



HAL
open science

BACKWARD DIFFERENTIATION FORMULA FINITE DIFFERENCE SCHEMES FOR DIFFUSION EQUATIONS WITH AN OBSTACLE TERM

Olivier Bokanowski, Kristian Debrabant

► **To cite this version:**

Olivier Bokanowski, Kristian Debrabant. BACKWARD DIFFERENTIATION FORMULA FINITE DIFFERENCE SCHEMES FOR DIFFUSION EQUATIONS WITH AN OBSTACLE TERM. 2019. hal-01686742v1

HAL Id: hal-01686742

<https://inria.hal.science/hal-01686742v1>

Preprint submitted on 18 Jan 2018 (v1), last revised 6 Mar 2020 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HIGH ORDER FINITE DIFFERENCE SCHEMES FOR SOME NONLINEAR DIFFUSION EQUATIONS WITH AN OBSTACLE TERM

OLIVIER BOKANOWSKI AND KRISTIAN DEBRABANT

ABSTRACT. Old and novel finite difference schemes, using Backward Differentiation Formula (BDF), are studied for the approximation of one-dimensional non-linear diffusion equations with an obstacle term, of the form

$$\min(v_t - a(t, x)v_{xx} + b(t, x)v_x + r(t, x)v, v - \varphi(t, x)) = f(t, x).$$

A new inconditionnal stability result is obtained for one of the scheme, second order consistent in both space and time. Our study consider in particular the “generic” case when v_{xx} is bounded but $x \rightarrow v_{xx}(t, x)$ may have isolated discontinuities. Numerical examples show second order convergence in both space and time, inconditionnally on the ratio of the mesh steps. An L^2 error estimate of order $\frac{1}{2}$ is furthermore obtained. Application to the American option problem in mathematical finance is given and used throughout the paper. Also a Crank-Nicolson finite difference scheme is revisited to better explain its behavior, that may switch from second order to first order, depending on the mesh parameters. In the analysis, an equivalence of the obstacle equation with an Hamilton-Jacobi-Bellman equation is also mentionned in the case when there is no time dependency in the coefficients. We also consider two academic problems with explicit solutions for parabolic equations with an obstacle term, in order to study the relevance of the proposed schemes.

Keywords: diffusion equation, obstacle equation, viscosity solution, numerical methods, finite difference scheme, Crank Nicolson scheme, Backward Differentiation Formula, high order schemes.

1. INTRODUCTION

We consider a second order partial differential equation with an obstacle term, of the following form:

$$\min(v_t + \mathcal{A}v, v - \varphi(t, x)) = f(t, x), \quad t \in (0, T), \quad x \in \Omega, \quad (1a)$$

$$v(0, x) = v_0(x), \quad x \in \Omega. \quad (1b)$$

with

$$\mathcal{A}v := -\frac{1}{2}\sigma^2(t, x)v_{xx} + b(t, x)v_x + r(t, x)v \quad (2)$$

where b, r are continuous and bounded functions, σ is bounded, and Lipschitz continuous in x (uniformly with respect to t). The functions φ, f , and v_0 are Lipschitz continuous, with $v_0(x) \geq \varphi(0, x) + f(0, x)$ for compatibility reasons with (1a). The set Ω can be either $\Omega = \mathbb{R}$, or $\Omega = (X_{min}, X_{max})$. In the second case we will consider (1) complemented by boundary conditions of Dirichlet type:

$$v(t, X_{min}) = v_\ell(t), \quad t \in (0, T) \quad (3a)$$

$$v(t, X_{max}) = v_r(t), \quad t \in (0, T) \quad (3b)$$

The solution u can be defined as the unique continuous solution of (1) when $\Omega = \mathbb{R}$ (or (1)-(3) when $\Omega = (X_{min}, X_{max})$) in the viscosity sense (see [7]). A variational framework can also be used [11].

In the recent years there has been a lot of interest in the approximation of such obstacle problems. Related to optimal stochastic stopping time problems, we will consider in particular

$$\min(v_t - \frac{1}{2}\lambda^2 x^2 v_{xx} - rxv_x + rv, v - \varphi(x)) = 0 \quad t \in (0, T), \quad x \in \Omega, \quad (4a)$$

$$v(0, x) = \varphi(x) \quad x \in \Omega, \quad (4b)$$

with $\Omega = (0, \infty)$, with constant coefficients $\lambda > 0$, $r > 0$, $f = 0$, and with initial data identical to the obstacle function. The American put option problem in mathematical finance corresponds in particular to the case of the initial data (or "payoff" function) $\varphi(x) := \max(K - x, 0)$. For this problem it is known that the solution presents a singular point $x_s(t)$ moving with time, such that $v(t, x) = \varphi(x)$ for $x \leq x_s(t)$, and $v(t, x) > \varphi(x)$ for $x > x_s(t)$, and $t \rightarrow x_s(t)$ has a Hölder continuity behavior near $t = 0$ (see Remark 4.7 as well as [9], [2], and [1, Chap 6]). Some results on the structure of the interface related to (1) can also be found in [4] (and see related references).

A finite element scheme in the American option setting has been considered by Jaillet, Lamberton and Lapeyre in [15], where its convergence is also proved under conditions on the mesh steps. For a comprehensive study of finite difference schemes as well as finite element schemes in this context, we refer to Achdou and Pironneau [1]. In relation with the obstacle problem, a finite volume method is also studied in Berton and Eymard [3]. In connection with the present work, Oosterlee proposed in [18] to use second order backwards differencing formula (BDF2) schemes in the context of the American option problem, in combination with a multigrid approach (see also K. Oosterlee et al.[19]).

Now we remark that in the case of $v_0 \equiv \varphi$ (with $\Omega = \mathbb{R}$), and for an operator of the form

$$\mathcal{A}v := -\frac{1}{2}\sigma^2(x)v_{xx} + b(x)v_x + r(x)v, \quad (5)$$

i.e., with coefficients and source term $f = f(x)$ which do not depend on time, then the solution v or (1) is also the unique viscosity solution of the following Hamilton-Jacobi (HJ) equation:

$$v_t + \min(0, \mathcal{A}v) = f(x), \quad t \in (0, T), \quad x \in \Omega, \quad (6a)$$

$$v(0, x) = \varphi(x) + f(x), \quad x \in \Omega. \quad (6b)$$

The equivalence between (1) and (6) was signaled to us by R. Eymard. It was proved by C. Martini in [17]. A sketch of the proof of independent interest is given in Appendix B (see Remark B.1).

In this paper, in section 2, we first revisit some Crank-Nicolson schemes adapted to the obstacle problem (1). we consider two types of Crank-Nicolson schemes, that only differ in the way to treat the obstacle term. We explain why they may show second order convergence results even if a naive setting of the scheme is only first-order consistent with the PDE (1). Indeed we show that the scheme may be second order consistent with the PDE (6) depending on the mesh parameters. However we numerically observe that even fully second order consistent version of the CN scheme does not gives second order numerical convergence for bad ratio of the mesh parameters (corresponding to "high CFL numbers"). Stability results exists for the Crank-Nicolson scheme, in the L^∞ setting, for the approximation of the linear heat equation[20]. However, to the best of our knowledge, there is no convergence proof for the Crank-Nicolson finite difference schemes adapted to the nonlinear equation (1).

To circumvent these problems, we consider the use of the Backward Difference Formula (BDF) for the approximation of the time derivative v_t , adapted to the obstacle problem. In Section 2, we introduce implicit BDF schemes in the same way as Oosterlee [18]. In particular second- and third-order BDF schemes are considered. When formulated on the nonlinear PDE (1), these schemes are non-linear and implicit, but they can be solved by using a simple Newton-type algorithm. More importantly, a new unconditionnal L^2 stability estimate is obtained in the case of the second order BDF obstacle scheme (BDF2). This is obtained by using estimates similar to the ‘‘Gear’’ scheme for the linear case (for instance, [1]) with some new ingredients in order to deal with the non-linearity coming from the obstacle term. We then obtain also a new error estimate in an L^2 norm which is of order $(\tau + h)^{1/2}$ where τ and h are the time and space mesh steps. This estimate holds under some specific assumptions on the regularity on the exact solution v which allows v_{xx} bounded but possibly discontinuous at some finite number of singular points $(t, y_j(t))_{1 \leq j \leq p}$ that do not evolve too rapidly (see section 4.2).

Two academic models are also introduced, with explicit solutions, yet very close to the American option model. These models allow us to study precisely and more easily the numerical convergence and allows for a slightly smoother behavior of the interface (compared to the American option model). This allows also to observe third order behavior for a BDF3 scheme on a specific model with bounded u_{xxx} derivative at the free boundary.

We also shortly report, in Appendix D, on an other attempt where the whole term $u_t + \mathcal{A}u$ is approximated by a BDF approach. We are then led to a new scheme that requires only an implicit treatment of linear terms (for the approximation of the diffusion part) and an explicit treatment of the obstacle term. Even if we have no stability proof for these schemes, numerical results are given, showing good behavior of the schemes (stability and expected orders).

Our study concerns here only one-dimensional obstacle problems, but the proposed schemes based on BDF approximations can be extended to higher dimensions.

2. CRANK-NICOLSON FINITE DIFFERENCE SCHEMES REVISITED

In this section we revisit some Crank-Nicolson schemes and related approaches for a diffusion equation in presence of an obstacle term, and show some numerical difficulties that may arise when considering such schemes with ‘‘high CFL numbers’’.

Let us consider a domain $\Omega = [X_{min}, X_{max}]$ and a uniform mesh with $J \geq 1$ points inside:

$$x_j = X_{min} + jh, \quad j = 0, \dots, J + 1,$$

where $h := \frac{X_{max} - X_{min}}{J+1}$. Let $N \geq 1$, $\tau = \frac{T}{N}$ and $t_n = n\tau$.

We shall say that we have a ‘‘high CFL number’’ when $\frac{\tau}{h} \gg 1$ (or $\frac{J}{N} \gg 1$), compared to a situation where $\frac{\tau}{h} \simeq 1$ (or $\frac{J}{N} \simeq 1$).

Denoting u_j^n an approximation of $v(t_n, x_j)$, we consider a centered finite difference approximation for the operator $\mathcal{A}v$:

$$\begin{aligned} (\mathcal{A}v)(t_n, x_j) &\simeq \\ \frac{1}{2}\sigma^2(t_n, x_j) &\left(\frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{h^2} \right) + b(t_n, x_j) \frac{u_{j+1}^n - u_{j-1}^n}{2h} + r(t_n, x_j)u_j^n \end{aligned} \quad (7)$$

For simplicity we assume there is no time dependency in the coefficients of \mathcal{A} . The diffusion part will always dominate the advection part ($\frac{1}{2}\frac{\sigma^2}{h^2} \geq \frac{|b|}{2h}$) to avoid stability issues with the centered approximation. Note that for the American put option problem this requires $h \leq x_1\sigma^2/r$.

