Electronic Transactions on Numerical Analysis.
Volume 15, pp. 165-185, 2003.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

ON MULTIGRID FOR LINEAR COMPLEMENTARITY PROBLEMS WITH APPLICATION TO AMERICAN-STYLE OPTIONS*

C. W. OOSTERLEE ‡

Abstract. We discuss a nonlinear multigrid method for a linear complementarity problem. The convergence is improved by a recombination of iterants. The problem under consideration deals with option pricing from mathematical finance. Linear complementarity problems arise from so-called American-style options. A 2D convection-diffusion type operator is discretized with the help of second order upwind discretizations. The properties of smoothers are analyzed with Fourier two-grid analysis. Numerical solutions obtained for the option pricing problem are compared with reference results.

Key words. linear complementarity problems, American-style options, nonlinear multigrid, projected Gauss-Seidel, iterant recombination, second-order upwind discretization, Fourier analysis

AMS subject classifications. 65M55, 65F99, 90A09

1. Introduction. In this paper, we discuss multigrid methods for solving a time-dependent 2D partial differential equation (PDE) arising in option pricing theory. The problem considered is the computation of the value of an American-style option in a stochastic volatility setting. It leads to the solution of a convection-diffusion type PDE with a free boundary. In [23], it has been shown that for American-style options the theory of linear complementarity, as it was developed in the 1970's, applies. It is possible to rewrite the arising free boundary problem as a linear complementarity problem (LCP) of the form

$$Lu \le f_1 \quad \mathbf{x} \in \Omega \tag{1.1}$$

$$u \ge f_2 \quad \mathbf{x} \in \Omega \tag{1.2}$$

$$(u - f_2)(Lu - f_1) = 0 \quad \mathbf{x} \in \Omega, \tag{1.3}$$

plus boundary conditions, where L is a linear differential operator. The option pricing context and the discretization of the LCP are discussed in Section 2. The LCP formulation is beneficial for iterative solution, since the unknown boundary does not appear explicitly and can be obtained in a postprocessing step.

In 1983, Brandt and Cryer [3] proposed a multigrid method for LCPs arising from free boundary problems. The algorithm is a multigrid generalization of the projected SOR method [7]. Due to the nonlinear character of the problem, the multigrid method is based on the full approximation scheme [2], FAS, that is often used for solving nonlinear PDEs. The solution method has therefore been called the projected full approximation scheme (PFAS) in [3]. In the original paper, the operator L in (1.1), (1.3) was the nicely elliptic Laplace operator and fast convergence was presented. PFAS has already been successfully used in the financial community for American options with stochastic volatility in [6]. The smoother applied in [6] is, however, somewhat involved. It is based on the pointwise PSOR method for the detection of the free boundary, followed by a partial linewise step in order to deal with the stretched numerical grid occurring in the financial problem. As the free boundary is unknown and can be of general shape, the line relaxation may often need to be adapted. The need to change multigrid components like the smoother for optimal convergence of a new problem at hand is sometimes considered as unsatisfactory. This is, for example, stated in [25] where an alternative formulation for American options with a penalty function is proposed and an

* Received May 01, 2001. Accepted for publication October 21, 2001. Recommended by Irad Yavneh.
‡Delft University of Technology, Faculty of Information Technology and Systems, Delft, the Netherlands(email: c.w.oosterlee@math.tudelft.nl)

ILU preconditioned CG-type solver is considered. This solver can be considered as being of "black-box" type, but it is not hierarchical.

It is our aim with this paper to give more insight into the multigrid convergence for problems from option pricing. Fourier two-grid analysis [19] will be used to analyze favorable smoothers and coarse grid correction components for the discrete operator under consideration. The influence of different point- and linewise smoothers, or of under- and overrelaxation parameters can be analyzed quantitatively in this way, as presented in Section 3.2. At the same time, we introduce in Section 3 some recent developments in multigrid methods to the field of LCPs, making the algorithms more robust, like overrelaxation parameters and recombination techniques.

In the discretization of the operator arising in American-style option pricing by a second order upwind scheme, different complications for optimal multigrid convergence, such as anisotropies and positive off-diagonal stencil elements in space-dependent operator coefficients, occur simultaneously. When pointwise Gauss-Seidel-type smoothing is combined with a standard coarse grid correction, certain error components may remain large. In such a situation, it is possible to choose a more expensive smoother, such as linewise relaxation or to change the coarsening procedure, for example to semi-coarsening, or to improve the convergence by a Krylov subspace acceleration. We will choose the subspace acceleration approach here.

A well-known solution approach of the latter type for *nonlinear equations* is to apply global (Newton) linearization, solve the arising linear system with a Krylov subspace method, such as the GMRES method [18], and a multigrid preconditioner. This is not the approach followed in this paper. We apply to the nonlinear problems a solution method based on PFAS as the multigrid technique. The Krylov subspace acceleration can be interpreted as a technique, in which intermediate iterates are recombined in order to obtain an improved approximation, as explained in [20]. With this method many different nonlinear partial differential equations have been solved, see, for example, [19, 16]. The solver is related to the acceleration cycle presented in [4]. Here, we generalize this method to solving LCPs in Section 4. Numerical results are presented in Section 5.

2. Option basics and the Black-Scholes equation. Research in option pricing theory concerns, among many other issues, the computation of the value of an option during the life of an option contract. A famous equation for this is the Black-Scholes partial differential equation. It represents a simple model for the values of two basic options, the so-called put option and call option.

In the case of a European *put* option, the holder of the option may *sell at the expiration date*, a prescribed time in the future, certain assets, like shares or stocks, at the exercise price. The other party of the option contract, the writer, must buy the asset, if the holder decides to sell it. In the case of a European *call* option, the holder has the right to *purchase* an asset on a certain date at a prescribed amount. The writer is then obliged to sell the asset. In this paper, we concentrate on the put option.

Options have two main uses: speculation and hedging. Whereas speculation might be well-known, hedging needs some more explanation. Let's consider a portfolio with assets and put options. If the asset price s falls, (based on the definition above) the value of a put option increases: It is possible to sell assets at the expiration date for the exercise price, although the actual price falls. The value of the portfolio therefore depends on the ratio between the number of assets and the number of put options in the portfolio. A ratio exists, which results in no movement in the value of the portfolio. This ratio is instantaneously risk-free. Hedging means here reducing risk (for example against falling asset prices) by combining options with assets in a portfolio.

To derive the Black-Scholes equation, one assumes a geometric Brownian motion stochastic differential equation as a model generating asset prices s to be valid,

$$ds = s(\sigma dw + \mu dt) \ . \tag{2.1}$$

Here, μdt is a deterministic return. The volatility σ measures the standard deviation of the returns ds/s. The random variable dw is assumed to be a Wiener process with mean 0 and variance dt, so that the mean of ds is $\mu s dt$ and its variance is $\mathcal{E}[ds^2] - \mathcal{E}[ds]^2 = \sigma^2 s^2 dt$. Furthermore, the equation has the Markov property: It does not depend on the past history of the asset prices.

Although asset prices are valid for discrete time, the PDE models are set up in continuous time with limit $dt \to 0$. The value of the option is denoted by $u = u(s,t)$ here. The value u is influenced by the exercise price E, the expiration time T ($0 \le t \le T$), the interest rate r (here assumed to be constant), and the volatility σ.

Itô's lemma makes it possible to handle the random term dw as $dt \to 0$ (analogously to Taylor's expansion for deterministic variables). With the insight that the random walks in s and u are driven by the same random variable dw, one can add in a portfolio an option with value u to a number, Delta "Δ", of assets s in such a way that the portfolio is instantaneously risk free. By choosing $\Delta := \partial u/\partial s$, the deterministic portfolio increment is obtained. At this stage, the assumption that there are no arbitrage possibilities comes into play, which means that risk free profits that are greater than placing money at a bank are not allowed in the model. So, the instantaneously risk free portfolio must earn a rate of return that equals the interest rate [10, 23]. Together with the assumptions in this model's framework of constant volatility, of no transaction costs for hedging, no dividend payment and no taxes involved, this finally results in the Black-Scholes partial differential equation for the value of an option,

$$Lu := \frac{\partial u}{\partial t} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 u}{\partial s^2} + rs\frac{\partial u}{\partial s} - ru = 0 \tag{2.2}$$

In its basic form, (2.2) is a convection-diffusion-reaction type equation in one "space-like" dimension, the s direction. With the terms $s^j \partial^j u/\partial s^j$, one recognizes an Eulerian differential equation that can be transformed into a heat equation.

The Black-Scholes equation is a parabolic PDE with boundary and final conditions. At expiry a European put option has the value $E - s$ for $s < E$. It is worthless if $s > E$, see Figure 2.1. So,

$$u(s,T) = \max(E - s, 0) =: f_2(s). \tag{2.3}$$

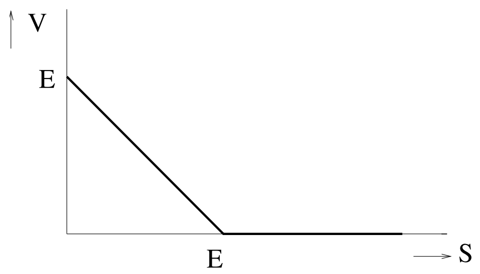

FIG. 2.1. *Final conditions for a put option.*

The boundary condition for a put (see, for example [10, 23]) at $s = 0$ is $u(0,t) = Ee^{-r(T-t)}$. It represents the exponential growth for receiving an amount E at $t = T$ with

ETNA
Kent State University
etna@mcs.kent.edu

constant interest rates. Furthermore, $u(s,t) \to 0$ as $s \to \infty$, because one obviously cannot gain by exercising the put option.

