Universiteit Antwerpen

Operator splitting schemes for American options under the two-asset Merton jump-diffusion model

Operator splitting schemes for American options under the two-asset Merton jump-diffusion model

Lynn Boen* and Karel J. in 't Hout*

December 17, 2019

Abstract

This paper deals with the efficient numerical solution of the two-dimensional partial integro-differential complementarity problem (PIDCP) that holds for the value of American-style options under the two-asset Merton jump-diffusion model. We consider the adaptation of various operator splitting schemes of both the implicit-explicit (IMEX) and the alternating direction implicit (ADI) kind that have recently been studied for partial integro-differential equations (PIDEs) in [3]. Each of these schemes conveniently treats the nonlocal integral part in an explicit manner. Their adaptation to PIDCPs is achieved through a combination with the Ikonen–Toivanen splitting technique [14] as well as with the penalty method [32]. The convergence behaviour and relative performance of the acquired eight operator splitting methods is investigated in extensive numerical experiments for American put-on-the-min and put-on-the-average options.

1 Introduction

The flexibility of American-style options, which allows the holder to exercise at any time up to and including maturity, renders them popular financial contracts. This flexibility poses a challenge in the numerical valuation of these options via partial differential equations (PDEs), since the early exercise feature leads to a nonlinear free boundary problem. When the underlying asset price process exhibits jumps, this free boundary problem forms a partial integro-differential complementarity problem (PIDCP) where the integral part is nonlocal. In addition, when there are multiple underlying assets, the PIDCP is multidimensional. The present paper is concerned with the efficient numerical solution of this advanced type of problems.

Classical methods for solving complementarity problems that arise in the valuation of American-style options are the Brennan–Schwartz algorithm [4] and the projected successive over-relaxation method. As it turns out, however, these two methods are of limited practical use as the former is only applicable under certain, restrictive conditions and the latter, iterative method converges too slowly. Over the recent years, a variety of effective numerical methods has been developed in the literature for P(I)DCPs that model American option values under various asset price processes. We provide a brief and nonexhaustive overview.

Clarke & Parrott [5] considered the two-dimensional PDCP for American option values under the Heston model and applied a multigrid method. This approach was next investigated in for example Oosterlee [24] and Toivanen & Oosterlee [31].

---

*Department of Mathematics, University of Antwerp, Middelheimlaan 1, B-2020 Antwerp, Belgium. Email: {lynn.boen,karel.inthout}@uantwerpen.be.

Zvan, Forsyth & Vetzal [32] proposed the penalty method for solving the Heston PDCP. The properties of this method were rigorously analyzed in Forsyth & Vetzal [9] for the one-dimensional Black–Scholes PDCP. The penalty method was generalized by d'Halluin et al. [7, 8] to one-dimensional PIDCPs for American option values, in particular under the Merton and Kou models, and subsequently applied by Clift & Forsyth [6] to two-dimensional PIDCPs. Here a fixed-point iteration is used to efficiently handle the integral part.

Ikonen & Toivanen [14] introduced an alternative approach for solving the one-dimensional Black–Scholes PDCP. Here the PDCP is reformulated by means of an auxiliary variable that facilitates, in each step of a given temporal discretization scheme, an effective splitting between the PDE part and the early exercise constraint. This IT splitting technique has next been employed in [15] for the Heston PDCP and in [30] for the Kou PIDCP. For treating the integral part in the latter case, an iterative method has been applied that is similar to a fixed-point iteration.

Haentjens et al. [10, 11] considered the Heston PDCP and combined alternating direction implicit (ADI) schemes for directional splitting of PDEs with the IT splitting technique for the early exercise constraint, defining the so-called ADI-IT methods. These methods were next studied in [20] for the one- and two-dimensional Black–Scholes PDCPs, where also a useful interpretation of this combined splitting approach was provided.

Complementary to this, the adaptation of ADI schemes to two-dimensional partial integro-differential equations (PIDEs) has recently been investigated by in 't Hout & Toivanen [19]. Here three novel adaptations of the well-established modified Craig–Sneyd (MCS) scheme [21] were analyzed and applied for the valuation of European options under the Bates model, where the mixed derivative term and the integral part are conveniently treated in an explicit fashion.

Boen & in 't Hout [3] subsequently studied seven operator splitting schemes of both the implicit-explicit (IMEX) and the ADI kind in the application to the two-dimensional Merton PIDE, where the two-dimensional integral part is always handled explicitly. It was concluded that, among the schemes considered, the adaptation introduced in [19] of the MCS scheme where the integral part is dealt with in a two-step Adams–Bashforth manner is preferable.

Finally, Heidarpour-Dehkordi & Christara [12] recently proposed an adaptation of the MCS scheme for the two-dimensional Black–Scholes PDCP by means of the penalty method and applied it to value American spread options.

The main aim of the present paper is to introduce and investigate adaptations of IMEX and ADI operator splitting schemes for the two-dimensional Merton PIDCP. We extend all second-order schemes from [3] by using the IT splitting technique and next study their performance in ample numerical experiments. In addition, we consider two penalty type methods that were proposed in [7] and [12].

This paper is organized as follows. In Section 2 the PIDCP is given that holds for American-style option values under the two-asset Merton jump-diffusion model. Section 3 describes the spatial discretization of this PIDCP. Section 4 deals with the temporal discretization of the obtained semidiscrete PIDCP and defines the adaptation of the pertinent IMEX and ADI schemes considered in [3] by means of IT splitting, which will be applied in an iterative manner. Also the two penalty type methods from [7] and [12] are formulated in this section. In Section 5 extensive numerical experiments are presented that yield detailed insight into the temporal convergence behaviour of all operator splitting methods from Section 4 and their relative performance. Section 6 gives our conclusions.

2   PIDC problem

The PIDE for the value $v = v(s_1, s_2, t)$ of a European-style option with maturity time $T > 0$ under the two-asset Merton jump-diffusion model is given by

$$\frac{\partial v}{\partial t} = \mathcal{D}v + \mathcal{J}v, \tag{2.1}$$

with differential and integral operators

$$\mathcal{D}v = \tfrac{1}{2}\sigma_1^2 s_1^2 \frac{\partial^2 v}{\partial s_1^2} + \rho\sigma_1\sigma_2 s_1 s_2 \frac{\partial^2 v}{\partial s_1 \partial s_2} + \tfrac{1}{2}\sigma_2^2 s_2^2 \frac{\partial^2 v}{\partial s_2^2} + (r - \lambda\zeta_1)s_1\frac{\partial v}{\partial s_1} + (r - \lambda\zeta_2)s_2\frac{\partial v}{\partial s_2} - (r + \lambda)v,$$

$$\mathcal{J}v = \lambda \int_0^\infty \int_0^\infty v(s_1 y_1, s_2 y_2, t) f(y_1, y_2) \mathrm{d}y_1 \mathrm{d}y_2.$$

Here $0 < t \leqslant T$ and $s_q > 0$ $(q = 1, 2)$ represents the price of asset q at time $\tau = T - t$. Next,

- r is the risk-free interest rate,
- σ_q $(q = 1, 2)$ is the volatility of asset q, conditional on the event that no jumps occur,
- ρ is the correlation between the two underlying standard Brownian motions,
- λ is the jump intensity of the underlying Poisson arrival process,
- ζ_q $(q = 1, 2)$ is the expected relative jump size for asset q.

Function f is the probability density function of a bivariate lognormal distribution,

$$f(y_1, y_2) = \frac{1}{2\pi\delta_1\delta_2\sqrt{1 - \widehat{\rho}^2}\, y_1 y_2} \exp\left(-\frac{\left(\frac{\ln(y_1) - \gamma_1}{\delta_1}\right)^2 + \left(\frac{\ln(y_2) - \gamma_2}{\delta_2}\right)^2 - 2\widehat{\rho}\left(\frac{\ln(y_1) - \gamma_1}{\delta_1}\right)\left(\frac{\ln(y_2) - \gamma_2}{\delta_2}\right)}{2(1 - \widehat{\rho}^2)}\right)$$

for $y_1 > 0$, $y_2 > 0$. Here γ_q and δ_q $(q = 1, 2)$ and $\widehat{\rho}$ are given real constants that can be interpreted as, respectively, the mean and standard deviation and correlation of a bivariate normal distribution. It holds that

$$\zeta_q = e^{\gamma_q + \frac{1}{2}\delta_q^2} - 1 \quad (q = 1, 2).$$

Let ϕ denote the payoff function of the option. Then we have the initial condition

$$v(s_1, s_2, 0) = \phi(s_1, s_2) \tag{2.2}$$

for $s_1 > 0$, $s_2 > 0$.

With the above notations and initial condition, the value of an American-style option satisfies the following PIDCP,

$$\begin{cases} \dfrac{\partial v}{\partial t}(s_1, s_2, t) \geqslant \mathcal{D}v(s_1, s_2, t) + \mathcal{J}v(s_1, s_2, t), \\[2ex] v(s_1, s_2, t) \geqslant \phi(s_1, s_2), \\[2ex] (v(s_1, s_2, t) - \phi(s_1, s_2))\left(\dfrac{\partial v}{\partial t}(s_1, s_2, t) - \mathcal{D}v(s_1, s_2, t) - \mathcal{J}v(s_1, s_2, t)\right) = 0 \end{cases} \tag{2.3}$$

whenever $s_1 > 0$, $s_2 > 0$, $0 < t \leqslant T$. Boundary conditions are given by imposing (2.3) for $s_1 = 0$ and $s_2 = 0$, respectively.

