

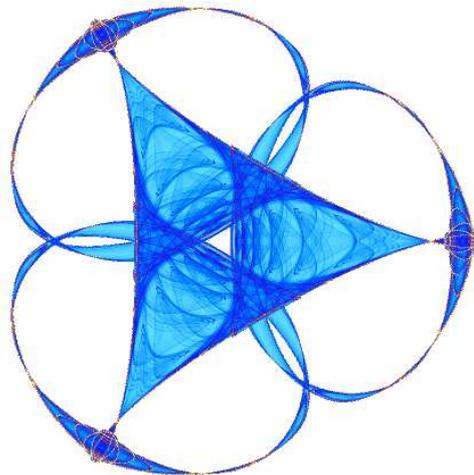
**MULTISCALE METHODS FOR THE VALUATION OF  
AMERICAN OPTIONS WITH STOCHASTIC VOLATILITY**

By

**Angela Kunoth**  
**Christian Schneider**  
and  
**Katharina Wiechers**

**IMA Preprint Series # 2382**

( September 2011 )



**INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS**

UNIVERSITY OF MINNESOTA  
400 Lind Hall  
207 Church Street S.E.  
Minneapolis, Minnesota 55455-0436

Phone: 612/624-6066 Fax: 612/626-7370

URL: <http://www.ima.umn.edu>

# MULTISCALE METHODS FOR THE VALUATION OF AMERICAN OPTIONS WITH STOCHASTIC VOLATILITY

ANGELA KUNOTH\*<sup>†</sup>, CHRISTIAN SCHNEIDER , AND KATHARINA WIECHERS

**Abstract.** This paper deals with the efficient valuation of American options. We adopt Heston’s approach for a model of stochastic volatility and derive a generalized Black Scholes equation. This leads to a parabolic boundary value problem with a free boundary, the optimal exercise price of the option. For its efficient numerical solution, we employ, among other multiscale methods, a monotone multigrid method based on linear finite elements in space and display corresponding numerical experiments.

**Key words.** American option pricing, stochastic volatility, parabolic boundary value problem, free boundary, monotone multigrid method, multigrid efficiency.

**AMS subject classifications.** 65M55, 35J86, 65N30, 65D07, 91B70.

**1. Introduction.** In Mathematical Finance, the fair pricing of an American option can in a standard model be formulated as a parabolic boundary value problem involving Black–Scholes’ equation [BS] with a free boundary. One seeks to compute the free boundary, the optimal exercise price of the option, together with the solution of the PDE, the value of the option. For American put options, one generally does not have closed form solutions so that one has to employ numerical schemes. In the simplest model, the volatility is assumed to be constant which, however, is usually not satisfied in practice. Particularly, an effect called volatility smile has been observed [Hu]. There are several approaches to estimate the volatility from observed stock data, see, e.g. [CRZ]. Here we adopt Heston’s approach [He] to model the volatility to satisfy a stochastic differential equation. Empirical evaluations show that Heston’s model for American options with stochastic volatility provides a substantial improvement for option pricing compared to the simple Black–Scholes’ model [AGG2]. Extensive information on models for stochastic volatility can be found in [FPS]. For Heston’s model and one asset, the standard Itô approach yields a variational inequality in two space variables with a nonsymmetric parabolic differential operator.

Numerical schemes for American option pricing are typically based on finite difference approaches, see, e.g., [CP, IT1, IT2, ItT, Oo, PS, RW]. In [IT1] an operator splitting method is used to separate the treatment of the early exercise constraint from the solution of the system of linear equations into separate fractional time steps. In [CP] and [Oo] multigrid methods are developed. In [RW] the focus is on high-dimensional problems which arise if one considers basket options. European basket options with stochastic volatility have been solved highly efficiently by sparse grid wavelet methods in [HKSU]. A comparison of the described methods can be found in [IT2]. In [ItT] a finite difference discretization in space and a Rannacher time-stepping scheme is proposed. The appearing discrete linear complementary problem is solved by a primal-dual active set approach.

A finite element approach for Heston’s model for stochastic volatility and Euro-

---

\*Supported in part by the Deutsche Forschungsgemeinschaft (SFB 611, Universität Bonn) and the Institute for Mathematics and its Applications (IMA) at the University of Minnesota with funds provided by the National Science Foundation (NSF).

<sup>†</sup>Corresponding Author. Institut für Mathematik, Universität Paderborn, Warburger Str. 100, 33098 Paderborn, Germany, [kunoth@math.uni-paderborn.de](mailto:kunoth@math.uni-paderborn.de), [www2.math.uni-paderborn.de/ags/kunoth/group/angelakunoth.html](http://www2.math.uni-paderborn.de/ags/kunoth/group/angelakunoth.html)

pean options was proposed in [AWW]. In [ZFV] finite elements for the second-order terms and finite volumes for the first order terms together with two penalty methods for linear complementary problems are considered to value American options with stochastic volatility. A rich source for numerical methods for option pricing problems is [AP1].

A somewhat different approach is the one proposed in [LS] which is based on simulation and the use of least-squares to estimate the expected value of the payoff to the option holder. This approach has been adopted and improved widely, see, e.g. [AGG1].

The purpose of this paper is twofold. First, we derive a variational inequality for the American option pricing problem which we will then discretize in terms of linear finite elements with respect to space. Second, the resulting linear inequality system is to be solved with optimal linear computational complexity. Based on a projective Gauss-Seidel scheme, we investigate different multiscale methods (nested iteration and a cascading type scheme) to accelerate the solution. Specific emphasis will be placed on an appropriate monotone multigrid method [Ko] for which appropriate prolongation and restriction operators need to be constructed. We conclude with corresponding numerical experiments.

## 2. Basics of Option Pricing.

**2.1. Definitions and the Basic Model.** We begin by introducing some basic notions and standard assumptions, see, e.g., [S]. An *option* is a financial contract which allows its holder to sell or buy an underlying asset at or before a specific time  $T \geq 0$  at a fixed *strike price*  $K \geq 0$ . Accordingly, one distinguishes between a *call option* (the right to buy) and a *put option* (the right to sell), and between an *European option* (exercise only at time  $T$ ) and an *American option* (exercise at any time  $t \in [0, T]$ ). The time-dependent stock underlying the option is denoted by  $S = S(t)$  for  $t \geq 0$  and is assumed to yield no dividends.

Let  $(\mathcal{A}, \Sigma, P)$  be a finite probability space. For a square integrable random number  $X$ , denote by  $E(X) := \sum_{z \in \mathcal{A}} X(z)P(z)$  its expectation value and by  $\text{Var}(X) = E((X - E(X))^2)$  its variance. The standard deviation (of the variance) or *volatility*  $\sigma$  is defined as  $\sigma := \sqrt{\text{Var}(X)}$ . The (real-valued and non-negative) variance underlying the stock is denoted by  $v = v(t) = v(X(t))$ . We wish to determine the *fair price* or *value* of the option  $V = V(t, S(t), v(t)) \in \mathbb{R}^+ (= [0, \infty))$  which is accepted by both the seller and buyer of the option.

The standard Black-Scholes model [BS] requires that the financial market is *frictionless*, i.e., there are no transaction costs or taxes; trade is possible at any time  $t \in [0, T]$  and that assets are available at any fraction. Moreover, there is a constant interest rate  $r \in \mathbb{R}^+$  and the market is *free of arbitrage*, e.g., immediate riskless gains are not possible.

For the convenience of the reader, we recall the modeling process first for the case of a *constant volatility*  $\sigma = \sqrt{v}$ , see, e.g., [Hu]. We assume that the random dynamic of a stock is described by a geometric *Brownian motion* or *Wiener process*  $W : \mathbb{R}^+ \times \mathcal{A} \rightarrow \mathbb{R}$ . Denote by  $\mu \in \mathbb{R}$  the *drift* (trend) of the stock. Then the process can be described by the stochastic differential equation (SDE)

$$dS(t) = \mu S(t)dt + \sigma S(t) dW(t). \quad (2.1)$$

For a Brownian motion  $W(t) = W(t, \omega)$ ,  $\omega \in \mathcal{A}$ ,  $t \geq 0$ , a stochastic process  $Y(t)$ ,  $t \geq 0$ , is called *Itô process* if there exist two functions  $a, b : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  with specific

properties, see, e.g., [Ok], such that

$$dY(t) = a(t, Y(t))dt + b(t, Y(t))dW(t) \quad (2.2)$$

is valid. Modifications of such processes can be described by the following (one-dimensional) *Itô formula*: If  $f \in \mathcal{C}^1(\mathbb{R}^+) \times \mathcal{C}^2(\mathbb{R})$  and  $Y(t)$  is an Itô process, then  $Z(t) := f(t, Y(t))$  is also an Itô-process, and one has

$$\begin{aligned} dZ(t) = & \left( \frac{\partial f(t, Y(t))}{\partial Y} a(t, Y(t)) + \frac{\partial f(t, Y(t))}{\partial t} + \frac{1}{2} \frac{\partial^2 f(t, Y(t))}{\partial Y^2} b^2(t, Y(t)) \right) dt \\ & + \frac{\partial f(t, Y(t))}{\partial Y} b(t, Y(t)) dW(t). \end{aligned} \quad (2.3)$$

Applying this formula to the option price function  $V(t, S(t))$  with  $S(t)$  according to (2.1) using  $a = \mu S$  and  $b = \sigma S$  leads to

$$dV = \left( \mu S \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \sigma S \frac{\partial V}{\partial S} dW(t). \quad (2.4)$$

Consider now a portfolio  $\Pi = \Pi(t) \in \mathbb{R}$ , consisting of an option with value  $V$  and a (possibly fractional) number  $\delta = \delta(t) \in \mathbb{R}$  of stocks with value  $S$ , composed as  $\Pi := V - \delta S$ . Using (2.1) and (2.4) yields

$$d\Pi = dV - \delta dS = \left( \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} - \delta \mu S \right) dt + \left( \frac{\partial V}{\partial S} - \delta \right) \sigma S dW. \quad (2.5)$$

In order to eliminate the stochastic influence in the last term, we choose for any  $t$  the term  $\delta = \frac{\partial V}{\partial S}$  to obtain

$$d\Pi = dV - \delta dS = \left( \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} \right) dt. \quad (2.6)$$

Since the change in the value of the portfolio must be equal to the change of a riskfree bond with value  $\Pi$  and interest rate  $r$ , due to the assumption that the market is free of arbitrage, it follows  $d\Pi = r\Pi dt$  and finally the *Black Scholes' equation*

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0. \quad (2.7)$$

**2.2. The Heston Model.** In financial markets, there is no reason to assume that the volatility is indeed constant. In fact, CIR processes are much more appropriate to describe the stochastic behaviour of volatility [CIR], see also [BR]. A stochastic process is called a *CIR (Cox-Ingersoll-Ross) process* if it satisfies an SDE

$$dY(t) = \kappa(\gamma - Y(t))dt + \xi \sqrt{Y(t)} dW(t), \quad (2.8)$$

where  $W$  is a Wiener process,  $\kappa$  denotes the mean reversion rate,  $\gamma$  the mean reversion level, and  $\xi \in \mathbb{R}^+$  is the volatility of the process. In addition to the volatility  $\sigma = \sqrt{v}$ , there appears here also the (constant) volatility  $\xi$  of the CIR process.

