NUMERICAL COMPUTATION OF THE VALUE FUNCTION OF OPTIMALLY CONTROLLED STOCHASTIC SWITCHING PROCESSES BY MULTI-GRID TECHNIQUES

Martina Bloß and Ronald H.W. Hoppe
Technische Universität Berlin, Fachbereich Mathematik
D-1000 Berlin 12 (F.R.G.)

ABSTRACT

By the dynamic programming principle the value function of an optimally controlled stochastic switching process can be shown to satisfy a boundary value problem for a fully nonlinear second-order elliptic differential equation of Hamilton-Jacobi-Bellman (HJB-) type. For the numerical solution of that HJB-equation we present a multi-grid algorithm whose main features are the use of nonlinear Gauss-Seidel iteration in the smoothing process and an adaptive local choice of prolongations and restrictions in the coarse-to-fine and fine-to-coarse transfers. Local convergence is proved by combining nonlinear multi-grid convergence theory and elementary subdifferential calculus. The efficiency of the algorithm is demonstrated for optimal advertising in stochastic dynamic sales response models of Vidale-Wolfe type.

1. INTRODUCTION

We consider a stochastic system operating in m different regimes under state constraints in the form of either an exit problem or reflecting boundary conditions. In the first case the regimes can be described by the diffusion processes

$$dy_x(t) = b^{\mu}(y_x(t))dt + \sigma^{\mu}(y_x(t))\,dw(t)\,, \quad 1 \le \mu \le m \tag{1.1a}$$

$$y_x(0) = x \tag{1.1b}$$

where $x \in \Omega$, Ω is a bounded smooth domain in Euclidean space $\mathbb{R}^d$, $d \in \mathbb{N}$, with boundary $\Gamma = \partial\Omega$, and w is a standard d-dimensional Wiener process. The drift $b^{\mu} = (b_1^{\mu},\dots,b_d^{\mu})^T$ and the diffusion $\sigma^{\mu} = (\sigma_{ij}^{\mu})_{i,j=1}^{d}$, $1 \le \mu \le m$, are assumed to be sufficiently smooth functions on $\mathbb{R}^d$. In the second case (1.1a) has to be replaced by the reflected diffusion process

$$dy_x(t) = b^{\mu}(y_x(t))\,dt + \sigma^{\mu}(y_x(t))\,dw(t) - \tag{1.1a$'$}$$
$$- \chi_{\Gamma}(y_x(t))\,\gamma^{\mu}(y_x(t))\,d\xi(t)\,, \quad 1 \le \mu \le m$$

where χ_{Γ} is the characteristic function of $\Gamma = \partial\Omega$, ξ is an increasing continuous adapted process and $\gamma^{\mu} = (\gamma_1^{\mu}, \dots, \gamma_d^{\mu})^T$, $1 \le \mu \le m$, is supposed to be continuous and bounded on $\mathbb{R}^d$ satisfying $\gamma^{\mu}(x)\cdot n(x) \ge \delta > 0$ for all $x \in \Gamma$ and all $1 \le \mu \le m$ where $n(x)$ denotes the unit outward normal in $x \in \Gamma$. Then, given smooth running costs $f^{\mu} = f^{\mu}(x,t)$, $x \in \Omega$, $t \ge 0$, $1 \le \mu \le m$, and nonnegative discount factors $c^{\mu} = c^{\mu}(x)$, $x \in \Omega$, $1 \le \mu \le m$, the control objective is to find an optimal switching control policy $V_{\mu,x} = (\tau_1, \mu_1; \tau_2, \mu_2; \dots)$ of random stopping times τ_i and regimes μ_i such that the total cost

$$J_{\mu,x} = E_{\mu,x}\left[\int_0^T f^{\mu(t)}\big(y_x(t)\big)\exp\left(-\int_0^t c^{\mu(s)}\big(y_x(s)\big)\,ds\right)dt\right] \tag{1.2}$$

is minimized, where $\mu(t) = \mu_i$, $\tau_i \le t < \tau_{i+1}$, with given $\mu(0) = \mu_0$, $E_{\mu,x}$ is the expectation and $T = \tau_x$ is the first exit time of the process for an exit problem while $T = +\infty$ in case of reflecting boundary conditions. The optimal cost function or value function $u(x)$, $x \in \Omega$, is given by

$$u(x) = \inf_{V_{\mu,x}} J_{\mu,x}\,. \tag{1.3}$$

We denote by A^μ, $1 \le \mu \le m$, the second-order elliptic differential operators

$$A^\mu = -\sum_{i,j=1}^{d} a_{ij}^\mu(x)\,\partial^2_{x_i x_j} - \sum_{i=1}^{d} b_i^\mu(x)\,\partial_{x_i} + c^\mu(x) \tag{1.4}$$

where $a^\mu = (a_{ij}^\mu)_{i,j=1}^{d}$, $a^\mu = \sigma^\mu(\sigma^\mu)^*/2$, $1 \le \mu \le m$.
Then, a formal application of the dynamic programming principle shows that the value function u satisfies the following Hamilton-Jacobi-Bellman (HJB-) equation

$$\max_{1\le\mu\le m} (A^\mu u(x) - f^\mu(x)) = 0, \quad x \in \Omega \tag{1.5a}$$

with homogeneous Dirichlet boundary conditions

$$u(x) = 0, \quad x \in \Gamma = \partial\Omega \tag{1.5b}$$

for exit problems and oblique derivative boundary conditions

$$\max_{1\le\mu\le m} (\gamma^\mu(x)\cdot\nabla u(x)) = 0, \quad x \in \Gamma = \partial\Omega \tag{1.5b)'$$

in case of reflecting boundary conditions.

Remark: If the functions f^μ, $1 \le \mu \le m$, in (1.2) represent performances or utilities, then the control objective is to maximize $J_{\mu,x}$. In this case, the optimal performance or utility function $u(x)$, $x \in \Omega$, is given by

$$u(x) = \sup_{V_{\mu,x}} J_{\mu,x} \tag{1.3)'$$

and the associated HJB-equation takes the form

$$\min_{1\le\mu\le m} (A^\mu u(x) - f^\mu(x)) = 0, \quad x \in \Omega \tag{1.5a)'$$

More details on controlled diffusion processes can be found in the textbooks by A. Bensoussan [1], A. Bensoussan and J.L. Lions [2],[3], W.H. Fleming and R. Rishel [13] and N.V. Krylov [22]. For a derivation of the HJB-equations in the context of viscosity solutions see the pioneering work by M. G. Crandall and P. L. Lions [12] and P. L. Lions [26],[27]. Regularity results have been obtained by various authors, for $C^{2,\alpha}(\overline{\Omega})$ regularity, $\alpha \in (0,1)$, in case of exit problems the reader is referred to L. Caffarelli, J. Kohn, L. Nirenberg and J. Spruck [8] and N.V. Krylov [23],[24],[25] while for reflecting boundary conditions $C^{2,\alpha}(\Omega) \cap C^{1,1}(\overline{\Omega})$ regularity has been proved by P.L. Lions and N.S. Trudinger [29].

As far as the numerical solution of HJB-equations is concerned, we mention the work done by Ph. Cortey-Dumont [11] and P. L. Lions and B. Mercier [28]. In particular, in [28] the authors consider an iterative scheme based on finite element discretizations of a HJB-equation of type (1.5a),(1.5b) where at each iteration step (1.5a) is linearized by locally choosing that $\mu \in \{1, \dots ,m\}$ for which the maximum is attained. Then, the resulting linear algebraic system can be solved by either direct methods or standard iterative solvers. However, it is well known that with decreasing step sizes h and thus increasing number N_h of unknowns direct methods suffer from an increasing computational work of order $O(N_h^3)$ while the convergence rates of standard iterative solvers deteriorate according to $1-O(h^2)$. These drawbacks can be overcome by the application of multi-grid methods where the computational work is directly proportional to the number of unknowns and the convergence rate is typically independent of the step sizes (cf. e.g. A. Brandt [5] and W. Hackbusch [16]). Multi-grid algorithms for HJB-equations with homogeneous Dirichlet boundary conditions, based on the iterative schemes given in [28] and using analogous finite difference discretizations of positive type with respect to a hierarchy of grids, have been developped by R.H.W. Hoppe in [17]. In this paper, we will present a more direct multi-grid approach which uses nonlinear Gauss-Seidel iteration applied to (1.5a) as a smoother and an adaptive local choice of prolongations and restrictions in the coarse-to-fine and fine-to-coarse transfers of the multi-grid cycles. That multi-grid algorithm will be detailly described in §2. Then, in §3 we will derive a local convergence result using nonlinear multi-grid convergence theory in the spirit of W. Hackbusch [14],[15], [16] and elementary subdifferential calculus as basic tools. The idea of proof is very similar to that one used by R. H. W. Hoppe in [18],[19] and by R. H. W. Hoppe and R. Kornhuber [20] concerning multi-grid algorithms for variational inequalities, complementarity problems and free boundary value problems, respectively. Finally, in §4 some numerical results will be given for HJB-equations with Dirichlet and Neumann boundary conditions the latter one representing the maximal utility of profits in optimal advertising for a stochastic dynamic sales response model of Vidale-Wolfe type.

2. THE MULTI-GRID ALGORITHM

In this section we will develop a multi-grid algorithm for the numerical solution of HJB-equations of type (1.5a) with either homogeneous Dirichlet boundary conditions (1.5b) or oblique derivative boundary conditions (1.5b)'.

