Stability and convergence of second order backward differentiation schemes for parabolic Hamilton-Jacobi-Bellman equations

Olivier Bokanowski, Athena Picarelli, Christoph Reisinger

STABILITY RESULTS FOR SECOND ORDER BACKWARD DIFFERENTIATION SCHEMES FOR PARABOLIC HAMILTON-JACOBI-BELLMAN EQUATIONS

OLIVIER BOKANOWSKI*, ATHENA PICARELLI†, AND CHRISTOPH REISINGER†

Abstract. We propose a new second order BDF (Backward Differentiation Formula) scheme for the numerical approximation of nonlinear HJB (Hamilton-Jacobi-Bellman) equations. The scheme under consideration is implicit, non-monotone, and second order accurate in time and space. The lack of monotonicity of the scheme prevents the use of well known convergence results for solutions in the viscosity sense. In this work, we establish rigorous stability results in a general nonlinear setting. While most results are derived for one-dimensional parabolic linear and HJB equations, some results are extended to multiple dimensions and to Isaacs equations. Some numerical tests are also included to validate the method.

1. Introduction. The aim of this paper is to provide stability results for a type of implicit finite difference scheme for the approximation of nonlinear parabolic equations using backward differentiation formulae (BDF).

In particular, we consider Hamilton-Jacobi-Bellman (HJB) equations of the following form:

$$v_t(t,x) + \sup_{a\in\Lambda}\Big\{\mathcal{L}^a[v](t,x) + r(t,x,a)v + \ell(t,x,a)\Big\} = 0, \tag{1}$$

where $(t,x) \in [0,T]\times\mathbb{R}^d$, $\Lambda \subset \mathbb{R}^m$ is a compact set and

$$\mathcal{L}^a[v](t,x) = -\frac{1}{2}\operatorname{tr}[\Sigma(t,x,a)D_x^2 v(t,x)] + b(t,x,a)D_x v(t,x)$$

is a second order differential operator. Here, $(\Sigma)_{ij}$ is symmetric non-negative definite. Linear parabolic equations, corresponding to the case $|\Lambda| = 1$, are a special case also discussed in the paper.

It is well known that for nonlinear, possibly degenerate equations the appropriate notion of solutions to be considered is that of viscosity solutions [5]. We assume throughout the whole paper the well-posedness of the problem, namely the existence and uniqueness of a solution in the viscosity sense.

Under such weak assumptions, convergence of numerical schemes can only be guaranteed if they satisfy certain monotonicity properties, in addition to the more standard consistency and stability conditions for linear equations [1]. This in turn reduces the obtainable consistency order to 1 in the general case [7].

On the other hand, in many cases – especially in non-degenerate ones – solutions exhibit higher regularity and are amenable to higher order approximations. Classical solutions which are $C^{1+\delta,2+\delta}$ in the interior exist under a strict ellipticity condition and certain regularity of the domain and the coefficients [10].

The higher order of convergence in both space and time of discontinuous Galerkin approximations is demonstrated theoretically and empirically in [12] for sufficiently regular solutions under a Cordes condition on the diffusion matrix, a measure of the ellipticity. More recently, it was shown empirically in [3] that schemes based on

* Laboratoire Jacques-Louis Lions Université Paris-Diderot (Paris 7), and Ensta ParisTech (boka@math.jussieu.fr).

†Mathematical Institute, University of Oxford, Andrew Wiles Building, Woodstock Rd, Oxford OX2 6GG, UK (athena.picarelli@gmail.com, christoph.reisinger@maths.ox.ac.uk)

first derivative approximations in time and space based on a second order backward differentiation formula (see, e.g., [13], Section 12.11, for the definition of BDF schemes for ODEs) have good convergence properties. In particular, in a non-degenerate controlled diffusion example where the Crank-Nicolson scheme, which is also of second order and also non-monotone, fails to converge, the BDF scheme shows second order convergence.

For constant coefficient parabolic PDEs, the L_2-stability and smoothing properties of the BDF scheme are a direct consequence of the strong A-stability of the scheme. Moreover, [2] shows that for the multi-dimensional heat equation the BDF time stepping solution and its first numerical derivative are stable in the maximum norm. The technique, which is strongly based on estimates for the resolvent of the discrete Laplacian, do not easily extend to variable coefficients or even the nonlinear case. The BDF time stepping scheme is analysed for the incompressible Navier–Stokes equations in [9] and for semilinear parabolic equations in [6].
The scheme we propose is constructed by using a second order BDF approximation for the first derivatives in both time and space. Combining this with the standard three-point central finite difference for the second spatial derivative in one dimension, the scheme is second order consistent by construction. The extension of the results to the multi-dimensional case is also analysed.

The contribution of this paper is to establish stability results in the L_2 and H^1 norms for variable coefficient linear parabolic PDEs and HJB (and Isaacs) equations, given certain regularity of the coefficients.

The rest of the paper is structured as follows. In Section 2, we define the BDF schemes studied in this paper and state the main results on well-posedness and stability in a discrete H^1 norm. In Section 3, we prove these results and outline an extension from HJB to Isaacs equations. In Section 4, we give stability results in the discrete L_2 norm, which are weaker in the sense that regularity of the diffusion coefficient is required, but stronger in the sense that they allow for degenerate diffusion and can be extended to two dimensions. Section 5 studies carefully two numerical examples, the Eikonal equation and a second order equation with controlled diffusion. Section 6 concludes.

2. Definition of the scheme and main result. We focus in the first instance on the one-dimensional equation

$$v_t + \max_{a\in\Lambda}\Big(-\frac{1}{2}\sigma^2(t,x,a)v_{xx} + b(t,x,a)v_x + r(t,x,a)v + \ell(t,x,a)\Big) = 0, \qquad t\in[0,T],\ x\in\mathbb{R}, \tag{2a}$$

$$v(0,x) = v_0(x) \quad x\in\mathbb{R}. \tag{2b}$$

It is known that with the following assumptions:
- Λ is a compact set,
- for some $C_0>0$ the functions $\phi\equiv\sigma,b,r,\ell : [0,T]\times\mathbb{R}\times\Lambda\to\mathbb{R}$ and $v_0:\mathbb{R}\to\mathbb{R}$ satisfy for any $t,s\in[0,T]$, $x,y\in\mathbb{R}$, $a\in\Lambda$

$$|v_0(x)| + |\phi(t,x,a)| \le C_0,$$
$$|v_0(t,x) - v_0(s,y)| + |\phi(t,x,a) - \phi(s,y,a)| \le C_0(|x-y| + |t-s|^{1/2}),$$

there exists a unique bounded continuous viscosity solution of (2).

2.1. A second order BDF scheme. For the approximation in the x variable, we will consider the PDE on a truncated domain $\Omega := [x_{\min}, x_{\max}]$, where $x_{\min} < x_{\max}$.

Let $N \in \mathbb{N}^* \equiv \mathbb{N}\backslash\{0\}$ the number of time steps, $\tau := T/N$ the time step size, and $t_n = n\tau$, $n = 1, \ldots, N$. Let $I \in \mathbb{N}^*$ the number of interior mesh points, and define a uniform mesh $(x_i)_{1\le i\le I}$ with mesh size h by

$$x_i := x_{\min} + ih, \quad i \in \mathbb{I} = \{1, \ldots, I\}, \quad \text{where} \quad h := \frac{x_{\max} - x_{\min}}{I+1}.$$

Hereafter we denote by u a numerical approximation of v, the solution of (1), i.e.

$$u_i^k \sim v(t_k, x_i).$$

For each time step t_k, the unkowns are the values $(u^k)_i$ for $i = 1, \ldots, I$.

Standard Dirichlet boundary conditions use the knowledge of the values at the boundary, $v(t, x_0) \equiv v(t, x_{\min})$ and $v(t, x_{I+1}) \equiv v(t, x_{\max})$. Here, as a consequence of the size of the stencil below, we will assume that two left and right mesh boundary values are given, that is, $v(t, x_j)$ for $j \in \{-1, 0\}$ as well as $j \in \{I+1, I+2\}$ are known (corresponding to the values at the points $(x_{-1}, x_0, x_{I+1}, x_{I+2}) \equiv (x_{\min} - h, x_{\min}, x_{\max}, x_{\max} + h)$), or we can assume that a sufficiently accurate approximation of these "boundary values" is available. Indeed, it is not the focus of the paper to investigate boundary approximations and this framework is assumed in order to avoid such a discussion.

We then consider the following scheme, for $k \ge 2$, $i \in \mathbb{I}$,

$$\begin{aligned} &\mathcal{S}^{(\tau,h)}(t_k, x_i, u_i^k, [u]_i^k) = \\ &\frac{3u_i^k - 4u_i^{k-1} + u_i^{k-2}}{2\tau} + \sup_{a\in\Lambda}\Big\{L^a[u^k](t_k, x_i) + r(t_k, x_i, a)u_i^k + \ell(t_k, x_i, a)\Big\} \;=\; 0, \end{aligned} \tag{3}$$

where we denote as usual by $[u]_i^k$ the numerical solution excluding at (t_k, x_i), and

$$L^a[u](t_k, x_i) := -\frac{1}{2}\sigma^2(t_k, x_i, a)D^2u_i + b^+(t_k, x_i, a)D^{1,-}u_i - \; b^-(t_k, x_i, a)D^{1,+}u_i,$$

$$D^2u_i := \frac{u_{i-1} - 2u_i + u_{i+1}}{h^2},$$

(the usual second order approximation of v_{xx}), $b^+ := \max(b, 0)$ and $b^- := \min(b, 0)$ denote the positive and negative part of b, respectively, and where a second order BDF approximation (with stencil shifted left or right) is used for the first derivative in space:

$$D^{1,-}u_i := \frac{3u_i - 4u_{i-1} + u_{i-2}}{2h} \quad \text{and} \quad D^{1,+}u_i := -\left(\frac{3u_i - 4u_{i+1} + u_{i+2}}{2h}\right). \tag{4}$$

Notice in particular the implicit form of the scheme (3). The solution of this nonlinear implicit scheme will be addressed later on.

We will also define the numerical Hamiltonian associated with the scheme:

$$H[u](t_k, x_i) := \sup_{a\in\Lambda}\Big\{L^a[u](t_k, x_i) + r(t_k, x_i, a)u_i + \ell(t_k, x_i, a)\Big\}.$$

As discussed above, the scheme is completed by the following boundary conditions:

$$u_i^k := v(t_k, x_i), \quad \forall i \in \{-1, 0\} \cup \{I+1, I+2\}. \tag{5}$$

Since (3) is a two-step scheme, for the first time step $k = 1$, $i \in \mathbb{I}$, we use a backward Euler step,

$$\mathcal{S}^{(\tau,h)}(t_1, x_i, u_i^1, [u]_i^1) = \frac{u_i^1 - u_i^0}{\tau} + \sup_{a\in\Lambda}\Big\{L^a[u^1](t_1, x_i) + r(t_1, x_i, a)u_i^1 + \ell(t_1, x_i, a)\Big\} = 0, \tag{6}$$

and

$$u_i^0 = v_0(x_i), \quad i \in \mathbb{I} \tag{7}$$

is given by the initial condition (2b).

Remark 1. As the backward Euler step is only used once, it does not affect the overall second order of the scheme.

Remark 2. Most of our results also apply to the scheme obtained by replacing the BDF approximation (4) of the drift term by a centred finite difference approximation:

$$\widetilde{D}^{1,\pm}u_i := \frac{u_{i+1} - u_{i-1}}{2h}. \tag{8}$$

However, numerical tests (see Section 5.1) show that the BDF upwind approximation as in (4) has a better behaviour in some extreme cases where the diffusion vanishes.

2.2. Basic definitions and main result. In the remainder of this paper, we prove various stability results for the scheme (3). We state in this section the first main well-posedness and stability results.

Let u denote the solution of (3) and let v be the solution of (1). The error associated with the scheme is then defined by

$$e_i^k := u_i^k - v(t_k, x_i).$$

For any function ϕ we will also use the notation $\phi_i^k := \phi(t_k, x_i)$ as well as $\phi^k := (\phi_i^k)_{1\le i\le I}$ and $[\phi]_i^k := (\phi_j^m)_{(j,m)\neq(i,k)}$, and the error vector at time t_k is defined by

$$E^k := (e_1^k, \dots, e_I^k)^T \;=\; u^k - v^k.$$

The consistency error will be denoted by $\mathcal{E}^k(\phi) := (\mathcal{E}_i^k(\phi))_{1\le i\le I} \in \mathbb{R}^I$ and is defined in the classical way as follows, for any smooth enough function ϕ:

$$\mathcal{E}_i^k(\phi) := \mathcal{S}^{(\tau,h)}(t_k, x_i, \phi_i^k, [\phi]_i^k) - \left(\phi_t + \sup_{a\in\Lambda}\Big\{\mathcal{L}^a[\phi](t_k, x_i) + r(t_k, x_i, a)\phi + \ell(t_k, x_i, a)\Big\}\right).$$

By extension, for the exact solution v of (1), we will simply define

$$\mathcal{E}_i^k(v) := \mathcal{S}^{(\tau,h)}(t_k, x_i, v_i^k, [v]_i^k). \tag{9}$$

Note that (9) is well-defined for any continuous function.

