Accepted Manuscript

High-order ADI finite difference schemes for parabolic equations in the combination technique with application in finance

Christian Hendricks, Christof Heuer, Matthias Ehrhardt, Michael Günther

PII: S0377-0427(16)30413-7
DOI: http://dx.doi.org/10.1016/j.cam.2016.08.044
Reference: CAM 10783

To appear in: *Journal of Computational and Applied Mathematics*

Received date: 14 December 2015
Revised date: 24 June 2016

Please cite this article as: C. Hendricks, C. Heuer, M. Ehrhardt, M. Günther, High-order ADI finite difference schemes for parabolic equations in the combination technique with application in finance, *Journal of Computational and Applied Mathematics* (2016), http://dx.doi.org/10.1016/j.cam.2016.08.044

High-order ADI finite difference schemes for parabolic equations in the combination technique with application in finance

Christian Hendricks*, Christof Heuer**, Matthias Ehrhardt* and Michael Günther*

Abstract

In this article we combine high-order (HO) finite difference discretisations with alternating direction implicit (ADI) schemes for parabolic partial differential equations with mixed derivatives in a sparse grid setting. In each implicit leg of the ADI schemes, we propose a high-order-compact (HOC) discretisation, such that only tridiagonal systems have to be solved. With the help of HO spatial discretisations and ADI schemes solutions with second order accuracy in time and fourth order accuracy in space can be computed. In order to reduce the number of involved grid points we use the combination technique to construct the so called sparse grid solution. The theoretical findings are illustrated by numerical examples with European basket options.

1. Introduction

High dimensional parabolic partial differential equations (PDEs) arise in many fields of science, for example in computational fluid dynamics or in computational finance for pricing derivatives, e.g., which are driven by a basket of underlying assets. The exponentially growing number of grid points in a tensor based grid makes it computationally demanding to solve problems in a high dimensional setting. This growing complexity leads to unreasonable long run-times and an excessive memory consumption. Even for a moderate number of spatial dimensions the so called *curse of dimensionality* shows its effects very clearly. In this article we combine three numerical methods in order to solve the equation numerically and to cope with the curse of dimensionality:

High-order-compact (HOC) finite difference schemes exploit the structure of the governing partial differential equation to achieve fourth order accuracy in space on the compact stencil. Starting with the early work by Gupta et al. [2] further effort has been spent on the derivation of HOC schemes, e.g. in [7, 8, 10, 16, 24, 30, 38, 39] to mention a few examples. In the field of computational finance HOC methods have been proposed by [7, 8, 10]. Compared to central standard high-order schemes relying on broad stencils, the compactness of HOC schemes significantly reduces the computational effort, while having the same order of consistency. However, in the high dimensional case the sparsity is deteriorated. With operator splitting techniques, e.g. Locally-One-Dimensional (LOD) or Alternating-Direction-Implicit (ADI) schemes, the discretisation matrix can be decomposed into tridiagonal systems, which can be solved sequentially in linear run-time. The classical ADI schemes were introduced by Peaceman and Rachford [31] and Douglas [6] for equations without mixed derivatives on rectangular domains. The original schemes exhibit order two in time if no mixed derivatives are present, otherwise the order is reduced to one. More sophisticated schemes are introduced in [1, 17, 23] and are able to deal with mixed derivatives while maintaining second order accuracy in time. Typically the ADI schemes are supplied with second order discretisations in space. HOC discretisations for ADI time stepping methods were considered by Mitchell and Fairweather [30] as well as Karaa and Zhang [24] for problems without mixed derivatives. Düring et al. [9] then derived high-order (HO) ADI methods for convection diffusion equations with mixed derivatives. Recently their work was extended in [16], where multi-dimensional diffusion equations

*Bergische Universität Wuppertal, Chair of Applied Mathematics and Numerical Analysis (AMNA), Gaußstrasse 20, 42119 Wuppertal, Germany. Email: {hendricks,ehrhardt,guenther}@math.uni-wuppertal.de
**Email: heuer.chr@googlemail.com

in a sparse grid setting are solved, and by Düring and Miles in [11], which considers a HO ADI discretisation for stochastic volatility models.

The combination technique can be employed to construct the *sparse grid solution* [12, 33], which only has $\mathcal{O}\left(h^{-1}\log_2(h^{-1})^{d-1}\right)$ nodes compared to $\mathcal{O}(h^{-d})$ nodes using the full grid, where d denotes the dimension of the problem. Thus, it suffers from the curse of dimensionality in a much lower extent. The combination technique is based on linearly combining an anisotropic sequence of solutions such that low order error terms cancel out. This results in a pointwise accuracy of $\mathcal{O}\left(h^{p}\log_2(h^{-1})^{d-1}\right)$ if a p-th order scheme is used to compute each sub-solution. The combination technique has successfully been applied to option pricing problems in [27, 33] by Reisinger and Leentvaar, who use spatial discretisations of order two and employ grid stretching as well as grid transformation techniques. In their numerical experiments the sparse grid solution outperforms the full grid solution and underlines the advantages of the combination technique in terms of accuracy versus computationally effort.

In this article we extend our previous work [16] to convection diffusion equations with mixed derivative terms. Therefore, we consider the following PDE on a d-dimensional spatial domain with space-dependent coefficients

$$\frac{\partial u}{\partial t} + L\,u = 0, \tag{1}$$

where L is an elliptic operator of the form

$$Lu = \sum_{i=1}^{d} a_i(\mathbf{x})\tfrac{\partial^2 u}{\partial x_i^2} + \sum_{\substack{i,j=1\\ i\neq j}}^{d} b_{i,j}(\mathbf{x})\tfrac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{d} c_i(\mathbf{x})\tfrac{\partial u}{\partial x_i}$$

on a rectangular domain $\Omega_d \times \Omega_t$ with suitable initial and boundary data. In the spatial domain we use a fourth order scheme, which provides highly accurate solutions. The time domain is discretised via ADI splitting, such that the discretisation matrix is efficiently decomposed into a sequence of sparse matrices. The linear systems occurring in the algorithm can then be solved in linear run-time. To further reduce the number of grid nodes, we construct a sparse grid solution using the combination technique. This significantly reduces the number of degrees of freedom, while maintaining a high accuracy. Based on the stability results for second-order ADI schemes in two spatial dimensions, we perform a von Neumann stability analysis for HO ADI schemes in the case of frozen coefficients. For the cases of three and four dimensional spatial domains we investigate the stability behaviour experimentally.

The article is organised as follows: in Section 2 we give an introduction to four well known ADI schemes. In Section 3 we carry over the idea of dimensional splitting to the semi-discrete (space discretisation) HO representation. Section 4 is devoted to the stability analysis of the HO ADI schemes with frozen coefficients. A brief introduction of the combination technique and sparse grids is followed by numerical examples. We apply the numerical schemes to the multi-dimensional Black-Scholes PDE to price European basket options.

2. ADI schemes

We consider the semi-discretisation of the PDE (1)

$$u'(t) = F(u(t)), \qquad t > 0,$$

with suitable initial and boundary data and u only depending on the time t. While fully implicit schemes result in a non-sparse linear equation system, which is very expensive to solve, explicit time-stepping suffers from restrictions on the time step-size to ensure stability. For an effective time discretisation of this semi-discrete problem setting ADI schemes have been discussed in the literature [1, 5, 6, 23, 22], relying on a decomposition of F

$$F(u(t)) = F_0(u(t)) + F_1(u(t)) + \ldots + F_d(u(t)).$$

In this paper F_0 stems from all mixed derivatives and F_i from all unidirectional contributions in the i-th coordinate direction of PDE (1) for $i = 1, ..., d$. In the case of a finite difference discretisation the mixed derivatives lead to broad stencils, which will be treated explicitly. The F_is for $i = 1, ..., d$ can be discretised with a compact three point stencil and will be treated implicitly. We consider four ADI schemes.

Douglas scheme (DO):

$$\begin{cases} Y_0 & = u_n + \Delta_t F(u_n), \\ Y_i & = Y_{i-1} + \theta\Delta_t \left(F_i(Y_i) - F_i(u_n)\right) \text{ for } i = 1, ..., d \\ u_{n+1} & = Y_d. \end{cases} \tag{2}$$

Craig-Sneyd scheme (CS):

$$\begin{cases} Y_0 & = u_n + \Delta_t F(u_n), \\ Y_i & = Y_{i-1} + \theta\Delta_t \left(F_i(Y_i) - F_i(u_n)\right) \text{ for } i = 1, ..., d \\ \tilde{Y}_0 & = Y_0 + \frac{1}{2}\Delta_t \left(F_0(Y_d) - F_0(u_n)\right) \\ \tilde{Y}_i & = \tilde{Y}_{i-1} + \theta\Delta_t \left(F_i(\tilde{Y}_i) - F_i(u_n)\right) \text{ for } i = 1, ..., d \\ u_{n+1} & = \tilde{Y}_d. \end{cases} \tag{3}$$

Modified Craig-Sneyd scheme (MCS):

$$\begin{cases} Y_0 & = u_n + \Delta_t F(u_n), \\ Y_i & = Y_{i-1} + \theta\Delta_t \left(F_i(Y_i) - F_i(u_n)\right) \text{ for } i = 1, ..., d \\ \hat{Y}_0 & = Y_0 + \theta\Delta_t \left(F_0(Y_d) - F_0(u_n)\right) \\ \tilde{Y}_0 & = \hat{Y}_0 + (\frac{1}{2} - \theta)\Delta_t \left(F(Y_d) - F(u_n)\right) \\ \tilde{Y}_i & = \tilde{Y}_{i-1} + \theta\Delta_t \left(F_i(\tilde{Y}_i) - F_i(u_n)\right) \text{ for } i = 1, ..., d \\ u_{n+1} & = \tilde{Y}_d. \end{cases} \tag{4}$$

Hundsdorfer-Verwer scheme (HV):

$$\begin{cases} Y_0 & = u_n + \Delta_t F(u_n), \\ Y_i & = Y_{i-1} + \theta\Delta_t \left(F_i(Y_i) - F_i(u_n)\right) \text{ for } i = 1, ..., d \\ \tilde{Y}_0 & = Y_0 + \frac{1}{2}\Delta_t \left(F(Y_d) - F(u_n)\right) \\ \tilde{Y}_i & = \tilde{Y}_{i-1} + \theta\Delta_t \left(F_i(\tilde{Y}_i) - F_i(Y_d)\right) \text{ for } i = 1, ..., d \\ u_{n+1} & = \tilde{Y}_d, \end{cases} \tag{5}$$

where $\Delta_t > 0$ is the step size in time, $u_n \sim u(n\Delta t)$ and $\theta > 0$. The Douglas scheme was initially developed for application to the heat equation, see [6], and exhibits order two in time if $\theta = \frac{1}{2}$ and $F_0 = 0$, order one otherwise. In the case of $\theta = \frac{1}{2}$ the method is known as the Douglas [6] and Brian [3] scheme, while it has also been considered by Douglas in [5] for $\theta = 1, F_0 = 0$. The Craig-Sneyd scheme and the modified Craig-Sneyd scheme can be seen as an extension of the Douglas scheme. The Craig-Sneyd scheme was introduced in [1] and has order two in time if $\theta = \frac{1}{2}$ independent of F_0. The modified Craig-Sneyd scheme was defined by in't Hout and Welfert in [23] and exhibits order two in time for any θ. Please note, that for $\theta = \frac{1}{2}$ the modified Craig-Sneyd scheme reduces to the Craig-Sneyd scheme. The Hundsdorfer-Verwer scheme was introduced in [17] and shows consistency of order two for arbitrary θ. A small θ value in general leads to a more accurate solution, but might cause instabilities if chosen too small. Hence, it is important to carefully determine bounds to ensure a high accuracy, as well as an unconditional stable bahaviour.

Let the scheme be given in one-step form

$$u_{n+1} = Ru_n,$$

with iteration matrix R and vectors u_n, u_{n+1} respectively. In order to analyse the stability of the schemes, they are applied to the linear scalar test equation

$$u'(t) = (\lambda_0 + \lambda_1 + ... + \lambda_d)\, u(t), \tag{6}$$

with complex values λ_i for $i = 0, 1, ..., d$. Then R reduces to the scalar factor $r(z_0, z_1, ..., z_d)$ with $z_i = \Delta_t \lambda_i$ for $i = 0, 1, ..., d$ and the numerical scheme is stable iff

$$|r| \leq 1$$

is fulfilled. Defining

$$z = z_0 + z_1 + ... + z_d, \qquad p = (1 - \theta z_1) \cdot (1 - \theta z_2) \cdot ... \cdot (1 - \theta z_d)$$

the scalar functions r for the ADI schemes (2) - (5) are given by

$$r_{DO}(z_0, z_1, ..., z_d) = 1 + \tfrac{z}{p}, \qquad r_{CS}(z_0, z_1, ..., z_d) = 1 + \tfrac{z}{p} + \tfrac{1}{2}\tfrac{z_0 z}{p^2},$$
$$r_{MCS}(z_0, z_1, ..., z_d) = 1 + \tfrac{z}{p} + \theta\tfrac{z_0 z}{p^2} + (\tfrac{1}{2} - \theta)\tfrac{z^2}{p^2}, \qquad r_{HV}(z_0, z_1, ..., z_d) = 1 + 2\tfrac{z}{p} - \tfrac{z}{p^2} + \tfrac{1}{2}\tfrac{z^2}{p^2}.$$

Conditions on θ ensuring unconditional stability have been derived in the pure diffusion case with two and three spatial dimensions [21, 23]. For higher dimensions necessary lower bounds could be given in [21]. In case of the HV scheme it has been proved in [23] that these are sufficient. For the other three schemes it is not clear if the lower bounds are sufficient. In [20, 22, 26] the stability of these four schemes is analysed for convection-diffusion equations. There the following conditions for two dimensional problems have been derived:

Assume conditions

$$Re(z_1) \leq 0, \qquad Re(z_2) \leq 0 \qquad \text{and} \qquad |z_0| \leq 2\sqrt{Re(z_1) \cdot Re(z_2)} \tag{7}$$

hold for $z_0, z_1, z_2 \in \mathbb{C}$ and $\theta \geq \frac{1}{2}$, then $|r_{DO}| \leq 1$ and $|r_{CS}| \leq 1$ holds. In 't Hout and Mishra [20] show for $F_0 = 0$ that for the modified Craig-Sneyd scheme it holds $|r_{MCS}(0, z_1, z_2)| \leq 1$ for all $z_1, z_2 \in \mathbb{C}$ with negative real part $Re(z_1) \leq 0, Re(z_2) \leq 0$ if and only if $\theta \geq \frac{1}{4}$. For $z_0 \in \mathbb{R}$ and $z_1, z_2 \in \mathbb{C}$ fulfilling (7), they derive the necessary stability condition $\theta \geq \frac{2}{5}$. However, the scheme has been applied successfully to convection-diffusion equations with mixed derivatives in [19] even for $\theta \geq \frac{1}{3}$. An experimental analysis of this observation can be found in [20]. Recently it has been shown by Mishra [29] that the MCS scheme is unconditionally stable for $\theta > \frac{1}{3}$ if the modulus of the mixed derivative coefficient is at most $\frac{2+\sqrt{10}}{6}$.

