Electronic Journal of Differential Equations, Vol. 2005(2005), No. 102, pp. 1–25.
ISSN: 1072-6691. URL: http://ejde.math.txstate.edu or http://ejde.math.unt.edu
ftp ejde.math.txstate.edu (login: ftp)

RATE OF CONVERGENCE OF FINITE-DIFFERENCE APPROXIMATIONS FOR DEGENERATE LINEAR PARABOLIC EQUATIONS WITH C^1 AND C^2 COEFFICIENTS

HONGJIE DONG, NICOLAI V. KRYLOV

Abstract. We consider degenerate parabolic and elliptic equations of second order with C^1 and C^2 coefficients. Error bounds for certain types of finite-difference schemes are obtained.

1. Introduction

Numerical and, in particular, finite-difference approximations of solutions to all kinds of *linear* partial differential equations is a well established and respected area. Concerning a general approach to these issues we refer to [3], [22], [23], and [24]. By the way, in [22] it is shown, in particular, how to prove the solvability in W_2^1 by using finite-difference schemes.

One studies the convergence of numerical approximations in spaces of summable or Hölder continuous functions. Discrete L_p theory of elliptic and parabolic equations (even of order higher than 2) can be found in [4], [25], [26] with analysis of convergence in discrete Besov spaces in [27] and in the references in these papers. Discrete $C^{2+\alpha}$ spaces approach also encompassing fully nonlinear equations can be found in [10], [20], and [21].

However, in all these references the equations are assumed to be uniformly nondegenerate. In this connection note that in [5], [12] and many other related papers degenerate and even nonlinear equations are considered. But the setting in these papers is such that *linear* elliptic and parabolic equations are included only if the leading coefficients are constant.

The authors got involved into finite-difference approximations while trying to establish the rate of convergence for fully-nonlinear elliptic Bellman equations. Nondegeneracy of such equations does not help much and, therefore, we considered degenerate equations. Also many such equations like the Monge-Ampère equation or equations in obstacle problems, say arising in mathematical finance, are degenerate. Therefore, the interest in degenerate equations is quite natural. Another

2000 *Mathematics Subject Classification.* 65M15, 35J60, 93E20.
Key words and phrases. Finite-difference approximations; linear elliptic and parabolic equations.

Submitted August 23, 2005. Published September 21, 2005.
N.V. Krylov was partially supported by grant DMS-0140405 from the NSF.

point is that for fully nonlinear degenerate Bellman equations the higher smoothness of the "coefficients" generally does not help get better smoothness of the true solutions. This is the reason why we concentrate mainly on equations with C^1 and C^2 coefficients.

The activity related to numerical approximations for fully nonlinear second order degenerate equations started few years ago with [17], [16], [15], and then was continued in [1], [2], [7], [11], and [13]. In many of these papers the idea is used that the approximating finite-difference equation and the original one should play symmetric roles. This led to somewhat restricted results since no estimates of smoothness of solutions of finite-difference equations were available. Nevertheless, restricted or not, even now we cannot believe that, for the moment, these are *the only* published results on the rate of convergence in the sup norm of finite-difference approximations even if the Bellman equation becomes a *linear* second order degenerate equation (variety of results for nondegenerate case can be found in [3], [23], [24] and references therein). One also has to notice that there is vast literature about other types of numerical approximations for linear degenerate equations such as Galerkin or finite-element approximations (see, for instance, [22], [23], and [24]). It is also worth noting that under variety of conditions the first *sharp* estimates for finite-difference approximations in linear one-dimensional degenerate case are proved in [28].

To explain our main idea note that linear even degenerate equations often possess smooth solutions, which one can substitute into the finite-difference scheme and then use, say the maximum principle to estimate the difference between the true solution and the approximation. We use precisely this quite standard and well-known method (see, for instance, [3], [23], [24]) giving up on the symmetry between the original and approximating equations.

To be able to apply this method one needs the true solution to have *four* spatial derivatives, which hardly often happens in fully nonlinear even uniformly nondegenerate case. But in the linear case, if the coefficients are not smooth enough, one can mollify them and get smooth solutions. However, then the idea described above would only lead to estimates for discretized equation with *mollified* coefficients. Therefore, the main problem becomes estimating the difference between the solutions of the initial finite-difference equation and the finite-difference equation constructed from mollified coefficients. We reduce this problem to estimating the Lipschitz constant of solutions to finite-difference equations and state the central result of this paper as Theorem 4.1.

In connection with the smoothness of solutions of finite-difference equations, we note that only recently in [13] the first result appeared for fully nonlinear elliptic and parabolic degenerate equations. Of course, the results of [13] are also valid for linear equations. However, the exposition in [13] is aimed at fully nonlinear equations and has many twists and turns which are not needed in the linear case. Understandably, it is desirable to use only "linear" methods while developing the theory of linear equations rather than appeal to a quite technical and much harder theory of fully nonlinear equations. Therefore, we decided to write the proofs for the linear case in the present article. We certainly hope that the methods developed here will be useful in other issues of the theory of linear equations. Restricting ourselves to the linear case also allows us to get sharper results and better rates of convergence for smoother (C^2 and C^4) coefficients. In connection with this restriction it is worth noting a peculiar issue. In a subsequent article we plan to to treat fully nonlinear equations in *domains*, and we are not able to make this treatment any easier if the equation is actually *linear*.

One of our results (Theorem 2.13) bears on the case of *Lipschitz* continuous coefficients and data and yields the *sup-norm* rate of convergence $h^{1/2}$, where h is the mesh size. Remark 2.20 shows that this result is sharp even for equations with constant coefficients. Although the proof of Theorem 2.13 based on Theorem 4.1 is new, its statement can be found in [11], and, actually, follows from Corollary 2.3 of [15]. However, this way of proving Theorem 2.13 uses the theory of fully nonlinear equations and *does not* allow to get better rates of convergence if the coefficients of the equation are smoother. In particular, in contrast with using Theorem 4.1, it will not lead to our results about the rates h and h^2 for linear equations with variable coefficients.

The article is organized as follows. Our main results, Theorems 2.12-2.19, are stated in Section 2 and proved in Section 5. Between these two sections we prove few auxiliary results the most important of which is Theorem 4.1. One of the auxiliary results, Lemma 3.3, is proved in Section 6. Section 7 contains a discussion of semidiscretization when only spatial derivatives are replaced with finite-differences and the final Section 8 contains some comments on possible extensions of our results.

To conclude the introduction, we set up some notation: $\mathbb{R}^d$ is a d-dimensional Euclidean space with $x = (x^1, x^2, \dots, x^d)$ to be a typical point in $\mathbb{R}^d$. As usual the summation convention over repeated indices is enforced unless specifically stated otherwise. For any $l = (l^1, l^2, \dots, l^d) \in \mathbb{R}^d$ and any differentiable function u on $\mathbb{R}^d$, we denote $D_l u = u_{x^i} l^i$ and $D_l^2 u = u_{x^i x^j} l^i l^j$, etc. By $D_t u$, $D_t^2 u, \dots$ we denote the derivatives of $u = u(t, x)$ in t, $D_x^j u$ is its generic derivative of order j in x.

We use the Hölder spaces $C^{1/2,1}$, $C^{1,2}$, $C^{2,4}, \dots$ of functions of $(t, x) \in \mathbb{R} \times \mathbb{R}^d$ defined in some subdomains of $\mathbb{R} \times \mathbb{R}^d$. More specifically $C^{1/2,1}$ is the space of bounded functions having finite Hölder constant of order $1/2$ in t and continuously differentiable in x with the derivatives being bounded; $C^{k,2k}$ is the space of functions having k derivatives in t and $2k$ derivatives in x, the functions themselves and their said derivatives are assumed to be bounded and continuous. These spaces are provided with natural norms: we use the notation $|\cdot|_{k,2k}$ in the case of functions given in $\mathbb{R} \times \mathbb{R}^d$ and $|\cdot|_{H,k,2k}$ in the case of functions given in $H \subset \mathbb{R} \times \mathbb{R}^d$.

Various constants are denoted by N in general and the expression $N = N(\dots)$ means (and means only) that the given constant N depends only on the contents of the parentheses.