In particular for the scheme (3) it is clear that we have second order consistency in space and time in smooth regions of ϕ, that is,

$$|\mathcal{E}_i^k(\phi)| \le c_1(\phi)\tau^2 + c_2(\phi)h^2. \tag{10}$$

More precisely, let $C^{p,q}$ denote the set of functions which are p times differentiable with respect to t and q times differentiable with respect to x. Assuming that $\phi \in C^{3,4}$, by Taylor expansion the constant $c_1(\phi)$ depends on ϕ_{3t} (the third derivative of ϕ with respect to the time variable), and the constant $c_2(\phi)$ depends on ϕ_{3x}, ϕ_{4x}, locally around the point of interest.

Throughout the paper, A will denote the finite difference matrix associated to the second order derivative, i.e.

$$A := \frac{1}{h^2}\begin{pmatrix} 2 & -1 & 0 & & \\ -1 & 2 & -1 & \ddots & \ddots \\ & -1 & \ddots & \ddots & 0 \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix}. \tag{11}$$

Let $\langle x, y\rangle_A := \langle x, Ay\rangle$. Then we consider the A-norm defined as follows:

$$|x|_A^2 := \langle x, Ax\rangle = \sum_{1\le i\le I+1} \left(\frac{x_i - x_{i-1}}{h}\right)^2 \tag{12}$$

(with the convention in (12) that $x_0 = x_{I+1} = 0$). Hence, $\sqrt{h}|x|_A$ approximates the H^1 norm in Ω. Similarly, we will consider later the standard Euclidean norm defined by $\|x\|^2 := \langle x, x\rangle$, such that $\sqrt{h}\|x\|$ approximates the L_2 norm.

Our first result concerns the solvability of the numerical scheme $\mathcal{S}^{\tau,h}$ (seen as an equation for u^k, with $[u]^k$ given) and is the following.

Assumption (A1) b and r are bounded functions.

THEOREM 3. *Let (A1) and the following CFL condition hold:*

$$\|b\|_\infty \frac{\tau}{h} < C. \tag{13}$$

Then, for τ small enough and $C = 3/2$ (resp. $C = 1$) there exists a unique solution of the scheme (3) for $k \ge 2$ (resp. $k = 1$, for scheme (6)).

Notice that the scheme is well-defined even if σ vanishes. A uniform ellipticity condition will be needed on σ for proving the stability of the scheme. We will provide a relaxation of this condition for stability in the Euclidean norm in Section 4.

Assumption (A2) There exists $\eta > 0$ such that

$$\inf_{t\in[0,T]} \inf_{x\in\Omega} \inf_{a\in\Lambda} \sigma^2(t,x,a) \;\ge\; \eta.$$

Our main result is the following.

THEOREM 4. *Assume (A1), (A2), as well as the CFL condition (13). Then there exists a constant $C \ge 0$ (independent of τ and h) and $\tau_0 > 0$ such that, for any $\tau \le \tau_0$,*

$$\max_{2\le k\le N} |E^k|_A^2 \le C\Big(|E^0|_A^2 + |E^1|_A^2 + \tau \sum_{2\le k\le N} |\mathcal{E}^k(v)|_A^2\Big). \tag{14}$$

The proof of Theorem 4 will be the subject of section 3.

Remark 5. As a consequence of the stability result and under further mild regularity assumptions on the data on the boundary of the domain one can also deduce that the scheme (3) is A-norm bounded,

$$\max_{2\leq k\leq N} |u^k|_A^2 \leq C, \tag{15}$$

where the constant C depends only on T and on the data but not on τ, h.

Remark 6. In general the exact solution v of (2) does not belong to $C^{3,4}$ which would give the estimate (10) and then from (14) a second order error bound immediately. Indeed, in degenerate cases, such as Test 1 in Section 5 for the Eikonal equation, it is well-known that v does not even belong to $C^{1,2}$ in general. However, at least in one-dimensional cases, it is very common that the solution is piecewise regular except at isolated singular points, and that v_t and v_{xx} may have jumps at these singular points and are otherwise *a.e.* bounded. Under such assumptions on the solution, it would be possible to obtain an error estimate by using the "strong" stability property of the scheme and by doing a direct error analysis.

More specific results in the Euclidean norm will be the focus of sections 4.

The extension of the presented results to other type of nonlinear operators (min, max min or min max) and corresponding equations will also be discussed.

3. Proof of main results.

In this section we prove Theorem 3 and Theorem 4.

3.1. Well-posedness of the scheme.

After multiplication by τ, the scheme (3) at time t_k (for $k \geq 2$) can be written in the following form:

$$\max_{a\in\Lambda}(M_a^k U - q_a^k) = 0,$$

where $q_a^k \in \mathbb{R}^I$ and $M_a^k \in \mathbb{R}^{I\times I}$ with the following non-zero entries:

$$(M_a^k)_{i,i} := \frac{3}{2} + \tau\left\{2\frac{\sigma^2}{h^2} + \frac{3b^+}{2h} + \frac{3b^-}{2h} + r\right\} \tag{16}$$

$$(M_a^k)_{i,i+1} := \tau\left\{-\frac{\sigma^2}{h^2} - \frac{4b^-}{2h}\right\}, \quad (M_a^k)_{i,i-1} := \tau\left\{-\frac{\sigma^2}{h^2} - \frac{4b^+}{2h}\right\} \tag{17}$$

$$(M_a^k)_{i,i+2} := \tau\frac{b^-}{2h} \quad (M_a^k)_{i,i-2} := \tau\frac{b^+}{2h} \tag{18}$$

with $\sigma \equiv \sigma(t_k, x_i, a)$, $b^\pm \equiv b^\pm(t_k, x_i, a)$ and $r \equiv r(t_k, x_i, a)$. For $k = 1$, the terms are different but the form (and analysis) is similar. The fact that the terms $(M_a)_{i,i\pm2}$ are nonnegative breaks the monotonicity of the scheme and therefore makes the analysis more difficult.

We will use the following lemma, whose proof is given in appendix A

LEMMA 7. *Asssume that Λ is a compact set, $(q_a)_{a\in\Lambda}$ is a family of vectors in $\mathbb{R}^I$, $(M_a)_{a\in\Lambda}$ is a family of matrices in $\mathbb{R}^{I\times I}$ such that:*

(i) for all $a \in \Lambda$,

$$(M_a)_{ii} > 0$$

(ii) (a form of diagonal dominance)

$$\sup_{a\in\Lambda}\max_{i\in\mathbb{I}} \frac{\sum_{j>i}|(M_a)_{ij}|}{|(M_a)_{ii}| - \sum_{j<i}|(M_a)_{ij}|} < 1. \tag{19}$$

Then there exists a unique solution X in $\mathbb{R}^n$ of

$$\max_{a\in\Lambda}(M_aX - q_a) = 0. \tag{20}$$

Remark 8. For a fixed $a \in \Lambda$, we have

$$\max_{i\in\mathbb{I}} \frac{\sum_{j>i}|(M_a)_{ij}|}{|(M_a)_{ii}| - \sum_{j<i}|(M_a)_{ij}|} < 1 \quad \Leftrightarrow \quad \min_{i\in\mathbb{I}}\left(|(M_a)_{ii}| - \sum_{j\neq i}|(M_a)_{ij}|\right) > 0.$$

However, condition (19) is not strictly equivalent to

$$\inf_{a\in\Lambda}\min_{i\in\mathbb{I}}\left(|(M_a)_{ii}| - \sum_{j\neq i}|(M_a)_{ij}|\right) > 0.$$

Proof of Theorem 3. We are going to prove properties (i) and (ii) in Lemma 7. Condition (i) is immediately verified, and we turn to proving (ii). We have (omitting the dependency on k and a in $\sigma, b^{\pm}, r$)

$$\mu_1 := \sum_{j>i}|(M_a)_{ij}| \leq \tau\left(\frac{\sigma_i^2}{h^2} + \frac{5b_i^-}{2h}\right)$$

and

$$\mu_2 := |(M_a)_{ii}| - \sum_{j<i}|(M_a)_{ij}| \geq \frac{3}{2} + \tau\left(\frac{\sigma_i^2}{h^2} - \frac{2b_i^+}{2h} + \frac{3b_i^-}{2h} + r\right).$$

By the CFL condition (13), there exists $\epsilon > 0$ such that $\frac{\tau\|b\|_\infty}{h} \leq \frac{3}{2} - \epsilon$. This implies in particular

$$\frac{3}{2} - \frac{\epsilon}{2} + \tau\left(-\frac{2b_i^+}{2h} + \frac{3b_i^-}{2h}\right) \geq \frac{\epsilon}{2} + \tau\frac{5b_i^-}{2h}$$

and therefore

$$\mu_1 \geq \left(\tau\frac{\sigma_i^2}{h^2} + \frac{\epsilon}{2} + \tau r\right) + \left(\tau\frac{5b_i^-}{2h} + \frac{\epsilon}{2}\right).$$

Then by using the fact that $\frac{a_1+a_2}{c_1+c_2} \leq \max\left(\frac{a_1}{c_1}, \frac{a_2}{c_2}\right)$ for numbers $a_i, c_i \geq 0$, we obtain

$$\frac{\mu_1}{\mu_2} \leq \max\left(\frac{\tau\frac{\sigma_i^2}{h^2}}{\tau\frac{\sigma_i^2}{h^2} + \frac{\epsilon}{2} + \tau r}, \frac{\tau\frac{5b_i^-}{2h}}{\tau\frac{5b_i^-}{2h} + \frac{\epsilon}{2}}\right).$$

Then taking τ small enough such that for instance $\frac{\epsilon}{2} + \tau r \geq \frac{\epsilon}{4}$, and since $b(.)$ and $\sigma(.)$ are bounded functions (by (A1)), we obtain the bound

$$\sup_{a\in A}\max_{i\in\mathbb{I}} \frac{\sum_{j>i}|(M_a)_{ij}|}{|(M_a)_{ii}| - \sum_{j<i}|(M_a)_{ij}|} \leq \max\left(\frac{\tau\frac{\|\sigma^2\|_\infty}{h^2}}{\tau\frac{\|\sigma^2\|_\infty}{h^2} + \frac{\epsilon}{4}}, \frac{\tau\frac{5\|b^-\|_\infty}{2h}}{\tau\frac{5\|b^-\|_\infty}{2h} + \frac{\epsilon}{2}}\right) < 1.$$

Since the last bound is a constant < 1, we can apply Lemma 7 to obtain the existence and uniqueness of the solution of the BDF scheme. □

3.2. Treatment of the nonlinearity. The proof consists of three main steps: first, we show a "linear" recursion for the error (Lemma 9); second, we pass from such a recursion involving the error in vector form to a scalar recursion (Lemma 10); finally, we show the stability estimate starting from this scalar recursion (Lemma 11).

By definition of the consistency error (9), one has (for $k \geq 2$, $1 \leq i \leq I$)

$$\frac{3v_i^k - 4v_i^{k-1} + v_i^{k-2}}{2\tau} + H[v^k](t_k, x_i) = \mathcal{E}_i^k. \tag{21}$$

The scheme simply reads

$$\frac{3u_i^k - 4u_i^{k-1} + u_i^{k-2}}{2\tau} + H[u^k](t_k, x_i) = 0. \tag{22}$$

Subtracting (21) from (22), denoting also $H[u^k] \equiv (H[u^k](t_k, x_i))_{1\leq i\leq I}$, the following recursion is obtained for the error in $\mathbb{R}^I$:

$$\frac{3E^k - 4E^{k-1} + E^{k-2}}{2\tau} + H[u^k] - H[v^k] = -\mathcal{E}^k. \tag{23}$$

For simplicity of presentation, we first consider the case when b and r vanish, i.e.

$$b(.) \equiv 0 \text{ and } r(.) \equiv 0.$$

In this case,

$$H[u^k]_i = \sup_{a\in\Lambda} \Big\{ -\frac{1}{2}\sigma^2(t_k, x_i, a)(D^2u^k)_i + \ell(t_k, x_i, a)\Big\}. \tag{24}$$

To simplify the presentation, we will assume that σ and ℓ are continuous functions of a so that the supremum is attained.[1] For each given k, i, let then $\bar{a}_i^k \in \Lambda$ denote an optimal control in (24).