For pure diffusion equations stability could be shown in the three dimensional case [21]:

$$\text{DO: } \theta \geq \max\left\{\tfrac{1}{2}, \tfrac{2(2\gamma+1)}{9}\right\}, \quad \text{CS: } \theta \geq \tfrac{1}{2},$$
$$\text{MCS: } \theta \geq \max\left\{\tfrac{1}{4}, \tfrac{2}{13}(2\gamma+1)\right\}, \quad \text{HV: } \theta \geq \max\left\{\tfrac{1}{4}, \tfrac{2\gamma+1}{4+2\sqrt{3}}\right\},$$

for a parameter $\gamma \in [0, 1]$, which describes the relative size of the mixed derivative coefficient

$$|a_{ij}| \leq \gamma\sqrt{a_{ii}a_{jj}} \text{ for all } i \neq j,$$

where $A = (a_{ij})$ denotes the symmetric, positive semi-definte diffusion coefficient matrix. In the higher dimensional case necessary bounds on θ could be found [21]:

$$\text{DO: } \theta \geq \max\left\{\tfrac{1}{2}, \tfrac{1}{2}(1 - \tfrac{1}{d})^{d-1}((d-1)\gamma + 1)\right\}, \quad \text{CS: } \theta \geq \max\left\{\tfrac{1}{2}, \tfrac{1}{2}(1 - \tfrac{1}{d})^{d} d\gamma\right\},$$
$$\text{MCS: } \theta \geq \max\left\{\tfrac{1}{4}, \tfrac{1}{2}\frac{(d-1)\gamma + 1}{1 + (1 + \frac{1}{d-1})^{d-1}}\right\}, \quad \text{HV: } \theta \geq \max\left\{\tfrac{1}{4}, \tfrac{1}{2}a_d((d-1)\gamma + 1)\right\},$$

where $a_d \in (0, \frac{1}{2})$ is the unique solution of the equation $2a_d\,(1 + \frac{1-a_d}{d-1})^{d-1} - 1 = 0$. In [14] the stability for convection-diffusion problems with three spatial dimensions was experimentally analysed. The bounds derived for pure diffusion equations turned out to lead to a stable behaviour in case of the DO, CS and MCS scheme. For the HV scheme with $\theta = \frac{1}{2} + \frac{1}{6}\sqrt{3}$ the error decayed monotonically with Δ_t. This θ-value was derived for two dimensional convection-diffusion equations without mixed derivatives in [26].

3. High-order ADI schemes

In recent years high-order-compact finite difference schemes have been proposed to solve elliptic [38] and parabolic [2, 8, 9, 24, 39] partial differential equations numerically. These schemes make use of the structure and smoothness of the solution of the problem to algebraically derive a fourth-order approximation while maintaining a compact stencil. This leads to a discretisation matrix with small bandwidth and hence to a low computational effort and low memory consumption. Combining this with ADI methods leads to an efficient time integrator. First, we introduce the finite difference operators used throughout this paper, afterwards we derive a HOC discretisation of the unidirectional contributions F_i for $i = 1, ..., d$ and then apply them to the ADI framework of the previous Section. Please note, that the derivation of the HO ADI scheme follows the approach by Düring et al. in [9].

3.1. Finite difference operators

In this Section we discretise the derivatives occurring in (1) by using central difference operators. Let $\boldsymbol{l} = (l_1, l_2, ..., l_d) \in \mathbb{N}^d$ denote a multi-index, then we can define a grid $\Omega_{\boldsymbol{l}}$ on Ω_d with mesh width $\boldsymbol{h} = (h_1, h_2, ..., h_d) = (c_1 \cdot 2^{-l_1}, c_2 \cdot 2^{-l_2}, ..., c_d \cdot 2^{-l_d})$, which consists of the grid nodes $x_{\boldsymbol{l},\boldsymbol{j}} = (x_{l_1,j_1}, x_{l_2,j_2}, ..., x_{l_d,j_d})$ with $x_{l_i,j_i} = j_i \cdot h_i$ for $j_i = 0, 1, ..., 2^{l_i}$ and c_i is the size of the computational domain in coordinate direction i for $i = 1, 2, ..., d$. With Taylor expansions under the assumption that u is sufficiently smooth, a second order approximation to the first and second derivative with respect to direction i at grid node $x_{\boldsymbol{l},\boldsymbol{j}}$ is given by

$$\delta_i^0 u(x_{\boldsymbol{l},\boldsymbol{j}}) = \tfrac{1}{2h_i}\left(u(x_{\boldsymbol{l},\boldsymbol{j}} + h_i e_i) - u(x_{\boldsymbol{l},\boldsymbol{j}} - h_i e_i)\right) = \frac{\partial u}{\partial x_i}(x_{\boldsymbol{l},\boldsymbol{j}}) + \mathcal{O}(h_i^2),$$
$$\delta_i^2 u(x_{\boldsymbol{l},\boldsymbol{j}}) = \tfrac{1}{h_i^2}\left(u(x_{\boldsymbol{l},\boldsymbol{j}} + h_i e_i) - 2u(x_{\boldsymbol{l},\boldsymbol{j}}) + u(x_{\boldsymbol{l},\boldsymbol{j}} - h_i e_i)\right) = \frac{\partial^2 u}{\partial x_i^2}(x_{\boldsymbol{l},\boldsymbol{j}}) + \mathcal{O}(h_i^2),$$

where e_i denotes the i-th unit vector. The mixed derivatives will be treated in an explicit way, hence we use a broad stencil to approximate the first derivative with order four

$$\begin{aligned}\tilde{\delta}_i^0 u(x_{\boldsymbol{l},\boldsymbol{j}}) &= \tfrac{1}{12h_i}\left(-u(x_{\boldsymbol{l},\boldsymbol{j}} + 2h_i e_i) + 8u(x_{\boldsymbol{l},\boldsymbol{j}} + h_i e_i) - 8u(x_{\boldsymbol{l},\boldsymbol{j}} - h_i e_i) + u(x_{\boldsymbol{l},\boldsymbol{j}} - 2h_i e_i)\right)\\ &= \frac{\partial u}{\partial x_i}(x_{\boldsymbol{l},\boldsymbol{j}}) + \mathcal{O}(h_i^4).\end{aligned}$$

Thus, we obtain a fourth order approximation of the cross derivative for $i \neq j$ with

$$\begin{aligned}\tilde{\delta}_i^0 \tilde{\delta}_j^0 u(x_{\boldsymbol{l},\boldsymbol{j}}) = \frac{1}{144 h_i h_j}\Big[&64\Big(u(x_{\boldsymbol{l},\boldsymbol{j}} + h_i e_i + h_j e_j) - u(x_{\boldsymbol{l},\boldsymbol{j}} - h_i e_i + h_j e_j)\\ &+ u(x_{\boldsymbol{l},\boldsymbol{j}} - h_i e_i - h_j e_j) - u(x_{\boldsymbol{l},\boldsymbol{j}} + h_i e_i - h_j e_j)\Big)\\ &+ 8\Big(-u(x_{\boldsymbol{l},\boldsymbol{j}} + 2h_i e_i + h_j e_j) - u(x_{\boldsymbol{l},\boldsymbol{j}} + h_i e_i + 2h_j e_j)\\ &\quad + u(x_{\boldsymbol{l},\boldsymbol{j}} - h_i e_i + 2h_j e_j) + u(x_{\boldsymbol{l},\boldsymbol{j}} - 2h_i e_i + h_j e_j)\\ &\quad - u(x_{\boldsymbol{l},\boldsymbol{j}} - 2h_i e_i - h_j e_j) - u(x_{\boldsymbol{l},\boldsymbol{j}} - h_i e_i - 2h_j e_j)\\ &\quad + u(x_{\boldsymbol{l},\boldsymbol{j}} + h_i e_i - 2h_j e_j) + u(x_{\boldsymbol{l},\boldsymbol{j}} + 2h_i e_i - h_j e_j)\Big)\\ &+ u(x_{\boldsymbol{l},\boldsymbol{j}} + 2h_i e_i + 2h_j e_j) - u(x_{\boldsymbol{l},\boldsymbol{j}} - 2h_i e_i + 2h_j e_j)\\ &+ u(x_{\boldsymbol{l},\boldsymbol{j}} - 2h_i e_i - 2h_j e_j) - u(x_{\boldsymbol{l},\boldsymbol{j}} + 2h_i e_i - 2h_j e_j)\Big].\end{aligned}$$

3.2. HO finite differences

In this Section we derive HO approximation of the summands F_i arising in the decomposition of (1). For brevity we streamline our notation and write $a_i := a_i(x_{\boldsymbol{l},\boldsymbol{j}})$ and $c_i := c_i(x_{\boldsymbol{l},\boldsymbol{j}})$, $b_{i,j} := b_{i,j}(x_{\boldsymbol{l},\boldsymbol{j}})$. The unidirectional contributions are given by

$$F_i(u) = a_i \tfrac{\partial^2 u}{\partial x_i^2} + c_i \tfrac{\partial u}{\partial x_i} = f \tag{8}$$

for $i = 1, ..., d$ and some arbitrary smooth right hand side f. Inserting the finite difference operators we obtain

$$F_i(u(x_{\boldsymbol{l},\boldsymbol{j}})) = a_i \delta_i^2 u(x_{\boldsymbol{l},\boldsymbol{j}}) - a_i \tfrac{h_i^2}{12} \tfrac{\partial^4 u}{\partial x_i^4}(x_{\boldsymbol{l},\boldsymbol{j}}) - a_i \tfrac{h_i^4}{360} \tfrac{\partial^6 u}{\partial x_i^6}(x_{\boldsymbol{l},\boldsymbol{j}}) + c_i \delta_i^0 u(x_{\boldsymbol{l},\boldsymbol{j}}) - c_i \tfrac{h_i^2}{6} \tfrac{\partial^3 u}{\partial x_i^3}(x_{\boldsymbol{l},\boldsymbol{j}}) - c_i \tfrac{h_i^4}{120} \tfrac{\partial^5 u}{\partial x_i^5}(x_{\boldsymbol{l},\boldsymbol{j}}) + \mathcal{O}(h_i^6) = f(x_{\boldsymbol{l},\boldsymbol{j}}). \tag{9}$$

Since the leading error term in (9) is of order two, we can derive a fourth-order compact approximation if the third and fourth derivative is approximated with second order on the compact stencil. In order to derive these approximations, we differentiate equation (8) once with respect to x_i and thus get

$$\tfrac{\partial a_i}{\partial x_i} \tfrac{\partial^2 u}{\partial x_i^2} + a_i \tfrac{\partial^3 u}{\partial x_i^3} + \tfrac{\partial c_i}{\partial x_i} \tfrac{\partial u}{\partial x_i} + c_i \tfrac{\partial^2 u}{\partial x_i^2} = \tfrac{\partial f}{\partial x_i}.$$

Hence, the third derivative is given by the auxiliary equation

$$\tfrac{\partial^3 u}{\partial x_i^3} = \tfrac{1}{a_i} \tfrac{\partial f}{\partial x_i} - \left(\tfrac{1}{a_i} \tfrac{\partial a_i}{\partial x_i} + \tfrac{c_i}{a_i} \right) \tfrac{\partial^2 u}{\partial x_i^2} - \tfrac{1}{a_i} \tfrac{\partial c_i}{\partial x_i} \tfrac{\partial u}{\partial x_i}. \tag{10}$$

In a similar fashion we obtain an expression for the fourth derivative by differentiating (8) twice with respect to x_i. The third and fourth derivative can then be approximated with second-order stencils via central difference operators. Replacing the truncation error in (9) leads to a fourth-order accurate approximation

$$\begin{aligned} &\left(a_i + \tfrac{h_i^2}{12} \tfrac{\partial^2 a_i}{\partial x_i^2} - \tfrac{h_i^2 c_i}{12 a_i} \tfrac{\partial a_i}{\partial x_i} - \tfrac{h_i^2}{6 a_i} \left[\tfrac{\partial a_i}{\partial x_i} \right]^2 + \tfrac{h_i^2 c_i^2}{12 a_i} + \tfrac{h_i^2}{6} \tfrac{\partial c_i}{\partial x_i} \right) \delta_i^2 u(x_{\boldsymbol{l},\boldsymbol{j}}) \\ &\quad + \left(c_i - \tfrac{h_i^2}{6 a_i} \tfrac{\partial a_i}{\partial x_i} \tfrac{\partial c_i}{\partial x_i} + \tfrac{h_i^2 c_i}{12 a_i} \tfrac{\partial c_i}{\partial x_i} + \tfrac{h_i^2}{12} \tfrac{\partial^2 c_i}{\partial x_i^2} \right) \delta_i^0 u(x_{\boldsymbol{l},\boldsymbol{j}}) + h_i^4 \tau_i \\ &\quad = f(x_{\boldsymbol{l},\boldsymbol{j}}) + \tfrac{h_i^2}{12} \delta_i^2 f(x_{\boldsymbol{l},\boldsymbol{j}}) + \left(\tfrac{h_i^2 c_i}{12 a_i} - \tfrac{h_i^2}{6 a_i} \tfrac{\partial a_i}{\partial x_i} \right) \delta_i^0 f(x_{\boldsymbol{l},\boldsymbol{j}}) \end{aligned} \tag{11}$$

on the compact stencil with

$$\begin{aligned} \tau_i = &\left(-\tfrac{1}{36 a_i} \tfrac{\partial^3 c_i}{\partial x_i^3} \tfrac{\partial a_i}{\partial x_i} + \tfrac{c_i}{72 a_i} \tfrac{\partial^3 c_i}{\partial x_i^3} + \tfrac{1}{144} \tfrac{\partial^4 c_i}{\partial x_i^4} \right) \tfrac{\partial u}{\partial x_i}(x_{\boldsymbol{l},\boldsymbol{j}}) \\ &+ \left(\tfrac{1}{144} \tfrac{\partial^4 a_i}{\partial x_i^4} + \tfrac{c_i}{72 a_i} \tfrac{\partial^3 a_i}{\partial x_i^3} - \tfrac{1}{12 a_i} \tfrac{\partial a_i}{\partial x_i} \tfrac{\partial^2 c_i}{\partial x_i^2} - \tfrac{1}{36 a_i} \tfrac{\partial a_i}{\partial x_i} \tfrac{\partial^3 a_i}{\partial x_i^3} + \tfrac{c_i}{24 a_i} \tfrac{\partial^2 c_i}{\partial x_i^2} + \tfrac{1}{36} \tfrac{\partial^3 c_i}{\partial x_i^3} \right) \tfrac{\partial^2 u}{\partial x_i^2}(x_{\boldsymbol{l},\boldsymbol{j}}) \\ &+ \left(\tfrac{1}{36} \tfrac{\partial^3 a_i}{\partial x_i^3} + \tfrac{c_i}{24 a_i} \tfrac{\partial^2 a_i}{\partial x_i^2} - \tfrac{1}{18 a_i} \tfrac{\partial a_i}{\partial x_i} \tfrac{\partial c_i}{\partial x_i} - \tfrac{1}{12 a_i} \tfrac{\partial a}{\partial x} \tfrac{\partial^2 a_i}{\partial x_i^2} + \tfrac{c_i}{36 a_i} \tfrac{\partial c_i}{\partial x_i} + \tfrac{1}{36} \tfrac{\partial^2 c_i}{\partial x_i^2} \right) \tfrac{\partial^3 u}{\partial x_i^3}(x_{\boldsymbol{l},\boldsymbol{j}}) \\ &+ \left(\tfrac{5}{144} \tfrac{\partial^2 a_i}{\partial x_i^2} + \tfrac{c_i}{48 a_i} \tfrac{\partial a_i}{\partial x_i} - \tfrac{1}{72 a_i} \left[\tfrac{\partial a}{\partial x_i} \right]^2 + \tfrac{c_i^2}{144 a_i} + \tfrac{1}{72} \tfrac{\partial c_i}{\partial x_i} \right) \tfrac{\partial^4 u}{\partial x_i^4}(x_{\boldsymbol{l},\boldsymbol{j}}) \\ &+ \tfrac{1}{80} c_i \tfrac{\partial^5 u}{\partial x_i^5}(x_{\boldsymbol{l},\boldsymbol{j}}) + \tfrac{1}{240} a_i \tfrac{\partial^6 u}{\partial x_i^6}(x_{\boldsymbol{l},\boldsymbol{j}}) + \mathcal{O}(h_i^2). \end{aligned} \tag{12}$$

From the truncation error τ_i we see that beside the solution u, also the coefficient functions a_i and c_i have to be sufficiently smooth, such that their fourth derivative is bounded.

