generating a sequence of vectors $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M)$; starting in $\mathbf{u}_0 := \mathbf{g}$. Each $\mathbf{u}_j$ is an approximation to $\mathbf{u}(t_j)$ at time points $t_j := jk$, $j = 1, \ldots, M$. The ℓ-th component of $\mathbf{u}_j$ is an approximation to the price V in (2.3) of the derivative

$$u(\mathbf{z_i}, t_j) = w(\boldsymbol{\phi}(\mathbf{z_i}), t_j) \approx V(\mathbf{x_i}, T - t_j)\ .$$

We describe two time-marching schemes Φ_k in the following.

3.2.1. *The Rannacher scheme.* The Rannacher scheme [Ran84] is a simple modification of the Crank-Nicolson scheme to deal with an un-smooth initial condition g as follows. It first applies $R > 0$ (R even) implicit Euler steps of step size $k/2$ followed by $M - R/2$ Crank-Nicolson steps of step size k. Thus, the Rannacher scheme generates the sequence $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M)$ as follows: For $\mathbf{u}_0 = \mathbf{g}$ do

$$\left\{ \begin{array}{rcll} \big(\mathbf{I} + k/2\mathbf{A}\big)\mathbf{u}_{(j+1)/2} & = & \mathbf{u}_{j/2}, & j = 0, 1, \ldots, R-1 \\ \big(\mathbf{I} + k/2\mathbf{A}\big)\mathbf{u}_{j+1} & = & \big(\mathbf{I} - k/2\mathbf{A}\big)\mathbf{u}_j, & j = R/2, \ldots, M-1 \end{array} \right. \tag{3.2}$$

Herein, we denote by $\mathbf{u}_{j/2}$ the approximation to $\mathbf{u}(t_{j/2})$, with $t_{j/2} = jk/2$. Typically, $R = 2$ initial implicit Euler steps are sufficient. Define the two mappings

$$\Phi^1_{k/2} : \mathbb{R}^n \to \mathbb{R}^n, \;\; \mathbf{u} \mapsto \Phi^1_{k/2}(\mathbf{u}) := (\mathbf{I} + k/2\mathbf{A})^{-1}\mathbf{u}$$
$$\Phi^2_{k} : \mathbb{R}^n \to \mathbb{R}^n, \;\; \mathbf{u} \mapsto \Phi^2_{k}(\mathbf{u}) := (\mathbf{I} + k/2\mathbf{A})^{-1}(\mathbf{I} - k/2\mathbf{A})\mathbf{u}$$

Then, the Rannacher scheme can be written as $\mathbf{u}_M = \Phi_R(\mathbf{u}_0)$ with

$$\Phi_R := \underbrace{\Phi^2_k \circ \cdots \circ \Phi^2_k}_{M-R/2 \text{ times}} \circ \underbrace{\Phi^1_{k/2} \circ \cdots \circ \Phi^1_{k/2}}_{R \text{ times}} .$$

The Rannacher scheme converges at best quadratically, depending on the smoothness of g, $\|\mathbf{u}_M - \mathbf{u}(T)\|_\infty \leq Ck^2$. Unfortunately, the Rannacher scheme is (very) slow, since it involves solving M sparse, but large-banded linear systems. We therefore apply an Alternating Direction Implicit (ADI) scheme, which avoids solving linear systems with large bandwidths.

3.2.2. *The Hundsdorfer-Verwer scheme.* The idea of any ADI time-stepping scheme is to split the differential operator $\widehat{\mathcal{B}}$ in (2.10) into $\widehat{\mathcal{B}} = \mathcal{B}_0 + \mathcal{B}_1 + \mathcal{B}_2 + \mathcal{B}_3$ with

$$\begin{aligned} \mathcal{B}_0 &:= \sum_{\substack{i,j=1 \\ j>i}}^{3} \widehat{a}_{i+j+1}(\mathbf{z})\partial_{z_i z_j} \\ \mathcal{B}_i &:= \widehat{a}_i(\mathbf{z})\partial_{z_i z_i} + \widehat{b}_i(\mathbf{z})\partial_{z_i} + \frac{1}{3}\widehat{c}(\mathbf{z}) \;, \end{aligned}$$