The Heston model developed in [He] can now be described as follows. Let the variance  $v$  be described by a CIR process (2.8). Let  $W_1, W_2$  be two Wiener processes

which are correlated by a parameter  $\rho \in [-1, 1]$ . Recalling the dynamics of the stock price (2.1), we obtain

$$dS = \mu S dt + \sqrt{v} S dW_1, \quad (2.9)$$

$$dv = \kappa(\gamma - v) dt + \xi \sqrt{v} dW_2. \quad (2.10)$$

For a multi-dimensional (geometric) Brownian motion  $\mathbf{W}(t) = (W_1(t), \dots, W_d(t))^T$ ,  $t > 0$ , where the components  $W_i, W_j$  are correlated with correlation coefficients  $\rho_{ij} \in [-1, 1]$ ,  $i, j = 1, \dots, d$ ,  $d \in \mathbb{N}$ , a stochastic process  $\mathbf{Y}(t)$ ,  $t > 0$ , with random variable  $\mathbf{Y}(t) \in \mathbb{R}^d$  is called *multi-dimensional Itô process* if there exist two functions  $\mathbf{a}, \mathbf{b} : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  with particular properties, see, e.g., [Ok, Wi], such that

$$dY_i(t) = a_i(t, \mathbf{Y}(t)) dt + b_i(t, \mathbf{Y}(t)) dW_i(t), \quad i = 1, \dots, d. \quad (2.11)$$

In this setting, modifications can be described by the *multi-dimensional Itô formula*: If  $f \in \mathcal{C}^1(\mathbb{R}) \times \mathcal{C}^2(\mathbb{R}^d)$  and  $\mathbf{Y}(t)$  is a multi-dimensional Itô process, then  $Z(t) := f(t, \mathbf{Y}(t))$  is also an Itô process, and one has

$$\begin{aligned} dZ(t) = & \left( \frac{\partial f(t, \mathbf{Y}(t))}{\partial t} + \sum_{i=1}^n \frac{\partial f(t, \mathbf{Y}(t))}{\partial Y_i} a_i(t, \mathbf{Y}(t)) \right. \\ & \left. + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f(t, \mathbf{Y}(t))}{\partial Y_i \partial Y_j} b_i(t, \mathbf{Y}(t)) b_j(t, \mathbf{Y}(t)) \rho_{ij} \right) dt \\ & + \sum_{i=1}^n \frac{\partial f(t, \mathbf{Y}(t))}{\partial Y_i} b_i(t, \mathbf{Y}(t)) dW_i(t). \end{aligned} \quad (2.12)$$

Applying this formula to  $V(t, S(t), v(t)) = V(t, \mathbf{Y}(t))$  with  $\mathbf{Y}(t) = (S(t), v(t))^T$  from (2.9) and (2.10) using  $a_1 = \mu S$ ,  $b_1 = \sqrt{v} S$ ,  $a_2 = \kappa(\gamma - v)$  and  $b_2 = \xi \sqrt{v}$  yields

$$\begin{aligned} dV = & \left( \frac{\partial V}{\partial t} + \frac{\partial V}{\partial S} \mu S + \frac{\partial V}{\partial v} \kappa(\gamma - v) + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} v S^2 + \frac{\partial^2 V}{\partial S \partial v} S v \xi \rho + \frac{1}{2} \frac{\partial^2 V}{\partial v^2} v \xi^2 \right) dt \\ & + \frac{\partial V}{\partial S} \sqrt{v} S dW_1 + \frac{\partial V}{\partial v} \sqrt{v} \xi dW_2. \end{aligned} \quad (2.13)$$

Consider now a portfolio  $\Pi = \Pi(t) \in \mathbb{R}$ , consisting of an option with value  $V$  and a (possibly fractional) number  $\delta = \delta(t) \in \mathbb{R}$  of stocks with value  $S$ , composed as  $\Pi := V - \delta S$ . Choosing  $\delta = \frac{\partial V}{\partial S}$  for every time  $t$  and using (2.9) and (2.13) leads to

$$\begin{aligned} d\Pi = & dV - \delta dS \quad (2.14) \\ = & \left( \frac{\partial V}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 V}{\partial S^2} + S v \xi \rho \frac{\partial^2 V}{\partial S \partial v} + \frac{1}{2} v \xi^2 \frac{\partial^2 V}{\partial v^2} + (\kappa(\gamma - v) - \xi \lambda \sqrt{v}) \frac{\partial V}{\partial v} \right) dt, \end{aligned} \quad (2.15)$$

where  $\lambda$  represents the market price of volatility risk, see, e.g. [CP]. Since the portfolio is now riskfree, the change in its value must be equal to the change of a bond with value  $\Pi$  and a riskfree interest rate  $r$ , due to the assumption that the market is free of arbitrage. Therefore, we have  $d\Pi = r\Pi dt$ . Finally, we arrive at the following parabolic partial differential equation with diffusion, convection and reaction terms, denoted as *Heston's equation*

$$\begin{aligned} \mathcal{L}V := & \frac{\partial V}{\partial t} + \frac{1}{2} \left( S^2 v \frac{\partial^2 V}{\partial S^2} + 2\rho \xi v S \frac{\partial^2 V}{\partial S \partial v} + \xi^2 v \frac{\partial^2 V}{\partial v^2} \right) + rS \frac{\partial V}{\partial S} + \kappa(\gamma - v) \frac{\partial V}{\partial v} - rV \\ = & 0. \end{aligned} \quad (2.16)$$

Here we have assumed like in [AWW, IT1] that the price of the volatility risk  $\lambda$  vanishes. For European options, a closed form solution can be found in [He]. Moreover, Merton's theorem, see, e.g., [Hu], says that the value of an American call option equals the value of a European call option.

Therefore, we will in the following concentrate on the most interesting case from a computational point of view: American put options satisfying (2.16) for which there is no closed form solution available. For this parabolic PDE to be solved backwards in time, we need to specify appropriate boundary and end conditions.

To this end, we denote by  $\mathcal{H}(S) := \max(K - S, 0)$  the so-called *payoff function* which is independent of  $v$ . If at final time  $T$  the value of the stock  $S$  is above the strike price  $K$ , one would not execute the option; the option is without value. On the other hand, if the value of  $S$  is below  $K$ , the value of the option is  $K - S$ . Therefore, at final time  $T$ , the value of the option is

$$V(T, S, v) = \mathcal{H}(S). \quad (2.17)$$

REMARK 2.1. *Note that, due to the same sign in front of the first order derivative in time and second order with respect to space, posing end conditions at  $T$  for (2.16) is a well-posed parabolic problem (in contrast to the backwards heat equation, see, e.g., [RR].)*

As for boundary conditions of  $V$  with respect to  $S$ , we argue as follows: If the option was exercised immediately, one would gain an amount of  $K - S$  which approaches  $K$  as  $S \rightarrow 0$ . The maximal achievable gain is  $K$ . Accordingly, it is best to exercise immediately to avoid discounting  $K$ . This yields

$$\lim_{S \rightarrow 0} V(t, S, v) = \lim_{S \rightarrow 0} \mathcal{H}(S) = K. \quad (2.18)$$

For  $S \rightarrow \infty$ , the option is worthless, so we have

$$\lim_{S \rightarrow \infty} V(t, S, v) = \lim_{S \rightarrow \infty} \mathcal{H}(S) = 0. \quad (2.19)$$

We want to mention that, in [CP], it was proposed to replace condition (2.19) by the Neumann boundary condition  $\lim_{S \rightarrow \infty} \frac{\partial V(t, S, v)}{\partial S} = 0$  'for practical reasons' not clear to us. The model treated in [Sch] has taken this modification into account. In the present context, we prefer to stay with (2.19)

Next we discuss boundary conditions for  $V$  with respect to  $v$ . Here one first has to take a look at the diffusion process  $v(t)$  and see if  $v = 0$  is an attainable boundary or not. The *Feller condition*, see e.g. [AP2], states that  $v = 0$  is attainable if  $\xi^2 > 2\kappa\gamma$  is fulfilled. If  $v = 0$  is an attainable boundary, it is possible to state a boundary condition here. If it is an unattainable boundary, one cannot state a boundary condition. So it is crucial to choose the parameters in (2.10) so that the Feller condition is fulfilled. Choosing the parameters according to [DPS], that is  $\xi = 0.61$ ,  $\kappa = 6.21$  and  $\gamma = 0.019$  for the Heston model is according to [DPS] the most reasonable. For this choice, the Feller condition is fulfilled and, thus, it is possible to state a boundary condition for  $v = 0$ . We choose the boundary condition

$$\lim_{v \rightarrow 0} V(t, S, v) = \mathcal{H}(S). \quad (2.20)$$

Finally, as in [CP], one expects that for  $v \rightarrow \infty$  the option price  $V$  would be insensitive to a change in the volatility, i.e.