2. The Setting and Main Results

Let $d_1, d \geq 1$ be integers, ℓ_k, $k = \pm 1, \dots, \pm d_1$ nonzero vectors in $\mathbb{R}^d$ and $\ell_k = -\ell_{-k}$. Suppose that we are given continuous real-valued functions $c(t, x)$, $f(t, x)$, $g(x)$, $\sigma_k(t, x)$, $b_k(t, x) \geq 0$, $k = \pm 1 \dots, \pm d_1$ satisfying

$$\sigma_k = \sigma_{-k}.$$

Introduce functions $\sigma(t, x)$, $a(t, x)$, and $b(t, x)$ taking values in the set of $d \times 2d_1$ and $d \times d$ matrices and $\mathbb{R}^d$, respectively, by

$$\sigma^{ik}(t, x) = \ell_k^i \sigma_k(t, x), \quad \sigma(t, x) = (\sigma^{ik}(t, x)),$$
$$a = (1/2)\sigma\sigma^*, \quad b(t, x) = \ell_r b_r(t, x)$$

with no summation with respect to k.

Assumption 2.1. For an integer $n \in \{1,2,4,6,\dots\}$ and some numbers $K_n \geq K_0 \geq 1, \lambda \geq 0$ we have

$$\sum_{|k|=1}^{d_1}(|\ell_k|+|\sigma_k|_0^2+|b_k|_0)+|c|_0+|f|_0+|g|_0 \leq K_0,$$

$$\sum_{|k|=1}^{d_1}(|\sigma_k|_{n/2,n}^2+|b_k|_{n/2,n})+|c|_{n/2,n}+|f|_{n/2,n}+|g|_n \leq K_n,$$

$$c(t,x) \geq \lambda.$$

Denote

$$L^0u(t,x)=a^{ij}(t,x)u_{x^ix^j}(t,x)+b^i(t,x)u_{x^i}(t,x),$$
$$Lu(t,x)=L^0u(t,x)-c(t,x)u(t,x).$$

Note that, for $a_k(t,x):=(1/2)|\sigma_k(t,x)|^2$, we have

$$a^{ij}(t,x)u_{x^ix^j}=a_k(t,x)D_{\ell_k}^2u.$$

Let $T \geq 0$ be a constant. We are interested in the following parabolic equation:

$$\frac{\partial}{\partial t}u(t,x)+Lu(t,x)+f(t,x)=0, \tag{2.1}$$

in $H_T:=[0,T)\times\mathbb{R}^d$ with terminal condition

$$u(T,x)=g(x), \quad x\in\mathbb{R}^d. \tag{2.2}$$

We know (see, for instance, [8]) that under the above conditions there is a unique bounded viscosity solution v of (2.1)-(2.2), which coincides with the probabilistic one given by

$$\begin{aligned} v(t,x)&=Eg(x_T)\exp(-\int_t^T c(s,x_s)\,ds)\\ &\quad+E\int_t^T f(s,x_s)\exp(-\int_t^s c(r,x_r)\,dr)\,ds, \end{aligned} \tag{2.3}$$

where $x_s=x_s(t,x)$ is defined as a solution of

$$x_s=x+\int_t^s\sigma_k(r,x_r)\ell_k\,dw_r^k+\int_t^s b_k(r,x_r)\ell_k\,dr, \quad s\geq t, \tag{2.4}$$

and w_r is a $2d_1$-dimensional Wiener process defined for $r \geq t$. Due to Assumption 2.1, we have

$$|v| \leq K_0(1-e^{-\lambda T})/\lambda+K_0e^{-\lambda T} \leq K_0(1+T\wedge\lambda^{-1}),$$

with natural interpretation if $\lambda=0$.

We use the following finite-difference approximations. For every $h>0$, $\tau>0$, $l\in\mathbb{R}^d$ and $(t,x)\in[0,T)\times\mathbb{R}^d$, introduce:

$$\delta_{h,l}u(t,x)=h^{-1}(u(t,x+hl)-u(t,x)), \quad \Delta_{h,l}=-\delta_{h,l}\delta_{h,-l},$$
$$\delta_\tau u(t,x)=\tau^{-1}(u(t+\tau,x)-u(t,x)),$$
$$\delta_\tau^T u(t,x)=\tau^{-1}(u(t+\tau_T(t),x)-u(t,x)), \quad \tau_T(t)=\tau\wedge(T-t).$$

Note that the first factor of $\delta^T_\tau u$ is τ^{-1} and not $(\tau_T(t))^{-1}$. Also note that

$$t + \tau_T(t) = (t+\tau) \wedge T,$$

so that to evaluate $\delta^T_\tau u(t,x)$ in H_T we only need to know the values of u in $\bar H_T$.

Let $\mathcal{B} = \mathcal{B}(\bar H_T)$ be the set of all bounded functions on $\bar H_T$. For every $h > 0$, we introduce two bounded linear operators L^0_h and $L_h : \mathcal{B} \to \mathcal{B}$:

$$L^0_h u = a_k(t,x)\Delta_{h,\ell_k} u + b_k(t,x)\delta_{h,\ell_k} u, \quad L_h u = L^0_h u - c(t,x)u. \tag{2.5}$$

The finite-difference approximations of v which we have in mind will be introduced by means of the equation

$$\delta^T_\tau u(t,x) + L_h u(t,x) + f(t,x) = 0, \ \ (t,x) \in H_T, \tag{2.6}$$

with terminal condition (2.2).

Remark 2.2. One may think that considering the operators L written in the form $a_k D^2_{\ell_k} + b_k D_{\ell_k} + c$ is a severe restriction. In this connection recall that according to the Motzkin-Wasov theorem any uniformly nondegenerate operator with bounded coefficients admits such representation. It is also easy to see (cf. [7]) that if we fix a finite subset $B \subset \mathbb{Z}^d$, such that $\operatorname{Span} B = \mathbb{R}^d$, and if an operator

$$Lu = a^{ij} u_{x^i x^j} + b^i u_{x^i} \tag{2.7}$$

admits a finite-difference approximation

$$L_h u(0) = \sum_{y \in B} p_h(y) u(hy) \to Lu(0) \quad \forall u \in C^2$$

and L_h are monotone, then automatically

$$L = \sum_{l \in B,\, l \neq 0} a_l D^2_l + \sum_{l \in B,\, l \neq 0} b_l D_l \tag{2.8}$$

for some $a_l \geq 0$ and $b_l \in \mathbb{R}$.

Problem (2.6)-(2.2) is actually a collection of disjoint problems given on each mesh associated with points $(t_0, x_0) \in [0,T) \times \mathbb{R}^d$:

$$\begin{gathered}\{\big((t_0 + j\tau) \wedge T, x_0 + h(i_1\ell_1 + \dots + i_{d_1}\ell_{d_1})\big) \ : \\ j = 0,1,2,\dots,\ i_k = 0, \pm 1, \pm 2, \dots,\ k = 1,2,\dots,d_1\}.\end{gathered}$$

For fixed $\tau, h > 0$ introduce

$$\begin{gathered}\bar{\mathcal{M}}_T = \{(t,x) \ : \ t = (j\tau) \wedge T, x = h(i_1\ell_1 + \dots + i_{d_1}\ell_{d_1}), \\ j = 0,1,2,\dots,\ i_k = 0, \pm 1, \pm 2, \dots,\ k = 1,2,\dots,d_1\}.\end{gathered}$$

Results obtained for equations on a subset of $\bar{\mathcal{M}}_T$ can be certainly translated into the corresponding results for all other meshes by shifts of the origin. Another important observation is that $\bar{\mathcal{M}}_T$ may lie in a subspace of $\mathbb{R}^{d+1}$.

Note a straightforward property of the above objects.

Lemma 2.3. *For any $h, \tau > 0$ we have $(\delta^T_\tau + L^0_h)1 = 0$. Furthermore, denote $p_{\tau,h} = K_0(h^{-2} + h^{-1} + \tau^{-1})$. Then for any $h, \tau > 0$ the operator*

$$u \to \delta^T_\tau u + L^0_h u + p_{\tau,h} u$$

is monotone, by which we mean that if $u_1, u_2 \in \mathcal{B}$ and $u_1 \leq u_2$, then

$$\delta^T_\tau u_1 + L^0_h u_1 + p_{\tau,h} u_1 \leq \delta^T_\tau u_2 + L^0_h u_2 + p_{\tau,h} u_2.$$

Let T' be the least point in the progression $\tau, 2\tau, \dots$, which is greater than or equal to T. Based on Lemma 2.3 and the contraction mapping theorem, we have the following several lemmas and corollaries (see [13]). The first one gives the existence and uniqueness of solutions to (2.6). The second one plays the role of comparison principle for finite-difference schemes.