compare with (2.10), and to apply the same to the finite difference operator $\widehat{B}^n_{\mathbf{h}}$, see e.g. [HV03] and the references therein. As consequence, the matrix $\mathbf{A}$ inherits the splitting and can thus be written as $\mathbf{A} = \mathbf{A}_0 + \mathbf{A}_1 + \mathbf{A}_2 + \mathbf{A}_3$. The ADI scheme then treats the matrix $\mathbf{A}_0$ with large bandwidths (corresponding to all mixed derivatives in $\widehat{\mathcal{B}}$) explicit. The matrices $\mathbf{A}_i$ are (or can be made by suitable permutations) small banded. For example, if we use Dirichlet boundary conditions on all faces of the cube $\widehat{G}$, these matrices have bandwidth $n + 1$, where $n \in \{2, 4\}$ is the order of the finite difference quotients used.

The ADI scheme proposed by Hundsdorfer and Verwer (HV) generates the sequence $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M)$ as follows: For $\mathbf{u}_0 = \mathbf{g}$, $\theta \in]0, 1]$ and $j = 0, \ldots, M - 1$ do

$$\left\{ \begin{array}{rcll} \mathbf{y}_0 &=& \big(\mathbf{I} - k\mathbf{A}\big)\mathbf{u}_j & \\ \big(\mathbf{I} + k\theta\mathbf{A}_i\big)\mathbf{y}_i &=& \mathbf{y}_{i-1} + k\theta\mathbf{A}_i\mathbf{u}_j, & i = 1, 2, 3 \\ \mathbf{z}_0 &=& \mathbf{y}_0 - \frac{1}{2}k\mathbf{A}(\mathbf{y}_3 - \mathbf{u}_j) & \\ \big(\mathbf{I} + k\theta\mathbf{A}_i\big)\mathbf{z}_i &=& \mathbf{z}_{i-1} + k\theta\mathbf{A}_i\mathbf{y}_3, & i = 1, 2, 3 \\ \mathbf{u}_{j+1} &=& \mathbf{z}_3 & \end{array} \right. \tag{3.3}$$

In the HV scheme, denote by Φ_k the function which maps $\mathbf{u}_j$ to $\mathbf{u}_{j+1}$, i.e., $\mathbf{u}_{j+1} = \Phi_k(\mathbf{u}_j)$. We call Φ_k a HV-step with step size k. Then, the HV scheme can be written as $\mathbf{u}_M = \Phi_{HV}(\mathbf{u}_0)$, with

$$\Phi_{HV} := \underbrace{\Phi_k \circ \cdots \circ \Phi_k}_{M \ \text{times}} .$$

Since one HV-step requires solving six linear systems, we need to solve in total $6M$ linear, small banded systems. As the Rannacher scheme, the HV scheme converges at best quadratically [I'HW15] and thus the maximal nodal error of the fully discrete scheme is (at best)

$$\max_{\mathbf{z_i} \in \mathcal{G}} |u_{\ell,M} - u(\mathbf{z_i}, T)| \leq C(h^n + k^2), \quad n \in \{2, 4\}$$

where $u_{\ell,M}$ is the ℓ-th component of the vector $\mathbf{u}_M$ (ℓ is corresponding to the grid point $\mathbf{z_i}$) and where $h = \max\{h_1, h_2, h_3\}$. Hence, if we apply a fourth order finite difference discretisation in space, the error of the time discretisation will dominate, unless with choose $k = \mathcal{O}(h^2)$, which however leads to an inefficient scheme. To avoid this, we combine the HV scheme with a Richardson extrapolation. Let $\widehat{\mathbf{u}}_{j+1} := \Phi_k(\mathbf{u}(t_j))$ and $\widetilde{\mathbf{u}}_{j+1} = (\Phi_{k/2} \circ \Phi_{k/2})(\mathbf{u}(t_j))$ be two approximations to $\mathbf{u}(t_{j+1})$. The approximation $\widehat{\mathbf{u}}_{j+1}$ is obtained by applying one HV step with step size k to $\mathbf{u}(t_j)$, whereas the approximation $\widehat{\mathbf{u}}_{j+1}$ is obtained by applying two HV steps with step size $k/2$ to $\mathbf{u}(t_j)$. Since the HV scheme is of second order, we have for some constant c

$$\begin{aligned} \mathbf{u}(t_{j+1}) &= \widehat{\mathbf{u}}_{j+1} + ck^2 + \mathcal{O}(k^3) \\ \mathbf{u}(t_{j+1}) &= \widetilde{\mathbf{u}}_{j+1} + c\Big(\frac{k}{2}\Big)^2 + \mathcal{O}(k^3) \end{aligned}$$