For notational convenience we will restrict ourselves to the two-dimensional case, i.e., we assume $\Omega \subset \mathbb{R}^2$, and we consider a hierarchy of grids $(\Omega_k)_{k=0}^{\ell}$ constructed in the following way: For step sizes $h_{k+1} = h_k/2$, $0 \leq k \leq \ell-1$, given some $h_0 > 0$, we define

$$\mathbb{R}_k^2 = \{x_\nu = (x_{\nu_1}, x_{\nu_2}) \mid x_{\nu_i} = \nu_i h_k,\ \nu_i \in \mathbb{Z},\ 1 \leq i \leq 2\}$$

and we set

$$\Omega_k = \Omega \cap \mathbb{R}_k^2, \quad 0 \leq k \leq \ell \tag{2.1}$$

which we refer to as the set of interior grid points (h_0 is assumed to be sufficiently small in order to guarantee $\Omega_0 \neq \emptyset$). Further, we define

$$\Gamma_k = \Gamma \cap \{(x_{\nu_1}, x_2), (x_1, x_{\nu_2}) \mid x_i \in \mathbb{R},\ x_{\nu_i} = \nu_i h_i,\ \nu_i \in \mathbb{Z},\ 1 \leq i \leq 2\}$$

as the set of boundary grid points. Then we discretize the elliptic differential operators A^μ, $1 \leq \mu \leq m$, with respect to $\overline{\Omega}_k = \Omega_k \cup \Gamma_k$,

$0 \le k \le \ell$, by standard finite difference approximations L_k^μ involving difference stars consisting of at most a grid point $x \in \Omega_k$ and its eight nearest neighbours. (If necessary, a Shortley-Weller type scheme will be used for discretizations near the boundary). Moreover, in boundary grid points $x \in \Gamma_k$ the gradient $\nabla u(x) = (\partial_{x_1} u(x), \partial_{x_2} u(x))^T$ will be approximated by $\nabla_k u_k(x) = (D_{k,x_1} u_k(x), D_{k,x_2} u_k(x))^T$ with suitably defined forward resp. backward difference quotients $D_{k,x_i} u_k(x)$, $1 \le i \le 2$. We further define the grid functions $f_k^\mu(x)$, $x \in \Omega_k$, and $\gamma_k^\mu(x)$, $x \in \Gamma_k$, $1 \le \mu \le m$, $0 \le k \le \ell$, by pointwise restriction of f^μ and γ^μ to Ω_k and Γ_k, respectively. Then, the discretized HJB-equation on Ω_k, $0 \le k \le \ell$, reads as follows

$$\max_{1 \le \mu \le m} \left(L_k^\mu u_k(x) - f_k^\mu(x) \right) = 0 , \quad x \in \Omega_k \tag{2.2a}$$

with boundary conditions given either by

$$u_k(x) = 0 , \quad x \in \Gamma_k \tag{2.2b}$$

or

$$\max_{1 \le \mu \le m} \left(\gamma_k^\mu(x) \cdot \nabla_k u_k(x) \right) = 0 , \quad x \in \Gamma_k . \tag{2.2b'}$$

Identifying grid functions on Ω_k with vectors in $\mathbb{R}^{N_k}$, $N_k = \operatorname{card} \Omega_k$, and incorporating the boundary conditions (2.2b) resp. (2.2b)', the discretized boundary value problem can be algebraically formulated as the nonlinear system

$$F_k(u_k) = \max_{1 \le \mu \le m} \left(A_k^\mu - f_k^\mu \right) = 0 \tag{2.3}$$

where the A_k^μ's, $1 \le \mu \le m$, represent the resulting $N_k \times N_k$ coefficient matrices. As far as these matrices are concerned, throughout the following we will assume:

The matrices A_k^μ, $0 \le k \le \ell$, $1 \le \mu \le m$, have all (2.4) positive diagonal elements, non-positive off-diagonal elements and are lower semistrictly diagonally dominant.

We recall that a matrix $A = (a_{ij})_{i,j}^{n}$ is called lower semistrictly diagonally dominant, if it is diagonally dominant and

$$|a_{ii}| > \sum_{j=1}^{i-1} |a_{ij}|, \quad 2 \le i \le n$$

(cf. e.g. A. Berman and R.J. Plemmons [4]).
In particular, under hypothesis (2.4) the A_k^{μ}'s are nonsingular M-matrices and hence, the functions $F_k : \mathbb{R}^{N_k} \to \mathbb{R}^{N_k}$ defined by means of (2.3) can be easily shown to be continuous surjective M-functions which in turn implies that for any choice of f_k^{μ}, $1 \le \mu \le m$, the HJB-equations (2.3) admit unique solutions $u_k^* \in \mathbb{R}^{N_k}$, $0 \le k \le \ell$ (cf. W.C. Rheinboldt [31]).

The multi-grid approach to the numerical solution of (2.3) on the highest level $k = \ell$ is based on the fact that the coarse grid corrections on the lower levels can also be formulated as HJB-type equations: Given u_ℓ^ν, $\nu \ge 0$, on level ℓ and having computed a smoothed iterate $\bar{u}_\ell^\nu$, the error $e_\ell^\nu = u_\ell^* - \bar{u}_\ell^\nu$ satisfies

$$\max_{1 \le \mu \le m} \left(A_\ell^\mu e_\ell^\nu - d_\ell^\mu \right) = 0$$

where d_ℓ^μ denotes the defect $d_\ell^\mu = f_\ell^\mu - A_\ell^\mu \bar{u}_\ell^\nu$, $1 \le \mu \le m$. This suggests to correct $\bar{u}_\ell^\nu$ according to

$$\bar{u}_\ell^{\nu,new} = \bar{u}_\ell^\nu + p_{\ell-1}^{\ell} \left(u_{\ell-1} - r_\ell^{\ell-1} \bar{u}_\ell^\nu \right) \tag{2.5}$$

where $p_{\ell-1}^{\ell}$ and $r_\ell^{\ell-1}$ is a suitable prolongation and restriction, respectively, and $u_{\ell-1} \in \mathbb{R}^{N_{\ell-1}}$ solves the lower dimensional HJB-equation

$$\max_{1 \le \mu \le m} \left(A_{\ell-1}^\mu u_{\ell-1} - g_{\ell-1}^\mu \right) = 0 \tag{2.6}$$

with $g^{\mu}_{\ell-1}$ given by $g^{\mu}_{\ell-1} = A^{\mu}_{\ell-1} r^{\ell-1}_{\ell} \overline{u}^{\nu}_{\ell} + r^{\ell-1}_{\ell} d^{\mu}_{\ell}$. Then, choosing $u^{\nu}_{\ell-1} = r^{\ell-1}_{\ell} \overline{u}^{\nu}_{\ell}$ as a startiterate on level ℓ-1 , the above process will be successively repeated until the coarsest grid k = 0 is reached. Finally, after returning to the highest level, several post-smoothing steps will be performed with respect to the startiterate $\overline{u}^{\nu,new}_{\ell}$ thus yielding a new iterate $u^{\nu+1}_{\ell}$.

We will now describe in detail both the smoothing process and the choice of restrictions and prolongations. As a smoother on levels $1 \le k \le \ell$ and as an iterative solver for the coarse grid correction on the lowest level k = 0 we choose nonlinear Gauss-Seidel iteration applied to the HJB-equations

$$\max_{1 \le \mu \le m} \left(A^{\mu}_{k} u_{k} - g^{\mu}_{k} \right) = 0 \tag{2.7}$$

where g^{μ}_{k} , $1 \le \mu \le m$, $0 \le k \le \ell$, is recursively defined by

$$g^{\mu}_{\ell} = f^{\mu}_{\ell} \tag{2.8a}$$

$$g^{\mu}_{k} = A^{\mu}_{k} r^{k}_{k+1} \overline{u}^{\nu}_{k} + r^{k}_{k+1} d^{\mu}_{k+1} , \quad 0 \le k \le \ell-1 . \tag{2.8b}$$

We remind that nonlinear Gauss-Seidel iteration is known as a convergent iterative procedure for nonlinear algebraic systems involving M-functions (cf. e.g. W. C. Rheinboldt [31]).

We denote by

$$A^{\mu}_{k} = D^{\mu}_{k} - L^{\mu}_{k} - U^{\mu}_{k} , \quad 1 \le \mu \le m , \quad 0 \le k \le \ell \tag{2.9}$$

the decomposition of the matrix A^{μ}_{k} into its diagonal, lower diagonal and upper diagonal part. Then, performing κ nonlinear Gauss-Seidel iterations with respect to a lexicographic ordering of grid points amounts to the successive solution of

$$v_k^{\iota+1} = S_k\left(v_k^\iota ; g_k^1, \dots, g_k^m\right) := \min_{1\le\mu\le m} \left((D_k^\mu - L_k^\mu)^{-1} (U_k^\mu v_k^\iota + g_k^\mu)\right), \quad 0 \le \iota \le \kappa-1 \tag{2.10}$$

where $v_k^0 = \bar{u}_k^\nu$ for pre-smoothing and $v_k^0 = \bar{u}_k^{\nu,new}$ for post-smoothing.

Remark: In case of an HJB-equation of type (1.5a)' the corresponding nonlinear Gauss-Seidel iteration is given by (2.10) with " $\min\limits_{1\le\mu\le m}$ " replaced by " $\max\limits_{1\le\mu\le m}$ " .

