Semi-Lagrangian schemes for linear and fully non-linear diffusion equations

Kristian Debrabant[a] and Espen R. Jakobsen[b]

[a]Technische Universität Darmstadt, Fachbereich Mathematik, Schloßgartenstr.7, D-64289 Darmstadt, Germany, debrabant@mathematik.tu-darmstadt.de

[b]Department of Mathematical Sciences, Norwegian University of Science and Technology,N-7491 Trondheim, Norway, erj@math.ntnu.no

TECHNISCHE UNIVERSITÄT DARMSTADT

Abstract

For linear and fully non-linear diffusion equations of Bellman-Isaacs type, we introduce a class of monotone approximation schemes relying on monotone interpolation. As opposed to classical numerical methods, these schemes converge for degenerate diffusion equations having general non-diagonal dominant coefficient matrices. Such schemes have to have a wide stencil in general. Besides providing a unifying framework for several known first order accurate schemes, our class of schemes also includes more efficient versions, and a new second order scheme that converges only for essentially monotone solutions. The methods are easy to implement and analyze, and they are more efficient than some other known schemes. We prove stability and convergence of the schemes in the general case, and provide error estimates in the convex case which are robust in the sense that they apply to degenerate equations and non-smooth solutions. The methods are extensively tested.

Keywords: Monotone approximation schemes, difference-interpolation methods, stability, convergence, error bound, degenerate parabolic equations, Hamilton-Jacobi-Bellman equations, viscosity solution.

2000 MSC: 65M12, 65M15, 65M06, 35K10, 35K55, 35K65, 49L25, 49L20.

1 Introduction

In this paper we introduce and analyze a class of monotone approximation schemes for fully non-linear diffusion equations of Bellman-Isaacs type,

$$u_t - \inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}}\Big\{L^{\alpha,\beta}[u](t,x) + c^{\alpha,\beta}(t,x)u + f^{\alpha,\beta}(t,x)\Big\} = 0 \qquad \text{in} \quad Q_T, \tag{1.1}$$

$$u(0,x) = g(x) \qquad \text{in} \quad \mathbb{R}^N, \tag{1.2}$$

where $Q_T := (0,T]\times\mathbb{R}^N$ and

$$L^{\alpha,\beta}[u](t,x) = \mathrm{tr}[a^{\alpha,\beta}(t,x)D^2u(t,x)] + b^{\alpha,\beta}(t,x)Du(t,x).$$

The coefficients $a^{\alpha,\beta} = \frac{1}{2}\sigma^{\alpha,\beta}\sigma^{\alpha,\beta\top}$, $b^{\alpha,\beta}$, $c^{\alpha,\beta}$, $f^{\alpha,\beta}$ and the initial data g take values respectively in $\mathbb{S}^N$, the space of $N\times N$ symmetric matrices, $\mathbb{R}^N$, $\mathbb{R}$, $\mathbb{R}$, and $\mathbb{R}$. We will only assume that $a^{\alpha,\beta}$ is positive semi-definite, thus the equation is allowed to degenerate and does not have any smooth solutions in general. Under suitable assumptions (see Section 2), the initial value problem (1.1)-(1.2) has a unique, bounded, Hölder continuous, viscosity solution u. This function is the upper or lower value of a stochastic differential game, or, if $\mathcal{A}$ or $\mathcal{B}$ is a singleton, the value function of a finite horizon, optimal stochastic control problem [25].

We introduce a family of schemes that we call Semi-Lagrangian (SL) schemes. They are a type difference-interpolation schemes and arise as time-discretizations of the following semi-discrete equation

$$u_t - \inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}}\Big\{L_k^{\alpha,\beta}[\mathcal{I}_h u](t,x) + c^{\alpha,\beta}(t,x)u + f^{\alpha,\beta}(t,x)\Big\} = 0 \quad \text{in} \quad X^h\times(0,T),$$

where $L_k^{\alpha,\beta}$ is a monotone difference approximation of $L^{\alpha,\beta}$ and $\mathcal{I}_h$ is a monotone interpolation operator on the spatial grid X^h. For more details see Section 3. Typically these scheme are first order wide-stencil schemes, and if the matrix $a^{\alpha,\beta}$ is bad enough, the stencil has to keep increasing as the grid is refined to have convergence. They include as special cases schemes from [16, 19, 7, 12], more efficient versions of these schemes, and a new second order compact stencil scheme. There are two main advantages of these schemes: (i) they are easy to understand and implement, and more importantly, (ii) they are consistent and monotone for every positive semi-definite diffusion matrix $a^{\alpha,\beta} = \frac{1}{2}\sigma^{\alpha,\beta}\sigma^{\alpha,\beta\top}$. The last point is important because monotone methods are known to converge to the correct solution [3], while non-monotone methods need not converge [21] or can even converge to false solutions [23].

Classical finite difference approximations (FDMs) of (1.1) are not monotone (of positive type) unless the matrix $a^{\alpha,\beta}$ satisfies additional assumptions like e. g. being diagonally dominant [18]. More general assumptions are given in e. g. [5, 14] but at the cost of increased stencil length. In fact, Dong and Krylov [14] proved that *no fixed stencil FDM* can approximate equations with a second derivative term involving a general positive semi-definite matrix function $a^{\alpha,\beta}$. Note that this type of result has been known for a long time, see e. g. [20, 12]. Some very simple examples of such "bad" matrices are given by

$$\begin{pmatrix} x_1^2 & \frac{1}{2}x_1x_2 \\ \frac{1}{2}x_1x_2 & x_2^2 \end{pmatrix} \text{ in } [0,1]^2, \quad \begin{pmatrix} \alpha^2 & \alpha\beta \\ \alpha\beta & \beta^2 \end{pmatrix} \text{ for } \mathcal{A}=\mathcal{B}=[0,1], \quad I - \frac{Du\otimes Du}{|Du|^2},$$

and these type of matrices appear in Finance, Stochastic Control Theory, and Mean Curvature Motion. The third example leads to quasi-linear equations and will not be considered here, we refer instead to [12].

To obtain convergent or monotone methods for problems involving non-diagonally dominant matrices, we know of two strategies: (i) The classical method of rotating the coordinate system locally to obtain diagonally dominant matrices $a^{\alpha,\beta}$, see e. g. Section 5.4 in [18], and (ii) the use of wide stencil methods. The two solutions seem to be somewhat related, but at least the defining ideas and implementation are different. Both ways lead to methods that have reduced order compared to standard schemes for diagonally dominant problems. But it seems to us that it is much more difficult to implement the first strategy.

In addition to the wide-stencil methods mentioned above, we also mention the method of Bonnans-Zidani [5] which is not an SL type scheme. Schemes for other types of equations related to our SL schemes have been studied by Crandall-Lions [12] for the Mean Curvature Motion equation, by Oberman [22] for Monge-Ampère equations, and by Camilli and the second author for non-local Bellman equations [8]. The terminology SL schemes is already used for schemes for transport equations, conservation laws, and first order Hamilton-Jacobi equations. In the Hamilton-Jacobi setting, these schemes go back to the 1983 paper [9] of Capuzzo-Dolcetta.

The rest of this paper is organized as follows. In the next Section we explain the notation and state a well-posedness and regularity result for (1.1)–(1.2). The SL schemes are motivated and defined in an abstract setting in Section 3, and in Section 4 we prove that they are consistent, monotone, L^∞-stable, and convergent. We provide several examples of SL schemes in Section 5, including the linear interpolation SL (LISL) scheme. This is the basic example of this paper, and it is a first order scheme that can be defined on unstructured grids.

Our SL schemes make use of monotone interpolation, and higher order interpolation is not monotone in general. But for essentially monotone solutions, we can use monotone cubic Hermite interpolation (see Fritsch and Carlson [17] and Eisenstat, Jackson and Lewis [15]) to obtain new second order schemes called monotone cubic interpolation SL (MCSL) schemes. These new schemes are defined in Section 6, and in contrast to the LISL schemes they are compact stencil schemes. Note well that in the special case of first order HJ-equations with monotone solutions, these schemes are consistent, monotone, second order schemes! To our knowledge, this is the first example of a monotone scheme which is more than first order accurate in the HJ-setting.

We discuss various issues concerning the SL schemes in Section 7. We compare the LISL scheme to the scheme of Bonnans-Zidani [5] and find that the LISL scheme is much easier to understand and implement, and in general it is much faster. However, on bounded domains the LISL scheme will in some cases over-step the boundaries and some ad hoc solution has to be found. This problem is avoided by the Bonnans-Zidani scheme at the cost of being less accurate near the boundary than in the interior. Finally, we explain that the SL schemes can be interpreted as collocation methods for derivative free equations, or as dynamic programming equations of discrete stochastic differential games or optimal control problems.

In Sections 8, 9 and Appendix B, we produce robust error estimates for convex equations (i. e. $\mathcal{B}$ is a singleton in (1.1)). These estimates are obtained through the regularization method of Krylov and apply to degenerate equations, non-smooth solutions, and both the LISL and MCSL schemes. Finally, in Section 10, our methods are extensively tested. In particular we find that the LISL and MCSL schemes yield much faster methods to solve the finance problem of Bruder, Bokanowski, Maroso, and Zidani [6].

2 Notation and well-posedness

In this section we introduce notation and assumptions, and give a well-posedness and comparison result for the initial value problem (1.1) – (1.2).

We denote by $\leq$ the component by component ordering in $\mathbb{R}^M$ and the ordering in the sense of positive semi-definite matrices in $\mathbb{S}^N$. The symbols $\wedge$ and $\vee$ denote the minimum respectively the maximum. By $|\cdot|$ we mean the Euclidean vector norm in any $\mathbb{R}^p$ type space (including the spaces of matrices and tensors). Hence if $X \in \mathbb{R}^{N,P}$, then $|X|^2 = \sum_{i,j} |X_{ij}|^2 = \mathrm{tr}(XX^\top)$ where $X^\top$ is the transpose of X.

If w is a bounded function from some set $Q' \subset \overline{Q}_\infty$ into either $\mathbb{R}$, $\mathbb{R}^M$, or the space of $N \times P$ matrices, we set

$$|w|_0 = \sup_{(t,y)\in Q'} |w(t,y)|.$$

Furthermore, for $\delta \in (0,1]$, we set

$$[w]_\delta = \sup_{(t,x)\neq(s,y)} \frac{|w(t,x) - w(s,y)|}{(|x-y| + |t-s|^{1/2})^\delta} \quad \text{and} \quad |w|_\delta = |w|_0 + [w]_\delta.$$

Let $C_b(Q')$ and $\mathcal{C}^{0,\delta}(Q')$, $\delta \in (0,1]$, denote respectively the space of bounded continuous functions on Q' and the subset of $C_b(Q')$ in which the norm $|\cdot|_\delta$ is finite. Note in particular the choices $Q' = Q_T$ and $Q' = \mathbb{R}^N$. In the following we always suppress the domain Q' when writing norms.

For simplicity, we will use the following assumptions on the data of (1.1)–(1.2):

(A1) For any $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$, $a^{\alpha,\beta} = \frac{1}{2}\sigma^{\alpha,\beta}\sigma^{\alpha,\beta\,\top}$ for some $N \times P$ matrix $\sigma^{\alpha,\beta}$. Moreover, there is a constant K independent of α, β such that

$$|g|_1 + |\sigma^{\alpha,\beta}|_1 + |b^{\alpha,\beta}|_1 + |c^{\alpha,\beta}|_1 + |f^{\alpha,\beta}|_1 \leq K.$$

These assumptions are standard and ensure comparison and well-posedness of (1.1)–(1.2) in the class of bounded x-Lipschitz functions.

Proposition 2.1. *Assume that* (A1) *holds. Then there exists a unique solution u of* (1.1)–(1.2) *and a constant C only depending on T and K from* (A1) *such that*

$$|u|_1 \leq C.$$

Furthermore, if u_1 and u_2 are sub- and supersolutions of (1.1) *satisfying $u_1(0,\cdot) \leq u_2(0,\cdot)$, then $u_1 \leq u_2$.*

The proof is standard. Assumption (A1) can be relaxed in many ways, e. g. using weighted norms, Hölder or uniform continuity, etc. But in doing so, solutions can become unbounded and less smooth, and the analysis becomes harder and less transparent. Therefore we will not consider such extensions in this paper.

By solutions in this paper we always mean viscosity solutions, see e. g. [11, 25].

3 Definition of SL schemes

In this section we propose a class of monotone approximation schemes for (1.1)–(1.2) which we call Semi-Lagrangian or SL schemes. This class includes (parabolic versions of) the "control schemes" of Menaldi [19] and Camilli and Falcone [7] and the monotone schemes of Crandall and Lions [12]. It also includes SL schemes for first order Bellman equations [9, 16], and it allows for more effective versions of these schemes as discussed in Section 5. For a motivation for the name, we refer to Remark 7.2. For the time discretization we propose a generalized mid-point rule that includes explicit, implicit, and a second order Crank-Nicolson like approximation. Note that the equation is non-smooth in general, so the usual way of defining a Crank-Nicolson scheme would only give a first order scheme in time.

The schemes will be defined on a possibly unstructured family of grids $\{G_{\Delta t,\Delta x}\}$ with

$$G = G_{\Delta t,\Delta x} = \{(t_n, x_i)\}_{n\in\mathbb{N}_0, i\in\mathbb{N}} = \{t_n\}_{n\in\mathbb{N}_0} \times X_{\Delta x},$$

for $\Delta t, \Delta x > 0$. Here $0 = t_0 < t_1 < \cdots < t_n < t_{n+1}$ satisfy

$$\max_n \Delta t_n \leq \Delta t \qquad \text{where} \qquad \Delta t_n = t_n - t_{n-1},$$

and $X_{\Delta x} = \{x_i\}_{i\in\mathbb{N}}$ is the set of vertices or nodes for a non-degenerate polyhedral subdivision $\mathcal{T}^{\Delta x} = \{T_j^{\Delta x}\}_{j\in\mathbb{N}}$ of $\mathbb{R}^N$. For some $\rho \in (0,1)$ the polyhedrons $T_j = T_j^{\Delta x}$ satisfy

$$\operatorname{int}(T_j \cap T_i) \underset{i\neq j}{=} \emptyset, \quad \bigcup_{j\in\mathbb{N}} T_j = \mathbb{R}^N, \quad \rho\Delta x \leq \sup_{j\in\mathbb{N}}\{\operatorname{diam} B_{T_j}\} \leq \sup_{j\in\mathbb{N}}\{\operatorname{diam} T_j\} \leq \Delta x,$$

where diam is the diameter of the set and B_{T_j} is the greatest ball contained in T_j.

To motivate the numerical schemes, we write $\sigma = (\sigma_1, \sigma_2, ..., \sigma_m, ..., \sigma_P)$ where σ_m is the m-th column of σ and observe that for $k > 0$ and smooth functions ϕ,

$$\begin{aligned}
\frac{1}{2}\operatorname{tr}[\sigma\sigma^\top D^2\phi(x)] &= \frac{1}{2}\sum_{m=1}^{P}\operatorname{tr}[\sigma_m\sigma_m^\top D^2\phi(x)] \\
&= \sum_{m=1}^{P}\frac{1}{2}\frac{\phi(x+k\sigma_m) - 2\phi(x) + \phi(x-k\sigma_m)}{k^2} + \mathcal{O}(k^2), \\
bD\phi(x) &= \frac{\phi(x+k^2 b) - \phi(x)}{k^2} + \mathcal{O}(k^2) \\
&= \frac{1}{2}\frac{\phi(x+k^2 b) - 2\phi(x) + \phi(x+k^2 b)}{k^2} + \mathcal{O}(k^2).
\end{aligned}$$

These approximations are monotone (of positive type) and the errors are bounded by $\frac{1}{48}P|\sigma|_0^4|D^4\phi|_0 k^2$ and $\frac{1}{2}|b|_0^2|D^2\phi|_0 k^2$ respectively. To relate these approximations to a grid G, we replace ϕ by its interpolant $\mathcal{I}\phi$ on that grid and obtain

$$\begin{aligned}
\frac{1}{2}\operatorname{tr}[\sigma\sigma^\top D^2\phi(x)] &\approx \sum_{m=1}^{P}\frac{1}{2}\frac{(\mathcal{I}\phi)(x+k\sigma_m) - 2(\mathcal{I}\phi)(x) + (\mathcal{I}\phi)(x-k\sigma_m)}{k^2}, \\
bD\phi(x) &\approx \frac{1}{2}\frac{(\mathcal{I}\phi)(x+k^2 b) - 2(\mathcal{I}\phi)(x) + (\mathcal{I}\phi)(x+k^2 b)}{k^2}.
\end{aligned}$$

If the interpolation is monotone (positive) then the full discretization is still monotone and represents a typical example of the discretizations we consider below.