Rewriting this scheme in terms of matrices or symbolic operators gives

$$A_{x_i} u = B_{x_i} f$$

for vectors u and f, where A_{x_i} corresponds to the left hand side of (11) and B_{x_i} to its right hand side. The semi-discrete scheme can then be written as

$$u'(t) = F_0(u) + B_{x_1}^{-1} A_{x_1} u + ... + B_{x_d}^{-1} A_{x_d} u + \mathcal{O}(h_1^4) + ... + \mathcal{O}(h_d^4) + \sum_{i,j} \mathcal{O}(h_i^4 h_j^4). \tag{13}$$

The mixed derivatives are approximated via

$$b_{i,j} \frac{\partial^2 u}{\partial x_i \partial x_j} = b_{i,j} \tilde{\delta}_i^0 \tilde{\delta}_j^0 u + h_i^4 \tilde{\tau}_i + h_j^4 \tilde{\tau}_j + h_i^4 h_j^4 \tau_{i,j}$$

with

$$\tilde{\tau}_i = b_{i,j} \frac{1}{30} \frac{\partial^6 u}{\partial x_i^5 \partial x_j}, \qquad \tilde{\tau}_j = b_{i,j} \frac{1}{30} \frac{\partial^6 u}{\partial x_i \partial x_j^5}, \qquad \tau_{i,j} = -b_{i,j} \frac{1}{900} \frac{\partial^{10} u}{\partial x_i^5 \partial x_j^5} \tag{14}$$

for $i \neq j$, $i,j = 1,...,d$. Thus, the mixed term is given by $F_0(u) = \sum_{i\neq j} b_{i,j}\tilde{\delta}_i^0\tilde{\delta}_j^0 u$ or in matrix notation by $F_0(u) = \sum_{i\neq j} A_{x_i,x_j} u$. Since the treatment of the mixed derivative requires to use a broad stencil, the spatial approximation is not defined on a compact stencil anymore. Nevertheless, the usage of ADI schemes allows us to treat the mixed derivatives explicitly. Therefore, the composition of the derived HO approximation and ADI time splitting is compact in each implicit step and non-compact in each explicit step.

3.3. HO ADI schemes

Using the HO formulation in each of the ADI schemes from Section 2 thus gives

HOC Douglas scheme (HDO):

$$\begin{cases} Y_0 & = u_n + \Delta_t \left(F_0(u_n) + B_{x_1}^{-1} A_{x_1} u_n + ... + B_{x_d}^{-1} A_{x_d} u_n\right) \\ \left(B_{x_i} - \theta\Delta_t A_{x_i}\right) Y_i & = B_{x_i} Y_{i-1} - \theta\Delta_t A_{x_i} u_n \text{ for } i = 1,...,d \\ u_{n+1} & = Y_d, \end{cases} \tag{15}$$

HOC Craig-Sneyd scheme (HCS):

$$\begin{cases} Y_0 & = u_n + \Delta_t \left(F_0(u_n) + B_{x_1}^{-1} A_{x_1} u_n + ... + B_{x_d}^{-1} A_{x_d} u_n\right) \\ \left(B_{x_i} - \theta\Delta_t A_{x_i}\right) Y_i & = B_{x_i} Y_{i-1} - \theta\Delta_t A_{x_i} u_n \text{ for } i = 1,...,d \\ \tilde{Y}_0 & = Y_0 + \frac{1}{2}\Delta_t \left(F_0(Y_d) - F_0(u_n)\right) \\ \left(B_{x_i} - \theta\Delta_t A_{x_i}\right) \tilde{Y}_i & = B_{x_i} \tilde{Y}_{i-1} - \theta\Delta_t A_{x_i} u_n \text{ for } i = 1,...,d \\ u_{n+1} & = \tilde{Y}_d, \end{cases} \tag{16}$$

HOC Modified Craig-Sneyd scheme (HMCS):

$$\begin{cases} Y_0 & = u_n + \Delta_t \left(F_0(u_n) + B_{x_1}^{-1} A_{x_1} u_n + ... + B_{x_d}^{-1} A_{x_d} u_n\right), \\ \left(B_{x_i} - \theta\Delta_t A_{x_i}\right) Y_i & = B_{x_i} Y_{i-1} - \theta\Delta_t A_{x_i} u_n \text{ for } i = 1,...,d \\ \hat{Y}_0 & = Y_0 + \theta\Delta_t \left(F_0(Y_d) - F_0(u_n)\right) \\ \tilde{Y}_0 & = \hat{Y}_0 + (\tfrac{1}{2} - \theta)\Delta_t \left(F(Y_d) - F(u_n)\right) \\ \left(B_{x_i} - \theta\Delta_t A_{x_i}\right) \tilde{Y}_i & = B_{x_i} \tilde{Y}_{i-1} - \theta\Delta_t A_{x_i} u_n \text{ for } i = 1,...,d \\ u_{n+1} & = \tilde{Y}_d. \end{cases} \tag{17}$$

and *HO Hundsdorfer-Verwer scheme (HHV):*

$$\begin{cases} Y_0 & = u_n + \Delta_t \left(F_0(u_n) + B_{x_1}^{-1} A_{x_1} u_n + ... + B_{x_d}^{-1} A_{x_d} u_n\right), \\ \left(B_{x_i} - \theta\Delta_t A_{x_i}\right) Y_i & = B_{x_i} Y_{i-1} - \theta\Delta_t A_{x_i} u_n \text{ for } i = 1,...,d \\ \tilde{Y}_0 & = Y_0 + \frac{1}{2}\Delta_t \left(F(Y_d) - F(u_n)\right) \\ \left(B_{x_i} - \theta\Delta_t A_{x_i}\right) \tilde{Y}_i & = B_{x_i} \tilde{Y}_{i-1} - \theta\Delta_t A_{x_i} Y_d \text{ for } i = 1,...,d \\ u_{n+1} & = \tilde{Y}_d. \end{cases} \tag{18}$$

In order to avoid the explicit computation of the inverse $B_{x_i}^{-1}$ for $i = 1,...,d$, one can rewrite the schemes by multiplying with B_{x_i} and introducing new variables $Z_i := \prod_{j=i+1}^{d} B_{x_j} Y_i$, $\tilde{Z}_i := \prod_{j=i+1}^{d} B_{x_j} \tilde{Y}_i$, and $\hat{Z}_0 := \prod_{j=1}^{d} B_{x_j} \hat{Y}_0$. For a detailed derivation we refer to [16]. All mixed derivatives are treated explicitly and the discretisation operator reads $B_{x_1} \cdot ... \cdot B_{x_d} F_0(u)$. This formulation leads to broad stencils since B_{x_i} as well as $\tilde{\delta}_i^0$ act in coordinate direction i. The involved discretisations in the single coordinate directions are given by $\delta_{x_i}^0\tilde{\delta}_i^0$ and $\delta_{x_i}^2\tilde{\delta}_i^0$

$$\begin{aligned} \delta_{x_i}^0\tilde{\delta}_i^0 u(x_{\boldsymbol{l},\boldsymbol{j}}) &= \tfrac{1}{24h_i^2}\big(- u(x_{\boldsymbol{l},\boldsymbol{j}} - 3h_i e_i) + 8u(x_{\boldsymbol{l},\boldsymbol{j}} - 2h_i e_i) \\ &\quad + u(x_{\boldsymbol{l},\boldsymbol{j}} - h_i e_i) - 16u(x_{\boldsymbol{l},\boldsymbol{j}}) + u(x_{\boldsymbol{l},\boldsymbol{j}} + h_i e_i) + 8u(x_{\boldsymbol{l},\boldsymbol{j}} + 2h_i e_i) - u(x_{\boldsymbol{l},\boldsymbol{j}} + 3h_i e_i)\big) \\ &= \tfrac{\partial^2 u}{\partial x_i^2}(x_{\boldsymbol{l},\boldsymbol{j}}) + \tfrac{h_i^2}{6}\tfrac{\partial^4 u}{\partial x_i^4}(x_{\boldsymbol{l},\boldsymbol{j}}) + \mathcal{O}(h_i^4), \\ \delta_{x_i}^2\tilde{\delta}_i^0 u(x_{\boldsymbol{l},\boldsymbol{j}}) &= \tfrac{1}{12h_i^3}\big(u(x_{\boldsymbol{l},\boldsymbol{j}} - 3h_i e_i) - 10u(x_{\boldsymbol{l},\boldsymbol{j}} - 2h_i e_i) + 17u(x_{\boldsymbol{l},\boldsymbol{j}} - h_i e_i) - 17u(x_{\boldsymbol{l},\boldsymbol{j}} + h_i e_i) \\ &\quad + 10u(x_{\boldsymbol{l},\boldsymbol{j}} + 2h_i e_i) - u(x_{\boldsymbol{l},\boldsymbol{j}} + 3h_i e_i)\big) \\ &= \tfrac{\partial^3 u}{\partial x_i^3}(x_{\boldsymbol{l},\boldsymbol{j}}) + \tfrac{h_i^2}{12}\tfrac{\partial^5 u}{\partial x_i^5}(x_{\boldsymbol{l},\boldsymbol{j}}) + \mathcal{O}(h_i^4). \end{aligned}$$

These operators are a discretisation of the second and third derivative with accuracy two. However, they are not optimal in the sense of width and therefore we propose to use smaller stencils with the same accuracy in our schemes

$$\frac{\partial^2 u}{\partial x_i^2} = \frac{1}{h_i^2}\left(u(x_{\boldsymbol{l},\boldsymbol{j}} + 2h_i e_i) - 2u(x_{\boldsymbol{l},\boldsymbol{j}}) + u(x_{\boldsymbol{l},\boldsymbol{j}} + 2h_i e_i)\right) + \mathcal{O}(h_i^2),$$
$$\frac{\partial^3 u}{\partial x_i^3} = -\frac{1}{2h_i^3}\left(u(x_{\boldsymbol{l},\boldsymbol{j}} - 2h_i e_i) - 2u(x_{\boldsymbol{l},\boldsymbol{j}} - h_i e_i) + 2u(x_{\boldsymbol{l},\boldsymbol{j}} + h_i e_i) - u(x_{\boldsymbol{l},\boldsymbol{j}} + 2h_i e_i)\right) + \mathcal{O}(h_i^2).$$

Please note, that this does not effect the consistency of the scheme. The ghost points arising in the approximation of the third derivative approximation or in $\tilde{\delta}_i^0$ can be computed via extrapolation [9]

$$u(x_{\boldsymbol{l},\boldsymbol{j}} \pm e_i) = 5u(x_{\boldsymbol{l},\boldsymbol{j}}) - 10u(x_{\boldsymbol{l},\boldsymbol{j}} \mp h_i e_i) + 10u(x_{\boldsymbol{l},\boldsymbol{j}} \mp 2h_i e_i) - 5u(x_{\boldsymbol{l},\boldsymbol{j}} \mp 3h_i e_i) + u(x_{\boldsymbol{l},\boldsymbol{j}} \mp 4h_i e_i) + \mathcal{O}(h_i^5).$$

4. Stability for frozen coefficients

In this Section we carry out a stability analysis of the HO ADI schemes (15) – (18). We consider the von Neumann stability analysis under the assumptions of frozen coefficients of the general convection-diffusion PDE

$$u_t = \operatorname{div}(A\nabla u) + c \cdot \nabla u \tag{19}$$

with $A = (a_{ij})$, $c = (c_1, c_2, ..., c_d)^\top$ and supplemented with periodic boundary conditions. Please note, that this causes all derivatives of the coefficients in our scheme to vanish. In the current setting the discretisation matrices are normal and commuting, such that it is sufficient to consider the scalar test equation (6) to prove stability in the von Neumann framework. For a detailed discussion we refer to [18]. Throughout this Section we will analytically investigate the stability for $d = 2$ and experimentally in higher dimensions. In a first step we rewrite our schemes to the one step form

$$u_{n+1} = R\, u_n.$$

For brevity we streamline our notation and introduce

$$Z_0 = \Delta_t \sum_{i \neq j} A_{x_i, x_j}, \qquad Z = Z_0 + \Delta_t B_{x_1}^{-1} A_{x_1} + ... + \Delta_t B_{x_d}^{-1} A_{x_d},$$
$$Q_{x_i} = B_{x_i} - \theta \Delta_t A_{x_i} \text{ for } i = 1, ..., d, \qquad P^{-1} = \prod_{i=1}^{d} Q_{x_i}^{-1} B_{x_i}.$$

The iteration matrices are then given by

$$R_{HDO} = I + P^{-1} Z, \qquad R_{HCS} = I + P^{-1} Z + \tfrac{1}{2} P^{-2} Z_0 Z,$$
$$R_{HMCS} = I + P^{-1} Z + \theta P^{-2} Z_0 Z + (\tfrac{1}{2} - \theta) P^{-2} Z^2, \quad R_{HHV} = I + 2P^{-1} Z - P^{-2} Z + \tfrac{1}{2} P^{-2} Z^2.$$

A detailed derivation can be found in [16]. One readily observes that the stability matrices exhibit the same structure like the stability functions in Section 2. Inserting Fourier modes into the discretisation operators we obtain the eigenvalues

$$\begin{aligned}
\tilde{z}_i &= 2\left(a_{ii} + \tfrac{h_i^2 c_i^2}{12 a_{ii}}\right) \tfrac{1}{h_i^2}(\cos\phi_i - 1) + c_i \tfrac{1}{h_i} I \sin\phi_i && \text{for } i = 1, ..., d,\\
\bar{z}_i &= 1 - \tfrac{1}{6}(1 - \cos\phi_i) + \tfrac{c_i}{12 a_{ii}} h_i I \sin\phi_i && \text{for } i = 1, ..., d,\\
z_0 &= -\sum_{i \neq j} a_{ij} \tfrac{1}{36} \tfrac{\Delta_t}{h_i h_j} (8\sin\phi_i - \sin 2\phi_i)(8\sin\phi_j - \sin 2\phi_j), && \text{(20)}
\end{aligned}$$

with imaginary unit I. The eigenvalues $\tilde{z}_i$ stem from A_{x_i} whereas $\bar{z}_i$ from B_{x_i} and z_0 from all cross derivatives. The angles ϕ_i are integer multiples of $2\pi/m_i$ with m_i being the dimension of the grid in x_i-direction for $i = 1, ..., d$. Defining

$$z_i = \Delta_t \tilde{z}_i / \bar{z}_i, \tag{21}$$

we obtain the scalar stability functions

$$r_{DO}(z_0,z_1,...,z_d) = 1+\tfrac{z}{p}, \qquad r_{CS}(z_0,z_1,...,z_d) = 1+\tfrac{z}{p}+\tfrac{1}{2}\tfrac{z_0 z}{p^2},$$
$$r_{MCS}(z_0,z_1,...,z_d) = 1+\tfrac{z}{p}+\theta\tfrac{z_0 z}{p^2}+(\tfrac{1}{2}-\theta)\tfrac{z^2}{p^2}, \quad r_{HHV}(z_0,z_1,...,z_d) = 1+2\tfrac{z}{p}-\tfrac{z}{p^2}+\tfrac{1}{2}\tfrac{z^2}{p^2},$$

with $p=(1-\theta z_1)\cdot ...\cdot(1-\theta z_d)$ and $z=z_0+z_1+...+z_d$. We see that the stability functions of the HO ADI schemes coincide with their standard ADI counterpart from Section 2. Thus, the proof of stability reduces to the analysis of the eigenvalues.