$$\lim_{v \rightarrow \infty} \frac{\partial V(t, S, v)}{\partial v} = 0. \quad (2.21)$$

**2.3. American Put Options in the Heston Model.** For American options (and fixed variance), it is well-known that there exists for each time  $t \in [0, T]$  always a stock price  $S$  for which early exertion before final time  $T$  is advantageous, see, e.g., [S]. One can show that these values define a continuous curve  $S_f(t)$ . It is a priori unknown and therefore defines a *free boundary*. Accordingly, one should exercise the option in the areas where  $S \leq S_f$  and hold the option if  $S > S_f$ . We are therefore finally led to the following option pricing problem:

PROBLEM 2.2. *Solve for  $V = V(t, S, v)$  the system*

$$\begin{aligned} \mathcal{L}V &= 0 & \text{for } S > S_f, v \geq 0, t \in [0, T] \\ V(t, S, v) &= \mathcal{H}(S) & \text{for } S \leq S_f, v \geq 0, t \in [0, T]. \end{aligned} \quad (2.22)$$

with boundary conditions

$$\begin{aligned} \lim_{S \rightarrow 0} V(t, S, v) &= \lim_{S \rightarrow 0} \mathcal{H}(S) = K, & \lim_{S \rightarrow \infty} V(t, S, v) &= \lim_{S \rightarrow \infty} \mathcal{H}(S) = 0, \\ \lim_{v \rightarrow 0} V(t, S, v) &= \mathcal{H}(S), & \lim_{v \rightarrow \infty} \frac{\partial V(t, S, v)}{\partial v} &= 0. \end{aligned} \quad (2.23)$$

Note that in addition to solving a parabolic PDE with a linear but nonsymmetric operator  $\mathcal{L}$ , a particular difficulty arises from the unknown free boundary  $S_f$ .

**2.4. Free Boundary Value and Linear Complementary Problems.** It will be useful to consider instead of Problem 2.2 the following equivalent formulation as a *linear complementary problem*, see e.g., [S] for a proof. This formulation has the advantage that the free boundary  $S_f$  does not explicitly appear in the problem formulation which later enables numerical discretization. Of course, the free boundary still has to be computed. We denote the domain on which the option price  $V$  lives by  $\Omega_{\mathcal{L}}^{\infty} := (0, T) \times (0, \infty) \times (0, \infty) \subset \mathbb{R}^3$ . (The upper index indicates an unbounded set which will later be appropriately cut.)

PROBLEM 2.3. *Solve for  $V = V(t, S, v)$ ,  $(t, S, v) \in \Omega_{\mathcal{L}}^{\infty}$  the system*

$$\begin{aligned} \mathcal{L}V(V - \mathcal{H}) &= 0 \\ \mathcal{L}V &\leq 0 \\ V - \mathcal{H} &\geq 0 \end{aligned} \quad (2.24)$$

with boundary conditions (2.23) and end condition  $V(T, S, v) = \mathcal{H}(S)$  for all  $(v, S) \in \mathbb{R}^+ \times \mathbb{R}^+$ .

**2.5. Transformation and Localization.** In view of the fact that  $\mathcal{L}$  defined in (2.16) is indeed a degenerate parabolic differential operator, one usually introduces a new variable  $x$  replacing the stock price  $S$  in terms of the *transformed stock variable*  $x = \ln(\frac{S}{K})$ . This results in the disappearance of the variable coefficient  $S$  in (2.16). The transformed value of the option  $y = y(t, x, v)$  then satisfies the transformed Heston equation

$$\frac{\partial y}{\partial t} + \frac{1}{2}\xi^2 v \frac{\partial^2 y}{\partial v^2} + \rho\xi v \frac{\partial^2 y}{\partial v \partial x} + \frac{1}{2}v \frac{\partial^2 y}{\partial x^2} + \kappa(\gamma - v) \frac{\partial y}{\partial v} + \left(r - \frac{1}{2}v\right) \frac{\partial y}{\partial x} - ry = 0 \quad (2.25)$$

on part of the domain  $\Omega_{\mathcal{L}}^{\infty}$ . Recall that all parameters except  $x, v$  are constant real values. Abbreviating

$$A := \frac{1}{2}v \begin{pmatrix} \xi^2 & \rho\xi \\ \rho\xi & 1 \end{pmatrix}, \quad \mathbf{b} := \begin{pmatrix} -\kappa(\gamma - v) + \frac{1}{2}\xi^2 \\ -r + \frac{1}{2}v + \frac{1}{2}\xi\rho \end{pmatrix}, \quad (2.26)$$

the parabolic PDE (2.27) can be rewritten as

$$\mathcal{Y}y := \frac{\partial y}{\partial t} + \nabla \cdot A \nabla y - \mathbf{b} \cdot \nabla y - ry = 0 \quad (2.27)$$

where  $\nabla y := (\frac{\partial y}{\partial x}, \frac{\partial y}{\partial v})^T \in \mathbb{R}^2$  is the gradient of  $y$ , and for vectors  $\mathbf{c}, \mathbf{b} \in \mathbb{R}^2$ ,  $\mathbf{c} \cdot \mathbf{b} = \mathbf{c}^T \mathbf{b}$  is the Euclidean inner product. The payoff function  $\mathcal{H}(S) = \max(K - S, 0)$  is transformed into

$$g(x) := \mathcal{H}(Ke^x) = \max(K - Ke^x, 0). \quad (2.28)$$

Moreover, the free boundary  $S_f$  in (2.22) is transformed to a function  $x_f = x_f(t, v)$ . We denote the transformed domain by  $\Omega_{\mathcal{Y}}^\infty := (0, T) \times \mathbb{R} \times \mathbb{R}^+ \subset \mathbb{R}^3$ . The transformed linear complementary problem then reads as follows.

PROBLEM 2.4. *Solve for  $y = y(t, x, v)$ ,  $(t, x, v) \in \Omega_{\mathcal{Y}}^\infty$ , the system*

$$\begin{aligned} \mathcal{Y}y(y - g) &= 0 \\ \mathcal{Y}y &\leq 0 \\ y - g &\geq 0 \end{aligned} \quad (2.29)$$

with boundary conditions

$$\begin{aligned} \lim_{x \rightarrow \pm\infty} y(t, x, v) &= \lim_{x \rightarrow \pm\infty} g(x), \\ y(t, x, 0) &= g(x), \quad \lim_{v \rightarrow \infty} \frac{\partial y(t, x, v)}{\partial v} = 0. \end{aligned}$$

and end condition  $y(T, x, v) = g(x)$  for all  $(x, v) \in \mathbb{R} \times \mathbb{R}^+$ .

The question of existence and uniqueness of a solution to this problem will be addressed later in the variational setting.

In order to enable numerical computations, we localize the unbounded domain  $\Omega_{\mathcal{Y}}^\infty$  by defining  $\Omega_{\mathcal{Y}} := (0, T) \times \Omega \subset \mathbb{R}^3$  with the spatial domain  $\Omega := (x_{\min}, x_{\max}) \times (v_{\min}, v_{\max}) \subset \mathbb{R}^2$  such that  $0 < v_{\min} < v_{\max}$  and  $x_{\min} < 0 < x_{\max}$ . Since the variance  $v$  is always positive, choosing such a strictly positive  $v_{\min}$  is well justified; this also avoids a degenerating coefficient  $A$  in (2.27). Typically, one chooses  $x_{\min} = -x_{\max}$  for symmetry reasons. The boundary conditions in Problem 2.4 must be adapted accordingly. Thus, defining

$$\begin{aligned} \Gamma &:= \{(x, v) \in \partial\Omega : v = v_{\min} \text{ or } x = x_{\min} \text{ or } x = x_{\max}\} \\ \Gamma_v &:= \partial\Omega \setminus \Gamma = \{(x, v) \in \partial\Omega : v = v_{\max}\}, \end{aligned} \quad (2.30)$$

the boundary conditions (2.23) read

$$y = g \quad \text{on } \Gamma, \quad \frac{\partial y}{\partial v} = 0 \quad \text{on } \Gamma_v, \quad (2.31)$$

and the end conditions are  $y(T, x, v) = g(x)$  for all  $(x, v) \in \Omega$ .

REMARK 2.5. *Often one performs an additional transformation with respect to time in the form  $\tilde{t} := T - t$ . This results in an opposite sign in front of  $\nabla \cdot A \nabla y$  in (2.27) and in a parabolic PDE forward in time with initial conditions at  $t = 0$  instead of end conditions at  $t = T$ . Specifically, one employs this transformation in case of a constant volatility since then (2.27) reduces to the heat equation.*

**3. Variational Formulation and Discretization.** Next we will derive an appropriate variational formulation of Problem 2.4. Let  $L_2(\Omega)$  be the usual space of Lebesgue measurable and square integrable functions on  $\Omega$ , and denote by  $H^1(\Omega)$  the Sobolev space of first order weak derivatives. We define

$$\tilde{\mathcal{K}} := \{\varphi \in H^1(\Omega) : \varphi \geq g, \varphi = g \text{ on } \Gamma\} \quad (3.1)$$

where the inequality sign means to hold pointwise for all  $(x, v) \in \Omega$ . Let  $\varphi \in \tilde{\mathcal{K}}$  be any test function and  $y \in \tilde{\mathcal{K}}$  be a solution of Problem 2.4. We multiply the second equation in (2.29) by  $\varphi - g$  (which does not change the sign) and integrate over  $\Omega$  yielding  $\int_{\Omega} \mathcal{Y}y(\varphi - g)d\Omega \leq 0$ . Subtraction of the first equation in (2.29), integrated over  $\Omega$ , i.e.,  $\int_{\Omega} \mathcal{Y}y(y - g)d\Omega = 0$  yields

$$\int_{\Omega} \mathcal{Y}y(\varphi - y)d\Omega \leq 0, \quad (3.2)$$

thereby eliminating  $g$ . Moreover, we apply a further transform  $u := y - g$  in order to achieve zero boundary conditions and inequality. For this, we need to assume for the moment that  $g$  is sufficiently smooth. The more general case which applies to  $g = g(x)$  defined in (2.28) is addressed in Remark 3.3 below.