To construct the general scheme, we generalize the above construction. We now consider general finite difference approximations of the differential operator $L^{\alpha,\beta}[\phi]$ in (1.1) defined as

$$L_k^{\alpha,\beta}[\phi](t,x) := \sum_{i=1}^{M} \frac{\phi(t,x+y_{k,i}^{\alpha,\beta,+}(t,x)) - 2\phi(t,x) + \phi(t,x+y_{k,i}^{\alpha,\beta,-}(t,x))}{2k^2}, \tag{3.1}$$

for $k > 0$ and some $M \geq 1$, where for all smooth functions ϕ,

$$|L_k^{\alpha,\beta}[\phi] - L^{\alpha,\beta}[\phi]| \leq C(|D\phi|_0 + \cdots + |D^4\phi|_0)k^2. \tag{3.2}$$

This approximation and interpolation yield a semi-discrete approximation of (1.1),

$$U_t - \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \Big\{ L_k^{\alpha,\beta}[\mathcal{I}U](t,x) + c^{\alpha,\beta}(t,x)U + f^{\alpha,\beta}(t,x) \Big\} = 0 \quad \text{in} \quad (0,T)\times X_{\Delta x},$$

and the final scheme can then be found after discretizing in time using a parameter $\theta \in [0,1]$,

$$\delta_{\Delta t_n} U_i^n = \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \Big\{ L_k^{\alpha,\beta}[\mathcal{I}\bar{U}_{\cdot}^{\theta,n}]_i^{n-1+\theta} + c_i^{\alpha,\beta,n-1+\theta}\bar{U}_i^{\theta,n} + f_i^{\alpha,\beta,n-1+\theta} \Big\} \tag{3.3}$$

in G, where $U_i^n = U(t_n,x_i)$, $f_i^{\alpha,\beta,n-1+\theta} = f^{\alpha,\beta}(t_{n-1}+\theta\Delta t_n, x_i)$, $\dots$ for $(t_n,x_i)\in G$,

$$\delta_{\Delta t}\phi(t,x) = \frac{\phi(t,x)-\phi(t-\Delta t,x)}{\Delta t}, \qquad \text{and} \qquad \bar{\phi}_{\cdot}^{\theta,n} = (1-\theta)\phi_{\cdot}^{n-1} + \theta\phi_{\cdot}^{n}.$$

As initial conditions we take

$$U_i^0 = g(x_i) \quad \text{in} \quad X_{\Delta x}. \tag{3.4}$$

Remark 3.1. For the choices $\theta = 0, 1$, and $1/2$ the time discretization corresponds to respectively explicit Euler, implicit Euler and midpoint rule. For $\theta = 1/2$, the full scheme can be seen as generalized Crank-Nicolson type discretizations.

4 Analysis of SL schemes

In this section we prove that the SL scheme (3.3) is consistent and monotone, and we present L^∞-stability, existence, uniqueness, and convergence results for these schemes. In Section 8 we also give error estimates when $\mathcal{B}$ is a singleton and equation (1.1) is convex.

For the approximation $L_k^{\alpha,\beta}$ and interpolation $\mathcal{I}$ we will assume that

$$\begin{cases} \sum_{i=1}^{M}[y_{k,i}^{\alpha,\beta,+} + y_{k,i}^{\alpha,\beta,-}] = 2k^2 b^{\alpha,\beta} + \mathcal{O}(k^4), \\ \sum_{i=1}^{M}[y_{k,i}^{\alpha,\beta,+}\otimes y_{k,i}^{\alpha,\beta,+} + y_{k,i}^{\alpha,\beta,-}\otimes y_{k,i}^{\alpha,\beta,-}] = k^2\sigma^{\alpha,\beta}\sigma^{\alpha,\beta\top} + \mathcal{O}(k^4), \\ \sum_{i=1}^{M}[\otimes_{j=1}^3 y_{k,i}^{\alpha,\beta,+} + \otimes_{j=1}^3 y_{k,i}^{\alpha,\beta,-}] = \mathcal{O}(k^4), \\ \sum_{i=1}^{M}[\otimes_{j=1}^4 y_{k,i}^{\alpha,\beta,+} + \otimes_{j=1}^4 y_{k,i}^{\alpha,\beta,-}] = \mathcal{O}(k^4). \end{cases} \tag{Y1}$$

There are $K \geq 0, r\in\mathbb{N}$ such that $|(\mathcal{I}\phi)-\phi|_0 \leq K|D^p\phi|_0\Delta x^p$ for all $\mathbb{N}\ni p\leq r$ and all smooth functions ϕ. (I1)

There is a non-negative basis of functions $\{w_j(x)\}_j$ such that $(\mathcal{I}\phi)(x) = \sum_j \phi(x_j)w_j(x)$, $w_i(x_j) = \delta_{ij}$, and $w_j(x)\geq 0$ for all $i,j\in\mathbb{N}$. (I2)

Under assumption (Y1), a Taylor expansion shows that $L_k^{\alpha,\beta}$ is a second order consistent approximation satisfying (3.2). If we assume also (I1) , it then follows that $L_k^{\alpha,\beta}[\mathcal{I}\phi]$ is a consistent approximation of $L^{\alpha,\beta}[\phi]$ if $\frac{\Delta x^r}{k^2} \to 0$. Indeed

$$|L_k^{\alpha,\beta}[\mathcal{I}\phi] - L^{\alpha,\beta}[\phi]| \leq |L_k^{\alpha,\beta}[\mathcal{I}\phi] - L_k^{\alpha,\beta}[\phi]| + |L_k^{\alpha,\beta}[\phi] - L^{\alpha,\beta}[\phi]|,$$

where $|L_k^{\alpha,\beta}[\phi] - L^{\alpha,\beta}[\phi]|$ is estimated in (3.2), and by (I1),

$$|L_k^{\alpha,\beta}[\mathcal{I}\phi] - L_k^{\alpha,\beta}[\phi]| \leq C|D^r\phi|_0 \frac{\Delta x^r}{k^2}.$$

Remark 4.1. Assumption (Y1) is similar to the local consistency conditions used in [18]. The $O(k^4)$ terms insure that the method is second order accurate as $k \to 0$. Convergence will still be achieved if we relax $O(k^4)$ to $o(k^2)$ as $k \to 0$.

Remark 4.2. An interpolation satisfying (I2) is said to be *positive* or *monotone* and preserves monotonicity of the data. Note that such an interpolation $\mathcal{I}\phi$ does not use (exact) derivatives to reconstruct the function ϕ. Typically $\mathcal{I}$ will be constant, linear, or multi-linear interpolation (i. e. $r \leq 2$ in (I1)) since higher order interpolation is not monotone in general. For later use we note that from (I1) and (I2) it follows that

$$r \geq 1 \Rightarrow \sum_i w_i(x) \equiv 1 \qquad \text{and} \qquad r \geq 2 \Rightarrow \sum_i x_i w_i(x) \equiv x. \tag{4.1}$$

Now we prove that the scheme (3.3)–(3.4) is consistent and monotone. The scheme is said to be *monotone* if it can be written as

$$\sup_\alpha \inf_\beta \left\{ B_{j,j}^{\alpha,\beta,n,n} U_j^n - \sum_{i \neq j} B_{j,i}^{\alpha,\beta,n,n} U_i^n - \sum_i B_{j,i}^{\alpha,\beta,n,n-1} U_i^{n-1} - F_j^{\alpha,\beta,n} \right\} = 0 \tag{4.2}$$

in G, where $B_{i,j}^{\alpha,\beta,n,m} \geq 0$ and $F_j^{\alpha,\beta,n}$ does not depend on U.

Lemma 4.1. *Assume* (I1)*,* (I2)*, and* (Y1) *hold.*

(a) The consistency error of the scheme (3.3) *is bounded by*

$$\frac{|1-2\theta|}{2}|\phi_{tt}|_0 \Delta t + C\Bigg(\Delta t^2 \left(|\phi_{tt}|_0 + |\phi_{ttt}|_0 + |D\phi_{tt}|_0 + |D^2\phi_{tt}|_0 \right) + |D^r\phi|_0 \frac{\Delta x^r}{k^2} + (|D\phi|_0 + \cdots + |D^4\phi|_0)k^2 \Bigg).$$

(b) The scheme (3.3) *is monotone if the following CFL condition holds,*

$$(1-\theta)\Delta t \left[\frac{M}{k^2} - c_i^{\alpha,\beta,n-1+\theta} \right] \leq 1 \;\; \text{and} \;\; \theta \Delta t \, c_i^{\alpha,\beta,n-1+\theta} \leq 1 \text{ for all } \alpha, \beta, n, i. \tag{4.3}$$

Remark 4.3. By parabolic regularity $D^2 \sim \partial_t$ which means that e. g. $|D^2\phi_{tt}|_0$ is proportional to $|\phi_{ttt}|_0$. When $\theta = 1/2$ ("Crank-Nicolson"), the scheme (3.3) is second order accurate in time.

Proof. It is immediate that the scheme (3.3) is consistent with (1.1) with a truncation error bounded by

$$\frac{|1-2\theta|}{2}|\phi_{tt}|_0\Delta t + \frac{1}{3}|\phi_{ttt}|_0 \Delta t^2 + \sup_{\alpha,\beta,n} \left\{ \left| L^{\alpha,\beta}[\bar{\phi}^{\theta,n}] - L_k^{\alpha,\beta}[\mathcal{I}\bar{\phi}^{\theta,n}] \right|_0 \right\}$$
$$+ \sup_{\alpha,\beta,n} \left\{ \left| L^{\alpha,\beta}[\phi^{n-1+\theta} - \bar{\phi}^{\theta,n}] \right|_0 + \left| c^{\alpha,\beta,n-1+\theta}(\phi^{n-1+\theta} - \bar{\phi}^{\theta,n}) \right|_0 \right\}$$

for smooth ϕ. By (I1) and (3.2), $|L^{\alpha,\beta}[\bar{\phi}^{\theta,n}] - L_k^{\alpha,\beta}[\mathcal{I}\bar{\phi}^{\theta,n}]|$ can be bounded by

$$C|D^r\phi|_0 \frac{\Delta x^r}{k^2} + C(|D\phi|_0 + \cdots + |D^4\phi|_0)k^2,$$

while $|L^{\alpha,\beta}[\phi^{n-1+\theta} - \bar{\phi}^{\theta,n}]|$ is bounded by

$$\Delta t^2\theta(1-\theta)\sup_{\alpha,\beta}|L^{\alpha,\beta}[\phi_{tt}]|_0 \leq C\Delta t^2\theta(1-\theta)\Big\{|D\phi_{tt}|_0 + |D^2\phi_{tt}|_0\Big\}.$$

Finally, $|c^{\alpha,\beta,n-1+\theta}(\phi^{n-1+\theta} - \bar{\phi}^{\theta,n})\big| \leq C\theta(1-\theta)\Delta t^2|\phi_{tt}|_0$. Hence part (a) follows.

To prove part (b), we note that since $\sum_i w_i \equiv 1$ we have

$$L_k^{\alpha,\beta}[\mathcal{I}\phi(t,\cdot)](t_{n-1+\theta}, x_j) = \sum_{i\in\mathbb{N}} l_{j,i}^{\alpha,\beta,n-1+\theta}\left[\phi(t,x_i) - \phi(t,x_j)\right],$$

where

$$l_{j,i}^{\alpha,\beta,n-1+\theta} = \sum_{l=1}^{M} \frac{w_i(x_j + y_{k,l}^{\alpha,\beta,+}(t_{n-1+\theta},x_j)) + w_i(x_j + y_{k,l}^{\alpha,\beta,-}(t_{n-1+\theta},x_j))}{2k^2}.$$

This quantity is non-negative by (I2), and $\sum_i l_{j,i}^{\alpha,\beta,n-1+\theta} = \frac{M}{k^2}$ by (I1), (I2), and $\sum_i w_i \equiv 1$. Therefore the only non-zero coefficients in (4.2) at (t_n, x_j) are

$$B_{j,j}^{\alpha,\beta,n,n} = 1 + \theta\Delta t_n\big(\tfrac{M}{k^2} - l_{jj}^{\alpha,\beta,n-1+\theta} - c_j^{\alpha,\beta,n-1+\theta}\big),$$

$$B_{j,j}^{\alpha,\beta,n,n-1} = 1 - (1-\theta)\Delta t_n\big(\tfrac{M}{k^2} - l_{jj}^{\alpha,\beta,n-1+\theta} - c_j^{\alpha,\beta,n-1+\theta}\big),$$

$$B_{j,i}^{\alpha,\beta,n,n} = \theta\Delta t_n l_{j,i}^{\alpha,\beta,n-1+\theta}, \qquad B_{j,i}^{\alpha,\beta,n,n-1} = (1-\theta)\Delta t_n l_{j,i}^{\alpha,\beta,n-1+\theta},$$

where $j \neq i$. These coefficients are positive if (4.3) holds. □

Existence, uniqueness, stability, and convergence results are given below. In the following, we denote by $c^{\alpha,\beta,+}$ the positive part of $c^{\alpha,\beta}$.

Theorem 4.2. *Assume* (A1), (I1), (I2), (Y1), *and* (4.3).

(a) There exists a unique bounded solution U_h *of* (3.3)–(3.4).

(b) The solution U_h *of* (3.3)–(3.4) *is* L^∞*-stable when* $2\theta\Delta t \sup_{\alpha,\beta}|c^{\alpha,\beta,+}|_0 \leq 1$*:*

$$|U^n|_0 \leq e^{2\sup_{\alpha,\beta}|c^{\alpha,\beta,+}|_0 t_n}\Big[|g|_0 + t_n \sup_{\alpha,\beta}|f^{\alpha,\beta}|_0\Big].$$

(c) U_h *converges uniformly to the solution* u *of* (1.1)–(1.2) *as* $\Delta t, k, \frac{\Delta x^r}{k^2} \to 0$.

Proof. Existence and *uniqueness* of bounded solutions follow by induction. Let $t = t_n$ and assume U^{n-1} is a known bounded function. For $\varepsilon > 0$ we define the operator T by

$$TU_j^n = U_j^n - \varepsilon\cdot(\text{left hand side of (4.2)}) \qquad \text{for all} \qquad j\in\mathbb{Z}^M.$$

Note that the fixed point equation $U^n = TU^n$ is equivalent to equation (3.3). By the definition and sign of the B-coefficients we see that

$$\begin{aligned}
&TU_j^n - T\tilde{U}_j^n\\
&\leq \sup_{\alpha,\beta}\Big\{\Big[1 - \varepsilon(1 + \Delta t_n\theta(\tfrac{M}{k^2} - c_j^{\alpha,\beta,n-1+\theta}))\Big](U_j^n - \tilde{U}_j^n) + \varepsilon\Delta t_n\theta\tfrac{M}{k^2}|U_\cdot^n - \tilde{U}_\cdot^n|_0\Big\}\\
&\leq (1 - \varepsilon[1 - \Delta t_n\theta\sup_{\alpha,\beta}|c^{\alpha,\beta,+}|_0])|U_\cdot^n - \tilde{U}_\cdot^n|_0
\end{aligned}$$

if ε is so small that $1-\varepsilon(1+\Delta t\theta(\frac{M}{k^2}-c_j^{\alpha,\beta,n-1+\theta}))\geq 0$ and $\varepsilon(1-\Delta t_n\theta\sup_{\alpha,\beta}|c^{\alpha,\beta,+}|_0)<1$ for all j,n,α,β. Taking the supremum over all j and interchanging the role of U and $\tilde{U}$ proves that T is a contraction on the Banach space of bounded functions on $X_{\Delta x}$ under the sup-norm. Existence and uniqueness then follows from the fixed point theorem (for U^n) and for all of U by induction since $U^0=g$ is bounded.