3 Spatial discretization

In this section we describe the spatial discretization of the PIDCP (2.3). As the first step, the unbounded spatial domain is truncated to $[0, S_{\max}] \times [0, S_{\max}]$ with sufficiently large value $S_{\max}$. On the two far sides $s_1 = S_{\max}$ and $s_2 = S_{\max}$ linear boundary conditions are taken, which are common in finance,

$$\frac{\partial^2 v}{\partial s_1^2} = 0 \ \ (\text{if } s_1 = S_{\max}) \quad \text{and} \quad \frac{\partial^2 v}{\partial s_2^2} = 0 \ \ (\text{if } s_2 = S_{\max}). \tag{3.1}$$

Let $K > 0$ denote the strike price of the option. As in [20], a smooth nonuniform Cartesian spatial grid is constructed such that K falls midway two successive grid points in each direction. This turns out to be beneficial for accuracy. Let parameter $d > 0$ and fix a subinterval $[S_{\text{left}}, S_{\text{right}}]$ of $[0, S_{\max}]$ such that $S_{\text{left}} + S_{\text{right}} = 2K < S_{\max}$. Define

$$\xi_{\min} = \sinh^{-1}\left(\frac{-S_{\text{left}}}{d}\right),$$

$$\xi_{\text{int}} = \frac{S_{\text{right}} - S_{\text{left}}}{d},$$

$$\xi_{\max} = \xi_{\text{int}} + \sinh^{-1}\left(\frac{S_{\max} - S_{\text{right}}}{d}\right).$$

Let $q \in \{1, 2\}$. For any given integer $\nu = \nu_q \geqslant 1$, let

$$\Delta\xi = \Delta\xi_q = \frac{\xi_{\text{int}} - 2\xi_{\min}}{\nu},$$

let $m = m_q > \nu$ be the smallest integer such that $m\Delta\xi \geqslant \xi_{\max} - \xi_{\min}$, reset $\xi_{\max}$ to $\xi_{\min} + m\Delta\xi$ and define equidistant points

$$\xi_j = \xi_{q,j} = \xi_{\min} + j\Delta\xi \quad (j = 0, 1, \ldots, m).$$

It is easily verified that $\xi_{\text{int}}/2$ forms the middle of $[\xi_0, \xi_\nu]$. Hence, it lies exactly halfway two successive ξ-grid points whenever ν is odd. The grid in the q-th spatial direction

$$0 = s_{q,0} < s_{q,1} < \ldots < s_{q,m}$$

is then defined through the transformation

$$s_{q,j} = \psi(\xi_j) \quad (j = 0, 1, \ldots, m),$$

where

$$\psi(\xi) = \begin{cases} S_{\text{left}} + d\sinh(\xi) & \text{for } \xi_{\min} \leqslant \xi \leqslant 0, \\ S_{\text{left}} + d\xi & \text{for } 0 < \xi \leqslant \xi_{\text{int}}, \\ S_{\text{right}} + d\sinh(\xi - \xi_{\text{int}}) & \text{for } \xi_{\text{int}} < \xi \leqslant \xi_{\max}. \end{cases}$$

This grid is uniform inside $[S_{\text{left}}, S_{\text{right}}]$ and nonuniform outside, where the mesh width inside equals $d\Delta\xi$ and is always smaller than the mesh widths outside. By construction, $K = \psi(\xi_{\text{int}}/2)$ falls halfway two successive grid points in each spatial direction whenever ν is odd. In this paper we heuristically select $d = K/3$ and $S_{\text{left}} = 0.8K$, $S_{\text{right}} = 1.2K$. Further, $S_{\max}$ is reset to $\psi(\xi_{\max})$, which is slightly larger than the original value.

The discretization of the convection-diffusion-reaction part $\mathcal{D}v$ in (2.3) is performed using finite differences. Let $u : [0, S_{\max}] \to \mathbb{R}$ be any given smooth function, let $0 = s_0 < s_1 < \ldots < s_m = S_{\max}$

be any given smooth nonuniform unidirectional grid and define mesh widths $h_j = s_j - s_{j-1}$ $(1 \leqslant j \leqslant m)$. We consider the following second-order central finite difference formulas for convection and diffusion:

$$u'(s_j) \approx \alpha_{j,-1} u(s_{j-1}) + \alpha_{j,0} u(s_j) + \alpha_{j,1} u(s_{j+1}),$$

$$u''(s_j) \approx \beta_{j,-1} u(s_{j-1}) + \beta_{j,0} u(s_j) + \beta_{j,1} u(s_{j+1}),$$

with

$$\alpha_{j,-1} = \frac{-h_{j+1}}{h_j(h_j + h_{j+1})} \qquad \alpha_{j,0} = \frac{h_{j+1} - h_j}{h_j h_{j+1}} \qquad \alpha_{j,1} = \frac{h_j}{h_{j+1}(h_j + h_{j+1})},$$

$$\beta_{j,-1} = \frac{2}{h_j(h_j + h_{j+1})} \qquad \beta_{j,0} = \frac{-2}{h_j h_{j+1}} \qquad \beta_{j,1} = \frac{2}{h_{j+1}(h_j + h_{j+1})}$$

for $1 \leqslant j \leqslant m - 1$. In view of the degeneracy of $\mathcal{D}v$ at the zero boundaries, no discretization is required if $j = 0$. If $j = m$, then we use the first-order backward finite difference formula for the first derivative and, by the linear boundary condition (3.1), the second derivative vanishes. Finally, the mixed derivative term in $\mathcal{D}v$ is handled by applying the finite difference formulas for the first derivative consecutively in the two spatial directions.

Let the vector $V(t) = (V_{0,0}(t), V_{1,0}(t), \ldots, V_{m_1-1,m_2}(t), V_{m_1,m_2}(t))^\top$ where entry $V_{i,j}(t)$ denotes the semidiscrete approximation to $v(s_{1,i}, s_{2,j}, t)$ for $0 \leqslant i \leqslant m_1$, $0 \leqslant j \leqslant m_2$. The semidiscrete version of the convection-diffusion-reaction part $\mathcal{D}v$ can then be written as

$$A^{(D)}V(t)$$

with matrix

$$A^{(D)} = A^{(M)} + A_1 + A_2,$$

where the matrix $A^{(M)}$ corresponds to the mixed derivative term and the matrix A_q corresponds to all derivative terms in the q-th spatial direction $(q = 1, 2)$. Further, the reaction term has been distributed equally across A_1 and A_2.

For the discretization of the integral part $\mathcal{J}v$ we consider a transformation to the log-price variable $x_q = \ln(s_q)$ $(q = 1, 2)$. This yields a two-dimensional cross-correlation,

$$\left(\overline{\mathcal{J}}\,\overline{v}\right)(x_1, x_2, t) := \lambda \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \overline{v}(z_1, z_2, t)\overline{f}(z_1 - x_1, z_2 - x_2)\mathrm{d}z_1\mathrm{d}z_2 = (\mathcal{J}v)(e^{x_1}, e^{x_2}, t). \qquad (3.2)$$

Here $\overline{v}(z_1, z_2, t) = v(e^{z_1}, e^{z_2}, t)$ and $\overline{f}(\eta_1, \eta_2) = f(e^{\eta_1}, e^{\eta_2})e^{\eta_1}e^{\eta_2}$ is the probability density function of a bivariate normal distribution. Let M_q be a given power of 2 such that $\Delta x_q = \ln(S_{\max})/M_q$ is smaller than the smallest mesh width in the nonuniform $\ln(s_q)$-grid $(q = 1, 2)$. Then the double integral (3.2) is approximated on the uniform Cartesian grid $(x_{1,k}, x_{2,l}) = (k\Delta x_1, l\Delta x_2)$ for $k = -M_1 + 1, \ldots, M_1$, $l = -M_2 + 1, \ldots, M_2$ by

$$\overline{J}_{k,l}(t) = \lambda \sum_{j=-M_2+1}^{M_2} \sum_{i=-M_1+1}^{M_1} \overline{V}_{i,j}(t)\overline{f}_{i-k,j-l}\Delta x_1 \Delta x_2.$$

Here $\overline{V}_{i,j}(t) \approx \overline{v}(x_{1,i}, x_{2,j}, t)$ and $\overline{f}_{i-k,j-l} = \overline{f}(x_{1,i} - x_{1,k}, x_{2,j} - x_{2,l})$. Define the vectors

$$\overline{J}(t) = (\overline{J}_{-M_1+1,-M_2+1}(t), \overline{J}_{-M_1+2,-M_2+1}(t), \ldots, \overline{J}_{M_1-1,M_2}(t), \overline{J}_{M_1,M_2}(t))^\top,$$
$$\overline{V}(t) = (\overline{V}_{-M_1+1,-M_2+1}(t), \overline{V}_{-M_1+2,-M_2+1}(t), \ldots, \overline{V}_{M_1-1,M_2}(t), \overline{V}_{M_1,M_2}(t))^\top,$$

then
$$\overline{J}(t) = \overline{A}^{(J)}\,\overline{V}(t),$$

where $\overline{A}^{(J)}$ is a given asymmetric two-level block-Toeplitz matrix of size $(4M_1M_2) \times (4M_1M_2)$.

Matrix-vector products with asymmetric multilevel block-Toeplitz matrices can be computed efficiently using the FFT algorithm introduced by Barrowes, Teixeira & Kong [1]. This algorithm employs two FFTs and one inverse FFT, all of them one-dimensional, reducing the computational cost of the matrix-vector product to $\mathcal{O}(M_1M_2\log(M_1M_2))$. The algorithm from [1] essentially embeds the block-Toeplitz matrix in a circulant matrix before applying the FFT, thus avoiding any wrap-around effect, which would have occurred when applying the FFT directly to the block-Toeplitz matrix.

Notice that the price grid on which the PIDCP (2.3) is discretized is nonuniform and in general does not define a uniform log-price grid. We therefore bilinearly interpolate the option value approximations between the s- and x-grids immediately before and after application of the FFT algorithm of [1], leading to the discretized version of $\mathcal{J}v$:

$$A^{(J)}V(t)$$

with given, fixed matrix $A^{(J)}$. More details are given in [3].

With the above discretizations of $\mathcal{D}v$ and $\mathcal{J}v$, the following semidiscrete PIDCP is obtained:

$$V'(t) \geqslant AV(t), \quad V(t) \geqslant V^0, \quad (V(t) - V^0)^\top (V'(t) - AV(t)) = 0 \tag{3.3}$$

for $0 < t \leqslant T$. Here inequalities for vectors are to be understood componentwise. The matrix

$$A = A^{(D)} + A^{(J)} = A^{(M)} + A_1 + A_2 + A^{(J)}$$

and the vector V^0 is directly given by the payoff function ϕ.

It can be readily seen that the maximum norm of the matrix $A^{(J)}$ is bounded by a moderate constant, independent of the spatial grid, whenever the jump intensity λ is moderate. Hence, the dense matrix $A^{(J)}$ forms a nonstiff part of the semidiscrete system, whereas the sparse matrix $A^{(D)}$ constitutes a stiff part.

