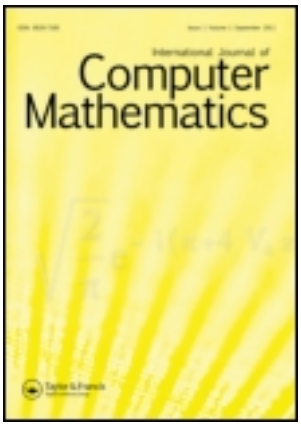
This article was downloaded by: [Universitatsbibliothek Paderborn]
On: 27 July 2012, At: 06:53
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Computer Mathematics

Publication details, including instructions for authors and subscription information:
http://tandfonline.com/loi/gcom20

Multiscale methods for the valuation of American options with stochastic volatility

Angela Kunoth [a], Christian Schneider [a] & Katharina Wiechers [a]

[a] Institut für Mathematik, Universität Paderborn, Warburger Str. 100, 33098, Paderborn, Germany

Version of record first published: 24 Apr 2012

To cite this article: Angela Kunoth, Christian Schneider & Katharina Wiechers (2012): Multiscale methods for the valuation of American options with stochastic volatility, International Journal of Computer Mathematics, 89:9, 1145-1163

To link to this article: http://dx.doi.org/10.1080/00207160.2012.672732

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

International Journal of Computer Mathematics
Vol. 89, No. 9, June 2012, 1145–1163

Taylor & Francis
Taylor & Francis Group

Multiscale methods for the valuation of American options with stochastic volatility

Angela Kunoth*, Christian Schneider and Katharina Wiechers

Institut für Mathematik, Universität Paderborn, Warburger Str. 100, 33098 Paderborn, Germany

(*Received 23 September 2011; revised version received 10 December 2011; accepted 27 February 2012*)

This paper deals with the efficient valuation of American options. We adopt Heston's approach for a model of stochastic volatility, leading to a generalized Black–Scholes equation called Heston's equation. Together with appropriate boundary conditions, this can be formulated as 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.

Keywords: American option pricing; stochastic volatility; Heston's model; parabolic boundary value problem; free boundary; monotone multigrid method; multigrid efficiency

2010 AMS Subject Classification: 65M55; 35J86; 65N30; 65D07; 91B70

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

1. Introduction

In Mathematical Finance, the fair pricing of an American option can be formulated as a parabolic boundary value problem involving the Black–Scholes equation [7] with a free boundary in a standard model. One seeks to compute the free boundary, the optimal exercise price of the option, together with the solution of the partial differential equation (PDE), the value of the option. For American put options, one generally does not have closed-form solutions, so numerical schemes have to be employed. 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 [29]. There are several approaches to estimate the volatility from observed stock data, see, e.g. [14]. Here, we adopt Heston's approach [25] to model the volatility to satisfy a stochastic differential equation (SDE). Empirical evaluations show that Heston's model for American options with stochastic volatility provides a substantial improvement for option pricing compared with the simple Black–Scholes model [3]. Extensive information on models for stochastic volatility can be found in [19]. For Heston's model and one asset, the standard Itō approach yields a variational inequality in two space variables with a non-symmetric parabolic differential operator.

*Corresponding author. Email: kunoth@math.uni-paderborn.de

ISSN 0020-7160 print/ISSN 1029-0265 online
© 2012 Taylor & Francis
http://dx.doi.org/10.1080/00207160.2012.672732
http://www.tandfonline.com

Numerical schemes for American option pricing are typically based on finite-difference approaches, see, e.g. [11,30–32,40–42]. In [30], an operator splitting method was 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 [11,40], multigrid methods were developed. In [42], the focus was 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 [26]. A comparison of the described methods can be found in [31]. In [32], a finite-difference discretization in space and a Rannacher time-stepping scheme were 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 European options was proposed in [5]. In [48], finite elements for the second-order terms and finite volumes for the first-order terms together with two penalty methods for linear complementary problems were considered to value American options with stochastic volatility. A rich source for numerical methods for option pricing problems is [1].

A somewhat different approach is the one proposed in [36], 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. [2].

The purpose of this paper is twofold. First, we derive a variational inequality for the American option pricing problem, which we then discretize in terms of linear finite elements with respect to space. Second, we solve the resulting linear inequality system with optimal linear computational complexity. Based on a projective Gauss–Seidel scheme, we investigate different multiscale methods (nested iteration and a cascadic type scheme) to accelerate the solution. Specific emphasis is placed on an appropriate monotone multigrid (MMG) method [34], 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 recalling some basic notions and standard assumptions, see, e.g. [46]. 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 a *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 $\mathrm{Var}\,(X) = E((X - E(X))^2)$ its variance. The standard deviation (of the variance) or *volatility* σ is defined as $\sigma := \sqrt{\mathrm{Var}\,(X)}$. The (real-valued and non-negative) variance underlying the stock is denoted by $v = v(t) = v(X(t))$. We would like 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 [7] requires that the financial market is *frictionless*, that is, there are no transaction costs or taxes, that 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*, for example, immediate riskless gains are not possible.

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

For *constant volatility* $\sigma = \sqrt{v}$, the value of the option, described by the option price function $V(t, S(t))$ depending on the stock $S(t)$, is given by *the Black–Scholes equation:*

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0; \tag{1}$$

see, e.g. [29].

2.2 *Heston's 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 [12], see also [6]. A stochastic process is called a *CIR* (*Cox–Ingersoll–Ross*) *process* if it satisfies an SDE:

$$\mathrm{d}Y(t) = \kappa(\gamma - Y(t))\,\mathrm{d}t + \xi\sqrt{Y(t)}\,\mathrm{d}W(t), \tag{2}$$

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

Now, let the variance v be described by a CIR process (2). From the resulting model, *Heston's equation* is a parabolic PDE with diffusion, convection and reaction terms of the form

$$\begin{aligned}\mathcal{L}V := \frac{\partial V}{\partial t} + \frac{1}{2}\left(S^2 v\frac{\partial^2 V}{\partial S^2} + 2\rho\xi v S\frac{\partial^2 V}{\partial S\,\partial v} + \xi^2 v\frac{\partial^2 V}{\partial v^2}\right) + rS\frac{\partial V}{\partial S} + \kappa(\gamma - v)\frac{\partial V}{\partial v} - rV \\ = 0,\end{aligned} \tag{3}$$

see [25] or e.g. [10,20]. For European options, a closed-form solution can be found in [25]. Moreover, Merton's theorem, see, e.g. [29], says that the value of an American call option equals the value of a European call option.

Therefore, in the following, we concentrate on the most interesting case from a computational point of view: American put options satisfying Equation (3) for which there is no closed-form solution available. For this parabolic PDE to be solved backwards in time, we need to specify additional constraints which uniquely determine the solution.