American options. Whereas for European options exercise is only possible at the expiration date, the exercise of *American options* is permitted at any time during the life of an option, $0 < t \leq T$. This leads to the following considerations. It is well-known that a European put option is, in a certain s range, less than the pay-off function, $u(s,t) < \max(E-s,0)$ [23]. In the American context, this would mean that buying the option for u, selling the asset immediately for E, and buying the asset on the market for s, would result in a risk-free profit of $E - u - s > 0$ in this s range, which contradicts the arbitrage concept [10, 23]. Therefore, when early exercise is permitted, a constraint

$$u(s,t) \geq \max(E-s,0) =: f_2(s) \tag{2.4}$$

must be imposed in order to avoid an arbitrage possibility. In this s region, the value of the American put option is raised compared to the value of the similar option of European type, due to the constraint (2.4). The addition of the constraint to the PDE (2.2) or to (2.7) gives rise to a free boundary problem: A special s value exists, the optimal exercise price s_f, with the following properties: on one side of s_f it is beneficial to hold the option, on the other side, it is advantageous to exercise the option. $s_f(t)$ is time dependent and not known in advance. (Of course, optimal exercise also depends on other important parameters and, most of all, on the option holder's market strategy.) For an American put, if $u > \max(E-s,0)$ equation (2.2) or (2.7) holds ($s > s_f(t)$); If $u = E - s$, an inequality,

$$Lu < 0 \tag{2.5}$$

holds ($s \leq s_f(t)$) with L as in (2.2) or in (2.7).

Based on the classical theory of linear complementarity, it is possible to formulate this free boundary problem into an LCP of type (1.1),(1.2),(1.3), so that the free boundary conditions need not be handled explicitly.

Stochastic volatility. A generalization of the Black-Scholes equation is obtained if the restriction of constant volatility is replaced by the assumption of a *stochastic volatility*. Instead of (2.1), the following stochastic differential equations are assumed to govern the asset price process s and its variance process y [11],

$$\begin{aligned} ds &= \mu s dt + \sqrt{y} s dw_1 \\ dy &= \alpha(\beta - y)dt + \gamma\sqrt{y} dw_2 \end{aligned} \tag{2.6}$$

where $w_1,\ w_2$ represent standard Brownian motion with correlation coefficient $\overline{\rho} \in [-1,1]$, γ is the volatility of the variance process, $\alpha,\ \beta > 0$ determine the mean reversion (so that the variance will drift back to some mean value β at a rate governed by α). The volatility $\sqrt{y}$ can be shown to be positive valued provided that $\gamma^2 \leq 2\alpha\beta$. The assumption of the stochastic processes (2.6) leads to a 2D PDE problem for the value of an option in which the variance y is, besides s and t, a third variable (degree of freedom), see [1, 25] for the derivation. The resulting PDE reads,

$$\begin{aligned} Lu := \frac{\partial u}{\partial t} + \frac{1}{2}[s^2 y \frac{\partial^2 u}{\partial s^2} + 2\overline{\rho}\gamma y s \frac{\partial^2 u}{\partial s \partial y} + \gamma^2 y \frac{\partial^2 u}{\partial y^2}] + \\ rs\frac{\partial u}{\partial s} + [\alpha(\beta - y) - \lambda\gamma\sqrt{y}]\frac{\partial u}{\partial y} - ru = 0, \quad \Omega = \{(s,y)|s \geq 0, y \geq 0\}, \end{aligned} \tag{2.7}$$

where λ is the so-called "market price of the risk function" (for foreign currencies a nonzero constant is parameter of choice). It has been set to 0 here, as in [6, 25]. For more details on the stochastic volatility concept, we refer to the financial literature [1, 10, 25]. The following boundary conditions are proposed for a put option in [6],

$$u(0, y, t) = f_2(0, t), \ \ \forall y \geq 0, \ \ t \in [0, T], \tag{2.8}$$
$$u(s, 0, t) = f_2(s, t), \ \ \forall s \geq 0, \ \ t \in [0, T], \tag{2.9}$$

with f_2 defined by (2.3). These boundary conditions imply an immediate exercise of the put, in the case of a zero asset price and in the case of zero volatility. As for the other two boundaries ($s \to \infty, \ y \to \infty$), the computational domain is commonly truncated at finite values $s_{\max}, \ y_{\max}$ respectively, at which Neumann boundary conditions are imposed [6],

$$\frac{\partial u(s_{\max}, y, t)}{\partial s} = 0, \ \ \forall y \in [0, y_{\max}], \ \ t \in [0, T] \tag{2.10}$$
$$\frac{\partial u(s, y_{\max}, t)}{\partial y} = 0, \ \ \forall s \in [0, s_{\max}], \ \ t \in [0, T]. \tag{2.11}$$

Typically, $y_{\max}$ is set to one, $s_{\max}$ to two times E. The computed values of the options are generally not very sensitive with respect to the size of the truncation, as discussed in [6].

Summarizing, for American style options with the stochastic volatility model (2.6),(2.7), the following LCP needs to be solved,

$$Lu(s, y, t) \leq f_1 \tag{2.12}$$
$$u(s, y, t) \geq f_2(s) \tag{2.13}$$
$$(u(s, y, t) - f_2(s))Lu(s, y, t) = 0 \tag{2.14}$$

with f_2 defined by (2.4), $f_1 = 0$, L given by (2.7) and boundary conditions (2.8)–(2.11).

2.1. The discretization. The stochastic volatility problem leads to the numerical solution of a 2D time-dependent problem with an operator L of convection-diffusion type. Here we outline the discretization of the operator in (2.7), which is in nonconservative form. Since we expect solutions without steep gradients, it is possible to choose standard well-known discretization schemes. After a transformation $t^* = T - t$, the equation which is backward in time with a final condition is transformed to an equation forward in time with an initial condition. This puts a minus sign in front of the $\partial/\partial t$ term in (2.7).

For the time discretization, we consider the so-called *backward difference formula BDF2* [8] scheme,

$$\frac{3u_h(s, y, t^* + \tau) - 4u_h(s, y, t^*) + u_h(s, y, t^* - \tau)}{2\tau} = L_h u_h(s, y, t^* + \tau) \ . \tag{2.15}$$

The time discretization accuracy of this implicit scheme is $O(\tau^2)$. We prefer this discretization over the well-known Crank-Nicolson scheme (also called trapezoid rule), because of its better stability characteristics. The Crank-Nicolson scheme is not *L-stable* (see, for example, [9]), whereas BDF2 is. BDF2 has more favorable damping properties than the Crank-Nicolson scheme. In a forthcoming paper discussing another type of option, we will show that the latter can result in undamped oscillations in important financial quantities, called the hedge parameters.

The mixed derivative term is discretized by the $O(h^2)$ four-point discretization. The second derivatives in both directions are handled by the usual three-point stencils. Linear second order upwind discretizations, like the upwind κ-discretizations [14], are sufficiently accurate for the discretization of the convective terms in the $s-$ and $y-$directions. A 1D

ETNA
Kent State University
etna@mcs.kent.edu

upwind κ-discretization can be written as a combination of a central difference discretization plus a second-order dissipation term, which is proportional to the third derivative of u:

$$(au_s)_h = a\left(\frac{u_h(s+h_s,y,t^*)-u_h(s-h_s,y,t^*)}{2h_s}\right)+\left(\frac{ah_s^2(1-\kappa)}{4}\right)\cdot \left(\frac{u_h(s-h_s,y,t^*)-3u_h(s,y,t^*)+3u_h(s+h_s,y,t^*)-u_h(s+2h_s,y,t^*)}{h_s^3}\right) \tag{2.16}$$

for $a<0$ (for example, $a=-rs$ after the transformation). In this paper, we use the Fromm scheme for the discretization (i.e., $\kappa=0$). The well-known central discretization scheme $\kappa=1$ again leads to oscillations in financial quantities, see for example [26]. The linear κ-scheme is a first choice for obtaining second-order accurate schemes with a convection term. The κ-schemes are, however, not monotone, which means that they have to be modified (with limiters) for problems containing strong gradients, like shocks or boundary layers. In that case, the BDF2 time discretization also needs to be modified in order to guarantee total variation diminishing (TVD) solutions in space and time with relatively large time steps. So-called partially implicit BDF blended schemes [12] from the family of implicit-explicit (IMEX) time discretization schemes are an alternative in this situation. We will use them in another financial setting in a forthcoming paper.

Since we will not encounter extremely steep gradients in this paper, the second order κ upwind schemes and BDF2 time discretization are fully satisfactory. Unrealistic oscillations in time or in (s,y)-space are not encountered.

