Computing and Visualization in Science manuscript No.
(will be inserted by the editor)

On Multigrid for Anisotropic Equations and Variational Inequalities

Pricing Multi-Dimensional European and American Options

Christoph Reisinger and Gabriel Wittum

Interdisciplinary Center for Scientific Computing, University of Heidelberg, INF 368, 69120 Heidelberg, Germany
(email: {christoph.reisinger, wittum}@iwr.uni-heidelberg.de)

Received: date / Revised version: date

Abstract. Partial differential operators in finance often originate in bounded linear stochastic processes. As a consequence, diffusion over these boundaries is zero and the corresponding coefficients vanish. The choice of parameters and stretched grids lead to additional anisotropies in the discrete equations or inequalities.

In this study various block smoothers are tested in numerical experiments for equations of Black-Scholes-type (European options) in several dimensions. For linear complementarity problems, as they arise from optimal stopping time problems (American options), the choice of grid transfer is also crucial to preserve complementarity conditions on all grid levels. We adapt the transfer operators at the free boundary in a suitable way and compare with other strategies including cascadic approaches and full approximation schemes.

Key words Multigrid, plane smoothers, linear complementarity problems, option pricing

1 Introduction

1.1 The Black-Scholes model and related equations

1.1.1 Options

In the simplest setting an option gives the holder the right — but not the obligation — to buy (call) or sell (put) an underlying risky asset S for a fixed price (the strike price) K at some fixed time (expiration) T in the future (European contract) or any time up to the expiration date (American contract). The classical example for an underlying asset is a single stock, where interest rate options and exchange rate options are further standard contracts.

We consider as underlying asset a contingent claim g dependent on a multivariate process $\boldsymbol{S}$ (e. g. d stocks). At expiration the price u of a, say, put option is clearly

$$u(\boldsymbol{S},T) = u_0(\boldsymbol{S}) := \max\left(K - g(\boldsymbol{S}), 0\right). \tag{1}$$

1.1.2 European options

If $\boldsymbol{S}$ is modelled as geometric Brownian motion $\mathrm{d}S_i = \mu_i S_i \mathrm{d}t + \sigma_i \mathrm{d}W_i$ with drifts μ_i, constant volatilities σ_i and Wiener processes W_i like in the original studies by Black, Merton and Scholes [1], the option price follows the PDE

$$\frac{\partial u}{\partial t} + \mathcal{L}_{BS} u = 0, \tag{2}$$

$$\mathcal{L}_{BS} u := \frac{1}{2}\sum_{i,j=1}^{n} \sigma_i \sigma_j \rho_{ij} S_i S_j \frac{\partial^2 u}{\partial S_i \partial S_j} + r \sum_{i=1}^{n} S_i \frac{\partial u}{\partial S_i} - ru,$$

which can be solved backwards in time from the terminal condition (1). Here ρ_{ij} are the asset corelations and r the risk-free interest rate, which are both assumed to be constant. Despite well-known deficiencies the Black-Scholes model is still most widely used among practitioners. Moreover, it can serve as a basis for homogenisation approaches [7].

As a multi-dimensional example we will consider in the following a basket of stocks, $g(\boldsymbol{S}) = \sum \mu_i S_i$, with positive weights μ_i. After time reversion the Black-Scholes-PDE forms the initial value problem

$$\frac{\partial u}{\partial t} - \mathcal{L}_{BS} u = 0 \qquad \forall (\boldsymbol{S}, t) \in \mathbf{R}^n_+ \times (0,T), \tag{3}$$

$$u(\mathbf{S}, 0) = u_0(\boldsymbol{S}) = \Big(K - \sum_{i=1}^{d} \mu_i S_i\Big)_+ \forall \mathbf{S} \in \mathbf{R}^d_+.$$

For $S_i = 0$ natural boundary conditions hold (note that all coefficients vanish as $S_i \to 0$), for $|\boldsymbol{S}| \to \infty$ asymptotical values can be set as Dirichlet conditions.

Note also that the equation is invariant under rescaling of the asset variables, so without loss of generality the domain can be taken as $[0,1]^d$. The only reference value is then the strike price K.

1.1.3 American options

In the case of American contracts the option price is subject to an inequality constraint and can be shown [17]

to be the solution of the linear complementarity problem

$$\frac{\partial u}{\partial t} - \mathcal{L}_{BS} u \le 0, \tag{4}$$

$$u \ge g, \tag{5}$$

$$\left(\frac{\partial u}{\partial t} - \mathcal{L}_{BS} u\right) \cdot (u - g) = 0 \tag{6}$$

with $g = u_0$ and initial and boundary values as in the European case.

1.2 Transformations

In the following we study various formulations of (3) in different coordinates. Due to the special structure of the Black-Scholes model, it can be transformed into a more convenient form. The widely used 'log price' $y_i := \ln S_i$ $(1 \le i \le d)$ leads to a problem with constant coefficients

$$\frac{\partial u}{\partial t} = \frac{1}{2} \sum_{i,j=1}^{d} \sigma_i \sigma_j \rho_{ij} \frac{\partial^2 u}{\partial y_i \partial y_j} + \sum_{i=1}^{d} (r - \frac{1}{2}\sigma_i^2) \frac{\partial u}{\partial y_i} - ru$$
$$\forall (\boldsymbol{y}, t) \in \mathbf{R}^n \times (0, T), \tag{7}$$
$$u(\mathbf{y}, 0) = u_0(\mathrm{e}^{\boldsymbol{y}}) = \Big(K - \sum_{i=1}^{d} \mu_i \mathrm{e}_i^{y}\Big)_+ \forall \mathbf{y} \in \mathbf{R}^d.$$

Now transformation of the diffusion operator to diagonal form, $\Lambda = \boldsymbol{Q} \Sigma \boldsymbol{Q}^T$, further simplifies the system via rotation and translation $\boldsymbol{z} := \boldsymbol{Q}\boldsymbol{y} - t\boldsymbol{b}$ ($b_i := \sum_{j=1}^{d} q_{ij}(r - \frac{1}{2}\sigma_j^2)$):

$$\frac{\partial u}{\partial t} = \frac{1}{2} \sum_{i=1}^{d} \lambda_i \frac{\partial^2 u}{\partial {z_i}^2} - ru \quad \forall (\boldsymbol{z}, t) \in \mathbf{R}^d \times (0, T), \tag{8}$$

$$u(\mathbf{z}, 0) = u_0(\mathrm{e}^{\boldsymbol{Q}^T \boldsymbol{z}}) \quad \forall \mathbf{z} \in \mathbf{R}^d.$$

This formulation is an ideal starting point for principal component analysis and dimension reduction techniques that are inevitable for high-dimensional problems [14].