Let us denote by $Au^n + q(t_n)$ the approximation of $\mathcal{A}v(t_n, \cdot)$ on a given set of grid points, with $u^n = (u_1^n, \dots, u_J^n)^T$ and with given Dirichlet boundary conditions,

for $n = 0, \dots, N$:

$$u_0^n = v_\ell(t_n) \quad \text{and} \quad u_{J+1}^n = v_r(t_n). \quad (8)$$

Note that the matrix A is also in general time-dependent, but we will remove this time-dependency for simplicity of presentation. The vector $q(t_n)$ may depend of the time also because of the time dependency in the boundary conditions (8). We have second order consistency in space, that is, denoting $v_i^n := v(t_n, x_i)$,

$$(Av^n + q(t_n))_j = (Av)(t_n, x_j) + O(h^2).$$

A first simple Crank Nicolson (CN) scheme is, for the obstacle equation (1), assuming $f \equiv 0$, for $n = 0, \dots, N - 1$,

$$\mathcal{S}_j^{1,n} := \min \left(\frac{u_j^{n+1} - u_j^n}{\tau} + \frac{1}{2}(Au^{n+1} + Au^n)_j + q(t_{n+1/2})_j, u_j^{n+1} - g_j \right) = 0 \quad 1 \leq j \leq J \quad (9)$$

where we have denoted

$$g_j := \varphi(x_j).$$

together with boundary conditions (8), and initialized with

$$u_j^0 := v_0(x_j), \quad 1 \leq j \leq J. \quad (10)$$

Looking now at equation (6) (assuming again $f \equiv 0$), an other possible CN scheme is

$$\mathcal{S}_j^{2,n} := \min \left(\frac{u_j^{n+1} - u_j^n}{\tau} + \frac{1}{2}(Au^{n+1} + Au^n)_j + q(t_{n+1/2})_j, \frac{u_j^{n+1} - u_j^n}{\tau} \right) = 0 \quad 1 \leq j \leq J \quad (11)$$

initialized with $u_j^0 := v_0(x_j)$. Because $\tau > 0$, this scheme is also equivalent to

$$\min \left(\frac{u_j^{n+1} - u_j^n}{\tau} + \frac{1}{2}(Au^{n+1} + Au^n)_j + q(t_{n+1/2})_j, u_j^{n+1} - u_j^n \right) = 0 \quad 1 \leq j \leq J. \quad (12)$$

Remark 2.1. For both schemes, the unknown u^{n+1} can be obtained by using fix point methods (there exists a unique solution of the obstacle problem $\min(Bx - b, x - g) = 0$ as soon as, for instance, B is diagonally dominant with $B_{ii} > 0$ (see [1]).

Remark 2.2. In the case when furthermore B is an M -matrix in the sense that $B_{ii} > \sum_{j \neq i} |B_{ij}|$, $\forall i$, and $B_{ij} \leq 0$, $\forall i \neq j$, then a Newton-like algorithm [5] can be implemented, which is particularly efficient for solving obstacle problems exactly (up to machine precision) in a few number of iterations. We refer also to [14] for convergence of semi-smooth Newton methods or related algorithms applied to solve discretized PDE obstacle problems.

Lemma 2.3. (i) *The scheme (9) is (only) first order consistent in time, second order in space.*

(ii) *The scheme (11) is second order consistent in time and space.*

Proof. (i) For the linear part $u_t + Au$, the CN scheme is second order consistent only precisely at time $t_{n+1/2}$, while the obstacle term $v_j^{n+1} - g_j$ is evaluated at time t_{n+1} so is only first order consistent with the value $v(t_{n+1/2}, x_j) - g_j$. So the consistency error is of order $O(\tau) + O(h^2)$ but no better.

(ii) Assuming that v is regular, and denoting $v_j^n := v(t_n, x_j)$:

$$\frac{v_j^{n+1} - v_j^n}{\tau} = v_t(t_{n+1/2}, x_j) + O(\tau^2 \|v_{3t}\|_\infty),$$

and

$$\begin{aligned} & \frac{v_j^{n+1} - v_j^n}{\tau} + \frac{1}{2}(Av^{n+1} + Av^n)_j + q(t_{n+1/2})_j \\ & = (v_t + \mathcal{A}v)(t_{n+1/2}, x_j) + O(\tau^2 \|v_{3t}\|_\infty) + O(h^2(\|v_{3x}\|_\infty + \|v_{4x}\|_\infty)). \end{aligned}$$

Hence we have the following second order estimate

$$S_j^{2,n}(v) = \min(v_t + \mathcal{A}v, v_t)(t_{n+1/2}, x_j) + O(\tau^2 + h^2).$$

□

However, the following assertions hold:

Lemma 2.4. (i) If $u^{n+1} \geq u^n$ for the scheme (9) (with $u_j^n \geq g_j$), then u^{n+1} is also solution of the scheme (11) (hence both schemes give identical values).

(ii) This is the case in particular when the following conditions hold:

- the vector $q(t) = q$ does not depend on time (that is the case of the American put option (4) with $\varphi(x) := \max(K - x, 0)$)
- the matrix $I - \frac{\tau}{2}A$ is positive componentwise (which is the case under the CFL condition $\frac{2}{\max_{j=1}^J \sigma^2(x_j)} \frac{\tau}{h^2} \leq 1$)
- the matrix $I + \frac{\tau}{2}A$ is an M -matrix (which is the case here).

The proof of this Lemma is given in appendix C.

This explains why the apparently first order consistent CN scheme (9) can give second order results for low CFL number, since then the scheme is equivalent to a second order consistent scheme with respect to the second equivalent PDE (6).

In order to verify the expected orders, we have tested the CN schemes numerically on the American put option model with initial data

$$\varphi(x) := \max(K - x, 0) \tag{13a}$$

and parameters

$$\lambda = 0.2, \quad r = 0.1, \quad T = 1, \quad K := 100. \tag{13b}$$

The errors were computed on a discrete set of 13 reference values in [80, 110], spaced by 2.5 (this interval also contains the singular point of the solution u). These reference values are computed using a BDF obstacle scheme of third order and with $N = J = 10240$ and that will be made precise in section 3. Details on domain and boundary conditions are given in Appendix A.

Results are given in Table 1 with discretization parameters $J = N$, and in Table 2 with discretization parameters $N = J/10$ (this second case corresponds to higher CFL numbers). Note that the quite restrictive CFL condition in part (ii) of Lemma 2.4 is not fulfilled.

Mesh		Error L^1		Error L^2		Error L^∞		CPU time
J	N	error	order	error	order	error	order	
80	80	1.74E-02	-	2.16E-02	-	4.49E-02	-	0.14
160	160	3.15E-03	2.47	3.73E-03	2.54	5.25E-03	3.10	0.25
320	320	8.42E-04	1.91	9.87E-04	1.92	1.39E-03	1.92	0.52
640	640	2.23E-04	1.92	2.58E-04	1.93	3.56E-04	1.97	1.54
1280	1280	6.06E-05	1.88	6.96E-05	1.89	9.15E-05	1.96	6.65
2560	2560	1.66E-05	1.86	1.91E-05	1.86	2.69E-05	1.77	37.66

TABLE 1. Crank-Nicolson scheme (9) for (1) and $N = J$. The Crank-Nicolson scheme (11) for (6) gives same error table in this case.

However, for lower N values (higher CFL numbers) the CN scheme is no more second order and goes back to first order behavior. This is illustrated in Table 2. In particular we observe that the pointwise inequality $u^{n+1} \geq u^n$ is no more true (due to the fact that the amplification matrix $(I + \frac{\tau}{2}A)^{-1}(I - \frac{\tau}{2}A)$ does not have only positive coefficients anymore).

Now, taking the obstacle function to be u^n instead of $g_j = (\varphi(s_j))$, hence solving the scheme (11), enforces that $u^{n+1} \geq u^n$. Results for the scheme (11) are similar that for scheme (9) for low CFL numbers (both schemes give identical values for the case $N = J$, here). However, for higher CFL numbers, results obtained with the scheme (11) differ and are given in Table 3. Although not second order, convergence is slightly improved for the L^2 and L^∞ norms. (We believe that the loss of second order is due to a loss of stability properties combined with a reduced consistency error near the singularity which affects the scheme both at t_n and t_{n+1} .)

Mesh		Error L^1		Error L^2		Error L^∞		time(s)
J	N	error	order	error	order	error	order	
80	8	1.83E-02	-	2.47E-02	-	4.90E-02	-	0.05
160	16	4.67E-03	1.97	7.23E-03	1.77	2.18E-02	1.17	0.07
320	32	1.68E-03	1.48	2.97E-03	1.28	9.90E-03	1.14	0.13
640	64	6.89E-04	1.28	1.38E-03	1.11	4.75E-03	1.06	0.38
1280	128	2.99E-04	1.20	6.59E-04	1.06	2.31E-03	1.04	1.99
2560	256	1.37E-04	1.13	3.22E-04	1.03	1.14E-03	1.02	16.04

TABLE 2. Crank-Nicolson scheme (9) for (1), with higher CFL number

Mesh		Error L^1		Error L^2		Error L^∞		time(s)
J	N	error	order	error	order	error	order	
80	8	2.68E-02	-	3.11E-02	-	4.90E-02	-	0.25
160	16	6.91E-03	1.96	8.28E-03	1.91	1.14E-02	2.11	0.13
320	32	2.53E-03	1.45	2.97E-03	1.48	3.91E-03	1.54	0.18
640	64	9.35E-04	1.43	1.11E-03	1.43	1.50E-03	1.38	0.41
1280	128	3.74E-04	1.32	4.44E-04	1.32	6.07E-04	1.31	1.66
2560	256	1.56E-04	1.26	1.86E-04	1.26	2.56E-04	1.25	11.36

TABLE 3. Crank-Nicolson scheme (11) for (6), with higher CFL number

3. BDF OBSTACLE SCHEMES

We now consider BDF type approximations for the first derivative u_t , leading to implicit schemes. We propose two implicit schemes (BDF2 and BDF3) which have the same complexity of the previous CN implicit schemes but that gives improved numerical results with respect to precision and to stability. Furthermore a stability and error analysis will be carried out for the BDF2 scheme.

3.1. BDF2 obstacle scheme. Our first scheme is therefore the following two-step implicit scheme (here after also referred as BDF2 obstacle scheme), for $n \geq 1$:

$$\mathcal{H}_j^{n+1}(u) := \min \left(\frac{3u_j^{n+1} - 4u_j^n + u_j^{n-1}}{2\tau} + (Au^{n+1} + q(t_{n+1}))_j, u_j^{n+1} - g_j \right) - f_j^{n+1} = 0$$

where $f_j^{n+1} := f(t_{n+1}, x_j)$, and initialized with appropriate u^0 and u^1 values (unless otherwise precised, u^1 is computed by an implicit Euler scheme). Such approximations for the linear term u_t , known as BDF approximations, are well known and used in various contexts [8, 12].

The use of a BDF scheme for a diffusion plus obstacle problem is not new (Oosterlee et al [18, 19], the idea was also suggested by Seydel in [21, see page 187 and 217]). To the best of our knowledge, a precise analysis of the scheme was missing.

By construction the scheme has the following consistency error, when v is regular, for $v_j^n = v(t_n, x_j)$:

$$\begin{aligned} \mathcal{H}_j^{n+1}(v) &= \min(v_t + \mathcal{A}v, v - \varphi)(t_{n+1}, x_j) - f(t_{n+1}, x_j) \\ &\quad + O(\tau^2 \|v_{3t}\|_\infty) + O(h^2(\|v_{3x}\|_\infty + \|v_{4x}\|_\infty)). \end{aligned} \quad (14)$$

Hence the scheme is second order consistent in time (and here also with space) with respect to the obstacle problem (1).