4 Temporal discretization

In the temporal discretization of the semidiscrete PIDCP (3.3) we employ two main techniques that have been proposed in the computational finance literature: the Ikonen–Toivanen (IT) splitting technique and the penalty method. The IT splitting technique was first considered for American option valuation in [14, 15, 30] and the penalty method in [9, 32].

4.1 IT splitting

IT splitting is combined with a temporal discretization scheme for semidiscrete PIDEs, so as to yield a temporal discretization of semidiscrete PIDCPs. In [3] a variety of contemporary operator splitting schemes has been studied for the temporal discretization of the semidiscrete two-dimensional Merton PIDE for European option values. Here, the integral part, which corresponds to the dense matrix $A^{(J)}$, is always conveniently treated in an explicit fashion.

In the following we define six temporal discretization methods for the semidiscrete two-dimensional Merton PIDCP (3.3) by combining six operator splitting schemes from [3] with the IT splitting technique. The pertinent operator splitting schemes are of the implicit-explicit (IMEX) and the alternating direction implicit (ADI) kind. The combination of IT splitting with IMEX schemes was introduced in [23] and next investigated in e.g. [18, 26, 28] for one- and two-dimensional PIDCPs. The combination of this technique with ADI schemes was introduced in [11] and subsequently studied in [10, 20] for two-dimensional PDCPs (without integral term).

A novel feature of our present application of the IT splitting technique is that it is used in an iterative manner in each time step. This turns out to yield a significant enhancement, as will be shown in Section 5 below. We denote the number of iterations by $\kappa \geqslant 1$ and the iterated version of the IT splitting technique by $\mathrm{IT}(\kappa)$. For $\kappa = 1$, this reduces to the original IT splitting.

Let integer $N \geqslant 1$ be given, step size $\Delta t = T/N$ and $\widehat{V}^0 = V^0$. Each of the following methods defines an approximation $\widehat{V}^n$ to $V(t^n)$ at the temporal grid point $t^n = n\Delta t$ successively for $n = 1, 2, \ldots, N$.

1. CNFI-IT(κ) method:

   This method can be viewed as a combination of the Crank–Nicolson scheme with fixed-point iteration (CNFI) for PIDEs, proposed by Tavella & Randall [29] and d'Halluin, Forsyth & Vetzal [8], with the $\mathrm{IT}(\kappa)$ splitting technique:

   $$\begin{cases} \left(I - \tfrac{1}{2}\Delta t A^{(D)}\right) Z_k = \left(I + \tfrac{1}{2}\Delta t A^{(D)}\right) \widehat{V}^{n-1} + \tfrac{1}{2}\Delta t A^{(J)}(\widehat{Z}_{k-1} + \widehat{V}^{n-1}) + \Delta t \lambda_{k-1}, \\ \widehat{Z}_k = \max\left\{ Z_k - \Delta t \lambda_{k-1}, V^0 \right\}, \quad \lambda_k = \max\left\{ 0, \lambda_{k-1} + (V^0 - Z_k)/\Delta t \right\}, \\ \text{for } k = 1, 2, \ldots, \kappa \text{ and } \widehat{V}^n = \widehat{Z}_\kappa, \quad \widehat{\lambda}^n = \lambda_\kappa. \end{cases} \quad (4.1)$$

   Here, and in the subsequent methods, $\widehat{Z}_0 = \widehat{V}^{n-1}$ and $\lambda_0 = \widehat{\lambda}^{n-1}$ with $\widehat{\lambda}^0$ taken to be the zero vector. Clearly, in each time step, κ matrix-vector products with the matrix $A^{(J)}$ occur.

2. IETR-IT(κ) method:

   The IETR scheme has been considered in [16] for PIDEs. It treats the convection-diffusion-reaction part using the implicit trapezoidal rule (Crank–Nicolson) and the integral part using the explicit trapezoidal rule. Combining this scheme with the $\mathrm{IT}(\kappa)$ splitting technique yields:

   $$\begin{cases} Y_0 = \widehat{V}^{n-1} + \Delta t\,(A^{(D)} + A^{(J)})\widehat{V}^{n-1} + \Delta t \lambda_{k-1}, \\ \bar{Y}_0 = Y_0 + \tfrac{1}{2}\Delta t A^{(J)}(Y_0 - \widehat{V}^{n-1}), \\ Y_1 = \bar{Y}_0 + \tfrac{1}{2}\Delta t A^{(D)}(Y_1 - \widehat{V}^{n-1}), \\ Z_k = Y_1, \\ \lambda_k = \max\left\{ 0, \lambda_{k-1} + (V^0 - Z_k)/\Delta t \right\}, \\ \text{for } k = 1, 2, \ldots, \kappa \text{ and} \\ \widehat{V}^n = \max\left\{ Z_\kappa - \Delta t \lambda_{\kappa-1}, V^0 \right\}, \\ \widehat{\lambda}^n = \lambda_\kappa. \end{cases} \quad (4.2)$$

   Noticing that $A^{(J)}\widehat{V}^{n-1}$ can be computed upfront, the IETR-IT(κ) method requires $\kappa + 1$ matrix-vector products with $A^{(J)}$ per time step.

3. CNAB-IT(κ) method:

   The CNAB scheme was proposed for PIDEs by Salmi & Toivanen [27]. It again treats the

7

convection-diffusion-reaction part by the Crank–Nicolson scheme, but the integral part is now handled in a two-step Adams–Bashforth manner. In Salmi, Toivanen & von Sydow [28] the CNAB scheme has been combined with IT splitting and applied to two-dimensional PIDCPs. The extension by IT(κ) splitting yields:

$$
\begin{cases}
\left(I - \tfrac{1}{2}\Delta t A^{(D)}\right) Z_k = \left(I + \tfrac{1}{2}\Delta t A^{(D)}\right) \widehat{V}^{n-1} + \tfrac{1}{2}\Delta t A^{(J)}(3\widehat{V}^{n-1} - \widehat{V}^{n-2}) + \Delta t \lambda_{k-1}, \\[2mm]
\lambda_k = \max\left\{0, \lambda_{k-1} + (V^0 - Z_k)/\Delta t\right\}, \\[2mm]
\text{for } k = 1, 2, \ldots, \kappa \text{ and} \\[2mm]
\widehat{V}^n = \max\left\{Z_\kappa - \Delta t \lambda_{\kappa-1}, V^0\right\}, \\
\widehat{\lambda}^n = \lambda_\kappa.
\end{cases}
\tag{4.3}
$$

Clearly, each time step of this method requires just one matrix-vector product with $A^{(J)}$.

The temporal discretization schemes that underlie the three methods (4.1), (4.2), (4.3) are all of the IMEX kind, making use of the operator splitting $A = A^{(D)} + A^{(J)}$. The next three methods are based upon ADI schemes, which employ the additional splitting $A^{(D)} = A^{(M)} + A_1 + A_2$.

4. MCS-IT(κ) method:

The modified Craig–Sneyd (MCS) scheme was introduced for PDEs containing mixed derivative terms by in 't Hout & Welfert [21]. Its direct adaptation to PIDEs has been investigated by in 't Hout & Toivanen [19]. Complementary to this, the MCS scheme has been combined with IT splitting by Haentjens et al. [10, 11], defining the MCS-IT method for PDCPs (without integral part). The following MCS-IT(κ) method generalizes all of these:

$$
\begin{cases}
Y_0 = \widehat{V}^{n-1} + \Delta t \left(A^{(D)} + A^{(J)}\right)\widehat{V}^{n-1} + \Delta t \lambda_{k-1}, \\
Y_j = Y_{j-1} + \theta \Delta t A_j (Y_j - \widehat{V}^{n-1}) \quad (j = 1, 2), \\
\bar{Y}_0 = Y_0 + \theta \Delta t \left(A^{(M)} + A^{(J)}\right)(Y_2 - \widehat{V}^{n-1}), \\
\widetilde{Y}_0 = \bar{Y}_0 + (\tfrac{1}{2} - \theta)\Delta t \left(A^{(D)} + A^{(J)}\right)(Y_2 - \widehat{V}^{n-1}), \\
\widetilde{Y}_j = \widetilde{Y}_{j-1} + \theta \Delta t A_j (\widetilde{Y}_j - \widehat{V}^{n-1}) \quad (j = 1, 2), \\[2mm]
Z_k = \widetilde{Y}_2, \\
\lambda_k = \max\left\{0, \lambda_{k-1} + (V^0 - Z_k)/\Delta t\right\}, \\[2mm]
\text{for } k = 1, 2, \ldots, \kappa \text{ and} \\[2mm]
\widehat{V}^n = \max\left\{Z_\kappa - \Delta t \lambda_{\kappa-1}, V^0\right\}, \\
\widehat{\lambda}^n = \lambda_\kappa.
\end{cases}
\tag{4.4}
$$

We make the standard choice $\theta = \tfrac{1}{3}$, which is prompted by stability and accuracy results in the literature for two-dimensional problems (see e.g. [17, 19, 21, 22]). Since $A^{(J)}\widehat{V}^{n-1}$ can be computed in advance and the explicit stages $\bar{Y}_0$, $\widetilde{Y}_0$ can be merged, the integral part is evaluated $\kappa + 1$ times per time step.

It is easily verified that in (4.4) the integral part and mixed derivative term are both handled by the explicit trapezoidal rule. The implicit stages Y_j, $\widetilde{Y}_j$ (for $j = 1, 2$) are often called stabilizing corrections and are unidirectional. The pertinent linear systems for these stages are tridiagonal and can therefore be solved very efficiently by using an a priori LU factorization.