In the case of basket options the following transformation, where the basket itself serves as a variable ($x_1 = \sum_i \mu_i S_i$) and the other coordinates $(x_2, \ldots, x_d)$ parametrise its level sets (see also Fig. 1), proved very useful:

$$S_1 = x_1 x_2, \tag{9}$$

$$S_i = x_1 x_{i+1} \prod_{k=2}^{i} (1 - x_k), \quad 2 < i < d, \tag{10}$$

$$S_d = x_1 \prod_{k=2}^{d} (1 - x_k). \tag{11}$$

This leads to a parabolic equation of the type

$$\frac{\partial u}{\partial t} + \sum_{i,j=1}^{d} a_{ij}(\boldsymbol{x}) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(\boldsymbol{x}) \frac{\partial u}{\partial x_i} + ru = 0 \tag{12}$$

with polynomial coefficients. For details see [14]. Depending on the dimensionality and the required accuracy all these formulations can be useful.

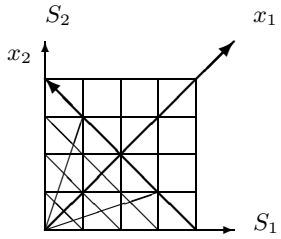

Fig. 1. $d = 2$: $x_1 = s_1 + S_2$, $x_2 = S_1/(S_1 + S_2)$

For other contracts or more general models for the asset price movement, e. g. with stochastic volatilities or interest rates, equations of similar structure arise, particularly diffusion (and often convection) over boundary faces vanish. This is a necessary (and sufficient) condition for the underlying stochastic processes to be bounded in the respective dimensions and guarantees uniqueness of the solution of the parabolic problem without boundary conditions [18]. In a typical setting

$$a_{ij}(\boldsymbol{x}) = 0 \ \text{ for } \ x_i, x_j \in \{0, 1\} \quad \forall i, j \in I, \tag{13}$$

$$b_i(\boldsymbol{x}) = 0 \ \text{ for } \ x_i \in \{0, 1\} \text{ or at least} \tag{14}$$

$$b_i(\boldsymbol{x}) n_i(\boldsymbol{x}) \ge 0 \ \text{ for } \ x_i \in \{0, 1\} \quad \forall i \in I, \tag{15}$$

where $n_i(\boldsymbol{x})$ is the outward normal and $I \subset \{1, \ldots, d\}$. An efficient robust solver for this class of equations and variational inequalities is desired.

1.3 Challenges and previous work

The objective of this paper are robust solvers of optimal complexity for anisotropic discrete problems resulting from the discretisation of strongly anisotropic continuous equations or variational inequalities of type (3) to (12), eventually on massively anisotropic grids as they are required e. g. for sparse grid extrapolation [14].

1.3.1 Anisotropic equations

Evidently, the anisotropies come here from three factors:

1. stretched grids (global),
2. the parameters (global), i. e. different volatilities for different components as in equation (2) or the decay of the spectrum as in (8),
3. vanishing coefficients along the boundary (local), as observed in equations (3) and (12) combined with (13), (14).

It is clear that 1. and 2. are roughly equivalent on a discrete level. This type of problem has been thoroughly studied in the literature, see e. g. [11] for an extensive survey of numerical results for different types of point, line and plane smoothers for the anisotropic three-dimensional Poisson equation or [12] for the application of plane smoothers for anisotropic Navier-Stokes equations.

In this paper we will focus on the last aspect.

1.3.2 Linear complementarity problems

If the solution is subject to complementarity conditions, this requires additional consideration. The crucial point is the construction of suitable transfer operators that preserve complementarity in the coarsening process.

Brandt and Cryer [3] study variants of Brandts full approximation scheme, where projected SOR-sweeps [5] are used as smoother. They choose essentially injection, which gives satisfactory results for porous flow through a dam, but due to the reduced order leads to strongly deteriorated convergence rates in problems with varying coefficients.

Hoppe, Kornhuber [8], and Wittum [10] modify the transfer operators at the free boundary such that prolongation and restriction are performed on active and inactive sets seperately. In other words, the coarse grid correction is switched off for the free boundary while still preserving the order of the transfer operators.

Braess et al. [2] avoid this problem altogether in a cascadic technique, where obviously no restriction is needed. However, in this 'one way' approach low-frequency errors can not be reduced efficiently once a finer level has been reached, which is a problem if the solution needs to be known very exactly (beyond discretisation error).

1.3.3 Applications in option pricing

Kŏcvara et al. [15] propose for American option pricing equations a combined (one level) iteration with a sequence of projected SSOR steps followed by CG steps restricted to the inactive set. This gives considerably improved convergence compared to a simple projected iteration and works well for sufficiently well-conditioned systems.

Clarke and Parrott [4] successfully apply a multigrid iteration with a hybrid-smoother that combines global projected point Gauss-Seidel with line smoothing on the inactive set for a one-factor model with stochastic volatility.

Oosterlee [13] studies the same problem and improves the results of Brandt and Cryer for the setting of varying coefficients by iterant recombination.

1.4 Outline

In section 2 we will study numerically the robustness of various block smoothers for problems of different dimensionality ($d = 2, 3, 4$). In d dimensions naturally blocks of dimensionality 0 to d can be chosen (2.2.1 to 2.2.3). For the (approximate) inversion of the blocks CG-type solvers and recursively applied multigrid methods are compared.

Section 3 confers this approach to obstacle problems. Additional attention is devoted to the grid transfer (3.3), where we introduce one-sided transfer operators at the free boundary. Full multigrid (3.4) provides a good initial approximation for the active set, where smoothing is also performed plane-wise (3.5). Finally we compare with a cascadic strategy (3.6).

2 Multigrid for anisotropic problems

In this section we are concerned with the efficiency of iterative solution methods for linear equations

$$\boldsymbol{A}_h \boldsymbol{u}_h = \boldsymbol{b}_h$$

arising from space — e. g. finite difference — and (implicit) time discretisations — e. g. the θ-scheme — of equations of the type introduced in the previous section. We will denote this problem by LP($\boldsymbol{A}_h$,$\boldsymbol{b}_h$). Particularly, we study multigrid methods with levels $l = 1, \ldots, L$, transfer operators I_l^{l-1} (restriction) and I_{l-1}^l (prolongation). The coarse grid matrices are obtained by the Galerkin ansatz $\boldsymbol{A}^{l-1} = I_l^{l-1} \boldsymbol{A}^l I_{l-1}^l$.