Remark 2.1 (a) Before doing so, we would like to point out that in contrast to many physical problems described by PDEs like the heat equation, see, e.g. [43], the issue of appropriate boundary and end conditions is still subject to discussions, see, e.g. [24,47], as well as the choice of the relation of the parameters ξ, κ and γ in *Feller's condition*.

(b) We consider the main contribution of this paper by providing a fast solution method for a parabolic free boundary value problem which requires that the theoretical set-up is such that the problem admits a unique solution. In view of this aspect, a particular choice of boundary or end conditions or parameters does not affect the design and performance of the scheme.

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

$$V(T, S, v) = \mathcal{H}(S). \tag{4}$$

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

Remark 2.2 Note that due to the same sign in front of the first-order derivative in time and second-order derivative with respect to space, posing end conditions at T for Equation (3) is a well-posed parabolic problem (in contrast to the backwards heat equation, see, e.g. [43]).

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 $S \to 0$. The maximal achievable gain is K. Accordingly, it is best to exercise immediately to avoid discounting K. This yields

$$\lim_{S\to 0} V(t, S, v) = \lim_{S\to 0} \mathcal{H}(S) = K. \tag{5}$$

For $S \to \infty$, the option is worthless, so we choose

$$\lim_{S\to\infty} V(t, S, v) = \lim_{S\to\infty} \mathcal{H}(S) = 0. \tag{6}$$

We would like to mention that in [11], replacement of condition (6) by the Neumann boundary condition $\lim_{S\to\infty}(\partial V(t, S, v)/\partial S) = 0$ was proposed 'for practical reasons', which has not been clear to us. As one referee pointed out, one can argue that 'we deal with computational domains of limited size', that is, for practical reasons, the maximal attainable value for S is 'usually not so big' and then a 'Neumann condition is simply more accurate, or even a second derivative set to zero.'

Next, we discuss boundary conditions for V with respect to v. This is an issue which is not so clear and which is still subject to discussion; in fact, one referee called this 'a well-known dilemma' in Heston's model. An extensive treatment from both the mathematical and the financial points of view has been provided recently in [47]. Some argue as follows: take a look at the diffusion process $v(t)$ and see whether $v = 0$ is an attainable boundary or not. *Feller's condition*, see, e.g. [4], states that $v = 0$ is attainable if $\xi^2 > 2\kappa\gamma$ is fullfilled. 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 Equation (3) so that Feller's condition is fullfilled. Choosing the parameters, that is, $\xi = 0.61$, $\kappa = 6.21$ and $\gamma = 0.019$ for Heston's model, is according to Duffie *et al.* [16] the most reasonable. For this choice, Feller's condition is fullfilled and, thus, it is possible to state a boundary condition for $v = 0$. We choose the boundary condition

$$\lim_{v\to 0} V(t, S, v) = \mathcal{H}(S). \tag{7}$$

The extensive discussion in [10,47], Chapter 2.2, citing [11,32], supports this choice and even proposes $\lim_{v\to 0} V(t, S, v) = 0$, which it is argued may be viewed as a simplified version of Equation (7). Other studies argue that it makes more sense from a financial point of view that Feller's condition is not satisfied. Then, however, the question as to how to actually choose the boundary conditions remains. We do not wish to enter into this discussion, but point to Remark 2.1(b) instead.

Finally, as in [11], one expects that for $v \to \infty$, the option price V would be insensitive to a change in the volatility, that is,

$$\lim_{v\to\infty} \frac{\partial V(t, S, v)}{\partial v} = 0. \tag{8}$$

2.3 *American put options in Heston's model*

For American options (and fixed variance), it is well known that for each time $t \in [0, T]$, there always exists a stock price S for which early exertion before final time T is advantageous, see, e.g. [46]. One can show that these values define a continuous curve $S_f(t)$. It is *a priori* unknown

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

and therefore defines a *free boundary*. Accordingly, one should exercise the option in the areas where $S \le S_f$ and hold the option if $S > S_f$. We are, therefore, finally led to the following option pricing problem:

PROBLEM 2.3 *Solve for $V = V(t, S, v)$ the system*

$$\begin{aligned} \mathcal{L}V = 0 \qquad & \text{for } S > S_f,\ v \ge 0,\ t \in [0, T) \\ V(t, S, v) = \mathcal{H}(S) \ & \text{for } S \le S_f,\ v \ge 0,\ t \in [0, T) \end{aligned} \tag{9}$$

with boundary conditions

$$\begin{aligned} &\lim_{S\to 0} V(t, S, v) = \lim_{S\to 0} \mathcal{H}(S) = K, \quad \lim_{S\to\infty} V(t, S, v) = \lim_{S\to\infty} \mathcal{H}(S) = 0, \\ &\lim_{v\to 0} V(t, S, v) = \mathcal{H}(S), \quad \lim_{v\to\infty} \frac{\partial V(t, S, v)}{\partial v} = 0. \end{aligned} \tag{10}$$

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

2.4 *Free boundary value and linear complementary problems*

It is useful to consider the following equivalent formulation as a *linear complementary problem* instead of Problem 2.3, see, e.g. [46] 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 later will be appropriately cut.)

PROBLEM 2.4 *Solve for $V = V(t, S, v)$, $(t, S, v) \in \Omega_{\mathcal{L}}^{\infty}$ the system*

$$\begin{aligned} \mathcal{L}V(V - \mathcal{H}) &= 0 \\ \mathcal{L}V &\le 0 \\ V - \mathcal{H} &\ge 0 \end{aligned} \tag{11}$$

with boundary conditions (10) 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 Equation (3) 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* $x = \ln(S/K)$. This results in the disappearance of the variable coefficient S in Equation (3). The transformed value of the option $y = y(t, x, v)$ then satisfies the transformed Heston equation:

$$\frac{\partial y}{\partial t} + \frac{1}{2}\xi^2 v \frac{\partial^2 y}{\partial v^2} + \rho\xi v \frac{\partial^2 y}{\partial v\, \partial x} + \frac{1}{2} v \frac{\partial^2 y}{\partial x^2} + \kappa(\gamma - v)\frac{\partial y}{\partial v} + \left(r - \frac{1}{2}v\right)\frac{\partial y}{\partial x} - ry = 0 \tag{12}$$

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

$$A := \frac{1}{2} v \begin{pmatrix} \xi^2 & \rho\xi \\ \rho\xi & 1 \end{pmatrix}, \quad \mathbf{b} := \begin{pmatrix} -\kappa(\gamma - v) + \frac{1}{2}\xi^2 \\ -r + \frac{1}{2}v + \frac{1}{2}\xi\rho \end{pmatrix}, \tag{13}$$