This consistency error justifies the introduction of BDF schemes, that precisely approximate $u_t + \mathcal{A}u$ at time t_{n+1} without the need of other particular requirements (it does not need that $v_t + \mathcal{A}v = 0$ at previous times $t < t_{n+1}$, which would not hold in the presence of an obstacle term).

Let $\min(X, Y) := (\min(x_j, y_j))_j$ denote the minimum of two vectors $X = (x_j), Y = (y_j)$ of \mathbb{R}^J . For convenience, the scheme will also be written as follows:

$$\min \left((I_d + \frac{2}{3}\tau A) u^{n+1} - \frac{4}{3}u^n + \frac{1}{3}u^{n-1} + \frac{2}{3}\tau q(t_{n+1}) - \tau f^{n+1}, u^{n+1} - g - f^{n+1} \right) = 0 \quad (15)$$

(After subtracting by f^{n+1} , a multiplication by $\frac{2\tau}{3} > 0$ of the left part of the min term does not change the equation.)

Remark 3.1 (Newton's method.). As already mentioned before, the scheme can be solved by a semi-smooth Newton's method. More precisely, denoting $B := I_d + \frac{2}{3}\tau A$ (a real valued $J \times J$ matrix) and $b := \frac{4}{3}u^n - \frac{1}{3}u^{n-1} - \frac{2}{3}\tau q(t_{n+1})$, the problem is to solve for $x \in \mathbb{R}^J$

$$\min(Bx - b, x - g) = 0 \quad \text{in } \mathbb{R}^J. \quad (16)$$

The matrix B satisfies the conditions of Remark (2.2) ensuring the convergence of Newton's algorithm.

Results for the BDF second order obstacle scheme (or BDF-2), are given in Table 4, with $N = J$, and in Table 5, with $N = J/10$ (high CFL numbers), using the same parameters as (13). These results show robustness of the scheme even with high CFL numbers, and also an improvement of the convergence with respect to the CN schemes (the order is closer to 2 even for high CFL numbers).

Mesh		Error L^1		Error L^2		Error L^∞		time(s)
J	N	error	order	error	order	error	order	
80	80	1.72E-02	-	2.14E-02	-	4.48E-02	-	0.13
160	160	3.02E-03	2.51	3.57E-03	2.58	5.06E-03	3.15	0.26
320	320	7.93E-04	1.93	9.30E-04	1.94	1.32E-03	1.93	0.49
640	640	2.06E-04	1.94	2.40E-04	1.96	3.39E-04	1.97	1.54
1280	1280	5.46E-05	1.92	6.30E-05	1.93	8.75E-05	1.95	6.43
2560	2560	1.44E-05	1.93	1.67E-05	1.92	2.64E-05	1.73	36.70

TABLE 4. BDF2 scheme for Eq.(1).

Mesh		Error L^1		Error L^2		Error L^∞		time(s)
J	N	error	order	error	order	error	order	
80	8	8.23E-03	-	1.25E-02	-	3.59E-02	-	0.06
160	16	9.61E-04	3.10	1.28E-03	3.29	2.20E-03	4.02	0.09
320	32	4.18E-04	1.20	5.41E-04	1.24	8.83E-04	1.32	0.13
640	64	1.54E-04	1.44	1.92E-04	1.49	2.99E-04	1.56	0.32
1280	128	4.77E-05	1.69	5.82E-05	1.73	8.71E-05	1.78	1.14
2560	256	1.14E-05	2.07	1.39E-05	2.06	2.01E-05	2.12	6.21

TABLE 5. BDF2 scheme for Eq.(1), with high CFL numbers.

3.2. BDF3 obstacle scheme. In the same way, we propose the following three-step (BDF3) implicit scheme, for $n \geq 2$:

$$\mathcal{H}_j^{n+1}(u) := \min \left(\frac{\frac{11}{6}u_j^{n+1} - 3u_j^n + \frac{3}{2}u_j^{n-1} - \frac{1}{3}u_j^{n-2}}{\tau} + (Au^{n+1} + q(t_{n+1}))_j, u_j^{n+1} - g_j \right) - f_j^{n+1} = 0$$

The scheme may be initialized by any second order approximation for the first two steps u^1 and u^2 (such as a Crank Nicolson scheme, which we have used here).

As we have done for the BDF2 scheme, we can multiply the left term by 6τ , define $B := 11 Id + 6\tau A$, and obtain then an equivalent scheme in the following form, for $n \geq 2$, in \mathbb{R}^J :

$$\min \left(Bu^{n+1} - 18u^n + 9u^{n-1} - 2u^{n-2} + 6\tau q(t_{n+1}) - \tau f^{n+1}, u^{n+1} - g - f^{n+1} \right) = 0.$$

The unknown u^{n+1} can be solved by using again a semi-smooth Newton's method.

For sake of completeness, numerical results with BDF3 scheme are presented in Table 6 on the same example. As expected, they do not show improvement with respect to the BDF2 scheme, since the spatial approximation is only second order, but at least this shows stability of the scheme. Moreover, the typical regularity for such obstacle problems is to have u_{xx} bounded but possibly discontinuous, in particular at the interface when there is a switch between a region where $v(t, x) - \varphi(x) = 0$ to a region where $v_t + \mathcal{A}v = 0$. Therefore this limits accuracy, in principle, to second order at best.

The performance of a BDF3 obstacle scheme (using a 4th order approximation in space) will be tested in section 5.3 on a model problem with a bounded u_{3x} derivative, showing third order in that case.

Mesh		Error L^1		Error L^2		Error L^∞		time(s)
J	N	error	order	error	order	error	order	
80	80	1.74E-02	-	2.16E-02	-	4.49E-02	-	0.29
160	160	3.16E-03	2.46	3.73E-03	2.53	5.26E-03	3.09	0.28
320	320	8.52E-04	1.89	9.99E-04	1.90	1.40E-03	1.91	0.55
640	640	2.24E-04	1.93	2.60E-04	1.94	3.57E-04	1.97	1.60
1280	1280	6.07E-05	1.88	6.97E-05	1.90	9.15E-05	1.96	6.30
2560	2560	1.64E-05	1.89	1.88E-05	1.89	2.68E-05	1.77	37.19

TABLE 6. BDF3 scheme for (1).

4. STABILITY AND ERROR ESTIMATE FOR THE BDF2 SCHEME

Throughout this section, we will consider the following assumptions:

Assumption (A1):

- $a \equiv \frac{1}{2}\sigma^2$, b and r are bounded functions.
- there exists $\eta_0 > 0$, $a(t, x) \geq \eta_0 > 0$ for all t, x
- a is Lipschitz continuous in x uniformly w.r.t. t , that is:

$$\exists L \geq 0, |a(t, x) - a(t, y)| \leq L|x - y|, \quad t \in (0, T), (x, y) \in \Omega^2. \quad (17)$$

4.1. Stability estimate. Let us first start by considering an abstract obstacle problem of the form

$$\min(Bx - b, x - g) = 0 \quad \text{for } x \in \mathbb{R}^J,$$

where B is a square matrix of size J and b, g are given vectors of \mathbb{R}^J . We will use the following elementary result.

Lemma 4.1. *For any matrix B , the following equivalence holds:*

$$\min(Bx - b, x - g) = 0 \Leftrightarrow x \geq g \text{ and } \left(\langle Bx - b, v - x \rangle \geq 0, \forall v \geq g \right) \quad (18)$$

Proof. It is known [6] that if B is a positive definite symmetric matrix, the following equivalences hold:

$$\min(Bx - b, x - g) = 0 \Leftrightarrow x \text{ solves } \min_{x \geq g} \frac{1}{2} \langle x, Bx \rangle - \langle b, x \rangle \quad (19)$$

$$\Leftrightarrow x \geq g \text{ and } \left(\langle Bx - b, v - x \rangle \geq 0, \forall v \geq g \right) \quad (20)$$

When B is not symmetric, the equivalence between the min equation and (20) is still true. \Rightarrow : For $v \geq g$, $\langle Bx - b, v - x \rangle = \langle Bx - b, v - g \rangle + \underbrace{\langle Bx - b, g - x \rangle}_{=0}$ so is

positive since $Bx - b \geq 0$ and $v - g \geq 0$.

\Leftarrow : By taking $v = x + \lambda e_j$ with $\lambda \rightarrow +\infty$ we get $(Bx - b)_j \geq 0$, hence $Bx - b \geq 0$. Then, $\langle Bx - b, x - g \rangle \geq 0$, and also $\langle Bx - b, x - g \rangle \leq 0$ by taking $v = g$ as a test function in the inequality. Hence $\langle Bx - b, x - g \rangle = 0$. Together with $Bx - b \geq 0$, $x - g \geq 0$, this implies that $\min(Bx - b, x - g) = 0$. \square

The idea now is to use the inequality of Lemma 4.1 in order to obtain a stability estimate. For the linear case, it is possible for the Gear scheme - or BDF2 scheme - to obtain stability estimates in the L^2 norm (see for instance [10]). We are going to obtain similar estimates for the obstacle problem.

The scheme (15) is

$$\min \left(\left(I + \frac{2\tau}{3} A \right) u^{n+1} - \frac{4}{3} u^n + \frac{1}{3} u^{n-1} + \frac{2\tau}{3} q_{n+1}, u^{n+1} - g \right) = 0. \quad (21)$$

Let $v(t, x)$ be a regular enough function, $v_i^n := v(t_n, x_j)$, and $\bar{\epsilon}^n \in \mathbb{R}^J$ be such that

$$\frac{1}{2\tau} (3v^{n+1} - 4v^n + v^{n-1}) + Av^{n+1} + q_{n+1} = ((v_t + \mathcal{A}v)(t_{n+1}, x_j))_{1 \leq i \leq I} + \bar{\epsilon}^n. \quad (22)$$

The term $\bar{\epsilon}^n$ corresponds to a ‘‘consistency error’’ for the linear part of the PDE, here written in discrete form on the grid mesh. Our only assumption here is that (22) makes sense, which is the case if for instance there exists left or right limits for $(v_t + \mathcal{A}v)(t_{n+1}, x_j^\pm)$ so that $\bar{\epsilon}_j^n$ be well defined. Then we have

$$\min \left(\frac{1}{2\tau} (3v^{n+1} - 4v^n + v^{n-1}) + Av^{n+1} + q_{n+1} - \bar{\epsilon}^n, v^{n+1} - g \right) = 0. \quad (23)$$

Therefore v^n satisfies a perturbed scheme, as follows:

$$\min \left(\left(I + \frac{2\tau}{3}A \right) v^{n+1} - \frac{4}{3}v^n + \frac{1}{3}v^{n-1} + \frac{2\tau}{3}q_{n+1} - \frac{2\tau}{3}\bar{\epsilon}^n, v^{n+1} - g \right) = 0. \quad (24)$$

Remark 4.2. Typically $\bar{\epsilon}^n$ is of order $O(\tau^2 + h^2)$ where v is regular.