For future reference and comparison with the complementarity problems we list here the standard V-cycle for linear problems.

Algorithm 1 ***V-Cycle***

procedure *V(l, $\boldsymbol{A}^l$, $\boldsymbol{b}^l$, $\boldsymbol{u}^l$)*
begin
  if $l = 1$ ***then***
    $\boldsymbol{u}^1 \leftarrow$ *LP($\boldsymbol{A}^1$, $\boldsymbol{b}^1$);*
  else begin
    $\boldsymbol{u}^l \leftarrow \mathcal{S}(\boldsymbol{A}^l, \boldsymbol{b}^l, \boldsymbol{u}^l)$; *(pre-smoothing)*
    $\boldsymbol{r}^l = \boldsymbol{b}^l - \boldsymbol{A}\boldsymbol{u}^l$;
    $\boldsymbol{r}^{l-1} = I_l^{l-1} \boldsymbol{r}^l$;
    $\boldsymbol{d}^{l-1} \leftarrow$ *V($l-1$, $\boldsymbol{A}^{l-1}$, $\boldsymbol{r}^{l-1}$, $\boldsymbol{d}^{l-1}$);*
    $\boldsymbol{d}^l = I_{l-1}^l \boldsymbol{d}^{l-1}$;
    $\boldsymbol{u}^l = \boldsymbol{u}^l + \boldsymbol{d}^l$;
    $\boldsymbol{u}^l \leftarrow \mathcal{S}(\boldsymbol{A}^l, \boldsymbol{b}^l, \boldsymbol{u}^l)$; *(post-smoothing)*
  end
return *V-Cycle*

There are several approaches to deal with inhomogeneities and anisotropies in a (here: geometric) multilevel context. We can point out three main classes, assigned to the different components of the multigrid method: block-smoothing (blocks with weak coupling are inverted collectively), semi-coarsening (coarsening of the grid in the direction of the anisotropy only) and matrix-dependent grid transfer. The latter is not required here as the discrete relative jumps of the coefficients on the grid that appear at the boundary are of moderate size.

In the following subsections we will investigate various block smoothers in greater detail.

2.1 Hyperplane smoothers

The degeneracy of the partial differential equation at the boundary, as described in 1.2 (see (13), (14)), strongly suggests the use of block-smoothers where planes orthogonal/parallel to the boundaries are inverted simultaneously. As can be expected, accurate solutions of these equations are not necessary and normally a reduction of the residual by a factor about 0.1 will suffice to provide smoothing properties. This can be achieved by some CG-type steps or a few (typically one) MG-cycles.

More precisely, if the index set of the scalar unknowns on a tensor product grid is

$$\mathcal{I} = \{\boldsymbol{1} \le \boldsymbol{i} \le \boldsymbol{n}\},$$

where $\boldsymbol{i}$ is a multiindex and $\boldsymbol{n} = (n_1, \dots, n_d)$ is the length of the grid in the different directions (yielding a total of $N = n_1 \cdots n_d$ unknowns) we define $(d-m)$-dimensional blocks

$$\mathcal{I}^{j_1,\dots,j_m}_{k_1,\dots,k_m} = \{\boldsymbol{1} \le \boldsymbol{i} \le \boldsymbol{n} : i_{k_1} = j_1, \dots, i_{k_m} = j_m\},$$

i. e. the coordinates in the m directions $k_1, \dots, k_m$ are kept fixed at $j_1, \dots, j_m$. The number of blocks is then $M = (n_1 \cdot \ldots \cdot n_m)/(n_{k_1} \cdot \ldots \cdot n_{k_m})$. For simplicity we denote for fixed $k_1, \dots, k_m$, $j_1, \dots, j_m$ the M blocks by $\mathcal{I}_1, \dots, \mathcal{I}_M$, so that $\mathcal{I} = \mathcal{I}_1 \dot\cup \dots \dot\cup \mathcal{I}_M$. This partitioning corresponds to blocks of vectors $\boldsymbol{u}_i$, $\boldsymbol{b}_i$ and matrices $\boldsymbol{A}_{ij}$, $1 \le i, j \le M$, and induces a block Gauß-Seidel algorithm of the form

$$\boldsymbol{u}_i^{k+1} = \boldsymbol{A}_{ii}^{-1}\Big(\boldsymbol{b}_i - \sum_{j<i} \boldsymbol{A}_{ij}\boldsymbol{u}_j^{k+1} - \sum_{j>i} \boldsymbol{A}_{ij}\boldsymbol{u}_j^{k}\Big). \tag{16}$$

For two and three dimensions the case $m = 1$ is known as line smoothing, the case $m = 2$ as plane smoothing. Generally speaking, we will study the extreme cases $m = d$ (direct solver), $m = d-1$ (hyper-plane smoother), $m = 1$ (line smoother) and $m = 0$ (point smoother).

If not indicated differently, all non-trivial block-systems in the following numerical tests are solved by Bi-CGSTAB [16] as are the coarse grid problems.

2.2 Numerical tests

For the numerical tests we study model problem (3) with a few uncorrelated assets ($d = 2, 3, 4$), $\sigma_i = 0.4$, $r = 0.05$, $T = 1$, $K = 0.25$. For a discretisation with central differences we perform one implicit Euler step covering the time interval.

Starting from an initial guess $\boldsymbol{u}^{(0)}$, the multigrid iteration generates a sequence $\boldsymbol{u}^{(k)}$, $(k \ge 0)$. For those iterates the average convergence rate is

$$\bar\rho_k = \left(\frac{\|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{u}^{(k)}\|}{\|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{u}^{(0)}\|}\right)^{\frac{1}{k}}. \tag{17}$$

Of most practical interest, however, is the number of work units $W(\epsilon)$ that is necessary for an error reduction by a factor of ϵ. As standard work unit we take one V(1,1)-cycle with a point Gauß-Seidel smoother on the respective grid. If the number of operations shall be included in the rate, we look at a weighted average rate of the type

$$\hat\rho = \bar\rho^{\frac{1}{W}}, \tag{18}$$

where W is the CPU-time of a cycle in standard units.

Note that for the 'direct' Bi-CG solver only selected iterations are plottet.