Lemma 1. *Let $d=2$ and HO ADI schemes (15)-(18) be applied to the convection-diffusion problem (19) with symmetric positive semi-definite coefficient matrix A. Then it holds for the eigenvalues, defined according to (20), (21), respectively,*

$$Re(z_1)\le 0, \qquad Re(z_2)\le 0 \qquad \text{and} \qquad |z_0|\le 2\sqrt{Re(z_1)\cdot Re(z_2)}.$$

Proof. The positive semi-definiteness of A is equivalent to

$$a_{11}\ge 0, \qquad a_{22}\ge 0, \qquad (a_{12}+a_{21})^2\le 4a_{11}a_{22}.$$

We first compute the real part of the eigenvalues z_i for $i=1,2$ and obtain

$$Re(z_i) = \Delta_t \frac{2\left(\frac{a_i}{h_i^2}+\frac{c_i^2}{12a_i}\right)(\cos\phi_i-1)\left(1-\frac{1}{6}(1-\cos\phi_i)\right)+\frac{c_i^2}{12a_i}\sin^2\phi}{\left(1-\frac{1}{6}(1-\cos\phi_i)\right)^2+\left(\frac{c_i h_i}{12a_i}\right)^2\sin^2\phi_i}.$$

With

$$\alpha_i := \tfrac{\cos\phi_i-1}{1-\frac{1}{6}(1-\cos\phi_i)},$$
$$\beta_i := h_i^2\tfrac{c_i^2}{a_{ii}^2},$$
$$\gamma_i := \sin^2\phi_i/\left(1-\tfrac{1}{6}(1-\cos\phi_i)\right)^2,$$

it holds

$$Re(z_i) = \Delta_t \frac{a_{ii}}{h_i^2} 2\alpha_i \,\frac{1+\frac{1}{24}\beta_i(2+\gamma_i/\alpha_i)}{1+\frac{1}{144}\beta_i\gamma_i}.$$

Please note that

$$\frac{1+\frac{1}{24}\beta_i(2+\gamma_i/\alpha_i)}{1+\frac{1}{144}\beta_i\gamma_i}\ge 1, \tag{22}$$

which can be verified by straightforward calculus

$$\frac{1+\frac{1}{24}\beta_i(2+\gamma_i/\alpha_i)}{1+\frac{1}{144}\beta_i\gamma_i}\ge 1 \quad\Leftrightarrow\quad \tfrac{1}{6}\le\frac{2\alpha_i+\gamma_i}{\alpha_i\gamma_i}=\frac{5+\cos\phi_i}{18\cos^2\phi_i/2}\in[\tfrac{1}{3},\infty).$$

Furthermore, the inequality

$$\tfrac{1}{36}\left(8\sin\phi_i-\sin 2\phi_i\right)^2\le -2\alpha_i \tag{23}$$

is fulfilled. Due to $\alpha_i\le 0$ and inequality (22) we directly observe that the real parts of the eigenvalues z_i lie on the left hand side of the complex plane. It remains to show $|z_0|\le 2\sqrt{Re(z_1)\cdot Re(z_2)}$. Due to the positive semi-definiteness of A we obtain

$$|z_0|^2\le 4a_{11}a_{22}\left(\frac{1}{36}\right)^2\frac{\Delta_t^2}{h_1^2h_2^2}\left(8\sin\phi_1-\sin 2\phi_1\right)^2\left(8\sin\phi_2-\sin 2\phi_2\right)^2$$

Exploiting conditions (23), (22), we obtain

$$\begin{aligned}|z_0|^2 &\leq 16 a_{11} a_{22} \frac{\Delta_t^2}{h_1^2 h_2^2} \alpha_1 \alpha_2 \\ &\leq 16 a_{11} a_{22} \frac{\Delta_t^2}{h_1^2 h_2^2} \alpha_1 \alpha_2 \frac{1 + \frac{1}{24}\beta_1 \left(2 + \gamma_1/\alpha_1\right)}{1 + \frac{1}{144}\beta_1 \gamma_1} \frac{1 + \frac{1}{24}\beta_2 \left(2 + \gamma_2/\alpha_2\right)}{1 + \frac{1}{144}\beta_2 \gamma_2} \\ &= 4 Re(z_1) \cdot Re(z_2).\end{aligned}$$

□

Remark:
The Lemma ensures the unconditional stability of schemes (15), (16) and the necessary condition on the lower bound for θ in scheme (17). For the HHV scheme (18) this leads to unconditional stability if $z_0 = 0$.

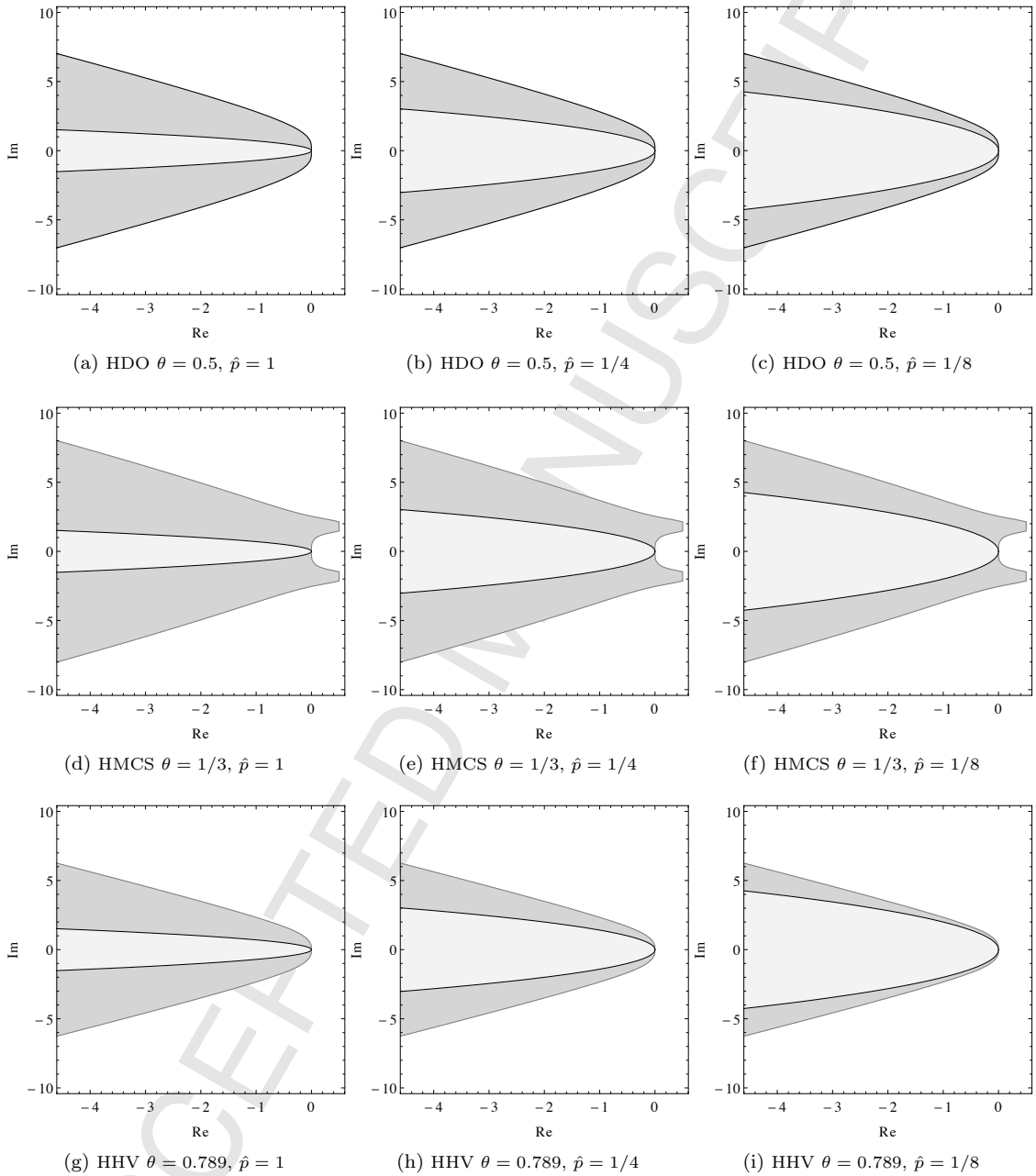
In Figures 1 and 2 we plot the stability regions of the HO ADI schemes for three and four spatial dimensions. The part in dark gray shows the stability region for the special choice $z_0 = 0$, $z_1 = z_2 = ... = z_d$. As in [16, 21] we consider the choice $z_1 = z_2 = ... = z_d$ to derive a necessary stability condition. The part in light gray shows the position of the eigenvalues z_i given by equation (21). The sample points have been computed for the parameter set $h_i = 10^{-1}$, $\Delta_t = 1$, $c_i = 1/2$ and $a_i = c_i \cdot \hat{p}$. This case is rather conservative as it considers a large parabolic mesh ratio. The parameter $\hat{p}$ determines the ratio between convection and diffusion and can be seen as the non-scaled reciprocal of the *Péclet number* [36]; the smaller $\hat{p}$, the stronger is the convection dominance. Note, in the case of $z_0 = 0$ the stability functions of the Douglas and the Craig-Sneyd scheme coincide. Therefore, we omit to plot the regions for the HCS scheme. The θ values have been chosen according to the results for ADI schemes with second order spatial disrectisation applied to diffusion equations without mixed derivative terms from literature, e.g. [21]. In [16] it was shown that these bounds are also valid for HO ADI schemes applied to pure diffusion problems. Hence, we also expect the HO ADI schemes for convection-diffusion equations to have similar stability properties as their second order counterpart. Both in the three (Figure 1) and four dimensional (Figure 2) case the eigenvalues (light gray) lie within the stability region (dark gray) in all plots except in Figure 2 (i). For strong convection dominance in plot (i) the HHV scheme is unstable. At the current state we have only investigated the stability if no mixed derivatives occur. In Section 6 we evaluate the behaviour of the schemes in the case of non-vanishing mixed derivatives.

5. Sparse grid combination technique

Solving high dimensional problems on a tensor based grid leads to a rapidly increasing complexity, the so called *curse of dimensionality*. As the number of degrees of freedom grows with $\mathcal{O}(h^{-d})$ the memory consumption quickly reaches the limit of available memory. Sparse grids and the sparse grid combination technique can reduce the number of necessary grid points significantly, but also maintain a rather high rate of accuracy. The sparse grid approach goes back to Smolyak [37], who developed it for numerical integration. Zenger [42], Bungartz et al. [4] and Schiekofer [35] transferred this idea to solve PDEs in the context of finite elements, finite volumes and finite differences. These methods in general require hierarchical, tree-like data structures, which makes the data structure management more complicated than in the full grid case [13]. With the help of the sparse grid combination technique [12] this problem can be overcome. Here, full grid solutions are linearly combined to construct the sparse grid solution. This allows to use standard full grid PDE solvers. Hence, this approach is very versatile and broadly applicable. Furthermore, each sub-solution can be computed independently, which makes it easily parallelisable.

In the following we give a brief introduction to the sparse grid combination technique in two dimensions. A more detailed derivation can be found in the literature [33] . We consider a problem on the unit square $\Omega = [0, 1]^2$ and assume that our numerical solution u_{h_1,h_2} has a second order error of the form

$$u - u_{h_1,h_2} = h_1^2 w_1(h_1) + h_2^2 w_2(h_2) + h_1^2 h_2^2 w_{1,2}(h_1, h_2).$$

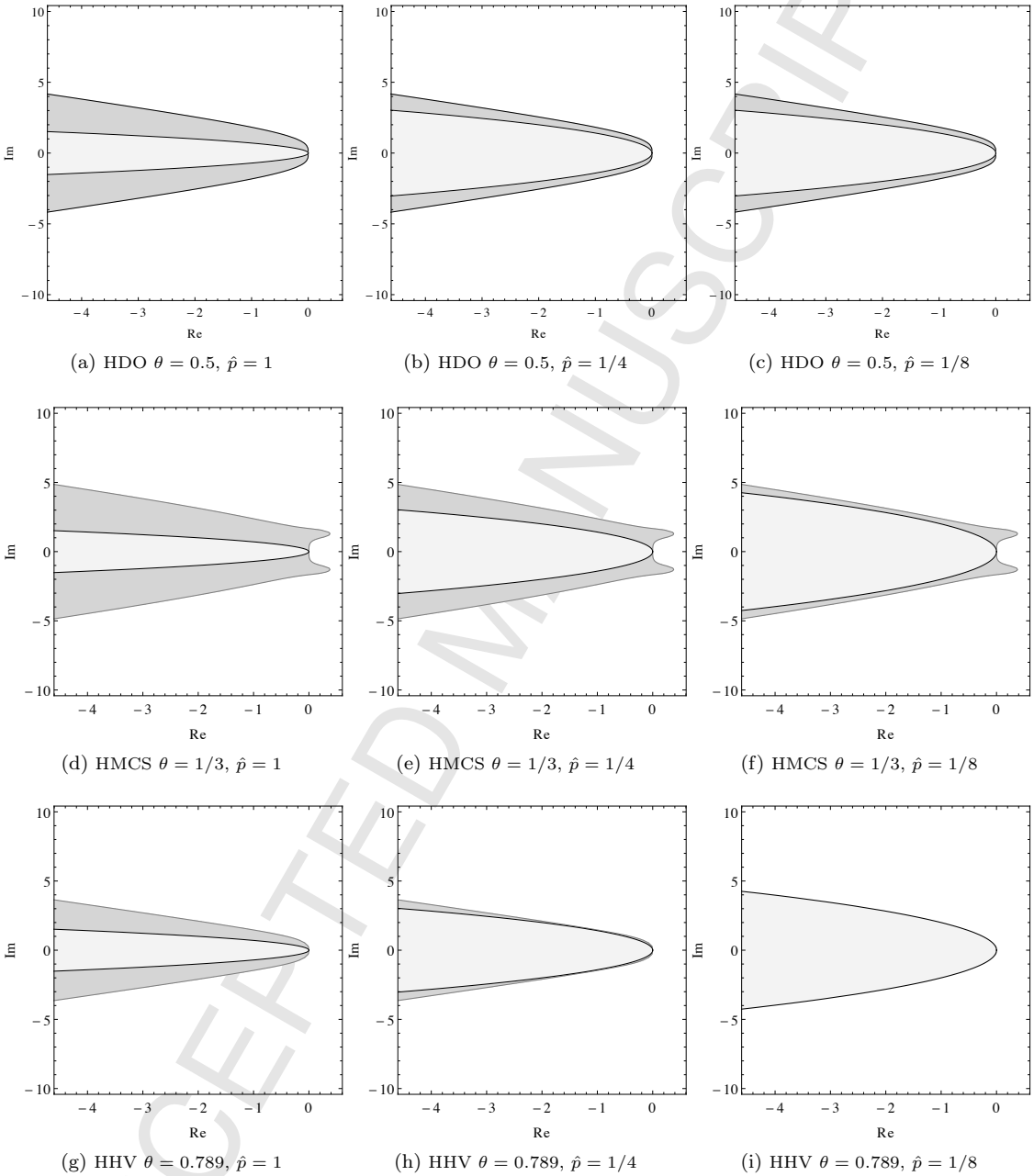


(a) HDO $\theta = 0.5$, $\hat{p} = 1$ (b) HDO $\theta = 0.5$, $\hat{p} = 1/4$ (c) HDO $\theta = 0.5$, $\hat{p} = 1/8$