In the same way, let $\bar{b}_i^k$ denote an optimal control for $H[v^k]_i$. By using the optimality of $\bar{a}_i^k$, it holds

$$\begin{aligned}
&H[u^k]_i - H[v^k]_i \\
&= -\frac{1}{2}\sigma^2(t_k, x_i, \bar{a}_i^k)(D^2u^k)_i + \ell(t_k, x_i, \bar{a}_i^k) - \sup_{a\in\Lambda}\Big\{-\frac{1}{2}\sigma^2(t_k, x_i, a)(D^2v^k)_i + \ell(t_k, x_i, a)\Big\} \\
&\leq -\frac{1}{2}\sigma^2(t_k, x_i, \bar{a}_i^k)(D^2u^k)_i - \Big(-\frac{1}{2}\sigma^2(t_k, x_i, \bar{a}_i^k)(D^2v^k)_i\Big) \\
&= -\frac{1}{2}\sigma^2(t_k, x_i, \bar{a}_i^k)(D^2E^k)_i
\end{aligned} \tag{25}$$

and, in the same way,

$$H[u^k]_i - H[v^k]_i \geq -\frac{1}{2}\sigma^2(t_k, x_i, \bar{b}_i^k)(D^2E^k)_i. \tag{26}$$

Therefore, combining (25) and (26), we see that $H[u^k]_i - H[v^k]_i$ is a convex combination of $-\frac{1}{2}\sigma^2(t_k, x_i, \bar{a}_i^k)(D^2E^k)_i$ and $-\frac{1}{2}\sigma^2(t_k, x_i, \bar{b}_i^k)(D^2E^k)_i$. In particular, we can write

$$H[u^k]_i - H[v^k]_i = -\frac{1}{2}\tilde{\sigma^2}(t_k, x_i)(D^2E^k)_i \tag{27}$$

[1] The general case is obtained easily by considering sequences of ϵ-optimal controls and letting $\epsilon \to 0$, such that (28) below still holds for a suitably defined $\tilde{\sigma^2}$, $\tilde{b^+}$, $\tilde{b^-}$, $\tilde{r}$.

where $\tilde{\sigma^2}(t_k, x_i)$ is a convex combination of $\sigma^2(t_k, x_i, \bar{a}_i^k)$ and $\sigma^2(t_k, x_i, \bar{b}_i^k)$. Obviously, in the general case (i.e. $b, r \not\equiv 0$) one gets

$$H[u^k]_i - H[v^k]_i = -\frac{1}{2}(\tilde{\sigma^2})_i^k D^2 E_i^k + (\tilde{b^+})_i^k D^{1,-} E_i^k - (\tilde{b^-})_i^k D^{1,+} E_i^k + \tilde{r}_i^k E_i^k \quad (28)$$

where, for $\phi = \sigma^2, b, r$,

$$\tilde{\phi}_i^k := \gamma_i^k \phi(t_k, x_i, \bar{a}_i^k) + (1 - \gamma_i^k)\phi(t_k, x_i, \bar{b}_i^k)$$

for some $\gamma_i^k \in [0, 1]$. We have proved the following:

LEMMA 9. *Let u be the solution of scheme (3) and v the solution of equation (2). The error $E^k = u^k - v^k$ satisfies for any $k \geq 2$ and $i \in \mathbb{I}$:*

$$\frac{3E_i^k - 4E_i^{k-1} + E_i^{k-2}}{2\tau} - \frac{1}{2}(\tilde{\sigma^2})_i^k D^2 E_i^k + (\tilde{b^+})_i^k D^{1,-} E_i^k - (\tilde{b^-})_i^k D^{1,+} E_i^k + \tilde{r}_i^k E_i^k = -\mathcal{E}_i^k. \quad (29)$$

3.3. From a vector recursion to a scalar recursion. Let us start again by considering the case $b, r \equiv 0$. Let

$$\Delta^k := \frac{1}{2} \operatorname{diag}\left((\tilde{\sigma^2})_i^k\right)_{i \in \mathbb{I}}. \quad (30)$$

We remark that for $x \in \mathbb{R}^I$,

$$-D^2 x = Ax,$$

where A is the finite difference matrix defined in (11). Hence the following equality holds in $\mathbb{R}^I$: $H[u^k] - H[v^k] = \Delta^k A E^k$, and therefore, from (27), one has the following vector recursion:

$$\frac{3E^k - 4E^{k-1} + E^{k-2}}{2\tau} + \Delta^k A E^k = -\mathcal{E}^k. \quad (31)$$

Taking the scalar product of (31) with AE^k, and using that

$$\langle \Delta^k A E^k,\ A E^k \rangle \geq 0,$$

since Δ^k has positive coefficients, the following inequality holds:

$$\langle 3E^k - 4E^{k-1} + E^{k-2},\ E^k \rangle_A \leq -2\tau \langle \mathcal{E}^k,\ E^k \rangle_A. \quad (32)$$

By using the identity $2(a - b, a)_A = |a|_A^2 + |a - b|_A^2 - |b|_A^2$, one has:

$$\begin{aligned}
&\langle 3E^k - 4E^{k-1} + E^{k-2}, E^k \rangle_A \\
&= 4\langle E^k - E^{k-1}, E^k \rangle_A - \langle E^k - E^{k-2}, E^k \rangle_A \\
&= \frac{1}{2}\left(4|E^k|_A^2 + 4|E^k - E^{k-1}|_A^2 - 4|E^{k-1}|_A^2\right) - \frac{1}{2}\left(|E^k|_A^2 + |E^k - E^{k-2}|_A^2 - |E^{k-2}|_A^2\right) \\
&= \frac{1}{2}\left(3|E^k|_A^2 - 4|E^{k-1}|_A^2 + |E^{k-2}|_A^2 + 4|E^k - E^{k-1}|_A^2 - |E^k - E^{k-2}|_A^2\right) \\
&\geq \frac{1}{2}\left(3|E^k|_A^2 - 4|E^{k-1}|_A^2 + |E^{k-2}|_A^2\right) + |E^k - E^{k-1}|_A^2 - |E^{k-1} - E^{k-2}|_A^2, \quad (33)
\end{aligned}$$

where we have also used $|a+b|^2 \leq 2|a|^2 + 2|b|^2$. In the end we obtain the following scalar recursion, for $k \geq 2$:

$$\frac{1}{2}\Big(3|E^k|_A^2 - 4|E^{k-1}|_A^2 + |E^{k-2}|_A^2\Big) + |E^k - E^{k-1}|_A^2 - |E^{k-1} - E^{k-2}|_A^2 \leq 2\tau |E^k|_A \, |\mathcal{E}^k|_A \tag{34}$$

Now let us come back to the more general case when drift and growth terms ($b(.)$ and $r(.)$) are present. For simplicity of presentation we will assume that b has constant positive sign. The terms coming from the negative part of b can be treated in a similar way.
By (28), we get the following recursion

$$\frac{3E^k - 4E^{k-1} + E^{k-2}}{2\tau} + \Delta^k A E^k + F^k B E^k + R^k E^k = -\mathcal{E}^k, \tag{35}$$

where

$$B = \frac{1}{2h}\begin{pmatrix} 3 & 0 & & & \\ -4 & 3 & 0 & & \\ 1 & -4 & \ddots & \ddots & \\ 0 & \ddots & \ddots & \ddots & 0 \\ \ddots & \ddots & 1 & -4 & 3 \end{pmatrix}, \qquad F^k = \mathrm{diag}(\tilde{b})_i^k, \qquad R_k = \mathrm{diag}(\tilde{r})_i^k.$$

We obtain a similar bound as in (34), except for a modification of the factor of $|E^k|_A$:

LEMMA 10. *Let assumptions (A1) and (A2) in Theorem 4 be satisfied. Then there exists a constant $C \geq 0$ such that*

$$\frac{1}{2}\Big((3 - C\tau)|E^k|_A^2 - 4|E^{k-1}|_A^2 + |E^{k-2}|_A^2\Big) + |E^k - E^{k-1}|_A^2 - |E^{k-1} - E^{k-2}|_A^2 \\ \leq 2\tau |E^k|_A \, |\mathcal{E}^k|_A. \tag{36}$$

Proof. We form the scalar product of (35) with AE^k. Using the fact that $(\sigma^2)_i^k \geq \eta > 0$ for all k, i, one has

$$\langle \Delta^k A E^k, A E^k \rangle \geq \frac{\eta}{2} \|A E^k\|^2, \tag{37}$$

where $\|\cdot\|$ denotes the canonical Euclidean norm in $\mathbb{R}^I$.

In order to estimate the drift component, let us introduce the notation

$$\delta e := (e_i - e_{i-1})_{1 \leq i \leq I} \tag{38}$$

with the convention that $e_i = 0$ for all indices i which are not in $\mathbb{I}$. It holds:

$$\begin{aligned} |\langle F^k B E^k, A E^k \rangle| &= \left| \frac{1}{2h} \langle F^k (3e_i^k - 4e_{i-1}^k + e_{i-2}^k)_{i \in \mathbb{I}}, A E^k \rangle \right| \\ &= \left| \frac{1}{2h} \langle F^k \big(3\delta E^k - \delta^2 E^k\big), \; A E^k \rangle \right| \\ &\leq \frac{1}{2h} \Big\{ 3 \|F^k \delta E^k\| \, \|A E^k\| + \|F^k \delta^2 E^k\| \, \|A E^k\| \Big\}. \end{aligned}$$

By using the boundedness of the drift term, and the fact that $\|\delta E^k\|, \|\delta^2 E^k\| \leq h|E^k|_A$, one gets

$$\begin{aligned}|\langle F^k B E^k, A E^k\rangle| &\leq \frac{\|b\|_\infty}{2h}\Big\{3\|AE^k\|\|\delta E^k\| + \|AE^k\|\|\delta^2 E^k\|\Big\}\\ &\leq 2\|b\|_\infty \, \|AE^k\| \, |E^k|_A. \end{aligned} \tag{39}$$

For the last term, by using the boundedness of r and the Cauchy-Schwartz inequality, we obtain

$$|\langle R^k E^k, A E^k\rangle| \leq \|r\|_\infty \|E^k\|\|AE^k\|. \tag{40}$$

Therefore, putting (37), (39) and (40) together,

$$\begin{aligned}&\langle \Delta^k A E^k + F^k B E^k + R^k E^k, A E^k\rangle\\ &\qquad \geq \frac{\eta}{2}\|AE^k\|^2 - 2\|b\|_\infty \|AE^k\| \, |E^k|_A - \|r\|_\infty \|AE^k\|\|E^k\|. \end{aligned} \tag{41}$$

Elementary calculus shows that the minimal eigenvalue of A is $\lambda_{\min}(A) = \frac{4}{h^2}\sin^2(\frac{\pi h}{2})$, and that $\lambda_{\min}(A) \geq 4$. Hence $\langle X, AX\rangle \geq 4\langle X, X\rangle$ and therefore $\|X\| \leq \frac{1}{2}|X|_A$. In the same way, we have also $|X|_A \leq \frac{1}{2}\|AX\|$.

Hence it holds

$$\langle \Delta^k A E^k + F^k B E^k + R^k E^k, A E^k\rangle \geq \frac{\eta}{2}\|AE^k\|^2 - C_1\|AE^k\||E^k|_A \tag{42}$$

with $C_1 := 2\|b\|_\infty + \frac{1}{2}\|r\|_\infty$. By using $C_1\|AE^k\||E^k|_A \leq \frac{\eta}{2}\|AE^k\|^2 + \frac{1}{2\eta}C_1^2|E^k|_A^2$, we finally obtain

$$\langle \Delta^k A E^k + F^k B E^k + R^k E^k, A E^k\rangle \geq -\frac{1}{2\eta}C_1^2|E^k|_A^2. \tag{43}$$

Then, combining (33) and the last bound (43), we obtain the desired inequality with the constant $C := \frac{1}{2\eta}C_1^2$. ∎

3.4. Stability lemma. In the following, it is assumed that $|\cdot|$ is any vectorial norm. We will use the result for the canonical Euclidean norm $|\cdot| \equiv \|\cdot\|$ and the A-norm $|\cdot| \equiv |\cdot|_A$.

In order to prove the following Lemma 11, we will exploit some properties of matrices of the following form:

$$M_\tau := \begin{pmatrix} (3-C\tau) & -4 & 1 & 0 & \\ 0 & (3-C\tau) & -4 & \ddots & \ddots \\ & 0 & \ddots & \ddots & 1 \\ & & \ddots & \ddots & -4 \\ & & & 0 & (3-C\tau) \end{pmatrix}, \tag{44}$$

in particular the fact that $(M_\tau)^{-1} \geq 0$ for τ small enough.