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

the parabolic PDE (14) can be rewritten as

$$\mathcal{Y}y := \frac{\partial y}{\partial t} + \nabla \cdot A\nabla y - \mathbf{b} \cdot \nabla y - ry = 0, \tag{14}$$

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

$$g(x) := \mathcal{H}(Ke^x) = \max(K - Ke^x, 0). \tag{15}$$

Moreover, the free boundary S_f in Equation (9) 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 2.5 *Solve for $y = y(t, x, v)$, $(t, x, v) \in \Omega_{\mathcal{Y}}^{\infty}$ the system*

$$\begin{aligned} \mathcal{Y}y(y - g) &= 0 \\ \mathcal{Y}y &\le 0 \\ y - g &\ge 0 \end{aligned} \tag{16}$$

with boundary conditions

$$\lim_{x \to \pm\infty} y(t, x, v) = \lim_{x \to \pm\infty} g(x),$$

$$y(t, x, 0) = g(x), \quad \lim_{v \to \infty} \frac{\partial y(t, x, v)}{\partial v} = 0$$

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

The question of the 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 Equation (14). Typically, one chooses $x_{\min} = -x_{\max}$ for symmetry reasons. The boundary conditions in Problem 2.5 must be adapted accordingly. Thus, defining

$$\begin{aligned} \Gamma &:= \{(x, v) \in \partial\Omega : v = v_{\min} \text{ or } x = x_{\min} \text{ or } x = x_{\max}\} \\ \Gamma_v &:= \partial\Omega \setminus \Gamma = \{(x, v) \in \partial\Omega : v = v_{\max}\}, \end{aligned} \tag{17}$$

the boundary conditions (10) read

$$y = g \quad \text{on } \Gamma, \qquad \frac{\partial y}{\partial v} = 0 \quad \text{on } \Gamma_v, \tag{18}$$

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

Remark 2.6 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 Equation (14) 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 the case of constant volatility since then Equation (14) reduces to the heat equation.

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

3. Variational formulation and discretization

In this section, we derive an appropriate variational formulation of Problem 2.5. Let $L_2(\Omega)$ be the usual space of Lebesgue measurable and square integrable functions on Ω and denote by $H^1(\Omega)$ the Sobolev space of first-order weak derivatives. We define

$$\tilde{\mathcal{K}} := \{\varphi \in H^1(\Omega) : \varphi \geq g,\ \varphi = g \text{ on } \Gamma\}, \tag{19}$$

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 2.5. We multiply the second equation in (16) by $\varphi - g$ (which does not change the sign) and integrate over Ω, yielding $\int_\Omega \mathcal{Y}y(\varphi - g)\,\mathrm{d}\Omega \leq 0$. Subtraction of the first equation in (16), integrated over Ω, that is, $\int_\Omega \mathcal{Y}y(y - g)\,\mathrm{d}\Omega = 0$, yields

$$\int_\Omega \mathcal{Y}y(\varphi - y)\,\mathrm{d}\Omega \leq 0, \tag{20}$$

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 Equation (15) is addressed in Remark 3.3.

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

$$\mathcal{K} := \{\varphi \in H^1(\Omega) : \varphi \geq 0,\ \varphi = 0 \text{ on } \Gamma\}. \tag{21}$$

Thus, instead of Problem 2.5, we consider the following.

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

$$u(t, \cdot, \cdot) \in \mathcal{K}, \quad \frac{\partial u}{\partial v} = 0 \quad \textit{on } \Gamma_v, \quad u(T, x, v) = 0 \quad \textit{on } \Omega, \tag{22}$$

the inequality

$$\int_\Omega \mathcal{Y}u(\varphi - u)\,\mathrm{d}\Omega \leq -\int_\Omega \mathcal{Y}g(\varphi - u)\,\mathrm{d}\Omega =: \ell(\varphi - u) \tag{23}$$

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

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

For discretization, we first discretize the time derivative in Equation (20) with a standard finite-difference scheme. We decompose the time interval into equidistant points $0 =: t^{(0)} < t^{(1)} < \cdots < t^{(M)} := T$ 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 θ-scheme for $\theta \in [0, 1]$

$$\begin{aligned}\ell(\varphi - u^{(k)}) \geq{}& \int_\Omega \frac{u^{(k+1)} - u^{(k)}}{\tau}(\varphi - u^{(k)})\,\mathrm{d}\Omega \\ &+ (1 - \theta)\int_\Omega (\nabla \cdot A\nabla u^{(k+1)} - \mathbf{b} \cdot \nabla u^{(k+1)} - ru^{(k+1)})(\varphi - u^{(k)})\,\mathrm{d}\Omega \\ &+ \theta \int_\Omega (\nabla \cdot A\nabla u^{(k)} - \mathbf{b} \cdot \nabla u^{(k)} - ru^{(k)})(\varphi - u^{(k)})\,\mathrm{d}\Omega.\end{aligned} \tag{24}$$

We recall that we would like to solve Equation (14) *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 Euler scheme and $\theta = 0$ yields 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

$$\mathcal{Y}^{(k)}u^{(k)} := -\theta\nabla\cdot A\nabla u^{(k)} + \theta\mathbf{b}\cdot\nabla u^{(k)} + (\theta r + \tau^{-1})u^{(k)} \tag{25}$$

and right-hand side

$$f^{(k+1)} := (1-\theta)\nabla\cdot A\nabla u^{(k+1)} - (1-\theta)\mathbf{b}\cdot\nabla u^{(k+1)} - ((1-\theta)r - \tau^{-1})u^{(k+1)} + \mathcal{Y}g, \tag{26}$$

Equation (24) is just

$$\int_\Omega \mathcal{Y}^{(k)}u^{(k)}(\varphi - u^{(k)})\,\mathrm{d}\Omega \geq \int_\Omega f^{(k+1)}(\varphi - u^{(k)})\,\mathrm{d}\Omega. \tag{27}$$

Applying Green's formula to the left-hand side with respect to the variables x and v, we obtain

$$\begin{aligned}\int_\Omega \mathcal{Y}^{(k)}u^{(k)}(\varphi - u^{(k)})\,\mathrm{d}\Omega = \theta\int_\Omega (A\nabla u^{(k)}\cdot\nabla(\varphi - u^{(k)}) + \mathbf{b}\cdot\nabla u^{(k)}(\varphi - u^{(k)}))\,\mathrm{d}\Omega \\ + \int_\Omega (\theta r + \tau^{-1})u^{(k)}(\varphi - u^{(k)})\,\mathrm{d}\Omega - \int_{\partial\Omega} A\frac{\partial u^{(k)}}{\partial\mathbf{n}}(\varphi - u^{(k)})\,\mathrm{d}s,\end{aligned} \tag{28}$$

where $\mathbf{n}$ denotes the outward normal derivative at $\partial\Omega$. Due to the choice of $\mathcal{K}$ from which φ and $u^{(k)}$ are chosen, the boundary terms on Γ vanish. The remaining boundary term on Γ_v vanishes due to the Neumann condition (22) for $u^{(k)}$.