Lemma 2.4. *Take a nonempty set*

$$Q \subset \mathcal{M}_T := \bar{\mathcal{M}}_T \cap H_T.$$

Let $\phi(t,x)$ be a bounded function on $\bar{\mathcal{M}}_T$. Then there is a unique bounded function u defined on $\bar{\mathcal{M}}_T$ such that equation (2.6) *holds in Q and $u = \phi$ on $\bar{\mathcal{M}}_T \setminus Q$.*

Lemma 2.5. *Let u_1, u_2 be functions on $\bar{\mathcal{M}}_T$ and $f_1(t,x)$,$f_2(t,x)$ functions on $\mathcal{M}_T$. Assume that in Q*

$$\delta_\tau^T u_1(t,x) + L_h u_1(t,x) + f_1(t,x) \ge \delta_\tau^T u_2(t,x) + L_h u_2(t,x) + f_2(t,x).$$

Let $h \le 1$ and $u_1 \le u_2$ on $\bar{\mathcal{M}}_T \setminus Q$ and assume that $u_i e^{-\mu|x|}$ are bounded on $\mathcal{M}_T$, where $\mu \ge 0$ is a constant. Then there exists a constant $\tau^* > 0$, depending only on K, d_1, and μ, such that if $\tau \in (0,\tau^*)$ then on $\bar{\mathcal{M}}_T$

$$u_1 \le u_2 + T' \sup_Q (f_1 - f_2)_+, \tag{2.9}$$

and, if in addition $\lambda \ge 1$, we have

$$u_1 \le u_2 + \sup_Q (f_1 - f_2)_+. \tag{2.10}$$

Furthermore, $\tau^*(K, d_1, \mu) \to \infty$ as $\mu \downarrow 0$ and if u_1, u_2 are bounded on $\bar{\mathcal{M}}_T$, so that $\mu = 0$, then (2.9), (2.10) *hold without any constraints on h and τ.*

Remark 2.6. In the sense of viscosity solutions, Lemma 2.5 is also well known to be true with the differential operator L in place of L_h.

Corollary 2.7. *Let $c_0 \ge 0$ be a constant such that*

$$\tau^{-1}(e^{c_0\tau} - 1) \le \lambda.$$

Then

$$|v_{\tau,h}(t,x)| \le K_0 \lambda^{-1}(1 - e^{-\lambda(T+\tau)}) + e^{-c_0(T-t)}|g|_0$$

on $\bar{H}_T$ with natural interpretation of this estimate if $c_0 = \lambda = 0$, that is

$$|v_{\tau,h}| \le K_0(T+\tau) + |g|_0.$$

Corollary 2.8. *Let u_1 and u_2 be bounded solutions of* (2.6) *in H_T with terminal condition $g_1(x)$ and $g_2(x)$, where g_1 and g_2 are given bounded functions. Then under the condition of Corollary 2.7, in $\bar{H}_T$ we have*

$$u_1(t,x) \le u_2(t,x) + e^{-c_0(T-t)} \sup (g_1 - g_2)_+.$$

Corollary 2.9. *Assume that there is a constant R such that $f(t,x) = g(x) = 0$ for $|x| \ge R$. Then*

$$\lim_{|x|\to\infty} \sup_{[0,T]} |v_{\tau,h}(t,x)| = 0.$$

Lemma 2.10. *Let function f_n and g_n, $n = 1, 2, \dots$, satisfy the same conditions as f, g with the same constants and let $v^n_{\tau,h}$ be the unique bounded solutions of problem (2.6)-(2.2) with f_n and g_n in place of f and g, respectively. Assume that on $\bar{H}_T$*

$$\lim_{n\to\infty}(|f - f_n| + |g - g_n|) = 0.$$

Then pointwisely $v^n_{\tau,h} \to v_{\tau,h}$ on $\bar{H}_T$.

Remark 2.11. In many cases Lemma 2.10 allows us to concentrate only on compactly supported f and g.

Here come the main results of this article.

Theorem 2.12. *Under Assumption 2.1 with $n = 1$, there is a constant N_1, depending only on d, d_1, T and K_1 (but not on h and τ) such that*

$$|v - v_{\tau,h}| \le N_1(\tau^{1/4} + h^{1/2})$$

in H_T. In addition, there exists a constant N_2 depending only on d, d_1, and K_1, such that if $\lambda \ge N_2$, then N_1 is independent of T.

Theorem 2.13. *Under the assumption of Theorem 2.12 suppose that σ, b, c, f are independent of t and $\lambda \ge N_2$, where N_2 is taken form Theorem 2.12. Let $\tilde{v}(x)$ be a probabilistic or the unique bounded viscosity solution of*

$$Lu(x) + f(x) = 0$$

in $\mathbb{R}^d$. Let $\tilde{v}_h$ be the unique bounded solution of

$$L_h u(x) + f(x) = 0 \tag{2.11}$$

in $\mathbb{R}^d$. Then

$$|\tilde{v} - \tilde{v}_h| \le N h^{1/2}$$

in $\mathbb{R}^d$, where N depends only on d, d_1, and K_1.

Theorem 2.14. *Under Assumption 2.1 with $n = 2$, there is a constant N_3, depending only on d, d_1, T and K_2 (but not on h and τ) such that*

$$|v - v_{\tau,h}| \le N_3(\tau^{1/2} + h)$$

in H_T. In addition, there exists a constant N_4 depending only on d, d_1, and K_1, such that if $\lambda \ge N_4$, then N_3 is independent of T.

Theorem 2.15. *Under the assumption of Theorem 2.14 suppose that σ, b, c, f are independent of t and $\lambda \ge N_4$, where N_4 is taken form Theorem 2.14. Then*

$$|\tilde{v} - \tilde{v}_h| \le N h$$

in $\mathbb{R}^d$, where N depends only on d, d_1, and K_2.

Theorem 2.16. *Under Assumption 2.1 with $n = 4$, there is a constant N_5, depending only on d, d_1, T and K_4 (but not on h and τ) such that*

$$|v - v_{\tau,h}| \le N_5(\tau + h) \tag{2.12}$$

in H_T. In addition, there exists a constant N_6 depending only on d, d_1, and K_1, such that if $\lambda \ge N_6$, then N_5 is independent of T.

In the case of $n = 4$ we can get an even better estimate with a different scheme. To state the result we need to introduce one more assumption and somewhat change our notation.

Assumption 2.17. We have $|b_k(t,x)| \le Ka_k(t,x)$ for all $k = \pm 1, \dots, \pm d_1$.

This time we again introduce L_h as in (2.5) but use a different formula for L_h^0:

$$L_h^0 u(t,x) = a_k(t,x)\Delta_{h,\ell_k} u(t,x) + b_k(t,x)\delta_{2h,\ell_k} u(t, x - h\ell_k). \tag{2.13}$$

As above Lemmas 2.3, 2.4 and 2.5 hold true but this time only if

$$Kh \le 2. \tag{2.14}$$

Of course, in Lemma 2.4 and in Theorems 2.14 and 2.19 below by $v_{\tau,h}$ and $\tilde{v}_h$ we mean the unique bounded solutions of the corresponding equations with new operators L_h and we assume (2.14).

Theorem 2.18. *Under Assumption 2.1 with $n = 4$ and Assumption 2.17, there is a constant N_7, depending only on d, d_1, T and K_4 (but not on h and τ) such that*

$$|v - v_{\tau,h}| \le N_7(\tau + h^2)$$

in H_T. In addition, there exists a constant N_8 depending only on d, d_1, and K_1, such that if $\lambda \ge N_8$, then N_7 is independent of T.

Theorem 2.19. *Under the assumptions of Theorem 2.18 suppose that σ, b, c, f are independent of t and $\lambda \ge N_8$, where N_8 is taken form Theorem 2.18. Then*

$$|\tilde{v} - \tilde{v}_h| \le Nh^2$$

in $\mathbb{R}^d$, where N depends only on d, d_1, and K_4.

Remark 2.20. The rate in Theorems 2.12 and 2.13 is sharp at least in what concerns h. The reader will see that Theorem 2.13 is derived from Theorem 2.12 in such a way that if one could improve the rate in Theorem 2.12, then the same would happen with Theorem 2.13 if λ is large enough. So to prove the sharpness we may only concentrate on the time independent case.

Take $d = 2$ and consider the equation

$$v_x + v_y - \lambda v = -g(|x - y|)$$

in $\mathbb{R}^2 = \{(x,y)\}$, where $g(t) = |t| \wedge 1$ and $\lambda > 0$.