Our aim is now to show a stability estimate in order to control the error $\|u^n - v^n\|_2^2$ in terms of $\sum_{1 \leq k \leq n-1} \tau \|\bar{\epsilon}^k\|^2$.

For a vector $x = (x_j)_{1 \leq j \leq J}$, let

$$N(x) := \left(\sum_{j=1}^{J+1} |x_j - x_{j-1}|^2 \right)^{1/2} \quad (25)$$

(with the convention $x_0 := 0$ and $x_{J+1} := 0$).

We will suppose that the following coercivity assumption holds: there exist constants $\eta > 0$, $\gamma \geq 0$ such that for all $e \in \mathbb{R}^J$:

$$\langle e, Ae \rangle \geq \eta N(e/h)^2 - \gamma \|e\|_2^2. \quad (26)$$

Lemma 4.3. *Under assumption (A1), there exists $\eta, \gamma \geq 0$ such that (26) holds true.*

Proof. Considering $\mathcal{A}u = -a(t, x)u_{xx} + b(t, x)u_x + r(t, x)u$ and hereafter without explicitly mentioning the time variable, it holds

$$A = \frac{1}{h^2} \text{tridiag}(-a_i, 2a_i, -a_i) + \frac{1}{2h} \text{tridiag}(-b_i, 0, b_i) + \text{diag}(r_i)$$

where $a_i = a(x_i)$, $b_i = b(x_i)$ and $r_i = r(x_i)$. By straightforward calculations,

$$\langle e, Ae \rangle = \frac{1}{h^2} \sum_{i=1}^{J+1} (a_i e_i - a_{i-1} e_{i-1})(e_i - e_{i-1}) + \frac{1}{2h} \sum_{i=1}^J b_i (e_{i+1} - e_{i-1}) e_i + \sum_{i=1}^J r_i e_i^2 \quad (27)$$

Now we make use of $|a_i - a_{i-1}| \leq Ch$ for some constant $C \geq 0$ (since $a(\cdot)$ is Lipschitz continuous), and $a_i \geq \eta_0$, to obtain:

$$\begin{aligned} \frac{1}{h^2} \sum_{i=1}^{J+1} (a_i e_i - a_{i-1} e_{i-1})(e_i - e_{i-1}) &\geq \frac{1}{h^2} \sum_{i=1}^{J+1} \eta_0 (e_i - e_{i-1})^2 - \frac{1}{h} \sum_{i=1}^J C |e_i| |e_i - e_{i-1}| \\ &\geq \eta_0 N(e/h)^2 - C \|e\|_2 N(e/h) \end{aligned} \quad (28)$$

We have also, by using $e_{i+1} - e_{i-1} = (e_{i+1} - e_i) + (e_i - e_{i-1})$:

$$\sum_{i=1}^J |b_i (e_{i+1} - e_{i-1}) e_i| \leq \|b\|_\infty 2N(e) \|e\|_2.$$

Hence a lower bound in form:

$$\langle e, Ae \rangle \geq \eta_0 N(e/h)^2 - (C + \|b\|_\infty) N(e/h) \|e\|_2 - C \|e\|_2^2$$

for some constant C . Denoting $C_1 := C + \|b\|_\infty$, using the inequalities $C' N(e/h) \|e\|_2 \leq \frac{\eta_0}{2} N(e/h)^2 + \frac{C'^2}{2\eta_0} \|e\|_2^2$, we finally obtain

$$\langle e, Ae \rangle \geq \frac{\eta_0}{2} N(e/h)^2 - \left(C + \frac{C'^2}{2\eta_0} \right) \|e\|_2^2$$

which gives the desired lower bound with $\eta = \eta_0/2$ and $\gamma = C + \frac{C'^2}{2\eta_0}$. \square

From now on we shall denote the error by

$$e^n := v^n - u^n.$$

Proposition 4.4. *Consider the scheme (21), and a perturbed scheme (24). Let $\tau > 0$ be sufficiently small. Then there exist a constant C_1 independent of n and a constant $\bar{\gamma} > 0$ such that for all $t_n \leq T$*

$$e^{-\bar{\gamma}t_n} \|e^{n+1}\|_2^2 + \tau\eta \sum_{k=1}^n e^{-\bar{\gamma}t_k} N(e^{k+1}/h)^2 \leq C_1 \left(\|e^0\|_2^2 + \|e^1\|_2^2 + \tau \sum_{k=1, \dots, n} e^{-\bar{\gamma}t_k} \|\bar{\epsilon}^k\|_2^2 \right) \quad (29)$$

where, by (25), $N(e^k/h)^2 = \sum_{j=1}^{J+1} \left| \frac{e_j^k - e_{j-1}^k}{h} \right|^2$.

Proof of Proposition 4.4. Let

$$B := I + \frac{2\tau}{3}A,$$

and vectors b_u, b_v be such that

$$b_u := \frac{4}{3}u^n - \frac{1}{3}u^{n-1} - \frac{2\tau}{3}q_{n+1}$$

and

$$b_v := \frac{4}{3}v^n - \frac{1}{3}v^{n-1} - \frac{2\tau}{3}q_{n+1}.$$

Then, by Lemma 4.1, the min equation (21) for the exact scheme is equivalent to $u^{n+1} \geq g$ and

$$\langle Bu^{n+1} - b_u, w - u^{n+1} \rangle \geq 0, \quad \forall w \geq g. \quad (30)$$

The min equation (24) for the perturbed scheme is equivalent to $v^{n+1} \geq g$ and

$$\langle Bv^{n+1} - (b_v + \frac{2}{3}\tau\bar{\epsilon}^n), w - v^{n+1} \rangle \geq 0, \quad \forall w \geq g. \quad (31)$$

Taking $w = v^{n+1}$ in (30) gives

$$\langle Bu^{n+1} - b_u, v^{n+1} - u^{n+1} \rangle \geq 0,$$

and $w := u^{n+1}$ in (31) gives

$$\langle Bv^{n+1} - (b_v + \frac{2}{3}\tau\bar{\epsilon}^n), u^{n+1} - v^{n+1} \rangle \geq 0.$$

Combining the last two relations gives

$$\langle Be^{n+1} - \frac{4}{3}e^n + \frac{1}{3}e^{n-1} - \frac{2\tau}{3}\bar{\epsilon}^n, e^{n+1} \rangle \leq 0 \quad (32)$$

and therefore

$$\langle 3e^{n+1} - 4e^n + e^{n-1}, e^{n+1} \rangle + 2\tau \langle e^{n+1}, Ae^{n+1} \rangle \leq 2\tau \langle \bar{\epsilon}^n, e^{n+1} \rangle. \quad (33)$$

Now let x_n, y_n and z_n be defined by

$$x_n := \|e^n\|_2^2, \quad y_n := \|e^{n+1} - e^n\|_2^2, \quad z_n := 2\tau \langle Ae^{n+1}, e^{n+1} \rangle.$$

The following estimate holds:

$$3x_{n+1} - 4x_n + x_{n-1} + 2y_n + 2z_n \leq 2y_{n-1} + 4\tau \langle \bar{\epsilon}^n, e^{n+1} \rangle. \quad (34)$$

To prove (34), we first use the properties $\langle a - b, a \rangle = \frac{1}{2}(\|a\|_2^2 + \|a - b\|_2^2 - \|b\|_2^2)$ as well as $\frac{1}{2}\|a + b\|_2^2 \leq \|a\|_2^2 + \|b\|_2^2$, to obtain

$$\begin{aligned} & 2\langle 3e^{n+1} - 4e^n + e^{n-1}, e^{n+1} \rangle \\ &= 2(4\langle e^{n+1} - e^n, e^{n+1} \rangle - \langle e^{n+1} - e^{n-1}, e^{n+1} \rangle) \\ &= 4(x_{n+1} + y_n - x_n) - (x_{n+1} + \|e^{n+1} - e^{n-1}\|_2^2 - x_{n-1}) \\ &\geq 4(x_{n+1} + y_n - x_n) - (x_{n+1} + 2(y_n + y_{n-1}) - x_{n-1}) \\ &\geq 3x_{n+1} - 4x_n + x_{n-1} + 2y_n - 2y_{n-1} \end{aligned}$$

and we conclude by using (33).

Let

$$w_n := 4\tau\eta N(e^{n+1}/h)^2.$$

By using the bound $2\tau\langle \bar{\epsilon}^n, e^{n+1} \rangle \leq 2\tau\|\bar{\epsilon}^n\|\|e^{n+1}\| \leq \tau\|\bar{\epsilon}^n\|_2^2 + \tau x_{n+1}$ and the coercivity (26), we obtain, for $n \geq 1$:

$$3x_{n+1} - 4x_n + x_{n-1} + 2y_n + w_n \leq 2y_{n-1} + 2\tau\|\bar{\epsilon}^n\|_2^2 + (2\tau + 4\tau\gamma)x_{n+1}. \quad (35)$$

Let

$$\bar{\gamma} := 2 + 4\gamma \quad \text{and} \quad \beta := \tau\bar{\gamma}.$$

It follows

$$(3 - \beta)x_{k+1} - 4x_k + x_{k-1} + 2y_k + w_k \leq 2y_{k-1} + 2\tau\|\bar{\epsilon}^k\|_2^2, \quad k \geq 1. \quad (36)$$

We multiply (36) by $e^{-k\beta}$ and sum up the inequalities from $k = 1$ to $n \geq 1$. Let $f(x) = x^2 - 4x + 3 - \beta$ and notice that for τ small enough (and therefore small $\beta > 0$), $f(e^{-\beta}) \sim \beta > 0$. We deduce that for some constant C_{01}

$$\begin{aligned} & e^{-n\beta}((3 - \beta)x_{n+1} - (4 - (3 - \beta)e^\beta)x_n) + \sum_{k=1}^n e^{-k\beta}w_k \\ &+ \sum_{k=2}^{n-1} e^{-(k-1)\beta}f(e^{-\beta})x_k + 2\sum_{k=1}^n e^{-k\beta}y_k \\ &\leq C_{01}(x_0 + x_1) + 2\sum_{k=1}^n e^{-k\beta}y_{k-1} + 2\sum_{k=1}^n e^{-k\beta}\tau\|\bar{\epsilon}^k\|_2^2. \end{aligned} \quad (37)$$

Using that $e^{-k\beta}y_{k-1} \leq e^{-(k-1)\beta}y_{k-1}$ and $f(e^{-\beta}) > 0$ we deduce that

$$\begin{aligned} & e^{-n\beta}((3 - \beta)x_{n+1} - (4 - (3 - \beta)e^\beta)x_n) + \sum_{k=1}^n e^{-k\beta}w_k \\ &\leq C_{01}(x_0 + x_1) + 2e^{-\beta}y_0 + \sum_{k=1}^n e^{-k\beta}2\tau\|\bar{\epsilon}^k\|_2^2 \\ &\leq C_{02}(x_0 + x_1 + \tau\sum_{k=1}^n e^{-k\beta}\|\bar{\epsilon}^k\|_2^2) =: Q \end{aligned} \quad (38)$$

with $C_{02} := C_{01} + 4$ (where we have used that $y_0 \leq 2(x_0 + x_1)$).