Analogously, we obtain for the right-hand side,

$$\begin{aligned}\int_\Omega f^{(k+1)}(\varphi - u^{(k)})\,\mathrm{d}\Omega = -(1-\theta)\int_\Omega A\nabla u^{(k+1)}\cdot\nabla(\varphi - u^{(k)})\,\mathrm{d}\Omega \\ + \int_\Omega (-(1-\theta)\mathbf{b}\cdot\nabla u^{(k+1)} - ((1-\theta)r - \tau^{-1})u^{(k+1)})(\varphi - u^{(k)})\,\mathrm{d}\Omega \\ + \int_\Omega \frac{\partial g}{\partial t}(\varphi - u^{(k)})\,\mathrm{d}\Omega - \int_\Omega A\nabla g\cdot\nabla(\varphi - u^{(k)})\,\mathrm{d}\Omega \\ - \int_\Omega (\mathbf{b}\cdot\nabla g + rg)(\varphi - u^{(k)})\,\mathrm{d}\Omega + \int_{\partial\Omega} A\frac{\partial g}{\partial\mathbf{n}}(\varphi - u^{(k)})\,\mathrm{d}s.\end{aligned} \tag{29}$$

Defining the right-hand side of Equation (28) as the evaluation of a bilinear form $a(\cdot,\cdot)$, that is,

$$\begin{aligned}a(u^{(k)}, \varphi - u^{(k)}) := \theta\int_\Omega (A\nabla u^{(k)}\cdot\nabla(\varphi - u^{(k)}) + \mathbf{b}\cdot\nabla u^{(k)}(\varphi - u^{(k)}))\,\mathrm{d}\Omega \\ + \int_\Omega (\theta r + \tau^{-1})u^{(k)}(\varphi - u^{(k)})\,\mathrm{d}\Omega,\end{aligned} \tag{30}$$

we derived the following *variational inequality*.

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

$$a(u^{(k)}, \varphi - u^{(k)}) \geq \langle f^{(k+1)}, \varphi - u^{(k)}\rangle, \tag{31}$$

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

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

Remark 3.3 For the derivation of Problem 3.2, we assumed that the lower bound g is sufficiently smooth. In fact, the formulation (31) reveals that $g \in H^1(\Omega)$ is sufficient, which holds for g defined in Equation (15).

It is well known from the treatments on variational inequalities, see, e.g. [18,22,33], 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 Ω A is positive definite and bounded. This question and detailed conditions on the relation of the different coefficients have been investigated in detail in [5] 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. [8]. We assume that Ω 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 Γ 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 two-dimensional case) generalized hat functions.

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

$$u_h^{(k)} = u_h^{(k)}(x, v) := \sum_{i=1}^{N} u_i^{(k)} N_i(x, v) =: (\mathbf{u}^{(k)})^{\mathrm{T}} \mathbf{N}. \tag{32}$$

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

$$\varphi_h = \varphi_h(x, v) := \sum_{i=1}^{N} \varphi_i N_i(x, v) =: \boldsymbol{\varphi}^{\mathrm{T}} \mathbf{N}. \tag{33}$$

Inserting these representations into the left-hand side in Equation (31), we obtain

$$a(u^{(k)}, \varphi - u^{(k)}) = (\boldsymbol{\varphi} - \mathbf{u}^{(k)})^{\mathrm{T}} B \mathbf{u}^{(k)}, \tag{34}$$

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 Equation (31) is expressed in vector form as $(\boldsymbol{\varphi} - \mathbf{u}^{(k)})^{\mathrm{T}} \boldsymbol{f}$, where $\boldsymbol{f}$ 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 as follows: find* $\boldsymbol{u} \in \mathbb{R}^N$ *with* $\boldsymbol{u} \geq \mathbf{0}$ *such that*

$$(\boldsymbol{\varphi} - \boldsymbol{u})^{\mathrm{T}} (B\boldsymbol{u} - \boldsymbol{f}) \geq 0, \tag{35}$$

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 $\boldsymbol{u} \in \mathbb{R}^N$ *satisfying*

$$\boldsymbol{u}^{\mathrm{T}} (B\boldsymbol{u} - \boldsymbol{f}) = 0, \quad B\boldsymbol{u} - \boldsymbol{f} \geq 0, \quad \boldsymbol{u} \geq \mathbf{0}. \tag{36}$$

Here, an inequality like $\mathbf{u} \geq \mathbf{0}$ is to be understood componentwise. Again, the 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.

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

Recall from [9] for variational inequalities of the form (31), the error estimate

$$\|u^{(k)} - u_h^{(k)}\|_{H^1(\Omega)} = \mathcal{O}(h), \tag{37}$$

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

$$\|u - u_h\|_{L_2(\Omega_\mathcal{Y})} = \mathcal{O}(h^2 + \tau^2). \tag{38}$$

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 (36), 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 (36) for symmetric B, see, e.g. [23]. We 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 is a strictly upper diagonal matrix. The iterative scheme is then as follows: given an initial vector $\mathbf{w}^0 \in \mathbb{R}^N$, iterate for $j = 1, 2, \ldots$ 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 (36) (with B replaced by a symmetric positive definite G for the moment), one adds after the iteration a componentwise 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 Equation (21).

ALGORITHM 4.1 The projective Gauss–Seidel method *Set the initial vector* $\boldsymbol{w}^0 \in \mathbb{R}^N$.
Iterate for $j = 1, 2, \ldots$

$$\begin{aligned} \tilde{\mathbf{w}}^j &= \mathbf{w}^{j-1} + (U + D)^{-1}(\mathbf{g} - G\mathbf{w}^{j-1}) \\ w_i^j &= \max(0, \tilde{w}_i^j), \quad i = 1, \ldots, N. \end{aligned} \tag{39}$$

We shortly write

$$\mathbf{w}^j = \mathtt{pGS}\,(\mathbf{w}^{j-1}), \quad j = 1, 2, \ldots \tag{40}$$

for Algorithm 4.1. Generalizations of this method are known as PSOR (projected successive overrelaxation) schemes for which Equation (40) is a special case. Convergence of this method to the unique solution of Equation (36) is known for any initial vector $\mathbf{w}^0$ provided that G is symmetric positive definite and the convex set $\mathbf{K}$ is factorial, for example, $\mathbf{K} = \prod_{i=1}^N K_i$ [13]. While the latter is clearly satisfied, the matrix B defined in Equation (34) is not symmetric. However, as long as B is not 'too' asymmetric in a sense that can be made precise, and as long as $\frac{1}{2}(B + B^{\mathrm{T}})$ is positive definite, the PSOR method still converges [39]. We also observed this numerically for B defined in Equation (34), see Section 5.