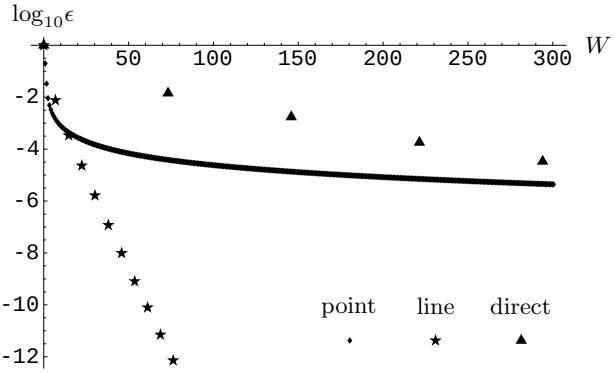

Fig. 2. 2D Black-Scholes, 263169 points: Error reduction ϵ vs. work units W for point Gauß-Seidel smoother, alternating line Gauß-Seidel smoother and direct (Bi-CGSTAB) solver

Table 1. 2D Black-Scholes: Average convergence rate $\bar\rho$, solution time T plus the relative work units W and relative convergence rate $\hat\rho$ for grid levels $l = 7, \dots, 10$ (16641 to 1050625 points) for point Gauß-Seidel smoother and alternating line Gauß-Seidel smoother

l		7	8	9	10
point	$\bar\rho$	0.9681	0.9825	0.9858	0.9866
	T	14.6	106	525	3220
line	$\bar\rho$	0.0200	0.0254	0.0289	0.0308
	W	6.70	6.60	7.30	5.20
	$\hat\rho$	0.557	0.587	0.616	0.512
	T	0.92	3.6	15.9	65.1

Table 2. 3D Black-Scholes, plane smoother: average rate $\bar\rho$, cycle time $\bar t$ and overall solution time T (sec) for $\epsilon_2 = 10^{-k/3}$

k	1	2	3	4	5	10
$\bar\rho$	0.085	0.019	0.0070	0.0066	0.0066	0.0075
$\bar t$	9.6	11.7	13.1	14.6	16.5	27.5
T	115	82	79	87	99	165

2.2.1 2-dim basket

It is clear that lines can be inverted in linear complexity by a direct tridiagonal solver. From Table 1 and Fig. 2 we deduce that for the point smoother multigrid efficieny is lost, whereas for the alternating line smoother there is an asymptotical (grid independent) convergence rate.

2.2.2 3-dim basket

First we check the impact of the accuracy of the plane solutions on the overall convergence speed. Table 2 shows the avarage convergence rate $\bar\rho$, the average CPU-time $\bar t$ for a multigrid cycle, and total time T for an error reduction of $\epsilon = 10^{-12}$ at a 65^3 grid. The block-systems are hereby solved by Bi-CGSTAB with varying accuracy ϵ_2. We see that an error reduction below 0.1 is not reasonable. For the following simulations we therefore choose $\epsilon_2 = 0.1$.

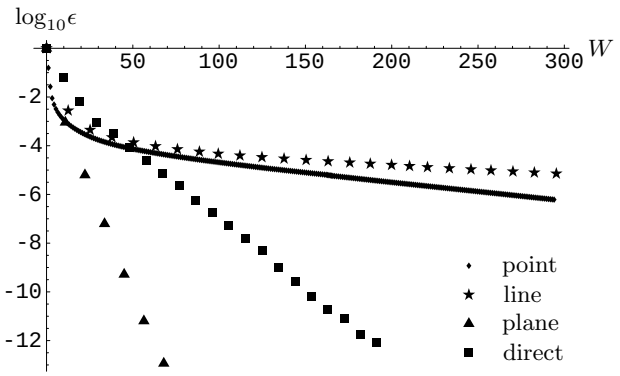

Fig. 3. 3D Black-Scholes: Error reduction ϵ vs. work units W for point Gauß-Seidel smoother, alternating line Gauß-Seidel smoother, alternating plane Gauß-Seidel smoother and direct (Bi-CGSTAB) solver

Table 3. 3D Black-Scholes: Average convergence rate $\bar{\rho}$, work units per cycle W and relative rate $\hat{\rho}$ for grid levels $l = 4, \ldots, 7$ (4913 to 2146691 points) for point Gauß-Seidel smoother, alternating line Gauß-Seidel smoother and alternating plane Gauß-Seidel smoother

	l	4	5	6	7
point	$\bar{\rho}$	0.74	0.93	0.98	
line	$\bar{\rho}$	0.2276	0.6710	0.9031	
	W	11.6	11.1	12.3	
	$\hat{\rho}$	0.88	0.96	0.99	
plane	$\bar{\rho}$	0.0006	0.026	0.0070	0.0158
	W	6.5	5.4	11.3	
	$\hat{\rho}$	0.31	0.51	0.65	

Next we compare the different block smoothers available in three dimensions. The alternating line smoother no longer shows multigrid efficiency and in spite of improved convergence rates compared to the point smoother is less efficient due to the higher cost involved (see Fig. 3, Table 3). Only the alternating plane smoother is robust and shows satisfactory rates.

However, since the number of CG-iterations required for the approximate solution of the plane systems increases proportional to the number of points in each direction (i. e. the square root of the condition number of the block systems), the solver still shows no linear complexity.

In fact, if we have n points in each of d directions, we have to solve $d \cdot n$ block systems per alternating smoothing step. The number of points in each hyperplane is n^{d-1}, the grid size n^{-1} yields a condition number proportional to n^2. Consequently, for an error reduction by a fixed factor ϵ_2, $\mathcal{O}(n)$ Bi-CGSTAB iterations are required with a cost of $\mathcal{O}(n^{d-1})$ each. In total $\mathcal{O}(n^{d+1})$ work units will be necessary for an iteration, that is for an error reduction of ϵ

$$W(N, \epsilon) \sim |\log \epsilon| \cdot N^{(d+1)/d}, \tag{19}$$

when the total number of unknowns is N. The superlinear behaviour gets less pronounced for higher dimensional problems.