Let us prove that $e^{-\bar{\gamma}t_n}x_n \leq x_1 + C_{02}Q$, which will give the desired bound. By using $k\beta = k\tau\bar{\gamma} = \bar{\gamma}t_k$, we deduce from (38)

$$x_{k+1} \leq \rho x_k + \frac{e^{\bar{\gamma}t_n}Q}{3 - \beta}, \quad 1 \leq k \leq n,$$

where $\rho := \frac{4-(3-\beta)e^\beta}{3-\beta} \sim \frac{1}{3}$ as $\beta = \tau\bar{\gamma} \rightarrow 0$. By recursion we get for $1 \leq k \leq n$:

$$\begin{aligned} x_k &\leq \rho^{k-1}x_1 + \frac{e^{\bar{\gamma}t_n}Q}{3-\beta}(1 + \rho + \dots + \rho^{k-2}) \\ &\leq x_1 + \frac{e^{\bar{\gamma}t_n}Q}{3-\beta} \frac{1}{1-\rho}. \end{aligned}$$

By using this bound for x_n into (38), we obtain the desired result (29) with a possibly different universal constant C_1 . This concludes the proof of Proposition 4.4. \square

Remark 4.5 (BDF3 scheme). The previous stability estimate does not extend to the BDF3 obstacle scheme. Indeed, it is known that BDF3 is not A -stable (as well as any BFD k for $k \geq 3$, see [13]) which prevents the same stability analysis to apply for diffusion equations.

4.2. Error estimate for the BDF2 scheme. The following assumptions will be used.

Assumption (A2). We assume that there exists an integer $p \geq 0$ and continuous functions $t \rightarrow y_i(t)$ for $i = 1, \dots, p$ with $y_i(t) \in [X_{min}, X_{max}]$, such that, setting

$$\Omega_T := \{(t, x) \in (0, T) \times \Omega, x \notin (y_i(t))_{1 \leq i \leq p}\}$$

(Ω is a subdomain of $(0, T) \times \Omega$), the following holds:

- (i) $(t, x) \rightarrow v(t, x)$ is regular (i.e $C^{3,4}$) on Ω_T
- (ii) there exists constants $(C_j)_{1 \leq j \leq 4}$, $C_j \geq 0$, such that:

$$\begin{aligned} \|v_t\|_{L^\infty(\Omega_T)} &\leq C_1, & \|v_{2x}\|_{L^\infty(\Omega_T)} &\leq C_2, \\ \|v_{3t}\|_{L^\infty(\Omega_T)} &\leq C_3, & \|v_{4x}\|_{L^\infty(\Omega_T)} &\leq C_4 \end{aligned}$$

where $(C_j)_{1 \leq j \leq 4}$ are constants independent of (t, x) .

Notice that in order that (A2) holds, $\|(v_0)_{xx}\|_\infty$ must be bounded.

Assumption (A3). There exists $\alpha \in (0, 1]$ such that

- (i) for all i , $t \rightarrow y_i(t)$ is α -Hölder continuous on $[0, T]$;
- (ii) for all i , for any $\epsilon > 0$, y_i is Lipschitz continuous on $[\epsilon, T]$, and

$$\max_{[\epsilon, T]} |\dot{y}_i(t)| \leq C \epsilon^{\alpha-1}$$

for some constant $C \geq 0$ independent of ϵ .

Remark 4.6. Note that assumption (A2) allows for $v_{xx}(t, \cdot)$ to have ‘‘jumps’’ at the singular points $x = y_k(t)$.

Remark 4.7. For the american put option problem (4) with $\varphi(x) := (K - x)_+$, it is known that there is a unique singular point $y_1(t) \equiv x_s(t)$ such that $v(t, x) = \varphi(x)$ for $x < x_s(t)$, $v(t, x) > \varphi(x)$ for $x > x_s(t)$, and that $1 - x_s(t)/K \sim \lambda(t) |\ln(t)|^{1/2}$ for $t \rightarrow 0^+$ (see [2], and [1, Chap 6], as well as [9]).

Furthermore, a function of the form of $(t |\ln(t)|)^{1/2}$ satisfies assumption (A3) for any $\alpha < 1/2$.

This problem does not satisfies (A2) since this would imply that φ_{xx} to be bounded, and here φ_x is discontinuous.

However, assumptions (A2)-(A3) allow for closely related problems where φ is C^1 and piecewise C^2 (with φ_{xx} possibly discontinuous), and by allowing some rapidly moving singularities $y_i(t)$ with respect to t as $t \rightarrow 0$.

The following definition is introduced as a discrete equivalent of $\|u^n - v(t_n, \cdot)\|_{L^2(\Omega)}$:

$$\|e^n\|_{\ell^2(\Omega)} := \left(\sum_i h |e_i^n|^2 \right)^{1/2}.$$

Theorem 4.8. (error estimate). *Assume that the exact solution v of (1) satisfies (A1), (A2) and (A3). Then for the BDF2 scheme, the following bound holds as $(\tau, h) \rightarrow 0$:*

$$\max_{1 \leq n \leq N} \|e^n\|_{\ell^2(\Omega)} \leq C(\tau + h)^{1/2}$$

for some constant $C \geq 0$ independent of (τ, h) .

Proof. We are going to estimate the consistency error $\bar{\epsilon}$ as it appears in (23).

Let us first consider the approximation in the x variable. Let $\bar{\epsilon}^{n,1}$ be a consistency error in space, defined by

$$(\mathcal{A}v)(t_{n+1}, x_i) = (\mathcal{A}v^{n+1} + q_{n+1})_i + \bar{\epsilon}_i^{n,1}. \quad (39)$$

(If (t_{n+1}, x_i) corresponds to a singular point of v , we consider for $\mathcal{A}v(t_{n+1}, x_i)$ definitions of v_t and v_{xx} which satisfy the bounds $|v_t| \leq \|v_t\|_{L^\infty}$ and $|v_{xx}| \leq \|v_{xx}\|_{L^\infty}$.) In the region where $x \rightarrow v(t_{n+1}, x)$ is regular, assuming that $v_{4x}(t_{n+1}, \cdot)$ is bounded on the interval $[x_{i-1}, x_{i+1}]$, by using Taylor expansions it holds

$$\bar{\epsilon}_i^{n,1} = C_{4x}O(h^2).$$

On the contrary in a region $[x_{i-1}, x_{i+1}]$ that may encounter a singularity $y_j(t)$, we have no more than a bounded second order derivative ($v_{xx} \in L^\infty$), $|(v(t_{n+1}, x_{i+1}) - 2v(t_{n+1}, x_i) + v(t_{n+1}, x_{i-1}))/h^2| \leq \|v_{xx}\|_{L^\infty}$, and we only have

$$\bar{\epsilon}_i^{n,1} = C_{2x}O(1).$$

Moreover,

$$\text{Card}\left\{i, [x_{i-1}, x_{i+1}] \cap \{y_j(t_{n+1})\}_{1 \leq j \leq p} \neq \emptyset\right\} \leq 3p.$$

Then clearly

$$\|\bar{\epsilon}^{n,1}\|^2 = \sum_i |\bar{\epsilon}_i^{n,1}|^2 = \sum_{i, \text{regular}} |\bar{\epsilon}_i^{n,1}|^2 + \sum_{i, \text{singular}} |\bar{\epsilon}_i^{n,1}|^2 \quad (40)$$

$$\leq \sum_{i, \text{regular}} (C_{4x}h^2)^2 + 3p(C_{2x}O(1))^2 \quad (41)$$

$$\leq C \quad (42)$$

for some constant C that depends only of C_{2x}, C_{4x} and p . Hence the quantities $\|\bar{\epsilon}^{n,1}\|_2$ are uniformly bounded with respect to n . Therefore, we obtain

$$\tau \sum_{k=1}^n \|\bar{\epsilon}^{k,1}\|^2 \leq C. \quad (43)$$

Now we consider the approximation in time. Let $\bar{\epsilon}_i^{n,2}$ be such that

$$\frac{3v_i^{n+1} - 4v_i^n + v_i^{n-1}}{2\tau} = v_t(t_{n+1}, x_i) + \bar{\epsilon}_i^{n,2}. \quad (44)$$

If $t \rightarrow v(t, x_i)$ is regular on $[t_{n-1}, t_{n+1}]$ with bounded v_{3t} derivative, elementary Taylor expansions give

$$|\bar{\epsilon}_i^{n,2}| \leq C_3\tau^2.$$

On the other hand when $t \rightarrow v(t, x_i)$ is only Lipschitz continuous on the interval $[t_{n-1}, t_{n+1}]$, using the Lipschitz bound and Taylor expansions it holds

$$|\bar{\epsilon}_i^{n,2}| \leq 2\|v_t\|_{L^\infty} \leq 2C_1.$$

Let us introduce a set of singular indexes as follows:

$$\mathcal{I}_s := \left\{i, x_i \in \bigcup_{j=1, \dots, p} y_j([t_{n-1}, t_{n+1}])\right\}. \quad (45)$$

We have

$$\|\bar{\epsilon}^{n,2}\|^2 = \sum_i |\bar{\epsilon}_i^{n,2}|^2 = \sum_{i \notin \mathcal{I}_s} |\bar{\epsilon}_i^{n,2}|^2 + \sum_{i \in \mathcal{I}_s} |\bar{\epsilon}_i^{n,2}|^2 \quad (46)$$

$$\leq \sum_{i \notin \mathcal{I}_s} (C_3 \tau^2)^2 + \sum_{i \in \mathcal{I}_s} (2C_1)^2 \quad (47)$$

$$\leq C \frac{\tau^4}{h} + (2C_1)^2 \text{Card}(\mathcal{I}_s). \quad (48)$$

For $n \geq 2$ and for $t \in \Theta_n := [t_{n-1}, t_{n+1}]$,

$$\begin{aligned} |y_j(t) - y_j(t_{n+1})| &\leq \|\dot{y}_j\|_{L^\infty(\Theta_n)} 2\tau \\ &\leq C(t_{n-1})^{\alpha-1} 2\tau \\ &\leq 2C\tau^\alpha (n-1)^{\alpha-1}. \end{aligned}$$

So if $x_i \in y_j(\Theta_n)$ then $|x_i - y_j(t_{n+1})| \leq 2C\tau^\alpha (n-1)^{\alpha-1}$. Then, for any $A > 0$, the number of integers i such that $x_i \in [-A+c, A+c]$ is bounded by $2A/h + 1$. Hence, for $n \geq 2$, $\text{Card}(\mathcal{I}_s) \leq p(4C\frac{\tau^\alpha}{h}(n-1)^{\alpha-1} + 1)$. We deduce a bound in the form

$$\text{Card}(\mathcal{I}_s) \leq C \frac{\tau^\alpha}{h} n^{\alpha-1} + 1 \quad (49)$$

for some constant $C \geq 0$.

In the case when $n = 1$, $t_{n+1} = t_2 = 2\tau$ and for any $t \in [0, t_2]$,

$$|y_j(t) - y_j(t_2)| \leq |y_j(t) - y_j(0)| + |y_j(t_2) - y_j(0)| \leq 2C(2\tau)^\alpha.$$

Hence, for $n = 1$, we deduce a bound for $\text{Card}(\mathcal{I}_s)$ which is similar to (49).