As for the convergence speed of the projective Gauss–Seidel scheme, as in the unrestricted case, this depends 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

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

number is known to behave like $\kappa_2(B) = \mathcal{O}(h^{-2})$ as $h \to 0$, see, e.g. [23]. Consequently, the speed of any iterative method deteriorates as the grid size becomes smaller. Remedies to overcome this problem are the subject of the following 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

$$V_0 \subset V_1 \subset \cdots \subset V_\ell \subset \cdots V_L \subset H^1(\Omega) \tag{41}$$

comes into play, where the grid spacing h_ℓ 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 ℓ 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. [8,15,23,35].

A multigrid method adapted to the restricted problem in Equation (36) is described in the next section.

4.3 *An MMG method*

The method of choice for discretized variational inequalities is the monotone multigrid (MMG) method [34,37], which was set up for piecewise linear Ansatz functions. An extension to higher order B-splines can be found in [27] with an application to the highly accurate computation of the Greeks for American options with constant volatility in [28].

The MMG 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 (41) with the highest level L denoting the discretization level on which Equation (36) is to be solved,

$$\begin{aligned} B_L \mathbf{u}_L &\geq \boldsymbol{f}_L, \\ \mathbf{u}_L &\geq \mathbf{0}, \\ \mathbf{u}_L^{\mathrm{T}}(B_L \mathbf{u}_L - \boldsymbol{f}_L) &= 0. \end{aligned} \tag{42}$$

Let $\mathbf{u}_L^\nu \in V_L$ be the approximation in the νth 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 $\mathbf{u}_L^{\nu,1} := \mathtt{pGS}\,(\mathbf{u}_L^\nu)$ and the exact solution $\mathbf{u}_L$ can be approximated without essential loss of information on a coarser grid Δ_{L-1}. For the linear complementary problem (42), this is realized for two grids Δ_L and Δ_{L-1} as follows. Introducing the defect $\mathbf{d}_L := \boldsymbol{f}_L - B_L \mathbf{u}_L^{\nu,1}$, Equation (42) can be written as

$$\begin{aligned} B_L \mathbf{v}_L &\geq \mathbf{d}_L, \\ \mathbf{v}_L &\geq -\mathbf{u}_L^{\nu,1}, \\ (\mathbf{v}_L + \mathbf{u}_L^{\nu,1})^{\mathrm{T}}(B_L \mathbf{v}_L - \mathbf{d}_L) &= 0. \end{aligned} \tag{43}$$

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

On a coarser grid Δ_{L-1}, the defect problem can now be approximated by

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

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 $r, \tilde{r} : V_L \to V_{L-1}$. 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 would like to point out that it is important that the restriction $\tilde{r}$ be chosen such that the new iterate satisfies the constraint

$$\mathbf{u}_L^{\nu,2} := \mathbf{u}_L^{\nu,1} + p\mathbf{v}_{L-1} \geq \mathbf{g}_L := 0 \tag{44}$$

on the fine grid. Applying this idea recursively on several different grids, one obtains the MMG method for linear complementary problems (36) as follows.

ALGORITHM 4.2 $\mathbf{MMG}_\ell$ (*νth cycle on level $\ell \geq 1$*)
Let $\boldsymbol{u}_\ell^\nu \in V_\ell$ be a given approximation.

(1) *A priori smoothing and projection*:

$$\mathbf{u}_\ell^{\nu,1} := (\mathtt{pGS}\,(\mathbf{u}_\ell^\nu))^{\eta_1}.$$

(2) *Coarse grid correction*:

$$\mathbf{d}_{\ell-1} := r(\boldsymbol{f}_\ell - B_\ell \mathbf{u}_\ell^{\nu,1}),$$
$$\mathbf{g}_{\ell-1} := \tilde{r}(\mathbf{g}_\ell - \mathbf{u}_\ell^{\nu,1}),$$
$$B_{\ell-1} := rB_\ell p.$$

If $\ell = 1$, exactly solve the linear complementary problem (42) *and set $\mathbf{v}_{\ell-1} := \mathbf{v}$.*
If $\ell > 1$, perform ζ steps of $\mathbf{MMG}_{\ell-1}$ 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*:

$$\mathbf{u}_\ell^{\nu,3} := (\mathtt{pGS}\,(\mathbf{u}_\ell^{\nu,2}))^{\eta_2}.$$

Set $\mathbf{u}_\ell^{\nu+1} := \mathbf{u}_\ell^{\nu,3}$.

The numbers of *a priori* and *a posteriori* smoothing steps are denoted by η_1 and η_2, respectively. For $\zeta = 1$, one obtains a V cycle and for $\zeta = 2$ a W cycle.

Condition (44) 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 [34]. A corresponding construction for B-splines of general order k can be found in [27].

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

5. Numerical results

Finally, we would like 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 and (3) convergence behaviour of the iterative methods in view of the asymmetry of the system matrix B.

5.1 *Validation of the model*

First, we would like to investigate whether our computed solution can be validated. Since there exists no exact solution of Problem 2.3 in closed form, we compare our solution with the situation from the study of Clarke and Parrot [11]. 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 used in [16], see the remarks before (7), for which we are, however, not aware of values for comparison.)

We consider an American put option for a stock price of today as $S(0) = 10$, that is, 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 found 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
256×256	0.796	[40]
177×103	0.7961	[48]
320×128	0.7581	[30]

In [30], on this grid, computation has been done for 64 time steps on the interval $[0, T] = [0, 0.25]$. The other studies do not explicitly mention the amount of time steps. Since it is not known which of these solutions is indeed the 'best', we took as reference the value $u_{\text{ref}} = u(0, 0.25) = 0.7581$ from [30].

The discretized solution of Problem 3.1 has been computed on the domain $\Omega = (-5, 5) \times (0.0025, 0.4975)$, that is, we chose $x_{\max} = 5$, $v_{\min} = 0.0025$ and $v_{\max} = 0.4975$. We used the Crank–Nicolson scheme for $\theta = \frac{1}{2}$. Since the convection term in Equation (30) often causes stability problems, we used a *graded grid* like in [5,30]. Here, in the direction of the transformed stock price x, we used a grid which is four times as large as the grid for the variance v. The solution has been computed with an MMG method on level $L = 4$.