(d) HMCS $\theta = 1/3$, $\hat{p} = 1$ (e) HMCS $\theta = 1/3$, $\hat{p} = 1/4$ (f) HMCS $\theta = 1/3$, $\hat{p} = 1/8$

(g) HHV $\theta = 0.789$, $\hat{p} = 1$ (h) HHV $\theta = 0.789$, $\hat{p} = 1/4$ (i) HHV $\theta = 0.789$, $\hat{p} = 1/8$

Figure 1: 3d: stability region (dark gray) for $z_1 = z_2 = z_3$ and eigenvalues z_i (light gray) for special parameter choices.



(a) HDO $\theta = 0.5$, $\hat{p} = 1$ (b) HDO $\theta = 0.5$, $\hat{p} = 1/4$ (c) HDO $\theta = 0.5$, $\hat{p} = 1/8$

(d) HMCS $\theta = 1/3$, $\hat{p} = 1$ (e) HMCS $\theta = 1/3$, $\hat{p} = 1/4$ (f) HMCS $\theta = 1/3$, $\hat{p} = 1/8$

(g) HHV $\theta = 0.789$, $\hat{p} = 1$ (h) HHV $\theta = 0.789$, $\hat{p} = 1/4$ (i) HHV $\theta = 0.789$, $\hat{p} = 1/8$

Figure 2: 4d: stability region (dark gray) for $z_1 = z_2 = z_3 = z_4$ and eigenvalues z_i (light gray) for special parameter choices.

The mesh widths h_1, h_2 are independent of one another and the functions w_1, w_2 only depend on either h_1 or h_2. Hence, we can subtract two numerical solutions with the same mesh width in one dimensions from another, such that one w_i term for either $i = 1$ or $i = 2$ cancels out. Pursuing this idea further leads to the two dimensional combination technique

$$u_n^s = \sum_{|\boldsymbol{l}|_1=n+1} u_{2^{-l_1},2^{-l_2}} - \sum_{|\boldsymbol{l}|_1=n} u_{2^{-l_1},2^{-l_2}}$$

with multi-index $\boldsymbol{l} = (l_1, l_2)$. Inserting the error structure from above all low order error terms cancel out and one obtains

$$\begin{aligned} u_n^s = u &+ 2^{-2(n+1)} w_1(2^{-(n+1)}) + 2^{-2(n+1)} w_2(2^{-(n+1)}) \\ &+ 2^{-2(n+1)} \sum_{i=0}^{n+1} w_{1,2}(2^{-i}, 2^{-(n+1-i)}) - 2^{-2n} \sum_{i=0}^{n} w_{1,2}(2^{-i}, 2^{-(n-i)}). \end{aligned}$$

If we assume that w_1, w_2 and $w_{1,2}$ are bounded by $K \in \mathbb{R}^+$, we get a pointwise error

$$|u_n^s - u| = \mathcal{O}(n\, 2^{-2n}),$$

$$|u_n^s - u| = \mathcal{O}(h^2 \log_2(h^{-1})) \text{ for } h = 2^{-n},$$

respectively. Compared to the original second order scheme the accuracy is deteriorated by $\mathcal{O}(\log_2(h^{-1}))$, but the loss of accuracy is compensated by a lower number of grid points. The sparse grid has $\mathcal{O}(2^{d-1} \cdot 2^n)$ degrees of freedom, which is equal to $\mathcal{O}(h^{-1} \log_2(h^{-1})^{d-1})$ for $h = 2^{-n}$, whereas the full grid has $\mathcal{O}(h^{-d})$ nodes. Thus, the combined sparse grid can achieve a high accuracy with a low number of grid points.

The general d dimensional combination technique for an order p scheme assumes an error structure of the form

$$u - u_{\boldsymbol{l}} = \sum_{k=1}^{d} \sum_{\substack{\{j_1,\dots,j_k\} \\ \subseteq\{1,\dots,d\}}} w_{j_1,\dots j_k}(.; h_{j_1}, \dots, h_{j_k}) h_{j_1}^p \cdots h_{j_k}^p$$

with multi-index $\boldsymbol{l} = (l_1, l_2, \dots, l_d)$, step sizes $\boldsymbol{h} = (2^{-l_1}, 2^{-l_2}, \dots, 2^{-l_d})$ and bounded functions $|w| \leq K$. The combination technique reads

$$u_n^s = \sum_{q=0}^{d-1} (-1)^q \binom{d-1}{q} \sum_{|\boldsymbol{l}|_1=n-q} u_{\boldsymbol{l}}$$

with a pointwise accuracy $\mathcal{O}(h^p \log_2(h^{-1})^{d-1})$, see [34]. The crucial point is the existence of such an error splitting structure. In [34], Reisinger investigated under which conditions such a splitting can be shown and notes the following properties, which have to be fulfilled:

Properties:

1. *The scheme has a truncation error of the form*

$$(A - A_h)u = \sum_{k=1}^{d} \sum_{\substack{\{j_1,\dots,j_k\} \\ \subseteq\{1,\dots,d\}}} \nu_{j_1,\dots j_k}(.; h_{j_1}, \dots, h_{j_k}) h_{j_1}^p \cdots h_{j_k}^p,$$

   where A is a differential operator (e.g. $\frac{\partial}{\partial t} - L$) and A_h its discrete finite difference approximation.
2. *The discretisation scheme has to be stable.*
3. *The initial data has to be sufficiently smooth and compatible boundary data is required, such that the mixed derivatives of required order are bounded.*

In option pricing one is in general interested in an accurate solution at the present time $t = 0$. Therefore, we intent to construct a space sparse grid at the final time slice of the numerical solution. In Section 3.2 we derived the truncation errors for the spatial discretisation, hence we have in the spatial domain

$$(L - L_h)u = \sum_{i=1}^{d} h_i^4 \nu_i(.; h_i) + \sum_{\substack{i,j=1 \\ i \neq j}}^{d} h_i h_j \nu_{i,j}(.; h_i, h_j),$$

where $\nu_i = \tau_i + \tilde{\tau}_i$ and $\nu_{i,j} = \tau_{i,j}$ with τ, $\tilde{\tau}$ given by equations (12), (14), respectively. Thus, the spatial error is of the desired form. In the case of second order finite differences it was shown in [34] that if bounded mixed derivatives $\frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} ... \partial x_d^{\alpha_d}}$ with $\alpha = (\alpha_1, ..., \alpha_d)$ and $\alpha_i \in \{0, 1, ..., 4\}$ exist, then such a splitting exists for the Poisson equation. For a fourth order scheme the analogue condition for the mixed derivatives with $\alpha_i \in \{0, 1, ..., 6\}$ was derived in [15].

6. Numerical experiments

In this Section we apply the numerical schemes to the multivariate Black-Scholes PDE

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sum_{i,j=1}^{d} \rho_{ij} \sigma_i \sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i S_j} + \sum_{i=1}^{d} r S_i \frac{\partial V}{\partial S_i} - rV = 0 \tag{24}$$

in the space-time cylinder $\Omega_d \times \Omega_t$ with $\Omega_d = [0, S_1^{\max}] \times ... \times [0, S_d^{\max}]$, $\Omega_t = [0, T]$. The volatility of the i-th asset is denoted by σ_i for $i = 1, 2, ..., d$, whereas the correlation between assets i and j is ρ_{ij}. The risk-free interest rate is given by r. The option value at the maturity T is defined by its payoff profile $g(S_1, ..., S_d)$. In the remainder of this article we restrict ourselves to basket put options with

$$g(S_1, ..., S_d) = \max\{K - \sum_{i=1}^{d} S_i, 0\}.$$

As the spatial domain of each asset S_i is truncated at $[0, S_i^{\max}]$, boundary values have to be prescribed. At $S_i = 0$ the PDE reduces to a lower dimensional PDE. Hence, we solve the lower dimensional PDE at each boundary. This is called the *natural boundary condition* [27, 33]. At the upper boundary we imply Dirichlet boundary conditions and set the option value to zero. This means that the computational domain has to be chosen large enough, such that the option is far out of the money at the upper boundary and the introduced error is negligible small.

We investigate the accuracy both in the temporal and spatial domain. Consider the vector space $\mathbb{R}^m$ with $\hat{h} = 1/m$ and discrete norms $\|v\|_2 = (\hat{h} \sum_{i=1}^{m} |v_i|^2)^{1/2}$, $\|v\|_\infty = \max\limits_{1 \leq i \leq m} |v_i|$ for $v \in \mathbb{R}^m$. Then we define the errors

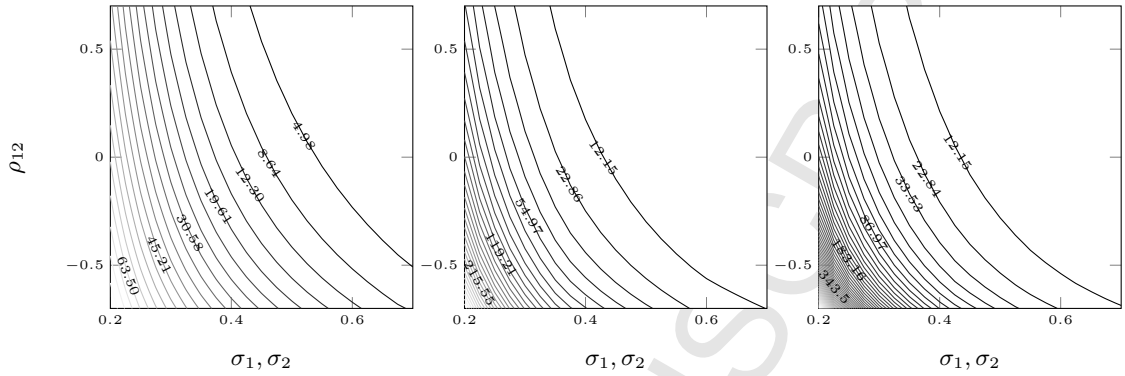
$$\text{err}_2 = \|u_h^{\Delta_t} - u_{\tilde{h}}^{\tilde{\Delta}_t}\|_2, \qquad \text{err}_\infty = \|u_h^{\Delta_t} - u_{\tilde{h}}^{\tilde{\Delta}_t}\|_\infty,$$

where $u_h^{\Delta_t}$ denotes a highly accurate reference solution with step size Δ_t in time and spatial mesh width h. The numerical approximation is denoted by $u_{\tilde{h}}^{\tilde{\Delta}_t}$ with $\tilde{\Delta}_t \to \Delta_t$ and $\tilde{h} \to h$. The solutions are compared at the final time level on the spatial grid of the reference solution. Therefore, we extend the approximation via multivariate cubic spline interpolation to the grid of the reference solution. Note, that the interpolant is fourth order accurate and hence does not have any negative effect on the rate of convergence. For the experiments in the spatial domain we employ the full grid solver to compute the reference solution in the two dimensional case. In the higher dimensional case the combination technique is used to compute the reference solution and then extended to the full grid. The experiments in the temporal domain are always performed on the full grid.

In our numerical experiments we investigate the behaviour of the sparse grid combination technique in comparison to the full grid solver. Recall that according to Section 5 the combination technique

	σ_1	σ_2	σ_3	σ_4	ρ_{12}	ρ_{13}	ρ_{14}	ρ_{23}	ρ_{24}	ρ_{34}
A	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	0.2
B	0.4	0.4	0.4	0.4	0.2	0.2	0.2	0.2	0.2	0.2
C	0.6	0.6	0.6	0.6	-0.5	0.5	-0.25	-0.25	-0.25	-0.25

Table 1: Parameter sets for numerical experiments



(a) 33^2 grid nodes. (b) 65^2 grid nodes. (c) 129^2 grid nodes.

Figure 3: Mixed fourth derivative $\|\delta_1^2\delta_2^2 u_h^{\Delta_t}\|_\infty$ for a decreasing mesh width h in the two dimensional case.

has stronger regularity requirements than the full grid solution. In the full grid case the derivatives in the truncation error need to be bounded, namely

$$\frac{\partial^6 u}{\partial x_i^6}, \quad \frac{\partial^{10} u}{\partial x_i^5 \partial x_j^5}, \quad \frac{\partial^6 u}{\partial x_i^5 \partial x_j} \quad \text{for } i, j = 1, \ldots, d \text{ and } i \neq j$$

arising in (12) and (14), respectively. The combination technique requires the mixed derivatives $\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}}$ with $\alpha_i \in \{0, 1, ..., 6\}$ to be bounded, see e.g. [15] for a fourth order approximation to the Poisson equation. These derivatives arise due to the anisotropic splitting within the combination technique and do not stem from mixed derivatives in PDE (1). A detailed discussion of the error splitting can be found in [15, 34]. In order to relate the option parameters to the expected smoothness of the solution, we compute $\|\delta_1^2\delta_2^2 u_h^{\Delta_t}\|_\infty$ for a decreasing mesh width as a measure of smoothness. Figure 3 suggests that the solution at the final time level becomes smoother for large diffusion and positive correlation. If strong diffusion is present already a resolution of 65 grid points in each dimension seems to be sufficient to capture the maximum of the derivative. For smaller diffusion a higher resolution is required to capture the maximum. If one assumes that analogue results hold for the higher derivatives, arising in the truncation errors of the full grid solver as well as in the combination technique, the theoretical asymptotic rate of convergence will be reached much faster for high diffusion with positive correlation than for small diffusion and negative correlation. Please note, that also a longer time to maturity leads to a smoother solution at $t = 0$.