In the coarse-to-fine and fine-to-coarse transfers a natural choice for p_{k-1}^k and r_k^{k-1} , $1 \le k \le \ell$, are prolongations based on bilinear interpolation and full weighted nine-point restrictions (cf. e.g. W. Hackbusch [16]). However, in the present situation full weighted restriction cannot be used globally, because otherwise false information would be transferred to the coarser grid causing non-convergence of the multi-grid algorithm. To see this, for $x \in \Omega_k$ we denote by $N_k(x)$ the set of grid points consisting of x itself and its eight nearest neighbours in $\bar{\Omega}_k$; i.e.,

$$N_k(x) = \{x, x \pm h_k e_\sigma \mid 1 \le \sigma \le 4\} \cap \bar{\Omega}_k$$

where $e_1 = (1,0)$, $e_2 = (0,1)$, $e_3 = e_1 + e_2$, $e_4 = e_1 - e_2$ (with appropriate modifications for grid points near the boundary Γ_k). Then, for a smoothed iterate $\bar{u}_k^\nu = v_k^\kappa$ we denote by Ω_k^μ the set of all grid points for which the minimum in (2.9) is attained at $\mu \in I_m = \{1, \dots, m\}$, i.e.

$$\Omega_k^\mu\left(\bar{u}_k^\nu\right) = \left\{ x \in \Omega_k \;\middle|\; \bar{u}_{k,i(x)}^\nu = \left((D_k^\mu - L_k^\mu)^{-1} (U_k^\mu v_k^{\kappa-1} + g_k^\mu)\right)_{i(x)} \right\} \tag{2.11}$$

where i stands for the bijective map which to each $x \in \Omega_k$ assigns its corresponding index $i(x) \in I_{N_k} = \{1, \dots, N_k\}$. Further, we define the grid point sets

$$\Omega_k^{\mu,0}\left(\bar{u}_k^\nu\right) = \left\{ x \in \Omega_k^\mu\left(\bar{u}_k^\nu\right) \;\middle|\; N_k(x) \cap \Omega_k \subset \Omega_k^\mu\left(\bar{u}_k^\nu\right) \right\}, \tag{2.12a}$$

$$\Omega_k^{\mu,1}\left(\bar{u}_k^\nu\right) = \Omega_k^{\mu}\left(\bar{u}_k^\nu\right) \setminus \Omega_k^{\mu,0}\left(\bar{u}_k^\nu\right), \tag{2.12b}$$

Now, let us suppose that we have a coarse grid point $x \in \Omega_{k-1} \cap \Omega_k^{\mu,1}(\bar{u}_k^\nu)$, $\mu \in I_m$ .Then there exists at least one fine grid neighbour $y \in N_k(x) \cap \Omega_k^{\bar{\mu}}(\bar{u}_k^\nu)$ for some $\bar{\mu} \in I_m$, $\bar{\mu} \neq \mu$. Since the defect d_k^μ in y is not a reliable indicator for the accuracy of $\bar{u}_k^\nu$ in x, full weighted nine-point restriction would result in a transfer of false information from $N_k(x)$ to $x \in \Omega_{k-1}$. In particular, the solution $\bar{u}_k^\nu$ on level ℓ in general would not be a fixed point of the multi-grid iteration. A possible remedy is the use of global pointwise restriction. But having in mind the situation analyzed above it is sufficient to use pointwise restriction only there where problems might occur, namely for grid points $x \in \Omega_{k-1} \cap \Omega_k^{\mu,1}(\bar{u}_k^\nu)$. Thus, denoting by $\mathring{r}_k^{k-1}$ and $\hat{r}_k^{k-1}$ pointwise and full weighted restriction, respectively, we advocate the following local choice of the restrictions r_k^{k-1}, $1 \leq k \leq \ell$,

$$\left(r_k^{k-1} v_k\right)(x) = \begin{cases} \left(\hat{r}_k^{k-1} v_k\right)(x), & x \in \Omega_{k-1} \cap \Omega_k^{\mu,0}(\bar{u}_k^\nu), \quad \mu \in I_m \\ \left(\mathring{r}_k^{k-1} v_k\right)(x), & x \in \Omega_{k-1} \cap \Omega_k^{\mu,1}(\bar{u}_k^\nu), \quad \mu \in I_m \end{cases} \tag{2.13}$$

As far as the prolongations are concerned, the global use of bilinear interpolation may also cause instability in a vicinity of grid points associated with different $\mu \in I_m$. For that reason we propose the following local choice of p_{k-1}^k, $1 \leq k \leq \ell$,

$$\left(p_k^{k-1} v_{k-1}\right)(x) = \begin{cases} \left(\hat{p}_{k-1}^k v_{k-1}\right)(x), & x \in \Omega_k^{\mu,0}(\bar{u}_k^\nu), \quad \mu \in I_m \\ 0, & x \in \Omega_k^{\mu,1}(\bar{u}_k^\nu), \quad \mu \in I_m \end{cases} \tag{2.14}$$

where $\hat{p}_k^{k-1}$ denotes prolongation based on bilinear interpolation.

The following procedure $\mathrm{MGHJB}(\ell, u_\ell, g_\ell^1, \dots, g_\ell^m)$ written in quasi-Algol describes a complete multi-grid cycle for $\ell+1$ levels $k = 0,1, \dots, \ell$ where $u_\ell = u_\ell^\nu$ before and $u_\ell = u_\ell^{\nu+1}$ after the execution of the cycle.

$\mathrm{MGHJB}(\ell, u_\ell, g_\ell^1, \dots, g_\ell^m)$, integer i,ℓ ; array $u_\ell, g_\ell^1, \dots, g_\ell^m$;
if $\ell = 0$ then
  for $i := 1$ step 1 until κ_3 do $u_\ell := S_\ell(u_\ell; g_\ell^1, \dots, g_\ell^m)$ else

begin array $u_{\ell-1}$, $d_\ell^1, \dots, d_\ell^m$, $g_{\ell-1}^1, \dots, g_{\ell-1}^m$;
  for $i := 1$ step 1 until κ_1 do $u_\ell := S_\ell(u_\ell; g_\ell^1, \dots, g_\ell^m)$;
  for $i := 1$ step 1 until m do $d_\ell^i := g_\ell^i - A_\ell^i u_\ell$;
  $u_{\ell-1} := r_\ell^{\ell-1} u_\ell$;
  for $i := 1$ step 1 until m do $g_{\ell-1}^i := A_{\ell-1}^i u_{\ell-1} + r_\ell^{\ell-1} d_\ell^i$;
  for $i := 1$ step 1 until $\gamma_{\ell-1}$ do $\mathrm{MGHJB}(\ell-1, u_{\ell-1}, g_{\ell-1}^1, \dots, g_{\ell-1}^m)$;
  $u_\ell := u_\ell + p_\ell^{\ell-1}(u_{\ell-1} - r_\ell^{\ell-1} u_\ell)$;
  for $i := 1$ step 1 until κ_2 do $u_\ell := S_\ell(u_\ell; g_\ell^1, \dots, g_\ell^m)$;
end MGHJB.

Note that at each level $1 \le k \le \ell$ within the cycle $\kappa_1 \ge 0$ pre-smoothings and $\kappa_2 \ge 0$ post-smoothings are performed while the number of *nonlinear Gauss-Seidel iterations* for the approximate solution of the correction HJB on the lowest level $k = 0$ is $\kappa_3 \ge 0$. The structure of the cycle is determined by γ_k , $1 \le k \le \ell-1$. (For $\gamma_k = 1$ we have a "V" -cycle and for $\gamma_k = 2$ a "W" -cycle).

A suitable startiterate u_ℓ^0 on the highest level $k = \ell$ can be obtained by nested iteration, i.e., using suitable prolongations $\tilde{p}_{k-1}^k$, $1 \le k \le \ell$, an approximation u_0 on the lowest level $k = 0$ is prolongated onto the level $k = 1$ yielding $u_1^0 = \tilde{p}_0^1 u_0$ which is then used as a start-iterate for the execution of one or several cycles $\mathrm{MGHJB}(1, u_1, g_1^1, \dots, g_1^m)$. This process will be repeated until the highest level $k = \ell$ is reached. Since nested iteration is a fairly standard procedure in the multi-grid business, for more details we refer to W. Hackbusch [16].

3. CONVERGENCE RESULTS

In this section we will prove local convergence of the multi-grid algorithm $\mathrm{MGHJB}(\ell, u_\ell, g_\ell^1, \dots, g_\ell^m)$ by using elementary subdifferential calculus and fundamental ingredients of nonlinear multi-grid convergence theory in the spirit of W. Hackbusch [14],[15].

We start with some basic assumptions on the continuous problem (1.5):

The second order elliptic differential operators A^μ, $1 \le \mu \le m$, given by (1.4) are supposed to be uniformly elliptic with smooth coefficients a_{ij}^μ, $b_i^\mu, c^\mu \in C^2(\overline{\Omega})$, $1 \le i, j \le 2$. Further, we assume $f^\mu \in C^2(\overline{\Omega})$ and, in case of reflecting boundary conditions, $\gamma_i^\mu \in C^2(\overline{\Omega})$, $1 \le i \le 2$.

As mentioned in the Introduction, under the above hypotheses the value function $u(x)$, $x \in \overline{\Omega}$, of the optimally controlled stochastic switching process under consideration can be shown to be the unique solution in $C^{2,\alpha}(\overline{\Omega})$ resp. $C^{2,\alpha}(\Omega) \cap C^{1,1}(\overline{\Omega})$ for some $\alpha \in (0,1)$ of the HJB-equation (1.5a) under the boundary conditions (1.5b) resp. (1.5b)' (cf. [8],[25],[29]).

If $u^*(x)$, $x \in \overline{\Omega}$, is the unique solution to (1.5), for $\mu \in I_m$ we set

$$\Omega^\mu = \{x \in \Omega \mid (A^\mu u^* - f^\mu)(x) = \max_{1 \le \sigma \le m} ((A^\sigma u^* - f^\sigma)(x))\},$$

$$\Gamma^\mu = \overline{\partial\Omega^\mu} \cap \left(\overline{\Omega \setminus \Omega^\mu}\right).$$

We assume that the boundary value problem (1.5) is nondegenerate in the sense that

$$\Omega^\mu \cap \Omega^\sigma = \emptyset, \quad \mu, \sigma \in I_m, \quad \mu \ne \sigma.$$

We define

$$\Gamma^* = \bigcup_{\mu \in I_m} \Gamma^\mu \tag{3.2}$$

as the continuous internal free boundary. Likewise, in view of the definition of the sets $\Gamma_k^{\mu,0}$ by (2.12b) we refer to

$$\Gamma_k^* = \bigcup_{\mu\in I_m} \Gamma_k^{\mu,0}(u_k^*), \quad 0 \le k \le \ell \tag{3.3}$$

as the discrete internal free boundary, and we also suppose nondegeneracy of the discrete problems (2.3), i.e.,

$$\Omega_k^{\mu}(u_k^*) \cap \Omega_k^{\sigma}(u_k^*) = \emptyset, \quad \mu,\sigma \in I_m, \quad \mu \ne \sigma, \quad 0 \le k \le \ell . \tag{3.4}$$

Concerning the internal free boundaries, in the sequel we will assume:

(i) The continuous internal free boundary Γ^* is a one-dimensional manifold admitting a Lipschitzian parametrization. (3.5a)

(ii) The discrete internal free boundaries Γ_k^*, $0 \le k \le \ell$, are situated in an $O(h_k)$ -neighbourhood of the continuous internal free boundary Γ^*, i.e.,

$$\max_{x\in\Gamma_k^*} \operatorname{dist}(x,\Gamma^*) = O(h_k) \quad (h_k \to 0) \tag{3.5b}$$

Remarks: (i) For a model obstacle problem, discretized by piecewise linear finite elements, under fairly general hypotheses F. Brezzi and L. Caffarelli [7] have established convergence of the discrete internal free boundaries to the continuous one of an order which is approximately the square root of the L^∞ convergence rate for the convergence of the solutions of the discrete problems to the continuous one. Taking into account both the relationship of HJB-equations of type (1.5a) to implicit obstacle problems and the optimal L^∞ convergence rate established by Ph. Cortey-Dumont [11] and F. Conrad and Ph. Cortey-Dumont [10], there is evidence that (3.5b) holds true in the situation considered in this paper.