We plot the discretized solution of Problem 3.1 in Figures 1 and 2. In Figure 1, we can see 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 can recognize the characteristic bend of this termination condition. In the right graphic, we can 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 clearly visible for the value of the option depending on the non-transformed stock prize in Figure 2. For better visualization, here the axis for the stock price S is

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

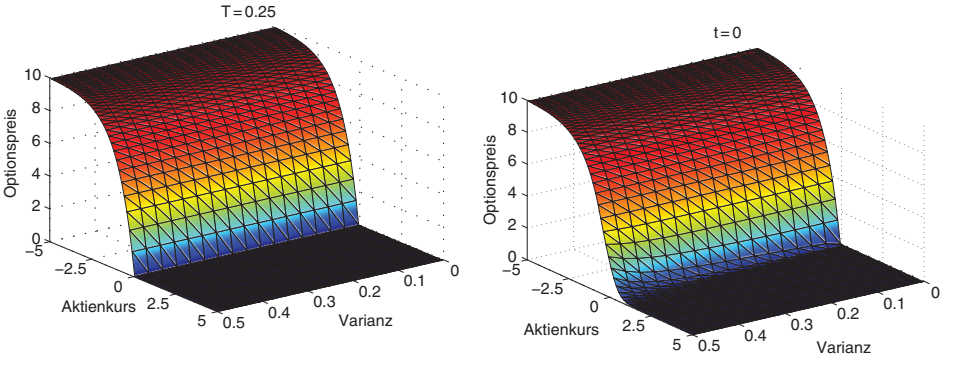


Figure 1. Left: transformed payoff function $g(x)$ according to Equation (15) of the American put option at time $T = 0.25$. Right: option price of today at $t = 0$.

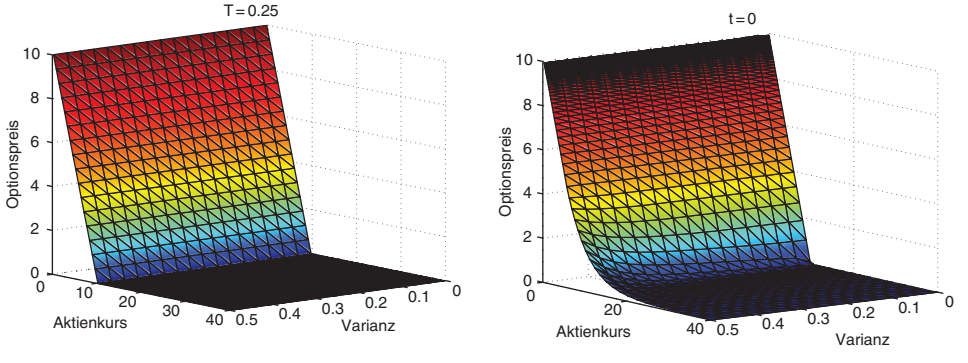


Figure 2. Left: payoff function $\mathcal{H}$ of the American put option at time $T = 0.25$. Right: option price for today $t = 0$ without transformation.

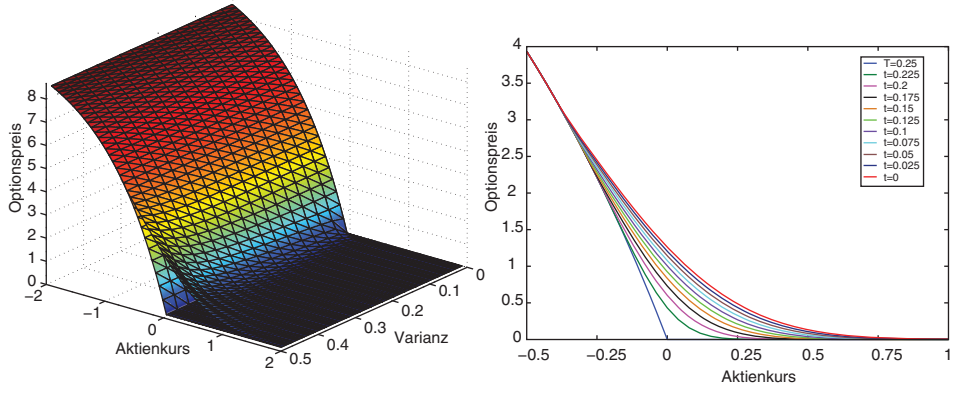


Figure 3. Left: visualization of the computed solution at $t = 0$ (upper plot) and transformed payoff function g (lower plot). Right: evolution of the option price for different times t for constant variance $v = 0.4975$.

slightly shortened. The left graphic shows the payoff function $\mathcal{H}$ at end time $T = 0.25$, and the right graphic shows the option price at current time $t = 0$. We can clearly see the diffusion effect.

Next, we would like to investigate the free boundary introduced in Equation (9). In Figure 3, in the left graphic, we can 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

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

$0 \le t^{(k)} < T$,

$$u(t^{(k)}, x, v) = g(x) \quad \text{for } x \le x_f(t^{(k)}, v),\ v \ge 0. \tag{45}$$

As we can see in the figure, the option and the obstacle g coincide for $x < x_f$. For growing x, we can 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 $x \le x_f(t^{(k)})$, $v \ge 0$, and every $0 \le t^{(k)} < T$. We can observe this projection even more clearly in the right graphic in Figure 3. Here, we can 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 $x = 0$; therefore, here the characteristic bend appears at $x = 0$. We can clearly see how this discontinuity is smoothed as time goes backward from $T = 0.25$ to $t = 0$. We can 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 employed 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, that is,

$$\|C\mathbf{u} - \boldsymbol{f}\|_2, \tag{46}$$

which corresponds to an *a priori* estimation:

$$\|u^{(k)} - u_h^{(k)}\|_{L_2(\Omega)} = \mathcal{O}(h^2). \tag{47}$$

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 $\texttt{tol} = 10^{-3}2^{-\ell}$ provided the most accurate and most efficient results. For level ℓ, we have $N_\ell = (2^\ell + 1)(2^{\ell+2} + 1)$ grid points in our triangulation using the graded mesh. The coarsest grid for $\ell = 2$ had 85 and the finest grid for $\ell = 7$ had 66,177 grid points.

In Table 1, we present the results for computing the American put option, using just the simple projective Gauss–Seidel method from Algorithm 4.1. Tables 1 and 2 are set up as follows. The first column shows the discretization level ℓ 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 show the absolute and relative errors between our computed solution and the reference value,

$$e_{\ell,\text{abs}} := |u_\ell^{(0)} - u_{\text{ref}}|, \quad e_{\ell,\text{rel}} := \frac{|u_\ell^{(0)} - u_{\text{ref}}|}{u_{\text{ref}}}. \tag{48}$$

Table 1. The results obtained using the projected Gauss–Seidel method from Algorithm 4.1 for $\text{tol} = 10^{-3}2^{-\ell}$.