5. MCS2-IT(κ) method:

An alternative adaptation of the MCS scheme to PIDEs has been proposed in [19] where the integral part is treated in a two-step Adams–Bashforth manner. Combining this adaptation with IT(κ) splitting leads to:

$$
\begin{cases}
X_0 = \widehat{V}^{n-1} + \Delta t A^{(D)} \widehat{V}^{n-1} + \Delta t \lambda_{k-1}, \\
Y_0 = X_0 + \frac{1}{2}\Delta t A^{(J)}(3\widehat{V}^{n-1} - \widehat{V}^{n-2}), \\
Y_j = Y_{j-1} + \theta \Delta t A_j(Y_j - \widehat{V}^{n-1}) \quad (j = 1, 2), \\
\bar{Y}_0 = Y_0 + \theta \Delta t A^{(M)}(Y_2 - \widehat{V}^{n-1}), \\
\widetilde{Y}_0 = \bar{Y}_0 + (\frac{1}{2} - \theta)\Delta t A^{(D)}(Y_2 - \widehat{V}^{n-1}), \\
\widetilde{Y}_j = \widetilde{Y}_{j-1} + \theta \Delta t A_j(\widetilde{Y}_j - \widehat{V}^{n-1}) \quad (j = 1, 2), \\
Z_k = \widetilde{Y}_2, \\
\lambda_k = \max\left\{0, \lambda_{k-1} + (V^0 - Z_k)/\Delta t\right\}, \\
\text{for } k = 1, 2, \ldots, \kappa \text{ and} \\
\widehat{V}^n = \max\left\{Z_\kappa - \Delta t \lambda_{\kappa-1}, V^0\right\}, \\
\widehat{\lambda}^n = \lambda_\kappa.
\end{cases}
\tag{4.5}
$$

We choose again $\theta = \frac{1}{3}$. It is clear that the MCS2-IT(κ) method requires only one evaluation of the integral part per time step.

6. SC2A-IT(κ) method:

The stabilizing correction two-step Adams-type (SC2A) scheme is a prominent member of the class of stabilizing correction multistep methods that has been investigated by Hundsdorfer & in 't Hout [13] for the numerical solution of PDEs. Its direct adaptation to PIDEs has been studied by Boen & in 't Hout [3] and treats the integral part and mixed derivative term jointly in a two-step Adams–Bashforth fashion. The combination with IT(κ) splitting yields:

$$
\begin{cases}
X_0 = \widehat{V}^{n-1} + \Delta t \left(A_1 + A_2\right) \sum_{i=1}^{2} b_i^0 \widehat{V}^{n-i} + \Delta t \lambda_{k-1}, \\
Y_0 = X_0 + \Delta t \left(A^{(M)} + A^{(J)}\right) \sum_{i=1}^{2} b_i^1 \widehat{V}^{n-i}, \\
Y_j = Y_{j-1} + \theta \Delta t A_j(Y_j - \widehat{V}^{n-1}) \quad (j = 1, 2), \\
Z_k = Y_2, \\
\lambda_k = \max\left\{0, \lambda_{k-1} + (V^0 - Z_k)/\Delta t\right\}, \\
\text{for } k = 1, 2, \ldots, \kappa \text{ and} \\
\widehat{V}^n = \max\left\{Z_\kappa - \Delta t \lambda_{\kappa-1}, V^0\right\}, \\
\widehat{\lambda}^n = \lambda_\kappa.
\end{cases}
\tag{4.6}
$$

Here the coefficients are $(b_1^1, b_2^1) = \left(\frac{3}{2}, -\frac{1}{2}\right)$ and $(b_1^0, b_2^0) = \left(\frac{3}{2} - \theta, -\frac{1}{2} + \theta\right)$ and, following [13], we select $\theta = \frac{3}{4}$. The SC2A-IT(κ) method also requires just one evaluation of the integral part per time step.

For each of the methods (4.1)–(4.6), the underlying IMEX or ADI scheme has order of consistency equal to two for fixed, nonstiff systems of ordinary differential equations (ODEs), provided $\kappa \geqslant 2$ for method (4.1).

In view of the nonsmoothness of the initial (payoff) function ϕ, the first two time steps of each of the six methods above are replaced by four damping steps with step size $\Delta t/2$ using the backward Euler scheme with fixed-point iteration and IT(κ) splitting, that is, the BEFI-IT(κ) method. With the full step size Δt, this method reads

$$
\begin{cases}
\left(I - \Delta t A^{(D)}\right) Z_k = \widehat{V}^{n-1} + \Delta t A^{(J)} \widehat{Z}_{k-1} + \Delta t \lambda_{k-1}, \\[2mm]
\widehat{Z}_k = \max\left\{ Z_k - \Delta t \lambda_{k-1}, V^0 \right\}, \quad \lambda_k = \max\left\{ 0, \lambda_{k-1} + (V^0 - Z_k)/\Delta t \right\}, \\[2mm]
\text{for } k = 1, 2, \ldots, \kappa \text{ and } \widehat{V}^n = \widehat{Z}_\kappa, \quad \widehat{\lambda}^n = \lambda_\kappa.
\end{cases}
$$

At present the convergence theory for time stepping methods based on IT splitting is still under development. In [10] a useful relevant result was proved for the BE-IT method applied to PDCPs. This result generalizes straightforwardly to the BEFI-IT(1) method applied to PIDCPs,

$$
\begin{cases}
(I - \Delta t A^{(D)})Z = (I + \Delta t A^{(J)})\widehat{V}^{n-1} + \Delta t \widehat{\lambda}^{n-1}, \\[2mm]
\widehat{V}^n = \max\{Z - \Delta t \widehat{\lambda}^{n-1}, V^0\}, \quad \widehat{\lambda}^n = \max\{0, \widehat{\lambda}^{n-1} + (V^0 - Z)/\Delta t\}.
\end{cases}
\tag{4.7}
$$

The above can be regarded as the IMEX Euler-IT method. The corresponding method for PIDCPs without IT splitting is the IMEX Euler method, which can be written as

$$
\begin{cases}
(I - \Delta t A^{(D)})V^n = (I + \Delta t A^{(J)})V^{n-1} + \Delta t \lambda^n, \\[2mm]
V^n \geqslant V^0, \quad \lambda^n \geqslant 0, \quad (V^n - V^0)^\top \lambda^n = 0.
\end{cases}
\tag{4.8}
$$

Let $p = (m_1 + 1)(m_2 + 1)$. For any given diagonal matrix $\mathcal{D} \in \mathbb{R}^{p \times p}$ with positive diagonal entries, define the scaled inner product by

$$
\langle U, V \rangle_{\mathcal{D}} = V^\top \mathcal{D}\, U \quad \text{whenever } U, V \in \mathbb{R}^p
$$

and let $\|\cdot\|_{\mathcal{D}}$ denote both the induced vector and matrix norms. The following theorem forms a direct generalization of [10, Thm. 3.1] and is stated without proof.

Theorem 4.1. *Consider the processes* (4.7) *and* (4.8). *Assume there exists a positive diagonal matrix $\mathcal{D}$ such that*
$$
\mathcal{D}A^{(D)} + (A^{(D)})^\top \mathcal{D} \text{ is negative semidefinite.}
$$
Assume there are real constants μ, ν independent of the spatial and temporal grids such that
$$
\|A^{(J)}\|_{\mathcal{D}} \leqslant \mu
$$
and
$$
\|\lambda^1\|_{\mathcal{D}} + \sum_{n=2}^{N} \|\lambda^n - \lambda^{n-1}\|_{\mathcal{D}} \leqslant \nu.
$$
Then
$$
\max_{1 \leqslant n \leqslant N} \|V^n - \widehat{V}^n\|_{\mathcal{D}} \leqslant \nu e^{\mu T} \Delta t
$$
whenever $\Delta t = T/N$, integer $N \geqslant 1$.

Theorem 4.1 yields the useful result that the sequence $\{\widehat{V}^n\}$ generated by (4.7) is $\mathcal{O}(\Delta t)$ close to the sequence $\{V^n\}$ defined by the basic method (4.8). For further results and a discussion of the assumptions in this theorem, we refer to [2, 10].

10

4.2 Penalty method

We next consider the penalty method in combination with temporal discretization schemes for PIDEs, defining temporal discretization methods for PIDCPs.

7. **CNFI-P method:**

   This method forms a combination of the CNFI scheme with the penalty method and has been introduced by d'Halluin, Forsyth & Labahn [7]:

$$
\begin{cases}
(I - \frac{1}{2}\Delta t^n A^{(D)} + P_{k-1})Z_k = (I + \frac{1}{2}\Delta t^n A^{(D)})\widehat{V}^{n-1} + \frac{1}{2}\Delta t^n A^{(J)}(Z_{k-1} + \widehat{V}^{n-1}) + P_{k-1}V^0 \\
\text{for } k = 1, 2, \ldots, \kappa \text{ and } \widehat{V}^n = Z_\kappa.
\end{cases}
\tag{4.9}
$$

Here $Z_0 = \widehat{V}^{n-1}$ is the starting value for the penalty iteration and P_{k-1} denotes the diagonal matrix with l-th diagonal entry

$$
(P_{k-1})_{l,l} = \begin{cases} Large & \text{if } (Z_{k-1})_l < (V^0)_l, \\ 0 & \text{otherwise.} \end{cases}
$$

The common convergence criterion is

$$
\max_l \frac{|(Z_\kappa)_l - (Z_{\kappa-1})_l|}{\max\{1, |(Z_\kappa)_l|\}} < tol.
\tag{4.10}
$$

We choose as natural values $tol = 10^{-7}$ and $Large = 10^7$.

There are κ matrix-vector products with $A^{(J)}$ in each time step of the CNFI-P method. Notice that κ depends on the time step number in view of the dynamic convergence criterion (4.10).

The CNFI-P method is applied with suitable nonuniform temporal grid points, which has been shown in e.g. [7, 9] to improve the temporal convergence behaviour. Following [15, 25] we take

$$
t^n = \left(\frac{n}{N}\right)^2 T \quad (n = 0, 1, 2, \ldots, N)
\tag{4.11}
$$

and set $\Delta t^n = t^n - t^{n-1}$.