LEMMA 11. *Assume that there exists a constant $C \geq 0$ such that $\forall k = 2, \ldots, N$:*

$$\frac{1}{2}\Big((3-C\tau)|E^k|^2 - 4|E^{k-1}|^2 + |E^{k-2}|^2\Big) + |E^k - E^{k-1}|^2 - |E^{k-1} - E^{k-2}|^2 \leq 2\tau|E^k| \, |\mathcal{E}^k|. \tag{45}$$

Then there exists a constant $C_1 \geq 0$ and $\tau_0 > 0$ such that $\forall 0 < \tau \leq \tau_0$, $\tau N = T$ fixed, $\forall n \leq N$:

$$\max_{2\leq k\leq n} |E^k|^2 \leq C_1\Big(|E^0|^2 + |E^1|^2 + \tau \sum_{2\leq j\leq n} |\mathcal{E}^j|^2\Big).$$

Proof. Let us denote

$$x_k := |E^k|^2 \qquad \text{and} \qquad y_k := |E^k - E^{k-1}|^2,$$

so that (45) reads

$$\Big((3 - C\tau)x_k - 4x_{k-1} + x_{k-2}\Big) \leq 2(y_{k-1} - y_k) + 4\tau|E^k|\ |\mathcal{E}^k|. \tag{46}$$

For a given $\tau > 0$ and given k, let $M_\tau \in \mathbb{R}^{(k-1)\times(k-1)}$ as defined in (44). Let $z, w \in \mathbb{R}^{k-1}$ be defined by

$$z := (x_k, x_{k-1}, \dots, x_2)^T \quad \text{and} \qquad w := (2(y_{j-1} - y_j) + 4\tau|E^j|\ |\mathcal{E}^j|)_{j=k,\dots,2}.$$

By (46), we have

$$M_\tau z \leq w. \tag{47}$$

We notice that $M_\tau = (3 - C\tau)I - 4J + J^2$ with

$$J := \text{tridiag}(0, 0, 1).$$

Hence, with

$$\lambda_1 = 2 + \sqrt{1 + C\tau} \quad \text{and} \quad \lambda_2 = 2 - \sqrt{1 + C\tau},$$

the roots of $\lambda^2 - 4\lambda + (3 - C\tau) = 0$ for $3 - C\tau \geq 0$, we can write

$$M_\tau = (\lambda_1 I - J)(\lambda_2 I - J) = \lambda_1\lambda_2 \left(I - \frac{J}{\lambda_1}\right)\left(I - \frac{J}{\lambda_2}\right).$$

Furthermore, since $J^{k-1} = 0$, it holds

$$\begin{aligned} M_\tau^{-1} &= \frac{1}{\lambda_1\lambda_2}\left(I - \frac{J}{\lambda_1}\right)^{-1}\left(I - \frac{J}{\lambda_2}\right)^{-1} \\ &= \frac{1}{\lambda_1\lambda_2}\left(\sum_{0\leq q\leq k-2}\left(\frac{J}{\lambda_1}\right)^q\right)\left(\sum_{0\leq q\leq k-2}\left(\frac{J}{\lambda_2}\right)^q\right) = \sum_{p=0}^{k-2} a_p J^p, \end{aligned}$$

where

$$a_p := \sum_{i=0}^{p} \frac{1}{\lambda_1^{j+1}\lambda_2^{p-j+1}} = \frac{1}{\lambda_2^{p+2}} \sum_{i=0}^{p} \left(\frac{\lambda_2}{\lambda_1}\right)^{j+1}.$$

Therefore $M_\tau^{-1} \geq 0$ componentwise (for $\tau < 3/C$), and using (47) it holds $z \leq M_\tau^{-1} w$.

It is possible to prove that there exists $\tau_0 > 0$ and a constant $C_0 \geq 0$ (depending only on T) such that $\forall 0 < \tau \leq \tau_0$ and $\forall p \geq 0$:

$$0 \leq a_p \leq C_0 \quad \text{and} \quad a_p - a_{p-1} \geq 0. \tag{48}$$

We postpone the proof of (48) to the end of the proof. For the first component of z, we deduce, for all $k \geq 2$:

$$\begin{aligned} x_k &\leq a_0 w_1 + a_1 w_2 + \cdots + a_{k-2} w_{k-1} \\ &\leq 2a_0(y_{k-1} - y_k) + 2a_1(y_{k-2} - y_{k-1}) + \cdots + 2a_{k-2}(y_1 - y_2) + 4C_0\tau \sum_{j=2}^{k} |E^j|\,|\mathcal{E}^j| \\ &= -2a_0 y_k + 2(a_0 - a_1)y_{k-1} + \cdots + 2(a_{k-3} - a_{k-2})y_2 + 2a_{k-2}y_1 + 4C_0\tau \sum_{j=2}^{k} |E^j|\,|\mathcal{E}^j|, \end{aligned} \tag{49}$$

where, for (49), we have used the fact that $a_p \leq C_0$. Since $y_j \geq 0$, $\forall j$, by definition, $a_{k-2} \leq C_0$, $a_0 = \frac{1}{\lambda_1\lambda_2} \geq 0$ and $a_j - a_{j-1} \geq 0$, $\forall j$, we obtain

$$x_k \leq 2C_0 y_1 + 4C_0\tau \sum_{j=2}^{k} |E^j|\,|\mathcal{E}^j|. \tag{50}$$

Recalling the definition of x_k and y_k, for any $2 \leq k \leq n$ one has:

$$\begin{aligned} |E^k|^2 &\leq 2C_0|E^1 - E^0|^2 + 4C_0\tau \sum_{j=2}^{k} |E^j|\,|\mathcal{E}^j| \\ &\leq 4C_0(|E^0|^2 + |E^1|^2) + 4C_0\tau\Big(\max_{2\leq k\leq n} |E^k|\Big) \sum_{j=2}^{n} |\mathcal{E}^j| \\ &\leq 4C_0(|E^0|^2 + |E^1|^2) + \frac{1}{2}\Big(\max_{2\leq k\leq n} |E^k|\Big)^2 + 8C_0^2\tau^2\Big(\sum_{j=2}^{n} |\mathcal{E}^k|\Big)^2 \end{aligned}$$

(where we made use of $2ab \leq \frac{a^2}{K} + Kb^2$ for any $a, b \geq 0$ and $K > 0$). Hence, we obtain

$$\frac{1}{2}\Big(\max_{2\leq k\leq n} |E^k|\Big)^2 \leq 4C_0(|E^0|^2 + |E^1|^2) + 8C_0^2\tau^2\Big(\sum_{j=2}^{n} |\mathcal{E}^k|\Big)^2,$$

which leads to

$$\begin{aligned} \Big(\max_{2\leq k\leq n} |E^k|\Big)^2 &\leq 8C_0|E^0|^2 + 8C_0|E^1|^2 + 16C_0^2\tau^2\Big(\sum_{j=2}^{n} |\mathcal{E}^k|\Big)^2 \\ &\leq C_1\Big(|E^0|^2 + |E^1|^2 + \tau\sum_{j=2}^{n} |\mathcal{E}^k|^2\Big) \end{aligned}$$

with $C_1 := \max(8C_0, 16C_0^2T)$ (we used $\Big(\sum_{j=2}^{n} |\mathcal{E}^k|\Big)^2 \leq n\sum_{j=2}^{n} |\mathcal{E}^k|^2$ and $n\tau \leq T$).

It remains to prove (48). From the definition of a_p one has

$$a_p = \frac{1}{\lambda_2^{p+2}} \sum_{i=0}^{p} \left(\frac{\lambda_2}{\lambda_1}\right)^{j+1} \leq \frac{1}{\lambda_2^{p+2}} \left(1 - \frac{\lambda_2}{\lambda_1}\right)^{-1}$$

for $p = 0, \dots, k-2$. Observing that $\frac{\lambda_2}{\lambda_1} \leq \frac{1}{3}$, it follows that

$$a_p \leq \frac{3}{2\lambda_2^{p+2}} \leq \frac{3}{2(2-\sqrt{1+C\tau})^n}.$$

Notice that $\sqrt{1+C\tau} \leq 1 + C\tau$, and also that $e^{-x} \leq 1 - x/2$, $\forall x \in [0,1]$. Hence $(2-\sqrt{1+C\tau})^n \geq (2-(1+C\tau))^n = (1-C\tau)^n \geq (e^{-2C\tau})^n = e^{-2Ct_n}$ for $C\tau \leq \frac{1}{2}$, and therefore $a_p \leq \frac{3}{2}e^{2Ct_n}$. The desired result follows with $C_0 := \frac{3}{2}e^{2CT}$ and $\tau_0 := \frac{1}{2C}$.

Moreover, one has

$$a_p - a_{p-1} = \frac{1}{\lambda_2^{p+1}} \left(\frac{1}{\lambda_2} \sum_{i=0}^{p} \left(\frac{\lambda_2}{\lambda_1} \right)^{j+1} - \sum_{i=0}^{p-1} \left(\frac{\lambda_2}{\lambda_1} \right)^{j+1} \right),$$

which is nonnegative for τ small enough thanks to the fact that $\lambda_1, \lambda_2 \geq 0$ and $\lambda_2 \leq 1$. □

3.5. Proof of Theorem 4. The strong stability result (14) follows directly from Lemma 9, Lemma 10 and Lemma 11.

3.6. Isaacs equations. The same technique used in Section 3.2 to deal with the nonlinear operator applies also for Isaacs equations, i.e. equations of the following form

$$v_t + \sup_{a \in \Lambda_1} \inf_{b \in \Lambda_2} \Big\{ -\mathcal{L}^{(a,b)}[v](t,x) + r(t,x,a,b)v + \ell(t,x,a,b) \Big\} = 0, \tag{51}$$

where $(t,x) \in [0,T] \times \mathbb{R}^d$, $\Lambda_1, \Lambda_2 \subset \mathbb{R}^m$ are compact sets and

$$\mathcal{L}^{(a,b)}[v](t,x) = \frac{1}{2}\sigma^2(t,x,a,b)v_{xx} + b(t,x,a,b)v_x.$$

To simplify the presentation, let us consider $b, r, \ell \equiv 0$. We get, by analogous definitions and reasoning to Section 3.2,

$$\frac{3E^k - 4E^{k-1} + E^{k-2}}{2\tau} + H[u^k] - H[v^k] = -\mathcal{E}^k,$$

where

$$H[u^k]_i = \sup_{a \in \Lambda_1} \inf_{b \in \Lambda_2} \Big\{ -\frac{1}{2}\sigma^2(t,x,a,b)(D_x^2 u^k)_i \Big\}. \tag{52}$$

Let $(\bar{a}_i^k, \bar{b}_i^k) \in \Lambda_1 \times \Lambda_2$ denote an optimal control in (52).[2] One has

$$H[u^k]_i = \sup_{a \in \Lambda_1} \Big\{ -\frac{1}{2}\sigma^2(t,x,a,\bar{b}_i^k)(D_x^2 u^k)_i \Big\} = \inf_{b \in \Lambda_2} \Big\{ -\frac{1}{2}\sigma^2(t,x,\bar{a}_i^k,b)(D_x^2 v^k)_i \Big\}.$$

Therefore

$$\begin{aligned}
&H[u^k]_i - H[v^k]_i \\
&= \sup_{a \in \Lambda_1} \Big\{ -\frac{1}{2}\sigma^2(t,x,a,\bar{b}_i^k)(D_x^2 u^k)_i \Big\} - \sup_{a \in \Lambda_1} \inf_{b \in \Lambda_2} \Big\{ -\frac{1}{2}\sigma^2(t_k,x_i,a,b)(D^2 v^k)_i \Big\} \\
&\geq \sup_{a \in \Lambda_1} \Big\{ -\frac{1}{2}\sigma^2(t,x,a,\bar{b}_i^k)(D_x^2 u^k)_i \Big\} - \sup_{a \in \Lambda_1} \Big\{ -\frac{1}{2}\sigma^2(t_k,x_i,a,\bar{b}_i^k)(D^2 v^k)_i \Big\} \\
&\geq \inf_{a \in \Lambda_1} \Big\{ -\frac{1}{2}\sigma^2(t,x,a,\bar{b}_i^k)(D_x^2 E^k)_i \Big\}.
\end{aligned} \tag{53}$$

[2] Or, if not attained, use an approximation argument as in Section 3.2.

Analogously, one can prove

$$H[u^k]_i - H[v^k]_i \leq \sup_{b\in\Lambda_2}\Big\{-\frac{1}{2}\sigma^2(t,x,\bar a_i^k,b)(D_x^2E^k)_i\Big\} \tag{54}$$

(here, we also use $\inf(a)-\inf(b)\leq\sup(a-b)$ and $\sup(a)-\sup(b)\geq\inf(a-b)$). At this point, it is sufficient to take for $\hat a_i^k\in\Lambda_1$ and $\hat b_i^k\in\Lambda_2$ optimal controls in (53) and (54), respectively, to be able to write $H[u^k]_i-H[v^k]_i$ as a convex combination of $-\frac{1}{2}\sigma^2(t,x,\hat a_i^k,\bar b_i^k)(D_x^2E^k)_i$ and $-\frac{1}{2}\sigma^2(t,x,\bar a_i^k,\hat b_i^k)(D_x^2E^k)_i$.

From this, an equation exactly as in (29) can be derived, with a suitable convex combination $(\tilde{\sigma^2})_i^k$ of diffusion coefficients, and similar for the drift and other terms.

4. Euclidean norm: some specific results. The fundamental stability result given by Lemma 11, applies to any vectorial norm. In view of this, in this section we are going to discuss some special cases where (45) can be obtained for $|\cdot|=\|\cdot\|$, the Euclidean norm. We also saw that applying the arguments in Section 3.2 (see Lemma 9 and Section 3.6) we can pass from a nonlinear equation to a linear one and for this reason we start here by considering the linear case.