(ii) Note that under hypotheses (3.5a),(3.5b) the grid point sets $\Omega_k^{\mu}(u_k^*)$, $1 \le \mu \le m$, $0 \le k \le \ell$, satisfy "property C" in the sense of W. Hackbusch [14].

Moreover, throughout the following we will assume that $\|\cdot\|_p$, $\|\cdot\|_p$, $0 \le p \le 3$, are norms on $\mathbb{R}^{N_k}$ (not necessarily all different) such that $\|\cdot\|_2$ is a monotone vector norm and $\| v_k \|_p \le \| v_k \|_{p+1}$, $v_k \in \mathbb{R}^{N_k}$, $0 \le p \le 2$. For a linear map $T_k^{\bar{k}} : \mathbb{R}^{N_k} \to \mathbb{R}^{N_{\bar{k}}}$ the associated matrix norms will be denoted by $\| T_k^{\bar{k}} \|_{p,q} = \sup\{ \| T_k^{\bar{k}} v_k \|_q / \| v_k \|_p \mid v_k \in \mathbb{R}^{N_k}, v_k \neq 0 \}$. Since A_k^μ, $0 \le k \le \ell$, $1 \le \mu \le m$, result from finite difference discretizations of second order uniformly elliptic operators, it is convenient to choose $\|\cdot\|_p$, $0 \le p \le 3$, as discrete analogues $|\cdot|_s$ of the Sobolev norms in $H^s(\Omega)$, $s \in \mathbb{R}^+$, $s + \frac{1}{2} \notin \mathbb{N}$. In particular, a suitable choice is $\|\cdot\|_0 = |\cdot|_{-2}$, $\|\cdot\|_1 = |\cdot|_{\beta-1}$, $\|\cdot\|_2 = |\cdot|_0$ and $\|\cdot\|_3 = |\cdot|_{1+\beta}$ for some appropriate $\beta \in [0,1]$ (cf. W. Hackbusch [14]).

We now consider the nonlinear functions $F_k : \mathbb{R}^{N_k} \to \mathbb{R}^{N_k}$, $0 \le k \le \ell$, defined by (2.3). These mappings are not differentiable in the usual sense, but admit generalized Jacobians $\partial F_k(u_k)$, $u_k \in \mathbb{R}^{N_k}$, in the sense of F.H. Clarke [9] satisfying

$$\partial F_k (u_k) \subseteq \partial F_{k,1} (u_k) \times \partial F_{k,2} (u_k) \times \dots \times \partial F_{k,N_k}(u_k) \tag{3.6}$$

where the right-hand side denotes the set of matrices whose i-th row is an element of the generalized gradient $\partial F_{k,i} (u_k)$ of the i-th component of F_k. Since F_k is defined by means of a max-function with a finite number of arguments, the generalized gradients can be easily computed: Denoting by $x_k(i)$ that grid point in Ω_k uniquely associated to $i \in \{1, \dots, N_k\}$, we have

$$\partial F_{k,i}(u_k) = \mathrm{co}\left(A_{k,i}^{\mu_1}, \dots, A_{k,i}^{\mu_p}\right), \tag{3.7}$$

if $x_k(i) \in \bigcup_{j=1}^{p} \Omega^{\mu_j}(u_k) \setminus (\bigcup_{\sigma \neq \mu_j} \Omega^{\sigma}(u_k))$, $1 \le p \le m$, where $A_{k,i}^{\mu_j}$ denotes the i-th row of the matrix $A_k^{\mu_j}$ and the right-hand side in (3.7) is the convex hull of the vectors $A_{k,i}^{\mu_j}$, $1 \le j \le p$.

Moreover, we have the generalized mean-value theorem (cf. [9; Prop. 2.6.5]):

$$F_k u_k - F_k v_k \in \mathrm{co}\, \partial F_k ([u_k, v_k]) (u_k, v_k) \tag{3.8}$$

where the right-hand side stands for the convex hull of all vectors of the form $DF_k(u_k-v_k)$ with $DF_k \in \partial F_k(w_k)$, $w_k \in [u_k, v_k] = \{z_k \mid z_k = tu_k + (1-t)v_k,\ t \in [0,1]\}$. In particular, from (3.8) we deduce the existence of a matrix $D_kF_k[u_k, v_k] \in \partial F_k([u_k, v_k])$ (not necessarily unique) such that

$$F_k u_k - F_k v_k = DF_k[u_k, v_k](u_k, v_k).$$

Now, in view of (3.4), (3.6) and (3.7) the generalized gradients $\partial F_{k,i}(u_k^*)$, $1 \le i \le N_k$, and hence $\partial F_k(u_k^*)$ are single valued. In particular $\partial F_k(u_k^*)$ is given by

$$\partial F_{k,i}(u_k^*) = A_{k,i}^{\mu}, \quad \text{if} \quad x_k(i) \in \Omega_k^{\mu}(u_k^*), \quad 1 \le i \le N_k. \qquad (3.9)$$

Furthermore, in view of the upper semicontinuity of $\partial F_k(\cdot)$ (cf. [9; Prop. 2.6.2(c)]), for all $\varepsilon_k > 0$ there exists $\delta_k > 0$ such that

$$\| \partial F_k^{-1}(u_k^*)\, DF_k[u_k, v_k] - I_k \|_{2,2} < \varepsilon_k \qquad (3.10)$$

for all u_k, v_k such that $\| u_k - u_k^* \|_2 < \delta_k$ and $\| v_k - u_k^* \|_2 < \delta_k$.

Analogous results as above apply to the nonlinear mappings $\widetilde{F}_k$ and S_k defining the coarse grid HJB-Equations (2.7) on levels $0 \le k \le \ell-1$ and the nonlinear Gauss-Seidel iteration (2.10) on levels $0 \le k \le \ell$, the latter being defined by means of a min-function with m arguments. In particular, nondegeneracy of (2.3) on level $k+1$ implies nondegeneracy of (2.7) on level k and thus $\partial \widetilde{F}_k(r_{k+1}^k u_{k+1}^*)$ is single-valued with

$$\partial \widetilde{F}_{k,i}(r_{k+1}^k u_{k+1}^*) = A_{k,i}^{\mu}, \quad \text{if} \quad x_k(i) \in \Omega_k^{\mu}(r_{k+1}^k u_{k+1}^*), \quad 1 \le i \le N_k. \qquad (3.11)$$

Moreover, $\partial S_k(u_k^*)$ is given by

$$\partial S_{k,i}(u_k^*) = ((D_k^{\mu} - L_k^{\mu})^{-1} U_k^{\mu})_i, \quad \text{if} \quad x_k^{(i)} \in \Omega_k^{\mu}(u_k^*), \quad 1 \le i \le N_k. \qquad (3.12)$$

Now, we are in a position to establish local convergence of the multi-grid algorithm MGHJB. We will proceed in the spirit of W. Hackbusch's convergence theory in [14], [15], [16] and first obtain a two-grid convergence result by linearization of the two-grid iteration

operator at the solution and then verifying a basic approximation and smoothness property for that linearized operator.
The first partial result deals with the smoothing process:

Proposition 3.1. Let u_ℓ^ν , $\nu \geq 0$, be the ν–th multi-grid iterate and let further $\bar{u}_\ell^\nu$ be the result of the smoothing process obtained by κ nonlinear Gauss-Seidel iterations (2.10) applied to the HJB-equation (2.3) on level ℓ . Then, under assumptions (2.4) and (3.4), there exists a matrix $D^\kappa S_\ell[u_\ell^\nu, u_\ell^*]$ such that

$$\bar{u}_\ell^\nu - u_\ell^* = D^\kappa S_\ell [u_\ell^\nu, u_\ell^*] (u_\ell^\nu - u_\ell^*) \ . \tag{3.13}$$

Moreover, for all $\varepsilon_\ell > 0$ there exists $\delta_\ell > 0$ such that

$$\| D^\kappa S_\ell [u_\ell^\nu, u_\ell^*] - (\partial S_\ell)^\kappa (u_\ell^*) \|_{2,2} < \varepsilon_\ell \tag{3.14}$$

for all u_ℓ^ν with $\| u_\ell^\nu - u_\ell^* \| < \delta_\ell$.

Proof: In view of (2.10) a matrix $D^\kappa S_\ell [u_\ell^\nu, u_\ell^*]$ satisfying (3.13) can be chosen according to

$$D^\kappa S_\ell [u_\ell^\nu, u_\ell^*] = \prod_{\iota=0}^{\kappa-1} D S_\ell [v_\ell^{\kappa-\iota}, u_\ell^*]$$

where $DS_\ell [v_\ell^{\kappa-\iota}, u_\ell^*] \in \mathrm{co}\, \partial S_\ell ([v_\ell^{\kappa-\iota}, u_\ell^*])$, $0 \leq \iota \leq \kappa-1$. Since each $DS_{\ell,i} [v_\ell^\iota, u_\ell^*]$, $1 \leq i \leq N_\ell$, $1 \leq \iota \leq \kappa$, is a convex combination of $((D_\ell^{\mu_j} - L_\ell^{\mu_j})^{-1} U_\ell^{\mu_j})_i$, $\mu_j \in I_m$, $1 \leq j \leq p$, $1 \leq p \leq m$, where $A_\ell^\mu = D_\ell^\mu - L_\ell^\mu - U_\ell^\mu$ satisfies (2.4) for all $1 \leq \mu \leq m$, we easily deduce that the spectral radius of $DS_{\ell,i}[v_\ell^\iota, u_\ell^*]$ is less than one, and thus $\| v_\ell^\iota - u_\ell^* \|_2 \to 0$. Hence, (3.14) is an immediate consequence of the upper semicontinuity of $\partial S_\ell (\cdot)$.