A similar argument using (4.2) proves L^∞-*stability* for $\theta\Delta t\sup_{\alpha,\beta}|c^{\alpha,\beta,+}|_0\leq\frac{1}{2}$:

$$
\begin{aligned}
|U^n|_0 &\leq \Big(\frac{1+(1-\theta)\Delta t\sup_{\alpha,\beta}|c^{\alpha,\beta,+}|_0}{1-\theta\Delta t\sup_{\alpha,\beta}|c^{\alpha,\beta,+}|_0}\Big)\Big[|U^{n-1}|_0+\Delta t_n\sup_{\alpha,\beta}|f^{\alpha,\beta}|_0\Big]\\
&\leq e^{2\sup_{\alpha,\beta}|c^{\alpha,\beta,+}|_0 t_n}\Big[|g|_0+t_n\sup_{\alpha,\beta}|f^{\alpha,\beta}|_0\Big].
\end{aligned}
$$

In view of this estimate, *convergence* of U_h to the solution u of (1.1)–(1.2) follows from the Barles-Souganidis result in [3]. □

5 Examples of SL schemes

5.1 Examples of approximations $L_k^{\alpha,\beta}$

We present several examples of approximations of the term $L^{\alpha,\beta}[\phi]$ of the form $L_k^{\alpha,\beta}[\phi]$, including previous approximations that have appeared in [16, 19, 7, 12] plus more computational efficient variants.

1. The approximation of Falcone [16] (see also [9]),

$$b^{\alpha,\beta}D\phi\approx\frac{\mathcal{I}\phi(x+hb^{\alpha,\beta})-\mathcal{I}\phi(x)}{h},$$

   corresponds to our $L_k^{\alpha,\beta}$ if $k=\sqrt{h}$, $y_k^{\alpha,\beta,\pm}=k^2b^{\alpha,\beta}$.

2. The approximation of Crandall-Lions [13],

$$\frac{1}{2}\mathrm{tr}[\sigma^{\alpha,\beta}\sigma^{\alpha,\beta\top}D^2\phi]\approx\sum_{j=1}^{P}\frac{\mathcal{I}\phi(x+k\sigma_j^{\alpha,\beta})-2\mathcal{I}\phi(x)+\mathcal{I}\phi(x-k\sigma_j^{\alpha,\beta})}{2k^2},$$

   corresponds to our $L_k^{\alpha,\beta}$ if $y_{k,j}^{\alpha,\beta,\pm}=\pm k\sigma_j^{\alpha,\beta}$ and $M=P$.

3. The corrected version of the approximation of Camilli-Falcone [7] (see also [19]),

$$
\begin{aligned}
&\frac{1}{2}\mathrm{tr}[\sigma^{\alpha,\beta}\sigma^{\alpha,\beta\top}D^2\phi]+b^{\alpha,\beta}D\phi\\
&\approx\sum_{j=1}^{P}\frac{\mathcal{I}\phi(x+\sqrt{h}\sigma_j^{\alpha,\beta}+\frac{h}{P}b^{\alpha,\beta})-2\mathcal{I}\phi(x)+\mathcal{I}\phi(x-\sqrt{h}\sigma_j^{\alpha,\beta}+\frac{h}{P}b^{\alpha,\beta})}{2h},
\end{aligned}
$$

   corresponds to our $L_k^{\alpha,\beta}$ if $k=\sqrt{h}$, $y_{k,j}^{\alpha,\beta,\pm}=\pm k\sigma_j^{\alpha,\beta}+\frac{k^2}{P}b^{\alpha,\beta}$ and $M=P$.

4. The new approximation obtained by combining approximations 1 and 2,

$$
\begin{aligned}
&\frac{1}{2}\mathrm{tr}[\sigma^{\alpha,\beta}\sigma^{\alpha,\beta\top}D^2\phi]+b^{\alpha,\beta}D\phi\\
&\approx\frac{\mathcal{I}\phi(x+k^2b^{\alpha,\beta})-\mathcal{I}\phi(x)}{k^2}+\sum_{j=1}^{P}\frac{\mathcal{I}\phi(x+k\sigma_j^{\alpha,\beta})-2\mathcal{I}\phi(x)+\mathcal{I}\phi(x-k\sigma_j^{\alpha,\beta})}{2k^2},
\end{aligned}
$$

   corresponds to our $L_k^{\alpha,\beta}$ if $y_{k,j}^{\alpha,\beta,\pm}=\pm k\sigma_j^{\alpha,\beta}$ for $j\leq P$, $y_{k,P+1}^{\alpha,\beta,\pm}=k^2b^{\alpha,\beta}$ and $M=P+1$.

5. The new, more efficient version of approximation 3,

$$\frac{1}{2}\mathrm{tr}[\sigma^{\alpha,\beta}\sigma^{\alpha,\beta\top}D^2\phi]+b^{\alpha,\beta}D\phi$$
$$\approx\sum_{j=1}^{P-1}\frac{\mathcal{I}\phi(x+k\sigma_j^{\alpha,\beta})-2\mathcal{I}\phi(x)+\mathcal{I}\phi(x-k\sigma_j^{\alpha,\beta})}{2k^2}$$
$$+\frac{\mathcal{I}\phi(x+k\sigma_P^{\alpha,\beta}+k^2b^{\alpha,\beta})-2\mathcal{I}\phi(x)+\mathcal{I}\phi(x-k\sigma_P^{\alpha,\beta}+k^2b^{\alpha,\beta})}{2k^2},$$

corresponds to our $L_k^{\alpha,\beta}$ if $y_{k,j}^{\alpha,\beta,\pm}=\pm k\sigma_j^{\alpha,\beta}$ for $j<P$, $y_{k,P}^{\alpha,\beta,\pm}=\pm k\sigma_P^{\alpha,\beta}+k^2b^{\alpha,\beta}$ and $M=P$.

Approximation 5 is always more efficient than 3 in the sense that it requires fewer arithmetic operations. The most efficient of approximations 3, 4, and 5, is 4 when $\sigma^{\alpha,\beta}$ does not depend on α,β but $b^{\alpha,\beta}$ does, and 5 in the other cases.

5.2 Linear interpolation SL scheme (LISL)

To keep the scheme (3.3) monotone, linear or multi-linear interpolation is the most accurate interpolation one can use in general. In this typical case we call the full scheme (3.3)–(3.4) the LISL scheme, and we will now summarize the results of Section 4 for this special case.

Corollary 5.1. *Assume that* (A1) *and* (Y1) *hold.*

(a) The LISL scheme is monotone if the CFL conditions (4.3) *hold.*

(b) The consistency error of the LISL scheme is $O(|1-2\theta|\Delta t+\Delta t^2+k^2+\frac{\Delta x^2}{k^2})$, *and hence it is first order accurate when* $k=O(\Delta x^{1/2})$ *and* $\Delta t=O(\Delta x)$ *for* $\theta\neq\frac{1}{2}$ *or* $\Delta t=O(\Delta x^{1/2})$ *for* $\theta=\frac{1}{2}$.

(c) If $2\theta\Delta t\sup_{\alpha,\beta}|c^{\alpha,\beta,+}|_0\leq 1$ *and* (4.3) *hold, then there exists a unique bounded and* L^∞*-stable solution* U_h *of the LISL scheme converging uniformly to the solution* u *of* (1.1)–(1.2) *as* $\Delta t,k,\frac{\Delta x}{k}\to 0$.

From this result it follows that the scheme is at most *first order accurate*, has *wide and increasing stencil* and a *good CFL condition*. From the consistency error and the definition of $L_k^{\alpha,\beta}$ the stencil is wide since the scheme is consistent only if $\Delta x/k\to 0$ as $\Delta x\to 0$ and has stencil length proportional to

$$l:=\frac{\max_{t,x,\alpha,\beta,i}|y_{k,i}^{\alpha,\beta,-}|\vee|y_{k,i}^{\alpha,\beta,+}|}{\Delta x}\sim\frac{k}{\Delta x}\to\infty\quad\text{as}\quad\Delta x\to 0.$$

Here we have used that if (A1) holds and $\sigma\not\equiv 0$, then typically $y_{k,i}^{\alpha,\beta,\pm}\sim k$. Also note that if $k=\Delta x^{1/2}$, then $l\sim\Delta x^{-1/2}$. Finally, in the case $\theta\neq 1$ the CFL condition for (3.3) is $\Delta t\leq Ck^2\sim\Delta x$ when $k=O(\Delta x^{1/2})$, and it is much less restrictive than the usual parabolic CFL condition, $\Delta t=O(\Delta x^2)$.

6 A second order SL scheme for monotone solutions

In this section we introduce new second order SL schemes of the form (3.3)–(3.4) for non-degenerate grids of tensor product type. These schemes are based on monotone cubic Hermite interpolation [17, 15], and are consistent for monotone solutions of the scheme. We will call these schemes monotone cubic interpolation SL schemes or MCSL schemes in short.

To define monotone cubic Hermite interpolation, we start by considering a 1D function ϕ. For each sub-interval $[x_i,x_{i+1}]$, $i\in\mathbb{Z}$, we construct a cubic Hermite interpolant

$$(\mathcal{I}\phi)(x)=c_0+c_1(x-x_i)+c_2(x-x_i)^2+c_3(x-x_i)^3$$

fulfilling

$$(\mathcal{I}\phi)(x_i)=\phi_i,\qquad(\mathcal{I}\phi)'(x_i)=d_i,$$
$$(\mathcal{I}\phi)(x_{i+1})=\phi_{i+1},\qquad(\mathcal{I}\phi)'(x_{i+1})=d_{i+1},$$

where $\phi_i = \phi(x_i)$ and d_i is an estimate of the derivative of ϕ at x_i. It follows that

$$c_0 = \phi_i, \qquad c_1 = d_i, \tag{6.1a}$$

$$c_2 = \frac{3\Delta_i - d_{i+1} - 2d_i}{\Delta x}, \qquad c_3 = -\frac{2\Delta_i - d_{i+1} - d_i}{\Delta x^2}, \tag{6.1b}$$

where $\Delta_i = \frac{\phi_{i+1}-\phi_i}{\Delta x}$. To obtain a fourth order accurate interpolant, ϕ_i' must be at least third order accurate. We will use the symmetric fourth order approximation

$$d_i = \frac{\phi_{i-2} - 8\phi_{i-1} + 8\phi_{i+1} - \phi_{i+2}}{12\Delta x}, \quad i \in \mathbb{Z}. \tag{6.2}$$

However, the resulting interpolation is not monotone. Necessary and sufficient conditions for monotonicity were found by Fritsch and Carlson [17] (see also [24]): If $\Delta_i = 0$, then monotonicity follows if and only if $d_i = d_{i+1} = 0$, and if

$$\alpha_i = \frac{d_i}{\Delta_i} \qquad \text{and} \qquad \beta_i = \frac{d_{i+1}}{\Delta_i},$$

then monotonicity for $\Delta_i \neq 0$ follows if and only if $(\alpha_i, \beta_i) \in \mathcal{M} = \mathcal{M}_e \cup \mathcal{M}_b$ where

$$M_e = \{(\alpha, \beta) : \ (\alpha-1)^2 + (\alpha-1)(\beta-1) + (\beta-1)^2 - 3(\alpha+\beta-2) \leq 0\},$$
$$M_b = \{(\alpha, \beta) : \ 0 \leq \alpha \leq 3, \ 0 \leq \beta \leq 3\}.$$

Eisenstat, Jackson and Lewis [15] give an algorithm that modifies the derivative approximation d_i such that the above conditions are fulfilled, and for monotone data the resulting interpolant is a C^1 fourth order approximation. We will only consider C^0 interpolants, and in that case their algorithm simplifies to:

Step 1 Compute the initial d_i using (6.2).

Step 2 Compute Δ_i. If $\Delta_i \neq 0$ compute α_i and β_i, else set $\alpha_i = \beta_i = 1$.

Step 3 Set $\alpha_i := \max\{\alpha_i, 0\}$ and $\beta_i := \max\{\beta_i, 0\}$.

Step 4 If $(\alpha_i, \beta_i) \notin \mathcal{M}$, modify (α_i, β_i) as follows:

- If $\alpha_i \geq 3$ and $\beta_i \geq 3$, set $\alpha_i = \beta_i = 3$,
- else if $\beta_i > 3$ and $\alpha_i + \beta_i \geq 4$, decrease β_i such that $(\alpha_i, \beta_i) \in \partial\mathcal{M}$,
- else if $\beta_i > 3$ and $\alpha_i + \beta_i < 4$, increase α_i such that $(\alpha_i, \beta_i) \in \partial\mathcal{M}$ or $\alpha_i = 4 - \beta_i$, in the last case subsequently decrease β_i until $(\alpha_i, \beta_i) \in \partial\mathcal{M}$,
- else if $\alpha_i > 3$ and $\alpha_i + \beta_i \geq 4$, decrease α_i such that $(\alpha_i, \beta_i) \in \partial\mathcal{M}$,
- else if $\alpha_i > 3$ and $\alpha_i + \beta_i < 4$, increase β_i such that $(\alpha_i, \beta_i) \in \partial\mathcal{M}$ or $\beta_i = 4 - \alpha_i$, in the last case subsequently decrease α_i until $(\alpha_i, \beta_i) \in \partial\mathcal{M}$.

In each sub-interval $[x_i, x_{i+1}]$, we replace d_i by $\alpha_i \Delta_i$ and d_{i+1} by $\beta_i \Delta_i$ in (6.1). Multidimensional interpolation operators are obtained as tensor products of one-dimensional interpolation operators, i. e. by interpolating dimension by dimension.

Lemma 6.1. *The above monotone cubic interpolation is always monotone and satisfies* (I2). *If the interpolated function is strictly monotone between grid points, then* (I1) *holds with $r = 4$ and the method is fourth order accurate.*

Proof. Assumption (I2) holds by construction. The error estimate follows from [15], since the above algorithm coincides with the two sweep algorithm given there when $n = 1$ interval is considered. In [15] it is proved that this algorithm gives third order accurate approximations to the exact derivatives and hence the cubic Hermite polynomial constructed using this approximation is fourth order accurate. □

Remark 6.1. Carlson and Fritsch [10] constructed an alternative monotone bicubic interpolation algorithm for $\mathbb{R}^2$. We are not aware of any work on high order monotone interpolation on unstructured grids.

By the Lemma 6.1 and the results in Section 4 we have the following result:

Corollary 6.2. *Assume that* (A1), (Y1) *hold, and that for all $h \in (0,1)$, there exists a bounded solution U_h of the MCSL scheme such that $\mathcal{I}U_h$ is strictly x-monotone between points in the x-grid $X_{\Delta x}$.*

(a) The MCSL scheme is monotone if the CFL condition (4.3) *hold.*

(b) The consistency error of the MCSL scheme is $O(|1-2\theta|\Delta t+\Delta t^2+k^2+\frac{\Delta x^4}{k^2})$, and hence the scheme is second order accurate when $k=O(\Delta x)$ and $\Delta t=O(\Delta x^2)$ for $\theta \neq \frac{1}{2}$ or $\Delta t = O(\Delta x)$ for $\theta=\frac{1}{2}$.