By setting $s_i=i\cdot h_s,\ y_j=j\cdot h_y,\ \ i,j:0,\dots,n$, (i.e. choosing an equidistant grid G_h), the stencil for the second order accurate discretization of (2.7) transformed into an equation forward in time reads in semi-discrete form

$$\frac{\partial u_h(s,y,t^*)}{\partial t^*}+\begin{bmatrix} & a^{(2)}_{0,2} & & \\ a^{(2)}_{-1,1} & a^{(2)}_{0,1} & a^{(2)}_{1,1} & \\ a^{(2)}_{-1,0} & a^{(2)}_{0,0} & a^{(2)}_{1,0} & a^{(2)}_{2,0} \\ a^{(2)}_{-1,-1} & a^{(2)}_{0,-1} & a^{(2)}_{1,-1} & \\ & a^{(2)}_{0,-2} & & \end{bmatrix} u_h(s,y,t^*)=0, \qquad (s,y,t^*)\in\Omega_h\times(0,T], \tag{2.17}$$

with elements $a^{(2)}_{\mu,\nu}$ given by

$$\begin{aligned}
a^{(2)}_{-1,1} &= \overline{\rho}\gamma i\cdot j/4, & a^{(2)}_{1,-1} &= \overline{\rho}\gamma i\cdot j/4,\\
a^{(2)}_{1,1} &= -\overline{\rho}\gamma i\cdot j/4, & a^{(2)}_{-1,-1} &= -\overline{\rho}\gamma i\cdot j/4,\\
a^{(2)}_{2,0} &= (1-\kappa)ri/4, & a^{(2)}_{-1,0} &= -i^2\cdot jh_y/2+(1+\kappa)ri/4,\\
a^{(2)}_{0,-2} &= \max(c(1-\kappa)/4,0), & a^{(2)}_{0,2} &= -\min(c(1-\kappa)/4,0),\\
a^{(2)}_{1,0} &= -i^2\cdot jh_y/2-1/4(5-3\kappa)ri, & &\\
a^{(2)}_{0,-1} &= -\gamma^2 j/h_y/2-\min(0,c(1-\kappa)/4)-\max(0,c(5-3\kappa)/4), & &\\
a^{(2)}_{0,1} &= -\gamma^2 j/h_y/2+\min(0,c(5-3\kappa)/4)+\max(0,c(1-\kappa)/4), & &\\
a^{(2)}_{0,0} &= r+i^2\cdot jh_y+\gamma^2 j/h_y+3/4(1-\kappa)ri+3|c|(1-\kappa)/4, & &
\end{aligned}$$

$$\text{with}\quad c=-(\alpha(\beta-jh_y)-\lambda\gamma\sqrt{jh_y})/h_y \tag{2.18}$$

from the (s, y) discretization explained above (with discrete boundary conditions and the time discretization). An interesting aspect is that all powers of h_s vanish after the discretization of (2.7), which means that different mesh sizes in the s-direction do not play a role for the multigrid convergence factors. This is not typical for general convection-diffusion-reaction equations.

3. The solution method. We will discuss the multigrid solver in detail in different subsections, starting with a discussion on suitable smoothing methods for the operator (2.7) with discretization (2.17) in the LCP setting.

3.1. Smoother for convection-diffusion type operators. We use a projected version of the well-known lexicographic pointwise Gauss-Seidel method as a smoother. Such an iteration consists of two partial steps. In a first step, a lexicographic pointwise Gauss-Seidel iteration is applied to (2.12) with the equality sign. In the second partial step, the solution is projected, so that the constraint (2.13) is satisfied,

$$\widehat{u}(s_i, y_j, t^*) = \frac{1}{a_{00}^{(2)}} \Bigg\{ f_1(s_i, y_j, t^*) - \sum_{\substack{\mu_1 \in J_s, \mu_1 \le 0 \\ (\mu_1,\mu_2)\ne(0,0)}} \sum_{\substack{\mu_2 \in J_y, \\ \mu_2 \le 0}} a_{\mu_1\mu_2}^{(2)} \overline{u}(s_i + \mu_1, y_j + \mu_2, t^*)$$

$$- \sum_{\substack{\mu_1 \in J_s, \mu_1 \ge 0 \\ (\mu_1,\mu_2)\ne(0,0)}} \sum_{\substack{\mu_2 \in J_y, \\ \mu_2 \ge 0}} a_{\mu_1\mu_2}^{(2)} u(s_i + \mu_1, y_j + \mu_2, t^*) \Bigg\}, \tag{3.1}$$

$$\overline{u}(s_i, y_j, t^*) = \max\{f_{2,h}(s_i, y_j, t^*), \widehat{u}(s_i, y_j, t^*)\} \qquad \forall (s_i, y_j) \in \Omega_h, \tag{3.2}$$

where $\overline{u}$ denotes an unknown after a relaxation and $\widehat{u}$ an unknown after a partial relaxation step, J_s, J_y are the integer index sets related to the nonzero stencil elements in (2.17).

A Gauss-Seidel iteration will not give a good smoother for *convection-dominated* convection-diffusion problems discretized with κ-schemes. Multistage smoothers, defect-correction approaches or KAPPA smoothers [15] are commonly used for second order accurate upwind discretizations in convection-dominated problems. Here, however, for the problem at hand the well-known Gauss-Seidel relaxation methods can be used as smoothers in multigrid, as we will show by Fourier analysis in the following section. As the problem is not really convection-dominated, the above mentioned convergence difficulties do not occur.

3.2. Two-grid Fourier analysis. An important analysis tool for multigrid methods is Fourier analysis, see, for example [2], [19], [21]. It is, in fact, the main multigrid analysis possibility for nonsymmetric problems. We will perform a Fourier two-grid analysis to study the properties of a smoothing method and of the other multigrid components in a two-grid method. The error $v_h^m = u_h^m - u_h$ is transformed by the $(m+1)$-st two-grid cycle as follows:

$$v_h^{m+1} = M_h^H v_h^m, \; M_h^H = S_h^{\nu_2} C_h^H S_h^{\nu_1}; \; C_h^H = I_h - I_H^h (L_H)^{-1} I_h^H L_h \;. \tag{3.3}$$

The spectral radius $\rho_h(M_h^H)$ of the linear two-grid operator M_h^H gives an indication of the asymptotic rate of the multigrid convergence.

On a grid $G_h := \{\boldsymbol{x} = (k_s h, k_y h) : k_s, k_y \in \mathbb{Z}\}$, we consider functions that are linear combinations of the Fourier components

$$\varphi_h(\boldsymbol{\theta}, \boldsymbol{x}) = e^{i\boldsymbol{k\theta}} = e^{i(k_s\theta_s + k_y\theta_y)}$$

ETNA
Kent State University
etna@mcs.kent.edu

with $\boldsymbol{x} \in G_h$, $\boldsymbol{k} = (k_s, k_y) \in \mathbb{Z}^2$ and frequencies $\boldsymbol{\theta} = (\theta_s, \theta_y) \in \mathbb{R}^2$.

The Fourier space $\boldsymbol{\varepsilon}^h = \text{span}\{e^{i\boldsymbol{k}\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta = (-\pi, \pi]^2\}$ contains any infinite grid function on G_h [19]. The discrete solution u_h, the current approximation u_h^m and the error v_h^m (3.3) can be represented as linear combinations of the basis functions $e^{i\boldsymbol{k}\boldsymbol{\theta}} \in \boldsymbol{\varepsilon}^h$.

$\boldsymbol{\varepsilon}^h$ can be divided into four-dimensional sub-spaces, *the harmonics* (see Figure 3.1):

$$\begin{aligned}(3.4)\quad \boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}} &= \text{span}\{\varphi(\boldsymbol{\theta}^{\alpha_s\alpha_y}, \boldsymbol{x}) = e^{i\boldsymbol{k}\boldsymbol{\theta}^{\alpha_s\alpha_y}}; \alpha_s, \alpha_y \in \{0,1\}\},\ \boldsymbol{x} \in G_h,\\ &\boldsymbol{\theta}^{00} \in \Theta^{00} := (-\pi/2, \pi/2]^2,\ \boldsymbol{\theta}^{\alpha_s\alpha_y} := (\theta_s - \alpha_s\text{sign}(\theta_s)\pi, \theta_y - \alpha_y\text{sign}(\theta_y)\pi)\end{aligned}$$

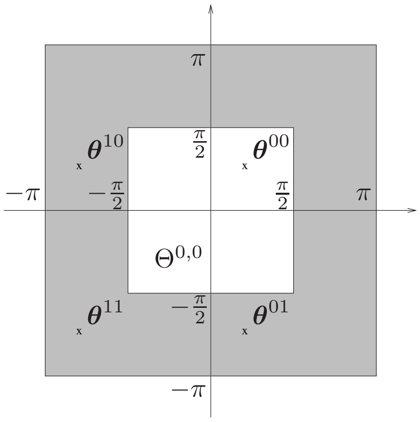

FIG. 3.1. *High (shaded region) and low frequency regions of ε^h with four harmonics*

The 2D coarse grid correction operator C_h^H leaves the 4-dimensional space of harmonics $\boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}}$ (3.4) with an arbitrary $\boldsymbol{\theta} \in \tilde{\Theta}^{00} = \Theta^{00} \setminus \{\boldsymbol{\theta} : \tilde{L}_H(2\boldsymbol{\theta}^{00}) = 0\}$ invariant [19]. The same invariance property is true for the smoothers considered,

$$C_h^H : \boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}} \to \boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}},\quad S_h : \boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}} \to \boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}}\ (\boldsymbol{\theta} \in \tilde{\Theta}^{00}).$$

Hence M_h^H is orthogonally equivalent to a block matrix consisting of 4×4 blocks, which will be denoted by $\tilde{M}_h^H(\boldsymbol{\theta}) := M_h^H\,|_{\boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}}}\quad (\boldsymbol{\theta} \in \tilde{\Theta}^{00})$. We can determine the spectral radius $\rho_h(M_h^H)$ by calculating the spectral radii of 4×4 matrices:

$$(3.5)\qquad \rho_h(M_h^H) = \max_{\boldsymbol{\theta} \in \tilde{\Theta}^{00}} \rho_h(\tilde{M}_h^H(\boldsymbol{\theta}))\ .$$

To obtain the representation of the $4 \times 4-$blocks

$$\tilde{M}_h^H(\boldsymbol{\theta}) = \tilde{S}^{\nu_2}(I - \tilde{P}_h(\tilde{L}_H)^{-1}\tilde{R}_h\tilde{L}_h)\tilde{S}^{\nu_1},$$

the Fourier symbols of the multigrid operators for the harmonics in $\boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}}$ are calculated.