Before we analyze the coarse grid correction, we remark that by definition of the prolongations and restrictions (cf. (2.13),(2.14)) and in view of (3.5a),(3.5b) we may assume that there exist neighbourhoods $\mathcal{U}_k (u_k^*)$ such that for all $\bar{u}_k^\nu \in \mathcal{U}_k (u_k^*)$ (note that p_{k-1}^k and r_{k-1}^k are defined locally in dependence on $\overline{u}_k^{\nu}$)

$$C^{-1}\, \| v_{k-1} \|_2 \;\le\; \| p_{k-1}^{k} v_{k-1} \|_2 \;\le\; C\, \| v_{k-1} \|_2 , \quad 1 \le k \le \ell \qquad (3.15a)$$

$$\| r_k^{k-1} \|_{p,p} \;\le\; C , \quad 0 \le p \le 3 , \quad 1 \le k \le \ell \qquad (3.15b)$$

(cf. W. Hackbusch [14; p. 430]).

Moreover, we will impose the following conditions on the Jacobians $\partial F_k (u_k^*)$, $0 \le k \le \ell$, and $\partial \widetilde{F}_{k-1} (r_k^{k-1} u_k^*)$, $1 \le k \le \ell$:

$$\| \partial F_k (u_k^*) \|_{2,0} \le C , \quad \| \partial \widetilde{F}_{k-1} (r_k^{k-1} u_k^*) \|_{2,0} \le C \qquad (3.16)$$

$$\| (\partial F_k)^{-1} (u_k^*) \|_{p,p+2} \le C , \quad \| (\partial \widetilde{F}_{k-1})^{-1} (r_k^{k-1} u_k^*) \|_{p,p+2} \le C, \quad p = 0,1 \qquad (3.17)$$

In view of (3.9), (3.11) and the fact that the A_k^{μ}'s , $1 \le \mu \le m$, result from finite difference approximations of uniformly elliptic second order differential operators with smooth coefficients, the estimates (3.16),(3.17) can be expected to hold true (cf. W. Hackbusch [14]).

Next, we provide the following preparatory stability result for (2.3) which is also of interest in its own:

<u>Lemma 3.2</u>. Let u_k , $\widetilde{u}_k$ be solutions to the HJB-equations (2.3) with data f_k^{μ} and $\tilde{f}_k^{\mu}$, $1 \le \mu \le m$, $0 \le k \le \ell$, respectively. Then, under assumptions (2.4), (3.1), (3.4) and (3.17) we have

$$\| u_k - \widetilde{u}_k \|_2 \;\le\; C \max_{1 \le \mu \le m} \| f_k^{\mu} - \tilde{f}_k^{\mu} \|_0 . \qquad (3.18)$$

<u>Proof</u>: For all $1 \le i \le N_k$ it follows from (2.3) that

$$(A_k^{\mu} (u_k - \widetilde{u}_k))_i \;\le\; f_k^{\mu} - \tilde{f}_k^{\mu} , \quad x_k(i) \in \Omega_k^{\mu}(\widetilde{u}_k)$$

$$(A_k^{\mu} (u_k - \widetilde{u}_k))_i \;\ge\; f_k^{\mu} - \tilde{f}_k^{\mu} , \quad x_k(i) \in \Omega_k^{\mu}(u_k)$$

and hence, taking into account that in view of (2.4) and (3.9) both $\partial F_k (u_k)$ and $\partial F_k (\widetilde{u}_k)$ are nonsingular M -matrices

$$(\partial F_k)^{-1} (u_k) (f_k^{\mu} - \tilde{f}_k^{\mu}) \;\le\; u_k - \widetilde{u}_k \;\le\; (\partial F_k)^{-1} (\widetilde{u}_k) (f_k^{\mu} - \tilde{f}_k^{\mu}) .$$

Then, using the monotonicity of $\|\bullet\|_2$ and (3.17) (with u_k^* replaced by u_k and $\tilde{u}_k$, respectively), (3.18) follows instantly from the preceding inequalities.

We now compute the two-grid iteration operator where for simplicity we assume that only pre-smoothing is performed:

Proposition 3.3. Suppose that u_ℓ^ν, $\nu \geq 0$, are the iterates obtained by application of the multi-grid algorithm MGHJB(ℓ, u_ℓ, $g_\ell^1, \ldots, g_\ell^m$) in case of two grids with $\kappa_1 = \kappa$ pre-smoothings and $\kappa_2 = 0$ post-smoothings. Then, under assumptions (2.4), (3.1), (3.4), (3.15), (3.16) and (3.17) we have

$$u_\ell^{\nu+1} - u_k^* = (M_\ell^{\ell-1} + Z_\ell)(u_\ell^\nu - u_k^*) \tag{3.19}$$

where

$$M_\ell^{\ell-1} = [(\partial F_\ell)^{-1}(u_\ell^*) - p_\ell^{\ell-1}(\partial\tilde{F}_{\ell-1})^{-1}(r_\ell^{\ell-1}u_\ell^*)\, r_\ell^{\ell-1}] \bullet \\ \bullet [\partial F_\ell\ (u_\ell^*)(\partial S_\ell)^\kappa (u_\ell^*)] \tag{3.20}$$

and

$$\| Z_\ell \|_{2,2} \leq C(\kappa)\, \eta^{(\nu)} \tag{3.21}$$

with $C(\kappa) > 0$ and $\eta^{(\nu)} \to 0$ as $\|u_\ell^\nu - u_\ell^*\|_2 \to 0$.

Proof: First, we define a map $\tilde{F}_\ell^{\ell-1} : \mathbb{R}^{N_\ell} \to \mathbb{R}^{N_{\ell-1}}$ by

$$\tilde{F}_\ell^{\ell-1} v_\ell = \max_{1 \leq \mu \leq m} (r_\ell^{\ell-1}(A_\ell^\mu v_\ell - g_\ell^\mu)).$$

Then there exist matrices $D\tilde{F}_\ell^{\ell-1}[r_\ell^{\ell-1}\bar{u}_\ell^\nu, u_{\ell-1}] \in \mathrm{co}\,\partial\tilde{F}_{\ell-1}([r_\ell^{\ell-1}\bar{u}_\ell^\nu, u_{\ell-1}])$ and $D\tilde{F}_\ell^{\ell-1}[\bar{u}_\ell^\nu, u_\ell^*] \in \mathrm{co}\,\partial\tilde{F}_\ell^{\ell-1}([\bar{u}_\ell^\nu, u_\ell^*])$ such that

$$D\tilde{F}_{\ell-1}[r_\ell^{\ell-1}\bar{u}_\ell^\nu, u_{\ell-1}](r_\ell^{\ell-1}\bar{u}_\ell^\nu - u_{\ell-1}) = \\ \tilde{F}_{\ell-1} r_\ell^{\ell-1}\bar{u}_\ell^\nu - \tilde{F}_{\ell-1}u_{\ell-1} = \tilde{F}_\ell^{\ell-1}\bar{u}_\ell^\nu - \tilde{F}_\ell^{\ell-1}u_\ell^* = D\tilde{F}_\ell^{\ell-1}[\bar{u}_\ell^\nu - u_\ell^*](\bar{u}_\ell^\nu - u_\ell^*)$$

where we have used the fact that $\tilde{F}_\ell^{\ell-1} u_\ell^* = 0$.

Taking advantage of (3.13) in Proposition 3.1, we get

$$r_\ell^{\ell-1}\,\overline{u}_\ell^\nu - u_{\ell-1} = [\,(D\widetilde{F}_{\ell-1})^{-1}\,[r_\ell^{\ell-1}\,\overline{u}_\ell^\nu\,,\,u_{\ell-1}\,]\,D\widetilde{F}_\ell^{\ell-1}\,[\overline{u}_\ell^\nu\,,\,u_\ell^*\,]\,\bullet \tag{3.22}$$

$$\bullet\,D^\kappa S_\ell\,[u_\ell^\nu\,,\,u_\ell^*]]\,(u_\ell^\nu - u_\ell^*)\,.$$

Then, if we set

$$X_{\ell-1} = (\partial\widetilde{F}_{\ell-1})^{-1}\,(r_\ell^{\ell-1}\,u_\ell^*)\,D\widetilde{F}_{\ell-1}\,[r_\ell^{\ell-1}\,\overline{u}_\ell^\nu\,,\,u_{\ell-1}] - I_{\ell-1}$$

$$X_\ell = (\partial\widetilde{F}_\ell^{\ell-1})^{-1}\,(u_\ell^*)\,D\widetilde{F}_\ell^{\ell-1}\,[\,\overline{u}_\ell^\nu, u_\ell^*\,] - I_\ell$$

$$Y_\ell = D^\kappa S_\ell\,[\,u_\ell^\nu, u_\ell^*\,] - (\partial S_\ell)^\kappa(\,u_\ell^*)$$

and observe that by the chain rule $\partial\widetilde{F}_\ell^{\ell-1}(\,u_\ell^*) = r_\ell^{\ell-1}\,\partial F_\ell^{\ell-1}(\,u_\ell^*)$, it follows from (2.5) and (3.22) that

$$u_\ell^{\nu+1} - u_\ell^* = [(\partial F_\ell\,(\,u_\ell^*)\,(I_\ell + X_\ell))^{-1} - \tag{3.23}$$

$$- p_{\ell-1}^\ell\,(\partial\widetilde{F}_{\ell-1}(r_\ell^{\ell-1}\,u_\ell^*)\,(I_{\ell-1} + X_{\ell-1}))^{-1}\,r_\ell^{\ell-1}\,]\,\bullet$$

$$\bullet\,[\partial F_\ell\,(u_\ell^*)\,(I_\ell + X_\ell)\,((\partial S_\ell\,)^\kappa\,(u_\ell^*) + Y_\ell)]\,(u_\ell^\nu - u_\ell^*)\,.$$