Multiplying the second equation by 4 and subtracting from the result the first equations yields $3\mathbf{u}(t_j) = 4\widetilde{\mathbf{u}}_{j+1} - \widehat{\mathbf{u}}_{j+1} + \mathcal{O}(k^3)$. Whence, the linear combination

$$\bar{\mathbf{u}}_{j+1} = \frac{4}{3}\widetilde{\mathbf{u}}_{j+1} - \frac{1}{3}\widehat{\mathbf{u}}_{j+1} \tag{3.4}$$

is a third-order approximation to $\mathbf{u}(t_j)$, i.e, $\mathbf{u}(t_{j+1}) = \bar{\mathbf{u}}_{j+1} + \mathcal{O}(k^3)$. The price we pay for the increased order by one is that we have to solve in total $18M$ linear systems (in contrast to $6M$ systems to be solved in the HV scheme without Richardson extrapolation). We now could iterate this procedure yielding

$$\mathbf{u}(t_{j+1}) = \frac{8}{7}\widetilde{\bar{\mathbf{u}}}_{j+1} - \frac{1}{7}\bar{\mathbf{u}}_{j+1} + \mathcal{O}(k^4)$$

where $\widetilde{\bar{\mathbf{u}}}_{j+1}$ denotes the Richardson extrapolate with step size $k/2$. To obtain this approximation, we need to solve in total $42M$ linear systems.

4. Numerical experiments

We consider a European call option with $K \in \{40, 100, 160\}$ and maturity $T = 5$ written on a stock with price $s_0 = 100$ that pays no dividend $q = 0$. The risk free is $r = 0$. Furthermore, we assume the initial variance v_0 and initial correlation z_0 to be $v_0 = 0.02$ and $z_0 = -0.4$, respectively. The correlation process is chosen to be a

Jacobi process (compare with table 1), and we let the remaining model parameters be

$$(\kappa, m, \delta, \kappa_Z, m_Z, \delta_Z, \rho_1, \rho_2) = (2.1, 0.03, 0.2, 3.5, -0.55, 0.18, \rho_1, 0)\ .$$

Here, $\rho_1 \in \{-0.2, 0, 0.2\}$ is the correlation of the Brownian motions that drive the stock price process and the correlation process, respectively, compare with (2.2). To obtain the price $V(\mathbf{x}_0, 0)$ of the option, we solve the PDE (2.9) for five different discretisation settings which we summarise in table 5 below. In any setting, the PDE is solved subject to the following boundary conditions. On the faces $\mathcal{F}_{z_i^l}$, $i = 1, 2$, (corresponding to $\{s = 0\}$ and $\{v = 0\}$) no boundary conditions are specified and whence the PDE is solved also on these. On the remaining four faces we set the second derivate of u to zero. For $n \in \{2, 4\}$, the nth-order finite difference discretisation with respect to space (compare with section 3) is performed with $N_i = 2^{L_i} - 1$ inner grid points with respect to the i-th coordinate direction. Thus, for the used boundary conditions, the matrix $\mathbf{A}$ in (3.2) and/or (3.3) has dimension $N \times N$, with $N := (N_1 + 1)(N_2 + 1)N_3 = 2^{L_1+L_2}(2^{L_3} - 1)$. We collect the L_i's in the vector $\mathbf{L} := (L_1, L_2, L_3)$. The number M of time steps we choose in the Rannacher (3.2) or the HV scheme with parameter $\theta \in]0, 1]$ in (3.3) is set to $M := \lceil \eta \max\{N_1, N_2, N_3\} \rceil$ for some $\eta > 0$. In the column “ts” (time stepping) of table 5, “R” stands for the Rannacher scheme and “HV” for Hundsdorfer-Verwer scheme. We realise a grid stretching as follows. We let $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3)$ with ϕ_i, $i = 1, 2$ as in (2.11) where $\xi_1 = K$ is the strike of the option and $\xi_2 = 0$ (a concentration around zero variance). Furthermore, we let the grid stretching parameters (somewhat arbitrary) be $\gamma_1 = 5 \cdot 10^{-1}$ and $\gamma_2 = 10^{-2}$. In any case, we do not stretch the grid with respect to the correlation coordinate, whence $x_3 = \phi_3(z_3) = z_3$. If do not stretch the grid at all, then $x_i = \phi(z_i) = z_i$, $i = 1, 2, 3$. This case is indicated in the column “gs ” (grid stretching) of table 5 as “no”. The domain G of the PDE (2.7) is chosen to be $G =]0, 4s_0[\times]0, 0.5[\times] - 1, 1[$, whence the domain $\widehat{G}$ of the transformed PDE (2.9) becomes $\widehat{G} =]0, 1[\times]0, 1[\times] - 1, 1[$ (if “gs=yes”). We might apply a Richardson extrapolation (3.4) to increase the accuracy of the time stepping scheme. This is indicated in column “Re” (Richardson extrapolation).