Combining the previous bounds we obtain, for some other constant $C \geq 0$,

$$\tau \sum_{k=1}^n \|\bar{\epsilon}^{k,2}\|^2 \leq n\tau C \frac{\tau^4}{h} + C\tau \sum_{k=2}^n \left(\frac{\tau^\alpha}{h} k^{\alpha-1} + 1\right) \quad (50)$$

$$\leq C \left(\frac{\tau^{1+\alpha}}{h} \sum_{k=1}^n k^{\alpha-1} + 1\right). \quad (51)$$

By using $\sum_{k=1}^n k^{\alpha-1} \leq Cn^\alpha \simeq C/\tau^\alpha$, we obtain

$$\tau \sum_{k=1}^n \|\bar{\epsilon}^{k,2}\|^2 \leq C \left(\frac{\tau}{h} + 1\right). \quad (52)$$

Using the stability estimate (29), the fact that $e^0 = 0$ and that $\|e^1\|_2$ is clearly bounded (similar consistency and error estimates hold for ϵ^1 and then for e^1 , this is left to the reader), we deduce that

$$\|e^n\|_2^2 + \tau \sum_{k=1}^n N(e^k/h)^2 \leq C \left(\frac{\tau}{h} + 1\right).$$

Hence $\|e^n\|_{\ell^2(\Omega)} := (h\|e^n\|_2^2)^{1/2} \leq C(\tau+h)^{1/2}$, which gives the desired estimate. \square

5. NUMERICAL RESULTS ON TWO MODEL TEST PROBLEMS

5.1. Two model test problems. We first define two model test problems. In the case of (1) we do not know about exact solutions. Therefore we introduce a modified problem with a source term in order to construct simple model obstacle problems with explicit solutions (or solution that can be easily computed with machine precision) and also with the main features of the one-dimensional American option problem.

This is obtained by choosing an explicit function $v = v(t, x)$ and adding a corresponding source term $f = f(t, x)$ to the original PDE, thus considering

$$\min \left(v_t - \frac{\lambda^2}{2} x^2 v_{xx} - rxv_x + rv, v - g(x) \right) = f(t, x) \quad (53)$$

More precisely, let $g(x) := \max(K - x, 0)$ denote the payoff function. Let K, X_{max}, c_0 and T be given constants such that $0 < K < X_{max}, c_0 > 0, T > 0$ and such that $K - c_0 T > 0$. Let $x_s(t)$ be defined by

$$x_s(t) := K - c_0 t \quad (54)$$

for $t \in [0, T]$ (in particular $x_s(t) \in]0, K[$ for $t \in]0, T[$).

We construct explicit function $v(t, x)$ defined for $x \in [0, X_{max}]$ and such that

- (i) $v(t, x) = g(x) = K - x$ for $x \leq x_s(t)$,
- (ii) $v(t, x) > g(x) = \max(K - x, 0)$ for $x \in]x_s(t), X_{max}]$,
- (iii) v is at least C^1 on $[0, X_{max}]$,
- (iv) $v(t, X_{max}) = 0$.

Note that requirement (iii) implies in particular $v_x(t, x_s(t)) = g'(x_s(t)) = -1$ for $t > 0$.

Model 1 Let $v = v(t, x)$ be the function defined by:

$$v(t, x) := \begin{cases} g(x) & \text{for } x < x_s(t) \\ g(x_s(t)) - \frac{x - x_s(t)}{1 - (x - x_s(t))/C(t)} & \text{otherwise} \end{cases} \quad (55)$$

where $C(t) > 0$ is such that $v(t, X_{max}) = 0$:

$$C(t) := \left(\frac{1}{X_{max} - x_s(t)} - \frac{1}{g(x_s(t))} \right)^{-1}.$$

Then the requirements (i) – (iv) are satisfied.

Model 2 Let $v = v(t, x)$ be the function defined by:

$$v(t, x) := \begin{cases} g(x) & \text{for } x < x_s(t) \\ g(x_s(t)) - C(t) \operatorname{atan} \left(\frac{x - x_s(t)}{C(t)} \right) & \text{otherwise} \end{cases} \quad (56)$$

for a given $C(t) > 0$. Notice that $v(t, x)$ is a non-increasing function of the variable x . This function will satisfy requirements (i) – (iv) if furthermore $C(t)$ is such that

$$\frac{g(x_s(t))}{C(t)} = \operatorname{atan} \left(\frac{X_{max} - x_s(t)}{C(t)} \right). \quad (57)$$

Letting $a := X_{max} - x_s(t)$ and $b = g(x_s(t)) = K - x_s(t)$ it is clear that $0 < b < a$ and therefore there exists a unique $\theta > 0$ such that $b\theta = \operatorname{atan}(a\theta)$. This value can be numerically obtained by using a fixed-point method. We then define $C(t) := 1/\theta$ to obtain a solution of (57). Therefore the function v is in explicit form but for the computation of the $C(t)$ function which can be computed to arbitrary precision.

Remark 5.1. For model 2, in order to compute $v_t(t, x)$ the derivative of $C(t)$, $\dot{C}(t)$, is needed. Denoting $a = a(t) = X_{max} - x_s(t)$ and $b = b(t) = g(x_s(t))$, and $\theta = \theta(t) = 1/C(t)$, by derivation of $b\theta = \operatorname{atan}(a\theta)$ we obtain $\dot{C}/C = -\dot{\theta}/\theta = (q\dot{b} - \dot{a})/(qb - a)$ where $q = 1 + (a/C)^2$, with $\dot{a} = -\dot{x}_s(t)$ and $\dot{b} = g'(x_s(t))\dot{x}_s(t)$.

Remark 5.2. The main difference between the two models is the regularity of the data near the singularity $x = x_s(t)$. More precisely, for the first model there is a jump in the second derivative: v is of class C^1 and v_{xx} is discontinuous. For the second model, v is of class C^2 and there is a jump in the third derivative v_{3x} .

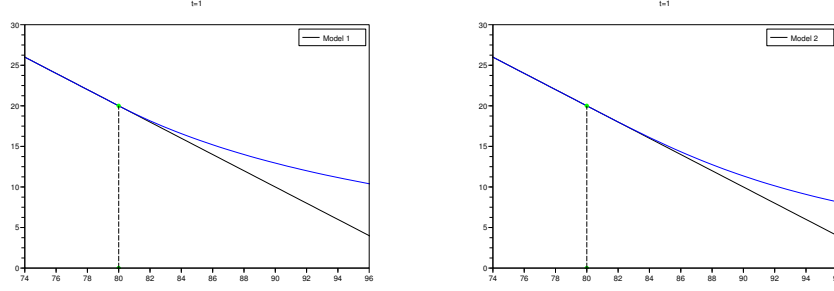


FIGURE 1. Zooming around the singular point $(x_s, g(x_s))$ for model 1 (left) and 2 (right).

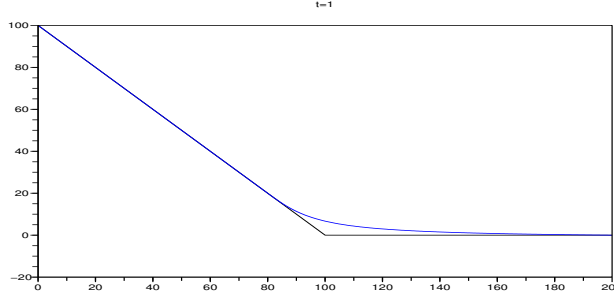


FIGURE 2. Model 2.

5.2. A 4th order approximation of the spatial operator \mathcal{A} . We furthermore introduce a 4th order numerical matrix approximation of the \mathcal{A} operator in order to better observe the time discretization error.

Let $D^2u_j := \frac{-u_{j-1} + 2u_j - u_{j+1}}{h^2}$. By using Taylor expansions, if $u_j = u(x_j)$, we have

$$-u_{xx}(x_j) = \frac{-u_{j-1} + 2u_j - u_{j+1}}{h^2} + \frac{1}{12h^2}(u_{j-2} - 4u_{j-1} + 6u_j - 4u_{j+1} + u_{j+2}) + O(h^4)$$

and also,

$$u_x(x_j) = \frac{u_{j+1} - u_{j-1}}{2h} + \frac{1}{12h}(u_{j-2} - 2u_{j-1} + 2u_{j+1} - u_{j+2}) + O(h^4).$$

Therefore we have a 4-th order approximation of the spatial derivatives. At the boundaries, we use $u_{-1} = K - X_{\min} - h$, $u_0 = K - X_{\min}$ and $u_{J+1} = u_{J+2} = 0$ (the left boundary condition is consistent with fourth order because we expect that $v(t, x) = g(x) = K - x$ near the left boundary, also the right boundary condition is consistent with the fact that the exact solution $v(t, x)$ (obtained for the case $X_{\max} = \infty$) decays faster than any polynomial as $x \rightarrow \infty$).

The BDF2 and BDF3 schemes are otherwise unchanged concerning the time discretization. Hence for $\tau \equiv c h$ we expect to mainly see the error of the time discretization. In particular we aim to observe, whenever possible, second or third order behavior in time.

5.3. Numerical results for models 1 and 2. For testing on models 1 and 2 the parameters used are $\lambda = 0.3$, $r = 0.1$, $T = 1$, $X_{\min} = 50$ and $X_{\max} = 250$.

Numerical results for Model 1 with the BDF2 (resp. BDF3) schemes are given in Tables 7 and 8. We observe that in Table 8 even with the third order scheme the numerical order is not greater than two. This can be explained by the fact that the solution has bounded second derivatives but no more.

Numerical results for Model 2 with the BDF2 (resp. BDF3) schemes are given in Tables 9 and 10. Now the last Table clearly shows third order. For this model, by construction, the exact solution has bounded third derivatives.

Mesh		Error L^1		Error L^2		Error L^∞		time(s)
J	N	error	order	error	order	error	order	
80	8	2.56E+00	-	5.93E-01	-	1.94E-01	-	0.10
160	16	6.48E-01	1.98	1.48E-01	2.00	5.11E-02	1.93	0.11
320	32	1.77E-01	1.87	4.00E-02	1.89	1.33E-02	1.95	0.24
640	64	4.72E-02	1.90	1.06E-02	1.92	3.41E-03	1.96	0.48
1280	128	1.25E-02	1.92	2.77E-03	1.94	8.60E-04	1.99	1.28
2560	256	3.24E-03	1.95	7.14E-04	1.96	2.16E-04	2.00	4.64
5120	512	8.06E-04	2.01	1.78E-04	2.00	5.39E-05	2.00	19.30

TABLE 7. (Model 1) BDF2 - second order - scheme with 4th order spatial approximation for Eq.(1), using high CFL numbers.

Mesh		Error L^1		Error L^2		Error L^∞		time(s)
J	N	error	order	error	order	error	order	
80	8	2.50E+00	-	5.99E-01	-	2.09E-01	-	0.09
160	16	6.42E-01	1.96	1.48E-01	2.01	5.34E-02	1.96	0.13
320	32	1.69E-01	1.93	3.85E-02	1.95	1.35E-02	1.99	0.24
640	64	4.41E-02	1.94	9.96E-03	1.95	3.36E-03	2.00	0.53
1280	128	1.15E-02	1.94	2.57E-03	1.96	8.43E-04	2.00	1.66
2560	256	2.96E-03	1.96	6.55E-04	1.97	2.11E-04	2.00	7.27
5120	512	7.33E-04	2.01	1.63E-04	2.01	5.28E-05	2.00	39.23

TABLE 8. (Model 1) BDF3 - third order - scheme with 4th order spatial approximation for (1), using high CFL numbers.