From (3.10) and (3.14) we deduce

$$\|\,X_\ell\,\|_{2,2} \le C(\kappa)\eta^{(\nu)}, \quad \|\,Y_\ell\,\|_{2,2} \le C(\kappa)\eta^{(\nu)}\,. \tag{3.24a}$$

In order to establish the analogous result for $X_{\ell-1}$, i.e.,

$$\|\,X_{\ell-1}\,\|_{2,2} \le C(\kappa)\eta^{(\nu)} \tag{3.24b}$$

according to (3.10) we have to verify

$$\|\,r_\ell^{\ell-1}\,\overline{u}_\ell^\nu - r_\ell^{\ell-1}\,u_\ell^*\,\|_2 \to 0 \tag{3.25a}$$

$$\text{as } \|\,u_\ell^\nu - u_\ell^*\,\|_2 \to 0$$

$$\|\,u_{\ell-1} - r_\ell^{\ell-1}\,u_\ell^*\,\|_2 \to 0\,. \tag{3.25b}$$

Assertion (3.25a) readily follows from (3.15b) and Proposition 3.1. On the other hand, for the difference $u_{\ell-1} - r_\ell^{\ell-1}\,u_\ell^*$ we get by application of Lemma 3.2

$$\| u_{\ell-1} - r_\ell^{\ell-1} u_\ell^* \|_2 \le$$

$$\le C \max_{1\le\mu\le m} \| A_{\ell-1}^\mu (r_\ell^{\ell-1} \bar{u}_\ell^\nu - r_\ell^{\ell-1} u_\ell^*) - (r_\ell^{\ell-1} A_\ell^\mu \bar{u}_\ell^\nu - r_\ell^{\ell-1} A_\ell^\mu u_\ell^*) \|_0 .$$

Observing $\| A_k^\mu \|_{2,0} \le C$, $k = \ell-1, \ell$, and (3.15b) for $p = 0,2$, (3.24b) is a consequence of the preceding inequality and Proposition 3.1.
Now, using (3.24a),(3.24b) together with the estimates (3.15), (3.16), (3.17) and $\| (\partial S_\ell)^\kappa (u_\ell^*) \|_{2,2} \le C(\kappa)$, assertions (3.19), (3.20) and (3.21) can be easily deduced from (3.23).

From the representation (3.23) of the two-grid iteration operator it follows that we do have two-grid convergence, if both the approximation property

$$\| (\partial F_k)^{-1} (u_k^*) - p_{k-1}^k (\partial \tilde{F}_{k-1})^{-1} (r_k^{k-1} u_k^*) r_k^{k-1} \|_{1,2} \le C h_k^\alpha \tag{3.26}$$

and the smoothing property

$$\| \partial F_k (u_k^*) (\partial S_k)^\kappa (u_k^*) \|_{2,1} \le C_0(\kappa) h_k^{-\alpha}, \quad 1 \le \kappa \le \kappa_{max}(h_k) \tag{3.27}$$

are satisfied, where $\alpha = 1+\beta$, $C_0(\kappa) \to 0$ as $\kappa \to \infty$ and $\kappa_{max}(h_k) \to \infty$ as $h_k \to 0$.

Lemma 3.4. Let $M_\ell^{\ell-1}$ be the two-grid iteration operator as given by (3.20). Then, under assumptions (3.26) and (3.27) we have

$$\| M_\ell^{\ell-1} \|_{2,2} \le C C_0(\kappa), \quad 1 \le \kappa \le \kappa_{max}(h_\ell) \tag{3.28}$$

For the validity of the approximation property (3.26) we assume the operators A_{k-1}^μ to approximate A_k^μ in the sense that

$$\hat{r}_k^{k-1} A_k^\mu \hat{p}_{k-1}^k = A_{k-1}^\mu + \delta_{k-1}^\mu, \quad 1 \le k \le \ell, \ 1 \le \mu \le m \tag{3.29a}$$

where

$$\| \delta_{k-1}^\mu \|_{3,0} \le C h_{k-1}^\alpha . \tag{3.29b}$$

Then, setting

$$\delta_{k-1} = r_k^{k-1} \partial F_k (u_k^*) p_{k-1}^k - \partial \tilde{F}_k (r_k^{k-1} u_k^*)$$

by means of (3.9), (3.11) we have $\delta_{k-1,i} = \delta^{\mu}_{k-1,i}$ for $x_{k-1}(i) \in \Omega^{\mu,0}_{k-1}(r_k^{k-1} u_k^*)$, $1 \le i \le N_{k-1}$, while $\delta_{k-1,i}$ may contain terms of order $0(h_{k-1}^{-2})$ for $x_{k-1}(i) \in \Gamma^{\mu}_{k-1}(r_k^{k-1} u_k^*)$. Then, in view of hypotheses (3.5a), (3.5b) the estimate

$$\| \delta_{k-1} \|_{3,0} \le C\, h_{k-1}^{\alpha} \tag{3.30}$$

can be deduced from W. Hackbusch [14; Note 3 (p. 431)].
Similar arguments can be used to prove

$$\| I_k - p_{k-1}^{k} r_k^{k-1} \|_{3,2} \le C\, h_k^{\alpha} \tag{3.31}$$

(cf. W. Hackbusch [14; Corollary 1 (p. 430)]).
Now, under assumptions (3.15), (3.16) and (3.17) the approximation property (3.26) is a consequence of (3.30) and (3.31) (cf. W. Hackbusch [14; Criterion 1 (p. 428)]).

As far as the smoothing property (3.27) is concerned we remark that according to (3.9), (3.12) $\partial S_k(u_k^*)$ is the Gauss-Seidel iteration operator associated with $\partial F_k(u_k^*)$. Recalling the fact that the A_k^{μ}'s are finite difference approximations of uniformly elliptic differential operators with coefficients in class $C^2(\overline{\Omega})$, there is evidence that (3.27) holds true, in particular if we use red-black ordering of grid points instead of the lexicographic one (cf. [14; Chapter 4]).

Using Lemma 3.4 we immediately obtain:

Proposition 3.5. Under the hypotheses of Proposition 3.3 and Lemma 3.4 there holds

$$\| u_{\ell}^{\nu+1} - u_{\ell}^{*} \|_2 \le [C\, C_0(\kappa) + C(\kappa)\, \eta^{(\nu)}]\, \| u_{\ell}^{\nu} - u_{\ell}^{*} \|_2 \,. \tag{3.32}$$

Proof: The estimate (3.32) is a direct consequence of (3.19), (3.21) and (3.28).

Remark: The preceding results can be easily modified in order to cover the case where also post-smoothing is performed, i.e., $\kappa_2 > 0$ in MGHJB$(\ell, u_{\ell}, g_{\ell}^{1}, \dots, g_{\ell}^{m})$ (cf. e.g. W. Hackbusch [15]).

For more than two grids the multi-grid iteration operator can be recursively defined by means of the two-grid iteration operators $M_k^{k-1} + Z_k$ on levels $1 \le k \le \ell$. Then, the following local convergence result can be established:

Theorem 3.6. Let $\overset{\nu}{u}_\ell$, $\nu \ge 0$, be the iterates obtained by the multi-grid algorithm $MGHJB(\ell, u_\ell, \overset{1}{g}_\ell, \dots, \overset{m}{g}_\ell)$ for a hierarchy of $\ell+1$ grids $k = 0,1, \dots ,\ell$ assuming a "W" -cycle structure (i.e. $\gamma = 2$). Then, under the same hypotheses as in Proposition 3.5, there exists $\kappa_{min} \ge 1$ such that for all $\kappa_{min} \le \kappa \le \kappa_{max}\,(h_1)$ the estimate (3.32) holds true. In particular, if the startiterate $\overset{0}{u}_\ell$ is chosen in an appropriate neighbourhood of the solution u_ℓ^*, we have $\| \overset{\nu}{u}_\ell - u_\ell^* \|_2 \to 0$ as $\nu \to \infty$.

Proof: Using Proposition 3.5 the result follows by arguing along the same lines as in W. Hackbusch [15; Thms. 3.4, 3.13].

4. NUMERICAL RESULTS

We have tested the efficiency of the multi-grid algorithm MGHJB by applying it to an HJB-equation both with Dirichlet and Neumann boundary conditions.

The first, rather academic example is an HJB-equation for two uniformly elliptic operators A^1, A^2 under homogeneous Dirichlet data which has already served as a numerical test example in [17] and [28] . The operators A^1, A^2 and right-hand sides f^1, f^2 are chosen according to

$$A^1 = -\frac{\partial^2}{\partial x^2} - 0.5\frac{\partial^2}{\partial x\,\partial y} - \frac{\partial^2}{\partial y^2}, \quad A^2 = -0.5\frac{\partial^2}{\partial x^2} - 0.1\frac{\partial^2}{\partial x\,\partial y} - \frac{\partial^2}{\partial y^2} \tag{4.1}$$

$$f^1 = f^2 = \max(A^1 u, A^2 u), \quad u = x(1-x)\,y(1-y) \tag{4.2}$$

so that u is the exact solution of the Dirichlet problem (1.5a),(1.5b) on $\Omega = (0,1)\times(0,1)$.
With respect to a hierarchy of equidistant grids Ω_k with step sizes $h_{k+1} = h_k/2$, $0 \le k \le \ell-1$ ($h_0 = 0.5$), we have discretized A^1, A^2 by using

the standard difference approximations $h_k^{-2} D_{k,x}^{+} D_{k,x}^{-}$, $h_k^{-2} D_{k,y}^{+} D_{k,y}^{-}$ and $h_k^{-2} [D_{k,x}^{+} D_{k,y}^{+} + D_{k,x}^{-} D_{k,y}^{-}] / 2.0$ for the partial derivatives $\partial^2/\partial x^2$, $\partial^2/\partial y^2$ and $\partial^2/\partial x \partial y$, where $D_{k,x}^{\pm} D_{k,y}^{\pm}$ denote the forward resp. backward difference in x resp. y on the grid Ω_k. It is easy to check that these discretizations yield difference schemes of positive type (cf. [17],[28]).