Setting	n	$\mathbf{L}$	gs	ts	θ	η	RE	Speed up
I	2	$(6, 5, 5)$	no	R		0.1	no	1
II	2	$(6, 5, 5)$	no	HV	1	0.5	no	12.3
III	4	$(6, 5, 5)$	yes	R		0.1	no	0.26
IV	4	$(6, 5, 5)$	yes	HV	1	0.1	yes	7.2
V	4	$(5, 4, 3)$	yes	HV	1	0.2	yes	214

Table 5. The discretisation settings.

As the setting I (second-order finite difference discretisation in space, Rannacher time stepping) is the standard method, we declare it to be the benchmark. To compare it with the other settings (with respect to CPU time), we calculate for every setting $j \in \{\text{I,II,III,IV,V}\}$ the ratio $t^{\text{I}}_{\text{CPU}}/t^{j}_{\text{CPU}}$ (we average over different strikes and correlations, compare with table 6). These ratios can be found in the last column "Speed up" of table 5. A value larger (smaller) than 1 means that the considered setting is faster (slower) than the benchmark method I. All computations are performed in Matlab (R2014b) on a MacBook Pro equipped with a 2 GHz Intel Core i5 processor and 8 GB of RAM. Note that setting V is by far the fastest one. In particular, it delivers the option price in about 0.3 seconds.

In table 6, we list the implied volatilities generated by the calculated option prices (by inverting the Black-Scholes formula) for strikes $K \in \{40, 100, 160\}$ and correlations $\rho_1 \in \{-0.2, 0, 0.2\}$ and compare them with the implied volatilities from [TEG16]. Here, the authors obtain the values in the row "MC" by a Monte Carlo simulation (described in [TEG17]) based on generating 10^5 paths (each with $20T$ discretisation steps) of the process $\mathbf{X}(t)$ defined in (2.1)-(2.2). In the second row "FTM" of table 6, we additionally state the implied volatilities from the same authors [TEG16], which they obtain by applying a Fourier transform method to an approximation of the characteristic function $\mathbb{E}[e^{i\mathbf{u}^\top \mathbf{X}(t)}]$.

	$K = 40$			$K = 100$			$K = 160$		
Method	-0.2	0	0.2	-0.2	0	0.2	-0.2	0	0.2
MC	19.27	19.25	19.33	16.75	16.71	16.79	15.16	15.41	15.46
FTM	19.02	19.03	19.04	16.84	16.83	16.82	15.35	15.36	15.36
I	19.57	19.58	19.59	16.65	16.64	16.64	15.56	15.57	15.58
II	19.53	19.54	19.55	16.66	16.65	16.65	15.57	15.59	15.60
III	19.33	19.34	19.35	16.70	16.70	16.69	15.55	15.56	15.57
IV	19.27	19.28	19.29	16.72	16.71	16.70	15.57	15.58	15.60
V	19.33	19.34	19.36	16.75	16.75	16.74	15.62	15.63	15.64

Table 6. Implied volatilities obtained from different PDE solvers in comparison to a Monte Carlo and a Fourier transform method. The largest deviations of the PDE values from the MC values can be observed for options far out of the money. Still, these values are in the interval $[\text{MC} - \text{Std}, \text{MC} + \text{Std}]$ with a standard deviation of the MC method being $\text{Std} = 0.30$, see [TEG16].

We observe that the discretisation setting IV yields the smallest deviations to the MC values, whereas setting V delivers slightly larger differences, but within a CPU time which is 30 times lower.

References

[Bat00] D. S. Bates, *Post-'87 crash fears in the S&P 500 futures option market*, Journal of Econometrics **94** (2000), 181–238.