4.1. One dimensional case. We consider the following linear equation

$$v_t-\frac{1}{2}\sigma^2(t,x)v_{xx}+b(t,x)v_x+r(t,x)v+\ell(t,x)=0. \tag{55}$$

Let us make the following assumptions:
Assumption (A3) There exists $L\geq 0$ such that

$$|\sigma^2(t,x)-\sigma^2(t,y)|\leq L|x-y|\quad\forall x,y\in\Omega,t\in[0,T].$$

PROPOSITION 12. *Let assumptions (A1),(A2) and (A3) be satisfied. Then* (45) *holds for* $|\cdot|=\|\cdot\|$.

Proof. We recall that (35) holds for E^k with

$$\Delta^k=\frac{1}{2}\operatorname{diag}\left((\sigma_i^k)^2\right),\quad F^k=\operatorname{diag}\left(b_i^k\right),\quad R^k=\operatorname{diag}\left(r_i^k\right).$$

Scalarly multiplying by E^k one gets

$$\langle E^k,3E^k-4E^{k-1}+E^{k-2}\rangle+2\tau\langle E^k,\Delta_kAE^k+F_kBE^k+R^kE^k\rangle=-2\tau\langle E^k,\mathcal{E}^k\rangle. \tag{56}$$

By the same passages used in Section 3.3 one can obtain

$$\begin{aligned}&\langle E^k,3E^k-4E^{k-1}+E^{k-2}\rangle \\ &\geq\frac{1}{2}\left(3\|E^k\|^2-4\|E^{k-1}\|^2+\|E^{k-2}\|^2\right)+\|E^k-E^{k-1}\|^2-\|E^{k-1}-E^{k-2}\|^2.\end{aligned} \tag{57}$$

Also,

$$\langle E^k,R^kE^k\rangle\geq-\|r\|_\infty\|E^k\|^2. \tag{58}$$

Using the Lipschitz continuity of σ^2 and the zero boundary conditions, one has

$$\begin{aligned}\langle E^k, \Delta^k A E^k\rangle &= \sum_{i\in\mathbb{I}} \frac{(\sigma_i^k)^2}{2h^2}(-e_{i+1}^k + 2e_i^k - e_{i-1}^k)e_i^k \\ &= \sum_{i\in\mathbb{I}} \frac{(\sigma_{i-1}^k)^2}{2h^2}(e_{i-1}^k - e_i^k)e_{i-1}^k + \sum_{i\in\mathbb{I}} \frac{(\sigma_i^k)^2}{2h^2}(e_i^k - e_{i-1}^k)e_i^k \\ &= \sum_{i\in\mathbb{I}} \frac{(\sigma_{i-1}^k)^2}{2h^2}(e_{i-1}^k - e_i^k)^2 + \sum_{i\in\mathbb{I}}(e_{i-1}^k - e_i^k)\left(\frac{(\sigma_{i-1}^k)^2}{2h^2} - \frac{(\sigma_i^k)^2}{2h^2}\right)e_i^k \\ &\geq \frac{\eta}{2h^2}\sum_{i\in\mathbb{I}}(e_{i-1}^k - e_i^k)^2 - \frac{L}{2h}\sum_{i\in\mathbb{I}}|e_{i-1}^k - e_i^k||e_i^k|.\end{aligned} \tag{59}$$

Therefore, by Cauchy-Schwarz inequality, one obtains

$$\langle E^k, \Delta_k A E^k\rangle \geq \frac{\eta}{2h^2}\|\delta E^k\|^2 - \frac{L}{2h}\|\delta E^k\|\|E^k\|, \tag{60}$$

where δE^k is defined by (38). Moreover, for the first order term one has

$$\begin{aligned}\langle E^k, F^k B E^k\rangle &= \sum_{i\in\mathbb{I}} \frac{b_i}{2h}(3e_i^k - 4e_{i-1}^k + e_{i-2}^k)e_i^k \\ &\geq -\frac{3\|b\|_\infty}{2h}\sum_{i\in\mathbb{I}}|e_i^k - e_{i-1}^k||e_i^k| - \frac{\|b\|_\infty}{2h}\sum_{i\in\mathbb{I}}|e_{i-1}^k - e_{i-2}^k||e_i^k| \\ &= -\frac{2\|b\|_\infty}{h}\|\delta E^k\|\|E^k\|,\end{aligned} \tag{61}$$

where for the last equality we have used that $\|\delta E^k\| = \|\delta^2 E^k\|$. Putting together estimates (58),(60) and (61) and denoting $C_1 = L + 4\|b\|_\infty$ we get

$$\begin{aligned}\langle E^k, \Delta_k A E^k + F_k B E^k + R^k E^k\rangle &\geq \frac{\eta}{2h^2}\|\delta E^k\|^2 - \frac{C_1}{2h}\|\delta E^k\|\|E^k\| - \|r\|_\infty\|E^k\|^2 \\ &= \frac{\eta}{4h^2}\|\delta E^k\|^2 - \frac{C_1^2}{4\eta}\|E^k\|^2 - \|r\|_\infty\|E^k\|^2.\end{aligned}$$

Hence, together with (57), this gives (45) with $|\cdot| = \|\cdot\|$ and $C = \frac{C_1^2}{2\eta} + 2\|r\|_\infty$. ∎

Remark 13. In the case of controlled equations, assumptions (A1), (A2) and (A3) have to be satisfied also when the dependence of the coefficients on the control is taken into account. In assumptions (A1) and (A2), the bound is already considered uniformly with respect to the control, whereas the adaptation of assumption (A3) to the controlled case would impose some Lipschitz regularity of the control with respect to the state variable. Such regularity of the control cannot usually be expected (see for instance the tests proposed in Section 5). However, the case where the controls only appear in the drift term would be a reasonable setting for assumption (A3).

The next result concerns the case of a possibly degenerate diffusion term. This requires stricter assumptions on the drift and diffusion. Indeed, in this case, one

cannot count on the positive term coming from the non-degenerate diffusion which, in the proof of Proposition 12, is used to compensate the negative correction terms coming from the drift term and the Lipschitz rescaling of the diffusion coefficient. This leads us to consider the following assumptions:

Assumption (A4) r is bounded, b has constant sign and there exist $L, L_1 \geq 0$ such that

$$\begin{aligned} |b(t,x) - b(t,y)| &\leq L|x-y| && \forall x,y \in \Omega, t \in [0,T]; \\ |D_x(\sigma^2)(t,x) - D_x(\sigma^2)(t,y)| &\leq L_1|x-y| && \forall x,y \in \Omega, t \in [0,T]. \end{aligned}$$

PROPOSITION 14. *Let assumption (A4) be satisfied. Then (45) holds for $|\cdot| = \|\cdot\|$.*

Proof. We consider again the scalar recursion (56). Clearly, $\langle E^k, 3E^k - 4E^{k-1} + E^{k-2}\rangle$ and $\langle E^k, R^k E^k\rangle$ work as in Proposition 12.

We begin by observing that, for the diffusive part,

$$\langle E^k, \Delta_k A E^k\rangle \geq -\frac{L_1}{2}|E^k|^2 \quad \Leftrightarrow \quad \sum_{i,j\in\mathbb{I}} e_i^k \, (\Delta_k A)_{ij} \, e_j^k \geq -\frac{L_1}{2}\sum_{i\in\mathbb{I}}(e_i^k)^2$$

for all $E^k = (e_i^k)_{1\leq i\leq I} \in \mathbb{R}^I$.

By the assumption on σ,

$$\left|\frac{\sigma^2(t,x_i+h) - 2\sigma^2(t,x_i) + \sigma^2(t,x_i-h)}{h^2}\right| \leq \frac{L_1}{2}.$$

It then follows that, using $a^2+b^2 \geq ab$ in the second line and the boundary conditions for the index shift in the third,

$$\begin{aligned} \langle E^k, \Delta_k A E^k\rangle &= \frac{1}{h^2}\sum_{i\in\mathbb{I}} e_i^k\sigma_i^2\left(-e_{i-1}^k + 2e_i^k - e_{i+1}^k\right) \\ &\geq \frac{1}{h^2}\sum_{i\in\mathbb{I}} \sigma_i^2\left(-\frac{(e_{i-1}^k)^2 + (e_i^k)^2}{2} + 2(e_i^k)^2 - \frac{(e_{i+1}^k)^2 + (e_i^k)^2}{2}\right) \\ &= \frac{1}{h^2}\sum_{i\in\mathbb{I}} -\frac{1}{2}(e_i^k)^2\left(\sigma_{i+1}^2 + 2\sigma_i^2 - \sigma_{i-1}^2\right) \\ &\geq -\frac{L_1}{2}\sum_{i\in\mathbb{I}}(e_i^k)^2, \end{aligned}$$

as we wanted to show.

Moreover, one has

$$\begin{aligned}\langle E^k, F^k B E^k\rangle &= \sum_{i\in\mathbb{I}} \frac{b_i^k}{2h}(3e_i^k - 4e_{i-1}^k + e_{i-2}^k)e_i^k \\ &= \sum_{i\in\mathbb{I}} \frac{4b_i^k}{2h}(e_i^k - e_{i-1}^k)e_i^k - \frac{b_i^k}{2h}(e_i^k - e_{i-2}^k)e_i^k \\ &= \sum_{i\in\mathbb{I}} \frac{b_i^k}{h}\Big((e_i^k)^2 + (e_i^k - e_{i-1}^k)^2 - (e_{i-1}^k)^2\Big) - \frac{b_i^k}{4h}\Big((e_i^k)^2 + (e_i^k - e_{i-2}^k)^2 - (e_{i-2}^k)^2\Big) \\ &\geq \sum_{i\in\mathbb{I}} \frac{b_i^k}{4h}\Big(3(e_i^k)^2 - 4(e_{i-1}^k)^2 + (e_{i-2}^k)^2\Big) \pm \frac{b_{i-1}^k}{h}(e_{i-1}^k)^2 \pm \frac{b_{i-2}^k}{4h}(e_{i-2}^k)^2 \\ &\quad + \frac{b_i^k}{2h}\Big((e_i^k - e_{i-1}^k)^2 - (e_{i-1}^k - e_{i-2}^k)^2\Big) \pm \frac{b_{i-1}^k}{2h}(e_{i-1}^k - e_{i-2}^k)^2.\end{aligned}$$

Observing that (using the asymptotic zero boundary conditions)

$$\sum_{i\in\mathbb{I}} b_{i-j}^k (e_{i-j}^k)^2 = \sum_{i\in\mathbb{I}} b_i^k (e_i^k)^2,\ j=1,2, \quad \sum_{i\in\mathbb{I}} b_{i-1}^k (e_{i-1}^k - e_{i-2}^k)^2 = \sum_{i\in\mathbb{I}} b_i^k (e_i^k - e_{i-1}^k)^2,$$

one obtains

$$\begin{aligned}\langle E^k, F^k B E^k\rangle &\geq \sum_{i\in\mathbb{I}} \Big(\frac{b_{i-1}^k}{h} - \frac{b_i^k}{h}\Big)(e_{i-1}^k)^2 + \Big(\frac{b_i^k}{4h} - \frac{b_{i-2}^k}{4h}\Big)(e_{i-2}^k)^2 \\ &\quad + \Big(\frac{b_{i-1}^k}{2h} - \frac{b_i^k}{2h}\Big)(e_{i-1}^k - e_{i-2}^k)^2 \\ &\geq -L\sum_{i\in\mathbb{I}}(e_{i-1}^k)^2 - \frac{L}{2}\sum_{i\in\mathbb{I}}(e_{i-2}^k)^2 - L\sum_{i\in\mathbb{I}}(e_{i-1}^k)^2 - L\sum_{i\in\mathbb{I}}(e_{i-2}^k)^2 \\ &\geq -\frac{7L}{2}\|E^k\|^2.\end{aligned}$$

Therefore, inequality (45) is obtained with $C = 7L + 2\|r\|_\infty + L_1$. ∎

Applying now Lemma 11 we obtain the following result:

THEOREM 15. *Let either assumption (A1),(A2),(A3) or (A4) be satisfied, as well as the CFL condition (13). Then there exists a constant $C \geq 0$ (independent of τ and h) and some $\tau_0 > 0$ such that, for any $\tau \leq \tau_0$*

$$\max_{2\leq k\leq N} \|E^k\|^2 \leq C\Big(\|E^0\|^2 + \|E^1\|^2 + \tau \sum_{2\leq k\leq N} \|\mathcal{E}^k\|^2\Big). \tag{62}$$