The two-grid convergence properties of stencil (2.17) with the coefficients depending on s and y are analyzed. Fourier analysis is, however, only exact for linear operators with constant (or frozen) coefficients. Therefore, we locally freeze the $s-$ and $y-$terms in front of the derivatives in (2.7) and check the two-grid convergence factors for several relevant values of these quantities. For this purpose, we divide the unit square in 256 intervals $h_s = h_y = 1/256$ and vary i and j in (2.17) by 8 units, $i, j : 0, \ldots, 256$. For each (i, j)-set, we compute $\rho(M_H^h)$ (3.5), which brings many two-grid factors.

In the analysis, we consider the steady equation ($\partial/\partial t^* = 0$), since this represents a worst case for multigrid convergence. Implicit time discretization brings a positive addition to the main diagonal operator element, which is beneficial for the smoothing properties. The parameter set considered in (2.7) is the following:

$$\alpha = 5,\ \beta = 0.16,\ \gamma = 0.9,\ \overline{\rho} = 0.1,\ \lambda = 0,\ r = 0.1, \tag{3.6}$$

We start the analysis of the lexicographical point Gauss-Seidel smoothing method and the coarse grid correction consisting of injection as the restriction operator (the reason for this is explained in Section 3.3.1 in the LCP context), bilinear interpolation as the prolongation and a direct discretization of the PDE on a standard coarsened grid, with $H = 2h$. The number of smoothing iterations is set to 2 here.

Remark: Lexicographic smoothing means that for each index the iteration proceeds from 0 to the maximal value. The direction of the information propagation is, however, opposite. This does not affect our results here. On much finer grids, however, one may expect multigrid convergence difficulties for small y- and large s-values, since then the convection term becomes more dominant. In this situation, an "anti lexicographic" order for smoothing is appropriate.

The computed (s, y)-dependent two-grid convergence factors are graphically presented in the left-hand part of Figure 3.2. Near the domain boundary $s = 0$, $\rho(M_H^h) \approx 1$ is observed which is clearly unsatisfactory. For larger s-values, much better two-grid factors are obtained. The right-hand side picture in Figure 3.2 shows the two-grid convergence results with s-line Gauss-Seidel relaxation (i.e., lines with $j = const.$). In this case, only an isolated large factor close to $(s, y) = (0, 0)$ is observed; most factors are about 0.38. These two-grid

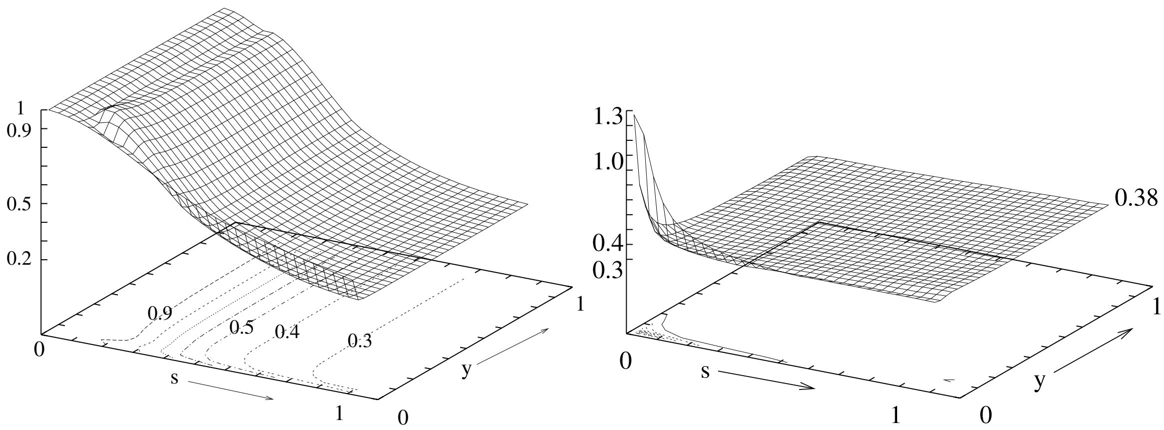

FIG. 3.2. *Fourier two-grid convergence factors $\rho(M_H^h)$ for different (s, y)-values; left: lexicographic point Gauss-Seidel, right: s-line Gauss-Seidel.*

factors are confirmed by numerical multigrid experiments with the components used in the two-grid analysis for equation (2.7) with Dirichlet boundary conditions on different domains Ω. On $\Omega_h = [0, 1]^2$, $h_s = h_y = 1/256$, we find poor multigrid convergence with the lexicographic point Gauss-Seidel smoother and a very satisfactory multigrid convergence factor of 0.35 with the s-line Gauss-Seidel smoother (see Table 3.1). The isolated large value in the latter case can be considered as a boundary effect, due to the lack of ellipticity at the corner point, that is not observed in the actual numerical experiment with Dirichlet conditions at the particular boundary. For other computational domains Ω_h, away from the $s = 0$ axis, we obtain a much improved multigrid convergence with the point Gauss-Seidel smoother, see Table 3.1. We can expect this from the $\rho(M_H^h)$-values in the left-hand side picture in Figure 3.2. Table 3.1 also presents the effect of an overrelaxation parameter $\omega = 1.2$ on

TABLE 3.1
Multigrid convergence of the PDE (2.7) on a 256^2-grid with varying domains and overrelaxation for point- and linewise Gauss-Seidel iteration ($\nu_1 = \nu_2 = 1$).

Smoother:	lex. point Gauss-Seidel				s-line GS
Overrelaxation / Domain Ω:	$[0,1]^2$	$[0.2,1]^2$	$[0.4,1]^2$	$[0.6,1]^2$	$[0,1]^2$
$\omega = 1$	0.994	0.85	0.55	0.30	0.35
$\omega = 1.2$	0.995	0.77	0.42	0.33	0.44

the multigrid convergence, since this can bring improved convergence for several elliptic PDE problems of anisotropic-type [22, 24]. The numerical convergence with the relaxation parameter has also been validated by Fourier analysis.

The reason for concentrating interest so much on the results with the point smoother is the following. Due to the occurring free boundary in the LCP problem, large parts at the left-hand side of a domain will not be processed explicitly by the multigrid smoothing and coarse grid correction. At these parts, the second constraint (2.13) holds with equality sign. For the grid points, where the smoothing will typically be applied for the option pricing problem under consideration, the very satisfactory two-grid convergence factors are found in the left-hand side picture of Figure 3.2 and Table 3.1. The convergence factors in this region are, in fact, very similar for pointwise and linewise smoothing. A problem with linewise smoothing for LCPs is, as explained in [6], that the lines should end at the free boundary for good smoothing and convergence. Therefore a detection mechanism must be incorporated into the smoothing method, which is in [6] the pointwise Gauss-Seidel iteration. For equidistant grids and the LCP problem setting, however, the pointwise smoother is fully satisfactory. From (2.18), we already observed that all powers of h_s vanish in the discretization of (2.7). For the multigrid convergence this means that also on severely stretched but equidistant grids in the s-direction, as we find them here with $s_{\max} = 20$ or $s_{\max} = 200$ in (2.10),(2.11), very similar convergence factors are observed by Fourier analysis and by the numerical experiments as the ones in Figure 3.2 and Table 3.1.

The convergence factors increase drastically, however, if we deal with larger values of h_y and substantially fewer grid points in y- than in s-direction. Figure 3.3 shows the two-grid factors for 32 points in y-direction, $h_y = 1/32$ ($h_s = 1/256$). This case is a limit case for the convergence of the point smoother; more points in y-direction give satisfactory convergence, less points worse convergence. Figure 3.3 also shows that the s-line smoother will give worse convergence in this situation, similar to the point smoother. Domains other than $[0, 1]^2$ will not improve the convergence here, since the worst factors are now found at the right-hand side domain boundary. An alternating line Gauss-Seidel smoother, for which $\rho(M_H^h)$ is presented in Figure 3.3c, brings good convergence. These Fourier analysis results are all confirmed by numerical experiments.