According to this transformation, we define a new constraint space as

$$\mathcal{K} := \{\varphi \in H^1(\Omega) : \varphi \geq 0, \varphi = 0 \text{ on } \Gamma\}. \quad (3.3)$$

Thus, instead of Problem 2.4, we consider

PROBLEM 3.1. *Solve for  $u = u(t, x, v) \in H^1(\Omega_{\mathcal{Y}})$  with*

$$u(t, \cdot, \cdot) \in \mathcal{K}, \quad \frac{\partial u}{\partial v} = 0 \text{ on } \Gamma_v, \quad u(T, x, v) = 0 \text{ on } \Omega, \quad (3.4)$$

*the inequality*

$$\int_{\Omega} \mathcal{Y}u(\varphi - u)d\Omega \leq - \int_{\Omega} \mathcal{Y}g(\varphi - u)d\Omega =: \ell(\varphi - u) \quad (3.5)$$

for all  $\varphi \in \mathcal{K}$  and  $t \in (0, T)$ .

Note that  $\frac{\partial y}{\partial v} = 0$  implies also  $\frac{\partial u}{\partial v} = 0$  since  $g$  does not depend on  $v$ .

For discretization, we first discretize the time derivative in (3.2) with a standard finite difference scheme. We decompose the time interval into equidistant points  $0 =: t^{(0)} < t^{(1)} < \dots < t^{(M)} := T$  with time step  $\tau := t^{(k)} - t^{(k-1)}$  and abbreviate the semi-discrete in time solution by  $u^{(k)} := u(t^{(k)}, x, v)$ . We approximate

$$\left. \frac{\partial u(t, x, v)}{\partial t} \right|_{t=t^{(k)}} \approx \frac{u^{(k+1)} - u^{(k)}}{\tau},$$

yielding the familiar  $\theta$ -scheme for  $\theta \in [0, 1]$

$$\begin{aligned} \ell(\varphi - u^{(k)}) &\geq \int_{\Omega} \frac{u^{(k+1)} - u^{(k)}}{\tau} (\varphi - u^{(k)}) d\Omega \\ &+ (1 - \theta) \int_{\Omega} \left( \nabla \cdot A \nabla u^{(k+1)} - \mathbf{b} \cdot \nabla u^{(k+1)} - r u^{(k+1)} \right) (\varphi - u^{(k)}) d\Omega \\ &+ \theta \int_{\Omega} \left( \nabla \cdot A \nabla u^{(k)} - \mathbf{b} \cdot \nabla u^{(k)} - r u^{(k)} \right) (\varphi - u^{(k)}) d\Omega. \end{aligned} \quad (3.6)$$

We recall that we wish to solve (2.27) *backward* in time. Thus, the terms indexed by  $k + 1$  are known while the terms with index  $k$  are to be determined. Choosing  $\theta = 1$  yields an explicit,  $\theta = 0$  an implicit Euler scheme, and  $\theta = \frac{1}{2}$  corresponds to the Crank-Nicolson scheme. Naturally, here we select the latter due to its consistency and convergence error rate  $\mathcal{O}(\tau^2)$ . Defining the operator

$$\mathcal{Y}^{(k)}u^{(k)} := -\theta\nabla \cdot A\nabla u^{(k)} + \theta\mathbf{b} \cdot \nabla u^{(k)} + (\theta r + \tau^{-1})u^{(k)} \quad (3.7)$$

and right hand side

$$f^{(k+1)} := (1-\theta)\nabla \cdot A\nabla u^{(k+1)} - (1-\theta)\mathbf{b} \cdot \nabla u^{(k+1)} - ((1-\theta)r - \tau^{-1})u^{(k+1)} + \mathcal{Y}g, \quad (3.8)$$

equation (3.6) is just

$$\int_{\Omega} \mathcal{Y}^{(k)}u^{(k)}(\varphi - u^{(k)})d\Omega \geq \int_{\Omega} f^{(k+1)}(\varphi - u^{(k)})d\Omega. \quad (3.9)$$

Applying Green's formula to the left hand side with respect to the variables  $x, v$  yields

$$\begin{aligned} \int_{\Omega} \mathcal{Y}^{(k)}u^{(k)}(\varphi - u^{(k)})d\Omega &= \theta \int_{\Omega} \left( A\nabla u^{(k)} \cdot \nabla(\varphi - u^{(k)}) + \mathbf{b} \cdot \nabla u^{(k)}(\varphi - u^{(k)}) \right) d\Omega \\ &\quad + \int_{\Omega} (\theta r + \tau^{-1})u^{(k)}(\varphi - u^{(k)})d\Omega \\ &\quad - \int_{\partial\Omega} A \frac{\partial u^{(k)}}{\partial \mathbf{n}}(\varphi - u^{(k)})ds \end{aligned} \quad (3.10)$$

where  $\mathbf{n}$  denotes the outward normal derivative at  $\partial\Omega$ . Due to the choice of  $\mathcal{K}$  from which  $\varphi, u^{(k)}$  are chosen, the boundary terms on  $\Gamma$  vanish. The remaining boundary term on  $\Gamma_v$  vanishes due to the Neumann condition (3.4) for  $u^{(k)}$ .

Analogously, we obtain for the right hand side

$$\begin{aligned} \int_{\Omega} f^{(k+1)}(\varphi - u^{(k)})d\Omega &= -(1-\theta) \int_{\Omega} A\nabla u^{(k+1)} \cdot \nabla(\varphi - u^{(k)})d\Omega \\ &\quad + \int_{\Omega} \left( -(1-\theta)\mathbf{b} \cdot \nabla u^{(k+1)} - ((1-\theta)r - \tau^{-1})u^{(k+1)} \right) (\varphi - u^{(k)})d\Omega \\ &\quad + \int_{\Omega} \frac{\partial g}{\partial t}(\varphi - u^{(k)})d\Omega - \int_{\Omega} A\nabla g \cdot \nabla(\varphi - u^{(k)})d\Omega \\ &\quad - \int_{\Omega} (\mathbf{b} \cdot \nabla g + rg)(\varphi - u^{(k)})d\Omega + \int_{\partial\Omega} A \frac{\partial g}{\partial \mathbf{n}}(\varphi - u^{(k)})ds. \end{aligned} \quad (3.11)$$

Defining the right hand side of (3.10) as the evaluation of a bilinear form  $a(\cdot, \cdot)$ , i.e.,

$$\begin{aligned} a(u^{(k)}, \varphi - u^{(k)}) &:= \theta \int_{\Omega} \left( A\nabla u^{(k)} \cdot \nabla(\varphi - u^{(k)}) + \mathbf{b} \cdot \nabla u^{(k)}(\varphi - u^{(k)}) \right) d\Omega \\ &\quad + \int_{\Omega} (\theta r + \tau^{-1})u^{(k)}(\varphi - u^{(k)})d\Omega \end{aligned} \quad (3.12)$$

we have derived the following *variational inequality*.

PROBLEM 3.2. *Solve for  $u^{(k)} \in \mathcal{K}$  and any  $\varphi \in \mathcal{K}$*

$$a(u^{(k)}, \varphi - u^{(k)}) \geq \langle f^{(k+1)}, \varphi - u^{(k)} \rangle \quad (3.13)$$

where  $\langle \cdot, \cdot \rangle$  is the dual form on  $(H^1(\Omega))' \times H^1(\Omega)$ .

REMARK 3.3. For the derivation of Problem 3.2, we have assumed that the lower bound  $g$  is sufficiently smooth. In fact, the formulation (3.13) reveals that  $g \in H^1(\Omega)$  is sufficient which holds for  $g$  defined in (2.28).

It is well-known from treatments on variational inequalities, see, e.g., [EO, GLT, KS] that Problem 3.2 has a unique solution (by a generalized theorem of Lax-Milgram) if the bilinear form  $a(\cdot, \cdot)$  is continuous and coercive on  $H^1(\Omega) \times H^1(\Omega)$  and the right hand side defines a linear continuous functional on  $H^1(\Omega)$ . This is clearly the case here since on  $\Omega$   $A$  is clearly positive definite and bounded. This question and detailed conditions on the relation of the different coefficients has been investigated in detail in [AWW] for the case of European options, leading to a variational equality, and introducing additional splits for the first order term for stability.

Next we discretize Problem 3.2 with respect to the space variables  $(x, v) \in \Omega$  in terms of linear *finite elements*, see, e.g., [Br]. We assume that  $\Omega$  is decomposed into uniform triangles of diameter proportional to a space discretization parameter  $h$  and that the decomposition is admissible. Let  $V_h \subset H^1(\Omega)$  with vanishing boundary values at  $\Gamma$  and  $\dim V_h =: N$  denote the finite-dimensional space spanned by the Courant finite elements (conforming  $P_1$  element) denoted by  $N_i$ . We always assume that the discretization is *conforming* which means that  $V_h$  is a subset of  $H^1(\Omega)$  and  $V_{2h} \subset V_h$ . The  $N_i$  may be viewed as (with respect to the twodimensional case) generalized hat functions.