Mesh		Error L^1		Error L^2		Error L^∞		time(s)
J	N	error	order	error	order	error	order	
80	8	2.91E-01	-	5.31E-02	-	1.32E-02	-	0.07
160	16	5.80E-02	2.33	1.07E-02	2.31	2.72E-03	2.28	0.09
320	32	1.29E-02	2.16	2.42E-03	2.15	6.22E-04	2.13	0.23
640	64	3.07E-03	2.07	5.77E-04	2.07	1.48E-04	2.07	0.50
1280	128	7.46E-04	2.04	1.40E-04	2.04	3.62E-05	2.04	1.38
2560	256	1.84E-04	2.02	3.46E-05	2.02	8.92E-06	2.02	5.12
5120	512	4.55E-05	2.01	8.58E-06	2.01	2.21E-06	2.01	20.44

TABLE 9. (Model 2) BDF2 - second order - scheme with 4th order spatial approximation for (1), using high CFL numbers.

Mesh		Error L^1		Error L^2		Error L^∞		time(s)
J	N	error	order	error	order	error	order	
80	8	6.94E-02	-	1.31E-02	-	3.76E-03	-	0.10
160	16	1.96E-03	5.14	5.61E-04	4.54	2.75E-04	3.77	0.13
320	32	4.00E-04	2.29	7.59E-05	2.89	2.20E-05	3.65	0.25
640	64	7.85E-05	2.35	1.50E-05	2.34	4.07E-06	2.43	0.62
1280	128	1.14E-05	2.79	2.15E-06	2.80	5.70E-07	2.84	1.86
2560	256	1.51E-06	2.91	2.81E-07	2.93	7.39E-08	2.95	7.95
5120	512	1.99E-07	2.92	3.63E-08	2.95	9.50E-09	2.96	41.00

TABLE 10. (Model 2) BDF3 - third order - scheme with 4th order spatial approximation for (1), using high CFL numbers.

APPENDIX A. NUMERICAL SETTINGS FOR THE AMERICAN OPTION MODEL

For the one-dimensional American put option model we have used the parameters (13).

In this setting, we observe that the singular point $x_s(T)$ is greater than 80, so for the numerical approximation we have considered the subdomain $\Omega = (X_{min}, X_{max}) \equiv (75, 275)$ and boundary conditions of Dirichlet type:

$$u(t, X_{min}) = K - X_{min}, \quad 0 < t < T,$$

and

$$u(t, X_{max}) = 0, \quad 0 < t < T.$$

We have numerically estimated that the truncation error from the right, using $X_{max} = 275$ instead of $X_{max} = +\infty$, is less than 10^{-8} .

APPENDIX B. AN HJB EQUATION FOR OBSTACLE PROBLEMS

This appendix is devoted to a sketch of proof for the equivalence between PDE (1) and PDE (6) in the case the coefficients are not time dependent.

In order to simplify the presentation we assume that $f \equiv 0$. We consider the problem (1) after a change of variable $t \rightarrow T - t$:

$$\min(-v_t + \mathcal{A}v, v - \varphi(x)) = 0, \quad t \in (0, T), \quad x \in \Omega, \quad (58a)$$

$$v(T, x) = \varphi(x), \quad x \in \Omega \quad (58b)$$

where, for simplification we consider here also $\Omega = \mathbb{R}$. We aim to prove that v is also a viscosity solution of (6).

(i) By uniqueness of the continuous solutions of (1), v is also given by the expectation formula

$$v(t, x) = \sup_{\tau \in \mathcal{T}_{[t, T]}} \mathbb{E}(e^{-\int_t^\tau r ds} \varphi(X_\tau^{t, x}) | \mathcal{F}_t).$$

(see for instance [16]) where we have considered a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a filtration $(\mathcal{F}_t)_{t \geq 0}$, $\mathcal{T}_{[t, T]}$ is the set of stopping times taking values a.s in $[t, T]$, $X_\tau := X_\tau^{t, x}$ is the strong solution of the stochastic differential equation (SDE):

$$dX_s = b(X_s)ds + \sigma(X_s)dW_s, \quad s \geq t,$$

with $X_t = x$, W_s denotes an \mathcal{F}_t -adapted Brownian motion on \mathbb{R} , and the ‘‘sup’’ is an essential supremum over $\mathcal{T}_{[t, T]}$. First one can use the semi-Martingale property

$$v(t, x) \leq \mathbb{E}(e^{-rh} v(t+h, X_{t+h}^{t, x}) | \mathcal{F}_t),$$

in order to deduce (in the viscosity sense), that

$$-v_t + \mathcal{A}v \geq 0.$$

(ii) Then we aim to show that $v(t, x) \geq v(t + h, x)$, for any $h > 0$. This will imply $-v_t \geq 0$ (in the viscosity sense). By definition,

$$v(t + h, x) = \sup_{\tau \in \mathcal{T}_{[t+h, T]}} \mathbb{E}(e^{-\int_{t+h}^{\tau} r ds} \varphi(X_{\tau}^{t+h, x}) | \mathcal{F}_{t+h}) \quad (59)$$

$$= \sup_{\tau \in \mathcal{T}_{[t, T-h]}} \mathbb{E}(e^{-\int_{t+h}^{\tau+h} r ds} \varphi(X_{\tau+h}^{t+h, x}) | \mathcal{F}_{t+h}) \quad (60)$$

$$= \sup_{\tau \in \mathcal{T}_{[t, T-h]}} \mathbb{E}(e^{-\int_t^{\tau} r ds} \varphi(X_{\tau}^{t, x}) | \mathcal{F}_t). \quad (61)$$

We have used the fact that the process $X^{t, x}$ satisfies an SDE with no time dependency in the coefficients, and also, since $\tau \in \mathcal{T}_{[t, T-h]}$, $X_{\tau+h}$ a.s. stops before time T , the fact that $\mathbb{E}(X_{\tau+h}^{t+h, x} | \mathcal{F}_{t+h}) = \mathbb{E}(X_{\tau}^{t, x} | \mathcal{F}_t)$ - which corresponds to an averaging during a period of time $T - (t + h)$. Then, in particular,

$$v(t + h, x) \leq \sup_{\tau \in \mathcal{T}_{[t, T]}} \mathbb{E}(e^{-\int_t^{\tau} r ds} \varphi(X_{\tau}^{t, x}) | \mathcal{F}_t) = u(t, x). \quad (62)$$

At this point we therefore have shown that

$$\min(-v_t + \mathcal{A}v, -v_t) \geq 0.$$

(iii) Let us assume that $-v_t(t, x) > 0$ (in the viscosity sense), and $t < T$. It implies that $v(t, x) > v(t + h, x)$ for all $h > 0$ small enough. Because $v(t, x) > v(t + h, x) \geq \varphi(x)$, we have $v(t, x) > \varphi(x)$. The following dynamic programming principle holds:

$$v(t, x) = \mathbb{E}(e^{-\int_t^{\tau_{t,x}^*} r ds} \varphi(X_{\tau_{t,x}^*}^{t, x}) | \mathcal{F}_t) = \mathbb{E}(e^{-\int_t^{\tau_{t,x}^*} r ds} v(\tau_{t,x}^*, X_{\tau_{t,x}^*}^{t, x}) | \mathcal{F}_t)$$

where $\tau_{t,x}^*$ is the optimal stopping time for the obstacle problem, defined by

$$\tau_{t,x}^* = \inf \left\{ \theta \geq t, v(\theta, X_{\theta}^{t, x}) = \varphi(X_{\theta}^{t, x}) \right\}.$$

It can be shown that $\tau_{t,x}^* > t$ a.s. (since $v(t, x) > \varphi(x)$, these functions being continuous). Then we obtain that $-v_t + \mathcal{A}v = 0$ at (t, x) in the viscosity sense. By using Ito's formula between t and $\tau_{t,x}^*$, and from the dynamic programming principle, we deduce that

$$0 = \mathbb{E} \left(\int_t^{\tau_{t,x}^*} e^{-\int_t^{\theta} r ds} (v_t - \mathcal{A}v)_{(\theta, X_{\theta}^{t, x})} d\theta \mid \mathcal{F}_t \right).$$

We already have proved that $v_t - \mathcal{A}v \leq 0$ a.s., so we deduce that $(v_t - \mathcal{A}v)(\theta, x) = 0$ a.e. for $\theta \in (t, \tau_{t,x}^*)$. For some random parameter w we have $t^* := \tau_{t,x}^*(w) > t$, from which it is deduced that $(v_t - \mathcal{A}v)(t, x) = 0$. Therefore we have proved in this case that $\min(-v_t(t, x) - (\mathcal{A}v)(t, x), -v_t) = 0$.

(iv) Conversely, we can use a uniqueness argument for the solutions of (6) in order to conclude the equivalence between (1) and (6).

Remark B.1. In the same way, it can be proved that the following PDE with source term and x -dependent coefficients in the operator \mathcal{A} :

$$\min(-u_t + \mathcal{A}u, u - \varphi(x)) = f(x), \quad t \in (0, T), \quad x \in \Omega, \quad (63a)$$

$$u(T, x) = \varphi(x) + f(x), \quad x \in \Omega \quad (63b)$$

is equivalent to the following Hamilton-Jacobi-Bellman equation

$$-u_t + \min(\mathcal{A}u, 0) = f(x), \quad t \in (0, T), \quad x \in \Omega, \quad (64a)$$

$$u(T, x) = \varphi(x) + f(x), \quad x \in \Omega \quad (64b)$$

Problem (63) is associated with the following stopping time problem

$$u(t, x) = \sup_{\tau \in \mathcal{T}_{[t, T]}} \mathbb{E} \left(e^{-\int_t^\tau r ds} \varphi(X_\tau^{t, x}) + \int_t^\tau e^{-\int_t^\theta r ds} f(X_\theta^{t, x}) d\theta | \mathcal{F}_t \right).$$

APPENDIX C. PROOF OF LEMMA 2.4

Proof of (i): Let $c_n := \frac{1}{2}(Au^{n+1} + Au^n)_j + q(t_{n+1/2})_j$, so that the first scheme (9) reads $\min(\frac{u_j^{n+1} - u_j^n}{\tau} + c_j^n, u_j^{n+1} - g_j) = 0$. Assuming that u^{n+1} is solution of scheme (9), with $u^{n+1} \geq u^n$, if $u_j^{n+1} = g_j$, then it is clear that also $u_j^n = g_j = u_j^{n+1}$ and therefore since $\frac{u_j^{n+1} - u_j^n}{\tau} + c_j^n \geq 0$ it can be deduced that $S_j^{2,n}(u) = 0$. On the other hand, if $u_j^{n+1} > g_j$, then it implies $\frac{u_j^{n+1} - u_j^n}{\tau} + c_j^n = 0$, from which we conclude to $S_j^{2,n}(u) = 0$. so u^{n+1} is also a solution of the scheme (11).