Then $v(x,y) = \lambda^{-1} g(|x-y|)$ and $v(0,0) = 0$. It is not hard to show that for $\ell_1 = (1,0)$ and $\ell_2 = (0,1)$, we have

$$v_h(x,y) = u_h(x-y), \quad u_h(x) = \frac{h}{2+\lambda h} \sum_{n=0}^{\infty} \big(\frac{2}{2+\lambda h}\big)^n Eg(x + h\xi_n),$$

where

$$\xi_n = \sum_{i=1}^{n} \eta_i$$

and $\eta_1, \eta_2, \dots$ are independent random variables such that

$$P(\eta_i = \pm 1) = \frac{1}{2}.$$

The sequence $w_n := \xi_n/\sqrt{n}$ is asymptotically normal with zero mean and variance 1. Therefore, for all big n and small h such that $h^{-2} \ge n$ we have $(h\sqrt{n})^{-1} \ge 1$ and

$$Eg(h\xi_n) = E((h|\xi_n|) \wedge 1) = h\sqrt{n} E[|w_n| \wedge (h\sqrt{n})^{-1}] \ge \gamma h\sqrt{n}$$

with constant $\gamma > 0$ independent of n, h.

It follows that for all small h

$$
\begin{aligned}
u_h(0) &\geq \gamma \frac{h^2}{2+\lambda h} \sum_{h^{-1}\leq n\leq h^{-2}} \big(\frac{2}{2+\lambda h}\big)^n \sqrt{n} \\
&\geq \gamma \frac{h^{3/2}}{2+\lambda h} \sum_{h^{-1}\leq n\leq h^{-2}} \big(\frac{2}{2+\lambda h}\big)^n \\
&\geq \gamma\lambda^{-1} h^{1/2} I(h) J(h),
\end{aligned}
$$

where

$$
I(h) = \big(\frac{2}{2+\lambda h}\big)^{2/h}, \quad J(h) = 1 - \big(\frac{2}{2+\lambda h}\big)^{1/(2h^2)}.
$$

This shows that the rate under discussion is sharp indeed since $I(h) \to e^{-\lambda}$ and $J(h) \to 1$ as $h \downarrow 0$.

Remark 2.21. The estimates in Theorem 2.16 and 2.18 are sharp even for $d = 1$. An example is the following parabolic equation

$$
v_t + v_{xx} + v_x = 0, \quad x \in \mathbb{R} \times [0, 1),
$$

with periodic terminal condition $g(x) = \sin(2\pi x)$. By a standard Fourier method, we know that for the scheme in Theorem 2.16 ($L_2([0, 1])$ is L_2 in x),

$$
\|v(1, \cdot) - v_{\tau,h}(1, \cdot)\|_{L_2([0,1])} = O(\tau + h).
$$

And for the scheme in Theorem 2.18 with symmetric first-order differences, we have

$$
\|v(1, \cdot) - v_{\tau,h}(1, \cdot)\|_{L_2([0,1])} = O(\tau + h^2).
$$

Thus, with the sup norm the errors are at least $O(\tau+h)$ and $O(\tau+h^2)$ respectively.

3. Smoothness of solutions to (2.1)

In this section, we state some known results about the smoothness of solutions to degenerate parabolic equations. The following two lemmas can be found in [18] and [16] or else in Chapter V of [19].

Lemma 3.1. *Under Assumption 2.1 with $n = 1$, the solution v of (2.1) given by (2.3) is bounded and continuous on $\bar{H}_T$ and differentiable with respect to x continuously in (t, x). Moreover, there exist constants M, N depending only on K_1, d, and d_1, such that,*

$$
|v(t, y) - v(s, x)| \leq N e^{(M-\lambda)_+ T} (|t - s|^{1/2} + |y - x|),
$$

for all $(t, y), (s, x) \in \bar{H}_T$ with $|t - s| \leq 1$.

Lemma 3.2. *Under Assumption 2.1 with $n = 2$, the solution v is twice continuously in (t, x) differentiable with respect to x and continuously in (t, x) differentiable with respect to t. Moreover, there exist constants N depending only on K_2, d and d_1, and M depending only on K_1, d, and d_1 such that*

$$
|v|_{\bar{H}_T,1,2} \leq N e^{(M-\lambda)_+ T}
$$

The proofs in [19] of the above lemmas are given by using probabilistic methods and moment estimates. By the same methods, we can get the following general result (cf. [19], [6]). For the sake of completeness, we give a proof of it in Section 6.

Lemma 3.3. (i) *Let $m \geq 1$ be an integer. Under Assumption 2.1 with $n = 2m$, the solution* (2.3) *is $2m$ times continuously differentiable with respect to x on $\bar{H}_T$ and m times continuously differentiable with respect to t on $\bar{H}_T$. Moreover, there exist constants N depending only on m, K_{2m}, d, and d_1, and M depending only on m, K_1, d, and d_1, such that*

$$|v|_{\bar{H}_T,m,2m} \leq Ne^{(M-\lambda)_+T}.$$

(ii) *If there is a constant N_0 and integer $l \geq 1$ such that*

$$K_j \leq N_0/\varepsilon^{(j-l)_+}, \quad j \leq 2m,$$

for some positive number $\varepsilon \leq 1$, then we have

$$|v|_{\bar{H}_T,m,2m} \leq N(N_0, d_1, d, m)e^{(M-\lambda)_+T}/\varepsilon^{(2m-l)_+}.$$

In what follows we will only use Lemma 3.3 for $m = 1, 2$ and $l = 1, 2$. The next theorem is about continuous dependence of solutions with respect to the coefficients and the terminal conditions.

Theorem 3.4. *Let σ_k, b_k, c, λ, f, g, $\hat{\sigma}_k$, $\hat{b}_k$, $\hat{c}$, $\hat{\lambda}$, $\hat{f}$, $\hat{g}$ satisfy Assumption 2.1 with $n = 1$ and $\hat{\lambda} = \lambda$. Let v and $\hat{v}$ be the corresponding solutions of* (2.1)-(2.2). *Assume*

$$\gamma := \sup_{H_T,k} \big(|\sigma_k - \hat{\sigma}_k| + |b_k - \hat{b}_k| + |c - \hat{c}| + |f - \hat{f}|\big) + \sup_x |g - \hat{g}| < \infty.$$

Then there are constants N and M depending only on K_1, d, and d_1 such that on $\bar{H}_T$

$$|v - \hat{v}| \leq N\gamma e^{(M-\lambda)_+T}$$

Proof. Consider $\mathbb{R}^d$ as a subspace of

$$\mathbb{R}^{d+1} = \{x = (x', x^{d+1}) : x' \in \mathbb{R}^d, x^{d+1} \in \mathbb{R}\}.$$

Introduce

$$\bar{H}_T(d+1) = \{(t, x', x^{d+1}) : (t, x') \in \bar{H}_T, x^{d+1} \in \mathbb{R}\},$$
$$H_T(d+1) = \{(t, x) \in \bar{H}_T(d+1) : t < T\}.$$

Let $\eta \in \mathcal{C}^1_b(\mathbb{R})$ be a function such that

$$\eta(-1) = 1,\ \eta(0) = 0,\ \eta'(p) = \eta'(q) = 0 \quad \text{for} \quad p \leq -1, q \geq 0.$$

Set

$$\tilde{\sigma_k}(t, x) = \hat{\sigma_k}(t, x')\eta(x^{d+1}/\gamma) + \sigma_k(t, x')(1 - \eta(x^{d+1}/\gamma))$$

and similarly introduce $\tilde{b}_k$, $\tilde{c}$, $\tilde{f}$, and $\tilde{g}$.