On this discretization, we approximate the semidiscrete solution  $u^{(k)}$  by

$$u_h^{(k)} = u_h^{(k)}(x, v) := \sum_{i=1}^N u_i^{(k)} N_i(x, v) =: (\mathbf{u}^{(k)})^T \mathbf{N}. \quad (3.14)$$

Likewise, the test functions  $\varphi(x, v)$  are approximated by

$$\varphi_h = \varphi_h(x, v) := \sum_{i=1}^N \varphi_i N_i(x, v) =: \boldsymbol{\varphi}^T \mathbf{N}. \quad (3.15)$$

Inserting these representations into the left hand side in (3.13) yields

$$a(u^{(k)}, \varphi - u^{(k)}) = (\boldsymbol{\varphi} - \mathbf{u}^{(k)})^T B \mathbf{u}^{(k)} \quad (3.16)$$

where the *system matrix*  $B \in \mathbb{R}^{N \times N}$  is defined by  $B_{ji} = a(N_i, N_j)$ . Note that  $B$  is invertible but not symmetric. Accordingly, the right hand side in (3.13) is expressed in vector form as  $(\boldsymbol{\varphi} - \mathbf{u}^{(k)})^T \mathbf{f}$  where  $\mathbf{f}$  has entries defined accordingly in terms of  $N_i$ . This means that for every time step (starting with end time  $t = T$ ), we need to solve the following *discrete variational inequality*.

THEOREM 3.4. The finite element discretization of Problem 3.2 is: find  $\mathbf{u} \in \mathbb{R}^N$  with  $\mathbf{u} \geq \mathbf{0}$  such that

$$(\boldsymbol{\varphi} - \mathbf{u})^T (B\mathbf{u} - \mathbf{f}) \geq 0, \quad (3.17)$$

for all  $\boldsymbol{\varphi} \in \mathbb{R}^N$  with  $\boldsymbol{\varphi} \geq \mathbf{0}$ .

This problem is equivalent to the discrete linear complementary problem to find  $\mathbf{u} \in \mathbb{R}^N$  satisfying

$$\mathbf{u}^T (B\mathbf{u} - \mathbf{f}) = 0, \quad B\mathbf{u} - \mathbf{f} \geq 0, \quad \mathbf{u} \geq \mathbf{0}. \quad (3.18)$$

Here an inequality like  $\mathbf{u} \geq \mathbf{0}$  is to be understood component-wise. Again, existence and uniqueness of a solution of either one of these problems is guaranteed by a generalized Lax-Milgram theorem, applied to finite-dimensional spaces.

Recall from [BHR] for variational inequalities of the form (3.13) the error estimate

$$\|u^{(k)} - u_h^{(k)}\|_{H^1(\Omega)} = \mathcal{O}(h), \quad (3.19)$$

provided that  $f^{(k+1)} \in L_2(\Omega)$  which is the case here. Thus, with respect to the  $L_2(\Omega)$  norm, we expect for the Crank-Nicolson scheme and  $\tau = h$  an error

$$\|u - u_h\|_{L_2(\Omega_y)} = \mathcal{O}(h^2 + \tau^2). \quad (3.20)$$

**4. Iterative Solution of Inequality Systems.** According to Problem 3.2 and Theorem 3.4, we need to solve in each time step the inequality system (3.18), starting with  $t^{(M)} = T$ . A standard method which will serve as a basic iterative method in each time step is the following.

**4.1. The Projective Gauss-Seidel Method.** This scheme is an adaptation of the familiar Gauss-Seidel method for the iterative solution of linear systems  $G\mathbf{w} = \mathbf{g}$  where  $G \in \mathbb{R}^{N \times N}$  is symmetric positive definite,  $\mathbf{g} \in \mathbb{R}^N$  is given, to problems (3.18) for symmetric  $B$ , see, e.g., [Ha]. We will first briefly describe the scheme and then comment on the symmetry requirement of the matrix.

For the Gauss-Seidel scheme, one decomposes  $G$  additively into  $G = L + D + U$  where  $D$  is a diagonal,  $L$  is a strictly lower diagonal and  $U$  a strictly upper diagonal matrix. The iterative scheme is then: given an initial vector  $\mathbf{w}^0 \in \mathbb{R}^N$ , iterate for  $j = 1, 2, \dots$  the relation  $\mathbf{w}^j = \mathbf{w}^{j-1} + (U + D)^{-1}(\mathbf{g} - G\mathbf{w}^{j-1})$ . For the solution of a discrete complementary problem (3.18) (with  $B$  replaced by a symmetric positive definite  $G$  for the moment), one adds after the iteration a component-wise projection step into the convex set  $\mathbf{K} := \{\mathbf{w} \in \mathbb{R}^N : \mathbf{w} \geq \mathbf{0}\}$  which is the discretized version of the convex set  $\mathcal{K}$  defined in (3.3).

ALGORITHM 4.1 (Projective Gauss-Seidel Method). *Set the initial vector  $\mathbf{w}^0 \in \mathbb{R}^N$ .*

*Iterate for  $j = 1, 2, \dots$ :*

$$\begin{aligned} \tilde{\mathbf{w}}^j &= \mathbf{w}^{j-1} + (U + D)^{-1}(\mathbf{g} - G\mathbf{w}^{j-1}) \\ w_i^j &= \max(0, \tilde{w}_i^j), \quad i = 1, \dots, N. \end{aligned} \quad (4.1)$$

We will shortly write

$$\mathbf{w}^j = \text{pGS}(\mathbf{w}^{j-1}), \quad j = 1, 2, \dots \quad (4.2)$$

for Algorithm 4.1. Generalizations of this method are known as PSOR (projected successive overrelaxation) schemes for which (4.2) is a special case. Convergence of this method to the unique solution of (3.18) is known for any initial vector  $\mathbf{w}^0$  provided that  $G$  is symmetric positive definite and the convex set  $\mathbf{K}$  is factorial, e.g.,  $\mathbf{K} = \prod_{i=1}^N K_i$  [Cr]. While the latter is clearly satisfied, the matrix  $B$  defined in (3.16) is not symmetric. However, as long as  $B$  is not ‘too’ unsymmetric in a sense that can be made precise, and as long as  $\frac{1}{2}(B + B^T)$  is positive definite, the PSOR method still converges [Ni]. We have also observed this numerically for  $B$  defined in (3.16), see Section 5.

As for the convergence speed of the projective Gauss-Seidel scheme, this depends as in the unrestricted case on the spectral condition number of  $B$ . For the finite

element (or any other finite difference) discretization on a uniform grid of size  $h$  employed in Section 3, the spectral condition number is known to behave like  $\kappa_2(B) = \mathcal{O}(h^{-2})$  as  $h \rightarrow 0$ , see, e.g., [Ha]. Consequently, the speed of any iterative method deteriorates as the grid size becomes smaller. Remedies to overcome this problem are the subject of the next subsections.

**4.2. Multiscale Accelerations.** In order to accelerate a basic iterative scheme like the Gauss-Seidel method in the unrestricted case, one typically employs a *preconditioner*  $C$  for which  $\kappa_2(CB) \ll \kappa_2(B)$  and which is of cost  $\mathcal{O}(N)$  to set up and apply. Optimal preconditioners achieve  $\kappa_2(CB) \sim 1$  independent of  $h$ . All classes for which this is possible are based on multiscale approaches. This means that a hierarchy of grids and approximation spaces

$$V_0 \subset V_1 \subset \dots \subset V_\ell \subset \dots \subset V_L \subset H^1(\Omega) \quad (4.3)$$

comes into play, where the grid spacing  $h_\ell$  is proportional to  $2^{-\ell}$  and  $\dim V_\ell =: N_\ell = \mathcal{O}(2^{2\ell})$  since  $\Omega \subset \mathbb{R}^2$ . We denote by  $\ell$  the *level of multiresolution* or simply *level*. In the unrestricted case, multigrid and multiscale preconditioners like the BPX preconditioner and a wavelet preconditioner are among this class, see, e.g., [Br, D, Ha, Ku].

A multigrid method adapted to the restricted problem in (3.18) will be described in the next section.

**4.3. A Monotone Multigrid Method (MMG).** The method of choice for discretized variational inequalities is the monotone multigrid method (MMG) [Ko, Ma] which was set up for piecewise linear Ansatz functions. An extension to higher order B-Splines can be found in [HK1] with an application to the highly accurate computation of the Greeks for American options with constant volatility in [HK2].

The monotone multigrid method can be implemented as a variant of a standard multigrid scheme by adding a projection step and employing specific restriction operators. Consider the nested sequence of finite-dimensional spaces (4.3) with the highest level  $L$  denoting the discretization level on which (3.18) is to be solved,

$$\begin{aligned} B_L \mathbf{u}_L &\geq \mathbf{f}_L, \\ \mathbf{u}_L &\geq \mathbf{0}, \\ \mathbf{u}^T (B_L \mathbf{u}_L - \mathbf{f}_L) &= 0. \end{aligned} \quad (4.4)$$

Let  $\mathbf{u}_L^\nu \in V_L$  be the approximation in the  $\nu$ -th iteration of the MMG method. The basic multigrid idea is that the error  $\mathbf{v}_L := \mathbf{u}_L - \mathbf{u}_L^{\nu,1}$  between the smoothed iterate  $\mathbf{u}_L^{\nu,1} := \text{pGS}(\mathbf{u}_L^\nu)$  and the exact solution  $\mathbf{u}_L$  can be approximated without essential loss of information on a coarser grid  $\Delta_{L-1}$ . For the linear complementary problem (4.4), this is realized for two grids  $\Delta_L$  and  $\Delta_{L-1}$  as follows. Introducing the defect  $\mathbf{d}_L := \mathbf{f}_L - B_L \mathbf{u}_L^{\nu,1}$ , (4.4) can be written as

$$\begin{aligned} B_L \mathbf{v}_L &\geq \mathbf{d}_L, \\ \mathbf{v}_L &\geq -\mathbf{u}_L^{\nu,1}, \\ (\mathbf{v}_L + \mathbf{u}_L^{\nu,1})^T (B_L \mathbf{v}_L - \mathbf{d}_L) &= 0. \end{aligned} \quad (4.5)$$