3.3. The Multigrid method for linear complementarity problems. The fundamental idea of multigrid for *nonlinear PDEs* of the form

$$Nu = f \tag{3.7}$$

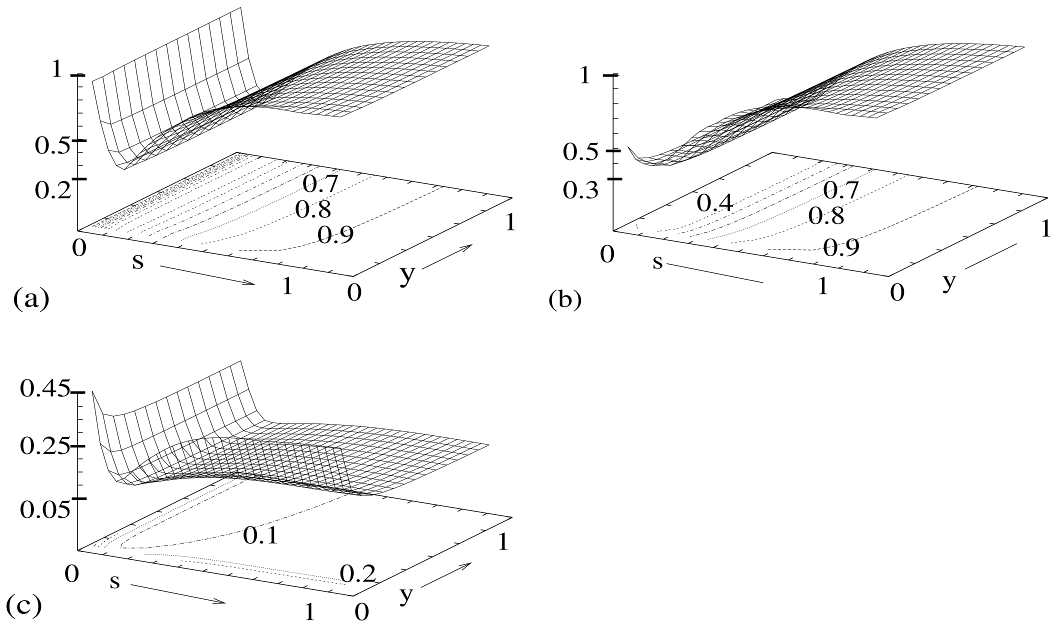

FIG. 3.3. *Fourier two-grid factors for a grid with $h_s = 1/256,\ \ h_y = 1/32$, (a) with point Gauss-Seidel smoother, (b) with s-line Gauss-Seidel, (c) alt. line Gauss-Seidel.*

is the same as that for linear equations. First, the errors of the solution have to be smoothed so that they can be approximated on a coarser grid. Then, a nonlinear analog of the linear defect equation is transferred to the coarse grid. In the nonlinear case, the (exact) defect equation on Ω_h is given by

$$N_h\left(\overline{u}_h^m + v_h^m\right) - N_h\overline{u}_h^m = d_h^m, \tag{3.8}$$

with $\overline{u}_h^m$ the approximation of the solution after relaxation in the mth multigrid cycle, v_h^m the error and d_h^m the corresponding defect. This equation is approximated on Ω_H by

$$N_H\left(\overline{u}_H^m + \widehat{v}_H^m\right) - N_H\overline{u}_H^m = d_H^m, \tag{3.9}$$

where $\widehat{v}_H^m$ is the coarse grid approximation of the error v_h^m. Not only is the defect d_h^m transferred to the coarse grid by some restriction operator I_h^H, but also the relaxed approximation $\overline{u}_h^m$ itself by a restriction operator $\widehat{I}_h^H$. On the coarse grid Ω_H, we deal with the problem $N_H w_H = f_H$, where $w_H = \overline{u}_H^m + \widehat{v}_H^m$ and where the right-hand side f_H is defined by $f_H := I_h^H(f_h - N_h\overline{u}_h^m) + N_H\widehat{I}_h^H\overline{u}_h^m$.

The *coarse grid corrections* $\widehat{v}_H$ are interpolated back to the fine grid, where the fine grid errors are smoothed again. The generalization from two grids to multigrid is done recursively. If N_h and N_H are linear operators, the FAS method is equivalent to the (linear) multigrid correction scheme. It was shown in [3], that a variant of FAS, the projected full approximation scheme, PFAS can be used to solve linear complementarity problems of the form (2.12),(2.13),(2.14). PFAS is based on the projected Gauss-Seidel smoother (3.1),(3.2). We now explain coarse grid correction parts for the problem (2.12),(2.13), (2.14) in some more detail.

ETNA
Kent State University
etna@mcs.kent.edu

3.3.1. The LCP coarse grid correction. Suppose that the error $v_h^m := u_h - \overline{u}_h^m$ is smooth after relaxation. The following LCP holds for v_h^m:

$$\begin{aligned} L_h v_h^m &\le d_h^m, \quad \mathbf{x} \in \Omega,\\ v_h^m + \overline{u}_h^m &\ge f_{2,h}, \quad \mathbf{x} \in \Omega,\\ (v_h^m + \overline{u}_h^m - f_{2,h})(L_h v_h^m - d_h^m) &= 0, \quad \mathbf{x} \in \Omega, \end{aligned}$$

with defect: $d_h^m = f_{1,h} - L_h \overline{u}_h^m$. A smooth error v_h^m can be approximated on a coarse grid without any essential loss of information. The LCP coarse grid equation for the coarse grid approximation of the error $\widehat{v}_H^m$ is therefore defined in PFAS by:

$$\begin{aligned} L_H \widehat{v}_H^m &\le I_h^H d_h^m,\\ \widehat{v}_H^m + \widehat{I}_h^H \overline{u}_h^m &\ge f_{2,H},\\ (\widehat{v}_H^m + \widehat{I}_h^H \overline{u}_h^m - f_{2,H})(L_H \widehat{v}_H^m - I_h^H d_h^m) &= 0. \end{aligned}$$

Since the problem is nonlinear and we are solving inequalities, we solve for a full approximation $u_H^m := \widehat{v}_H^m + I_h^H \overline{u}_h^m$ but interpolate only $\widehat{v}_H^m$ back to the fine grid.

A relevant difference between multigrid methods for equations and inequalities follows from the requirement that, in the case of a converged solution on a fine grid $\overline{u}_h^m \equiv u_h$, corrections from the coarse grid equation should be zero. Then, we need

$$I_H^h \widehat{v}_H^m = I_H^h (u_H^m - \widehat{I}_h^H u_h^m) = 0,$$

leading to $u_H^m = \widehat{I}_h^H u_h^m$ (assuming operator I_H^h keeps nonzero quantities nonzero). If for a fine grid LCP with a converged solution we consider the coarse grid correction, it leads to the following requirements [3] on the restriction operators,

$$\begin{aligned} I_h^H (f_{1,h} - L_h u_h^m) &\ge 0,\\ \widehat{I}_h^H u_h^m &\ge f_{2,H},\\ (\widehat{I}_h^H u_h^m - f_{2,H})^T I_h^H (f_{1,h} - L_h u_h^m) &= 0. \end{aligned} \tag{3.10}$$

For equations, since $f_{1,h} - L_h u_h^m \equiv 0$ in this situation, these requirements are satisfied for any reasonable choice of I_h^H and $\widehat{I}_h^H$. For LCPs, we choose for both restriction operators straight injection in order to satisfy the requirements (3.10). The injection operator is in a certain sense "constraint preserving". (These requirements do not, however, prevent us from using any residual transfer in the interior, away from the free boundary).

The prolongation operator I_H^h, bilinear interpolation, is applied only for unknowns on the "inactive" points, as this resulted in the best convergence in [3] (for detail, see M6 in section 5 of [3]):

$$\begin{aligned} &u_h^m \Leftarrow \overline{u}_h^m + I_H^h \widehat{v}_H^m \quad \text{if} \quad \overline{u}_h^m > f_{2,h},\\ &u_h^m \Leftarrow \overline{u}_h^m \quad \text{elsewhere} \quad (\overline{u}_h^m = f_{2,h}). \end{aligned}$$

This combination of restriction and prolongation does not satisfy the well-known rule [21, 19] on the orders of the transfer operators, for example, for PDEs with second derivatives. Moreover, it is known that problems with Neumann boundary conditions, as they appear for the option pricing problem in the stochastic volatility setting, converge rapidly with so-called modified full weighting operators at the boundary [19]. These boundary transfer operators are not used here, because of the requirements (3.10). From these points of view, the coarse

ETNA
Kent State University
etna@mcs.kent.edu

grid correction part might be not as powerful as it is commonly for multigrid for elliptic PDEs. Extra investment in smoothing (more iterations, line or ILU smoothers, etc.), or other approaches for convergence improvement might be necessary for fast and robust convergence.

The robustness of PFAS is also discussed in [13], where a convergence proof for a PFAS competitor, the so-called monotone multigrid method is presented. The example on which this discussion is based is not, however, a decisive example for possible robustness problems of PFAS for the LCPs as we will show in Section 5.1. Notwithstanding this discussion, we present an improvement for the robustness of PFAS by using it as a "preconditioner" for an acceleration method in Section 4.

4. Acceleration of Multigrid by Iterate Recombination (Multigrid as a Preconditioner). In this section, we discuss multigrid used as a preconditioner for Krylov subspace methods. From the multigrid point of view, multigrid as a preconditioner can also be interpreted as an acceleration of multigrid by iterate recombination. This interpretation easily allows generalizations, for example, to nonlinear problems and also to linear complementarity problems.