8. **MCS-P method:**

   Heidarpour-Dehkordi & Christara [12] combined the MCS scheme with the penalty method, defining the MCS-P method for PDCPs. We consider here its direct adaptation to PIDCPs:

$$
\begin{cases}
Y_0 = \widehat{V}^{n-1} + \Delta t\,(A^{(D)} + A^{(J)})\widehat{V}^{n-1}, \\
Y_j = Y_{j-1} + \theta\Delta t A_j(Y_j - \widehat{V}^{n-1}) \quad (j = 1, 2), \\
\bar{Y}_0 = Y_0 + \theta\Delta t\,(A^{(M)} + A^{(J)})(Y_2 - \widehat{V}^{n-1}), \\
\widetilde{Y}_0 = \bar{Y}_0 + (\frac{1}{2} - \theta)\Delta t\,(A^{(D)} + A^{(J)})(Y_2 - \widehat{V}^{n-1}), \\
\widetilde{Y}_1 = \widetilde{Y}_0 + \theta\Delta t A_1(\widetilde{Y}_1 - \widehat{V}^{n-1}), \\[4pt]
\text{and next} \\[4pt]
(I - \theta\Delta t A_2 + P_{k-1})Z_k = \widetilde{Y}_1 - \theta\Delta t A_2 \widehat{V}^{n-1} + P_{k-1}V^0 \\[4pt]
\text{for } k = 1, 2, \ldots, \kappa \text{ and } \widehat{V}^n = Z_\kappa.
\end{cases}
\tag{4.12}
$$

11

We select again $\theta = \frac{1}{3}$. Clearly, the penalty iteration is introduced in the last implicit stage of the MCS scheme. The same starting value Z_0, penalty matrix P_{k-1}, penalty factor *Large*, convergence criterion (4.10) and tolerance *tol* are employed as for the CNFI-P method. Since the penalty iteration does not involve the matrix $A^{(J)}$, the number of matrix-vector products with this matrix per time step of (4.12) is equal to two.

The first two time steps of the CNFI-P and MCS-P methods are replaced by four damping steps with half the original step size(s) applying the BEFI scheme combined with the penalty method, that is, the BEFI-P method.

5   Numerical results

In this section, ample numerical experiments are performed to gain insight into the convergence behaviour of the eight operator splitting methods formulated in Section 4 in the numerical solution of the semidiscrete two-dimensional Merton PIDCP (3.3). To this purpose, the *temporal discretization error* is considered at $t = T$ on a region of interest (ROI) in the spatial domain,

$$\widehat{E}^{ROI}(m, N) = \max \left\{ |\widehat{V}_{i,j}^{N'} - V_{i,j}(T)| : \Delta t = T/N' \text{ and } s_L < s_{1,i}, s_{2,j} < s_U \right\}. \qquad (5.1)$$

Here s_L and s_U denote given, fixed numbers with $0 < s_L < s_U$ that define the ROI. We take $m_1 = m_2 = m$ for the spatial discretization. The number of time steps N' is chosen in function of N and the splitting method so that the total number of matrix-vector products with the matrix $A^{(J)}$ over the whole time interval $[0, T]$ is (essentially) the same for all methods. This is done to have a fair comparison between the eight methods, as determining these matrix-vector products forms the dominating computational part of each time step. Accordingly, we take $N' = \lceil 2N/\kappa \rceil$ for method (4.1), $N' = \lceil 2N/(\kappa+1) \rceil$ for methods (4.2) and (4.4), $N' = 2N$ for (4.3), (4.5) and (4.6) and $N' = N$ for (4.12). For the CNFI-P method (4.9) the number of matrix-vector products with $A^{(J)}$ per time step is not known a priori, since a dynamic convergence criterion (4.10) is employed. However, numerical experiments reveal that convergence of the penalty iteration is attained after on average four iterations. In view of this, we consider method (4.9) with $N' = \lceil N/2 \rceil$. A reference solution $V(T)$ to the semidiscrete PIDCP (3.3) has been computed by applying the CNFI-P method[1] and $N' = 10N$ time steps.

We deal with two types of options: an American put-on-the-min option and an American put-on-the-average option. Their payoff functions, with given strike K, are

$$\phi_{\text{put-on-min}}(s_1, s_2) = \max(0, K - \min(s_1, s_2))$$

and

$$\phi_{\text{put-on-average}}(s_1, s_2) = \max \left(0, K - \frac{s_1 + s_2}{2} \right).$$

Three financial parameter sets are considered, which are specified in Table 1. They are identical to those chosen in Boen & in 't Hout [3] for the case of European options. The first parameter set is given in Clift & Forsyth [6]. Here λT is small, indicating a low expected number of jumps. The second set has the same diffusion parameters as in Zvan, Forsyth & Vetzal [33] and jump parameters are taken where λT is about the same size as for the first set. The third set was introduced in [3] and here λT is quite large. Note further that for all three sets the correlation coefficients ρ and $\widehat{\rho}$ are nonzero.

---

[1]Using e.g. the CNAB-IT(2) method leads to visually identical Figures 1–5.

	σ_1	σ_2	ρ	λ	γ_1	γ_2	$\widehat{\rho}$	δ_1	δ_2	r	K	T
Set 1	0.12	0.15	0.30	0.60	-0.10	0.10	-0.20	0.17	0.13	0.05	100	1
Set 2	0.30	0.30	0.50	2	-0.50	0.30	-0.60	0.40	0.10	0.05	40	0.5
Set 3	0.20	0.30	0.70	8	-0.05	-0.20	0.50	0.45	0.06	0.05	40	1

Table 1: Parameter sets for the two-asset Merton jump-diffusion model and American option.

Figure 1 displays in grey the approximated early exercise regions (EERs) for the American put-on-the-min and put-on-the-average options under the three parameter sets from Table 1. In the following we study, for all eight methods formulated in Section 4, the temporal discretization error (5.1) on both a large and a small ROI. These regions are indicated in Figure 1 with blue and red, respectively. The large ROI is given by $[(1/2)K, (3/2)K]$ and has a nonempty intersection with the EERs for parameter Sets 1 and 2. The small ROI is given by $[(7/8)K, (9/8)K]$ in the case of the put-on-the-min option and $[(9/10)K, (11/10)K]$ in the case of the put-on-the-average option. From Figure 1 we observe that the small ROI does not intersect the EERs, that is, it lies fully within the continuation regions.

We start by considering the methods (4.2)-(4.6) with IT(κ) splitting and $\kappa = 1$. Only the CNFI-IT(κ) method (4.1) is used with $\kappa = 2$, so that the order of consistency of the underlying IMEX scheme is also equal to two, compare Subsection 4.1. As discussed above, for a fair comparison, the methods (4.1), (4.2), (4.4) are applied in this case with $N' = N$ time steps and the methods (4.3), (4.5), (4.6) with $N' = 2N$ time steps. We take $N = m$ and consider a range of values m with $10 \leqslant m \leqslant 200$. The temporal errors of all eight methods are shown (versus $1/m$) in Figure 2 for the put-on-the-min option and in Figure 3 for the put-on-the-average option, under the three parameter sets given in Table 1. Here the left column corresponds to the large ROI and the right column the small ROI. As a positive observation, all eight temporal discretization methods show a regular, monotonic convergence behaviour for parameter Set 3. However, with the favourable exception of the CNFI-IT(2) method, each of the other methods can yield a less regular convergence behaviour under Sets 1 and 2. In particular, the temporal errors can level off as m increases. This is found notably in the cases where the ROI overlaps with, or lies close to, the EER.

Since the CNFI-IT(2) method appears to be more robust, we next consider the methods (4.2)-(4.6) with IT(κ) splitting for $\kappa = 2$ as well. For a fair comparison, (4.2) and (4.4) are now applied with $N' = \lceil 2N/3 \rceil$ time steps (and the others as above). In Figures 4 and 5 the temporal errors are displayed for, respectively, the put-on-the-min option and the put-on-the-average option, under the three parameter sets from Table 1 and for both the large and small ROI. The positive result is clearly observed that if $\kappa = 2$, then each of the six methods (4.1)-(4.6) always shows a regular, monotonic convergence behaviour. Tables 2, 3, 4 display the numerical orders of convergence of all methods for Sets 1, 2, 3, respectively, on the small ROI. They have been computed based on the ten largest values of m under consideration. For the methods (4.1)-(4.6) with IT(2) splitting, an order of convergence is obtained which lies between 1.2 and 1.6 for Set 1, between 1.7 and 1.9 for Set 2, and is about equal to 1.9 for Set 3. We conjecture that the relatively lower orders of convergence for Set 1 are related to the small ROI lying close by the EER, with the premise that at the boundary of the EER the option value function is less smooth. The CNAB-IT(2) and MCS2-IT(2) methods, which both treat the integral part in a two-step Adams–Bashforth fashion, always gave rise to the smallest error constant in our experiments among all considered methods with IT(2) splitting.

We mention that additional experiments have been performed with an extra iteration in the methods (4.1)-(4.6), that is $\kappa = 3$, but this did not lead to a further significant improvement of the temporal convergence behaviour.

For the CNFI-P method (4.9), as already alluded to above, the convergence behaviour can be less

13

regular for Sets 1 and 2. For each given $N = m$, the temporal error of this method is generally found to lie in, or near to, the range of those obtained with all IT(2) splitting type methods together.

The MCS-P method (4.12) often yields relatively large temporal errors, notably for Set 1. We also observed this for the values of θ, *tol*, *Large* from [12], for the nonuniform temporal grid (4.11) and when applied to the two-dimensional Black–Scholes PDCP (without integral part).

For future reference, approximations to the values of the American put-on-the-min option, respectively put-on-the-average option, under the three parameter sets are summarized in Table 5, respectively Table 6. Here the number of spatial grid points has been chosen such that the smallest spatial mesh width is about equal to 0.40 and the MCS2-IT(2) method has been applied with time step size $\Delta t = 0.01$. We find a good agreement with the approximations for the American put-on-the-min values in the case of Set 1 obtained in [6, p.766]. The maximal absolute error in the approximations given by Tables 5 and 6 is estimated to be less than 0.01.

6   Conclusion

We have investigated a variety of temporal discretization methods in the numerical solution of the two-dimensional time-dependent partial integro-differential complementarity problem for the values of American-style options under the two-asset Merton jump-diffusion model. The methods under consideration are constructed as adaptations of modern operator splitting schemes of the IMEX and ADI kind for partial integro-differential equations, relevant to European-style two-asset options. Here the two-dimensional integral part is always conveniently treated in an explicit fashion. Their adaptation to American options is achieved by a combination with two popular techniques from the computational finance literature, the IT splitting technique and the penalty method. In the present paper we propose the novel approach where IT splitting is employed in an iterative manner in each time step. We refer to this as IT(κ) splitting, where $\kappa \geqslant 1$ denotes the number of iterations.

Six IMEX and ADI schemes that have recently been studied for European two-asset options in [3] are applied with IT(κ) splitting: CNFI, IETR, CNAB, MCS, MCS2 and SC2A. In addition, the CNFI and MCS schemes are combined with the penalty method, as introduced in [7] and [12], respectively.