(c) If $2\theta\Delta t \sup_{\alpha,\beta}|c^{\alpha,\beta}|_0 \leq 1$, then the solution U_h is unique, L^∞-stable, and converges uniformly to the solution u of (1.1)–(1.2) *as $\Delta t, k, \frac{\Delta x^4}{k^2} \to 0$.*

Remark 6.2. The MCSL scheme is always monotone, but for non-monotone solutions the scheme is not consistent when $k=O(\Delta x)$ since the consistency error then is $O(\Delta t + k + \frac{\Delta x^2}{k^2})$. Moreover, it is easy to see that the MCSL scheme has strictly monotone solutions (between grid points) whenever the collocation equation (7.1) (see Section 7) has strictly monotone solutions (between grid points). To prove such type of results one can use comparison principle arguments, and we refer to Appendix A for results concerning equation (1.1). Since (7.1) satisfies the comparison principle (the proof is essentially a simplified version of the proof in Appendix B), the argument proving Lemma A.1 shows that its solutions are monotone under assumption (A2) in Appendix A. Under this assumption, existence of a monotone solution follows from Theorem 4.2 in Section 4.

Remark 6.3. It is well known that consistent monotone methods for first order equations are at most first order accurate in general. However, the MCSL scheme is an example showing that second order consistent monotone schemes are possible if the solutions have special structure: When $\sigma^{\alpha,\beta} \equiv 0$, the MCSL scheme is a monotone second order scheme for a first order equation when solutions are monotone. Experts we have talked to seem to be surprised by this fact.

7 Discussion

7.1 Comparison with the scheme of Bonnans-Zidani (BZ)

In the paper [5] (see also [4, 6]) Bonnans and Zidani suggest an alternative approach to discretize non-linear degenerate diffusion equations. Their idea is to approximate the diffusion matrix $a^{\alpha,\beta}$ by a nicer matrix $a_k^{\alpha,\beta}$ which admits monotone finite difference approximations. For every $k \in \mathbb{N}$ they find a stencil

$$\mathcal{S}_k \subset \{\xi = (\xi_1,\dots,\xi_N) \in \mathbb{Z}^N : \ 0 < \max_{i=1}^{N} |\xi_i| \leq k, \ i=1,\dots,N\}$$

and positive numbers $a_{k,\xi}^{\alpha,\beta}$ such that

$$a^{\alpha,\beta} \approx a_k^{\alpha,\beta} := \sum_{\xi \in \mathcal{S}_k} a_{k,\xi}^{\alpha,\beta} \xi\xi^\top.$$

This new diffusion matrix $a_k^{\alpha,\beta}$ gives a diffusion term that can be decomposed into a linear combination of directional derivatives and these are again approximated by central difference approximations,

$$\mathrm{tr}[a^{\alpha,\beta}D^2\phi] \approx \mathrm{tr}[a_k^{\alpha,\beta}D^2\phi] = \sum_{\xi \in \mathcal{S}_k} a_{k,\xi}^{\alpha,\beta} D_\xi^2 \phi \approx \sum_{\xi \in \mathcal{S}_k} a_{k,\xi}^{\alpha,\beta} \Delta_\xi \phi,$$

where $D_\xi^2 = \mathrm{tr}[\xi\xi^T D^2] = (\xi \cdot D)^2$ and

$$\Delta_\xi w(x) = \frac{1}{|\xi|^2 \Delta x^2}\{w(x+\xi\Delta x) - 2w(x) + w(x-\xi\Delta x)\}.$$

This approximation is monotone by construction and respects the grid. In two space dimensions, $a_k^{\alpha,\beta}$ can be chosen such that $|a^{\alpha,\beta} - a_k^{\alpha,\beta}| = O(k^{-2})$ (cf. [4]), and then it is easy to see that the consistency error is

$$O(k^{-2} + k^2\Delta x^2).$$

When $b^{\alpha,\beta} \equiv 0$, the BZ scheme can be obtained from (3.3) by replacing our $L_k^{\alpha,\beta}$ by the above Bonnans-Zidani diffusion approximation. This scheme shares many properties with the LISL scheme, it is at most *first order accurate* (take $k \sim$

$\Delta x^{-1/2}$), it has a similar *wide and increasing stencil*, and it has a similar *good CFL condition* $\Delta t \leq Ck^2\Delta x^2$ ($\sim \Delta x$ when $k \sim \Delta x^{-1/2}$). To understand why the stencil is wide, simply note that k by definition is the stencil length and that the scheme is consistent only if $k \to \infty$ and $k\Delta x \to 0$. The typical stencil length is $k \sim \Delta x^{-1/2}$, just as it was for the LISL scheme.

The main drawback of this method is that it is costly since we must compute the matrix $a_k^{\alpha,\beta}$ for every x, t, α, β in the grid. The LISL scheme is easier to understand and implement and runs faster. Later we will see numerical indications that LISL runs at least 10 times faster than the BZ scheme on some test problems. The BZ scheme has the advantage that it is easy to modify to prevent it from leaving the domain (accuracy is then reduced or monotonicity is lost). It is less natural to do this for the LISL scheme. However, in many problems it is not necessary to do any modification near the boundary as we will see below.

The MCSL scheme in the typical case when $k = \Delta x$, is a *second order accurate* and *compact stencil* scheme having the usual (not so good) *CFL conditions* for parabolic problems $\Delta t \sim k^2 = \Delta x^2$. It is far more efficient than the other two schemes as can be seen in Section 10, but it is only guaranteed to converge when the computed solutions are essentially monotone. The other two schemes "always" converge. We have summarized our findings in the table below.

	order	CFL	wide stencil	boundary conditions	convergence	efficiency
BZ	1	$\Delta t \sim \Delta x$	yes	OK	always	worst
LISL	1	$\Delta t \sim \Delta x$	yes	special treatment	always	
MCSL	2	$\Delta t \sim \Delta x^2$	no	OK	monotone solutions	best

Table 1: Comparison between the BZ, LISL, and MCSL schemes.

7.2 Boundary conditions

When solving PDEs on bounded domains, the SL schemes (and the BZ schemes) may exceed the domain if they are not modified near the boundary. The reason is of course the wide stencil. This may or may not be a problem depending on the equation and the type of boundary condition: (i) For Dirichlet conditions the scheme needs to be modified near the boundary or boundary conditions must be extrapolated. This may result in a loss of accuracy or monotonicity near the boundary. (ii) Homogeneous Neumann conditions can be implemented exactly by extending in the normal direction the values of the solution on the boundary to the exterior. (iii) If the boundary has no regular points, no boundary conditions can be imposed. In this case the SL schemes may not leave the domain because the normal diffusion tends to zero fast enough when the boundary is approached. Typical examples are Black-Scholes type of equations.

7.3 Interpretation as a collocation method

The scheme (3.3)–(3.4) can be interpreted as a collocation method for a derivative free equation, this is essentially the approach of Falcone et al. [16, 7]. The idea is that if

$$W^{\Delta x}(Q_T) = \{u : u \text{ is a function on } Q_T \text{ satisfying } u \equiv \mathcal{I}u \text{ in } Q_T\}$$

denotes the interpolant space associated to the interpolation $\mathcal{I}$, equation (3.3) can be stated in the following equivalent way: Find $U \in W^{\Delta x}(Q_T)$ solving

$$\delta_{\Delta t_n} U_i^n = \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \left\{ L_k^{\alpha,\beta}[\bar{U}^{\theta,n}]_i^{n-1+\theta} + c_i^{\alpha,\beta,n-1+\theta}\bar{U}_i^{\theta,n} + f_i^{\alpha,\beta,n-1+\theta} \right\} \text{ in } G. \tag{7.1}$$

In general $W^{\Delta x}$ can be any space of approximations which is interpolating on the grid $X_{\Delta x}$, e. g. a space of splines. We do not consider this case here.

7.4 Stochastic game/control interpretation

In general the scheme (3.3)–(3.4) can be interpreted as the dynamical programming equation of a discrete stochastic differential game. We will now try to explain this in the less technical case when $\mathcal{B}$ is a singleton and the game simplifies to an optimal stochastic control problem.

Assume that (A1) holds, and for simplicity, that $c^\alpha(t,x) \equiv 0$ and the other coefficients are independent of t. Then it is well-known (cf. [25]) that the (viscosity) solution u of equation (1.1)–(1.2) is the value function of the stochastic control problem:

$$u(T-t,x) = \min_{\alpha(\cdot)\in A} E\Big[\int_t^T f^{\alpha(s)}(X_s)\,ds + g(X_T)\Big], \tag{7.2}$$

where A is a set of admissible $\mathcal{A}$-valued controls and the diffusion process $X_s = X_s^{t,x,\alpha(\cdot)}$ is constrained to satisfy the SDE

$$X_t = x \qquad \text{and} \qquad dX_s = \sigma^{\alpha(s)}(X_s)\,dW_s + b^{\alpha(s)}\,ds \quad \text{for} \quad s > t. \tag{7.3}$$

This result is a consequence of dynamical programming (DP) and (1.1) is called the DP equation of the control problem (7.2)–(7.3). Similarly, the schemes (3.3)–(3.4) are DP equations (at least in the explicit case) of suitably chosen discrete time and space control problems approximating (7.2)–(7.3). We refer to [18] for more details.

We take the slightly different approach explored in [9, 16, 19, 7] to show the relation to control theory. The idea is to write the SL scheme in collocation form (7.1) and show that (7.1) is the DP equation of a discrete time continuous space optimal control problem. We illustrate this approach by deriving an explicit scheme involving L_k^α as defined in part 4 Section 5.1. Let $\{t_0 = 0, t_1, \dots, t_M = T\}$ be discrete times and consider the discrete time approximation of (7.2)–(7.3) given by

$$\tilde{u}(T-t_m,x) = \min_{\alpha\in A_M} E\Big[\sum_{k=m}^{M-1} f^{\alpha_k}(\tilde{X}_k)\,\Delta t_{k+1} + g(\tilde{X}_M)\Big], \tag{7.4}$$

$$\tilde{X}_m = x, \quad \tilde{X}_n = \tilde{X}_{n-1} + \sigma^{\alpha_n}(\tilde{X}_{n-1})\,k_n\,\xi_n + b^{\alpha_n}(\tilde{X}_{n-1})\,k_n^2\,\eta_n, \; n > m, \tag{7.5}$$

where $k_n = \sqrt{(P+1)\Delta t_n}$, $A_M \subset A$ is an appropriate subset of piecewise constant controls, and $\xi_n = (\xi_{n,1},\dots,\xi_{n,P})^\top$ and η_n are mutually independent sequences of i. i. d. random variables satisfying

$$P\Big((\xi_{n,1},\dots,\xi_{n,P},\eta_n) = \pm e_j\Big) = \frac{1}{2(P+1)} \quad \text{if } j \in \{1,\dots,P\},$$
$$P\Big((\xi_{n,1},\dots,\xi_{n,P},\eta_n) = e_{P+1}\Big) = \frac{1}{P+1},$$

(e_j denotes the j-th unit vector) and all other values of $(\xi_{n,1},\dots,\xi_{n,P},\eta_n)$ have probability zero. Here we have used a weak Euler approximation of the SDE coupled with a quadrature approximation of the integral. Note that $\tilde{X} \approx X$ and $\tilde{u} \approx u$ when Δt is small. By DP

$$\tilde{u}(T-t_m,x) = \min_{\alpha\in A_M} E\Big[\sum_{k=m}^{n-1} f^{\alpha_k}(\tilde{X}_k)\,\Delta t_{k+1} + \tilde{u}(T-t_n,\tilde{X}_n)\Big] \quad \text{for all} \quad n > m,$$

and taking $n = m+1$, $s_{M-m} = T - t_m$, $\Delta s_m = s_m - s_{m-1}$, $\bar{k}_m = k_{M-m}$, and evaluating the expectation using (7.5), we see that

$$\tilde{u}(s_{M-m},x) =$$
$$\min_{\alpha\in\mathcal{A}} \Big\{ f^\alpha(x)\Delta s_{M-m} + \frac{\bar{k}^2_{M-m-1}}{P+1} L^\alpha_{\bar{k}_{M-m-1}}[\tilde{u}](s_{M-m-1},x) + \tilde{u}(s_{M-m-1},x)\Big\},$$

where L_k^α is as in Section 5.1 part 4. If we subtract $\tilde{u}(s_{M-m-1},x)$ from both sides and divide by $\Delta s_{M-m} = \frac{\bar{k}^2_{M-m-1}}{P+1}$, we find (7.1) with $\theta = 0$.

In [7], a similar argument is given in the stationary case for schemes involving the L_k^α of part 3 Section 5.1. In fact it is possible to identify all L_k^α's appearing in Section 5.1 with DP equations of suitably chosen discrete time continuous space control problems. However assumption (Y1) is not strong enough for this approach to work for the general L_k^α defined in Sections 3 and 4.

Remark 7.1. A DP approach naturally leads to explicit methods for time dependent PDEs. But implicit methods can also be derived using a trick. Discretize the PDE in time by backward Euler to find a (sequence of) stationary PDEs and use the DP approach on each stationary PDE. For stationary problems the DP equation is always implicit, so the result is an implicit iteration scheme.

Remark 7.2. By the definition of L_k^α and (Y1), $x + y_{i,k}^{\alpha,\pm}$ can be seen as a short time approximation of (7.3). Hence the scheme (3.3) tracks particle paths approximately. In fact by the above discussion we might say that the scheme follows particles in the mean because of the expectation. For first order PDEs, schemes defined in this way are called SL schemes by e. g. Falcone. Moreover, in this case our schemes will coincide with the SL schemes of Falcone [16] in the explicit case. This explains why we choose to call these schemes SL schemes also in the general case.

8 Error estimates in the convex case

We will derive error bounds in the case when $\mathcal{B}$ is a singleton and (1.1) is convex. It is not known how to prove such results in the general case. Here and in the following, we do not indicate the trivial β dependence any more. For simplicity we also take a uniform time-grid, letting $G = \Delta t\,\{0,1,\dots,N_T\} \times X_{\Delta x}$ in this section. Let $Q_{\Delta t,T} := \Delta t\,\{0,1,\dots,N_T\} \times \mathbb{R}^N$ and consider the intermediate equation

$$\delta_{\Delta t} V^n(x) = \tag{8.1}$$

$$\inf_{\alpha\in\mathcal{A}} \left\{ L_k^\alpha[\bar{V}^{\theta,n}](t,x) + c^\alpha(t,x)\bar{V}^{\theta,n}(x) + f^\alpha(t,x) \right\}_{t=t_{n-1+\theta}} \quad \text{in} \quad Q_{\Delta t,T},$$

$$V(0,x) = g(x) \quad \text{in} \quad \mathbb{R}^N. \tag{8.2}$$

The first step is now to find a bound on $|U - V|$.