ℓ	$u_\ell^{(0)}$	$e_{\ell,\text{abs}}$	$e_{\ell,\text{rel}}$	∅(# PSOR)	ECO_ℓ	Time (s)
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

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

Table 2. The results obtained using the MMG method with two pre-smoothing and three post-smoothing steps.

ℓ	$u_\ell^{(0)}$	$e_{\ell,\mathrm{abs}}$	$e_{\ell,\mathrm{rel}}$	ECO_ℓ	Time (s)
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

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

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

$$\varnothing(\#\ \mathtt{PSOR}) := \frac{\#\ \text{all iterations in every time step}}{10}. \tag{49}$$

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*:

$$\mathrm{ECO}_\ell := \frac{\ln|e_{\ell,\mathrm{abs}} - e_{\ell-1,\mathrm{abs}}|}{\ln(2)}, \tag{50}$$

which describes the expected order. The last column shows the total time in seconds that 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 also tested two simple multiscale methods, which can be found in [8]. One method called *Nested Iteration* is based on iterating on level ℓ only up to discretization error accuracy given by Equation (37) and starting the iteration on level $\ell + 1$ with the solution vector from level ℓ, prolongated to the 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 in the performance of the projective Gauss–Seidel scheme for both these methods is negligible. So, we conclude that for the present situation, these methods do not have an effect.

In Table 2, we present the results obtained with the MMG method from Algorithm 4.2. We used $\eta_1 = 2$ pre-smoothing steps and $\eta_2 = 3$ post-smoothing steps and a V cycle. On the lowest level $\ell = 2$, we solved the linear complementary problem with the projected Gauss–Seidel method. We would like to point out that the same accuracy can be achieved with the MMG method in less than $\frac{1}{30}$ of the time needed for the projected Gauss–Seidel method.

5.3 *Convergence behaviour of the iterative schemes*

Finally, we present two graphics showing the convergence behaviour of the schemes considered. Figure 4 shows on the left the development of the absolute error $e_{\ell,\mathrm{abs}}$ on the vertical axis over the time on the horizontal axis in a log–log plot. We can observe that the projected Gauss–Seidel method as well as the simple multiscale variants show a very similar behaviour, while the MMG scheme is substantially faster. In fact, one can observe a multigrid efficiency which is optimal or close to it. We believe this to be 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 side of this figure 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). Here, we can see that the MMG method also outperforms all the other methods.

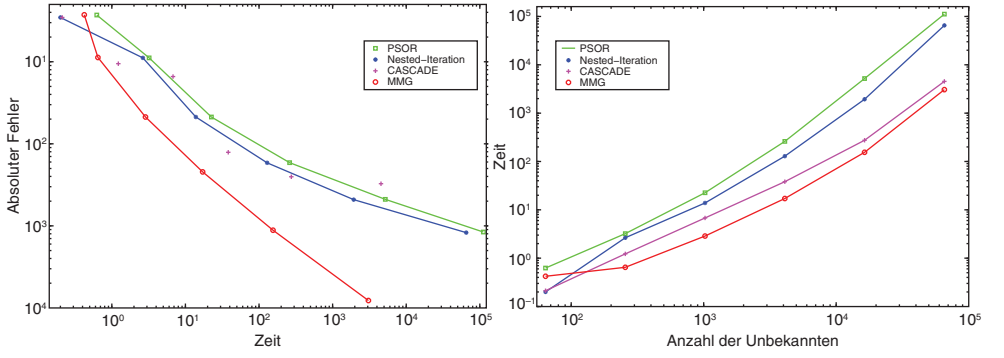


Figure 4. Left: absolute error (vertical) over time (horizontal) as a log–log plot. Right: time to solve a linear system over number of unknowns as a log–log plot.

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

The next step for the present model would be to develop MMG methods of higher order to compute the Greeks with high precision as in [27,28]. More modern models other than the classical Black–Scholes equations are jump-diffusion price processes. In these processes, the Wiener process in the Black–Scholes model is replaced by a Lévy process. According to Eberlein [17], such processes allow for a more realistic modelling of the dynamics of asset prices, see [21] 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 valuated by means of Galerkin methods in space and continuous piecewise linear spline wavelets in [38]. 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 the convergent adaptive wavelet method of optimal complexity along the lines of [45] would be our method of choice. However, it is not clear how to develop a convergent adaptive scheme for the treatment of inequalities, see [44] for the first steps in this direction for elliptic variational inequalities.

Acknowledgements

We thank Christoph Schwab, who originally pointed out the problem of stochastic volatility after completion of the papers [27,28] dealing with MMG methods with higher order B-splines. We also thank Ernst Eberlein, Andrej Palczweski and two anonymous referees on their remarks on modelling issues concerning boundary and Feller's condition, see Remark 2.1. We express our gratitude to two anonymous referees for their useful remarks on the whole manuscript. This work was 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).

References

[1] Y. Achdou and O. Pironneau, *Computational Methods for Option Pricing*, SIAM Series: Frontiers in Applied Maths., SIAM, Philadelphia, 2005.