Ample numerical experiments have been performed to investigate the convergence behaviour of the temporal discretization errors of the acquired eight methods. To render a fair comparison, the number of time steps for each method is selected such that over the whole time interval $[0, T]$ the methods employ the same number of evaluations of the, computationally dominant, integral part. As a main result we find that taking $\kappa = 2$ leads to a regular, monotonic convergence behaviour for the methods based on IT splitting, as opposed to the standard application of this technique where $\kappa = 1$. The obtained numerical orders of convergence with $\kappa = 2$ lie in between 1.2 and 1.9.

The CNFI-P method leads to temporal errors that are similar to those obtained with all IT(2) splitting methods together, but may show a less regular convergence behaviour. The MCS-P method often yields relatively large temporal errors.

In view of the above and taking into account the size of the error constants, the CNAB-IT(2) and MCS2-IT(2) methods merit our preference for the efficient and stable temporal discretization of the two-dimensional Merton PIDCP.
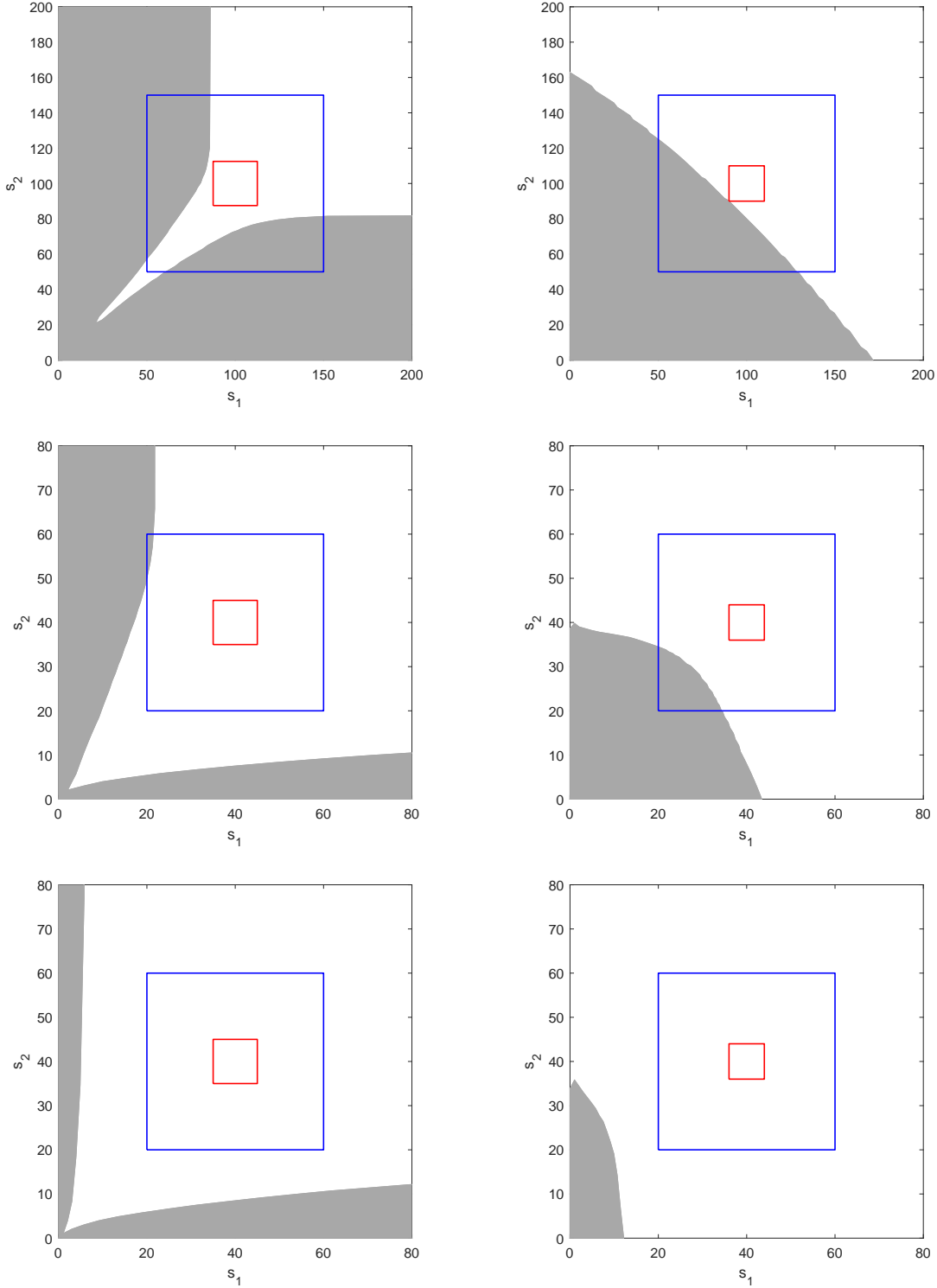
14

Acknowledgements

Figure 1: Early exercise regions for the American put-on-the-min (left) and put-on-the-average (right) options under the three parameter sets from Table 1 together with the regions of interest $[(1/2)K, (3/2)K]$ (blue) and $[(7/8)K, (9/8)K]$, respectively $[(9/10)K, (11/10)K]$ (red).

Figure 2: Temporal errors (5.1) of the eight operator splitting methods in the case of the American put-on-the-min option under the two-asset Merton jump-diffusion model with $\kappa = 2$ for method (4.1) and $\kappa = 1$ for methods (4.2)-(4.6). Displayed are the errors on the large ROI (left) and the small ROI (right) and for parameter Set 1 (top), Set 2 (mid) and Set 3 (bottom) from Table 1.