Let u_h^0 be an initial approximation for solving $L_h u_h = f_h$ and $d_h^0 = f_h - L_h u_h^0$ its defect. The *Krylov subspace* K_h^m is defined by $K_h^m := \text{span}[d_h^0, L_h d_h^0, \ldots, L_h^{m-1} d_h^0]$. This subspace can also be represented by

$$\begin{aligned} K_h^m &= \text{span}[u_h^1 - u_h^0, u_h^2 - u_h^1, \ldots, u_h^m - u_h^{m-1}] \\ &= \text{span}[u_h^0 - u_h^m, u_h^1 - u_h^m, \ldots, u_h^{m-1} - u_h^m] \ , \end{aligned} \tag{4.1}$$

where the $u_h^m = (I_h - L_h)u_h^{m-1} + f_h$ are the iterates from the well-known Richardson's iteration. These representations are easily obtained by induction using

$$u_h^1 - u_h^0 = d_h^0 \ , \quad u_h^{i+1} - u_h^i = (I_h - L_h)(u_h^i - u_h^{i-1}) \ .$$

The Krylov subspace approximation

$$u_{h,\text{acc}}^m \in u_h^0 + K_h^m = u_h^0 + \text{span}[d_h^0, L_h d_h^0, \ldots, L_h^{m-1} d_h^0]$$

is then characterized by finding approximations $u_{h,\text{acc}}^m$ for $m = 1, 2, \ldots$ *with minimal defect* in a suitable norm. The various Krylov subspace methods differ in the way the minimization is carried out. If we use the $||\cdot||_2$ norm for minimization, we obtain GMRES [18]. In the same way as the classical single grid iterative methods can be used as preconditioners, it is also possible to use a linear multigrid method (algebraic multigrid, for example) as a preconditioner.

Multigrid acceleration by iterate recombination and multigrid preconditioning lead to very similar algorithms: We search for an improved solution based on the second representation of the subspace in (4.1). In order to find an improved approximation $u_{h,\text{acc}}^m$, we consider a linear combination of the $\widetilde{m}+1$ latest approximations u_h^{m-i}, $i = 0, \cdots, \widetilde{m}$ (assuming $m \ge \widetilde{m}$),

$$u_{h,\text{acc}}^m = u_h^m + \sum_{i=1}^{\widetilde{m}} \overline{\alpha}_i (u_h^{m-i} - u_h^m). \tag{4.2}$$

For linear equations, the corresponding defect, $d_{h,\text{acc}}^m = f_h - L_h u_{h,\text{acc}}^m$, is given by

$$d_{h,\text{acc}}^m = d_h^m + \sum_{i=1}^{\widetilde{m}} \overline{\alpha}_i (d_h^{m-i} - d_h^m), \tag{4.3}$$

where $d_h^{m-i} = f_h - L_h u_h^{m-i}$. In order to obtain an improved approximation $u_{h,\text{acc}}^m$, the parameters $\overline{\alpha}_i$ are determined in such a way that the defect $d_{h,\text{acc}}^m$ is minimized, for example with respect to the l_2-norm $||\cdot||_2$. This is a classical defect minimization problem. Here, we solve the system of normal equations. The work for solving the minimization problem is small. In general, for such minimization problems, it may happen that one has to deal with an extremely ill-conditioned matrix. However, in this setting $\widetilde{m}$ can often be chosen small, for example, 5 or less. Such small matrices are usually still satisfactorily conditioned, so that the matrix can be solved directly. In the linear case, it is not important whether the accelerated unknowns are chosen for u (4.2) or the multigrid iterates. It merely implies different weights $\overline{\alpha}_i$ in the recombination process. We always use the accelerated iterates. This saves some storage and it is beneficial in the nonlinear case discussed below. The resulting iterative method is outlined in Figure 4.1, where the current approximation u_h^m is replaced by

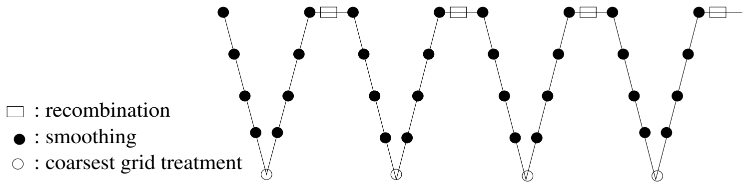

FIG. 4.1. *Recombination of multigrid iterates*

$u_{h,\text{acc}}^m$. With this approximation replacement, the next multigrid cycle is performed resulting in a new iterate u_h^{m+1}. The recombination (4.2) is again carried out with the latest iterates $u_h^{m+1-i}, \; i = 0, \cdots, \widetilde{m}$ and so on.

It is possible to generalize the idea of iterate recombination to nonlinear situations, where a nonlinear multigrid method, such as FAS is used. In this case, the defect relation (4.3) does not hold exactly, due to the nonlinearity. u_h^m is then only substituted by $u_{h,\text{acc}}^m$, if the defect $d_{h,\text{acc}}^m$ of $u_{h,\text{acc}}^m$ is not much larger than the defect d_h^m (as described below).

Here, we generalize the defect minimization approach to linear complementarity problems. This is done very similarly to the generalization to nonlinear PDEs. PFAS will be used as the underlying multigrid method, whose iterates are recombined. The defect minimization for obtaining an improved solution is based on *equation* (1.3) in the system of linear complementarity,

$$d_{h,\text{acc}}^m = (u_{h,\text{acc}}^m - f_{2,h})(Lu_{h,\text{acc}}^m - f_{1,h}).$$

We obtain solutions with an improved defect that satisfy constraint (1.2), as all recombined iterates do. The points where $u = f_2$ in the linear combination are the 'intersection' of the sets of these contact points amongst all the vectors participating in the linear combination with nonzero coefficients. The linear combination cannot increase the number of contact points. The convergence of the method to a solution that also satisfies (1.1) is still based on the convergence of PSOR in [7]. If all the vectors participating in the linear combination satisfy (1.1) at any particular point, then so does any linear combination (by the linearity of L).

The resulting method can be interpreted as a projected Gauss-Seidel method, accelerated in terms of the coarse grid correction by a projected FAS multigrid method and accelerated further in terms of an outer iteration by a recombination technique.

Due to the nonlinearity, criteria for replacing the PFAS solution u_h^m by the accelerated solution $u_{h,\text{acc}}^m$ are needed. We use the following selection criteria:

1. The norm of $d^m_{h,\text{acc}}$ is not larger than d^m_h and the $\widetilde{m}$ intermediate defects d^{m-i}_h :

$$\|d^m_{h,\text{acc}}\|_2 < \min_{1\le i\le\widetilde{m}} (\|d^{m-i}_h\|_2, \|d^m_h\|_2) \tag{4.4}$$

2. $u^m_{h,\text{acc}}$ is not too close to any of the intermediate solutions unless a considerable decrease of the defect norm is achieved.

Criterion 2 is necessary to prevent stagnation in the convergence. The same criteria are used for restarting the Krylov subspace, as is presented in [20]. Restarting takes place if the criteria are not satisfied in two consecutive iterations.

5. Numerical Results. In this section, we present some numerical experiments. We measure the size of the defect of equation (2.14) in the LCP system. In addition to this defect, we also checked the difference between the previous and a current solution, as in [3]. Both measures show a very similar convergence history. Indicated in the Tables 5.3 and 5.4 below is the average number of cycles per time step to reduce the defect in (2.14) in the infinity norm by five orders of magnitude,

$$\frac{|d^m_h(s,y,t)|_\infty}{|d^0_h(s,y,t)|_\infty} \le 10^{-5} \quad . \tag{5.1}$$

This quantity is also presented if the recombination is applied (in that case, the defect after the recombination is given by d^m_h).

As an appetizer, before we consider the option pricing problem, we briefly discuss an LCP from [13] based on the Laplace operator. In [13], a very poor PFAS convergence was presented. We will consider this problem with PFAS in more detail also including the recombination technique and show a very satisfactory convergence.

5.1. Linear complementarity model problem based on Laplace operator. In [13], the following optimization problem modelling elasto-plastic torsion of a cylindrical bar in a model region $\Omega = (0,1) \times (0,1)$ is considered,

$$\begin{aligned} u_{ss} + u_{yy} \le{}& 2C, && \mathbf{x} \in \Omega, \\ u \le{}& \mathrm{d}(\mathbf{x}, \partial\Omega), && \mathbf{x} \in \Omega, \\ (u - \mathrm{d}(\mathbf{x},\partial\Omega))(u_{ss} + u_{yy} - 2C) ={}& 0, && \mathbf{x} \in \Omega, \\ \text{with boundary cond.}\quad u ={}& 0, && \mathbf{x} \in \partial\Omega, \end{aligned} \tag{5.2}$$

where the $\mathrm{d}(\mathbf{x}, \partial\Omega)$-operator measures the distance from a grid point $\mathbf{x} = (s, y)$ to the domain boundary $\partial\Omega$.

With parameter $C = 10$, the problem (5.2) was considered to be very difficult in [13]. Figure 5.1 presents level curves of the solution (left-hand side figure) and the small region of "inactive" points, where $u_{ss} + u_{yy} = 2C$ is valid (right-hand side picture). This region, whose size depends on parameter C, represents the plastic region, whereas the active points, where the second constraint with equality sign is valid, represents the elastic region (see [17] for details). In [13] the V-cycle PFAS convergence with only one lexicographic projected Gauss-Seidel smoothing iteration has been presented. This convergence is not satisfactory as is confirmed by the upper curve at the top of Figure 5.2, where the same multigrid components as in [13] are used. This result is, however, not a real surprise. Even for linear elliptic PDEs, a multigrid V-cycle with one smoothing iteration, injection as the restriction operator and a direct discretization of the PDE on coarse grids must be considered with caution [19]. An injection-based coarse grid correction must be supplied with more smoothing iterations and with more coarse grid processing as confirmed in the upper picture in Figure 5.2: An F(2,0)-cycle with only slightly more work on coarser grids leads to very fast multigrid convergence.