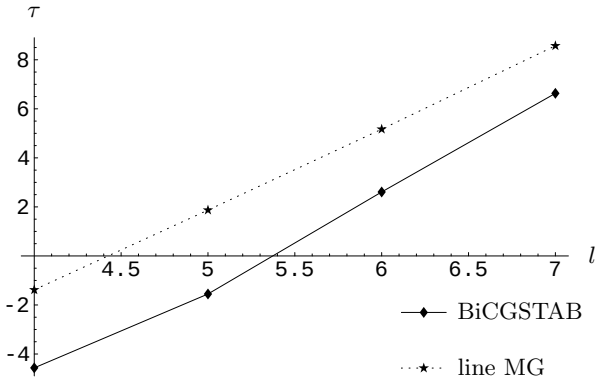

Fig. 4. 3D Black-Scholes: $\tau = -\frac{\log_2 T}{\log_{10} \epsilon}$ with CPU-time T vs. grid level l (4913 to 2146691 points) for error reduction $\epsilon \to 0$. Plane Gauß-Seidel smoothing with Bi-CGSTAB and line multigrid as block solvers

Table 4. 4D Black-Scholes: Average convergence rate $\bar{\rho}$ for grid levels $l = 3, 4, 5$ (6561, 83521 and 1185921 points) and soltution time T (secs) for BiCGSTAB solver and multigrid with alternating hyperplane Gauß-Seidel smoother

l		3	4	5
Bi-CGSTAB	T	0.36	18.5	809
MG	T	1.06	33.5	1292
	$\bar{\rho}$	$3 \cdot 10^{-7}$	$6.7 \cdot 10^{-5}$	$7 \cdot 10^{-4}$

To maintain multigrid efficiency for the iteration we substitute the CG-type block solver with an inner multigrid cycle with line Gauß-Seidel smoothing.

From the data plotted in Fig. 4 we can deduce an asymptotical order of about 4/3 for the complexity of the algorithm with plane smoothing (in accordance with the above considerations and (19)) and about 1 for the recursive multigrid method, both measured for an error reduction of 10^{-12}. For practical problems, though, the latter method is still too expensive due to the additional cost for the construction of a hierarchy of systems for the planes.

2.2.3 Higher dimensions

Already in four dimensions we see that for practically computable problems a plain conjugate gradient method is superior to the multigrid technique with blockwise smoothing on 3-dimensional hyperplanes (Table 4). The (uniform) grid size is not small enough to show the advantageous asymptotical properties of the multilevel approach.

However, for strongly anisotropic meshes, as they are encountered e. g. with the sparse grid combination technique [14], where often a large share of the points is distributed in lower-(1-,2-,3-)dimensional planes, multigrid applied recursively in plane smoothing is the method of choice.

Table 5. 3D transformed equation: Average convergence rate $\bar{\rho}$, number of iterations it and CPU-time T (sec) for an error reduction of 10^{-6} on grid levels $l = 4, \dots, 7$ (4913 to 2146691 points) for the plane smoother

l	4	5	6	7
$\bar{\rho}$	$3.8 \cdot 10^{-4}$	$5.3 \cdot 10^{-2}$	$2.9 \cdot 10^{-2}$	$6.1 \cdot 10^{-2}$
it	2	3	4	5
T	0.25	3.3	45	591

2.2.4 Transformed equations

To look further into the robustness of the plane smoother we check the convergence for the transformed equation (12). Here only $x_1 = 1$ is a Dirichlet boundary, on all other faces the equation degenerates, i. e. reduces to a lower-dimensional problem on these faces, which decouple from the rest of the region, and is arbitrarily anisotropic in their vicinity. The asset volatilities were chosen as $\sigma_1 = 0.01$, $\sigma_2 = 0.1$, $\sigma_3 = 0.5$, which cover the range of typical values e. g. for stocks. Table 5 shows the convergence results.

3 Multigrid for anisotropic obstacle problems

On a discrete level we seek a solution of problems of the form

$$\boldsymbol{A}_h \boldsymbol{u}_h \le \boldsymbol{b}_h, \tag{20}$$

$$\boldsymbol{u}_h \ge \boldsymbol{c}_h, \tag{21}$$

$$(\boldsymbol{A}_h \boldsymbol{u}_h - \boldsymbol{b}_h) \star (\boldsymbol{u}_h - \boldsymbol{c}_h) = \boldsymbol{0}, \tag{22}$$

where $\star$ symbolises componentwise multiplication. We denote the problem (20) to (22) as LCP($\boldsymbol{A}_h$, $\boldsymbol{b}_h$, $\boldsymbol{c}_h$).

We define for $\boldsymbol{u} \ge \boldsymbol{c}$ a projected residual $\boldsymbol{r}^+$ by

$$r_j^+ := \begin{cases} r_j & u_j \ge c_j, \\ \min(r_j, 0) & u_j = c_j, \end{cases}$$

and measure the error as the norm thereof.

By projection it is always guaranteed that the iterates lie in the admissible set $\{\boldsymbol{u} \ge \boldsymbol{c}\}$.

3.1 The projected multigrid cycle

We assign to (20)-(22) a coarse grid problem

$$\boldsymbol{A}_H \boldsymbol{d}_H \le \boldsymbol{b}_H := \bar{I}_H^h(\boldsymbol{A}_h \boldsymbol{u}_h - \boldsymbol{b}_h), \tag{23}$$

$$\boldsymbol{d}_H \ge \boldsymbol{c}_H := \hat{I}_H^h(\boldsymbol{c}_h - \boldsymbol{u}_h), \tag{24}$$

$$(\boldsymbol{A}_H \boldsymbol{d}_H - \boldsymbol{b}_H) \star (\boldsymbol{d}_H - \boldsymbol{c}_H) = \boldsymbol{0}, \tag{25}$$

where $\bar{I}_H^h$, $\hat{I}_H^h$ are restriction operators and the solution is updated by $u_h \leftarrow u_h + \bar{I}_h^H d_H$ with some prolongation $\bar{I}_h^H$.

Problem (23) to (25) defines a multilevel iteration of a structure similar to the linear case (see Algorithm 1), which is outlined in the following template algorithm (to this end we drop the subscripts h, H and instead indicate by superscript l the multigrid level).