In the following we consider three different test cases given in Table 1. In case A we choose a parameter set with large diffusion and a positive correlation. In the second case the diffusion coefficients are reduced. Compared to A the correlation in C is decreased, which leads, according to Figure 3, to a larger mixed fourth derivative. Thus, case A seems to be the most favorable for the combination technique, while the other two test cases are expected to be more difficult due to the reduced smoothness. If not mentioned otherwise, we choose $T = 1$, $K = 20$ and $r = 0.025$ in the numerical experiments.

6.1. Non-smooth initial data

The analysis of consistency for numerical methods typically relies on smoothness assumptions of the initial data. However, in practice and especially in financial option pricing the payoff function usually exhibits discontinuities: plain vanilla options have a discontinuity in the first derivative at

the strike price, while for example digital options have a discontinuous payoff profile. This leads to a maximal error in the at-the-money region in the numerical solution. Since the option values close to the strike price are in general of highest interest from the viewpoint of practitioners in finance this is a severe problem. In [32] several methods have been discussed to overcome this issue and to recover a high rate of accuracy.

An intuitive approach is to place more grid points in the region of interest. For example Sydow et al. [28] solve a sub-problem with $\mathcal{O}(h^{-2})$ grid nodes around the strike price to gain sixth order accuracy in space. Kreiss et al. [25] propose to smooth the initial condition. With this averaging a high rate of convergence can be recovered, while the initial condition converges to the original initial condition as the grid spacing goes to zero. This approach was successfully applied to option pricing problems in one dimension [32] and two dimensions [10]. An additional method to cope with the non-smooth initial data was given by Wahlbin [41], where the initial payoff profile undergoes an L_2 projection onto a set of basis functions. Besides these techniques in [40] grid shifting is suggested. Here the grid is sequentially shifted, such that the discontinuity falls midway between two successive grid nodes. The discrete payoff for the shifted grid reveals that this method can be interpreted as a kind of smoothing.

The latter three approaches have been investigated in [32] for one dimensional option pricing problems in the case of second order accuracy: in the numerical test all techniques showed the desired order of convergence. Furthermore, they give a brief outlook how to apply these methods to higher dimensional problems: at the current state it is not clear if grid shifting is possible for higher dimensions since the grid is not allowed to coincide with the discontinuity. In the case of the projection Wahlbin reports technical difficulties if the discontinuities do not match with the grid nodes.

In the remainder of this article we restrict ourself to smooth the initial condition according to [25] via a convolution operator. This approach can be easily extended for arbitrary dimension via a tensor product of one dimensional convolutions. Compared to the first method it can be computed during an offline phase and the smoothed data can be reused to price options with different parameter sets.

In the following we consider the d-dimensional formulation of the smoothing operator used in [10]. Let

$$\hat{\Phi}_4(\omega) = \frac{\sin^4 \frac{1}{2}\omega + \frac{2}{3}\sin^6 \frac{1}{2}\omega}{(\frac{1}{2}\omega)^4},$$

then the smoothed initial data $\tilde{g}$ is given by

$$\tilde{g}(x_1,...,x_d) = h_1^{-1}...h_d^{-1} \int_{-3h_1}^{3h_1} ... \int_{-3h_d}^{3h_d} \Phi_4(h_1^{-1}\tilde{x}_1)...\Phi_4(h_d^{-1}\tilde{x}_d) g(x_1-\tilde{x}_1,...,x_d-\tilde{x}_d) d\tilde{x}_1...d\tilde{x}_d, \qquad (25)$$

where Φ_4 is the inverse Fourier transform of $\hat{\Phi}_4(\omega)$ and g denotes the original initial condition.

6.2. European basket put options

In this Section we solve the multi-dimensional Black-Scholes equation (24). We apply the following coordinate transformations $x_i = \log(S_i)$ for $i = 1,...,d$, $\tau = T - t$, $u = e^{r\tau}V$ and obtain

$$\frac{\partial u}{\partial \tau} - \frac{1}{2}\sum_{i,j=1}^{d} \rho_{ij}\sigma_i\sigma_j \frac{\partial^2 u}{\partial x_i \partial x_j} - \sum_{i=1}^{d} \left(r - \tfrac{1}{2}\sigma_i^2\right) \frac{\partial u}{\partial x_i} = 0.$$

The payoff transforms to $g(x_1,...,x_d) = \max\{K - \sum_{i=1}^{d} e^{x_i}, 0\}$. The HOC formulation of F_i for $i = 1,...,d$ according to Section 3.2 can be derived by inserting $a_i := \frac{1}{2}\sigma_i^2$ and $c_i := r - \frac{1}{2}\sigma_i^2$ into equation (11). In a first experiment we investigate the spatial accuracy of the uniform full grid, as well as of the sparse grid, for a basket put with two underlying assets. In Figure 4 (a) - (c) the full grid solution exhibits an accuracy close to order four. However, compared to case A one observes a slightly lowered rate of convergence in the more difficult cases B and C. Due to the stronger regularity requirements of the sparse grid combination technique, the rate of convergence

shows a greater sensitivity to the smoothness of the solution. Nevertheless, the sparse grid is more accurate per grid node than the full grid in two of the three test cases: only in case C the sparse grid is outperformed.

In Figure 5 all numerical schemes show the desired rate of convergence in the time domain. The HDO scheme exhibits order one in time, while the HCS, HMCS and HHV scheme show order two. According to the stability results in Section 4 we see a stable behaviour.

In Figures 6, 8 we numerically analyse the performance for basket options with three assets. The full grid solution in Figure 6 states convergence close to order four. Similar to the two dimensional problems the rate of convergence of the sparse grid solution is more sensitive to the parameter changes in cases B and C than the full grid solver. If we consider the accuracy per grid node, the sparse grid error is lower than the full grid error. Figure 7 indicates that the run-time per node is only slightly higher for the combination technique than for the full grid solver and thus the sparse grid is more efficient in the asymptotic. Both plots show that the time increases with an order of approximately one as one would expect due to the ADI time stepping.

For the parameter choices A-C there is no convection dominance ($|\hat{p}| \geq 36/31$). Hence, we expect from the stability region plots in Figure 1 a stable behaviour. The numerical experiments in Figure 8 confirm this and the error decreases monotonically with a rate of accuracy according to theory.

Remark:
During our numerical tests the initial condition is smoothed according to the convolutions described above. Düring and Heuer [10] suggest to only smooth the grid points around the discontinuity to reduce the computational workload. Doing so the full grid performed according to the theoretical results, but the sparse grid solution showed oscillations near the discontinuity in our numerical experiments. This issue could be cured by smoothing all grid points. In the case of the above given payoff an analytical solution to the integral (25) is available, if the domain of integration does not intersect the discontinuity. Thus, smoothing the initial condition on the entire grid does not introduce a large additional computational effort. In order to smooth the initial condition for our numerical experiments, we either solve the integral analytically if possible or use the Matlab® built-in routine *integral*2, *integral*3, respectively.

6.3. Powered European basket options

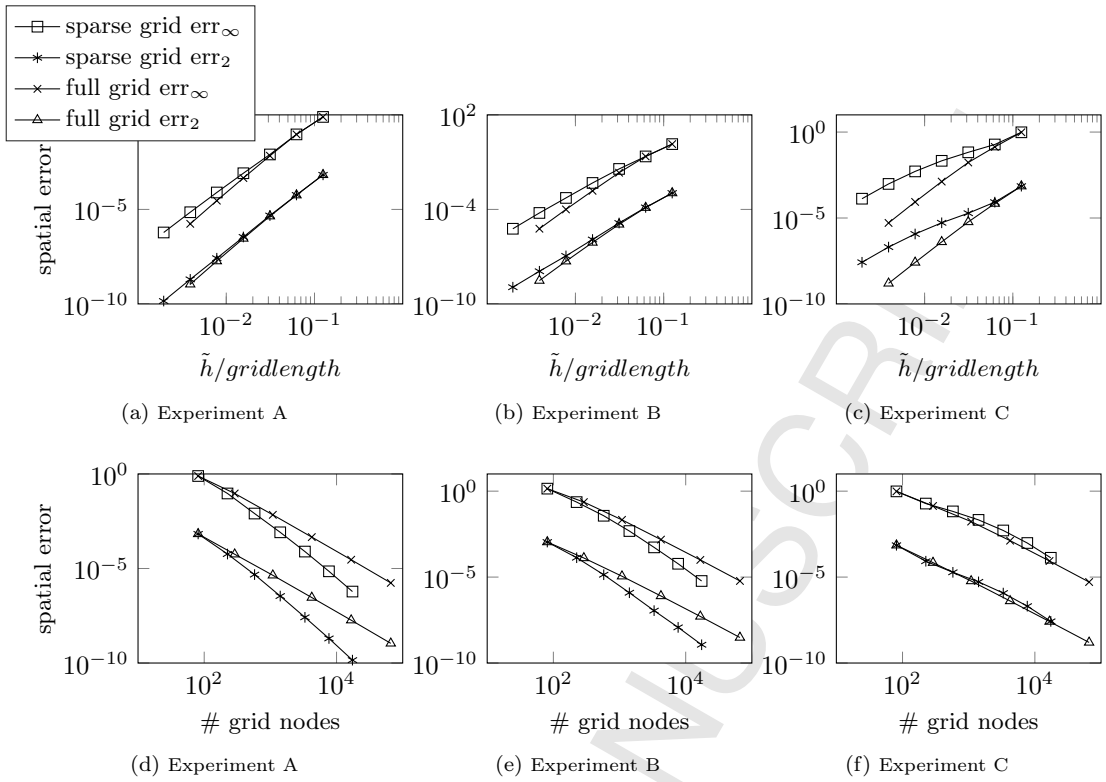
In this Section we compute powered European basket put options with three and four underlying assets. The power parameter $q \in \mathbb{N}$ allows to control the smoothness of the payoff profile and of the solution. Let the payoff under logarithmic transformed coordinates be given by

$$g(x_1, ..., x_d) = \max\{K - \sum_{i=1}^{d} e^{x_i}, 0\}^q,$$

then the initial condition fulfills $g \in C^{q-1}$. In the following we investigate the influence of the regularity on the rate of convergence.

Figure 9 shows the result for $q = 6$, $q = 9$ with three spatial dimensions and strike price $K = 1$. The sparse grid solution shows a significant better performance for a smoother initial condition. In contrast, the full grid solution shows no noticeable improvement. This underlines the stronger regularity requirements of the combination technique compared to standard full grid solvers. In Figure 10 we compute the error of the sparse grid for $q = 6$ and $q = 9$ for a four dimensional problem with strike price $K = 1$. The error decay again shows a strong sensitivity towards the smoothness of the initial condition. Please note, that we did not compute the full grid solution here, since due to the high dimensionality the memory constraints make it impossible to compute a full grid solution with a reasonable large number of grid nodes per coordinate direction on our test machine.

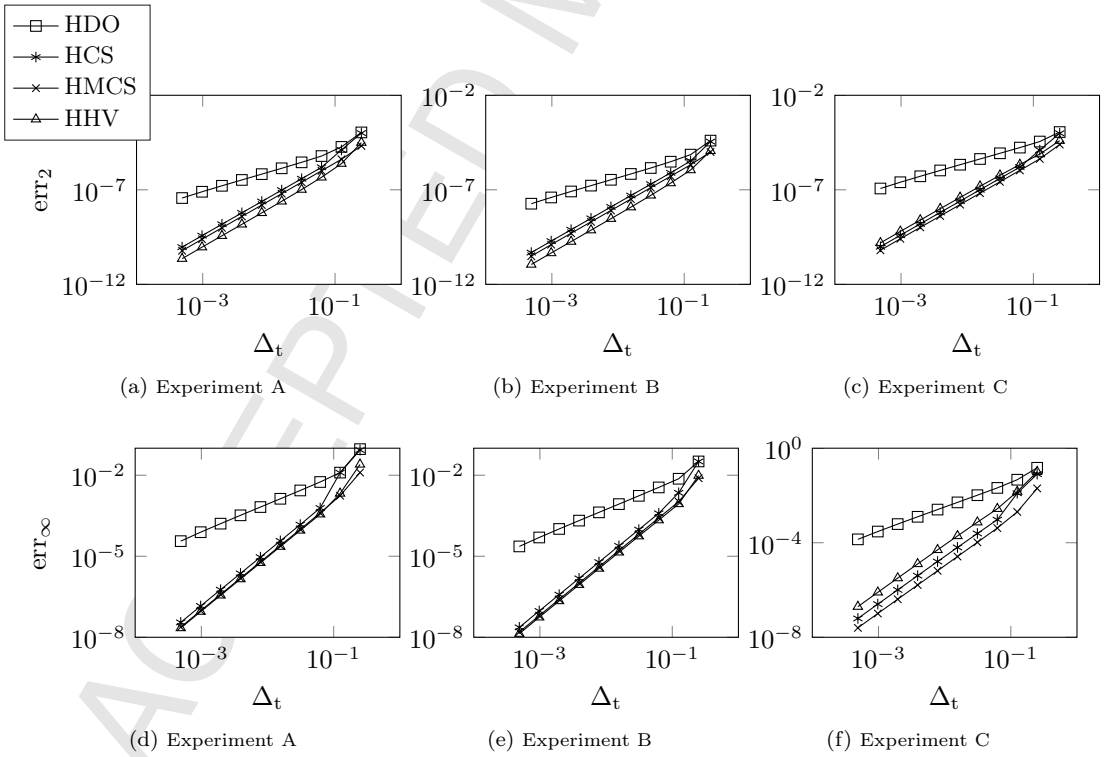
In Figure 11 we compare the rate of convergence in time of the four HO ADI schemes. The HDO and HCS scheme exhibit order one, while the HMCS and HHV show order two. Note, that the rate of the HCS scheme is reduced to order one due to $\theta \neq 1/2$. All schemes show a stable behaviour.