Choosing the restricted exact solution as a startiterate on the coarsest grid Ω_0, we have determined an initial iterate on the finest grid Ω_ℓ by nested iteration. We have then performed several multi-grid cycles until either machine accuracy has been reached or the total number of work units has exceeded 100. Here, a work unit corresponds to a symmetric nonlinear Gauss-Seidel iteration on the finest grid. Denoting by $\| e_\ell^\nu \|_{\ell,2}$, $\nu \geq 1$, the discrete L_2 -norm of the difference $e_\ell^\nu = u_\ell^\nu - u_\ell^{\nu-1}$ of two subsequent iterates, an asymptotic convergence rate, relating the gain in accuracy to the amount of work for implementation, has been computed by means of

$$\mu_\ell = (\| e_\ell^{\nu^*} \|_{\ell,2} / \| e_\ell^1 \|_{\ell,2}) ** (1 / [(\nu^* - 1) * N_{wu}]) \tag{4.3}$$

where N_{wu} is the number of work units for performing one multi-grid cycle and ν^* denotes the last iterate before either $\| e_\ell^\nu \|_{\ell,2} < eps$ or $(\nu\text{-}1)N_{wu} > 100$. Note that all computations reported in this section have been performed on a CRAY XMP-24 where $eps = 10^{-14}$.

Figures 1 and 2 represent the asymptotic convergence rates for W -cycles with a different number of pre-smoothings and no post-smoothing (Fig. 1) and for W -cycles with a different number of pre- and post-smoothings (Fig. 2). For comparison, in both figures we have also plotted the convergence rates of the corresponding single-grid nonlinear SOR-iteration with suboptimal choice of the relaxation parameter. The plots clearly demonstrate the expected $O(1 - h_\ell^2)$ -behavior of the single-grid convergence rates while the multi-grid convergence rates seem to approach a constant value for an increasing number of grids in the hierarchy thus confirming the theoretical results derived in the preceding section. Note that the multi-grid convergence rates are considerably higher than for standard linear second order elliptic boundary-value

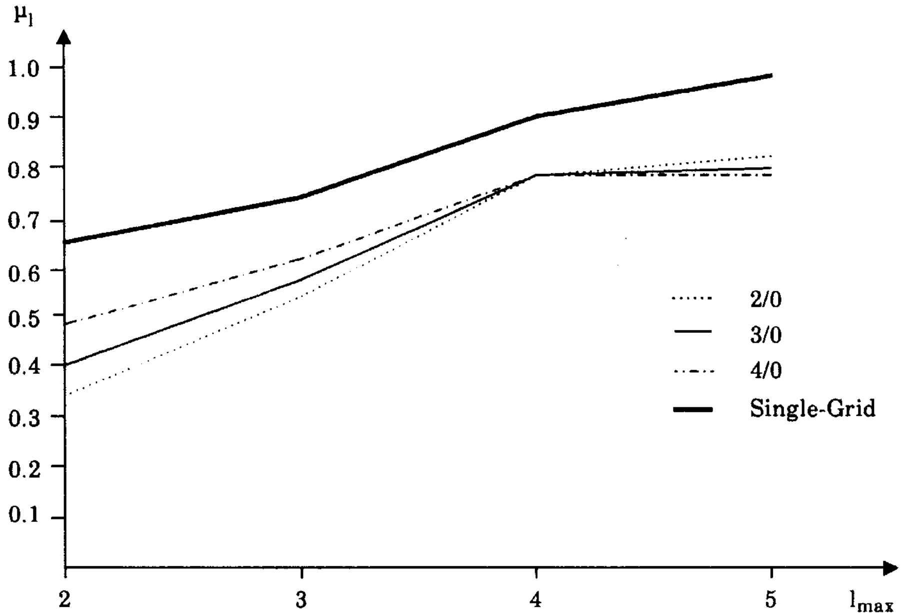

FIGURE 1

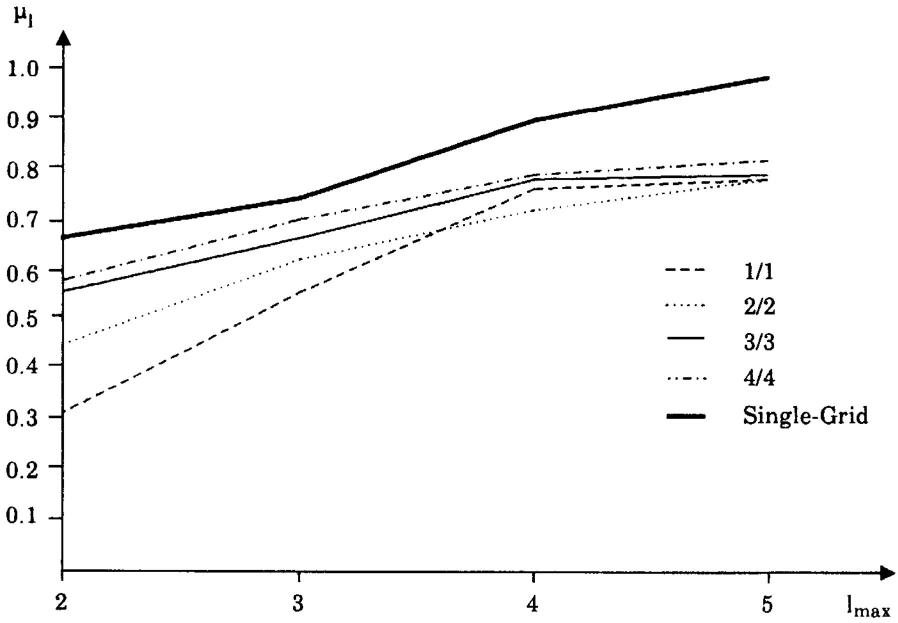

FIGURE 2

problems (cf. e.g. [16]) which is an effect commonly observed for free boundary problems. Indeed, the convergence rates are almost in the same range as those obtained for multi-grid algorithms applied to other types of free boundary problems (cf. e.g. [6],[19],[20] and [30]).

As a second, more realistic example we have dealt with the computation of the maximum utility of profits for a stochastic dynamic sales response model which has been analytically investigated by C.S. Tapiero, J. Eliashberg and Y. Wind in [32] and which represents the stochastic version of the classical deterministic Vidale-Wolfe model [33].

We consider a firm selling two products $z = (z_1,z_2)$ with market potentials $M = (M_1,M_2)$ at prices $p = (p_1,p_2)$ while spending $a = (a_1,a_2)$ for advertising. We denote by $y = (y_1,y_2)$ the market shares $y_i = z_i / M_i$, $1 \le i \le 2$, and by $m = (m_1,m_2)$ and $q = (q_1,q_2)$ the forgetting and advertising effectiveness effects, respectively. Modelling sales uncertainty by a diffusion term $\sigma(y,a) = (\sigma_{ij}(y,a))_{i,j=1}^{2}$ and including reflection constraints (i.e. continuing the process when y leaves the admissible region $\Omega = (0,1)\times(0,1)$), the market shares evolve according to the following stochastic diffusion process with reflection boundaries

$$dy = b(y,a,m,q)\,dt + \sigma(y,a)\,dw - \chi_\Gamma(y)\,n\,d\alpha \qquad (4.4a)$$

$$y(0) = x = (x_1,x_2) \in \Omega\ . \qquad (4.4b)$$

In (4.4a) the drift term $b(y,a,m,q)$ is given by

$$b_i(y,a,m,q) = -\, m_i y_i + q_i a_i (1-y_i)\ , \qquad 1 \le i \le 2 \qquad (4.5)$$

representing the deterministic Vidale-Wolfe model, $w = w(t)$ stands for a normalized two-dimensional Wiener process, χ_Γ is the characteristic function of $\Gamma = \partial\Omega$, n the outward normal at Γ and $d\alpha$ is a continuous adapted process (for details see [32]).

The admissible control set V consists of sequences $(\Theta,a) = (\Theta_\nu,a_\nu)_{\nu\in\mathbb{N}}$ of random times Θ_ν and advertising policies $a_\nu = (a_{\nu,1},a_{\nu,2})$ chosen from a finite set $A = \{a^\mu\}_{\mu=1}^{m}$. Then, denoting by $\pi(a) = \sum_{i=1}^{2} (p_i M_i y_i - a_i)$

the instantaneous profits and choosing a utility function $U = U(\pi(a))$, discounted at a constant discount factor $c > 0$, the optimal control problem is the following infinite horizon utility maximization problem

$$u(x) = \sup_{(\theta,a)\in V} E\left[\int_0^\infty U(\pi(a))\exp(-ct)\,dt\right]. \tag{4.6}$$

As utility function $U(\pi(a))$ and diffusion $\sigma(y,a)$ we have taken $U(\pi(a)) = \pi(a)$ and $\sigma(y,a) = \mathrm{diag}(\sigma_1(y,a),\sigma_2(y,a))$ with variances $\sigma_i^2(y,a) = (m_i y_i + q_i a_i (1-y_i))^{1/2}$, $1 \le i \le 2$. Moreover, we have chosen the following four admissible advertising policies

$$a^1 = (0,0)\,,\ a^2 = (a,0)\,,\ a^3 = (0,a)\,,\ a^4 = (a/2, a/2) \tag{4.7}$$

corresponding to "no advertising" (a^1) , "advertising for product 1" (a^2) , "advertising for product 2" (a^3) and "advertising for both products" (a^4) . The corresponding HJB-equation is then of type (1.5a)',(1.5b)' with

$$A^\mu = -\frac{1}{2}\sum_{i=1}^{2}\sigma_i^2(x,a^\mu)\frac{\partial^2}{\partial x_i^2} - \sum_{i=1}^{2} b_i(x,a^\mu,m,q)\frac{\partial}{\partial x_i} + c$$

$$f^\mu = \pi(a^\mu)\,,\quad \gamma^\mu \equiv 1\,,\quad 1 \le \mu \le 4\,.$$

We have discretized A^μ , $1 \le \mu \le 4$, with respect to the same hierarchy of grids Ω_k , $0 \le k \le \ell$, as in the first example. Using standard central difference quotients $h_k^{-2} D^+_{k,x_i} D^-_{k,x_i}$ for the second order derivatives $\partial^2/\partial x_i^2$ and the forward resp. backward difference quotient $h_k^{-1}\ D^+_{k,x_i}$ resp. $h_k^{-1}\ D^-_{k,x_i}$ for the first order derivatives $\partial/\partial x_i$ (according to the sign of $b_i(x,a^\mu,m,q)$ in $x \in \Omega_k$) , we get a discrete HJB-equation with coefficient matrices A_k^μ being lower semistrictly diagonally dominant M-matrices.