4.2. Two-dimensional case. Under suitable assumptions, the result of Theorem 15 can be extended to multi-dimensional equations. The nonlinearity can be treated exactly as in Section 3.2 and 3.6, so that we can focus on the linear case

$$v_t - \frac{1}{2}\operatorname{tr}[\Sigma(t,x,a)D_x^2 v] + b(t,x,a)D_x v + r(t,x,a)v + \ell(t,x,a) = 0,$$

with $a \in \Lambda$ given, and we will drop it in the following. For simplicity, consider the two-dimensional case, $d = 2$, with $r, \ell \equiv 0$ and omit the dependence on time of the coefficients. Given the covariance matrix Σ and the drift vector b by

$$\Sigma(x,y) = \begin{pmatrix} \sigma_1^2(x,y) & \rho\sigma_1\sigma_2(x,y) \\ \rho\sigma_1\sigma_2(x,y) & \sigma_2^2(x,y) \end{pmatrix} \quad \text{and} \quad b(x,y) = \begin{pmatrix} b_1(x,y) \\ b_2(x,y) \end{pmatrix},$$

with ρ the correlation parameter, the equation reads

$$v_t - \frac{1}{2}\sigma_1^2(x,y)v_{xx} - \rho\sigma_1\sigma_2(x,y)v_{xy} - \frac{1}{2}\sigma_2^2(x,y)v_{yy} + b_1(x,y)v_x + b_2(x,y)v_y = 0.$$

The computational domain is given by $\Omega := [x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}]$. We introduce the discretization in space defined by the steps $h_x, h_y > 0$ and we denote by $\mathcal{G}_{(h_x,h_y)}$ the associated mesh. In what follows, given any function ϕ of $(x,y) \in \Omega$, we will denote $\phi_{ij} = \phi(x_i, y_j)$ for $(i,j) \in \mathbb{I} := \mathbb{I}_1 \times \mathbb{I}_2$, where $\mathbb{I}_1 = \{1, \dots, I_1\}$, $\mathbb{I}_2 = \{1, \dots, I_2\}$. We consider a 7-point stencil for the second order derivatives (see [8, Section 5.1.4]):

$$v_{xx} \sim \frac{v_{i-1,j} - 2v_{ij} + v_{i+1,j}}{h_x^2} =: D_{xx}^2 v_{ij}, \qquad v_{yy} \sim \frac{v_{i,j-1} - 2v_{ij} + v_{i,j+1}}{h_y^2} =: D_{yy}^2 v_{ij}$$
$$v_{xy} \sim \frac{-v_{i,j-1} - v_{i,j+1} - v_{i-1,j} - v_{i+1,j} + v_{i-1,j-1} + v_{i+1,j+1} + 2v_{ij}}{2h_x h_y} =: D_{xy}^2 v_{ij}$$

and the BDF approximation of the first order derivatives

$$D_x^{1,-}u_{ij} := \frac{3u_{ij} - 4u_{i-1,j} + u_{i-2,j}}{2h_x} \quad \text{and} \quad D_x^{1,+}u_{ij} := -\left(\frac{3u_{ij} - 4u_{i+1,j} + u_{i+2,j}}{2h_x}\right),$$
$$D_y^{1,-}u_{ij} := \frac{3u_{ij} - 4u_{i,j-1} + u_{i,j-2}}{2h_y} \quad \text{and} \quad D_y^{1,+}u_{ij} := -\left(\frac{3u_{ij} - 4u_{i,j+1} + u_{i,j+2}}{2h_y}\right).$$

The scheme is therefore defined by

$$\frac{u_{ij}^k - 4u_{ij}^{k-1} + u_{ij}^{k-2}}{2\tau} - \frac{1}{2}\sigma_1^2(x_i,y_j)D_{xx}^2 u_{ij}^k - \rho\sigma_1\sigma_2(x_i,y_j)D_{xy}^2 u_{ij}^k - \frac{1}{2}\sigma_2^2(x_i,y_j)D_{yy}^2 u_{ij}^k$$
$$+ b_1^+(x_i,y_j)D_x^{1,-}u_{ij}^k - b_1^-(x_i,y_j)D_x^{1,+}u_{ij}^k + b_2^+(x_i,y_j)D_y^{1,-}u_{ij}^k - b_2^-(x_i,y_j)D_y^{1,+}u_{ij}^k = 0. \tag{63}$$

A straightforward calculation shows that the second order term also reads

$$\sigma_1^2(x_i,y_j)D_{xx}^2 u_{ij} + 2\rho\sigma_1\sigma_2(x_i,y_j)D_{xy}^2 u_{ij} + \sigma_2^2(x_i,y_j)D_{yy}^2 u_{ij}$$
$$= \alpha_{ij}D_{xx}^2 u_{ij} + \beta_{ij}D_{yy}^2 u_{ij}\gamma_{ij}\left(u_{i-1,j-1} - 2u_{ij} + u_{i+1,j+1}\right),$$

with α, β, γ defined by

$$\alpha_{ij} = \frac{\sigma_1(x_i,y_j)}{h_x}\left(\frac{\sigma_1(x_i,y_j)}{h_x} - \frac{\rho\sigma_2(x_i,y_j)}{h_y}\right),$$
$$\beta_{ij} = \frac{\sigma_2(x_i,y_j)}{h_y}\left(\frac{\sigma_2(x_i,y_j)}{h_y} - \frac{\rho\sigma_1(x_i,y_j)}{h_x}\right), \qquad \gamma_{ij} = \frac{\rho\sigma_1\sigma_2(x_i,y_j)}{h_y h_x}.$$

The scheme is completed with the following boundary conditions:

$$u_{i,j}^k = v(t_k, x_i, y_j), \quad \forall i \in \{-1, 0\} \cup \{I_1 + 1, I_1 + 2\},\ j \in \mathbb{I}_2,$$
$$u_{i,j}^k = v(t_k, x_i, y_j), \quad \forall j \in \{-1, 0\} \cup \{I_2 + 1, I_2 + 2\},\ i \in \mathbb{I}_1.$$

For simplicity, let us assume $h_x = h_y =: h$ (in this case, we write $\mathcal{G}_{(h_x,h_y)} = \mathcal{G}_h$). We consider the following assumptions:

(A1') $\|\sigma_i\|_\infty, \|b_i\|_\infty < \infty$ for $i = 1, 2$;

(A2') $\exists \eta > 0 : \sigma_i^2(x,y) - \rho\sigma_i(x,y)\sigma_j(x,y) \geq \eta \quad \forall (x,y) \in \mathbb{R}^2, i, j = 1, 2\ (i \neq j)$;

(A3') $\exists L \geq 0 : |\sigma_i(x,y) - \sigma_i(\bar{x},\bar{y})| \leq L(|x-\bar{x}|+|y-\bar{y}|) \quad \forall (x,y),(\bar{x},\bar{y}) \in \mathbb{R}^2, i = 1,2.$

Remark 16. If $h_x \neq h_y$ and for instance $h_y = Ch_x$ for some $C \geq 1$, (A2') has to hold with σ_2 replaced by σ_2/C as a result of the scaling properties of the scheme.

Remark 17. Observe that assumption (A2') is equivalent to requiring strong diagonal dominance of the covariance matrix.

PROPOSITION 18. *Let assumptions (A1'),(A2') and (A3') be satisfied. Then* (45) *is satisfied for* $|\cdot| = \|\cdot\|$.

Proof. Starting from (63) and using the notation introduced in Section 3, one can easily derive a vectorial recursion for the error $e^k_{ij} = u^k_{ij} - v(t_k, x_i, y_j)$, which, after a scalar multiplication by $E^k = (e^k_{ij})_{(i,j)\in\mathbb{I}}$, leads to

$$(I) = \tau\ (II) + 2\tau\ (III) + 2\tau\langle \mathcal{E}^k, E^k\rangle, \tag{64}$$

where

$$\begin{aligned}(I) =& \langle 3E^k - 4E^{k-1} + E^{k-2}, E^k\rangle,\\ (II) =& \sum_{(i,j)\in\mathbb{I}} \Big\{\alpha_{ij}\left(e^k_{i-1,j} - 2e^k_{ij} + e^k_{i+1,j}\right) + \beta_{ij}\left(e^k_{i,j-1} - 2e^k_{ij} + e^k_{i,j+1}\right)\\ &+ \gamma_{ij}\left(e^k_{i-1,j-1} - 2e^k_{ij} + e^k_{i+1,j+1}\right)\Big\} e^k_{ij}\end{aligned}$$

and, assuming without loss of generality that $b_1(x,y), b_2(x,y) \geq 0, \forall (x,y) \in \Omega$,

$$(III) = -\sum_{(i,j)\in\mathbb{I}} \left\{\frac{(b_1)_{ij}}{2h}\left(3e^k_{i,j} - 4e^k_{i-1,j} + e^k_{i-2,j}\right) + \frac{(b_2)_{ij}}{2h}\left(3e^k_{i,j} - 4e^k_{i,j-1} + e^k_{i,j-2}\right)\right\} e^k_{ij}.$$

The term (I) can be estimated exactly as in Proposition 12. Moreover, applying the same passages as in the proof of Proposition 12 in each direction $\xi \equiv (1,0),(0,1),(1,1)$ we can deduce that

$$\begin{aligned}(II) \leq& -\frac{\eta}{h^2}\sum_{(i,j)\in\mathbb{I}} (e^k_{i-1,j} - e^k_{ij})^2 + \frac{LM}{h}\sum_{(i,j)\in\mathbb{I}} |e^k_{i-1,j} - e^k_{ij}|\,|e^k_{ij}|\\ &-\frac{\eta}{h^2}\sum_{(i,j)\in\mathbb{I}} (e^k_{i,j-1} - e^k_{ij})^2 + \frac{LM}{h}\sum_{(i,j)\in\mathbb{I}} |e^k_{i,j-1} - e^k_{ij}|\,|e^k_{ij}|\\ &-\frac{\rho\eta}{h^2}\sum_{(i,j)\in\mathbb{I}} (e^k_{i-1,j-1} - e^k_{ij})^2 + \frac{\rho LM}{h}\sum_{(i,j)\in\mathbb{I}} |e^k_{i-1,j-1} - e^k_{ij}|\,|e^k_{ij}|,\end{aligned}$$

for a constant M that only depends on $\|\sigma_1\|_\infty$ and $\|\sigma_2\|_\infty$. Moreover, for (III), we can show that

$$(III) \leq \frac{C\|b\|_\infty}{h}\sum_{(i,j)\in\mathbb{I}} |e^k_{i-1,j} - e^k_{ij}|\,|e^k_{ij}| + |e^k_{i,j-1} - e^k_{ij}|\,|e^k_{ij}|.$$

Thanks to the non-degeneracy assumption in each direction, we can deduce (exactly as in Proposition 12) that for some constant $C \geq 0$

$$(II) + (III) \leq C\|E^k\|^2,$$

from which the result follows. ∎

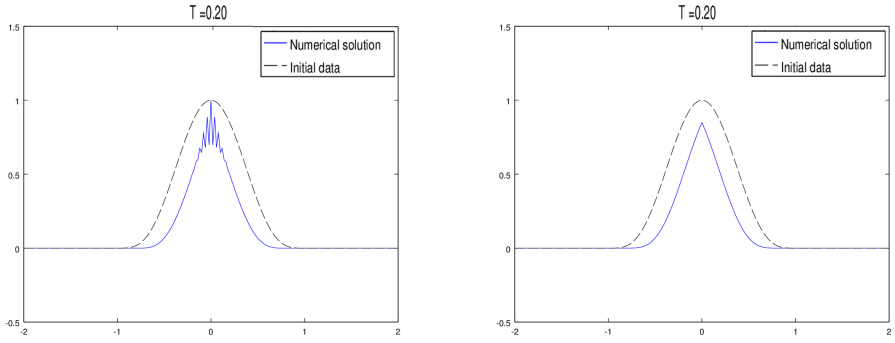

FIG. 1. *Test 1: Initial data (dashed line) and numerical solution at time $T = 0.2$ computed for $I + 1 = 200$ and $N = 20$ ($CFL = 0.5$) using BDF in time and centred approximation of the drift (left), BDF in time and space (right).*

Remark 19. When the strong diagonal dominance of the matrix Σ is not guaranteed, one can consider the generalized finite difference scheme in [4]. Given a set of directions $\mathcal{S}_{ij}^k \subset \mathbb{N} \times \mathbb{N}$ for each triplet (t_k, x_i, y_j), an approximation of the second order term is defined by

$$\frac{1}{2}\operatorname{tr}[\Sigma(t_k, x_i, y_j) D^2 v] \sim \sum_{\xi \in \mathcal{S}_{i,j}^k} \alpha_{i,j}^{\xi} D_{\xi}^2 v_{ij},$$

where

$$D_{\xi}^2 u_{ij} := u_{i+\xi_1, j+\xi_2} - 2u_{ij} + u_{i-\xi_1, j-\xi_2}$$

for given $\xi \equiv (\xi_1, \xi_2)$ and $\alpha_{i,j}^{\xi} \geq 0$ are chosen to make the scheme consistent. In the case of the seven-point stencil studied above, $\mathcal{S}_{ij}^k = \{(1,0), (0,1), (1,1)\}$ $\forall (i,j) \in \mathbb{I}, k \geq 1$ and $\alpha^{(1,0)} = \alpha$, $\alpha^{(0,1)} = \beta, \alpha^{(1,1)} = \gamma$. In order to apply directly the same argument as above, we would need all coefficients α^{ξ} to be Lipschitz continuous and non-degenerate (to compensate the drift term). This requirement does not seem easy to verify generically from the construction in [4].