Lemma 8.1. *Assume that (I1), (I2) and the CFL condition* (4.3) *hold and that* $\sup_n |V^n|_1 \le C_V$. *If* V *solves* (8.1)–(8.2) *and* U *solves* (3.3)–(3.4)*, then*

$$|U - V| \le C\frac{\Delta x}{k^2} \quad in \quad G.$$

Proof. Let $W = U - V$ and subtract the equation for V from the one for U to find

$$W_i^n \le W_i^{n-1} + \Delta t \sup_{\alpha\in\mathcal{A}} \Big\{ L_k^\alpha[\mathcal{I}\bar{W}_\cdot^{\theta,n}]_i^{n-1+\theta} + c_i^{\alpha,n-1+\theta}\bar{W}_i^{\theta,n}$$
$$+ L_k^\alpha[\mathcal{I}\bar{V}^{\theta,n} - \bar{V}_\cdot^{\theta,n}]_i^{n-1+\theta} \Big\} \quad \text{in} \quad G.$$

Assuming V^n has p bounded derivatives, we rearrange the equation and use (I1) to see that

$$\Big(1 + \theta\Delta t\Big(\frac{M}{k^2} - c_{W,i}^{n-1+\theta}\Big)\Big) W_i^n$$
$$\le W_i^{n-1} + \Delta t \sup_{\alpha\in\mathcal{A}} \Big\{ \theta\Big(L_k^\alpha[\mathcal{I}W_\cdot^n]_i^{n-1+\theta} + \frac{M}{k^2}W_i^n \Big)$$
$$+ (1-\theta)\Big(L_k^\alpha[\mathcal{I}W_\cdot^{n-1}]_i^{n-1+\theta} + c_i^{\alpha,n-1+\theta}W_i^{n-1} \Big) \Big\}$$
$$+ 2\Delta t K \sup_n |D^{r\wedge p}V^n|_0 \frac{\Delta x^{r\wedge p}}{k^2} \quad \text{in} \quad G$$

with $c_{W,i}^{n-1+\theta}W_i^n = \sup_\alpha c_i^{\alpha,n-1+\theta}W_i^n$. By the CFL condition (4.3), the coefficients of the above inequality are all non-negative. Hence since $W^n \le |W^n|_0 := \sup_i |W_i^n|$, we may replace W^n by $|W^n|_0$ on the right hand side. Moreover, since $\mathcal{I}|W^n|_0 = |W^n|_0$ and $L_k^\alpha[|W^n|_0] = 0$, the upper bound on the right hand side then reduces to

$$(1 + \Delta t(1-\theta)C_c)|W^{n-1}|_0 + \theta\Delta t\frac{M}{k^2}|W^n|_0 + \Delta t\tilde{K}\frac{\Delta x^{r\wedge p}}{k^2},$$

where $C_c = \max_\alpha |c^{\alpha,+}|_0$. The same bound holds if we replace W by $-W$, and hence we can conclude that

$$(1 + \Delta t\theta(\frac{M}{k^2} - C_c))|W^n|_0 \le (1 + \Delta t(1-\theta)C_c)|W^{n-1}|_0 + \theta\Delta t\frac{M}{k^2}|W^n|_0 + \Delta t\tilde{K}\frac{\Delta x^{r\wedge p}}{k^2}.$$

Since $W^0 \equiv 0$ in $X_{\Delta x}$, an iteration then reveals that

$$|W^n|_0 \le \Delta t\tilde{K}\frac{\Delta x^{r\wedge p}}{k^2}\sum_{m=0}^{n}\Big(\frac{1+\Delta t(1-\theta)C_c}{1-\Delta t\theta C_c}\Big)^m \le t_n\tilde{K}\frac{\Delta x^{r\wedge p}}{k^2}2e^{C_c t_n},$$

when Δt is small enough. Since V^n is Lipschitz ($p = 1$), the lemma follows. □

Next we want to estimate $|V-u|$ when u solves (1.1)–(1.2). This can be done using the regularization method of Krylov if we can find suitable continuity and continuous dependence results for the scheme. These results rely on the following additional (covariance-type) assumptions: Whenever two sets of data σ, b and $\tilde{\sigma}, \tilde{b}$ are given, the corresponding approximations $L_k^\alpha, y_{k,i}^{\alpha,\pm}$ and $\tilde{L}_k^\alpha, \tilde{y}_{k,i}^{\alpha,\pm}$ in (3.1) satisfy

$$\begin{cases} \sum_{i=1}^{M}[y_{k,i}^{\alpha,+}+y_{k,i}^{\alpha,-}]-[\tilde{y}_{k,i}^{\alpha,+}+\tilde{y}_{k,i}^{\alpha,-}] \leq 2k^2(b^\alpha-\tilde{b}^\alpha), \\ \sum_{i=1}^{M}[y_{k,i}^{\alpha,+}\otimes y_{k,i}^{\alpha,+}+y_{k,i}^{\alpha,-}\otimes y_{k,i}^{\alpha,-}]+[\tilde{y}_{k,i}^{\alpha,+}\otimes \tilde{y}_{k,i}^{\alpha,+}+\tilde{y}_{k,i}^{\alpha,-}\otimes \tilde{y}_{k,i}^{\alpha,-}] \\ \qquad -[y_{k,i}^{\alpha,+}\otimes \tilde{y}_{k,i}^{\alpha,+}+\tilde{y}_{k,i}^{\alpha,+}\otimes y_{k,i}^{\alpha,+}+y_{k,i}^{\alpha,-}\otimes \tilde{y}_{k,i}^{\alpha,-}+\tilde{y}_{k,i}^{\alpha,-}\otimes y_{k,i}^{\alpha,-}] \\ \leq 2k^2(\sigma^\alpha-\tilde{\sigma}^\alpha)(\sigma^\alpha-\tilde{\sigma}^\alpha)^\top+2k^4(b^\alpha-\tilde{b}^\alpha)(b^\alpha-\tilde{b}^\alpha)^\top, \end{cases} \tag{Y2}$$

when $\sigma, b, y_k^\pm$ are evaluated at (t,x) and $\tilde{\sigma}, \tilde{b}, \tilde{y}_k^\pm$ are evaluated at (t,y) for all t,x,y.

In Section 9 we will prove the following error estimate.

Theorem 8.2. *Assume that $\mathcal{B}$ is a singleton, that* (A1), (Y1), (Y2), *and the CFL conditions* (4.3) *hold, and that* $k \in (0,1)$ *and* $\Delta t \leq (2k_0 \wedge 2k_1)^{-1}$. *If u and V are bounded solutions of* (1.1)–(1.2) *and* (8.1)–(8.2), *then*

$$|V-u| \leq C(|1-2\theta|\Delta t^{1/4}+\Delta t^{1/3}+k^{1/2}) \quad in \quad Q_{\Delta t,T}.$$

It also follows from the regularity results in Section 9 (see Proposition 9.4) that $|V^n|_1 \leq 2C_T$, so by Lemma 8.1 and Theorem 8.2 we have the following result.

Corollary 8.3 (Error Bound)**.** *Under (I1), (I2), and the assumptions of Theorem 8.2, if u solves* (1.1)–(1.2) *and U solves* (3.3)–(3.4), *then*

$$|u-U| \leq |u-V|+|V-U| \leq C(|1-2\theta|\Delta t^{1/4}+\Delta t^{1/3}+k^{1/2}+\frac{\Delta x}{k^2}) \quad in \quad G.$$

This error bound applies to both the LISL and MCSL schemes, and it also holds for unstructured grids. If the solutions are more regular, it is possible to obtain better error estimates. But general and optimal results are not available. The best estimate in our case is $O(\Delta x^{1/5})$ which is achieved when $k=O(\Delta x^{2/5})$ and $\Delta t = O(k^2)$. Note that the CFL conditions (4.3) already imply that $\Delta t = O(k^2)$ if $\theta < 1$. Also note that the above bound does not show convergence when k is optimal for the LISL scheme ($k = O(\Delta x^{1/2})$) or the MCSL scheme ($k=O(\Delta x)$).

Remark 8.1. These results are consistent with results for LISL type schemes for stationary Bellman equations. In fact if all coefficients are independent of time and $c^\alpha(x) < -c < 0$, then by combining the results of [7] and [1], exactly the same error estimate is obtained for the solution of a particular stationary LISL scheme and the unique *stationary* Lipschitz solution of (1.1).

9 Proof of Theorem 8.2

We start by an existence and uniqueness result.

Lemma 9.1. *Assume that* (A1), (Y1), *and the CFL conditions* (4.3) *hold. Then there exists a unique solution* $U_h \in C_b(Q_{T,\Delta t})$ *of* (8.1)–(8.2).

The proof is similar to (but simpler than) the proof of Theorem 4.2 with the modification that the fixed point is achieved in the Banach space $C_b(\mathbb{R}^N)$ instead of the space of bounded functions on $X_{\Delta x}$.

We will now give a result comparing subsolutions of (8.1) to supersolutions of

$$\begin{aligned} \delta_{\Delta t}U^n(x) &= \inf_{\alpha\in\mathcal{A}}\left\{\tilde{L}_k^\alpha[\bar{U}^{\theta,n}](t,x)+\tilde{c}^\alpha(t,x)\bar{U}^{\theta,n}+\tilde{f}^\alpha(t,x)\right\}_{t=t_{n-1+\theta}} \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in } \mathbb{R}^N, n\geq 1, \\ U(0,x) &= \tilde{g}(x) \quad \text{in} \quad \mathbb{R}^N, \end{aligned} \tag{9.1}$$

where $\tilde{L}_k^\alpha$ is the operator defined in (3.1), (Y1), (Y2) when σ^α, b^α are replaced by $\tilde{\sigma}^\alpha, \tilde{b}^\alpha$.

Theorem 9.2. *Assume that* (A1)*,* (Y2)*,* (4.3) *hold for both* (8.1) *and* (9.1)*. If $U \in C(Q_{T,\Delta t})$ is a bounded above subsolution of* (8.1) *and $\tilde{U} \in C(Q_{T,\Delta t})$ a bounded below supersolution of* (9.1)*, then for all $k \in (0,1)$, $\Delta t \le (k_0 \wedge k_1)^{-1}$, $x, y \in \mathbb{R}^N$, $n \in \{0, 1, \dots, N_T\}$,*

$$\begin{aligned} U(t_n,x) - \tilde{U}(t_n,y) \le R_{k_0}(t_n)|(U(0,\cdot) - \tilde{U}(0,\cdot))^+|_0 \\ + R_{k_0}(t_n)R_{k_1}(t_n)(L_0 + t_n L)|x-y| \\ + t_n \sup_{\alpha\in\mathcal{A}} \left[|(f-\tilde{f})^+|_0 + R_{k_0}(t_n)(|U|_0 \wedge |\tilde{U}|_0)|c-\tilde{c}|_0\right] \\ + t_n^{1/2} 2K_T \sup_{\alpha\in\mathcal{A}} \left[|b-\tilde{b}|_0 + |\sigma - \tilde{\sigma}|_0\right] \end{aligned}$$

where $R_k(t) = 1/(1-k\Delta t)^{t/\Delta t}$, $K_T \le R_{k_0}(T)R_{k_1}(T)(L_0 + TL)$,

$$\begin{aligned} L_0 = |g|_1 \vee |\tilde{g}|_1, \quad L = (|c^\alpha|_1 \vee |\tilde{c}^\alpha|_1)(|U|_0 \wedge |\tilde{U}|_0) + |f^\alpha|_1 \vee |\tilde{f}^\alpha|_1, \\ k_0 = \sup_\alpha |c^{\alpha,+}|_0, \quad k_1 = 2\sup_\alpha \{|\sigma^\alpha|_1^2 + |b^\alpha|_1^2 + 1\}. \end{aligned}$$

Remark 9.1. The function $R_k(n\Delta t) = 1/(1-k\Delta t)^n$ satisfies

$$\delta_{\Delta t} R_k(t_n) = kR_k(t_n),$$

$R_k(0) = 1$, and $R_k(t_n) \le e^{2kt_n}$ when $\Delta t \le \frac{1}{2k}$.

This is a key result in this paper, and the proof is given in Appendix B. In the stationary case, results of this type have been obtained in [1, 8] for simpler schemes. The result is a joint uniqueness result (take $(\tilde{\sigma}, \tilde{b}, \tilde{c}, \tilde{f}, \tilde{g}) = (\sigma, b, c, f, g)$), continuous dependence result (take $x = y$), boundedness, and x-Lipschitz continuity result:

Corollary 9.3. *Under the assumptions of Theorem 9.2, if $k \in (0,1)$ and $\Delta t \le (2k_0 \wedge 2k_1)^{-1}$, then any bounded solution $U \in C_b(Q_{T,\Delta t})$ of* (8.1) *satisfies*

(i) $|U(t_n,\cdot)|_0 \le e^{2k_0 t_n}|g|_0 + t_n \sup_\alpha |f^\alpha|_0$,

(ii) $|U(t_n,x) - U(t_n,y)| \le e^{2(k_0+k_1)t_n}(L_0 + t_n L)|x-y|$,

where the constants, which are defined in Theorem 9.2, are independent of $k, \Delta t, \Delta x$.

Proof. Part (i) follows from Theorem 9.2 and Remark 9.1 since $\tilde{U} \equiv 0$ satisfies (9.1) with $(\tilde{\sigma}^\alpha, \tilde{b}^\alpha, \tilde{c}^\alpha, \tilde{f}^\alpha, \tilde{g}^\alpha) = (\sigma^\alpha, b^\alpha, c^\alpha, 0, 0)$. Part (ii) follows by taking $U = \tilde{U}$ and $x \neq y$. ☐

Now we extend the scheme (8.1) to the whole space Q_T. One way to do this and to obtain continuous in time solutions, is to pose initial conditions on $[0, \Delta t)$ by interpolating between $g(x)$ and $U(\Delta t, x)$ where U is the solution of (8.1)–(8.2).

$$\delta_{\Delta t} V(t,x) = \inf_{\alpha\in\mathcal{A}} \left\{ L_k^\alpha[\bar{V}^\theta(t,\cdot)](t^\theta,x) + c^\alpha(t^\theta,x)\bar{V}^\theta(t,x) + f^\alpha(t^\theta,x)\right\} \quad \text{in } (\Delta t, T] \times \mathbb{R}^N, \tag{9.2}$$

$$V(t,x) = \left(1 - \frac{t}{\Delta t}\right) g(x) + \frac{t}{\Delta t} U(\Delta t, x) \quad \text{in} \quad [0, \Delta t] \times \mathbb{R}^N. \tag{9.3}$$

where $\bar{V}^\theta(t,x) = (1-\theta)V(t-\Delta t, x) + \theta V(t,x)$ and $t^\theta = t - (1-\theta)\Delta t$.

From the previous results for U the existence, uniqueness, and properties of V easily follow.

Proposition 9.4. *Assume that* (A1)*,* (Y1)*,* (Y2)*, and the CFL conditions* (4.3) *hold, and that $k \in (0,1)$ and $\Delta t \le (2k_0 \wedge 2k_1)^{-1}$.*

(a) There exists a unique solution $V \in C_b(Q_T)$ of (9.2)–(9.3).

(b) There is a constant $C_T \ge 0$ independent of $k, \Delta t, \Delta x$ such that

(i) $|V|_0 \le C_T$,

(ii) $|V(t,x) - V(t,y)| \le C_T|x-y|$ *for all* $t \in [0,T],\ x,y,\in \mathbb{R}^N$,

(iii) $|V(s_1,x) - V(s_2,x)| \le C_T|s_1-s_2|^{1/2}$ *for all* $s_1,s_2 \in [0,T],\ x,\in \mathbb{R}^N$.

(c) Let $V \in C_b(Q_T)$ *and* $\tilde{V} \in C_b(Q_T)$ *be sub- and supersolutions of* (9.2)–(9.3) *corresponding to coefficients* $(\sigma^\alpha, b^\alpha, c^\alpha, f^\alpha, g)$ *and* $(\tilde{\sigma}^\alpha, \tilde{b}^\alpha, \tilde{c}^\alpha, \tilde{f}^\alpha, \tilde{g})$ *respectively. Then there is a constant* $C_T \ge 0$ *independent of* $k, \Delta t, \Delta x$ *such that for all* $t \in [0,T]$,

$$|V(t,\cdot) - \tilde{V}(t,\cdot)|_0 \le C_T\Big(|g-\tilde{g}|_0 + t\sup_\alpha[(|U|_0 \wedge |\tilde{U}|_0)|c^\alpha - \tilde{c}^\alpha|_0 + |f^\alpha - \tilde{f}^\alpha|_0] \\ + t^{1/2}\sup_\alpha[|\sigma^\alpha - \tilde{\sigma}^\alpha|_0 + |b^\alpha - \tilde{b}^\alpha|_0]\Big).$$

Proof. First note that the initial data on $[0,\Delta t]$ is uniformly bounded and Lipschitz continuous in x and t by construction and Corollary 9.3.