On a coarser grid  $\Delta_{L-1}$  the defect problem can now be approximated by

$$\begin{aligned} B_{L-1} \mathbf{v}_{L-1} &\geq \mathbf{d}_{L-1}, \\ \mathbf{v}_{L-1} &\geq \mathbf{g}_{L-1}, \\ (\mathbf{v}_{L-1} - \mathbf{g}_{L-1})^T (B_{L-1} \mathbf{v}_{L-1} - \mathbf{d}_{L-1}) &= 0, \end{aligned}$$

where  $\mathbf{d}_{L-1} := r \mathbf{d}_L$  and  $\mathbf{g}_{L-1} := \tilde{r}(-\mathbf{u}_L^{\nu,1})$  with (different) restriction operators  $r, \tilde{r} : V_L \rightarrow V_{L-1}$ . The solution  $\mathbf{v}_{L-1}$  of the coarse grid problem is now used as an approximation to the error  $\mathbf{v}_L$ . It is first transported back to the fine grid by a prolongation operator  $p$ . Then it is added to the approximation  $\mathbf{u}_L^{\nu,1}$ . We want to point out that it is important that the restriction  $\tilde{r}$  is chosen such that the new iterate satisfies the constraint

$$\mathbf{u}_L^{\nu,2} := \mathbf{u}_L^{\nu,1} + p\mathbf{v}_{L-1} \geq \mathbf{g}_L := 0 \quad (4.6)$$

on the fine grid. Applying this idea recursively on several different grids, one obtains the monotone multigrid method (MMG) for linear complementary problems (3.18) as follows.

**ALGORITHM 4.2. MMG $_\ell$**  ( $\nu$ -th cycle on level  $\ell \geq 1$ )

Let  $\mathbf{u}_\ell^\nu \in V_\ell$  be a given approximation.

1. *A priori smoothing and projection* :  $\mathbf{u}_\ell^{\nu,1} := (\text{pGS}(\mathbf{u}_\ell^\nu))^{\eta_1}$ .
2. *Coarse grid correction*:
 
$$\begin{aligned} \mathbf{d}_{\ell-1} &:= r(\mathbf{f}_\ell - B_\ell \mathbf{u}_\ell^{\nu,1}), \\ \mathbf{g}_{\ell-1} &:= \tilde{r}(\mathbf{g}_\ell - \mathbf{u}_\ell^{\nu,1}), \\ B_{\ell-1} &:= r B_\ell p. \end{aligned}$$

If  $\ell = 1$ , solve exactly the linear complementary problem (4.4) and set  $\mathbf{v}_{\ell-1} := \mathbf{v}$ .  
If  $\ell > 1$ , do  $\zeta$  steps of **MMG** $_{\ell-1}$  with initial value  $\mathbf{u}_{\ell-1}^0 := 0$  and solution  $\mathbf{v}_{\ell-1}$ .

Set  $\mathbf{u}_\ell^{\nu,2} := \mathbf{u}_\ell^{\nu,1} + p\mathbf{v}_{\ell-1}$ .

3. *A posteriori smoothing and projection* :  $\mathbf{u}_\ell^{\nu,3} := (\text{pGS}(\mathbf{u}_\ell^{\nu,2}))^{\eta_2}$ .  
Set  $\mathbf{u}_\ell^{\nu+1} := \mathbf{u}_\ell^{\nu,3}$ .

The number of a priori and a posteriori smoothing steps is denoted by  $\eta_1$  and  $\eta_2$ , respectively. For  $\zeta = 1$  one obtains a V-cycle, for  $\zeta = 2$  a W-cycle.

Condition (4.6) leads to an inner approximation of the solution set  $\mathbf{K}_L$  and ensures that the multigrid scheme is robust. Striving for optimal multigrid efficiency, satisfaction of the constraint should *not* be checked by interpolating  $\mathbf{v}_\ell$  back to the finest grid. Instead, special restriction operators  $\tilde{r}$  are needed for the obstacle function, see [Ko]. A corresponding construction for B-splines of general order  $k$  can be found in [HK1].

**5. Numerical Results.** Finally, we wish to present some numerical results which are supposed to answer the following questions: 1) Validation of the model concerning the quality of the fair prize and the free boundary; 2) Accuracy of the computed solution; 3) Convergence behaviour of the iterative methods in view of the unsymmetry of the system matrix  $B$ .

**5.1. Validation of the Model.** Firstly, we want to investigate whether our computed solution can be validated. Since there exists no exact solution of Problem 2.2 in closed form, we compare our solution with the situation from [CP]. This may by now be considered as a benchmark problem for numerically computing American put options with the following parameters: strike price  $K = 10$ , volatility of the volatility  $\xi = 0.9$ , price of the volatility risk  $\lambda = 0$ , mean reversion rate  $\kappa = 5$ , mean reversion level  $\theta = 0.16$ , correlation parameter  $\rho = 0.1$ , interest rate  $r = 0.1$  and end time  $T = 0.25$ . (Note that this choice of parameters is different from the one in [DPS], see the remarks before (2.20), for which we are, however, not aware of values for comparison.)

We consider an American put option for a stock price of today as  $S(0) = 10$ , i.e., the transformed stock price is  $x(0) = 0$ , and the variance of today  $v(0) = 0.25$ . Thus, the price of the option of today is  $u(0, 0, 0.25)$ . In the literature, we find the following prices computed on different uniform grids for the solution after the last time step at  $t = 0$ :

Grid for $(S, v)$	Price	Reference
$256 \times 256$	0.796	[Oo]
$177 \times 103$	0.7961	[ZFV]
$320 \times 128$	0.7581	[IT1]

In [IT1] one has computed on this grid 64 time steps on the interval  $[0, T] = [0, 0.25]$ . The other references do not explicitly mention the amount of time steps. Since it is not known which of these solutions is indeed the ‘best’, we take as reference the value  $u_{\text{ref}} = u(0, 0, 0.25) = 0.7581$  from [IT1].

The discretized solution of Problem 3.1 has been computed on the domain  $\Omega = (-5, 5) \times (0.0025, 0.4975)$ , i.e. we have chosen  $x_{\max} = 5$ ,  $v_{\min} = 0.0025$  and  $v_{\max} = 0.4975$ . We have used the Crank-Nicolson scheme for  $\theta = \frac{1}{2}$ . Since the convection term in (3.12) often causes stability problems, we have used a *graded grid* like in [AWW, IT1]. Here in the direction of the transformed stock price  $x$ , we have used a grid which is four times as large as the grid for the variance  $v$ . The solution has been computed with a monotone multigrid method on level  $L = 4$ .

We plot the discretized solution of Problem 3.1 in Fig. 5.1 and 5.2. We see in Fig. 5.1 the option price for the American put option depending on stock price  $S$  (axis to the left on the bottom) and the variance  $v$  (axis to the right on the bottom) for the two most relevant times  $t = T = 0.25$  (end time) and  $t = 0$  (today). The left graphic shows the transformed payoff function  $g$  at end time  $T = 0.25$ . We recognize the characteristic bend of this termination condition. In the figure on the right, we clearly see that the bend has been smoothed out which is a consequence of the parabolicity of the solution operator. Also, the diffusion effect is stronger in areas of larger variance.

This diffusion effect is even stronger visible for the value of the option depending on the non-transformed stock price in Fig. 5.2. For better visualization, here the axis for the stock price  $S$  is slightly shortened. The left graphic shows the payoff function  $\mathcal{H}$  at end time  $T = 0.25$ , and the right figure visualizes the option price at current time  $t = 0$ . We strongly see the diffusion effect.

Next we want to investigate the free boundary introduced in (2.22). In Fig. 5.3 in the left graphic, we see two functions: the option at time  $t = 0$  (upper function) and the transformed payoff function  $g$  (lower function). As we know from Section 3, for transformed stock prices which are worth less than the transformed free boundary  $x_f$ , we have at any time  $0 \leq t^{(k)} < T$

$$u(t^{(k)}, x, v) = g(x) \quad \text{for } x \leq x_f(t^{(k)}, v), v \geq 0. \quad (5.1)$$

As we can see in the figure, the option and the obstacle  $g$  coincide for  $x < x_f$ . For growing  $x$ , we see that  $g$  lies below the solution  $u(0, x, v)$ . This is in accordance with the theoretical results since the option price satisfies  $\mathcal{Y}^{(k)}u^{(k)} = 0$  for  $x \leq x_f(t^{(k)})$ ,  $v \geq 0$ , and every  $0 \leq t^{(k)} < T$ . We can observe this projection even more strongly in the right graphic in Fig. 5.3. Here we see the value of the option at different times for a constant variance  $v = 0.4975$ . The lowest plot displays the payoff function  $g$ . Since the stock price of today  $S(0)$  corresponds to the strike price  $K$ , we have for the transformed stock price  $x = 0$ ; the characteristic bend appears here therefore at

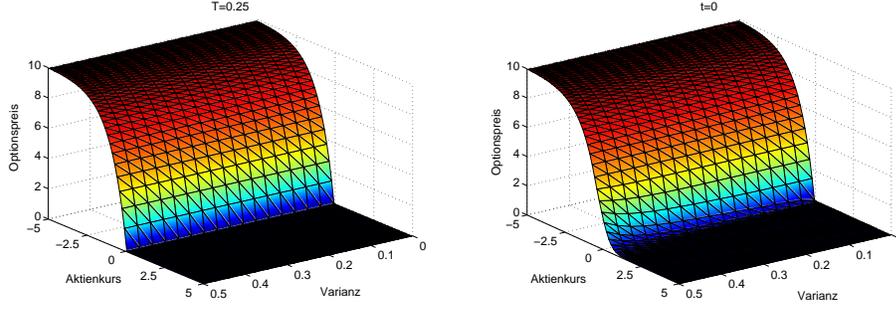


FIG. 5.1. Left: Transformed payoff function  $g(x)$  according to (2.28) of the American put option at time  $T = 0.25$ . Right: Option price of today at  $t = 0$ .