5. Numerical tests. We propose two examples comparing the performances of our BDF scheme with other second order finite differences schemes.

5.1. Test 1: Eikonal equation. The first example is based on a deterministic problem ($\sigma \equiv 0$) and aims to motivate the choice of the BDF approximation for the drift term in (4), compared to the more classical centered approximation (8). Notice that our theoretical analysis does not cover this example, since in the degenerate case assumption (H4) is required which is not satisfied in the present example.

We consider

$$\begin{cases} v_t + |v_x| = 0, & x \in [-2, 2],\ t \in [0, T], \\ v(0, x) = \max(0, 1 - x^2)^4, & x \in [-2, 2]. \end{cases}$$

Remark 20. The Eikonal equation can be written as $v_t + \sup_{a \in \{-1,1\}}(a v_x) = 0$.

The initial data appears in Figure 1 (dashed line). The exact solution is given by

$$v(t, x) = \min(v_0(x - t), v_0(x + t)).$$

In Figure 1 we report the results obtained at the terminal time $T = 0.2$ using schemes (3)-(8) (left) and (3)-(4) (right) with $CFL = \tau/h = 0.5$. We numerically observe that the centered approximation generates undesired oscillations, whereas our BDF scheme is stable. Furthermore, Table 1 shows convergence of order 2 in both time and space. As stated in Theorem 3, in the case of degenerate diffusion, a CFL condition of the form $\tau \leq Ch$ has to be satisfied in order to have well posedness of the BDF scheme.

N	$I+1$	H^1-norm		L_2-norm		L^∞-norm		CPU (s)
		error	order	error	order	error	order	
5	10	5.35E-01	-	1.25E-01	-	1.36E-01	-	0.094
10	20	2.42E-01	1.14	4.51E-02	1.47	6.83E-02	0.99	0.096
20	40	8.25E-02	1.55	1.55E-02	1.55	2.01E-02	1.77	0.126
40	80	2.38E-02	1.80	4.32E-03	1.84	5.23E-03	1.94	0.147
80	160	6.26E-03	1.92	1.11E-03	1.96	1.31E-03	2.00	0.194
160	320	1.61E-03	1.96	2.79E-04	1.99	3.24E-04	2.01	0.335
320	640	4.09E-04	1.98	7.10E-05	1.99	8.19E-05	2.00	0.759
640	1280	1.03E-04	1.99	1.78E-05	2.00	2.05E-05	2.00	2.306

TABLE 1

Test 1. Error and convergence rate to the exact solution for the BDF scheme with $CFL = \tau/h = 0.1$.

Remark 21. Although the solution is globally only Lipschitz, we note that the equation holds at $x = 0$ for both right-sided and left-sided limits of the x-derivative. As left- and right-sided finite differences are used for $x < 0$ and $x > 0$, respectively, and either side at $x = 0$, the truncation error is of order $\tau^2 + h^2$ globally, and hence the L_2 convergence order is 2. For the H^1-norm, we note that the truncation error can be written as either $h^2 D_x^3 v + \tau^2 D_t^3 \tau$ or $h D_x^2 v + \tau D_t^2 \tau$, where the derivatives are evaluated at suitable intermediate points. Moreover, $D_x^2 v$, $D_t^2 v$ and $D_t^3 v$ are all symmetric around $x = 0$, as follows by differentiation of the PDE, and hence the truncation error is symmetric and Lipschitz up to order h and τ^2. Therefore, assuming that a left-sided difference is used at x_{-1} and a right-sided difference at x_0 (but the opposite case follows similarly), $h^{-1}(\mathcal{E}_0^k - \mathcal{E}_{-1}^k) = O(h)$ and

$$|\mathcal{E}^k|_1^2 = |h^{-1}(\mathcal{E}_0^k - \mathcal{E}_{-1}^k)|^2 h + \sum_{i \neq 0} |h^{-1}(\mathcal{E}_i^k - \mathcal{E}_{i-1}^k)|^2 h \leq C(\tau^2 + h^2)^2,$$

from which again order 2 can be deduced.

5.2. Test 2: A simple controlled diffusion model equation. The second test we propose is a simple problem with controlled diffusion. We consider

$$\begin{cases} v_t + \sup_{\sigma \in \{\sigma_1, \sigma_2\}} \left(-\frac{1}{2}\sigma^2 v_{xx} \right) = 0 & x \in [-1,1], t \in [0,T] \\ v(0,x) = \sin(\pi x) & x \in [-1,1], \end{cases}$$

with parameters $\sigma_1 = 0.1$, $\sigma_2 = 0.5$, $T = 0.5$.

In spite of the apparent simplicity of the equation under consideration, in [11] an example of non-convergence of the Crank-Nicolson scheme is given for a similar optimal control problem. Applied in [3] to such a problem, the BDF scheme, in contrast, has shown good performance.

Figure 2 (top) shows the initial data and the value function at terminal time computed using the BDF scheme. The error and convergence rate in different norms

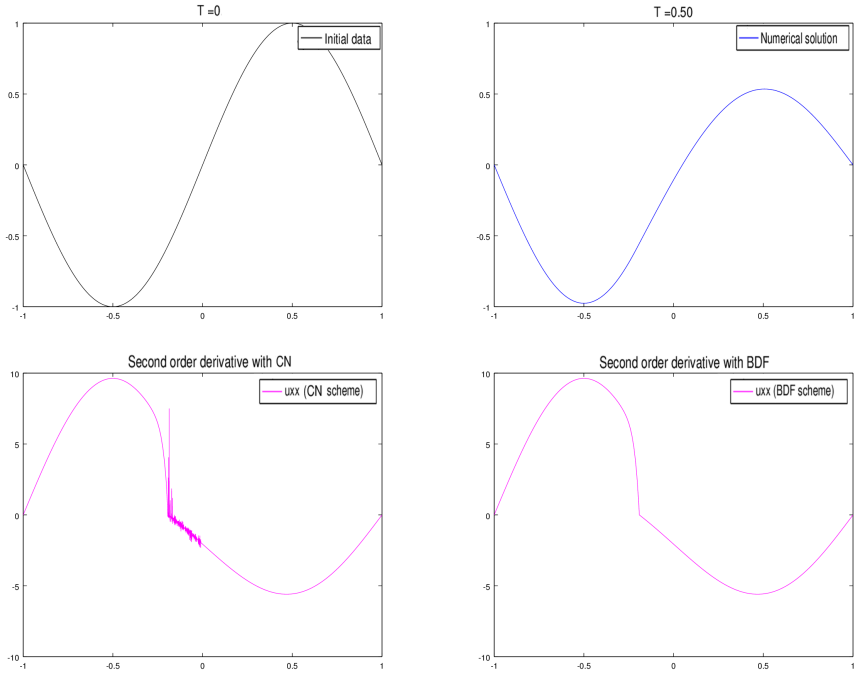

FIG. 2. *Test 2: Initial data (top, left), numerical solution at time $T = 0.5$ (top, right) computed using the BDF scheme, second order derivative computed with CN scheme (bottom, left) and BDF (bottom, right) computed for $N = 256$ and $I + 1 = 5120$.*

are reported in Table 2. Here, a numerical reference solution was computed by an implicit Euler scheme (in order to ensure convergence). Due to the low convergence order of the Euler scheme, the solution with the largest computationally feasible number of timesteps is not accurate enough to estimate precisely the order for the finest levels of the BDF scheme, resulting in a small bias for $N \in \{64, 128\}$.

N	$I+1$	H^1-norm		L_2-norm		L^∞-norm		CPU (s)
		error	order	error	order	error	order	
1	20	1.54E-01	-	5.11E-02	-	7.24E-02	-	0.030
2	40	5.53E-02	1.48	1.88E-02	1.45	2.63E-02	1.46	0.032
4	80	1.47E-02	1.91	5.17E-03	1.86	6.99E-03	1.91	0.041
8	160	3.58E-03	2.04	1.27E-03	2.03	1.66E-03	2.08	0.045
16	320	8.96E-04	2.00	3.13E-04	2.02	4.08E-04	2.02	0.063
32	640	2.24E-04	2.00	7.77E-05	2.01	1.01E-04	2.01	0.130
64	1280	5.44E-05	2.04	1.89E-05	2.04	2.46E-05	2.04	0.260
128	2560	1.21E-05	2.16	4.18E-06	2.18	5.39E-06	2.19	0.739

TABLE 2

Test 2. Error and convergence rate for the BDF scheme with high CFL $\tau = 5\Delta x$. A reference numerical solution computed by the implicit Euler scheme (7) with $I + 1 = 10 \times 2^{14}, N = (I + 1)/2$ is used for comparison.

Taking $\tau \sim \Delta x$, the BDF scheme converges with second order, see Table 2. This is not the case for CN as shown in Tables 3 and 4, which is especially visible in the

N	$I+1$	H^1-norm		L_2-norm		L^∞-norm		CPU (s)
		error	order	error	order	error	order	
1	20	4.11E-02	-	7.01E-03	-	9.44E-03	-	0.031
2	40	7.82E-03	2.39	1.45E-03	2.27	2.30E-03	2.04	0.037
4	80	1.98E-03	1.98	3.88E-04	1.91	5.64E-04	2.03	0.035
8	160	5.18E-04	1.93	1.03E-04	1.92	1.46E-04	1.94	0.043
16	320	1.11E-04	2.23	2.74E-05	1.91	3.88E-05	1.92	0.067
32	640	3.10E-05	1.83	7.88E-06	1.80	1.09E-05	1.83	0.133
64	1280	9.04E-06	1.78	2.74E-06	1.52	3.66E-06	1.58	0.300
128	2560	9.44E-05	-3.38	2.01E-05	-2.87	3.50E-05	-3.26	1.250

TABLE 3

Test 2. Error and convergence rate for the CN scheme with high CFL $\tau = 5\Delta x$. A reference numerical solution computed by the implicit Euler scheme (7) with $I+1 = 10 \times 2^{14}, N = (I+1)/2$ is used for comparison.

N	$I+1$	BDF2		CN	
		value	"order"	value	"order"
1	20	-0.57968	-	-0.57666	-
2	40	-0.57669	-	-0.57581	-
4	80	-0.57555	1.38	-0.57534	0.81
8	160	-0.57518	1.63	-0.57511	1.09
16	320	-0.57508	1.97	-0.57505	1.92
32	640	-0.57505	1.51	-0.57504	2.33
64	1280	-0.57504	1.91	-0.57504	2.00
128	2560	-0.57504	1.94	-0.57505	-2.44
256	5120	-0.57504	1.97	-0.57511	-1.88

TABLE 4

Test 2. Convergence table for the BDF and CN scheme comparing the value of the numerical solution at point $\bar{x} = 0.5$. The CN shows a loss of convergence.

		H^2-norm				$\|D_x^2 E\|_\infty$			
N	$I+1$	BDF2		CN		BDF2		CN	
		error	order	error	order	error	order	error	order
1	20	5.90E-01	-	3.96E-01	-	7.55E-01	-	1.01E+00	-
2	40	2.58E-01	1.19	1.50E-01	1.40	8.05E-01	-0.09	3.97E-01	1.34
4	80	4.49E-02	2.53	4.97E-02	1.59	7.11E-02	3.50	2.28E-01	0.80
8	160	1.12E-02	2.00	1.84E-02	1.43	3.24E-02	1.13	1.66E-01	0.46
16	320	3.15E-03	1.83	6.27E-03	1.56	1.22E-02	1.41	5.81E-02	1.51
32	640	8.43E-04	1.90	2.08E-03	1.59	3.25E-03	1.91	7.18E-02	-0.31
64	1280	2.13E-04	1.99	3.44E-03	-0.72	8.18E-04	1.99	4.83E-02	0.57
128	2560	5.00E-05	2.09	2.30E-02	-2.74	1.98E-04	2.05	3.05E-01	-2.66

TABLE 5

Test 2. Convergence table in the H^2-norm and in the L^∞-norm of the second order derivative for the BDF and CN scheme. The error is computed in the region $[-1,1] \setminus (-0.25,-0.15)$, i.e. avoiding the singular point $\bar{x} \sim -0.2$.

last lines where the ratio $\tau/\Delta x^2$ is high. One can verify that by taking a small CFL number, i.e. $\tau \sim \Delta x^2$ (where the scheme is monotone), the CN scheme shows second

order of convergence. Instabilities in the second order derivative of CN with a high CFL number are clearly seen in Figure 2 (bottom, left) (this is analogous to the effect found in [11]).

Remark 22. From Fig. 2 we conjecture that $D_x^2 v$ is Lipschitz in x, and from the PDE we deduce that v_t is also Lipschitz in t (the volatility is piecewise constant and jumps only where $D_x^2 v = 0$). The numerically observed second order convergence in both the L_2- and H^1-norms in spite of higher regularity can be rationalised as follows. The leading order term in the truncation error when expanding to second order in h, D_x^4, is a sum of a bounded function and a Dirac delta function, which prevents classical Taylor expansion to that order and does not result in a practical estimate as $\delta \notin L_2$. However, the recursion in Lemma 9 is still meaningful for a discrete approximation to δ because of the smoothing property of the strongly elliptic operator and of the BDF scheme; this Dirac delta gets smoothed over finite times and results in a bounded contribution to the L_2-norm. A similar argument can be made for the H^1-norm.