(a) Existence of a bounded x-continuous solution follows from repeated use of Lemma 9.1 since we have initial conditions on $[0,\Delta t]$. Continuity in time follows from Theorem 9.2 (with $x = y$) since the data is t-continuous.

(b) Part (i) and (ii) follow from Corollary 9.3 since the initial data is uniformly bounded and x-Lipschitz in $[0,\Delta t]$. To prove part (iii) we assume $s_1 < s_2$ and let $U(t,x)$ and $\tilde{U}(t,x)$ solve (9.2) with data

$$(\sigma^\alpha(t+s_1,x), b^\alpha(t+s_1,x), c^\alpha(t+s_1,x), f^\alpha(t+s_1,x), V(s_1,x)) \text{ and } (0,0,0,0,V(s_1,x))$$

respectively. Note that for $t \in [0, T-s_1]$, $\tilde{U}(t,x) \equiv V(s_1,x)$ and $U(t,x) \equiv V(t+s_1,x)$ where V is the unique solution of (9.2)–(9.3). By part (c) we then get

$$\begin{aligned}&|V(t+s_1,\cdot) - V(s_1,\cdot)|_0 = |U(t,\cdot) - \tilde{U}(t,\cdot)|_0 \\ &\le C_T\Big(0 + t\sup_\alpha[|f^\alpha|_0 + |V|_0|c^\alpha|_0] + t^{1/2}\sup_\alpha[|\sigma^\alpha|_0 + |b^\alpha|_0]\Big) \quad \text{for} \quad t > 0,\end{aligned}$$

and hence part (iii) follows.

(c) Note that by construction of the initial data and Theorem 9.2 with $x = y$, the result holds for $t \in [0,\Delta t]$, and then the result holds for any $t > \Delta t$ by another application of Theorem 9.2 with $x = y$. □

Using Krylov's method of shaking the coefficients, we will now find smooth subsolutions of (9.2). First we introduce the auxiliary equation

$$\delta_{\Delta t}V^\varepsilon(t,x) = \inf_{\substack{0\le s\le\varepsilon^2\\ |e|\le\varepsilon\\ \alpha\in\mathcal{A}}}\Big\{L_k^\alpha[\tau_{-e}\bar{V}^{\varepsilon,\theta}(t,\cdot)](r+s,x+e) \tag{9.4}$$

$$+ c^\alpha(r+s,x+e)\bar{V}^{\varepsilon,\theta}(t,x) + f^\alpha(r+s,x+e)\Big\}_{r=t^\theta-\Delta t-\varepsilon^2} \text{ in } (\Delta t, T] \times \mathbb{R}^N,$$

$$V^\varepsilon(t,x) = \Big(1 - \frac{t}{\Delta t}\Big)g(x) + \frac{t}{\Delta t}V^\varepsilon(\Delta t,x) \quad \text{in} \quad [0,\Delta t] \times \mathbb{R}^N, \tag{9.5}$$

where $\tau_e\phi(t,x) = \phi(t,x+e)$ and $V^\varepsilon(\Delta t,x)$ is obtained by first solving (9.4) for discrete times $t_n = n\Delta t$. For this equation to be well-defined for $t \in (\Delta t, T]$, the data and $y_{k,i}^{\alpha,\pm}$ must be defined for $t \in (-\Delta t - \varepsilon^2, T + \varepsilon^2]$. But this is ok since one can easily extend these functions to $t \in [-r, T+r]$ for any $r > 0$ in such a way that (A1), (Y1), (Y2) still hold. Also note that

$$L_k^\alpha[\tau_{-e}\bar{V}^{\varepsilon,\theta}(t,\cdot)](r+s,x+e) = \frac{1}{2k^2}\sum_{i=1}^M\Big\{\bar{V}^{\varepsilon,\theta}(t,x+y_{k,i}^{\alpha,+}(r+s,x+e)) \tag{9.6}$$

$$-2\bar{V}^{\varepsilon,\theta}(t,x) + \bar{V}^{\varepsilon,\theta}(t,x+y_{k,i}^{\alpha,-}(r+s,x+e))\Big\},$$

and hence (9.4) is an equation of the same type as (9.2) (with different $\mathcal{A}$ and shifted coefficients) satisfying (A1), (Y1), (Y2) whenever (9.2) does.

By Proposition 9.4 there is a unique solution V^ε of (9.4)–(9.5) in $[0, T+\Delta t+\varepsilon^2]\times\mathbb{R}^N$. Let $U^\varepsilon(t,x) := V^\varepsilon(t+\Delta t+\varepsilon^2, x)$ and define by convolution,

$$U_\varepsilon(t,x) = \int_{\mathbb{R}^N}\int_0^\infty U^\varepsilon(t-s, x-e)\rho_\varepsilon(s,e)\,ds\,de, \tag{9.7}$$

where $\varepsilon > 0$, $\rho_\varepsilon(t,x) = \frac{1}{\varepsilon^{N+2}}\rho(\frac{t}{\varepsilon^2}\frac{x}{\varepsilon})$, and

$$\rho\in C^\infty(\mathbb{R}^{N+1}),\quad \rho\ge 0,\quad \operatorname{supp}\rho\subset[0,1]\times\{|x|\le 1\},\quad \int\rho = 1.$$

Note that U_ε is well defined on the time interval $[-\Delta t, T]$. By the next result it is the sought after smooth subsolution of (9.2).

Proposition 9.5. *Under the assumptions of Proposition 9.4, the function U_ε defined in* (9.7) *satisfies*

(i) $U_\varepsilon\in C^\infty((-\Delta t, T)\times\mathbb{R}^N)$, $|U_\varepsilon|_1\le C$, $|D^m\partial_t^n U_\varepsilon|_0\le C\varepsilon^{1-m-2n}$ *for* $n, m\in\mathbb{N}$.

(ii) If V is the solution of (9.2)–(9.3)*, then* $|U_\varepsilon - V|\le C(\varepsilon+\Delta t^{1/2})$ *in* Q_T.

(iii) U_ε *is a subsolution of* (9.2) *in* Q_T.

Proof. The regularity estimates in (i) are immediate from properties of convolutions and the regularity of V^ε. The bound on $U_\varepsilon - V$ (in $[0,T]$) in (ii) follows from Proposition 9.4 (c) and (A1) which imply

$$|V^\varepsilon - V|_0\le C(\varepsilon+\Delta t^{1/2}),$$

and regularity of V^ε along with properties of convolutions,

$$|U_\varepsilon - V^\varepsilon|_0\le|U_\varepsilon - U^\varepsilon|_0 + |V^\varepsilon(\cdot+\Delta t+\varepsilon^2,\cdot) - V^\varepsilon|_0\le|V_\varepsilon|_1(\varepsilon+\Delta t^{1/2}).$$

To see that U_ε is a subsolution of (9.2), first note that from the definition of U^ε and (9.4) it follows that

$$\begin{aligned}\delta_{\Delta t}U^\varepsilon(t,x)\le{}& L_k^\alpha[\tau_{-e}\bar U^{\varepsilon,\theta}(t,\cdot)](t^\theta+s, x+e)\\ &+c^\alpha(t^\theta+s, x+e)\bar U^{\varepsilon,\theta}(t,x)+f^\alpha(t^\theta+s,x+e)\end{aligned}$$

for all $(t,x)\in[-\varepsilon^2, T]\times\mathbb{R}^N$, $|e|, s^2\le\varepsilon$, and $\alpha\in\mathcal{A}$. Now we change variables from $(t+s, x+e)$ to (t,x) to find that

$$\begin{aligned}\delta_{\Delta t}U^\varepsilon(t-s,x-e)\le{}& L_k^\alpha[\tau_{-e}\bar U^{\varepsilon,\theta}(t-s,\cdot)](t^\theta, x)\\ &+c^\alpha(t^\theta,x)\bar U^{\varepsilon,\theta}(t-s,x-e)+f^\alpha(t^\theta,x)\end{aligned}$$

for all $(t,x)\in[0,T]\times\mathbb{R}^N$, $|e|, s^2\le\varepsilon$, and $\alpha\in\mathcal{A}$. Then we multiply by $\rho_\varepsilon(s,e)$ and integrate w. r. t. (s,e). To see what the result is, note that

$$\begin{aligned}L_k^\alpha[\tau_{-e}U^\varepsilon(t-s,\cdot)](r,x) = \frac{1}{2k^2}\sum_{i=1}^M\Big\{&U^\varepsilon(t-s, x+y_{k,i}^{\alpha,+}(r,x)-e)\\ &-2U^\varepsilon(t-s,x-e)+U^\varepsilon(t-s,x+y_{k,i}^{\alpha,-}(r,x)-e)\Big\},\end{aligned}$$

and hence

$$\int\int L_k^\alpha[\tau_{-e}U^\varepsilon(t-s,\cdot)](r,x)\rho_\varepsilon(s,e)\,ds\,de = L^\alpha[U_\varepsilon(t,\cdot)](r,x).$$

For the whole equation we then have,

$$\delta_{\Delta t}U_\varepsilon(t,x)\le L_k^\alpha[\bar U_\varepsilon^\theta(t,\cdot)](t^\theta,x)+c^\alpha(t^\theta,x)\bar U_\varepsilon^\theta(t,x)+f^\alpha(t^\theta,x)$$

for all $(t,x)\in Q_T$ and $\alpha\in\mathcal{A}$. Since this inequality holds for all α, it follows that U_ε is a subsolution of (9.2) in all of Q_T. □

We are now in a position to prove the error estimate given in Theorem 8.2.

Proof of Theorem 8.2. Let U_ε be defined in (9.7). By Proposition 9.5 (i) and Lemma 4.1 (a),

$$\begin{aligned}
&\partial_t U_\varepsilon - \inf_{\alpha\in\mathcal{A}} \Big\{ L^\alpha[\bar{U}^{\varepsilon,\theta}(t,\cdot)](t^\theta,x) + c^\alpha(t^\theta,x)\bar{U}^{\varepsilon,\theta}(t,x) + f^\alpha(t^\theta,x) \Big\} \\
&\leq \frac{|1-2\theta|}{2} |\partial_t^2 U_\varepsilon|_0 \Delta t + C\Big\{ (|\partial_t^2 U_\varepsilon|_0 + |\partial_t^3 U_\varepsilon|_0 + |\partial_t^2 D U_\varepsilon|_0 + |\partial_t^2 D^2 U_\varepsilon|_0)\Delta t^2 \\
&\quad + (|DU_\varepsilon|_0 + \cdots + |D^4 U_\varepsilon|_0) k^2 \Big\} \\
&\leq C\Big\{ |1-2\theta|\varepsilon^{-3}\Delta t + \varepsilon^{-5}\Delta t^2 + \varepsilon^{-3}k^2 \Big\}
\end{aligned}$$

in Q_T. Moreover, by Proposition 9.5 (ii),

$$g(x) = U(0,x) \geq U_\varepsilon(0,x) - C(\varepsilon + \Delta t^{1/2}).$$

It follows that there is a constant $C \geq 0$ such that

$$U_\varepsilon - C e^{\sup_\alpha |c^\alpha|_0 t} \Big\{ \varepsilon + \Delta t^{1/2} + t\Big(|1-2\theta|\varepsilon^{-3}\Delta t + \varepsilon^{-5}\Delta t^2 + \varepsilon^{-3}k^2 \Big) \Big\}$$

is a classical subsolution of (1.1)–(1.2). By the comparison principle

$$U_\varepsilon - C e^{\sup_\alpha |c^\alpha|_0 t} \Big\{ \varepsilon + \Delta t^{1/2} + t\Big(|1-2\theta|\varepsilon^{-3}\Delta t + \varepsilon^{-5}\Delta t^2 + \varepsilon^{-3}k^2 \Big) \Big\} \leq u \text{ in } Q_T,$$

and hence by Proposition 9.5 (ii),

$$U - u = (U - U_\varepsilon) + (U_\varepsilon - u) \leq C\Big\{ \varepsilon + \Delta t^{1/2} + |1-2\theta|\varepsilon^{-3}\Delta t + \varepsilon^{-5}\Delta t^2 + \varepsilon^{-3}k^2 \Big\}.$$

We minimize w. r. t. ε and find that

$$u - U \leq \begin{cases} C(\Delta t^{1/4} + k^{1/2}) & \text{if } \theta \neq \frac{1}{2} \\ C(\Delta t^{1/3} + k^{1/2}) & \text{if } \theta = \frac{1}{2} \end{cases} \quad \text{in} \quad Q_T.$$

The lower bound on $u - U$ follows with symmetric – but much easier – arguments where a smooth supersolution of the equation (1.1) is constructed. Consistency and comparison for the scheme (9.2) is then used to conclude. In view of Lemma 4.1, the lower bound is a direct consequence of Theorem 3.1 (a) in [2]. □

10 Numerical results

In the following, we apply the LISL and MCSL schemes to linear and convex test problems in two space-dimensions. Hence all problems in this section are independent of β. For the LISL scheme, we choose $k = \sqrt{\Delta x}$ and a regular triangular grid, whereas for the MCSL scheme we choose $k = \Delta x$ and a regular rectangular grid. If not stated otherwise, we use $\theta = 0$ (explicit methods), CFL condition $\Delta t = k^2$, and approximation 5.1.5 for $L^{\alpha,\beta}$. As error measure we will always use the L^∞-norm, and the error rates are calculated as $r_i = \frac{\ln\|e_i\| - \ln\|e_{i-1}\|}{\ln\|\Delta x_i\| - \ln\|\Delta x_{i-1}\|}$. All calculations are done in Matlab, on an INTEL Core2 Duo Mobile T7700, 2.4Ghz Laptop.

10.1 Linear problem with smooth solution

Our first problem is taken from [4] and has exact solution $u(t,x) = (2-t)\sin x_1 \sin x_2$. The coefficients in (1.1) are given by

$$\begin{aligned}
f^\alpha(t,x) =& \sin x_1 \sin x_2 [(1+2\beta^2)(2-t) - 1] \\
&\quad - 2(2-t)\cos x_1 \cos x_2 \sin(x_1+x_2)\cos(x_1+x_2), \\
c^\alpha(t,x) =& 0, \qquad b^\alpha(t,x) = 0, \qquad \sigma^\alpha(t,x) = \sqrt{2}\begin{pmatrix} \sin(x_1+x_2) & \beta & 0 \\ \cos(x_1+x_2) & 0 & \beta \end{pmatrix}.
\end{aligned}$$

We consider $\beta^2 = 0.1$ and $\beta = 0$. Note that in the second case, the scheme considered in [4] is not consistent. Table 2 gives the (spatial) errors and rates obtained at $t = 1$ applying the LISL and the MCSL scheme. As expected for smooth solutions, in both cases we obtain order one for the LISL scheme and order two for the MCSL scheme. Here, we have chosen the grid points such that the solution is monotone in between. Without this, we would still obtain order one for the LISL scheme, but no convergence for the MCSL scheme. The reason is that for non-monotone data, the interpolation error of monotone cubic interpolation reduces to second order, and so the choice $k = \Delta x$ is not longer appropriate.