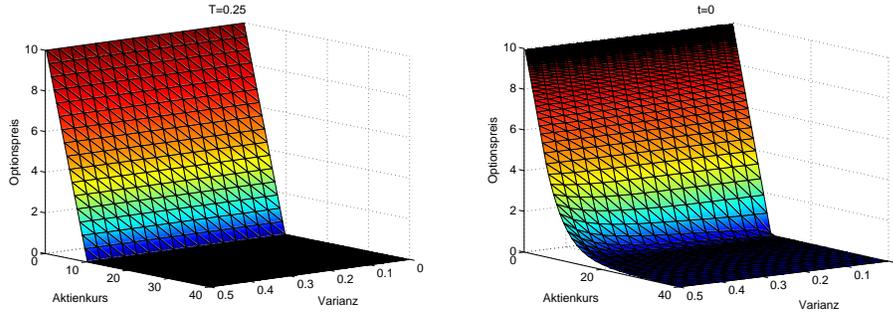


FIG. 5.2. Left: Payoff function  $\mathcal{H}$  of the American put option at time  $T = 0.25$ . Right: Option price for today  $t = 0$  without transformation.

$x = 0$ . We clearly see how this discontinuity is smoothed as time goes backward from  $T = 0.25$  to  $t = 0$ . We also see how in the different time steps the option price is projected to the payoff function.

**5.2. Accuracy of the Solution.** In all iterative methods, we employ a stopping parameter `tol` whose choice should be coupled to the discretization error of the finite element discretization. The residual of the resulting linear systems of equation is measured in the Euclidean norm, i.e.,

$$\|C\mathbf{u} - \mathbf{f}\|_2 \quad (5.2)$$

which corresponds to an a priori estimation

$$\|u^{(k)} - u_h^{(k)}\|_{L_2(\Omega)} = \mathcal{O}(h^2). \quad (5.3)$$

Accordingly, we should choose `tol` proportional to  $h^2 = 2^{-2\ell}$ . If the error is measured in the  $H^1(\Omega)$  norm, one should choose `tol` proportional to  $h = 2^{-\ell}$ . In our experiments, however, it turned out that `tol` =  $10^{-3} 2^{-\ell}$  provided the most accurate and most efficient results. For level  $\ell$ , we have  $N_\ell = (2^\ell + 1)(2^{\ell+2} + 1)$  grid points in our triangulation using the graded mesh. The coarsest grid for  $\ell = 2$  has 85 and the finest grid for  $\ell = 7$  has 66177 grid points.

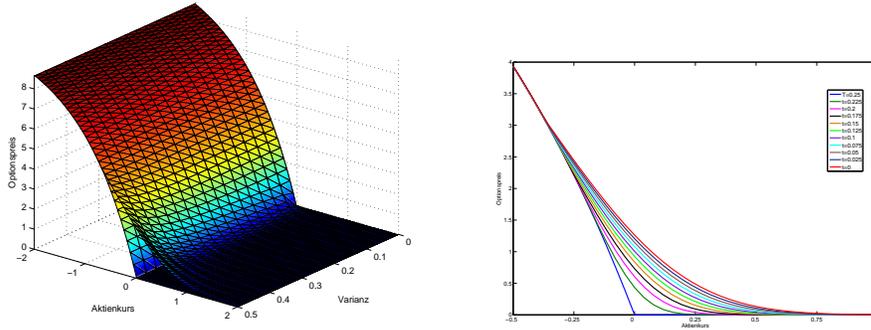


FIG. 5.3. Left: Visualization of computed solution at  $t = 0$  (upper plot) and transformed payoff function  $g$  (lower plot); Right: Evolution of the option price for different times  $t$  for constant variance  $v = 0.4975$ .

We display in Table 5.1 the results for computing the American put option, using just the simple projective Gauss-Seidel method from Algorithm 4.1. This and the following table is set up as follows. The first column shows the discretization level  $\ell$  on which the system is solved. The second column displays the approximate value of the option at time  $t = 0$  of today. The third and fourth columns shows the absolute and relative errors between our computed solution and the reference value,

$$e_{\ell,\text{abs}} := |u_{\ell}^{(0)} - u_{\text{ref}}|, \quad e_{\ell,\text{rel}} := \frac{|u_{\ell}^{(0)} - u_{\text{ref}}|}{u_{\text{ref}}}. \quad (5.4)$$

The next column displays the number of iterations, averaged over all time steps. Here we have  $T/\tau = 0.25/0.025 = 10$  time steps, and we define

$$\varnothing(\# \text{PSOR}) := \frac{\# \text{ all iterations in every time step}}{10}. \quad (5.5)$$

In view of the growth of the spectral condition number like  $\mathcal{O}(2^{-2\ell})$ , this number should quadruple when passing from one level to the next higher one. We also display the *experimental order of convergence*

$$\text{ECO}_{\ell} := \frac{\ln |e_{\ell,\text{abs}} - e_{\ell-1,\text{abs}}|}{\ln(2)} \quad (5.6)$$

which describes the expected order. The last column shows the total time in seconds the method needs. We remark that the method behaves exactly as predicted, with an experimental order of convergence around 1.5 for the chosen tolerance `tol`. We have also tested two simple multiscale methods which can be found in [Br]. One method called *Nested Iteration* is based on iterating on level  $\ell$  only up to discretization error accuracy given by (3.19), and starting the iteration on level  $\ell + 1$  with the solution vector from level  $\ell$ , prolonged to this finer level  $\ell + 1$ . A variant called *Cascade* is based on selecting a certain amount of iterations before continuing to the next level. In the numerical experiments, we can see that the difference of the projective Gauss-Seidel scheme to both of these methods perform is neglectible. So we conclude that for the present situation, these methods do not have an effect.

$\ell$	$u_\ell^{(0)}$	$e_{\ell,\text{abs}}$	$e_{\ell,\text{rel}}$	$\varnothing$ (# PSOR)	$\text{ECO}_\ell$	Time (sec.)
2	0.424602	$3.71208e-01$	$4.66691e-01$	6	—	0.624329
3	0.684257	$1.11553e-01$	$1.40175e-01$	11	1.6	3.21551
4	0.774593	$2.1217e-02$	$2.66608e-02$	26	2.4	22.5351
5	0.789919	$5.891e-03$	$7.40252e-03$	92	1.8	260.453
6	0.793706	$2.104e-03$	$2.64384e-03$	353	1.4	5197.12
7	0.794969	$8.41e-04$	$1.05678e-03$	1388	1.3	111705

TABLE 5.1

Results using the projected Gauss-Seidel method from Algorithm 4.1 for  $\text{tol} = 10^{-3} 2^{-\ell}$ .

$\ell$	$u_\ell^{(0)}$	$e_{\ell,\text{abs}}$	$e_{\ell,\text{rel}}$	$\text{ECO}_\ell$	Time (sec.)
2	0.422206	$3.73604e-01$	$4.69463e-01$	—	0.192924
3	0.682997	$1.12813e-01$	$1.41758e-01$	1.7	0.647031
4	0.771826	$2.3984e-02$	$3.01378e-02$	2.2	2.86817
5	0.791265	$4.545e-03$	$5.71116e-03$	2.3	17.1439
6	0.794925	$8.85e-04$	$1.11207e-03$	2.3	154.989
7	0.795687	$1.23e-04$	$1.54559e-04$	2.8	3064.73

TABLE 5.2

Results for the monotone multigrid method with 2 pre and 3 post smoothing steps.

In Table 5.2 we display the results obtained with the monotone multigrid method from Algorithm 4.2. We have used  $\eta_1 = 2$  pre and  $\eta_2 = 3$  post smoothing steps and a V cycle. On the lowest level  $\ell = 2$ , we solve the linear complementary problem with the projected Gauss-Seidel method. We wish to point out that the same accuracy can be achieved with the monotone multigrid method in less than 1/30 of the time needed for the projected Gauss-Seidel method.

**5.3. Convergence Behaviour of the Iterative Schemes.** Finally, we display two graphics showing the convergence behaviour of the schemes considered. Fig. 5.4 shows on the left the development of the absolute error  $e_{\ell,\text{abs}}$  on the vertical axis over the time on the horizontal axis in a log-log plot. We observe that the projected Gauss-Seidel method as well as the simple multiscale variants show a very similar behaviour while the monotone multigrid scheme (MMG) is substantially faster. In fact, one can observe a multigrid efficiency which is optimal or close to it. This we believe is a consequence of the non-symmetry of the system matrix  $B$  which, however, is not so bad that the methods would not converge any more. We show on the right of this graphic the time (on the vertical axis) which is needed by each method to solve one linear system of equations over the number of unknowns (horizontal axis). We see that the MMG method here also outperforms all the other methods.

All algorithms have been run on a Intel(R) Core(TM) 2 CPU with 2,40 GHz in Matlab 7.3.0 (R2006b).