6. Conclusion. We have proved the well-posedness and stability in L_2- and H^1-norms of a second order BDF scheme for HJB equations with enough regularity of the coefficients. The significance of the results is that this was achieved for a second order (and hence) non-monotone scheme. For smooth or piecewise smooth solutions, as is often the case, one can use the recursion we derived to bound the error of the numerical solution in terms of the truncation error of the scheme. The latter depends on the regularity of the solution and has to be estimated for individual examples.

The numerical tests demonstrate convergence at least as good as predicted by the theoretical results, and often better, due to symmetries of the solution or smoothing properties of the equation and the scheme. This is in contrast to some alternative second order schemes, such as the central spatial difference in the case of a first order equation, or the Crank-Nicolson time stepping scheme for a second order equation, which showed poor or no convergence.

Appendix A. Proof of Lemma 7.

In order to prove the existence and uniqueness of a solution to (20), we consider a fixed-point approach. The initial problem (20) can be written as follows:

$$\max_{a\in\Lambda}(L_a X - (q_a - U_a X)) = 0, \tag{65}$$

where L_a and U_a are two matrices such that $M_a \equiv L_a + U_a$. We consider in particular L_a to be the lower triangular part of M_a including the diagonal terms, $(L_a)_{ij} := (M_a)_{ij} 1_{i\geq j}$, and U_a the remaining upper triangular part, $(U_a)_{ij} := (M_a)_{ij} 1_{i<j}$.

For a given vector $c \in \mathbb{R}^I$, let $g(c) := X$ denote the (unique) solution of the following simplified problem:

$$\max_{a\in\Lambda}(L_a X - (q_a - U_a c)) = 0. \tag{66}$$

Indeed, because $(L_a)_{ii} = (M_a)_{ii} > 0$, denoting $v_a := q_a - U_a c$, it is easy to see that the unique solution of

$$\max_{a\in\Lambda}(L_a X - v_a) = 0$$

is given by

$$x_i := \min_{a\in\Lambda}\left((v_a)_i - \sum_{k=1}^{i-1}(L_a)_{ik} x_k/(L_a)_{ii}\right)$$

(the proof can be obtained by recursion on i). Therefore, solving (65) amounts to solving $g(X) = X$.

Let us show that g is δ-Lipschitz for the $\|.\|_\infty$ norm, with $\delta := \max_a \|(L_a)^{-1}U_a\|_\infty$. For a diagonally dominant matrix, the following classical estimate holds:

$$\|(L_a)^{-1}U_a\|_\infty \leq \max_{i\in\mathbb{I}} \frac{\sum_{j>i}|(M_a)_{ij}|}{|(M_a)_{ii}| - \sum_{j<i}|(M_a)_{ij}|}$$

(this is related to the Gauss-Seidel relaxation method). By using the assumptions on the matrices M_a, we have $\delta < 1$. Hence, g is a contraction mapping on $\mathbb{R}^I$ and therefore we obtain the existence and uniqueness of a solution of (65) as desired. □

Appendix B. Von Neumann analysis for constant coefficients. In this section we consider b and σ constant and $b \geq 0$.
At each point (t_n, x_k) we have

$$\frac{3u_k^{n+1} - 4u_k^n + u_k^{n-1}}{2\tau} - \frac{\sigma^2}{2}\frac{u_{k-1}^{n+1} - 2u_k^{n+1} + u_{k+1}^{n+1}}{h^2} + b\frac{3u_k^{n+1} - 4u_{k-1}^{n+1} + u_{k-2}^{n+1}}{2h} = 0.$$

We perform a Fourier analysis of the scheme taking into account solutions of the form

$$u_k^n = R^n e^{ik\theta}.$$

The condition for stability will be $|R| \leq 1$. Substituting into the scheme:

$$3R^{n+1}e^{ik\theta} - 4R^n e^{ik\theta} + R^{n-1}e^{ik\theta} - \frac{\sigma^2}{2}\frac{2\tau}{h^2}\Big(R^{n+1}e^{i(k-1)\theta} - 2R^{n+1}e^{ik\theta} + R^{n+1}e^{i(k+1)\theta}\Big)$$
$$+ b\frac{2\tau}{2h}\Big(3R^{n+1}e^{ik\theta} - 4R^{n+1}e^{i(k-1)\theta} + R^{n+1}e^{i(k-2)\theta}\Big) = 0.$$

Simplifying and defining $\rho = 1/R$ we obtain the following second order equation:

$$\rho^2 - 4\rho + 3 - \frac{\tau\sigma^2}{h^2}\Big(e^{-i\theta} - 2 + e^{i\theta}\Big) + \frac{\tau b}{h}\Big(3 - 4e^{-i\theta} + e^{-2i\theta}\Big) = 0$$

under the stability condition $|\rho| \geq 1$. One has

$$\rho_\pm = 2 \pm \sqrt{1 + \frac{\tau\sigma^2}{h^2}\Big(e^{-i\theta} - 2 + e^{i\theta}\Big) - \frac{\tau b}{h}\Big(3 - 4e^{-i\theta} + e^{-2i\theta}\Big)}.$$

We define

$$z := \Big\{1 + \frac{2\tau\sigma^2}{h^2}(\cos\theta - 1) - \frac{2\tau b}{h}(1 - \cos\theta)^2\Big\} - i\Big\{\frac{2\tau b}{h}\sin\theta(2 - \cos\theta)\Big\}.$$

We consider the following cases:
- **CASE** $b \equiv 0$:

$$|\rho_\pm| \geq 1 \Leftrightarrow \Big|2 \pm \sqrt{1 + \frac{2\tau\sigma^2}{h^2}(\cos\theta - 1)}\Big| \geq 1.$$

This gives:

$$\begin{cases} 1 + \frac{2\tau\sigma^2}{h^2}(\cos\theta - 1) \geq 0 \quad (\rho \text{ real}) \\ 1 + \frac{2\tau\sigma^2}{h^2}(\cos\theta - 1) \leq 1 \end{cases} \quad \text{or} \quad \begin{cases} 1 + \frac{2\tau\sigma^2}{h^2}(\cos\theta - 1) < 0 \\ \Big|2 \pm i\sqrt{\frac{2\tau\sigma^2}{h^2}(1 - \cos\theta) - 1}\Big| \geq 1 \end{cases}$$

which implies that in this case the scheme is unconditionally stable.
- **CASE** $\sigma \equiv 0$:
In this case we have

$$\rho_\pm = 2 \pm \sqrt{\Big(1 - \frac{2\tau b}{h}(1-\cos\theta)^2\Big) - i\Big(\frac{2\tau b}{h}\sin\theta(2-\cos\theta)\Big)} = 2 \pm \sqrt{z}$$

with

$$\Re(z) = 1 - \frac{2\tau b}{h}(1-\cos\theta)^2 \quad \text{and} \quad \Im(z) = -\frac{2\tau b}{h}\sin\theta(2-\cos\theta)$$

If $\sin\theta(2-\cos\theta) = 0$ we have

$$\begin{cases} 1 - \frac{2\tau b}{h}(1-\cos\theta)^2 \geq 0 \quad (\rho \text{ real}) \\ 1 - \frac{2\tau b}{h}(1-\cos\theta)^2 \leq 1 \end{cases} \quad \text{or} \quad \begin{cases} 1 - \frac{2\tau b}{h}(1-\cos\theta)^2 < 0 \\ \Big|2 \pm i\sqrt{\frac{2\tau b}{h}(1-\cos\theta)^2 - 1}\Big| \geq 1 \end{cases}$$

which is satisfied for any $\tau, h > 0$.
Let us now assume that $\sin\theta(2-\cos\theta) \neq 0$. One has

$$\Re(\sqrt{z})^2 = \frac{1}{2}\Big(\Re(z) + \sqrt{\Re(z)^2 + \Im(z)^2}\Big) \quad \text{and} \quad \Im(\sqrt{z})^2 = \frac{1}{2}\Big(-\Re(z) + \sqrt{\Re(z)^2 + \Im(z)^2}\Big)$$

therefore

$$\Re(\rho_\pm) = 2 \pm \sqrt{\frac{\Re(z) + \sqrt{\Re(z)^2 + \Im(z)^2}}{2}} \quad \text{and} \quad \Im(\rho_\pm) = \pm\sqrt{\frac{-\Re(z) + \sqrt{\Re(z)^2 + \Im(z)^2}}{2}}$$

We ask $\Re(\rho)^2 + \Im(\rho)^2 \geq 1$:

$$\begin{aligned}
&\Re(\rho_\pm)^2 + \Im(\rho_\pm)^2 \geq 1 \\
&\Leftrightarrow (2 - |\Re(\sqrt{z})|)^2 + \Im(\sqrt{z})^2 \geq 1 \\
&\Leftrightarrow (2 - |\Re(\sqrt{z})|)^2 + \Im(\sqrt{z})^2 \geq 1 \\
&\Leftrightarrow 4 - 4|\Re(\sqrt{z})| + \Re(\sqrt{z})^2 + \Im(\sqrt{z})^2 \geq 1 \\
&\Leftrightarrow 4 - 4\sqrt{\frac{\Re(z) + \sqrt{\Re(z)^2 + \Im(z)^2}}{2}} + \sqrt{\Re(z)^2 + \Im(z)^2} \geq 1 \\
&\Leftrightarrow 4\sqrt{\frac{\Re(z) + \sqrt{\Re(z)^2 + \Im(z)^2}}{2}} \leq \sqrt{\Re(z)^2 + \Im(z)^2} + 3 \\
&\Leftrightarrow 8\Big(\Re(z) + \sqrt{\Re(z)^2 + \Im(z)^2}\Big) \leq (\Re(z)^2 + \Im(z)^2) + 9 + 6\sqrt{\Re(z)^2 + \Im(z)^2} \\
&\Leftrightarrow 0 \leq 9 - 8\,\Re(z) + (\Re(z)^2 + \Im(z)^2) - 2\sqrt{\Re(z)^2 + \Im(z)^2} \\
&\Leftrightarrow 0 \leq 8\big(1 - \Re(z)\big) + (\Re(z)^2 + \Im(z)^2) - 2\sqrt{\Re(z)^2 + \Im(z)^2} + 1 \\
&\Leftrightarrow 0 \leq 8(1 - \Re(z)) + \Big(1 - \sqrt{\Re(z)^2 + \Im(z)^2}\Big)^2
\end{aligned}$$

One has

$$1 - \Re(z) = 1 - \Big(1 - \frac{2\tau b}{h}(1-\cos\theta)^2\Big) \geq 0.$$

So the scheme is unconditionally. Analogously one can treat the case $b, \sigma \neq 0$.

REFERENCES

[1] G. Barles and P.E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4:271–283, 1991.

[2] T. Beale. Smoothing properties of implicit finite difference methods for a diffusion equation in maximum norm. *SIAM J. Numer. Anal.*, 47(4):2476–2495, 2009.

[3] O. Bokanowski, A. Picarelli, and C. Reisinger. High-order filtered schemes for time-dependent second order HJB equations. *ESAIM Math. Model. Numer. Anal.*, 2017. To appear.

[4] J.F. Bonnans and H. Zidani. Consistency of generalized finite difference schemes for the stochastic HJB equation. *SIAM J. Numer. Anal.*, 41(3):1008–1021, 2003.

[5] M.G. Crandall, H. Ishii, and P.L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, 27(1):1–67, 1992.

[6] E. Emmrich. Stability and error of the variable two-step BDF for semilinear parabolic problems. *J. Appl. Math. & Computing*, 19(1-2):33–55, 2005.

[7] S. K. Godunov. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik*, 89(3):271–306, 1959.

[8] W. Hackbusch. *Elliptic Differential Equations: Theory and Numerical Treatment*, volume 18 of *Springer series in computational mathematics*. Springer, 2010.

[9] A. Hill and E. Süli. Approximation of the global attractor for the incompressible Navier–Stokes equations. *IMA J. Numer. Anal.*, 20(4):633–667, 2000.

[10] N.V. Krylov. Boundedly nonhomogeneous elliptic and parabolic equations. *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, 46(3):487–523, 1982.

[11] D.M. Pooley, P.A. Forsyth, and K.R. Vetzal. Numerical convergence properties of option pricing pdes with uncertain volatility. *IMA J. Numer. Anal.*, 23(2):241–267, 2003.

[12] I. Smears and E. Süli. Discontinuous Galerkin finite element methods for time-dependent Hamilton–Jacobi–Bellman equations with Cordes coefficients. *Numer. Math.*, 133(1):141–176, 2016.

[13] E. Süli and D.F. Mayers. *An Introduction to Numerical Analysis*. Cambridge University Press, 2003.