Algorithm 2 ***Projected V-Cycle***

procedure ***PV****(l, $\boldsymbol{A}^l$, $\boldsymbol{b}^l$, $\boldsymbol{c}^l$, $\boldsymbol{u}^l$)*
begin
  if $l = 1$ ***then***
    $\boldsymbol{u}^1 \leftarrow$ *LCP*($\boldsymbol{A}^1$, $\boldsymbol{b}^1$, $\boldsymbol{c}^1$);
  else begin
    $\boldsymbol{u}^l \leftarrow \mathcal{S}(\boldsymbol{A}^l, \boldsymbol{b}^l, \boldsymbol{c}^l, \boldsymbol{u}^l)$;
    $\boldsymbol{r}^l = \boldsymbol{b}^l - \boldsymbol{A}\boldsymbol{u}^l$;
    $\boldsymbol{r}^{l-1} = \bar{I}_l^{l-1} \boldsymbol{r}^l$;
    $\boldsymbol{c}^{l-1} = \hat{I}_l^{l-1} \left(\boldsymbol{c}^l - \boldsymbol{u}^l\right)$;
    $\boldsymbol{d}^{l-1} \leftarrow \boldsymbol{PV}(l-1, \boldsymbol{A}^{l-1}, \boldsymbol{r}^{l-1}, \boldsymbol{c}^{l-1}, \boldsymbol{d}^{l-1})$;
    $\boldsymbol{d}^l = \bar{I}_{l-1}^l \boldsymbol{d}^{l-1}$;
    $\boldsymbol{u}^l = \boldsymbol{u}^l + \boldsymbol{d}^l$;
    $\boldsymbol{u}^l \leftarrow \mathcal{S}(\boldsymbol{A}^l, \boldsymbol{b}^l, \boldsymbol{c}^l, \boldsymbol{u}^l)$;
  end
return *Projected V-Cycle*

3.2 Coarse grid solver

Projected Krylov-space techniques have been proposed by Braess et al. [2] as well as Dostal and Schöberl [6] for symmetric problems. A generalisation to the non-symmetric case seems not straightforward.

Therefore we employ a Bi-CG accelaration of projected Gauß-Seidel, similar to SSORP-PCG as proposed in [15] (for symmetric systems). In each iteration we perform m SOR-steps followed by s Bi-CGSTAB steps restricted to the inactive set. Typically we choose $m = 1$ and s such that the residual is reduced by a factor ϵ, e. g. $\epsilon = 0.3$.

3.3 Grid transfer

For the restriction operator a minimum requirement is that for the exact discrete solution ($\boldsymbol{u}_h$) of LCP($\boldsymbol{A}_h$, $\boldsymbol{b}_h$, $\boldsymbol{c}_h$)

$$\bar{I}_H^h(\boldsymbol{A}_h \boldsymbol{u}_h - \boldsymbol{b}_h) \ge \boldsymbol{0},$$

$$\hat{I}_H^h(\boldsymbol{c}_h - \boldsymbol{u}_h) \le \boldsymbol{0},$$

$$\bar{I}_H^h(\boldsymbol{A}_h \boldsymbol{u}_h - \boldsymbol{b}_h) \star \hat{I}_H^h(\boldsymbol{c}_h - \boldsymbol{u}_h) = \boldsymbol{0}, \tag{26}$$

i. e. $\boldsymbol{d}_H = \boldsymbol{0}$ solves the coarse grid system (23) to (25), which guarantees that $\boldsymbol{u}_h$ is a fixed point of the iteration. Particularly (26) is crucial and not satisfied in general, e. g. for full weighting. The use of injection, which clearly fulfils (26), implies deteriorated convergence.

Following [8,10] we introduce one-sided restriction operators at the free boundary as

$$\bar{I}_H^h \boldsymbol{u}_h(x_h) = \begin{cases} I_H^h \bar{\boldsymbol{u}}_h^{\mathcal{A}} & \text{if } x_h \in \mathcal{A}, \\ I_H^h \bar{\boldsymbol{u}}_h^{\mathcal{I}} & \text{if } x_h \in \mathcal{I}, \end{cases} \tag{27}$$

with full weighting I_H^h and

$$\bar{\boldsymbol{u}}_h^{\mathcal{A}/\mathcal{I}}(x_h) = \begin{cases} \boldsymbol{u}_h(x_h) & \text{if } x_h \in \mathcal{A}/\mathcal{I}, \\ 0 & \text{else.} \end{cases}$$

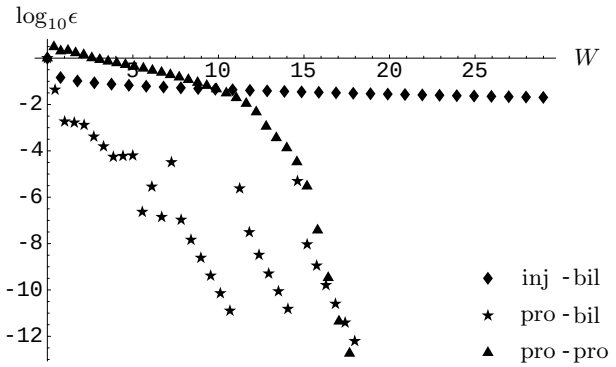

Fig. 5. 2D obstacle problem: Error reduction ϵ vs. work units W for V(1,1) cycles: injection and bilinear interpolation, projected full-weighting and bilinear interpolation, projected full weighting and projected bilinear interpolation

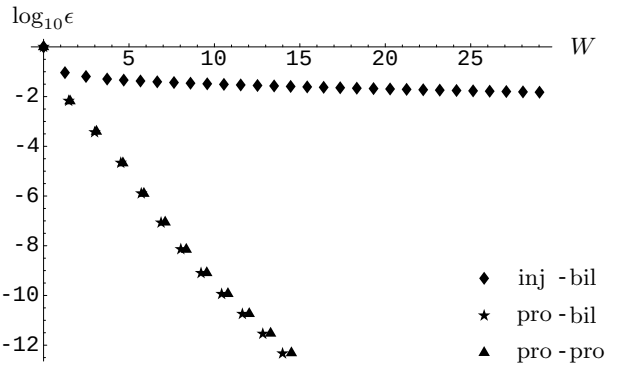

Fig. 6. 2D obstacle problem: Error reduction ϵ vs. work units W for V(1,1) cycles succeeding three F(1,1) steps: injection and bilinear interpolation, projected full-weighting and bilinear interpolation, projected full weighting and projected bilinear interpolation

Here $\mathcal{A}$ and $\mathcal{I}$ are the discrete active and inactive sets, respectively (for simplicity of notation the subscript h is supressed). The corresponding prolongation reads

$$\bar{I}_H^h \boldsymbol{u}^H(x_h) = \begin{cases} 0 & \text{if } \boldsymbol{c}_h(x_h) = 0, \\ I_H^h \boldsymbol{u}^H(x_h) & \text{else.} \end{cases} \tag{28}$$

(27) and (28) seperate active and inactive sets and preserve complementarity on all grid levels. Note that the coarse grid correction keeps the position of the free boundary fixed with this choice of transfer operators.