It is easy to check that Lemma 3.1 is applicable to our new objects, and we denote $\tilde{v}$ to be the solution of (2.1)-(2.2) with $\tilde{\sigma}_k$, $\tilde{b}_k$, $\tilde{c}$, $\tilde{f}$, $\tilde{g}$, $\tilde{f}$ in place of σ_k, b_k, c, f and g. Note this is actually a collection of disjoint problems parameterized by x^{d+1}. By the uniqueness of solutions, obviously for any $(t, x') \in \bar{H}_T$

$$\tilde{v}(t, x', -\gamma) = \bar{v}(t, x'), \quad \tilde{v}(t, x', 0) = v(t, x').$$

Therefore, the assertion of the theorem follows from Lemma 3.1. □

4. Some estimates for solutions to linear finite difference equations

In this section, we give several results about Lipschitz continuity of solution $u = v_{\tau,h}$ to linear finite difference equation. Firstly, observe that

$$\begin{aligned} \Delta_{h,l}(v^2) &= 2v\Delta_{h,l}v + (\delta_{h,l}v)^2 + (\delta_{h,-l}v)^2, \\ \delta_{h,l}(uv) &= u\delta_{h,l}v + v\delta_{h,l}u + h\delta_{h,l}u\delta_{h,l}v. \end{aligned} \tag{4.1}$$

We fix an $\varepsilon \in (0, h]$ and a unit vector $l \in \mathbb{R}^d$ and introduce

$$\bar{\mathcal{M}}_T(\varepsilon) := \{(t, x + i\varepsilon l) : (t, x) \in \bar{\mathcal{M}}_T, i = 0, \pm 1, \dots\}.$$

Let $Q \subset \bar{\mathcal{M}}_T(\varepsilon)$ be a nonempty finite set and u a function on $\bar{\mathcal{M}}_T(\varepsilon)$ satisfying (2.6) in $Q' = Q \cap H_T$.

We add two more directions $\ell_{d_1+1} := l$ and $\ell_{-d_1-1} := -l$ and let r be an index running through $\{\pm 1, \dots, \pm(d_1+1)\}$ and k through $\{\pm 1, \dots, \pm d_1\}$. Denote

$$h_k = h,\ k = \pm 1, \dots, \pm d_1,\ h_{\pm(d_1+1)} = \varepsilon.$$

Define the interior and boundary of Q:

$$Q^o_\varepsilon = \{(t, x) \in Q' : (t + \tau_T(t), x), (t, x \pm h_k \ell_k) \in Q, \forall k = 1, 2, \dots, d_1 + 1\}.$$
$$\partial_\varepsilon Q = Q \setminus Q^o_\varepsilon.$$

Theorem 4.1. *Under Assumption 2.1 with $n = 1$, suppose that there are constants $N_0, c_0 \ge 0$, $\gamma > 0$ such that*

$$K_1^2[(2d_1 + 15)K_1 + 6] \le -\gamma + \lambda + \tau^{-1}(1 - e^{-c_0\tau}) + N_0 \inf_{Q^0_\varepsilon, k} a_k, \tag{4.2}$$

which always holds with $N_0 = c_0 = 0$ if

$$K_1^2[(2d_1 + 15)K_1 + 6] \le \lambda - \gamma.$$

Then there is a constant $N \in (0, +\infty)$ depending only on N_0, K_1, d_1, d, and γ, such that on Q

$$|\delta_{\varepsilon,\pm l}u| \le Ne^{c_0(T+\tau)}\big(1 + \max_Q |u| + \max_{\partial_\varepsilon Q}(\max_r |\delta_{h_r,\ell_r}u|)\big).$$

Proof. Set $\xi(t) = e^{c_0 t}$ for $t < T$ and $\xi(T) = e^{c_0 T'}$, $w = \xi u$, $w_r = \delta_{h_r,\ell_r}w$, $w_\tau = \delta^T_\tau w$. Denote

$$W = \sum_r w_r^2, \quad M = N_0 + 1$$

and let (t_0, x_0) be a point at which

$$V := W + Mw^2$$

takes its maximum value on Q. It is easy to see that

$$\max_{Q,r} |w_r| \le V^{1/2}(t_0, x_0), \quad |\delta_{\varepsilon,\pm l}u| \le V^{1/2}(t_0, x_0) \tag{4.3}$$

on Q and we need only estimate $V(t_0, x_0)$. Furthermore, obviously

$$\begin{aligned} V^{1/2}(t, x) &\le 2(d_1 + 1)\max_r |w_r(t, x)| + \sqrt{M}|w(t, x)| \\ &\le e^{c_0(T+\tau)}[2(d_1 + 1)\max_r |\delta_{h_r,\ell_r}u(t, x)| + \sqrt{M}|u(t, x)|], \end{aligned}$$

so while estimating $V(t_0,x_0)$ we may assume that $(t_0,x_0)\in Q^0_\varepsilon$. Then for each $k=\pm1,\pm2,\dots,\pm d_1$ at (t_0,x_0) we have

$$0\geq \delta_{h,\ell_k}V=2\sum_r w_r\delta_{h,\ell_k}w_r+2Mww_k+h\sum_r(\delta_{h,\ell_k}w_r)^2+Mhw_k^2, \tag{4.4}$$

$$\begin{aligned}0\geq \Delta_{h,\ell_k}V&=2\sum_r w_r\Delta_{h,\ell_k}w_r+2Mw\Delta_{h,\ell_k}w\\&\quad+\sum_r(\delta_{h,\ell_k}w_r)^2+\sum_r(\delta_{h,\ell_{-k}}w_r)^2+Mw_k^2+Mw_{-k}^2,\end{aligned}\tag{4.5}$$

$$\begin{aligned}0\geq \delta^T_\tau V&=2\sum_r w_r\delta^T_\tau w_r+2Mww_\tau\\&\quad+\tau_T\sum_r(\delta^T_\tau w_r)^2+M\tau_Tw_\tau^2\geq 2\sum_r w_r\delta^T_\tau w_r+2Mww_\tau.\end{aligned}\tag{4.6}$$

It follows from (4.4) and (4.5) and our assumption: $a_k=a_{-k}\geq 0, b_k\geq 0$, that

$$0\geq 2w_rL^0_hw_r+2MwL^0_hw+(2a_k+b_kh)\big[\sum_r(\delta_{h,\ell_k}w_r)^2+Mw_k^2\big]. \tag{4.7}$$

On the other hand, due to (4.1),

$$\begin{aligned}-\delta_{h_r,\ell_r}f&=\delta^T_\tau(\xi^{-1}w_r)+\xi^{-1}\big[a_k\Delta_{h,\ell_k}w_r+(\delta_{h_r,\ell_r}a_k)\Delta_{h,\ell_k}w\\&\quad+h_r(\delta_{h_r,\ell_r}a_k)\Delta_{h,\ell_k}w_r+b_k\delta_{h,\ell_k}w_r+(\delta_{h_r,\ell_r}b_k)\delta_{h,\ell_k}w\\&\quad+h_r(\delta_{h_r,\ell_r}b_k)\delta_{h,\ell_k}w_r-cw_r-(\delta_{h_r,\ell_r}c)w-h_r(\delta_{h_r,\ell_r}c)w_r\big],\end{aligned}$$

where and in a few lines below there is no summation in r. Here (recall that $h\Delta_{h,\ell_k}u=\delta_{h,\ell_k}u+\delta_{h\ell_{-k}}u$)

$$\begin{gathered}h_r(\delta_{h_r,\ell_r}a_k)\Delta_{h,\ell_k}w_r=2h_rh^{-1}(\delta_{h_r,\ell_r}a_k)\delta_{h,\ell_k}w_r,\\ \delta^T_\tau(\xi^{-1}w_r)=\xi^{-1}\big(e^{-c_0\tau}\delta^T_\tau w_r-\tau^{-1}(1-e^{-c_0\tau})w_r\big).\end{gathered}$$

Hence,

$$\begin{aligned}-\xi\delta_{h_r,\ell_r}f&=e^{-c_0\tau}\delta^T_\tau w_r+L^0w_r+(\delta_{h_r,\ell_r}a_k)\Delta_{h,\ell_k}w\\&\quad+2h_rh^{-1}(\delta_{h_r,\ell_r}a_k)\delta_{h,\ell_k}w_r+(\delta_{h_r,\ell_r}b_k)\delta_{h,\ell_k}w+h(\delta_{h_r,\ell_r}b_k)\delta_{h,\ell_k}w_r\\&\quad-(c+\tau^{-1}(1-e^{-c_0\tau}))w_r-(\delta_{h_r,\ell_r}c)w-h_r(\delta_{h_r,\ell_r}c)w_r.\end{aligned}\tag{4.8}$$