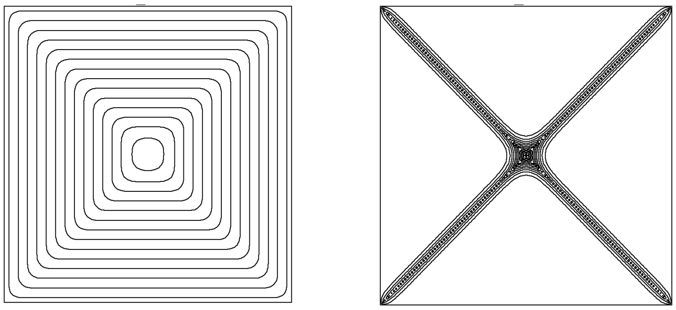

FIG. 5.1. *Level curves and inactive set for $C = 10$.*

These results are computed on a 256^2 grid. By only looking at the V(1,0)-cycle convergence, an incorrect impression of the quality of PFAS might easily lead to unnecessarily considering other solution methods.

In the lower picture in Figure 5.2, we present the convergence of accelerated multigrid, where 3 iterates are recombined after each multigrid cycle. A faster and more robust convergence is obtained in this case. Table 5.1 finally presents average reduction factors for different grid sizes and multigrid cycles. Especially the F-cycles show a very satisfactory convergence, even without the recombination technique.

TABLE 5.1
Multigrid and accelerated multigrid convergence ($\widetilde{m} = 3$) for the 2D model LCP and various multigrid cycles.

Grid	Cycles				
	V(2,0)	V(2,0) + acc.	F(2,0)	F(2,0) + acc.	F(1,1)
128^2	0.76	0.27	0.31	0.20	0.33
256^2	0.97	0.31	0.37	0.21	0.35
512^2	0.99	0.33	0.41	0.25	0.38

5.2. An American-style option. With the components introduced in Section 3, we now solve an LCP problem from American-style options with stochastic volatility. The PFAS multigrid method is based on the F(2,2)-cycle with pointwise lexicographic Gauss-Seidel smoothing.

The parameter set considered is again

$$\alpha = 5,\ \beta = 0.16,\ \gamma = 0.9,\ \overline{\rho} = 0.1,\ \lambda = 0,\ r = 0.1 \tag{5.3}$$

and exercise price $E = 10$. The expiration date is set to $T = 0.25$. For this set of parameters reference solutions for the American put option prices are presented in [5] and [25]. First, we discuss the accuracy of the discretization by comparing our numerical results on different grid sizes with the reference results.

A truncated domain $\Omega = [0, s_{\text{max}}] \times [0, y_{\text{max}}]$ is used with $s_{\text{max}} = 20$ and $y_{\text{max}} = 1$. The grids consist of 256 cells in the s-direction; in the y-direction four sizes are considered with $32,\ 64,\ 128$ and 256 cells. Table 5.2 compares the solution u_h obtained at $y = 0.0625$ and $y = 0.25$ respectively with the solutions in [5] and [25]. The results agree very well, especially with those in [25], in which the difference to the other reference result was discussed. This difference might be due to the stretched grid considered in [5] or due to the interpolation technique and the first order accurate discretization employed there in some parts of the domain.

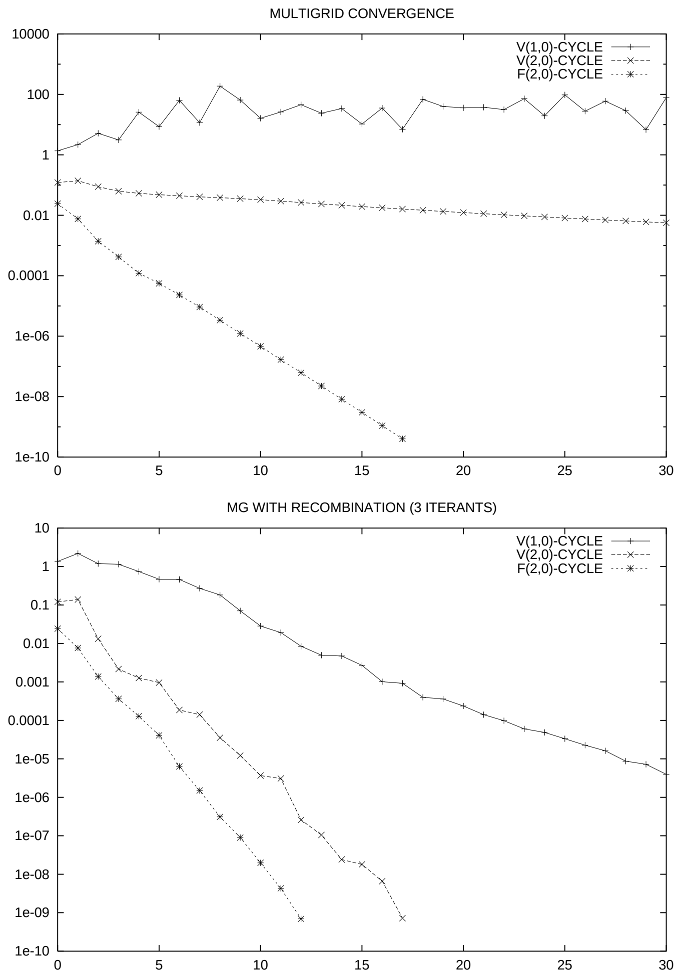

FIG. 5.2. *Multigrid convergence with different cycles (upper picture) and convergence of multigrid with acceleration (lower picture), 3 recombined iterates,* 256^2*-grid.*

Figure 5.3 presents the moving free boundary for this problem in pictures for the t values $0, 0.05, 0.1, 0.15, 0.2$. On the left-hand side of this free boundary, the active points are found, i.e., the constraint (2.13) with equality sign is valid.

Table 5.3 presents the average number of multigrid iterations necessary to reduce the initial defect each time step by 5 orders of magnitude on the different grid sizes. In the left-hand side part of the table, the multigrid convergence without the recombination is presented. It can be seen that for problem (2.7) with parameter set (5.3) fewer grid points in the y-direction lead to worse multigrid convergence as discussed in the Fourier analysis section. The overrelaxation parameter helps in this case (see the 256×32 results) to improve the convergence: instead of 46.5 iterations per time step, 27.6 are needed with $\omega = 1.4$. The right-hand side part of Table 5.3 shows the results with the same multigrid algorithm accelerated by a recombination of the 3 previous iterates. A much better convergence is obtained now, especially in the case with only 32 grid points in the y-direction, whereas the convergence improvement on the 256×256 grid is not impressive. This confirms the general impression that (almost)

ETNA
Kent State University
etna@mcs.kent.edu

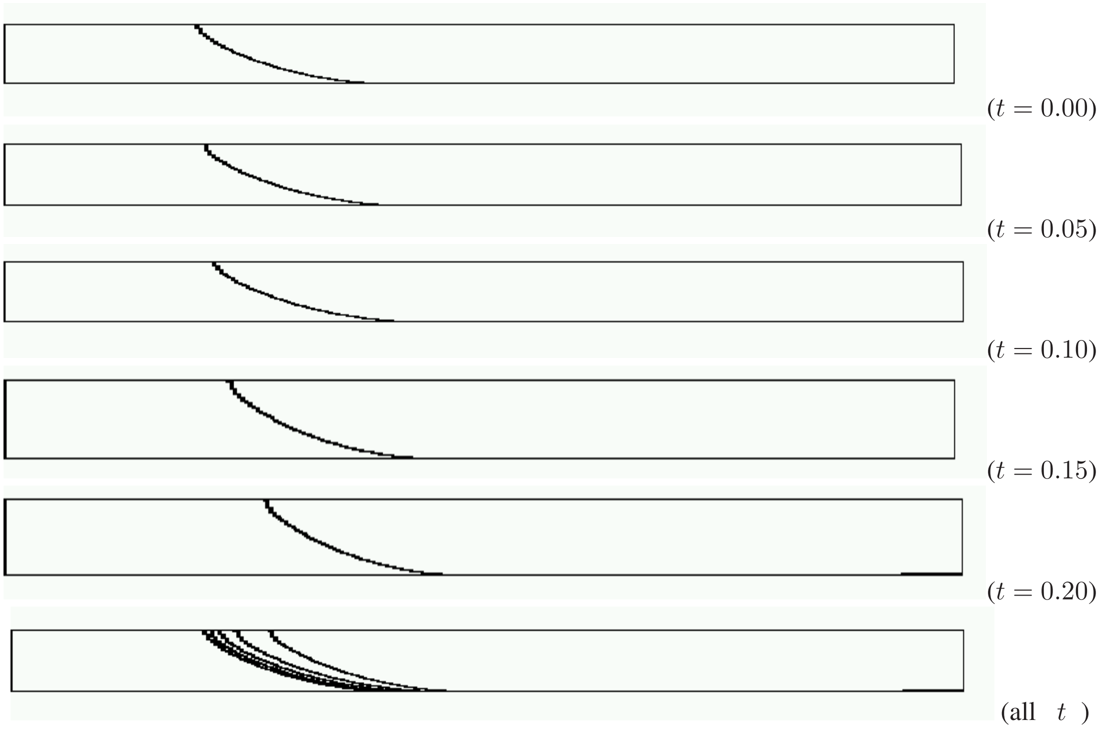

FIG. 5.3. *The moving free boundary for the American put option.*

optimal multigrid methods are difficult to improve by Krylov subspace acceleration, whereas in cases in which one of the multigrid components (here it is the smoother) is not optimal, the acceleration by the recombination technique can give a more robust and faster convergence. It is also shown that the combination of the best overrelaxation in the left-hand side of Table 5.3 with the recombination technique does not result in the best algorithm in the right-hand side. This has also been observed in [22], where a multigrid method with pointwise relaxation used as a preconditioner is analyzed by Fourier analysis for model anisotropic equations. The best choice here seems to be an overrelaxation parameter of 1.1 combined with the recombination technique.