(a) Experiment A (b) Experiment B (c) Experiment C

(d) Experiment A (e) Experiment B (f) Experiment C

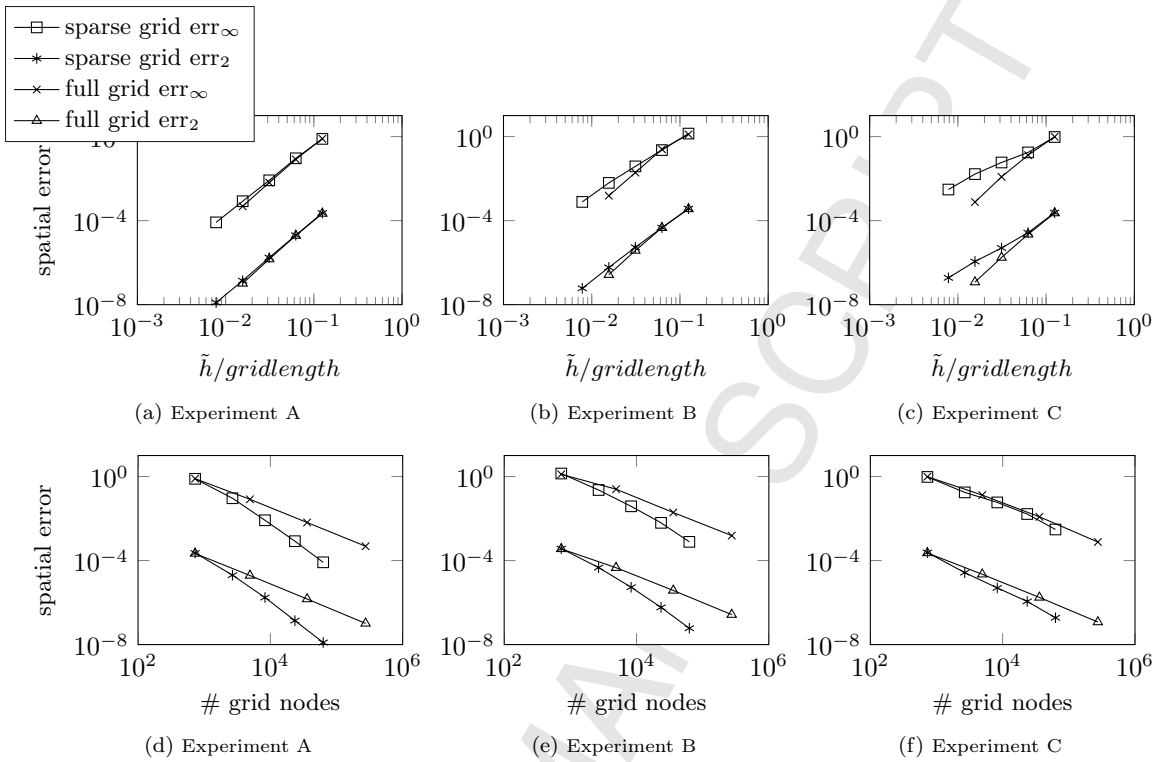
Figure 4: 2d spatial error for $\tilde{h} \to h = 2^{-9} \cdot gridlength$. In case of the sparse grid solution, the finest steps size is used. Order err_∞ on sparse grid: A: 3.39, B: 2.98, C: 2.05. Order err_2 on sparse grid: A: 3.70, B: 3.34, C: 2.34. Order err_∞ on full grid: A 3.79, B: 3.62, C: 3.52. Order err_2 on full grid: A: 3.85, B: 3.71, C: 3.77.



(a) Experiment A (b) Experiment B (c) Experiment C

(d) Experiment A (e) Experiment B (f) Experiment C

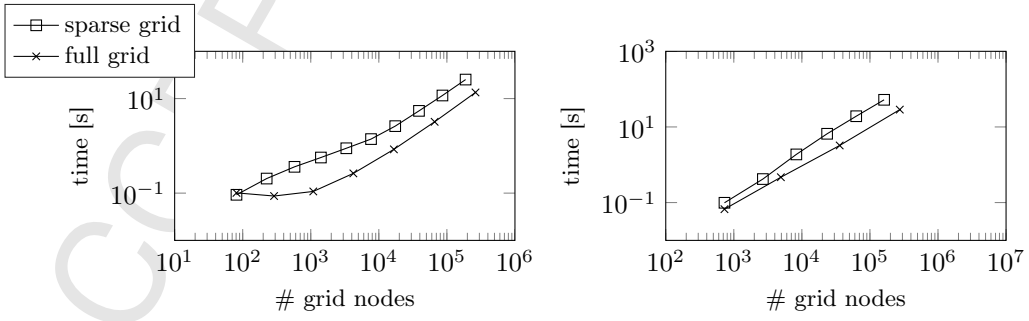
Figure 5: 2d temporal error for $\tilde{\Delta}_t \to \Delta_t = 2^{-14}$. The following values of θ were used: HDO $\theta = 0.5$, HCS $\theta = 0.5$, HMCS $\theta = 0.334$, HHV $\theta = 0.79$. The spatial discretisation is computed on a grid with 129 nodes in both coordinate directions. Order err_2: A 1.18 (HDO), 2.15 (HCS), 2.02 (HMCS), 2.16 (HHV); B 1.14 (HDO), 2.09 (HCS), 2.02 (HMCS) 2.14 (HHV); C 1.06 (HDO), 2.15 (HCS), 2.03 (HMCS), 1.97 (HHV). Order err_∞: A 1.15 (HDO), 2.28 (HCS), 2.06 (HMCS), 2.14 (HHV); B 1.09 (HDO), 2.15 (HCS), 2.05 (HMCS), 2.07 (HHV); C 1.07 (HDO), 2.19 (HCS), 2.10 (HMCS), 2.05 (HHV).



(a) Experiment A (b) Experiment B (c) Experiment C

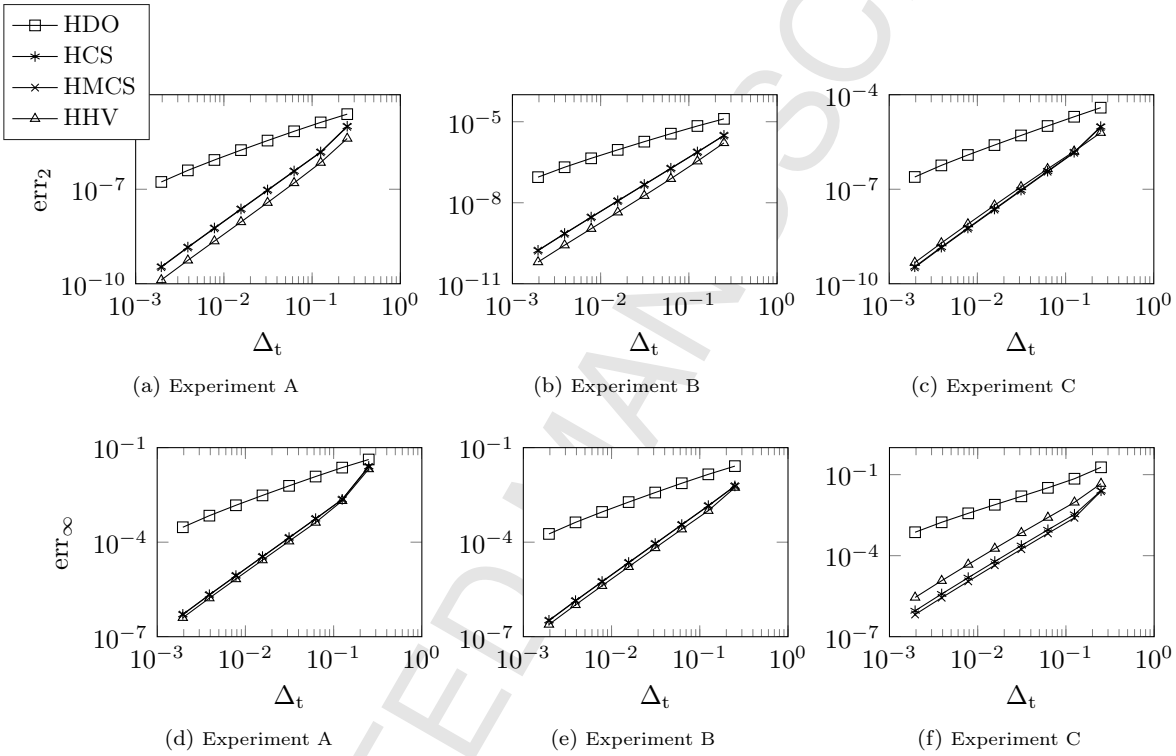
(d) Experiment A (e) Experiment B (f) Experiment C

Figure 6: 3d spatial error: the sparse grid solution at level 14 is used as a reference solution and evaluated on a full grid with 129^3 grid points. Order err_∞ on sparse grid: A: 3.32, B: 2.68, C: 2.01. Order err_2 on sparse grid: A: 3.55, B: 3.15, C: 2.52. Order err_∞ on full grid: A 3.57, B: 3.28, C: 3.44. Order err_2 on full grid: A: 3.69, B: 3.48, C: 3.64.



(a) 2-d (b) 3-d

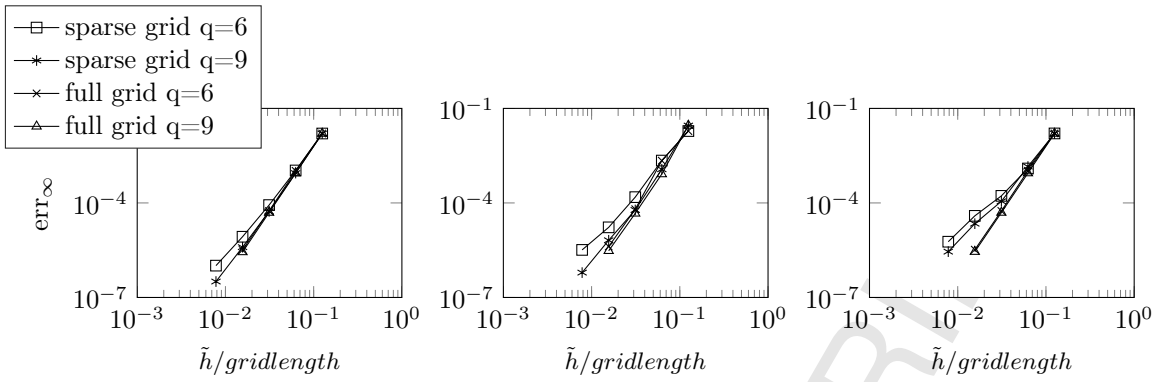
Figure 7: Number of grid nodes versus computation time in the case of 100 time steps in the HCS scheme.



(a) Experiment A (b) Experiment B (c) Experiment C

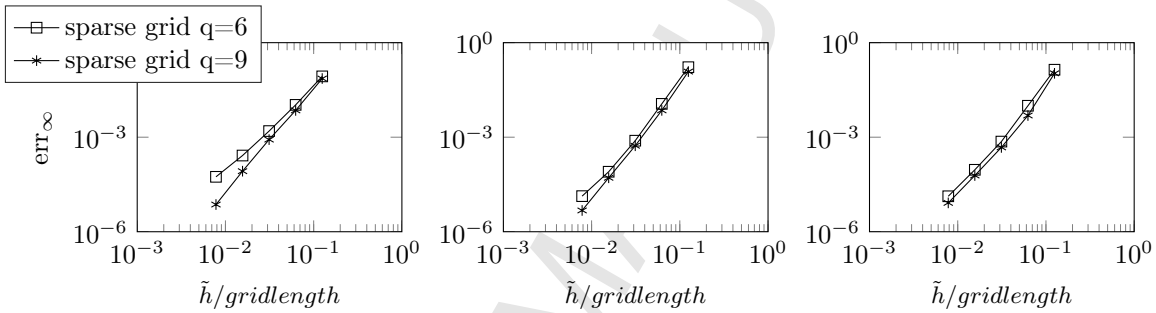
(d) Experiment A (e) Experiment B (f) Experiment C

Figure 8: 3d temporal error for $\tilde{\Delta}_t \to \Delta_t = 2^{-11}$. The following values of θ were used: HDO $\theta = 0.67$, HCS $\theta = 0.5$, HMCS $\theta = 0.462$, HHV $\theta = 0.79$. The spatial discretisation is computed on a grid with 65 nodes in all coordinate directions. Order err_2: A 1.03 (HDO), 2.07 (HCS), 2.07 (HMCS), 2.10 (HHV); B 1.02 (HDO), 2.01 (HCS), 2.01 (HMCS) 2.08 (HHV); C 1.04 (HDO), 2.07 (HCS), 2.07 (HMCS), 1.95 (HHV). Order err_∞: A 1.02 (HDO), 2.14 (HCS), 2.13 (HMCS), 2.15 (HHV); B 1.02 (HDO), 2.02 (HCS), 2.02 (HMCS), 2.03 (HHV); C 1.11 (HDO), 2.05 (HCS), 2.08 (HMCS), 1.97 (HHV).



(a) Experiment A, 3d (b) Experiment B, 3d (c) Experiment C, 3d

Figure 9: Spatial error for options with three underlying assets and powered payoff. The sparse grid solution at level 14 is used as a reference solution and evaluated on a full grid with 129^3 grid points. Order on sparse grid: A $q = 6$: 3.48, $q = 9$: 3.90, B $q = 6$: 3.21, $q = 9$: 3.74, C $q = 6$ 2.78, $q = 9$: 3.09. Order on full grid: A $q = 6$: 4.10 , $q = 9$ 4.16:, B $q = 6$: 4.16, $q = 9$: 4.30, C $q = 6$: 4.10 , $q = 9$: 4.15.



(a) Experiment A, 4d (b) Experiment B, 4d (c) Experiment C, 4d

Figure 10: Spatial error for options with four underlying assets and powered payoff. The sparse grid solution at level 17 is used as a reference solution and evaluated on a full grid with 33^4 grid points. Order: A $q = 6$: 2.65, $q = 9$: 3.30, B $q = 6$: 2.45, $q = 9$: 2.95, C $q = 6$: 2.60, $q = 9$: 2.83.

6.4. European basket options with space-dependent coefficients

In this Section we consider a basket option with space-dependent coefficients. Therefore, we assume a basket of two underlyings, driven by the system of stochastic differential equations

$$dS_{1,t} = \mu_1 S_{1,t} dt + S_1 \sigma_1(S_{1,t}) dW_{1,t},$$
$$dS_{2,t} = \mu_2 S_{2,t} dt + S_2 \sigma_2(S_{2,t}) dW_{2,t},$$

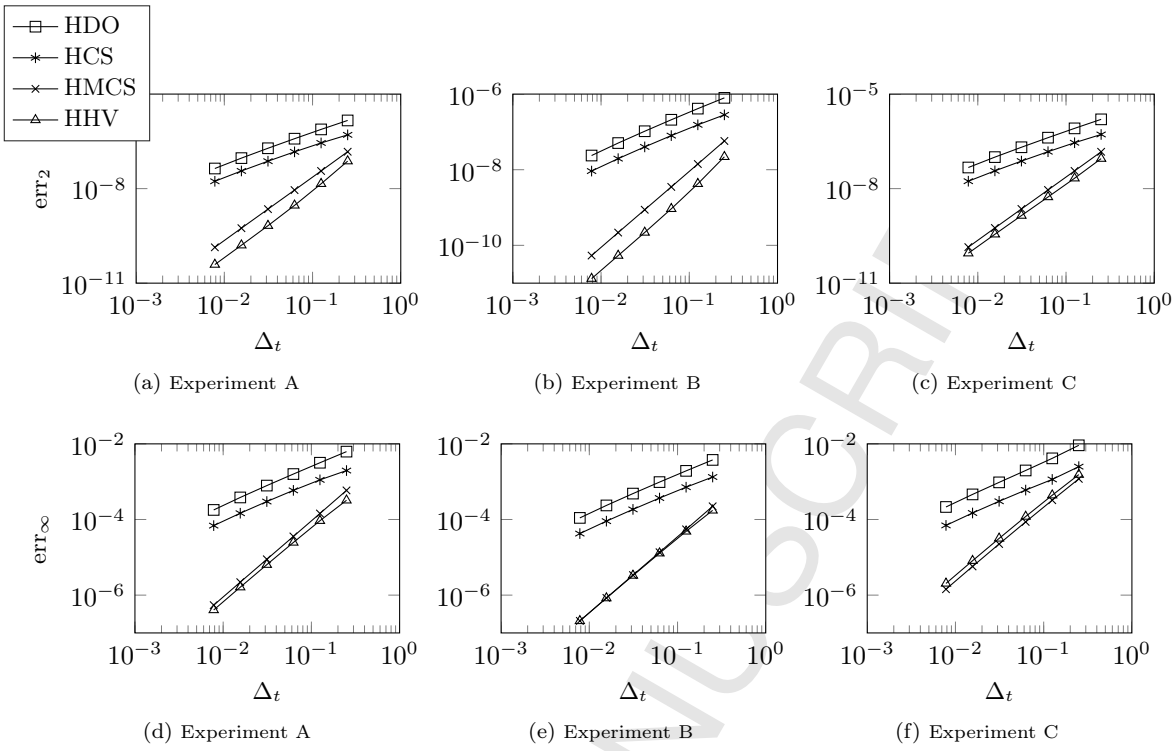
where the volatilities depend on the underlying and the Wiener increments are correlated, such that $\mathbb{E}(dW_{1,t} dW_{2,t}) = \rho_{12} dt$. Let $V(S_1, S_2, t)$ denote the option value, then by standard arguments, application of the Ito formula and considering the riskless portfolio, the option price dynamics can be described via the PDE