Δx	$\beta^2 = 0.1$ LISL error	rate	MCSL error	rate	$\beta = 0$ LISL error	rate	MCSL error	rate
3.93e-2	3.79e-2	0.86	1.03e-3		3.94e-2	0.87	1.03e-3	
1.96e-2	1.93e-2	0.97	2.57e-4	2.00	1.98e-2	0.99	2.57e-4	2.00
9.82e-3	9.45e-3	1.03	6.42e-5	2.00	9.94e-3	0.99	6.43e-5	2.00
4.91e-3	4.50e-3	1.07	1.61e-5	2.00	4.70e-3	1.08	1.61e-5	2.00
2.45e-3	2.43e-3	0.89	4.01e-6	2.00	2.45e-3	0.94	4.02e-6	2.00

Table 2: Results for the smooth linear problem at $t = 1$ with $\beta^2 = 0.1$ and $\beta = 0$, grid adapted to monotonicity

10.2 Linear problem with non-smooth solution

The second problem we test is a problem with non-smooth exact solution in $[-\pi, \pi]^2$ given by

$$u(t,x) = (1+t)\sin\frac{x_2}{2}\begin{cases}\sin\frac{x_1}{2} & \text{for } -\pi < x_1 < 0,\\ \sin\frac{x_1}{4} & \text{for } 0 < x_1 < \pi.\end{cases}$$

The coefficients in (1.1) are given by

$$\begin{aligned}f^\alpha(t,x) = {} & \sin\frac{x_2}{2}\begin{cases}\sin\frac{x_1}{2}\left(1+\frac{1+t}{4}(\sin^2 x_1+\sin^2 x_2)\right) & \text{for } -\pi < x_1 < 0\\ \sin\frac{x_1}{4}\left(1+\frac{1+t}{16}(\sin^2 x_1+4\sin^2 x_2)\right) & \text{for } 0 < x_1 < \pi\end{cases}\\ & -\sin x_1 \sin x_2 \cos\frac{x_2}{2}\begin{cases}\frac{1+t}{2}\cos\frac{x_1}{2} & \text{for } -\pi < x_1 < 0\\ \frac{1+t}{4}\cos\frac{x_1}{4} & \text{for } 0 < x_1 < \pi\end{cases},\\ c^\alpha(t,x) = {} & 0, \qquad b^\alpha(t,x) = 0, \qquad \sigma^\alpha(t,x) = \sqrt{2}\begin{pmatrix}\sin x_1\\ \sin x_2\end{pmatrix},\end{aligned}$$

and we pose Dirichlet boundary conditions. This is a monotone non-smooth problem, and we obtain order one half applying the LISL scheme and order one applying the MCSL scheme, i. e. reduced rates, see Table 3.

Δx	LISL error	rate	MCSL error	rate
3.90e-2	8.75e-3		4.19e-3	
1.96e-2	6.19e-3	0.50	2.20e-3	0.93
9.80e-3	4.38e-3	0.50	1.12e-3	0.97
4.90e-3	3.10e-3	0.50	5.69e-4	0.98
2.45e-3	2.19e-3	0.50	2.86e-4	0.99

Table 3: Results for the non-smooth linear problem at $t = 1$

10.3 Optimal control problems with smooth solutions

a) We test an example from [4] with exact solution $u(t,x_1,x_2) = \left(\frac{3}{2}-t\right)\sin x_1 \sin x_2$. The corresponding coefficients and control set in (1.1) are

$$\begin{aligned}f^\alpha = {} & \left(\frac{1}{2}-t\right)\sin x_1\sin x_2 + \left(\frac{3}{2}-t\right)\Bigg[\sqrt{\cos^2 x_1 \sin^2 x_2 + \sin^2 x_1 \cos^2 x_2}\\ & - 2\sin(x_1+x_2)\cos(x_1+x_2)\cos x_1\cos x_2\Bigg],\\ c^\alpha = {} & 0, \quad b^\alpha = \alpha, \quad \sigma^\alpha = \sqrt{2}\begin{pmatrix}\sin(x_1+x_2)\\ \cos(x_1+x_2)\end{pmatrix}, \quad \mathcal{A} = \{\alpha\in\mathbb{R}^2 : \alpha_1^2+\alpha_2^2 = 1\}.\end{aligned}$$

As σ^α does not depend on α but b^α does, we choose approximation 5.1.4 for $L^{\alpha,\beta}$ and thus need only about half of the number of interpolations we would need if we had chosen approximation 5.1.5.

b) The next test problem has exact solution $u(t,x_1,x_2) = (2-t)\sin x_1 \sin x_2$ and coefficients and control set given by

$$f^\alpha(t,x) = (1-t)\sin x_1 \sin x_2 - 2\alpha_1\alpha_2(2-t)\cos x_1 \cos x_2,$$
$$c^\alpha(t,x) = 0, \quad b^\alpha(t,x,\alpha) = 0, \quad \sigma^\alpha = \sqrt{2}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}, \quad \mathcal{A} = \{\alpha \in \mathbb{R}^2 : \alpha_1^2 + \alpha_2^2 = 1\}.$$

In both examples, the control is discretized on the unit circle by $\frac{4\pi}{h}$ grid points. The results at $t = 0.5$ are given in Table 4 , where again the grid is adapted to monotonicity. As expected for smooth solutions, the LISL scheme yields a numerical order of convergence of one, whereas the MCSL scheme yields order two.

	a)				b)			
Δx	LISL		MCSL		LISL		MCSL	
	error	rate	error	rate	error	rate	error	rate
3.93e-2	3.01e-2		8.40e-4		2.18e-2		5.14e-4	
1.96e-2	1.61e-2	0.91	2.12e-4	1.98	1.07e-2	1.03	1.29e-4	2.00
9.82e-3	8.03e-3	1.00	5.30e-5	2.00	5.45e-3	0.97	3.21e-5	2.00
4.91e-3	3.94e-3	1.03	1.33e-5	2.00	2.55e-3	1.10	8.03e-6	2.00
2.45e-3	2.03e-3	0.96	3.32e-6	2.00	1.34e-3	0.92	2.01e-6	2.00

Table 4: Results for optimal control problems at $t = 0.5$, grid adapted to monotonicity

10.4 Convergence test for a super-replication problem

We consider a test problem from [6] which was used to test convergence rates for numerical approximations of a super-replication problem from finance we will consider in Subsection 10.5. The corresponding PDE is

$$\inf_{\alpha_1^2+\alpha_2^2=1}\left\{\alpha_1^2 u_t(t,x) - \frac{1}{2}\operatorname{tr}\left(\sigma^\alpha(t,x)\sigma^{\alpha\top}(t,x)Du(t,x)\right)\right\} = f(t,x), \quad 0 \le x_1, x_2 \le 3 \tag{10.1}$$

with $\sigma^\alpha(t,x) = \begin{pmatrix}\alpha_1 x_1\sqrt{x_2}\\ \alpha_2\eta(x_2)\end{pmatrix}$ and $\eta(x) = x(3-x)$. We take $u(t,x) = 1 + t^2 - e^{-x_1^2 - x_2^2}$ as exact solution as in [6], and then f is forced to be

$$f(t,x) = \frac{1}{2}\Bigg(u_t - \frac{1}{2}x_1^2 x_2 u_{x_1x_1} - \frac{1}{2}x_2^2(3-x_2)^2 u_{x_2x_2}$$
$$-\sqrt{\left(-u_t + \frac{1}{2}x_1^2 x_2 u_{x_1x_1} - \frac{1}{2}x_2^2(3-x_2)^2 u_{x_2x_2}\right)^2 + \left(x_1\sqrt{x_2}^3(3-x_2)u_{x_1x_2}\right)^2}\Bigg).$$

In [6] $\eta(x) = x$, while we take $\eta(x) = x(3-x)$ to prevent the LISL scheme from overstepping the boundaries. Note that changing η does *not* change the solutions as long as $\eta > 0$ in the interior of the domain, see [6], and hence the above equation is equivalent to the equation used in [6]. The initial values and Dirichlet boundary values at $x_1 = 0$ and $x_2 = 0$ are taken from the exact solution. As in [6], at $x = 3$ and $y = 3$ homogeneous Neumann boundary conditions are implemented. To approximate the values of α_1, α_2, the Howard algorithm is used (see [6]), which requires an implicit time discretization, so we choose $\theta = 1$. The minimization is done over $\alpha_{1,k} + i\alpha_{2,k} = e^{2\pi ik/2N_{\Delta x}}$, $k = 1,\dots,N_{\Delta x}$, where $N_{\Delta x}$ is the number of space grid points, i. e. $N_{\Delta x} = 3/\Delta x$.

The results at $t = 1$ are given in Table 5. Again, the numerical order of convergence is approximately one when the LISL scheme is used and approximately two for the MCSL scheme. Note that compared to the results in [4], for comparable accuracies the LISL scheme is about ten times faster, the MCSL scheme about 100 to 1000 times faster.

10.5 A super-replication problem

We apply our method to solve a problem from finance, the super-replication problem under gamma constraints considered in [6]. It consists of solving equation (10.1) with $f \equiv 0$, Neumann boundary conditions and σ^α as in Subsection 10.4, and initial and Dirichlet conditions given by

$$u(t,x) = \max(0, 1-x_1), \quad t = 0 \quad or \quad x_1 = 0 \quad or \quad x_2 = 0.$$

(a) LISL

Δx	error	rate	time in s
1.50e-1	2.01e-1		0.71
7.50e-2	9.49e-2	1.08	6.76
3.75e-2	4.29e-2	1.15	75.73
1.87e-2	1.94e-2	1.15	1115.39

(b) MCSL

Δx	error	rate	time in s
3.00e-1	8.21e-2		1.17
1.50e-1	1.83e-2	2.16	11.58
7.50e-2	5.03e-3	1.86	149.24

Table 5: Results for the convergence test for the super-replication problem at $t = 1$

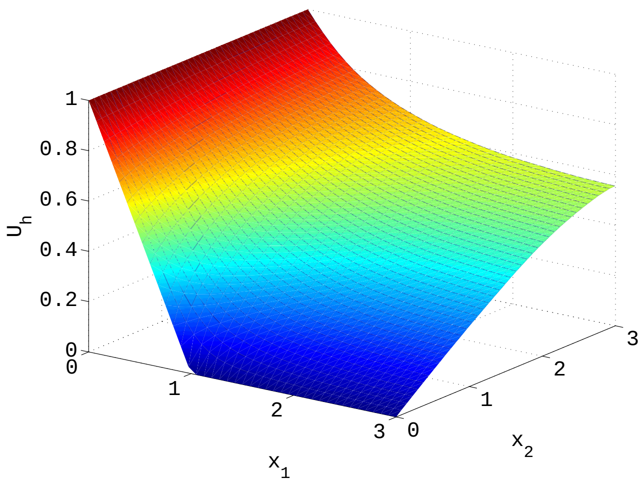

Figure 1: Numerical solution of super-replication problem at $t = 1$

The solution obtained with the LISL scheme is given in Figure 1 and coincides with the solution found in [6]. It gives the price of a put option of strike and maturity 1, and x_1 and x_2 are respectively the price of the underlying and the price of the forward variance swap on the underlying.

A Monotonicity of solutions of (1.1)

We will discuss a condition ensuring that the solution of (1.1)–(1.2) is monotone along some unit direction $e \in \mathbb{R}^N$.

(A2) Let $e \in \mathbb{R}^N$, $|e| = 1$. For all $x \in \mathbb{R}^N, \alpha \in \mathcal{A}, \beta \in \mathcal{B}, h > 0$

$$\begin{aligned} &a^{\alpha,\beta}(t, x+he) = a^{\alpha,\beta}(t,x), \quad b^{\alpha,\beta}(t, x+he) = b^{\alpha,\beta}(t,x), \\ &c^{\alpha,\beta}(t, x+he) \geq c^{\alpha,\beta}(t,x), \quad f^{\alpha,\beta}(t, x+he) \geq f^{\alpha,\beta}(t,x), \\ &g(x+he) \geq g(x). \end{aligned}$$

Lemma A.1. *Assume* (A1) *and* (A2). *If u is a viscosity solution of* (1.1)–(1.2)*, then*

$$u(t, x+he) - u(t,x) \geq 0 \qquad \textit{for all} \qquad h > 0,\ (t,x) \in \bar{Q}_T.$$

Proof. Assume that $u \geq 0$ and $c^{\alpha,\beta} \leq 0$ and let $v(t,x) = u(t, x+he)$. Since $v(t,x)$ satisfies (1.1)–(1.2) at the point $(t, x+he)$, an application of (A2) shows that it also is a supersolution at the point (t,x). By the comparison principle $u \leq v$ and the theorem is proved. In the general case consider $w = e^{-\sup_{\alpha,\beta}|c|_0 t}(u + |u|_0)$, and note that $w \geq 0$ and the corresponding $c^{\alpha,\beta}$-coefficient $c^{\alpha,\beta} - \sup_{\alpha,\beta}|c^{\alpha,\beta}|_0$ is non-positive. The first result then applies to w, and hence the theorem holds for u. □

Remark A.1. This result is not so far from optimal when $N > 1$ and the solution u is non-smooth (e.g. only Lipschitz continuous). To see that, we consider the linear case where $v = u_e := Du \cdot e$ satisfies

$$v_t = \mathrm{tr}[aD^2v] + bDv + cv + \underbrace{\mathrm{tr}[a_e D^2 u] + b_e Du + c_e u + f_e}_{\tilde{f}(t,x)} \quad \text{in} \quad Q_T$$

with e-directional derivatives a_e, b_e, c_e and f_e. If (A1) holds we can conclude from the comparison principle that

$$u_e = v \geq 0 \qquad \text{if} \qquad \tilde{f} \geq 0 \quad \text{and} \quad u_e(0,x) \geq 0.$$

If u is non-smooth, then $\tilde{f}$ is well-defined only if $a_e \equiv 0 \equiv b_e$, and the condition that $\tilde{f} \geq 0$ is essentially equivalent to assumption (A2). Of course, it is possible to relax (A2) if $N = 1$ or solutions are more regular.

Remark A.2. It is important to notice that the result of Lemma A.1 also holds for all PDEs that satisfy (A2) after (monotone) coordinate transformations. In finance there are many such equations, e. g. the Black-Scholes equation for a European option based on two stocks,

$$\begin{aligned} &u_t = \frac{1}{2}\sigma_1^2 x^2 u_{xx} + \rho\sigma_1\sigma_2 xy u_{xy} + \frac{1}{2}\sigma_2^2 y^2 u_{yy} + r(xu_x + yu_y) - ru, \quad t, x, y > 0, \\ &u(0,x,y) = \max(0, K - (x+y)), \quad x, y \geq 0. \end{aligned}$$

After the change of variables $(\bar{x},\bar{y}) = (\ln x, \ln y)$, this equation reduces to a constant coefficient equation. Since the initial data is decreasing in $\bar{x}$ and $\bar{y}$, the same is true for the solution u by Lemma A.1. Going back to (x, y) variables, we then find that u is decreasing also in x and y. (Strictly speaking we must extend $u(t,\cdot,\cdot)$ to $\mathbb{R}^2$ in a suitable way to apply Lemma A.1).