We multiply (4.8) by $2w_r$, sum up in r and use (4.6) and (4.7). Then at (t_0,x_0) we obtain

$$\begin{aligned}-2\xi w_r\delta_{h_r,\ell_r}f&\leq -e^{-c_0\tau}2Mww_\tau-2MwL^0_hw\\&\quad-(2a_k+b_kh)\sum_r(\delta_{h,\ell_k}w_r)^2-M(2a_k+b_kh)w_k^2\\&\quad-2(c+\tau^{-1}(1-e^{-c_0\tau}))\sum_r w_r^2+I,\end{aligned}\tag{4.9}$$

where

$$\begin{aligned}I:=\,&2w_r\big[\psi_{rk}\delta_{h_r,\ell_r}a_k+(\delta_{h_r,\ell_r}b_k)\delta_{h,\ell_k}w\\&+h_r(\delta_{h_r,\ell_r}b_k)\delta_{h,\ell_k}w_r-(\delta_{h_r,\ell_r}c)w-h_r(\delta_{h_r,\ell_r}c)w_r\big],\end{aligned}$$

$$\psi_{rk}=\Delta_{h,\ell_k}w+2h_rh^{-1}\delta_{h,\ell_k}w_r.$$

Now note that

$$L_h^0 w + e^{-c_0\tau} w_\tau = -\xi f + (c + \tau^{-1}(1 - e^{-c_0\tau}))w,$$

so that (4.9) becomes

$$\begin{aligned} -2Mw\xi f - 2\xi w_r \delta_{h_r,\ell_r} f &\le -2(c + \tau^{-1}(1 - e^{-c_0\tau}))(W + Mw^2) \\ &\quad - (2a_k + b_k h)\sum_r (\delta_{h,\ell_k} w_r)^2 - M(2a_k + b_k h)w_k^2 + I. \end{aligned} \tag{4.10}$$

To estimate I observe that, since $a_k = (1/2)\sigma_k^2$, we have

$$\delta_{h_r,\ell_r} a_k = \sigma_k \delta_{h_r,\ell_r}\sigma_k + (1/2)h_r(\delta_{h_r,\ell_r}\sigma_k)^2,$$

$$J := 2w_r\psi_{rk}\delta_{h_r,\ell_r} a_k = 2w_r(\delta_{h_r,\ell_r}\sigma_k)\psi_{rk}\sigma_k + w_r\psi_{rk}h_r(\delta_{h_r,\ell_r}\sigma_k)^2.$$

Furthermore, by using inequalities like $(a+b)^2 \le 2a^2 + 2b^2$, we get

$$\frac{1}{4}\sum_{r,k}\psi_{rk}^2\sigma_k^2 = \frac{1}{2}\sum_{r,k} a_k\psi_{rk}^2 \le 2(d_1+3)\sum_{r,k} a_k(\delta_{h,\ell_k} w_r)^2,$$

and for each k and r

$$|\psi_{rk}h_r|^2 \le 36\sup_{Q,p}|w_p|^2 \le 36\sup_Q\Big(\sum_p |w_p|^2 + Mw^2\Big) = 36V.$$

It follows that

$$\Big|\sum_{r,k} w_r\psi_{rk}h_r(\delta_{h_r,\ell_r}\sigma_k)^2\Big| \le 6V\sum_{k,r}(\delta_{h_r,\ell_r}\sigma_k)^2,$$

$$\begin{aligned} J &\le 2\sum_{r,k} a_k(\delta_{h,\ell_k} w_r)^2 + 4(d_1+3)\sum_{r,k} w_r^2(\delta_{h_r,\ell_r}\sigma_k)^2 + 6V\sum_{k,r}(\delta_{h_r,\ell_r}\sigma_k)^2 \\ &\le 2\sum_{r,k} a_k(\delta_{h,\ell_k} w_r)^2 + 2(2d_1+15)VK_1^3. \end{aligned}$$

This takes care of the first term in the definition of I.

Next, for the operator $T_l : u \to u(t, x+l)$ we have

$$\begin{aligned} &\big|2w_r[(\delta_{h_r,\ell_r} b_k)\delta_{h,\ell_k} w + h_r(\delta_{h_r,\ell_r} b_k)\delta_{h,\ell_k} w_r]\big| \\ &= \big|2w_r(\delta_{h_r,\ell_r} b_k)T_{h_r\ell_r} w_k\big| \\ &\le 2V\sum_{r,k}|\delta_{h_r,l_r} b_k| \le 6K_1^2 V, \end{aligned}$$

$$\begin{aligned} &\big|2w_r[-(\delta_{h_r,\ell_r} c)w - h_r(\delta_{h_r,\ell_r} c)w_r]\big| \\ &= \big|2w_r(\delta_{h_r,\ell_r} c)T_{h_r\ell_r} w\big| \\ &\le 2M^{-1/2}V\Big|\sum_r \delta_{h_r,l_r} c\Big| \le 6K_1^2 M^{-1/2}V. \end{aligned}$$

Thus we estimated all terms in I and from (4.10) conclude that

$$\begin{aligned} -2M\xi f - 2\xi w_r\delta_{h_r,\ell_r} f &\le -2(c + \tau^{-1}(1 - e^{-c_0\tau}))V \\ &\quad + 2K_1^2[(2d_1+15)K_1 + 3 + 3M^{-1/2}] \\ &\quad - 2MV\inf_{Q_\varepsilon^0,k} a_k + 2M^2w^2\inf_{Q_\varepsilon^0,k} a_k. \end{aligned}$$

Now since $M^{-1/2} \le 1$, condition (4.2) implies that

$$2\gamma V \le 2\xi w_r \delta_{h_r,\ell_r} f + 2Mw\xi f + 2M^2 w^2 \inf_{Q^0_\varepsilon,k} a_k \le \gamma V + N\xi^2 + N\xi^2 \max_Q |u|^2.$$

It follows that

$$\gamma V \le N\xi^2 + N\xi^2 \max_Q |u|^2,$$

which along with (4.3) brings the proof of Theorem 4.1 to an end. □

Remark 4.2. Condition (4.2) is obviously satisfied for any λ, c_0, and γ if our operator is uniformly nondegenerate so that $a_k \ge \mu$ for some constant $\mu > 0$.

On the basis of Theorem 4.1, Corollary 2.9 and Lemma 2.10, the following theorem can be proved in the same way as Theorem 5.6 is deduced from Theorem 5.2 in [13]. The method of proof is similar to that of Theorem 3.4 and consists of adding a new variable and considering the coefficients with hats (see below) as values of the corresponding coefficients for one value of the additional coordinate and the original coefficients as the values at a close value of the additional coordinate.

Theorem 4.3. *Let $\hat\sigma_k$, $\hat b_k$, $\hat c$, $\hat\lambda$, $\hat f$ satisfy the assumptions in Section 2 with $n = 1$ and let $\hat\lambda = \lambda$. Let u be a function on $\bar{\mathcal{M}}_T$ satisfying (2.6) in $\mathcal{M}_T$ and let $\hat u$ be a function on $\bar{\mathcal{M}}_T$ satisfying (2.6) in $\mathcal{M}_T$ with $\hat a_k$, $\hat b_k$, $\hat c$, $\hat f$ in place of a_k, b_k, c, f respectively. Assume that u and $\hat u$ are bounded on $\bar{\mathcal{M}}_T$ and*

$$|u(T,\cdot)|, |\hat u(T,\cdot)| \le K_1.$$

Introduce

$$\mu = \sup_{\mathcal{M}_T,k} \big(|\sigma_k - \hat\sigma_k| + |b_k - \hat{b_k}| + |c - \hat c| + |f - \hat f|\big).$$

Suppose that there exist constants $N_0, c_0 \ge 0$, $\gamma > 0$ such that (4.2) holds. Then there is a constant N depending only on N_0, K_1, d, γ, and d_1, such that

$$|u - \hat u| \le N\mu e^{c_0(T+\tau)} I$$

on $\bar{\mathcal{M}}_T$, where

$$I = \sup_{x\in\mathbb{R}^d} \big(1 + (\max_k |\delta_{h,\ell_k} u| + \max_k |\delta_{h,\ell_k} \hat u| + \mu^{-1}|u - \hat u|)(T,x)\big).$$

5. Proof of Theorems 2.12-2.19

Before starting we make a general comment on our proofs. While estimating $|v(t,x) - v_{\tau,h}(t,x)|$ we may fix $(t,x) \in H_T$ and since we can always shift the origin, we may confine ourselves to $t = 0$ and $x = 0$. In particular, it suffices to obtain estimates of $|v(t,x) - v_{\tau,h}(t,x)|$ for $t = 0$ and $(0,x) \in \mathcal{M}_T$. Next, if $h > 1$ our estimates become trivial since we are dealing with bounded functions. The same goes for τ. Therefore, we assume that

$$\tau + h \le 1,$$

and $\tau^{-1} \ge N(d_1, K_1)$ such that (4.2) is satisfied with $\gamma = 1$, $N_0 = 0$ and c_0 sufficiently large. We also remind the reader that T' is the least of $\tau, 2\tau, 3\tau, \ldots$ which is $\ge T$.