The obstacle is restricted via injection $\hat{I}_l^{l-1}$. The following numerical data were obtained with a smoother motivated by the one introduced in [4] (see subsection 3.5).

We clearly see from Fig. 5 the lost approximation order of injection. With one-sided (projected) transfer operators as introduced above the rates are roughly the same as in the unrestricted case once the active set has been found. However, in the initial phase convergence is retarded as the free boundary is only corrected on the finest grid. If bilinear prolongation is not restricted to the active set, this can be partly avoided. In Fig. 5 the jumps of the residual indicate the shifts of the free boundary.

A good initial approximation of the active set can be obtained from coarser levels by full multigrid cycles.

3.4 Full multigrid

Evidently a fast finding of the active set is vital for the overall convergence. For that purpose we start with an approximation at the coarsest mesh and construct a sequence of approximations on the finer levels.

Algorithm 3 ***Projected Full Multigrid***

procedure PFM($\boldsymbol{A}^L$, $\boldsymbol{b}^L$, $\boldsymbol{c}^L$, $\boldsymbol{u}^L$ ***)***
begin
  $\boldsymbol{u}^1 \leftarrow$ *LCP(*$\boldsymbol{A}^1$, $\boldsymbol{b}^1$, $\boldsymbol{c}^1$*);*
  for $l = 2, \ldots, L$ ***begin***
    $\boldsymbol{u}^l = I_{l-1}^l \boldsymbol{u}^{l-1}$;
    for $k = 1, \ldots, n_l$ ***begin***
      $\boldsymbol{u}^l \leftarrow$ ***PV(***l, $\boldsymbol{A}^l$, $\boldsymbol{b}^l$, $\boldsymbol{c}^l$, $\boldsymbol{u}^l$***);***
    end
  end
return *Projected Full Multigrid*

Since the topology of the active set is simple (see Fig. 7), a few full multigrid steps are sufficient. After that the problem reduces to a linear system on the inactive set. Globally bilinear and projected prolongation yield almost the same results (see Fig. 6).

3.5 Robust smoothers

The projected Gauß-Seidel algorithm

$$u_i^{k+\frac{1}{2}} = \frac{1}{a_{ii}} \left(b_i - \sum_{j<i} a_{ij} u_j^{k+1} - \sum_{j>i} a_{ij} u_j^k \right) \tag{29}$$

$$u_i^{k+1} = \max\left(u_i^{k+\frac{1}{2}}, c_i \right) \tag{30}$$

has been frequently applied as smoother for obstacle problems (e. g. [3]), but for the problem at hand convergence is poor. There are two evident ways to generalize the idea of plane smoothers to obstacle problems.

First one could think of a blockwise solution of obstacle problems,

$$\boldsymbol{u}_i^{k+1} = \mathcal{A}_{ii}^{-1}\big(\boldsymbol{b}_i - \sum_{j<i} \boldsymbol{A}_{ij} \boldsymbol{u}_j^{k+1} - \sum_{j>i} \boldsymbol{A}_{ij} \boldsymbol{u}_j^k, \boldsymbol{c}_i\big). \tag{31}$$

Here $\mathcal{A}_{ii}^{-1}(\boldsymbol{b}_i, \boldsymbol{c}_i)$ denotes the solution of the obvious linear complementarity problem involved on the planes. We will use the combined iteration from section 3.2 for the planewise solution.

Another choice was proposed in [4]. Here a projected pointwise Gauss-Seidel step is applied followed by line smoothing on the inactive sets.

Table 6 shows, which is also the generally observed tendency, that the hybrid smoother shows slightly better rates and has the additional bonus that only linear problems have to be solved in the smoothing steps.

Table 6. 2D obstacle problem: CPU-time T (sec) and average convergence rate $\bar{\rho}$ for an error reduction of 10^{-12} on grid levels $l = 5, 6, 7, 8, 9$ (1089 to 263169 points) for projected line smoother and hybrid smoother

l		5	6	7	8	9
line	T	0.31	1.3	4.9	23	138
	$\bar{\rho}$	0.047	0.03	0.043	0.059	0.099
hybrid	T	0.26	0.92	3.8	16	85
	$\bar{\rho}$	0.018	0.028	0.041	0.043	0.051

Table 7. 3D transformed obstacle problem: Average convergence rate $\bar{\rho}$, number of iterations it and CPU-time T for an error reduction of 10^{-6} on grid levels $l = 4, \ldots, 7$ (4913 to 2146691 points) for the hybrid smoother

l	4	5	6	7
$\bar{\rho}$	$1.8 \cdot 10^{-4}$	$4.6 \cdot 10^{-3}$	$2.4 \cdot 10^{-2}$	$8.9 \cdot 10^{-2}$
it	2	3	4	6
T	0.28	3.5	42	867

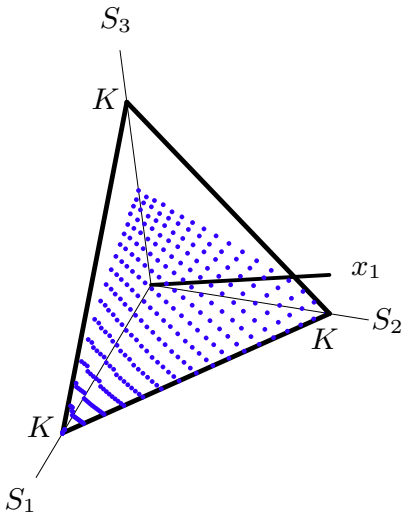

Fig. 7. Free boundary in the space spanned by three assets with $\sigma_1 = 0.01$, $\sigma_2 = 0.1$, $\sigma_3 = 0.5$. The exercise region for the basket put with strike K lies beneath this surface.

Now we go back to the transformed system and compare convergence with the linear case. The first three steps were always $F(1,1)$-cycles, the rest $V(1,1)$. One-sided restriction and prolongation operators were used. Table 7 shows comparable results to the linear case (Table 5). The free boundary is shown in Fig. 7.

3.6 Cascadic multigrid

In [2] Braess advocates a cascadic multigrid algorithm for variational inequalities. The idea is attractive since no restriction to coarser grid levels is required. Starting from the coarsest grid, the solution is prolonged to finer grid levels, where they are corrected by sufficiently many smoothing steps.