Providing a startiterate on the finest grid Ω_ℓ by nested iteration, we have computed the optimal utility of profits by successive application of MGHJB with W-cycle structure.

Figures 3 and 4 display the sets $\Omega_\ell^\mu(u_\ell^*)$ for $\ell = 5$ ($h_\ell = 1/64$) for the following market potentials M_i , prices p_i , sales-decay rates m_i and

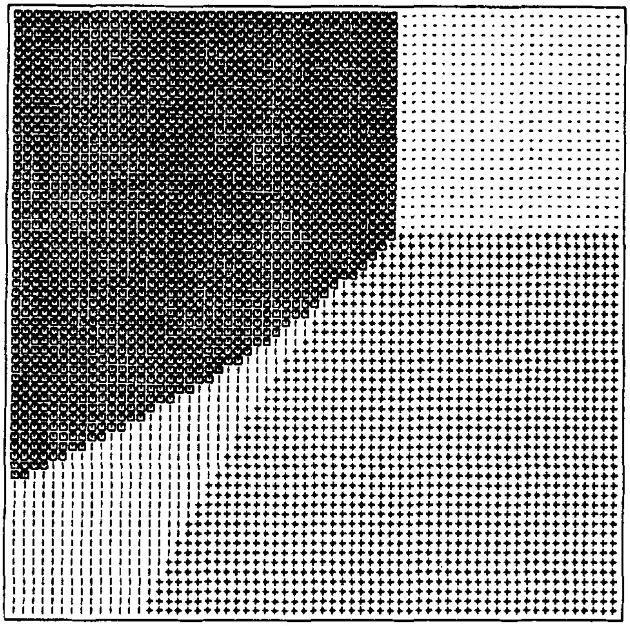

FIGURE 3

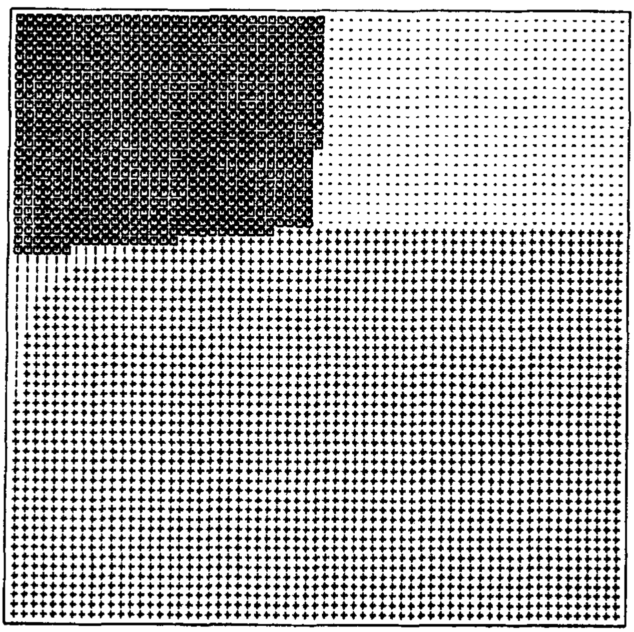

FIGURE 4

response constants q_i

(i) $M_1 = M_2 = 120\,000$, $p_1 = p_2 = 1$
$m_1 = m_2 = 0.05$, $q_1 = q_2 = 15$

(ii) $M_1 = 10\,000$, $M_2 = 120\,000$, $p_1 = 0.5$, $p_2 = 1.5$
$m_1 = 0.1$, $m_2 = 0.05$, $q_1 = 5$, $q_2 = 15$.

The constant a in (4.7) reflecting advertising expenses and the discount factor c have been chosen as $a = 1000$ and $c = 0.5$ both for (i) and (ii) .

Points $x \in \Omega_\ell$ belonging to $\Omega_\ell^1(u_\ell^*)$, $\Omega_\ell^2(u_\ell^*)$, $\Omega_\ell^3(u_\ell^*)$ and $\Omega_\ell^4(u_\ell^*)$ are marked by " – " , " ☐ " , " + " and " | " , respectively. The probabilistic interpretation is that for any initial state $x \in \Omega_\ell$ the markings tell us both the type and the value of control which has to be performed asymptotically (i.e. for $t \to \infty$). The figures display a certain risk-averse behaviour of the advertiser with essentially decreasing advertising expenses for increasing market shares.

REFERENCES

[1] A. Bensoussan, Stochastic control by functional analysis methods, North-Holland, Amsterdam, 1982

[2] A. Bensoussan and J.L. Lions, Applications of variational inequalities in stochastic control, North-Holland, Amsterdam, 1982

[3] A. Bensoussan and J.L. Lions, Impulse control and quasi-variational inequalities, Gauthier-Villars, Paris, 1984

[4] A. Berman and R.J. Plemmons, Nonnegative matrices in the mathematical sciences, Academic Press, New York, 1979

[5] A. Brandt, Multi-level adaptive solutions to boundary-value problems, Math. Comp. 31 , 1977, 333–390

[6] A. Brandt and C.W. Cryer, Multi-grid algorithms for the solution of linear complementarity problems arising from free boundary problems, SIAM J. Sci. Stat. Comput. 4 , 1983, 655–684

[7] E. Brezzi and L. Caffarelli, Convergence of the discrete free boundaries for finite element approximations, R.A.I.R.O. Analyse numérique / Numer. Anal. 17 , 1983, 385–395

[8] L. Caffarelli, J.J. Kohn, L. Nirenberg and J. Spruck, The Dirichlet problem for nonlinear second-order elliptic problems. II. Complex Monge-Ampère and uniformly elliptic equations, Comm. Pure Appl.

[9] F.H. Clarke, Optimization and nonsmooth analysis, John Wiley & Sons, New York 1983

[10] F. Conrad et Ph. Cortey-Dumont, Sur l'analyse numérique d'un problème de bifurcation dans les inéquations variationnelles, à paraître

[11] Ph. Cortey-Dumont, Sur l'analyse numérique des équations de Hamilton-Jacobi-Bellman, Math. Meth. in the Appl. Sci. 9 , 1987, 198–209

[12] M.G. Crandall and P.L. Lions, Viscosity solutions of Hamilton-Jacobi equations, Trans. Amer. Math. Soc. 277 , 1983, 1-42

[13] W.H. Fleming and R. Rishel, Deterministic and stochastic optimal control, Springer, Berlin, 1975

[14] W. Hackbusch, Convergence of multi-grid iterations applied to difference equations, Math. Comp. 34 , 1980, 425–440

[15] W. Hackbusch, On the convergence of multi-grid iterations, Beitr. Numer. Math. 9 , 1981, 213-239

[16] W. Hackbusch, Multi-grid methods and applications, Springer,

[17] R.H.W. Hoppe, Multi-grid methods for Hamilton-Jacobi-Bellman equations, Numer. Math. 49 , 1986, 239–254

[18] R.H.W. Hoppe, On the numerical solution of variational inequalities by multi-grid techniques, Proc. Int. Symp. on Numerical Analysis, Ankara, Turkey, Sept. 1–4, 1987 (eds.: B. Karasözen et al.), to appear 1988

[19] R.H.W. Hoppe, Une méthode multigrille pour la solution des problèmes d'obstacle, soumis à R.A.I.R.O., Modélisation Math. Anal. Numér.

[20] R.H.W. Hoppe and R. Kornhuber, Multi-grid solution of the two-phase Stefan problem, in: Multigrid Methods: Theory, Applications, and Supercomputing, Proc. 3rd Copper Mountain Conf., Copper Mountain, Colorado, April 6–10, 1987 (ed.: St. McCormick), Marcel Dekker, 1988, 267–297

[21] R.H.W. Hoppe and H.D. Mittelmann, A multi-grid continuation strategy for parameter-dependent variational inequalities, to appear in J. Comput. Appl. Math., 1989

[22] N.V. Krylov, Controlled diffusion processes, Springer, Berlin,

[23] N.V. Krylov, On degenerate nonlinear elliptic equations. I, Mat. Sb. 120 , 1983, 311–330

[24] N.V. Krylov, On degenerate nonlinear elliptic equations. II, Mat. Sb. 121 , 1983, 211–232

[25] N.V. Krylov, Boundedly nonhomogeneous elliptic and parabolic equations in a domain, Izv. Mat. Ser. 47 , 1983, 75–108

[26] P.L. Lions, Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. II, Commun. Partial Differ. Equations 8 , 1983, 1229–1276

[27] P.L. Lions, Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. III, in: Nonlinear Partial Differential Equations and their Applications, Collège de France Seminar, Vol. V (eds.: H. Brezis and J.L. Lions), Pitman, London, 1983

[28] P.L. Lions et B. Mercier, Approximation numérique des équations de Hamilton-Jacobi-Bellman, R.A.I.R.O. Analyse numérique / Numer. Anal. 14 , 1980, 369–393

[29] P.L. Lions and N.S. Trudinger, Linear oblique derivative problems for the uniformly elliptic Hamilton-Jacobi-Bellman equations, Math. Z. 191 , 1986, 1–15

[30] J. Mandel, A multilevel iterative method for symmetric, positive definite linear complementarity problems, Appl. Math. Optimization 11 , 1984, 77–95

[31] W.C. Rheinboldt, On M-functions and their application to nonlinear Gauss-Seidel iteration and to network flows, J. Math. Anal. Appl. 32 , 1970, 274–307

[32] C. Tapiero, J. Eliashberg and Y. Wind, Risk behaviour and optimum advertising with a stochastic dynamic sales response, Optim. Control Appl. Methods 8 , 1987, 299–304

[33] M.L. Vidale and H.B. Wolfe, An operations research study of sales response to advertising, Oper. Res. 5 , 1957, 370–381