Proof of Theorem 2.16. Due to Lemma 3.3, we have

$$|v|_{\bar{H}_T,2,4} \le N(K_4, d, d_1)e^{(M-\lambda)_+T}. \tag{5.1}$$

Set v_* to be the unique bounded viscosity solution of (2.1) in $\bar{H}_{T'}$ with terminal condition $v_*(T', x) = g(x)$. Then v_* satisfies the same inequality (5.1) with T' in place of T, and also since T' is a multiple of τ we have $\delta_\tau^{T'} = \delta_\tau$ on $\mathcal{M}_{T'}$. Thus by shifting the coordinates and using (5.1), for any $x \in \mathbb{R}^d$ we have

$$\begin{aligned}|v_*(T, x) - g(x)| &\le (T' - T)|v_*(T + \cdot, \cdot)|_{\bar{H}_{T'-T},1,2} \\ &\le N(K_4, d, d_1)e^{(M-\lambda)_+(T'-T)}(T' - T).\end{aligned}$$

Due to Corollary 2.8, on $\bar{H}_T$ we obtain

$$|v_*(t, x) - v(t, x)| \le N(K_4, d, d_1)e^{(M-\lambda)_+(T'-T)}\tau.$$

Also observe that if on $\mathcal{M}_{T'}$ $(= \mathcal{M}_T)$ we define $\bar{v}_{\tau,h} = v_{\tau,h}$ and let $\bar{v}_{\tau,h}(T', x) = v_{\tau,h}(T, x)$ $(= g(x))$, then on $\mathcal{M}_{T'}$

$$\delta_\tau^{T'} = \delta_\tau, \quad \delta_\tau^{T'}\bar{v}_{\tau,h} = \delta_\tau^T v_{\tau,h}, \quad \delta_\tau^{T'}\bar{v}_{\tau,h} + L_h\bar{v}_{\tau,h} + f = 0.$$

It follows by Taylor's formula that on $\mathcal{M}_{T'}$

$$\begin{aligned}&|\big(\delta_\tau^{T'} + L_h\big)\big(\bar{v}_{\tau,h} - v_*(t, x)\big)| \\ &= \big|\big(D_t + L\big)v_*(t, x) - \big(\delta_\tau + L_h\big)v_*(t, x)\big| \\ &\le N(d_1, K_4)(\tau \sup_{\bar{H}_{T'}}|D_t^2 v_*| + h \sup_{\bar{H}_{T'}}|D_x^2 v_*| + h^2 \sup_{\bar{H}_{T'}}|D_x^4 v_*|) \\ &\le N(d_1, d, K_4)e^{(M-\lambda)_+T'}(\tau + h).\end{aligned}$$

By using Lemma 2.5, we obtain on $\mathcal{M}_{T'}$

$$\begin{gathered}|v_* - v_{\tau,h}| = |v_* - \bar{v}_{\tau,h}| \le N(d_1, K_4)e^{(M-\lambda)_+T'}T'(\tau + h), \\ |v - v_{\tau,h}| \le |v - v_*| + |v_* - v_{\tau,h}| \le N(d_1, d, K_4)e^{(M-\lambda)_+T'}(T' + 1)(\tau + h).\end{gathered}$$

For $\lambda \ge 1 + M$, we use the assertion in Lemma 2.5 related to (2.10) and get

$$|v - v_{\tau,h}| \le N(d_1, d, K_4)(\tau + h).$$

Theorem 2.16 is proved. □

Proof of Theorem 2.12. We adopt the idea of mollification. Take a nonnegative function $\zeta \in C_0^\infty(\mathbb{R}^{d+1})$ with support in $(-1, 0) \times B_1$ and unit integral. For any bounded function u and $0 < \varepsilon \le 1$, we define the mollification of u by

$$u^{(\varepsilon)} = \varepsilon^{-d-2}\int_{\mathbb{R}^{d+1}} u(s, y)\zeta((t - s)/\varepsilon^2, (x - y)/\varepsilon)\, ds\, dy.$$

It is well known (cf. Lemma 5.1) that $u^{(\varepsilon)}$ is a smooth function on $\mathbb{R}^{d+1}$. And if $|u|_{1/2,1} \le K_0$, then for any integer $m \ge 1$ we have

$$|u^{(\varepsilon)}|_{m,2m} \le N(K_0, d, m)\varepsilon^{1-2m}, \tag{5.2}$$

$$|u^{(\varepsilon)} - u|_0 \le N(K_0, d)\varepsilon. \tag{5.3}$$

Next, let v^ε be the solution of (2.1)-(2.2) with $\sigma_k^{(\varepsilon)}$, $b_k^{(\varepsilon)}$, $c^{(\varepsilon)}$, $f^{(\varepsilon)}$, $g^{(\varepsilon)}$ in place of σ_k, b_k, c, f and g. By L_h^ε we denote the finite-difference operator corresponding to $\sigma_k^{(\varepsilon)}$, $b_k^{(\varepsilon)}$, $c^{(\varepsilon)}$, $f^{(\varepsilon)}$. Similarly, we introduce v_*^ε as in the proof of Theorem 2.16.

Also let $v^{\varepsilon}_{\tau,h}$ be the corresponding solution of the finite difference equation. Upon using (5.2) with $u = \sigma_k, b_k, c, f, g$ and Lemma 3.3 with $m = 1, 2$, we get

$$|v^{\varepsilon}|_{\bar{H}_T,1,2} \le \varepsilon^{-1} N(K_1, d, d_1) e^{(M-\lambda)_+ T}, \tag{5.4}$$

$$|v^{\varepsilon}|_{\bar{H}_T,2,4} \le \varepsilon^{-3} N(K_1, d, d_1) e^{(M-\lambda)_+ T}, \tag{5.5}$$

for some M depending only on K_1, d, and d_1. Two similar estimates hold for v^{ε}_* in place of v^{ε}. Therefore, by shifting the coordinates, we get

$$\begin{aligned} |v^{\varepsilon}_*(T, x) - g(x)| &\le (T' - T)|v^{\varepsilon}_*(T + \cdot, \cdot)|_{\bar{H}_{T'-T},1,2} \\ &\le \varepsilon^{-1} N(K_1, d, d_1) e^{(M-\lambda)_+ (T'-T)} \tau. \end{aligned}$$

Due to Corollary 2.8, on $\bar{H}_T$ we obtain

$$|v^{\varepsilon}_*(t, x) - v^{\varepsilon}(t, x)| \le \varepsilon^{-1} N(K_1, d, d_1) e^{(M-\lambda)_+ (T'-T)} \tau. \tag{5.6}$$

As before, we introduce $\bar{v}^{\varepsilon}_{\tau,h}$. By Taylor's formula, on $\mathcal{M}_{T'}$

$$\begin{aligned} &\big|\big(\delta^{T'}_{\tau} + L^{\varepsilon}_h\big)\big(\bar{v}^{\varepsilon}_{\tau,h}(t, x) - v^{\varepsilon}_*(t, x)\big)\big| \\ &= \big|\big(D_t + L^{\varepsilon}\big) v^{\varepsilon}_*(t, x) - \big(\delta_{\tau} + L^{\varepsilon}_h\big) v^{\varepsilon}_*(t, x)\big| \\ &\le N(d_1, d, K_1)(\tau \sup_{\bar{H}_{T'}} |D^2_t v^{\varepsilon}_*| + h \sup_{\bar{H}_{T'}} |D^2_x v^{\varepsilon}_*| + h^2 \sup_{\bar{H}_{T'}} |D^4_x v^{\varepsilon}_*|) \\ &\le N(d_1, d, K_1) e^{(M-\lambda)_+ T'} (\varepsilon^{-3}\tau + \varepsilon^{-1} h + \varepsilon^{-3} h^2). \end{aligned}$$

By using Lemma 2.5, we obtain

$$\begin{aligned} |v^{\varepsilon}_*(t, x) - v^{\varepsilon}_{\tau,h}(t, x)| &= |v^{\varepsilon}_*(t, x) - \bar{v}^{\varepsilon}_{\tau,h}(t, x)| \\ &\le N(d_1, d, K_1) e^{(M-\lambda)_+ T'} T'(\varepsilon^{-3}\tau + \varepsilon^{-1} h + \varepsilon^{-3} h^2) \end{aligned} \tag{5.7}$$

for any $(t, x) \in \mathcal{M}_{T'}$.