**6. Final Remarks and Outlook.** A next step for the present model would be to develop monotone multigrid methods of higher order to compute the Greeks with high precision as in [HK1, HK2]. More modern models beyond classical Black-Scholes equations are jump-diffusion price processes. In these, the Wiener process in the Black-Scholes model is replaced by a Lévy process. According to [Eb], such

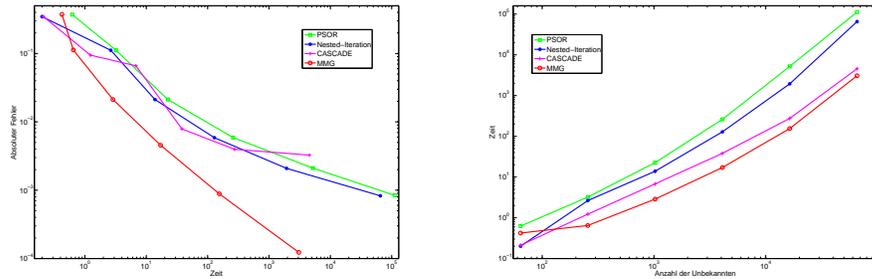


FIG. 5.4. *Left: Absolute error (vertical) over time (horizontal) as log-log plot. Right: Time to solve linear system over number of unknowns as log-log plot.*

processes allow for a more realistic modelling of the dynamics of asset prices, see [G] for the recent state-of-the art on the stochastic analysis models. American options on Lévy driven assets (with constant volatility) leading to integrodifferential equations were valued by means of Galerkin methods in space and continuous piecewise linear spline-wavelets in [MNS]. These discretizations have the advantage that the differential operator can be optimally preconditioned independent of the grid size like multigrid methods while simultaneously allowing for compression of the integral operator. Higher order spline-wavelets would again allow for a determination of the Greeks of high precision.

Ultimately, in order to avoid the time-stepping procedure, it would be interesting to formulate the variational inequality in full weak space-time form for which then convergent adaptive wavelet methods of optimal complexity along the lines of [SS] would be our method of choice. However, it is not clear how to develop a convergent adaptive scheme for the treatment of inequalities, see [RU] for first steps in this direction for elliptic variational inequalities.

**Acknowledgment.** We thank Christoph Schwab who originally pointed out the problem of stochastic volatility after completion of the papers [HK1, HK2] dealing with monotone multigrid methods with higher order B-Splines. We also thank Ernst Eberlein for some comments concerning the modelling of the problem and two anonymous referees for their useful remarks.

#### REFERENCES

- [AP1] Y. ACHDOU, O. PIRONNEAU, *Computational Methods for Option Pricing*, Siam, 2005.
- [AGG1] F. AITSAHLIA, M. GOSWAMI, S. GUHA, *American option pricing under stochastic volatility: an efficient numerical approach*, Computational Management Science, to appear, 2010.
- [AGG2] F. AITSAHLIA, M. GOSWAMI, S. GUHA, *American option pricing under stochastic volatility: an empirical evaluation*, Computational Management Science, to appear, 2010.
- [AP2] L. B. G. ANDERSEN, V. V. PITERBARG, *Moment explosions in stochastic volatility models*, Finance and Stochastics, Vol. 11, No. 1, pp. 29-50, 2007.
- [AWW] T. APEL, G. WINKLER, U. WYSTUP, *Valuation of options in Heston's stochastic volatility model using finite element methods*, Foreign Exchange Risk. Risk Publications, London, November 2001.
- [BR] C.A. BALL, A. ROMA, *Stochastic volatility option pricing*, J. Financial Quantitative Anal. 29, 1994, 589–607.
- [BS] F. BLACK, M. SCHOLES, *The pricing of options and corporate liabilities*, J. Political Economy 81, 1973, 637–654.
- [Br] D. BRAESS, *Finte Elements*, Cambridge University Press, Second edition, 2001.

- [BHR] F. BREZZI, W. HAGER, P.A. RAVIART, *Error estimates for the finite element solution of variational inequalities*, Numer. Math. 28, 1977, 431–443.
- [CP] N. CLARKE, K. PARROT, *Multigrid for American option pricing with stochastic volatility*, Applied Mathematical Finance 6, 1999, 177–195.
- [CIR] J.C. COX, J.E. INGERSOLL, S.A. ROSS, *A theory of the term structure of interest rate*, Econometrica 53, 1985, 385–407.
- [Cr] C.W. CRYER, *The solution of a quadratic programming problem using systematic overrelaxation*, Siam J. Contr. 9, 1971, 385–392.
- [CRZ] J. CVITANIĆ, B. ROZOVSKII, I. ZALIAPIN, *Numerical estimation of volatility values from discretely observed diffusion data*, J. Comp. Finance 9(4), 2006, 1–36.
- [D] W. DAHMEN, *Wavelet and multiscale methods for operator equations*, Acta Numerica, 1997, 55–228.
- [DPS] D. DUFFIE, J. PAN, K. SINGLETON, *Transform analysis and asset pricing for affine jump-diffusions*, Econometrica, Vol. 68, No. 6, pp. 1343–1376, 2000.
- [Eb] E. EBERLEIN, *Application of generalized hyperbolic Lévy motions to finance in Lévy processes: Theory and Applications*, O.E. Barndorff-Nielsen, T. Mikosch, S. Resnick (eds.), Birkhäuser 2001, 319–337.
- [EO] C.M. ELLIOTT, J.R. OCKENDON, *Weak and Variational Methods for Moving Boundary Value Problems*, Pitman, 1982.
- [FPS] J.-P. FOUQUE, G. PAPANICOLAOU, K. R. SIRCAR, *Derivatives in Financial Markets with Stochastic Volatility*, Cambridge University Press, Cambridge, 2000.
- [G] K. GLAU, *Feynman-Kac-Darstellungen zur Optionspreisbewertung in Lévy-Modellen (in German)*, Dissertation, Department of Mathematical Stochastics, University of Freiburg, 2010.
- [GLT] R. GLOWINSKI, J.-L. LIONS, R. TRÉMOLIÈRES, *Numerical Analysis of Variational Inequalities*, North-Holland, 1981.
- [Ha] W. HACKBUSCH, *Iterative Solution of Large Sparse Systems of Equations*, Springer, 1994.
- [HKS] N. HILBER, S. KEHTARI, C. SCHWAB, C. WINTER, *Wavelet finite element method for option pricing in highdimensional diffusion market models*, Research Report No. 2010-01, Seminar für Angewandte Mathematik, ETH Zürich, 2010.
- [HK1] M. HOLTZ, A. KUNOTH, *B-spline based monotone multigrid methods*, Siam J. Num. Anal., 45(1), 2007, 1175–1199.
- [HK2] M. HOLTZ, A. KUNOTH, *B-spline based monotone multigrid methods with an application to the pricing of American options*, In: Multigrid, Multilevel and Multiscale Methods, (Electr.) Proc. EMG, 2005, P. Wesseling, C.W. Oosterlee and P. Hemker (Eds.).
- [He] S.L. HESTON, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, The Review of Financial Studies 6(2), 1993, 327–343.
- [Hu] J.C. HULL, *Options, Futures and Other Derivates*, Prentice Hall International Editions, 2000.
- [IT1] S. IKONEN, J. TOIVANEN, *Operator splitting methods for American options with stochastic volatility*, European Congress on Computational Methods in Applied Sciences and Engineering, 2004.
- [IT2] S. IKONEN, J. TOIVANEN, *Efficient Numerical Methods for Pricing American Options Under Stochastic Volatility*, Wiley InterScience, 2007.
- [ItT] K. ITO, J. TOIVANEN, *Lagrange multiplier approach with optimized finite difference stencils for pricing American options under stochastic volatility*, Siam J. Scientif. Comp. 31(4), 2009, 2646–2664.
- [KS] D. KINDERLEHRER, G. STAMPACCHIA, *An Introduction to Variational Inequalities and Their Applications*, Academic Press, 1980.
- [Ko] R. KORNUBER, *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*, Teubner, 1997.
- [Ku] A. KUNOTH, *Optimized wavelet preconditioning*, in: Multiscale, Nonlinear and Adaptive Approximation, R.A.DeVore, A. Kunoth (eds.), Springer, 2009, 325–378.
- [LS] F.A. LONGSTAFF, E.S. SCHWARTZ, *Valuing American options by simulation: A simple least-squares approach*, Rev Fin 14, 2001, 113–147.
- [Ma] J. MANDEL, *A multilevel iterative method for symmetric, positive definite linear complementarity problems*, Appl. Math. Optimization 11, 1984, 77–95.
- [MNS] A.-M. MATACHE, P.-A. NITSCHKE, C. SCHWAB, *Wavelet Galerkin Pricing of American Options on Lévy Driven Assets*, Quantitative Finance 5(4), 2005, 403–424.
- [Ni] W. NIETHAMMER, *Relaxation bei unsymmetrischen Matrizen*, Math. Zeitschrift 85, 1964, 319–327.
- [OK] B. ØKSENDAL, *Stochastic Differential Equations*, Springer, 2003.

- [Oo] C.W. OOSTERLEE, *On multigrid for linear complementary problems with application to American-style options*, Electronics Transactions on Numerical Analysis 15, 2003, 165–185.
- [PS] J. PERSSON, L. VON SYDOW, *Pricing American options using a space-time adaptive finite difference method*, Mathematics and Computers in Simulation 80 (9), 2010, 1922–1935.
- [RR] M. RENARDY, R.C. ROGERS, *An Introduction to Partial Differential Equations*, Springer, 1993.
- [RW] C. REISINGER, G. WITTUM, *Efficient hierarchical approximation of high-dimensional option pricing problems*, Siam J. Scient. Comput. 29, 2007, 440–458.
- [RU] M. ROMETSCH, K. URBAN, *Adaptive wavelet methods for elliptic variational inequalities I: Analysis*, University of Ulm, Preprint, 2010.
- [Sch] CHR. SCHNEIDER, *Multiskalenmethoden zur Bewertung Amerikanischer Optionen mit stochastischer Volatilität* (in German), Diploma Thesis, Institut für Angewandte Mathematik, Universität Bonn, Germany, 2008.
- [SS] CH. SCHWAB, R. STEVENSON, *Space–time adaptive wavelet methods for parabolic evolution equations*, Math. Comp. 78, 2009, 1293–1318.
- [S] R. SEYDEL, *Tools for Computational Finance*, Springer, 2009.
- [Wi] P. WILMOTT, *Paul Wilmott introduces Quantitative Finance*, John Wiley and Sons Ltd., 2007
- [ZfV] R. ZVAN, P. A. FORSYTH, K.R. VETZAL, *Penalty methods for American options with stochastic volatility*, J. Comput. Appl. Math. 91, 1998, 199–218.