B The proof of Theorem 9.2

We will prove the result when $k_0 = 0$. The general case can be reduced to this case in a standard way by considering U/R_{k_0} and $\tilde{U}/R_{k_0}$ instead of U and $\tilde{U}$. We use doubling of variables techniques similar to those used to prove this type of results for equation (1.1). We take

$$\begin{aligned} &m_0 = |(U(0,\cdot) - \tilde{U}(0,\cdot))^+|_0, \\ &m = \sup_\alpha \Big[|(f^\alpha - \tilde{f}^\alpha)^+|_0 + (|U|_0 \wedge |\tilde{U}|_0)|c^\alpha - \tilde{c}^\alpha|_0 \Big], \\ &M^2 = 4 \sup_\alpha \big[|\sigma^\alpha - \tilde{\sigma}^\alpha|_0^2 + |b^\alpha - \tilde{b}^\alpha|_0^2 \big], \end{aligned}$$

where ϕ^+ denotes the positive part of ϕ, and define

$$\begin{aligned} &W(t,x,y) = U(t,x) - \tilde{U}(t,y), \\ &\phi(t,x,y) = m_0 + tm + \frac{1}{2\varepsilon} K_T t M^2 \\ &\qquad\qquad + \frac{1}{2} R_{k_1}(t)(L_0 + tL)(\varepsilon + \frac{1}{\varepsilon}|x-y|^2) + \delta(|x|^2 + |y|^2), \\ &\psi(t,x,y) = W(t,x,y) - \phi(t,x,y) - \eta(1+t), \\ &\tilde{m} = \sup_{\substack{t \in \Delta t\, \mathbb{N}_0 \\ x,y \in \mathbb{R}^N}} \psi(t,x,y) = \psi(\tilde{t},\tilde{x},\tilde{y}), \end{aligned}$$

for $\varepsilon, \delta, \eta > 0$ and a maximum point $(\tilde{t},\tilde{x},\tilde{y})$. A maximum point exists because of the δ-terms in ϕ. We will prove that for any sequence $\eta_l \to 0$, there is another sequence $\delta_l \to 0$ such that $\psi(\tilde{t}_l,\tilde{x}_l,\tilde{y}_l) \leq o(1)$ as $l \to \infty$. This implies Theorem 9.2 when $k_0 = 0$. To see this, fix $t > 0, x, y$ and note that for any $\varepsilon > 0$,

$$\begin{aligned} &U(t,x) - \tilde{U}(t,y) - m_0 - tm - \frac{1}{2\varepsilon} K_T t M^2 - \frac{1}{2} R_{k_1}(t)(L_0 + tL)(\varepsilon + \frac{1}{\varepsilon}|x-y|^2) \\ &\leq \psi(\tilde{t}_l,\tilde{x}_l,\tilde{y}_l) + \delta_l(|x|^2 + |y|^2) + \eta_l(1+t) \leq o(1) \quad \text{as} \quad l \to \infty. \end{aligned}$$

In this inequality we send $l \to \infty$ and choose

$$\varepsilon = |x-y| \vee t^{1/2} M$$

to find that

$$U(t,x) - \tilde{U}(t,y) \leq m_0 + tm + t^{1/2} K_T M + R_{k_1}(t)(L_0 + tL)|x-y|,$$

and hence Theorem 9.2 follows since $t > 0, x, y$ were arbitrary. We will not be explicit about the form of the δ-terms below. Their role is only to guarantee that the maximum is attained at a (finite) point $(\tilde{t},\tilde{x},\tilde{y})$, and their contribution will always be $o(1)$ as $\delta \to 0$ (see also Section 3 in [1]).

It is enough to prove that for every $\eta > 0$, $\psi(\tilde{t},\tilde{x},\tilde{y}) \leq o(1)$ as $\delta \to 0$. We proceed by contradiction assuming there is an $\eta > 0$ such that $\lim_{\delta\to 0}\psi(\tilde{t},\tilde{x},\tilde{y}) > 0$. By the definition of ψ we now have $W(\tilde{t},\tilde{x},\tilde{y}) > 0$ and $\tilde{t} > 0$ for all $\delta > 0$ small enough. The last statement is true since

$$\psi(0,\tilde{x},\tilde{y}) \leq m_0 + L_0|\tilde{x}-\tilde{y}| - m_0 - \frac{L_0}{2}(\varepsilon + \frac{1}{\varepsilon}|\tilde{x}-\tilde{y}|^2) - \eta < 0.$$

The rest of the proof will aim at getting a contradiction for the case $\tilde{t} > 0$. Even if we do not write it like that, what we show below is that $\frac{\psi(\tilde{t},\tilde{x},\tilde{y})-\psi(\tilde{t}-\Delta t,\tilde{x},\tilde{y})}{\Delta t} \leq o(1) - \eta$ as $\delta \to 0$, and this is impossible since $(\tilde{t},\tilde{x},\tilde{y})$ is a maximum point of ψ.

We proceed by defining the operator Π^α,

$$\Pi^\alpha[\phi(t,\cdot,\cdot)](r,x,y) = \sum_{i=1}^{M}\Big\{\phi(t,x+y_{k,i}^{\alpha,+}(r,x),y+\tilde{y}_{k,i}^{\alpha,+}(r,y)) \\ -2\phi(t,x,y) + \phi(t,x+y_{k,i}^{\alpha,-}(r,x),y+\tilde{y}_{k,i}^{\alpha,-}(r,y))\Big\}.$$

By the definition of L_k^α and $\tilde{L}_k^\alpha$ and (4.1), it follows that

$$\Pi^\alpha[W(t,\cdot,\cdot)](r,x,y) = 2k^2\Big\{L_k^\alpha[U(t,\cdot)](r,x) - \tilde{L}_k^\alpha[\tilde{U}(t,\cdot)](r,y)\Big\}.$$

We set $\lambda := \frac{\Delta t}{k^2}$ and subtract the inequalities defining U and $\tilde{U}$ (see (8.1) and (9.1)) to find that

$$\begin{aligned} W(t,x,y) &\leq W(t-\Delta t,x,y) \\ &+ \sup_\alpha\Big\{\frac{\lambda}{2}\Pi^\alpha[\bar{W}^\theta(t,\cdot,\cdot)](t^\theta,x,y) + \Delta t\, c^\alpha(t^\theta,x)\bar{W}^\theta(t,x,y)\Big\} \\ &+ \Delta t\, L|x-y| + \Delta t\, m \qquad \text{for } (t,x),(t,y)\in Q_T, \end{aligned}$$

where $\bar{W}^\theta(t,x,y) = (1-\theta)W(t-\Delta t,x,y) + \theta W(t,x,y)$ and $t^\theta = t-(1-\theta)\Delta t$. Note that this new "scheme" is still monotone by the definition of Π^α and the CFL condition. Hence we may replace W in the above inequality by any bigger function coinciding with W at (t,x,y). By the definition of $\tilde{m}$,

$$W \leq \phi + \eta(1+t) + \tilde{m} \quad \text{in} \quad \Delta t\,\mathbb{N}_0 \times \mathbb{R}^N \times \mathbb{R}^N,$$

and equality holds at $(\tilde{t},\tilde{x},\tilde{y})$. Therefore we find that

$$\begin{aligned} \phi(\tilde{t},\tilde{x},\tilde{y}) + \eta(1+\tilde{t}) &\leq \phi(\tilde{t}-\Delta t,\tilde{x},\tilde{y}) + \eta(1+\tilde{t}-\Delta t) \\ &+ \sup_\alpha \frac{\lambda}{2}\Pi^\alpha[\bar{\phi}^\theta(\tilde{t},\cdot,\cdot)](\tilde{t}^\theta,\tilde{x},\tilde{y}) + \Delta t\, L|\tilde{x}-\tilde{y}| + \Delta t\, m. \end{aligned} \tag{$\star$}$$

Here we also used the fact that $\Pi^\alpha[\eta(1+t)+\tilde{m}] = 0$ and $c^\alpha \leq 0$. Moreover we can Taylor expand to see that

$$\begin{aligned} \Pi^\alpha[\phi(t,\cdot,\cdot)](r,x,y) &= \sum_{i=1}^{M}\Big\{(Y_i^+ + Y_i^-)\cdot D_x\phi + (\tilde{Y}_i^+ + \tilde{Y}_i^-)\cdot D_y\phi \\ &+ \frac{1}{2}\mathrm{tr}[D_{xx}^2\phi\cdot(Y_i^+Y_i^{+\top} + Y_i^-Y_i^{-\top})] + \frac{1}{2}\mathrm{tr}[D_{yy}^2\phi\cdot(\tilde{Y}_i^+\tilde{Y}_i^{+\top} + \tilde{Y}_i^-\tilde{Y}_i^{-\top})] \\ &+ \frac{1}{2}\mathrm{tr}[D_{xy}^2\phi\cdot(Y_i^+\tilde{Y}_i^{+\top} + \tilde{Y}_i^+Y_i^{+\top} + Y_i^-\tilde{Y}_i^{-\top} + \tilde{Y}_i^-Y_i^{-\top})]\Big\}, \end{aligned}$$

where $Y_i^\pm = y_{k,i}^{\alpha,\pm}(r,x)$ and $\tilde{Y}_i^\pm = \tilde{y}_{k,i}^{\alpha,\pm}(r,y)$. Note that $YY^\top = Y\otimes Y$ for $Y\in\mathbb{R}^N$. Now we use (Y2) along with the definition of ϕ, to see that

$$\begin{aligned} \Pi^\alpha[\phi(t,\cdot,\cdot)](r,x,y) &\leq \frac{1}{\varepsilon}R_{k_1}(t)(L_0+tL)\Big\{2k^2(b^\alpha(r,x)-\tilde{b}^\alpha(r,y))(x-y) \\ &+ k^2\mathrm{tr}\Big[(\sigma^\alpha(r,x)-\tilde{\sigma}^\alpha(r,y))(\sigma^\alpha(r,x)-\tilde{\sigma}^\alpha(r,y))^\top\Big] \\ &+ k^4\mathrm{tr}\Big[(b^\alpha(r,x)-\tilde{b}^\alpha(r,y))(b^\alpha(r,x)-\tilde{b}^\alpha(r,y))^\top\Big]\Big\} + o(1), \end{aligned}$$

as $\delta \to 0$. These considerations lead to the following simplification of $(\star)$,

$$
\begin{aligned}
&\eta + \frac{\phi(\tilde{t},\tilde{x},\tilde{y}) - \phi(\tilde{t}-\Delta t,\tilde{x},\tilde{y})}{\Delta t} \\
&\le \theta \frac{1}{\varepsilon} R_{k_1}(\tilde{t})(L_0 + \tilde{t}L)(\frac{1}{2}M^2 + \frac{1}{2}k_1|\tilde{x}-\tilde{y}|^2) \\
&\quad + (1-\theta)\frac{1}{\varepsilon} R_{k_1}(\tilde{t}-\Delta t)(L_0 + (\tilde{t}-\Delta t)L)(\frac{1}{2}M^2 + \frac{1}{2}k_1|\tilde{x}-\tilde{y}|^2) \\
&\quad + L|\tilde{x}-\tilde{y}| + m + o(1) \\
&\le \frac{1}{\varepsilon} R_{k_1}(\tilde{t})(L_0 + \tilde{t}L)(\frac{1}{2}M^2 + \frac{1}{2}k_1|\tilde{x}-\tilde{y}|^2) + L|\tilde{x}-\tilde{y}| + m + o(1) := RHS,
\end{aligned}
$$

as $\delta \to 0$. Now we proceed to calculate

$$\delta_{\Delta t}\phi(t,x,y) := \frac{\phi(t,x,y) - \phi(t-\Delta t,x,y)}{\Delta t}.$$

To do that we note that

$$\delta_{\Delta t}(uv) = (\delta_{\Delta t}u)v + u\delta_{\Delta t}v - \Delta t(\delta_{\Delta t}u)(\delta_{\Delta t}v).$$

Since $\delta_{\Delta t}R_{k_1}(t) = k_1 R_{k_1}(t)$ we then see that

$$\delta_{\Delta t}[R_{k_1}(t)(L_0 + tL)] = k_1 R_{k_1}(t)(L_0 + tL) + R_{k_1}(t)L - \Delta t L k_1 R_{k_1}(t),$$

and hence

$$\delta_{\Delta t}\phi(\tilde{t},\tilde{x},\tilde{y}) = m + \frac{1}{2}K_T\frac{1}{\varepsilon}M^2 + \frac{1}{2}R_{k_1}(\tilde{t})\big[k_1(L_0 + \tilde{t}L) + L - \Delta t k_1 L\big](\varepsilon + \frac{1}{\varepsilon}|\tilde{x}-\tilde{y}|^2).$$

All of this leads to

$$\eta \le RHS - \delta_{\Delta t}\phi(\tilde{t},\tilde{x},\tilde{y}) \le o(1) \quad \text{as} \quad \delta \to 0.$$

The last inequality follows from the bound on K_T. We have our contradiction and the proof is complete.

Bibliography

[1] G. Barles and E. R. Jakobsen. On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations. *M2AN Math. Model. Numer. Anal.* 36(1):33–54, 2002.

[2] G. Barles and E. R. Jakobsen. Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations. *Math. Comp.* 76(240): 1861-1893, 2007.

[3] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.* 4(3):271–283, 1991.

[4] F. Bonnans, E. Ottenwaelter, and H. Zidani. A fast algorithm for the two dimensional HJB equation of stochastic control. *M2AN Math. Model. Numer. Anal.* 38(4):723–735, 2004.

[5] 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.

[6] B. Bruder, O. Bokanowski, S. Maroso, and H. Zidani. Numerical approximation for a superreplication problem under gamma constraints. *SIAM J. of Numer. Anal.* 47(3), 2009 (online).

[7] F. Camilli and M. Falcone. An approximation scheme for the optimal control of diffusion processes. *RAIRO Modél. Math. Anal. Numér.* 29(1): 97–122, 1995.

[8] F. Camilli and E. R. Jakobsen. A Finite Element like Scheme for Integro-Partial Differential Hamilton-Jacobi-Bellman Equations. *SIAM J. Numer. Anal.* 47(4): 2407-2431, 2009 (online).

[9] I. Capuzzo-Dolcetta. On a discrete approximation of the Hamilton-Jacobi equation of dynamic programming. *Appl. Math. Optim.* 10:367-377, 1983.

[10] R. E. Carlson and F. N. Fritsch. Monotone piecewise bicubic interpolation. *SIAM J. Numer. Anal.* 22(2):386-400, 1985.

[11] 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. (N.S.)*, 27(1):1–67, 1992.

[12] M. G. Crandall and P.-L. Lions. Convergent difference schemes for nonlinear parabolic equations and mean curvature motion. *Numer. Math.* 75 (1996), no. 1, 17–41.

[13] M. G. Crandall and P.-L. Lions. Two approximations of solutions of Hamilton-Jacobi equations. *Math. Comp.* 43(167):1–19, 1984.

[14] H. Dong and N. V. Krylov. On the rate of convergence of finte-difference approximations for Bellman equations with constant coefficients. *St. Petersburg Math. J.* 17(2): 295–313, 2006.

[15] S. C. Eisenstat, K. R. Jackson and J. W. Lewis. The order of monotone piecewise cubic interpolation. *SIAM J. Numer. Anal.* 22(6):1220-1237, 1985.

[16] M. Falcone. A numerical approach to the infinite horizon problem of deterministic control theory. *Appl. Math. Optim.* 15 (1987), no. 1, 1–13.

[17] F. N. Fritsch and R. E. Carlson. Monotone piecewise cubic interpolation. *SIAM J. Numer. Anal.* 17(2):238-246, 1980.

[18] H. J. Kushner and P. Dupuis. *Numerical methods for stochastic control problems in continuous time.* Springer-Verlag, New York, 2001.

[19] J.-L. Menaldi Some estimates for finite difference approximations. *SIAM J. Control Optim.* 27 (1989), no. 3

[20] T. S. Motzkin and W. Wasow. On the approximation of linear elliptic differential equations by difference equations with positive coefficients. *J. Math. Physics* 31, (1953). 253–259.

[21] A. M. Oberman. Convergent difference schemes for degenerate elliptic and parabolic equations: Hamilton-Jacobi equations and free boundary problems. *SIAM J. Numer. Anal.* 44 (2006), no. 2, 879–895

[22] A. Oberman. Wide stencil finite difference schemes for the elliptic Monge-Ampere equation and functions of the eigenvalues of the Hessian. *Discrete Contin. Dyn. Syst. Ser. B* 10 (2008), no. 1, 221–238.

[23] 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 (2003), no. 2, 241–267.

[24] P. J. Rasch and D. L. Williamson. On shape-preserving interpolation and semi-Lagrangian transport. *SIAM J. Sci. Statist. Comput.* 11(4):656–687, 1990.

[25] J. Young and X. Y. Zhou. *Stochastic Controls.* Spinger-Verlag, New York, 1999.