Table 8. Initial error $\|\boldsymbol{r}_0^{+,l}\|$, iterations n_l and final error $\|\boldsymbol{r}^{+,l}\|$ on all levels l with N_l degrees of freedom

l	N_l	$\|\boldsymbol{r}_0^{+,l}\|$	$\|\boldsymbol{r}^{+,l}\|$	n_l
3	81	$1.94\ 10^{1}$	$2.67\ 10^{-4}$	3
4	289	$7.98\ 10^{0}$	$1.33\ 10^{-3}$	3
5	1089	$3.31\ 10^{0}$	$1.18\ 10^{-3}$	3
6	4225	$1.30\ 10^{0}$	$1.04\ 10^{-3}$	3
7	16641	$4.94\ 10^{-1}$	$6.52\ 10^{-3}$	2
8	66049	$1.93\ 10^{-1}$	$2.03\ 10^{-3}$	2
9	263169	$9.15\ 10^{-2}$	$5.94\ 10^{-4}$	2

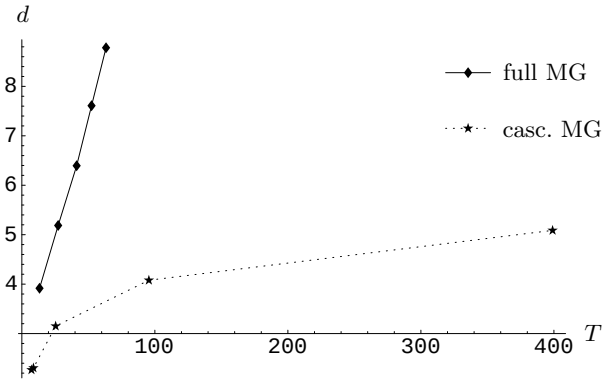

Fig. 8. 2D obstacle problem: CPU-time T required for d valid digits with full and cascadic multigrid (513^2 points)

Algorithm 4 ***Projected Cascadic Multigrid***

procedure ***PCM***$(\boldsymbol{A}^L, \boldsymbol{b}^L, \boldsymbol{c}^L, \boldsymbol{u}^L)$
begin
$\boldsymbol{u}^1 \leftarrow LCP(\boldsymbol{A}^1, \boldsymbol{b}^1, \boldsymbol{c}^1)$;
for $l = 2, \ldots, L$ ***begin***
  $\boldsymbol{u}^l = I_{l-1}^l \boldsymbol{u}^{l-1}$;
  while $\|\boldsymbol{r}^l\| > \epsilon_l \|\boldsymbol{r}_0^l\|$ ***begin***
    $\boldsymbol{u}^l \leftarrow \mathcal{S}(\boldsymbol{A}^l, \boldsymbol{b}^l, \boldsymbol{c}^l, \boldsymbol{u}^l)$;
  end
end
return *Projected Cascadic Multigrid*

Here $\boldsymbol{r}_0^l$ is the starting residual at level l. Since coarse grids are not revisited it is essential that error components corresponding to large wavelengths are sufficiently damped originally. In [2] the following stopping criterion is derived for the level l:

$$\epsilon_l = \epsilon_0 0.4^{L-l}. \tag{32}$$

The number of smoothing steps per iteration is then almost constant (Table 8 for $\epsilon_0 = 0.1$). Therefore cascadic multigrid shows linear complexity down to an error comparable to the discretisation error (see Fig. 8). For extrapolation methods like the sparse grid combination technique this is not sufficient though.

4 Conclusions

We have focussed on two major aspects that we encounter when applying multigrid methods to option pric-

ing problems: anisotropies and free boundaries. The numerical results point out clearly that plane smoothers are expensive, but robust with respect to anisotropies and degeneracy of the equation towards the boundary. In fact, they are efficient also for the solution of higher-dimensional problems on sparse grids [14]. For American-style contracts, adapted smoothers and careful grid transfer around the obstacle provide comparable results also in the presence of constraints on the solution.

References

1. F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. Pol. Econ.*, pages 637–659, 1973.
2. H. Blum, D. Braess, and F. T. Suttmeier. A cascadic multigrid algorithm for variational inequalities. *Computing and Visualization in Science*, 2003. submitted.
3. A. Brandt and C. W. Cryer. Multigrid algorithms for the solution of linear complementarity problems arising from free boundary problems. *SIAM J. Sci. Stat. Comput.*, 4(4), December 1983.
4. N. Clarke and K. Parrott. Multigrid for American option pricing with stochastic volatility. *Appl. Math. Finance*, 6:177–197, 1999.
5. C. W. Cryer. The solution of a quadratic programming problem using systematic overrelaxation. *SIAM J. Control*, 9:385–392, 1971.
6. Z. Dostál and J. Schöberl. Modified proportioning for bound constraint quadratic programming with the rate of convergence and finite termination. Technical report, Technical University Ostrava, 2002.
7. J.-P. Fouque, G. Papanicolaou, and R. Sircar. *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge University Press, 2000.
8. R. Hoppe and R. Kornhuber. Multi-grid methods for the two phase Stefan problem. Technical Report 171, TU Berlin, 1987.
9. J. C. Hull. *Options, Futures & Other Derivatives*. Prentice-Hall International, fourth edition, 2000.
10. B. Kawohl, J. Stará, and G. Wittum. Analysis and numerical studies of a shape design problem, 1990. Preprint. Sonderforschungsbereich 123, Stochastische Mathematische Modelle.
11. I. M. Llorente and N. D. Melson. Behavior of plane relaxation methods as multigrid smoothers. *Electronic Transactions on Numerical Analysis*, 10:92–114, 2000.
12. C. W. Oosterlee. A GMRES-based plane smoother in multigrid to solve 3D anisotropic fluid flow problems. *Journ. Comp. Physics*, 130:41–53, 1997.
13. C. W. Oosterlee. On multigrid for linear complementarity problems with application to American-style options. To appear in Electr. Trans. Num. Analysis, 2002.
14. C. Reisinger. *Numerische Methoden für hochdimensionale parabolische Gleichungen am Beispiel von Optionspreisaufgaben*. PhD thesis, Universität Heidelberg, 2004.
15. A. Siddiqi, P. Manchanda, and M. Kŏcvara. An iterative two step algorithm for American option pricing. working paper.
16. H. v. d. Vorst. Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. *SIAM J. Sci. Stat. Comp.*, 13(2):631–644, 1992.
17. P. Wilmott, J. N. Dewynne, and S. D. Howison. *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press, 1993.
18. Y. Zhu and J. Li. Multi-factor financial derivatives on finite domains. working paper.