Figure 3: Temporal errors (5.1) of the eight operator splitting methods in the case of the American put-on-the-average option under the two-asset Merton jump-diffusion model with $\kappa = 2$ for method (4.1) and $\kappa = 1$ for methods (4.2)-(4.6). Displayed are the errors on the large ROI (left) and the small ROI (right) and for parameter Set 1 (top), Set 2 (mid) and Set 3 (bottom) from Table 1.
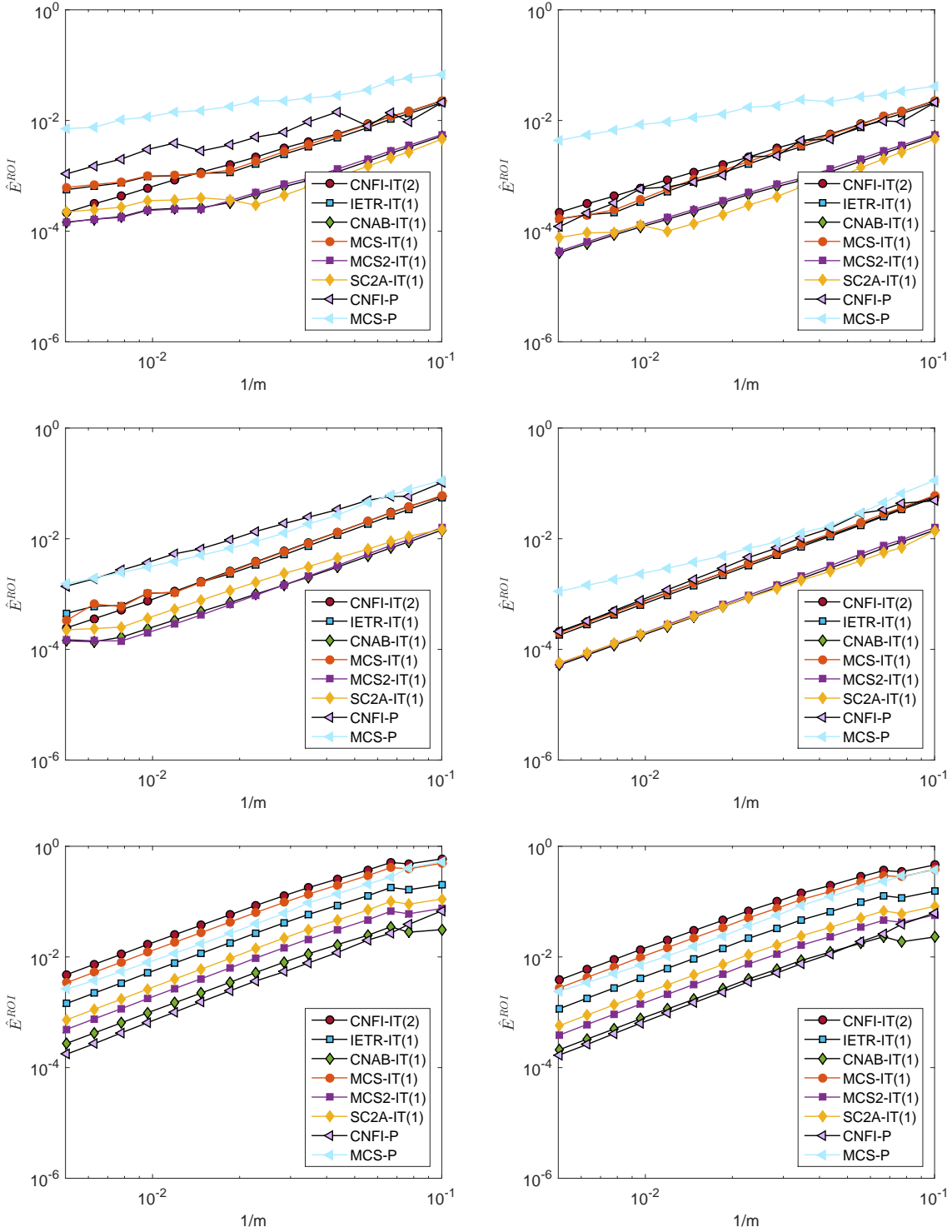
Figure 4: Temporal errors (5.1) of the eight operator splitting methods in the case of the American put-on-the-min option under the two-asset Merton jump-diffusion model with $\kappa = 2$ for all methods (4.1)-(4.6). Displayed are the errors on the large ROI (left) and the small ROI (right) and for parameter Set 1 (top), Set 2 (mid) and Set 3 (bottom) from Table 1.
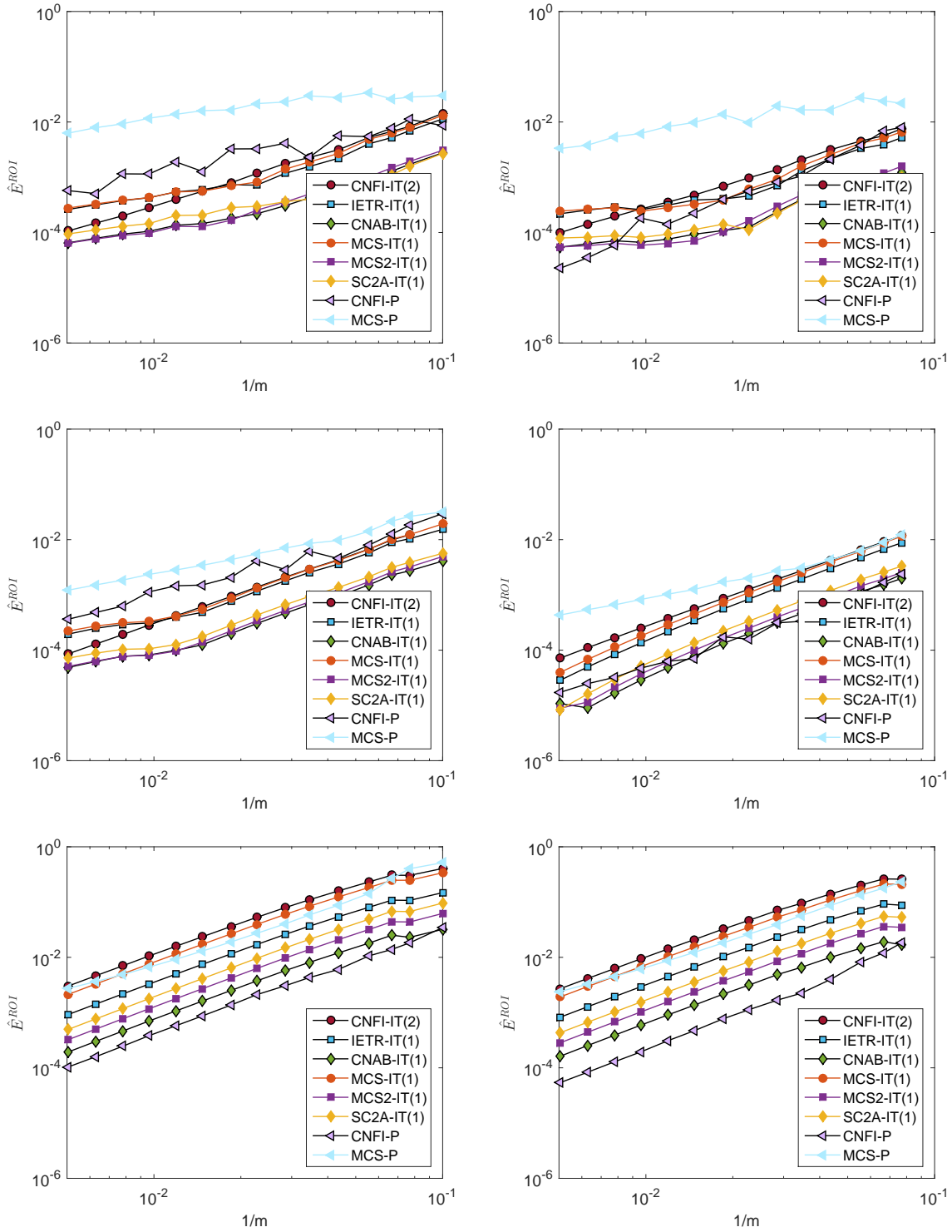
Figure 5: Temporal errors (5.1) of the eight operator splitting methods in the case of the American put-on-the-average option under the two-asset Merton jump-diffusion model with $\kappa = 2$ for all methods (4.1)-(4.6). Displayed are the errors on the large ROI (left) and the small ROI (right) and for parameter Set 1 (top), Set 2 (mid) and Set 3 (bottom) from Table 1.
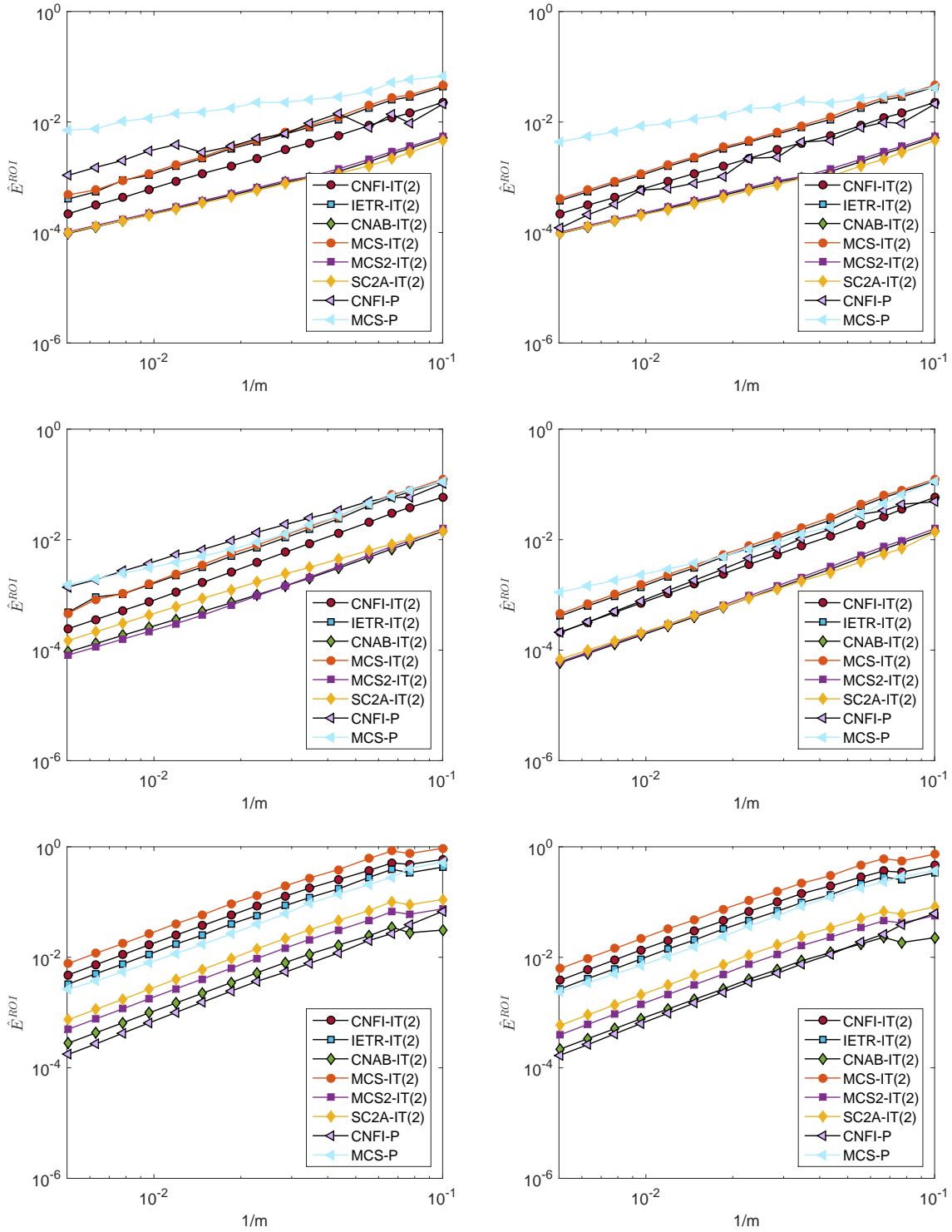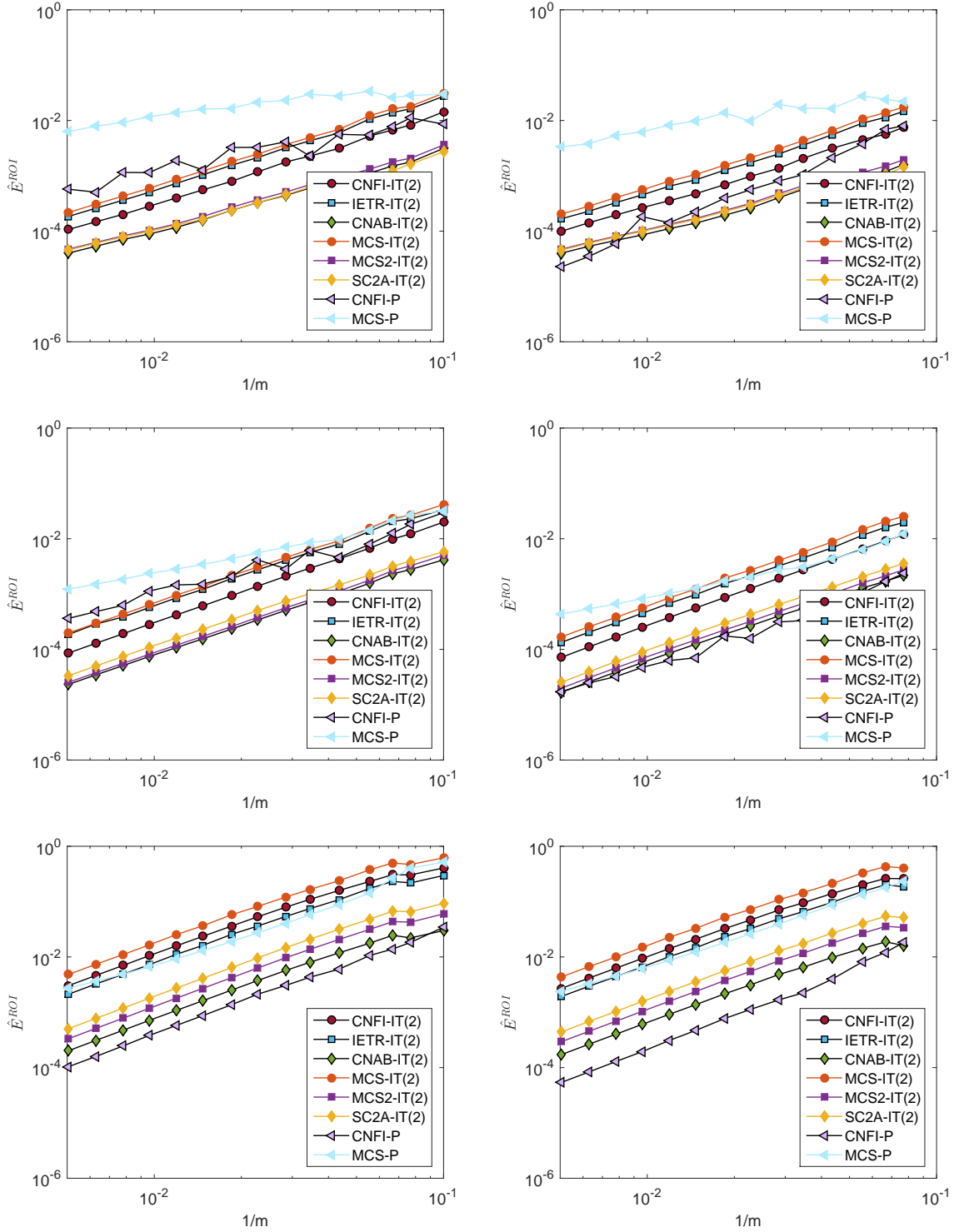
	put-on-the-min	put-on-the-average
CNFI-IT(2)	1.53	1.53
CNAB-IT(2)	1.22	1.34
IETR-IT(2)	1.60	1.58
MCS-IT(2)	1.59	1.58
MCS2-IT(2)	1.22	1.34
SC2A-IT(2)	1.17	1.32
CNFI-P	1.71	1.99
MCS-P	0.85	0.89