Proof of (ii): For the American option problem, the left boundary condition will be $u_\ell(t_n) = K - X_{min}$, and the right boundary condition $u_r(t_n) = 0$. So that the vector $q(t_{n+1/2}) = q$ does not depend on time. It is easy to see that $u^1 \geq u^0 = g$. Assuming $u^n \geq u^{n-1}$, for some $n \geq 1$, and denoting $F_n(x) := \min(Bx - b_n, x - g)$ where $B = I + \frac{\tau}{2}A$ and $b_n = (I - \frac{\tau}{2}A)u^n + \tau q$, the scheme is equivalent to $F_n(u^n) = 0$. Because B is an M -matrix, the function F_n is monotone in the sense that $F_n(x) \leq F_n(y) \Rightarrow x \leq y$ (componentwise). On the other hand, $u^{n-1} \leq u^n$ and assuming that $I + \frac{dt}{2}A \geq 0$, it holds $b_{n-1} \leq b_n$, and therefore $F_n(x) \leq F_{n-1}(x)$ for all x . In particular, $F_{n-1}(u_n) = 0 = F_n(u_{n+1}) \leq F_{n-1}(u_{n+1})$, and by the monotonicity of F_{n-1} we conclude to $u_n \leq u_{n+1}$.

APPENDIX D. BDF-MAX SCHEMES FOR OBSTACLE PROBLEMS

We aim here to report on an other attempt on a related BDF scheme in view of the treatment of an obstacle term in a parabolic problem, giving good numerical results even if the stability or convergence proof are still an issue.

We first consider the BDF idea for the direct treatment of the full operator $u_t + \mathcal{A}u$.

Consider a first order scheme denoted S_τ to approximate $u_t + \mathcal{A}u = 0$, of the form $u^{n+1} = S_\tau u^n$. Typically this can be an implicit Euler scheme, therefore solving

$$\frac{u^{n+1} - u^n}{\tau} + Au^{n+1} + q(t_{n+1}) = 0.$$

In that case $S_\tau u \equiv S_\tau^n u := (I_d + \tau A)^{-1}(u - \tau q(t_{n+1}))$. One could also consider a semi-Lagrangian scheme. The idea here is to use a scheme that is not too much CFL restrictive and easy to solve. Suppose there is furthermore a second (resp. third) order expansion for small τ of the form:

$$\frac{v_j^{n+1} - (\mathcal{S}_\tau v^n)_j}{\tau} = (v_t + \mathcal{A}v)(t_{n+1}, x_j) + a(t_{n+1}, x_j)\tau + O(\tau^2), \quad (65)$$

resp.

$$\frac{v_j^{n+1} - (\mathcal{S}_\tau v^n)_j}{\tau} = (v_t + \mathcal{A}v)(t_{n+1}, x_j) + a(t_{n+1}, x_j)\tau + b(t_{n+1}, x_j)\tau^2 + O(\tau^3) \quad (66)$$

(with $v_j^n := v(t_n, x_j)$ where v is a regular function). Then, using the idea of BDF approximations, we easily obtain the following orders, assuming (65):

$$\frac{3v_j^{n+1} - 4(\mathcal{S}_\tau v^n)_j + (\mathcal{S}_{2\tau} v^{n-1})_j}{2\tau} = (v_t + \mathcal{A}v)(t_{n+1}, x_j) + O(\tau^2), \quad (67)$$

or, assuming (66):

$$\frac{11v_j^{n+1} - 18(\mathcal{S}_\tau v^n)_j + 9(\mathcal{S}_{2\tau} v^{n-1})_j - 2(\mathcal{S}_{3\tau} v^{n-2})_j}{6\tau} = (v_t + \mathcal{A}v)(t_{n+1}, x_j) + O(\tau^3). \quad (68)$$

Moreover, it is not necessary to assume $v_t + \mathcal{A}v = 0$ in order to obtain the above approximations.

Then a second order in time-consistent scheme is directly obtained by considering u^{n+1} solution of

$$\min \left(\frac{3u^{n+1} - 4\mathcal{S}_\tau u^n + \mathcal{S}_{2\tau} u^{n-1}}{2\tau}, u^{n+1} - g \right) = f_j^{n+1}. \quad (69)$$

Since the left part of the min can be multiplied by $\frac{2\tau}{3}$, the u^{n+1} term can be solved explicitly to obtain an explicit scheme.

In this way, we can consider

- an “explicit” two-step max scheme for $n \geq 1$, hereafter referred as the “BDF2max” scheme:

$$u^{n+1} = \max \left(\frac{1}{3} (4\mathcal{S}_\tau^n u^n - \mathcal{S}_{2\tau}^{n-1} u^{n-1}) + \frac{2\tau}{3} f^{n+1}, g + f^{n+1} \right). \quad (70)$$

(with initialization $u^0 = g$ and u^1 : implicit Euler).

- an “explicit” three-step max scheme, hereafter referred as the “BDF3max” scheme, for $n \geq 2$:

$$u^{n+1} = \max \left(\frac{6}{11} (3\mathcal{S}_\tau^n u^n - \frac{3}{2}\mathcal{S}_{2\tau}^{n-1} u^{n-1} + \frac{1}{3}\mathcal{S}_{3\tau}^{n-2} u^{n-2}) + \frac{6\tau}{11} f^{n+1}, g + f^{n+1} \right) \quad (71)$$

(with initialization $u^0 = g$ and u^1 implicit Euler, u^2 Crank-Nicolson).

The “explicit” denomination is to emphasize the fact that the obstacle term is treated in an explicit way, to the contrary to the BDF schemes of the previous section.

Proposition D.1. *For regular data, the scheme (70) has a second order consistency error $O(\tau^2)$ with respect to PDE (1). Similarly, the scheme (71) has third order consistency error $O(\tau^3)$.*

If the operator \mathcal{A} is approximated by finite differences as in Section (2), then the spatial consistency error $O(h^2)$ is also present in both cases.

Here we show numerical results using implicit Euler scheme for the definition of S_τ . The BDF2-max scheme is tested on the American option model (4) (see also Appendix A). Results are given in Table 11. For the BDF3-max scheme results we observe similar second order convergence results (as expected when the spatial approximation is second order).

In conclusion, we observe a good numerical behavior of the BDF2-max and BDF3-max schemes, roughly of second order in the L^2 norm, unconditionnally in the mesh steps. Because of the discontinuity in v_{xx} (with bounded v_{xx}) we do not expect better than second order behavior. However it does not seem easy to obtain stability results for these schemes, because of the presence of the non-linear max operator coupled with the non-monotonicity of the linear part of the operator (70) or (71).

Mesh		Error L^1		Error L^2		Error L^∞		time(s)
J	N	error	order	error	order	error	order	
80	80	1.78E-02	-	2.22E-02	-	4.90E-02	-	0.11
160	160	5.15E-03	1.79	6.47E-03	1.78	1.48E-02	1.73	0.22
320	320	1.52E-03	1.76	2.14E-03	1.59	6.10E-03	1.27	0.40
640	640	3.88E-04	1.97	5.33E-04	2.01	1.51E-03	2.01	0.97
1280	1280	1.18E-04	1.72	1.62E-04	1.72	4.61E-04	1.72	2.60
2560	2560	3.25E-05	1.86	4.10E-05	1.98	1.01E-04	2.19	8.64

TABLE 11. second order BDF-max scheme for (1).

REFERENCES

- [1] Yves Achdou and Olivier Pironneau. *Computational methods for option pricing*, volume 30 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.
- [2] Guy Barles, Julien Burdeau, Marc Romano, and Nicolas Samsøe n. Estimation de la frontière libre des options américaines au voisinage de l'échéance. *C. R. Acad. Sci. Paris Sér. I Math.*, 316(2):171–174, 1993.
- [3] Julien Berton and Robert Eymard. Finite volume methods for the valuation of American options. *M2AN Math. Model. Numer. Anal.*, 40(2):311–330, 2006.
- [4] A. Blanchet, J. Dolbeault, and R. Monneau. On the continuity of the time derivative of the solution to the parabolic obstacle problem with variable coefficients. *J. Math. Pures Appl. (9)*, 85(3):371–414, 2006.
- [5] Olivier Bokanowski, Stefania Maroso, and Hasnaa Zidani. Some convergence results for Howard's algorithm. *SIAM J. Numer. Anal.*, 47(4):3001–3026, 2009.
- [6] Philippe G. Ciarlet. *Introduction à l'analyse numérique matricielle et à l'optimisation*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1982.
- [7] M.G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. American Mathematical Society (New Series)*, 27:1–67, 1992.
- [8] C. F. Curtiss and J. O. Hirschfelder. Integration of stiff equations. *Proc. Nat. Acad. Sci. U. S. A.*, 38:235–243, 1952.
- [9] J. N. Deynne, S. D. Howison, I. Rupf, and P. Wilmott. Some mathematical results in the pricing of American options. *European J. Appl. Math.*, 4(4):381–398, 1993.
- [10] Etienne Emmrich. Stability and error of the variable two-step bdf for semilinear parabolic problems. *J. Appl. Math. & Computing*, 19(1-2):33–55, 2005.
- [11] A. Friedman. *Variational principles and free-boundary problems*. Robert E. Krieger Publishing Co., Inc., Malabar, FL, second edition, 1988.
- [12] C. William Gear. *Numerical initial value problems in ordinary differential equations*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971.
- [13] E. Hairer and G. Wanner. Linear multistep method. *Scholarpedia*, 5(4):4591, 2010.
- [14] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13(3):865–888 (2003), 2002.
- [15] Patrick Jaillet, Damien Lamberton, and Bernard Lapeyre. Variational inequalities and the pricing of American options. *Acta Appl. Math.*, 21(3):263–289, 1990.
- [16] Damien Lamberton and Bernard Lapeyre. *Introduction to stochastic calculus applied to finance*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, second edition, 2008.
- [17] C. Martini. American Option Prices as Unique Viscosity Solutions to Degenerated Hamilton-Jacobi-Bellman Equations. Research Report RR-3934, INRIA, 2000.
- [18] C. W. Oosterlee. On multigrid for linear complementarity problems with application to American-style options. *Electron. Trans. Numer. Anal.*, 15:165–185 (electronic), 2003. Tenth Copper Mountain Conference on Multigrid Methods (Copper Mountain, CO, 2001).
- [19] C. W. Oosterlee, F. J. Gaspar, and J. C. Frisch. WENO and blended BDF discretizations for option pricing problems. In *Numerical mathematics and advanced applications*, pages 419–428. Springer Italia, Milan, 2003.

- [20] S. I. Serdjukova. Uniform stability of a six-point scheme of higher order accuracy for the heat equation. *Ž. Vyčisl. Mat. i Mat. Fiz.*, 7(1):214–218, 1967.
- [21] Rüdiger U. Seydel. *Tools for computational finance*. Universitext. Springer London, fifth edition, 2012.

LABORATOIRE JACQUES-LOUIS LIONS, UNIVERSITÉ PARIS-DIDEROT (PARIS 7),
AND ENSTA PARISTECH (*BOKA@MATH.JUSSIEU.FR*).

UNIVERSITY OF SOUTHERN DENMARK, DEPARTMENT OF MATHEMATICS
AND COMPUTER SCIENCE, ODENSE (*DEBRABANT@IMADA.SDU.DK*).