[Ber16] L. Bergomi, *Stochastic Volatility Modeling*, Financial Mathematics Series, Chapman&Hall/CRC, 2016.

[Bos14] S. Bossu, *Advanced Equity Derivatives*, Wiley Finance Series, Wiley, 2014.

[DS06] F. Delbaen and W. Schachermayer, *The Mathematics of Arbitrage*, Springer, 2006.

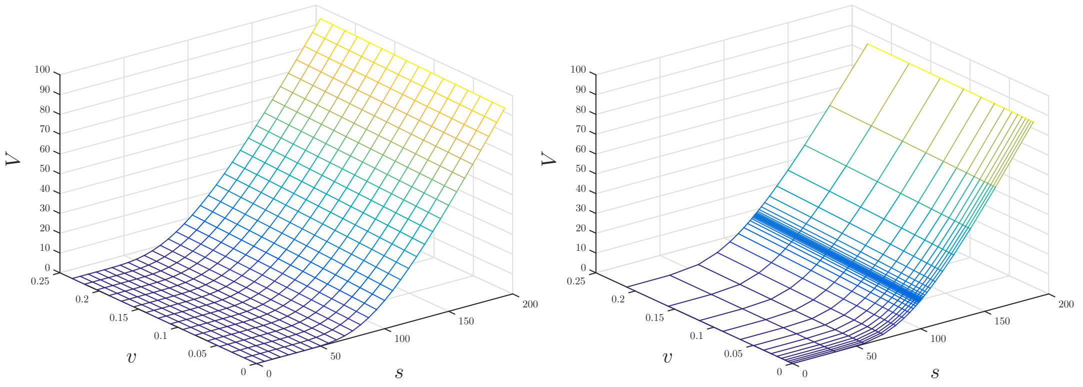

Figure 1. The graph of the option value $(s, v) \mapsto V(s, v, z_0, 0)$ for $z_0 = -0.4$ and $(s, v) \in [0, 200[\times[0, 0.2[$ of a European call option with strike $K = 100$ and $\rho_1 = 0$ in the Heston model with stochastic correlation. Left. The graph is obtained with the discretisation setting I. Right. The graph is obtained with the discretisation setting V. The grid stretching is clearly visible.

[ET10] E. Ekström and J. Tysk, *The black-scholes equation in stochastic volatility models*, Journal of Mathematical Analysis and Applications **368** (2010), no. 2, 498–507.

[Hes93] S. Heston, *A closed-form Solution for Options with Stochastic Volatility, with Applications to Bond and Currency Options*, The Review of Financial Studies **6** (1993), 327–343.

[Hil19] N. Hilber, *26 lines of code to price single factor derivates*, Working Paper (2019), 1–29.

[HS00] D. Heath and M. Schweizer, *Martingales versus PDEs in Finance: An Equivalence Result with Examples*, Journal of Applied Probability **37** (2000), no. 4, 947–957.

[HV03] W. Hundsdorfer and J.G. Verwer, *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*, Springer Series in Computational Mathematics, Springer, 2003.

[I'HW15] K.J. In 't Hout and M Wyns, *Convergence of the Hundsdorfer-Verwer scheme for two-dimensional convection-diffusion equations with mixed derivative term*, AIP Conference Proceedings **1648** (2015), no. 1, 850054, available at `https://aip.scitation.org/doi/pdf/10.1063/1.4913109`.

[Ran84] R. Rannacher, *Finite element solution of diffusion problems with irregular data*, Numerische Mathematik **43** (1984), no. 2, 309–327.

[Rou13] F. D. Rouah, *The Heston Model and Its Extensions in Matlab and C#*, Wiley finance series, Wiley, 2013.

[TEG16] L. Teng, M. Ehrhardt, and M. Günther, *On the Heston model with Stochastic Correlation*, International Journal of Theoretical and Applied Finance **19** (2016), no. 6.

[TEG17] ______, *Numerical Simulation of the Heston Model under Stochastic Correlation*, International Journal of Financial Studies **6** (2017), no. 1, 1–16.

[Tsa10] R. S. Tsay, *Analysis of financial time series*, 3rd ed., Wiley series in probability and statistics, Wiley, 2010.

School of Management and Law, ZHAW, CH-8400 Winterthur, Switzerland

Zurich University
of Applied Sciences

School of
Management and Law

St.-Georgen-Platz 2
P.O. Box
8401 Winterthur
Switzerland

www.zhaw.ch/sml