Table 2: Convergence orders of all methods, with $\kappa = 2$, on the small ROI under parameter Set 1.

	put-on-the-min	put-on-the-average
CNFI-IT(2)	1.87	1.89
CNAB-IT(2)	1.80	1.81
IETR-IT(2)	1.86	1.83
MCS-IT(2)	1.87	1.83
MCS2-IT(2)	1.83	1.82
SC2A-IT(2)	1.69	1.85
CNFI-P	2.04	1.59
MCS-P	1.23	1.04

Table 3: Convergence orders of all methods, with $\kappa = 2$, on the small ROI under parameter Set 2.

	put-on-the-min	put-on-the-average
CNFI-IT(2)	1.88	1.86
CNAB-IT(2)	1.91	1.91
IETR-IT(2)	1.87	1.85
MCS-IT(2)	1.85	1.84
MCS2-IT(2)	1.93	1.93
SC2A-IT(2)	1.93	1.93
CNFI-P	1.98	1.97
MCS-P	1.85	1.64

Table 4: Convergence orders of all methods, with $\kappa = 2$, on the small ROI under parameter Set 3.

	Set 1		
	$S_0^{(1)} = 90$	$S_0^{(1)} = 100$	$S_0^{(1)} = 110$
$S_0^{(2)} = 90$	16.391	13.999	12.758
$S_0^{(2)} = 100$	13.021	9.620	7.877
$S_0^{(2)} = 110$	11.443	7.227	5.132

	Set 2		
	$S_0^{(1)} = 36$	$S_0^{(1)} = 40$	$S_0^{(1)} = 44$
$S_0^{(2)} = 36$	15.467	14.564	13.794
$S_0^{(2)} = 40$	14.092	13.107	12.263
$S_0^{(2)} = 44$	12.921	11.877	10.982

	Set 3		
	$S_0^{(1)} = 36$	$S_0^{(1)} = 40$	$S_0^{(1)} = 44$
$S_0^{(2)} = 36$	21.742	20.908	20.167
$S_0^{(2)} = 40$	21.272	20.394	19.611
$S_0^{(2)} = 44$	20.892	19.983	19.166

Table 5: American put-on-the-min option value approximations under parameter Sets 1, 2, 3.

	Set 1		
	$S_0^{(1)} = 90$	$S_0^{(1)} = 100$	$S_0^{(1)} = 110$
$S_0^{(2)} = 90$	10.003	5.989	3.441
$S_0^{(2)} = 100$	6.030	3.442	1.887
$S_0^{(2)} = 110$	3.491	1.891	0.993

	Set 2		
	$S_0^{(1)} = 36$	$S_0^{(1)} = 40$	$S_0^{(1)} = 44$
$S_0^{(2)} = 36$	5.406	4.363	3.547
$S_0^{(2)} = 40$	4.214	3.339	2.669
$S_0^{(2)} = 44$	3.225	2.507	1.969

	Set 3		
	$S_0^{(1)} = 36$	$S_0^{(1)} = 40$	$S_0^{(1)} = 44$
$S_0^{(2)} = 36$	12.466	11.930	11.440
$S_0^{(2)} = 40$	11.434	10.943	10.495
$S_0^{(2)} = 44$	10.493	10.043	9.633

Table 6: American put-on-the-average option value approximations under parameter Sets 1, 2, 3.

References

[1] B. E. Barrowes, F. L. Teixeira, and J. A. Kong. Fast algorithm for matrix-vector multiply of asymmetric multilevel block-Toeplitz matrices in 3-D scattering. *Microw. Opt. Tech. Lett.*, 31:28–32, 2001.

[2] L. Boen. *Multivariate Financial Models and Partial Integro-Differential Equations for the Valuation of Financial Options*. University of Antwerp, 2019. PhD thesis.

[3] L. Boen and K. J. in 't Hout. Operator splitting schemes for the two-asset Merton jump-diffusion model. *J. Comp. Appl. Math.*, 2020. To appear. Available at arXiv: 1901.03839.

[4] M. J. Brennan and E. S. Schwartz. The valuation of American put options. *J. Finance*, 32:449–462, 1977.

[5] N. Clarke and K. Parrott. Multigrid for American option pricing with stochastic volatility. *Appl. Math. Fin.*, 6:177–195, 1999.

[6] S. S. Clift and P. A. Forsyth. Numerical solution of two asset jump diffusion models for option valuation. *Appl. Numer. Math.*, 58:743–782, 2008.

[7] Y. d'Halluin, P. A. Forsyth, and G. Labahn. A penalty method for American options with jump diffusion processes. *Numer. Math.*, 97:321–352, 2004.

[8] Y. d'Halluin, P. A. Forsyth, and K. R. Vetzal. Robust numerical methods for contingent claims under jump diffusion processes. *IMA J. Numer. Anal.*, 25:87–112, 2005.

[9] P. A. Forsyth and K. R. Vetzal. Quadratic convergence for valuing American options using a penalty method. *SIAM J. Sci. Comp.*, 23:2095–2122, 2002.

[10] T. Haentjens and K. J. in 't Hout. ADI schemes for pricing American options under the Heston model. *Appl. Math. Fin.*, 22:207–237, 2015.

[11] T. Haentjens, K. J. in 't Hout, and K. Volders. ADI schemes with Ikonen–Toivanen splitting for pricing American put options in the Heston model. In *AIP Conference Proceedings*, volume 1281, pages 231–234. 2010.

[12] V. Heidarpour-Dehkordi and C. C. Christara. Spread option pricing using ADI methods. *Int. J. Numer. Anal. Mod.*, 15:353–369, 2018.

[13] W. Hundsdorfer and K. J. in 't Hout. On multistep stabilizing correction splitting methods with applications to the Heston model. *SIAM J. Sci. Comp.*, 40:A1408–A1429, 2018.

[14] S. Ikonen and J. Toivanen. Operator splitting methods for American option pricing. *Appl. Math. Lett.*, 17:809–814, 2004.

[15] S. Ikonen and J. Toivanen. Operator splitting methods for pricing American options under stochastic volatility. *Numer. Math.*, 113:299–324, 2009.

[16] K. J. in 't Hout. *Numerical Partial Differential Equations in Finance Explained*. Palgrave Macmillan, 2017.

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

[18] K. J. in 't Hout and J. Toivanen. Application of operator splitting methods in finance. In *Splitting Methods in Communication, Imaging, Science, and Engineering*, pages 541–575. Springer, 2016.

[19] K. J. in 't Hout and J. Toivanen. ADI schemes for valuing European options under the Bates model. *Appl. Numer. Math.*, 130:143–156, 2018.

[20] K. J. in 't Hout and R. L. Valkov. Numerical study of splitting methods for American option valuation. In *Novel Methods in Computational Finance*, pages 373–398. Springer, 2017.

[21] 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.*, 59:677–692, 2009.

[22] K. J. in 't Hout and M. Wyns. Convergence of the Modified Craig–Sneyd scheme for two-dimensional convection-diffusion equations with mixed derivative term. *J. Comp. Appl. Math.*, 296:170–180, 2016.

[23] Y. Kwon and Y. Lee. A second-order tridiagonal method for American options under jump-diffusion models. *SIAM J. Sci. Comp.*, 33:1860–1872, 2011.

[24] C. W. Oosterlee. On multigrid for linear complementarity problems with application to American-style options. *Elec. Trans. Numer. Anal.*, 15:165–185, 2003.

[25] C. Reisinger and A. Whitley. The impact of a natural time change on the convergence of the Crank–Nicolson scheme. *IMA J. Numer. Anal.*, 34:1156–1192, 2014.

[26] S. Salmi and J. Toivanen. Comparison and survey of finite difference methods for pricing American options under finite activity jump-diffusion models. *Int. J. Comp. Math.*, 89:1112–1134, 2012.

[27] S. Salmi and J. Toivanen. IMEX schemes for pricing options under jump-diffusion models. *Appl. Numer. Math.*, 84:33–45, 2014.

[28] S. Salmi, J. Toivanen, and L. von Sydow. An IMEX-scheme for pricing options under stochastic volatility models with jumps. *SIAM J. Sci. Comp.*, 36:B817–B834, 2014.

[29] D. Tavella and C. Randall. *Pricing Financial Instruments*. Wiley, 2000.

[30] J. Toivanen. Numerical valuation of European and American options under Kou's jump-diffusion model. *SIAM J. Sci. Comp.*, 30:1949–1970, 2008.

[31] J. Toivanen and C. W. Oosterlee. A projected algebraic multigrid method for linear complementarity problems. *Numer. Math. Theor. Meth. Appl.*, 5:85–98, 2012.

[32] R. Zvan, P. A. Forsyth, and K. R. Vetzal. Penalty methods for American options with stochastic volatility. *J. Comp. Appl. Math.*, 91:199–218, 1998.

[33] R. Zvan, P. A. Forsyth, and K. R. Vetzal. A finite volume approach for contingent claims valuation. *IMA J. Numer. Anal.*, 21:703–731, 2001.