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma_1^2(S_1) S_1^2 \tfrac{\partial^2 V}{\partial S_1^2} + \tfrac{1}{2}\sigma_2^2(S_2) S_2^2 \tfrac{\partial^2 V}{\partial S_2^2} + \rho_{12}\sigma_1(S_1)\sigma_2(S_2) S_1 S_2 \tfrac{\partial^2 V}{\partial S_1 \partial S_2} + rS_1 \tfrac{\partial V}{\partial S_1} + rS_2 \tfrac{\partial V}{\partial S_2} - rV = 0. \tag{26}$$

In the remainder, we assume the volatilities to follow a cubic polynomial

$$\sigma_i(S_i) = a_{i,0} + a_{i,1} S_i + a_{i,2} S_i^2 + a_{i,3} S_i^3$$

for $i = 1, 2$. Note, that also more sophisticated models for the volatilities can be used if they are sufficiently smooth. In order to derive the numerical method to solve PDE (26) we apply the same

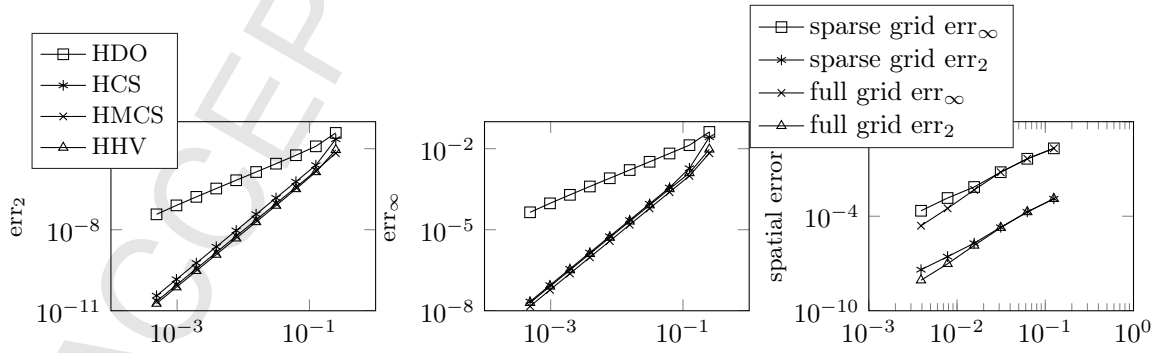


(a) Experiment A (b) Experiment B (c) Experiment C

(d) Experiment A (e) Experiment B (f) Experiment C

Figure 11: 4d temporal error for $\tilde{\Delta}_t \to \Delta_t = 2^{-10}$ and $q = 1$. The following values of θ were used: HDO $\theta = 0.85$, HCS $\theta = 0.663$, HMCS $\theta = 0.5935$, HHV $\theta = 0.79$. The first three choices are the lower bounds derived for pure diffusion equations, while the latter one is the bound for 2-d convection-diffusion equations without mixed derivative terms. The spatial discretisation is computed on a grid with 33 nodes in each coordinate direction. Order err_2 : A 1.01 (HDO), 0.98 (HCS), 2.01 (HMCS), 2.17 (HHV); B 1.01 (HDO), 0.99 (HCS), 2.01 (HMCS) 2.13 (HHV); C 1.01 (HDO), 0.99 (HCS), 2.01 (HMCS), 1.99 (HHV). Order err_∞ : A 1.02 (HDO), 0.97 (HCS), 2.01 (HMCS), 1.93 (HHV); B 1.02 (HDO), 0.99 (HCS), 2.01 (HMCS), 1.08 (HHV); C 1.08 (HDO), 1.02 (HCS), 1.94 (HMCS), 1.91 (HHV).

transformations as in Section 6.2. Figure 12 shows the numerical results for a two dimensional basket option, whose assets are correlated with $\rho_{12} = 0.5$. The temporal error decay in (a), (b) is monotone and of the desired order. In the spatial domain the sparse grid has a slightly lower rate of convergence than the full grid solution. The numerical results are in line with the theoretical findings and show a similar behaviour like in the frozen coefficients case.



(a) error in time (b) error in time (c) spatial error

Figure 12: Numerical example with space dependent coefficients $\sigma_1(S) = 0.5 - 0.008 \cdot S + 0.001 \cdot S^2 - 0.36 \cdot 10^{-6} \cdot S^3$, $\sigma_2(S) = 0.4 - 0.01 \cdot S + 0.0002 \cdot S^2 - 0.87 \cdot 10^{-6} \cdot S^3$. Order (a) err_2: 1.07 (HDO), 2.07 (HCS), 2.02 (HMCS), 2.06 (HHV). Order (b) err_∞: 1.06 (HDO), 2.14 (HCS), 2.05 (HMCS), 2.04 (HHV). Order (c) err_2: 3.04 (sparse grid), 3.50 (full grid); err_∞ 2.68 (sparse grid), 3.33 (full grid).

7. Conclusion

In this paper we introduced a fourth-order spatial discretisation for multi-dimensional convection-diffusion equations with space dependent coefficients. In the time domain we applied ADI schemes to achieve up to second order accuracy. Due to the ADI splitting the spatial discretisation could be decomposed in such a way, that all implicitly treated approximations are defined on the compact stencil, which results in tridiagonal linear systems. These systems can be solved very efficiently in linear run-time. Broad stencils only occur in the explicit steps of the algorithm. In the two dimensional case we were able to show that the stability regions coincide with the standard central second order ADI schemes. For the HDO and HCS scheme sufficient conditions on θ, ensuring unconditional stability, could be proven. For the HMCS scheme a necessary lower bound on θ was found, while for the HHV scheme a lower bound on θ for vanishing correlation was derived. In the three and four dimensional case the proposed schemes showed good stability properties. In order to reduce the computational complexity, we applied the sparse grid combination technique in the spatial domain. This allows to lower the number of grid nodes and to reduce the effects of the curse of dimensionality. In terms of accuracy per grid node the sparse grid performed better than the full grid solution in two of the three test cases. Only for negative correlation in experiment C the full grid showed a better accuracy in the 2-d case. Based on experiments with a powered payoff the regularity requirements of the combination technique could be compared to the ones of full grid solvers. The sparse grid showed a significant better performance with sufficiently smooth data, which underlines the stronger smoothness requirements of the combination technique.

Acknowledgement

The work of the authors was partially supported by the European Union in the FP7-PEOPLE-2012-ITN Programme under Grant Agreement Number 304617 (FP7 Marie Curie Action, Project Multi-ITN STRIKE – Novel Methods in Computational Finance).
Further the authors acknowledge partial support from the bilateral German-Spanish Project "HiPeCa High Performance Calibration and Computation in Finance" (Grant No. 57049700), Programme Acciones Conjuntas Hispano-Alemanas, financed by DAAD.

8. References

[1] I.J.D. Craig and A.D. Sneyd. An Alternating-Direction Implicit scheme for parabolic equations with mixed derivatives. *Comput. Math. Appl.*, 16(4):341–350, 1988.

[2] G. Berikelashvili, M. Gupta, and M. Mirianashvili. Convergence of fourth order compact finite difference schemes for three–dimensional convection-diffusion equations. *SIAM J. Numer. Anal.*, 45:443–455, 2007.

[3] P. Brian. A finite-difference method of high-order accuracy for the solution of three-dimensional transient heat conduction problems. *AIChE Journal*, 7:367–370, 1961.

[4] H. Bungartz and M. Griebel. Sparse grids. *Cambridge University Press*, pages 1–123, 2004.

[5] J. Douglas and H. Rachford. On the numerical solution of heat conduction problems in two and three space variables. *Trans. Amer. Math. Soc.*, 82:421–439, 1956.

[6] J. J. Douglas. Alternating Direction methods for three space variables. *Numer. Math.*, 4(1):41–63, 1962.

[7] B. Düring and M. Fournié. High-order compact finite difference scheme for option pricing in stochastic volatility models. *J. Comput. Appl. Math.*, 236(17):4462–4473, Nov. 2012.

[8] B. Düring, M. Fournie, and C. Heuer. High-order compact finite difference schemes for option pricing in stochastic volatility models on non-uniform grids. *J. Comput. Appl. Math.*, 271:247–266, 2014.

[9] B. Düring, M. Fournié, and A. Rigal. High-order ADI schemes for convection-diffusion equations with mixed derivative terms. In *Spectral and High Order Methods for Partial Differential Equations - ICOSAHOM 2012*, volume 95, pages 217–226. Springer International Publishing, 2014.

[10] B. Düring and C. Heuer. High-order compact schemes for parabolic problems with mixed derivatives in multiple space dimensions. *SIAM J. Numer. Anal*, 53(5):2113–2134, 2015.

[11] B. Düring and J. Miles. High-order ADI scheme for option pricing in stochastic volatility models. *Preprint*, 2015.

[12] M. Griebel, M. Schneider, and C. Zenger. A Combination Technique for the solution of sparse grid problems. *IMACS Elsevier, Iterative Methods in Linear Algebra*, 16:263–281, 1992.

[13] M. Griebel and V. Thurner. The efficient solution of fluid dynamics problems by the combination technique. *Int. J. Numer. Meth. Heat and Fluid Flow*, 5(3):251–269, 1995.

[14] T. Haentjens and K. J. in't Hout. ADI finite difference schemes for the Heston-Hull-White PDE. *J. Comput. Fin.*, 16:83–110, 2012.

[15] C. Hendricks, M. Ehrhardt, and M. Günther. Error splitting preservation for high order finite difference schemes in the combination technique. *to appear in Numer. Math. Theor. Meth. Appl.*, 2016.

[16] C. Hendricks, M. Ehrhardt, and M. Günther. High-order ADI schemes for diffusion equations with mixed derivatives in the combination technique. *Appl. Numer. Math.*, 101:36–52, 2016.

[17] W. Hundsdorfer. Accuracy and stability of splitting with stabilizing corrections. *Appl. Numer. Math.*, 42(1-3):213–233, 2002.

[18] W. Hundsdorfer and J. Verwer. *Numerical solution of time-dependent advection-diffusion-reaction equations*, volume 33. Springer, 2003.

[19] K. J. in't Hout and S. Foulon. ADI finite difference schemes for option pricing in the Heston model with correlation. *Int. J. Numer. Anal. Mod.*, 7:303–320, 2010.

[20] K. J. in't Hout and C. Mishra. Stability of the Modified Craig-Sneyd scheme for two-dimensional convection-diffusion equations with mixed derivative terms. *Math. Comp. Simul.*, 81:2540–2548, 2011.

[21] K. J. in't Hout and C. Mishra. Stability of ADI schemes for multidimensional diffusion equations with mixed derivative terms. *Appl. Numer. Math.*, 74:83–94, 2013.

[22] K. J. in't Hout and B. D. Welfert. Stability of ADI schemes applied to convection-diffusion equations with mixed derivative terms. *Appl. Numer. Math.*, 57(1):19–35, Jan. 2007.

[23] 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. *Appl. Numer. Math.*, 59(3-4):677–692, 2009.

[24] S. Karaa and J. Zhang. High-order ADI method for solving unsteady convection-diffusion problems. *J. Comput. Phys.*, 198(1):1–9, 2004.

[25] H. O. Kreiss, V. Thomée, and O. Widlund. Smoothing of initial data and rates of convergence for parabolic difference equations. *Commun. Pure Appl. Math.*, 23(2):241–259, 1970.

[26] D. Lanser, J. Blom, and J. Verwer. Time integration of the shallow water equations in spherical geometry. *J. Comput. Phys.*, 171:373–393, 2001.

[27] C. Leentvaar and C. Oosterlee. On coordinate transformation and grid stretching for sparse grid pricing of basket options. *J. Comput. Appl. Math,*, 222(1):193–209, 2008. Special Issue: Numerical Methods in Finance.

[28] G. Linde, J. Persson, and L. V. Sydow. High-order adaptive space discretizations for the BlackScholes equation. Master's thesis, 2005.

[29] C. Mishra. A new stability result for the Modified Craig-Sneyd scheme applied to two-dimensional convection-diffusion equations with mixed derivatives. *Appl. Math. Comput.*, 285(C):41–50, July 2016.

[30] A. R. Mitchell and G. Fairweather. Improved forms of the Alternating Direction methods of Douglas, Peaceman, and Rachford for solving parabolic and elliptic equations. *Numer. Math.*, 6(1):285–292, 1964.

[31] D. W. Peaceman and J. H. H. Rachford. The numerical solution of parabolic and elliptic differential equations. *J. Soc. Indust. Appl. Math.*, 3(1):28–41, 1955.

[32] D. M. Pooley, K. R. Vetzal, and P. A. Forsyth. Convergence remedies for non-smooth payoffs in option pricing. *J. Comput. Fin.*, 6(4):25–40, 2003.

[33] C. Reisinger. *Numerische Methoden für hochdimensionale parabolische Gleichungen am Beispiel von Optionspreisaufgaben.* PhD thesis, Ruprecht-Karls-Universität Heidelberg, 2004.

[34] C. Reisinger. Analysis of linear difference schemes in Sparse Grid Combination Technique. *IMA J. Numer. Anal.*, 33(2):544–581, 2013.

[35] Schiekofer. *Die Methode der Finiten Differenzen auf dünnen Gittern zur Lösung elliptischer und parabolischer partieller Differentialgleichungen.* PhD thesis, Universität Bonn, 1999.

[36] R. Seydel. *Tools for computational finance.* Springer, 2006.

[37] S. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Dokl. Akad. Nauk SSSR*, 148:1042–1045, 1963.

[38] W. Spotz. *High compact finite difference schemes for computational mechanics.* PhD thesis, University of Texas, Austin, 1995.

[39] W. Spotz and G. Carey. High-order compact scheme for the steady stream-function vorticity equations. *Int. J. for Numer. Meth. in Engin.*, 38:3497–3512, 1995.

[40] D. Tavella and C. Randall. *Pricing financial instruments: the finite difference method.* Wiley, New York, 2000.

[41] L. B. Wahlbin. A remark on parabolic smoothing and the finite element method. *SIAM J. Numer. Anal*, 17(1):33–38, 1980.

[42] C. Zenger. Sparse grids. Technical report, Institut für Informatik, Technische Universität München, Oct. 1990.