TABLE 5.2
American put option values with stochastic volatility for $y = 0.0625$ and $y = 0.25$ compared with reference values on different grid sizes at $t = 0$, $E = 10$.

$y = 0.0625$	Asset price				
Grid	8	9	10	11	12
256×32	2.00	1.104	0.494	0.198	0.0769
256×64	2.00	1.106	0.508	0.206	0.0793
256×128	2.00	1.107	0.515	0.210	0.0804
256×256	2.00	1.107	0.518	0.212	0.0809
256×256 (smaller time step)	2.00	1.107	0.517	0.212	0.0815
ref [25]	2.00	1.108	0.520	0.214	0.0821
ref [5]	2.00	1.108	0.532	0.226	0.0907

$y = 0.25$	Asset price				
Grid	8	9	10	11	12
256×32	2.079	1.334	0.796	0.449	0.242
256×64	2.079	1.335	0.797	0.449	0.243
256×128	2.079	1.335	0.797	0.449	0.243
256×256	2.079	1.335	0.797	0.449	0.243
256×256 (smaller time step)	2.079	1.334	0.796	0.449	0.243
ref [25]	2.078	1.334	0.796	0.448	0.243
ref [5]	2.073	1.329	0.799	0.454	0.250

TABLE 5.3
Average number of cycles needed for 5 orders of defect reduction over 20 time steps on different stretched meshes.

Grid	F(2,2)-cycle				F(2,2)- with recombination			
	$\omega = 1$	$\omega = 1.1$	$\omega = 1.2$	$\omega = 1.4$	$\omega = 1$	$\omega = 1.1$	$\omega = 1.2$	$\omega = 1.4$
256^2	5.3	5.1	6.0	8.0	5.0	5.0	5.6	7.5
256×128	5.9	6.6	6.6	8.1	5.4	5.4	5.5	6.6
256×64	16.7	15.5	14.5	11.9	7.6	7.8	7.7	8.3
256×32	46.5	41.5	36.7	27.6	12.2	11.5	11.8	14.0

Finally, Table 5.4 presents the accelerated multigrid convergence with $\widetilde{m} = 3$ for the parameter set (5.3) with $E = 100$ and, therefore, with a 10 times larger domain, $s_{\max} = 200$. As expected from the Fourier analysis, the convergence results are very similar to the results in Table 5.3.

TABLE 5.4
Average number of cycles needed for 5 orders of defect reduction over 20 time steps on different stretched meshes.

Method	Grid			
	256^2	256×128	256×64	256×32
F(2,2)- with recomb., $\widetilde{m} = 3, \omega = 1.1$	5.3	5.9	9.4	15.1

Remark: The method described here also works well for the call option with dividend payment. Although the free boundary is in a completely different part of the computational domain then, for the call option the solution is simply zero in the problematic part. It is, of course, still beneficial for the convergence to cut off the domain near $s = 0$ as it was shown in Table 3.1.

6. Conclusions. In this paper we have presented a multigrid solution method for linear complementarity problems. The method is based on the projected full approximation scheme. It is combined with a recombination of iterates convergence acceleration technique. By providing much detail about the different parts of the solution method, we hope to give insight into its expected convergence. The smoother and the other components in the multigrid method have been analyzed by means of Fourier analysis for the main discrete operator appearing in the LCP. The Gauss-Seidel lexicographic point smoother was chosen here.

With the acceleration technique, fast convergence is obtained for an option pricing problem on grids with different grid sizes. The error of the discretization is determined by comparison with reference solutions.

The smoother proposed in [6] consisting of a point SOR method for the free boundary detection followed by a modified line Gauss-Seidel smoother could also be employed here. The Fourier analysis results in Figure 3.3 show, however, that only a multigrid method based on an alternating line smoother is robust. With such a more involved smoother, a recombination technique is not necessary for fast convergence.

Although we only considered one parameter set (3.6) in this paper, we used the Fourier analysis (not shown here) to investigate the parameter range for which the conclusions remain valid. The sensitivity with respect to variations of the (important) parameter r is, for example in its relevant range, not at all significant. The same is true for the other parameters.

By retaining a point smoother for this 2D problem, we constructed a fast and cheap solver, that can serve as a basis for treating other, higher dimensional, problems in the option pricing context.

Acknowledgement The author gratefully acknowledges R. Wienands for providing his Fourier analysis software package $LFA00\_2D$ and his assistance in using it.

REFERENCES

[1] C.A. Ball and A. Roma, Stochastic volatility option pricing. *J. Financial Quantitative Anal.,* 29: 589-607, 1994.
[2] A. Brandt, Multi–level adaptive solutions to boundary–value problems. *Math. Comp.,* 31: 333-390, 1977.
[3] A. Brandt and C.W. Cryer, Multigrid algorithms for the solution of linear complementarity problems arising from free boundary problems, *SIAM J. Sci Comput.,* 4: 655-684, 1983.
[4] A. Brandt and V. Mikulinsky, On recombining iterants in multigrid algorithms and problems with small islands. *SIAM J. Sci. Comput.,* 16: 20-28, 1995.

[5] N. Clarke and K. Parrot, *The multigrid solution of two factor American put options.* Res. Report 96-16, Oxford Comp. Lab, Oxford, 1996.
[6] N. Clarke and K. Parrot, Multigrid for American option pricing with stochastic volatility. *Appl. Math. Finance,* 6: 177-197, 1999.
[7] C.W. Cryer, The solution of a quadratic programming problem using systematic overrelaxation, *SIAM J. Control* 9: 385-392, 1971.
[8] C.W. Gear, *Numerical initial value problemsfor ordinary differential equations.* Prentice-Hall, Englewood Cliffs NJ, 1971.
[9] E. Hairer and G. Wanner, *Solving ordinary differential equations II, Stiff and differential-algebraic problems.* Second Ed. Springer, Berlin, 1996.
[10] J.C. Hull, *Option, futures and other derivatives.* Prentice-Hall, Englewood Cliffs NJ, 1997.
[11] J.C. Hull and A. White, The pricing of options on assets with stochastic volatility. *J. of Finance* 42: 281-300, 1987.
[12] W. Hundsdorfer, Partially implicit BDF2 blends for convection dominated flows. *SIAM J. Numer. Anal.,* 38: 1763–1783, 2001.
[13] R. Kornhuber, Monotone multigrid methods for elliptic variational inequalities I, *Num. Math.* 69: 167-184, 1994.
[14] B. van Leer, Upwind-difference methods for aerodynamic problems governed by the Euler equations. *In:* B. Enquist, S. Osher, R. Somerville (eds.), *Large scale computations in fluid mechanics.* Lectures in Appl. Math. 22 II, Americ. Math. Soc., Providence, R.I., 327-336, 1985.
[15] C.W. Oosterlee, F.J. Gaspar, T. Washio and R. Wienands, Multigrid line smoothers for higher order upwind discretizations of convection-dominated problems. *J. Comp. Phys.,* 139: 274-307, 1998.
[16] C.W. Oosterlee and T. Washio, Krylov subspace acceleration of nonlinear multigrid with application to recircul ating flows. *SIAM J. Sci. Comput.* 21: 1670-1690, 2000.
[17] J-F. Rodrigues, *Obstacle problems in mathematical physics.* North-Holland, Math. Studies 134, Amsterdam, 1987.
[18] Y. Saad and M.H. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Comput.,* 7: 856-869, 1986.
[19] U. Trottenberg, C.W. Oosterlee and A. Schüller, *Multigrid*, Academic Press, London, 2000.
[20] T. Washio and C.W. Oosterlee, Krylov subspace acceleration for nonlinear multigrid schemes. *Electr. Trans. on Num. Analysis* 6: 271-290, 1997.
[21] P. Wesseling, *An introduction to multigrid methods*, Wiley Chicester, 1992.
[22] R. Wienands, C.W. Oosterlee and T. Washio, Fourier analysis of GMRES(m) preconditioned by multigrid. *SIAM J. Sci. Comput.,* 22: 582-603, 2000.
[23] P. Wilmott, J. Dewynne and S. Howison, *Option pricing*. Oxford Financial Press, 1993.
[24] I. Yavneh, On red-black SOR smoothing in multigrid. *SIAM J. Sci. Comput.,* 17:180-192, 1996.
[25] R. Zvan, P.A. Forsyth and K.R. Vetzal, Penalty methods for American options with stochastic volatility. *J. Comp. Appl. Math.* 91: 199-218, 1998.
[26] R. Zvan, P.A. Forsyth and K.R. Vetzal, Robust numerical methods for PDE models of Asian options. *J. Comput. Finance* 1: 39-78, 1998.