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

[2] F. AitSahlia, M. Goswami, and S. Guha, *American option pricing under stochastic volatility: An efficient numerical approach*, Comput. Manag. Sci. 7(2) (2010), pp. 171–187.
[3] F. AitSahlia, M. Goswami, and S. Guha, *American option pricing under stochastic volatility: An empirical evaluation*, Comput. Manag. Sci. 7(2) (2010), pp. 189–206.
[4] L.B.G. Andersen and V.V. Piterbarg, *Moment explosions in stochastic volatility models*, Finance Stoch. 11(1) (2007), pp. 29–50.
[5] T. Apel, G. Winkler, and U. Wystup, *Valuation of Options in Heston's Stochastic Volatility Model Using Finite Element Methods*, Foreign Exchange Risk, Risk Publications, London, 2001.
[6] C.A. Ball and A. Roma, *Stochastic volatility option pricing*, J. Financ. Quant. Anal. 29 (1994), pp. 589–607.
[7] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, J. Polit. Economy 81 (1973), pp. 637–654.
[8] D. Braess, *Finte Elements*, 3rd ed., Cambridge University Press, Cambridge, 2001.
[9] F. Brezzi, W. Hager, and P.A. Raviart, *Error estimates for the finite element solution of variational inqualities*, Numer. Math. 28 (1977), pp. 431–443.
[10] W.-T. Chen, *An extensive exploration of three key quantitative approaches for pricing various financial derivatives*, PhD thesis, University of Wollongong, Australia, 2011.
[11] N. Clarke and K. Parrot, *Multigrid for American option pricing with stochastic volatility*, Appl. Math. Finance 6 (1999), pp. 177–195.
[12] J.C. Cox, J.E. Ingersoll, and S.A. Ross, *A theory of the term structure of interest rate*, Econometrica 53 (1985), pp. 385–407.
[13] C.W. Cryer, *The solution of a quadratic programming problem using systematic overrelaxation*, Siam J. Contr. 9 (1971), pp. 385–392.
[14] J. Cvitanić, B. Rozovskii, and I. Zaliapin, *Numerical estimation of volatility values from discretely observed diffusion data*, J. Comput. Finance 9(4) (2006), pp. 1–36.
[15] W. Dahmen, *Wavelet and multiscale methods for operator equations*, Acta Numer. 6 (1997), pp. 55–228.
[16] D. Duffie, J. Pan, and K. Singleton, *Transform analysis and asset pricing for affine jump-diffusions*, Econometrica 68(6) (2000), pp. 1343–1376.
[17] E. Eberlein, *Application of generalized hyperbolic Lévy motions to finance*, in *Lévy Processes: Theory and Applications*, O.E. Barndorff-Nielsen, T. Mikosch, and S. Resnick, eds., Birkhäuser, Boston, MA, 2001, pp. 319–337.
[18] C.M. Elliott and J.R. Ockendon, *Weak and Variational Methods for Moving Boundary Value Problems*, Pitman, Boston, MA, 1982.
[19] J.-P. Fouque, G. Papanicolaou, and K. R. Sircar, *Derivatives in Financial Markets with Stochastic Volatility*, Cambridge University Press, Cambridge, 2000.
[20] J. Gatheral, *The Volatility Surface: A Practitioner's Guide*, John Wiley & Sons, Hoboken, NJ, 2006.
[21] K. Glau, *Feynman-Kac-Darstellungen zur Optionspreisbewertung in Lévy-Modellen* (in German), Ph.D. diss., Department of Mathematical Stochastics, University of Freiburg, 2010. Available at http://www.freidok.uni-freiburg.de/volltexte/7724/.
[22] R. Glowinski, J.-L. Lions, and R. Trémolières, *Numerical Analysis of Variational Inequalities*, North-Holland, 1981.
[23] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, Springer, New York, 1994.
[24] F.B. Hanson and G. Yan, *American put option pricing for stochastic-volatility, jump-diffusion models*, Proceedings of 2007 American Control Conference, 11 July 2007, pp. 384–389.
[25] S.L. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, Rev. Finan. Stud. 6(2) (1993), pp. 327–343.
[26] N. Hilber, S. Kehtari, C. Schwab, and C. Winter, *Wavelet finite element method for option pricing in highdimensional diffusion market models*, Research Rep. No. 2010-01, Seminar für Angewandte Mathematik, ETH Zürich, 2010.
[27] M. Holtz and A. Kunoth, *B-spline based monotone multigrid methods*, Siam J. Numer. Anal. 45(1) (2007), pp. 1175–1199.
[28] M. Holtz and 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, P. Wesseling, C.W. Oosterlee, and P. Hemker, eds., 2005.
[29] J.C. Hull, *Options, Futures and Other Derivates*, 8th ed., Pearson, Upper Saddle River, NJ, 2011.
[30] S. Ikonen and J. Toivanen, *Operator splitting methods for American options with stochastic volatility*, Numer. Math. 113(2) (2009), pp. 299–324.
[31] S. Ikonen and J. Toivanen, *Efficient Numerical Methods for Pricing American Options Under Stochastic Volatility*, Numer. Methods Partial Differential Equations 24(1) (2008), pp. 104–126.
[32] K. Ito and J. Toivanen, *Lagrange multiplier approach with optimized finite difference stencils for pricing American options under stochastic volatility*, Siam J. Sci. Comput. 31(4) (2009), pp. 2646–2664.
[33] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and their Applications*, Academic Press, New York, 1980.
[34] R. Kornhuber, *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*, Teubner, Stuttgart, 1997.
[35] A. Kunoth, *Optimized wavelet preconditioning*, in *Multiscale, Nonlinear and Adaptive Approximation*, R.A. DeVore and A. Kunoth, eds., Springer, Berlin, Heidelberg, 2009, pp. 325–378.
[36] F.A. Longstaff and E.S. Schwartz, *Valuing American options by simulation: A simple least-squares approach*, Rev. Finan. 14 (2001), pp. 113–147.
[37] J. Mandel, *A multilevel iterative method for symmetric, positive definite linear complementarity problems*, Appl. Math. Optim. 11 (1984), pp. 77–95.

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012

[38] A.-M. Matache, P.-A. Nitsche, and C. Schwab, *Wavelet Galerkin pricing of American options on Lévy driven assets*, Quant. Finance 5(4) (2005), pp. 403–424.
[39] W. Niethammer, *Relaxation bei unsymmetrischen Matrizen*, Math. Z. 85 (1964), pp. 319–327.
[40] C.W. Oosterlee, *On multigrid for linear complementary problems with application to American-style options*, Electron. Trans. Numer. Anal. 15 (2003), pp. 165–185.
[41] J. Persson and L. von Sydow, *Pricing American options using a space-time adaptive finite difference method*, Math. Comput. Simulation 80(9) (2010), pp. 1922–1935.
[42] C. Reisinger and G. Wittum, *Efficient hierarchical approximation of high-dimensional option pricing problems*, Siam J. Sci. Comput. 29 (2007), pp. 440–458.
[43] M. Renardy and R.C. Rogers, *An Introduction to Partial Differential Equations*, 2nd ed., Springer, New York, 2004.
[44] M. Rometsch and K. Urban, *Adaptive wavelet methods for elliptic variational inequalitites I: Analysis*, preprint (2010).
[45] Ch. Schwab and R. Stevenson, Space–time adaptive wavelet methods for parabolic evolution equations, Math. Comput. 78 (2009), pp. 1293–1318.
[46] R. Seydel, *Tools for Computational Finance*, 4th ed., Springer, Berlin, Heidelberg, 2009.
[47] S.-P. Zhu and W.-T Chen, *A predictor–corrector scheme based on ADI method for pricing American puts with stochastic volatility*, Comput. Math. Appl. 62 (2011), pp. 1–26.
[48] R. Zvan, P.A. Forsyth, and K.R. Vetzal, *Penalty methods for American options with stochastic volatility*, J. Comput. Appl. Math. 91 (1998), pp. 199–218.

Downloaded by [Universitatsbibliothek Paderborn] at 06:53 27 July 2012