Owing to (5.3) and (5.2) and recalling the definition of μ and I in Theorem 4.3 (with $\sigma^{(\varepsilon)}_k$ in place of $\hat{\sigma}_k$, etc.) we write

$$\mu \le N(K_1, d)\varepsilon, \quad I \le N(K_1, d).$$

Therefore, by the result of Theorem 4.3,

$$|v_{\tau,h}(t, x) - v^{\varepsilon}_{\tau,h}(t, x)| \le N(d, d_1, K_1)\varepsilon e^{c_0(T+\tau)} \tag{5.8}$$

for any $(t, x) \in \bar{\mathcal{M}}_T$. Similarly, by Theorem 3.4 we have

$$|v(t, x) - v^{\varepsilon}(t, x)| \le N(d, d_1, K_1)\varepsilon e^{(M-\lambda)_+ T} \tag{5.9}$$

for any $(t, x) \in \bar{\mathcal{M}}_T$. After combining (5.6)-(5.9), we reach

$$|v(t, x) - v_{\tau,h}(t, x)| \le N(d_1, d, K_1, T')(\varepsilon + \varepsilon^{-3}\tau + \varepsilon^{-1}\tau + \varepsilon^{-1} h + \varepsilon^{-3} h^2)$$

for any $(t, x) \in \mathcal{M}_T$. It only remains to put $\varepsilon = (\tau + h^2)^{1/4}$, and then the first part of Theorem 2.12 is proved. To prove the second part, it suffices to inspect the above argument and see that for $\lambda \ge M(K_1, d, d_1)$ we can get rid of the exponential factors in (5.7)-(5.9) and also the factor T' in (5.7). This proves the second part and the theorem. □

Proof of Theorem 2.14. We take the function ζ from the previous proof and additionally assume that ζ is symmetric with respect to x, i.e.

$$\zeta(t, x) = \zeta(t, -x), \quad \forall\, (t, x) \in \mathbb{R}^{d+1}.$$

□

Lemma 5.1. *With the function ζ as above, for any bounded function u on $\mathbb{R}^{d+1}$ such that $|u|_{1,2} \le K_0$, any integer $m \ge 0$ and any $\varepsilon \in (0,1]$,*

$$|u^{(\varepsilon)}|_{m,2m} \le N(K_0,d)\varepsilon^{2-2m}. \tag{5.10}$$

Moreover, on $\mathbb{R}^{d+1}$

$$|u^{(\varepsilon)} - u| \le N(K_0,d)\varepsilon^2. \tag{5.11}$$

Proof. Since ζ is symmetric with respect to x, for any nonnegative integers i,j satisfying $2i+j=2m$, we have

$$\begin{aligned} D_t^i D_x^j u^{(\varepsilon)} &= \varepsilon^{-(2i+j)} \int_{\mathbb{R}^{d+1}} u(t-\varepsilon^2 s, x-\varepsilon y) D_s^i D_y^j \zeta(s,y)\,ds\,dy \\ &= \frac{1}{2}\varepsilon^{-(2i+j)} \int_{\mathbb{R}^{d+1}} I_1 D_s^i D_y^j \zeta(s,y)\,ds\,dy, \end{aligned} \tag{5.12}$$

where $I_1 = u(t-\varepsilon^2 s, x-\varepsilon y) + u(t-\varepsilon^2 s, x+\varepsilon y) - 2u(t,x)$. Note that by Taylor's formula

$$\begin{aligned} |I_1| &\le |u(t-\varepsilon^2 s, x-\varepsilon y) + u(t-\varepsilon^2 s, x+\varepsilon y) - 2u(t-\varepsilon^2 s, x)| \\ &\quad + 2|u(t-\varepsilon^2 s, x) - u(t,x)| \le K_0\varepsilon^2(|y|^2 + 2|s|). \end{aligned}$$

Coming back to (5.12) yields (5.10). To prove (5.11), we only need to notice that

$$u^{(\varepsilon)}(t,x) - u(t,x) = \frac{1}{2}\int_{\mathbb{R}^d} I_1 \zeta(s,y)\,ds\,dy.$$

□

Due to the previous lemma and Lemma 3.3, instead of (5.4)-(5.5), we have

$$|v^\varepsilon|_{\bar{H}_T,1,2} \le N e^{(M-\lambda)_+ T}, \quad |v^\varepsilon|_{\bar{H}_T,2,4} \le \varepsilon^{-2} N e^{(M-\lambda)_+ T}, \tag{5.13}$$

where $N = N(K_2,d,d_1)$, $\varepsilon \in (0,1]$. Then as above

$$|v_*^\varepsilon(t,x) - v^\varepsilon(t,x)| \le N(K_2,d,d_1) e^{(M-\lambda)_+ (T'-T)}\tau. \tag{5.14}$$

for some M depending only on K_1, d and d_1. Also,

$$|\big(\delta_\tau^{T'} + L_h^\varepsilon\big)\big(v_*^\varepsilon(t,x) - \bar{v}_{\tau,h}^\varepsilon(t,x)\big)| \le N(d_1,K_2) e^{(M-\lambda)_+ T'}(\varepsilon^{-2}\tau + h + \varepsilon^{-2}h^2),$$

for any $(t,x) \in \mathcal{M}_T$ and $\varepsilon \in (0,1]$. Hence,

$$\begin{aligned} |v_*^\varepsilon(t,x) - v_{\tau,h}^\varepsilon(t,x)| &= |v_*^\varepsilon(t,x) - \bar{v}_{\tau,h}^\varepsilon(t,x)| \\ &\le N(d_1,d,K_2) e^{(M-\lambda)_+ T} T'(\varepsilon^{-2}\tau + h + \varepsilon^{-2}h^2), \end{aligned} \tag{5.15}$$

$$|v_{\tau,h}(t,x) - v_{\tau,h}^\varepsilon(t,x)| \le N(d,d_1,K_2)\varepsilon^2 e^{c_0(T+\tau)}, \tag{5.16}$$

$$|v(t,x) - v^\varepsilon(t,x)| \le N(d,d_1,K_2)\varepsilon^2 e^{(M-\lambda)_+ T}. \tag{5.17}$$

After combining (5.14)-(5.17), we obtain

$$|v(t,x) - v_{\tau,h}(t,x)| \le N(d_1,d,K_2,T')(\tau + \varepsilon^2 + \varepsilon^{-2}\tau + h + \varepsilon^{-2}h^2). \tag{5.18}$$

for any $(t,x) \in \mathcal{M}_T$. Again we put $\varepsilon = (\tau + h^2)^{1/4}$, and the first part of Theorem 2.14 is proved. As before, for $\lambda \ge M(K_1,d,d_1)$, we can make N in (5.18) to be independent of T.

Proof of Theorem 2.18. As we have already pointed out, under Assumption 2.17, Lemmas 2.3, 2.4, and 2.5 still hold true with the operator (2.13) for $h \le 2/K$. By Taylor's formula, for any three times continuously differentiable (in x) function u,

$$|\delta_{2h,\ell_k} u(x - h\ell_k) - D_{\ell_k} u(x)| \le h^2 \sup_{s\in[-h,h]} |D^3_{\ell_k} u(x + s\ell_k)|/6.$$

Therefore, this time

$$\begin{aligned}
&|\big(\delta^{T'}_\tau + L_h\big)(\bar v_{\tau,h}(t, x) - v_*(t, x)| \\
&= \big|\big(D_t + L\big)v_*(t, x) - \big(\delta_\tau + L_h\big)v_*(t, x)\big| \\
&\le N(d_1, d, K_4)(\tau \sup_{\bar H_{T'}} |D^2_t v_*| + h^2 \sup_{\bar H_{T'}} |D^3_x v_*|| + h^2 \sup_{\bar H_{T'}} |D^4_x v_*||) \\
&\le N(d_1, d, K_4)e^{(M-\lambda)_+T'}(\tau + h^2),
\end{aligned}$$

for any $(t, x) \in \mathcal{M}_T$. By using Lemma 2.5, we obtain that on $\mathcal{M}_T$ (we always assume that $h \le 2/K$)

$$|v_* - v_{\tau,h}| = |v_* - \bar v_{\tau,h}| \le N(d_1, d, K_4)e^{(M-\lambda)_+T'} T'(\tau + h^2),$$

and as few times above

$$|v - v_{\tau,h}| \le |v - v_*| + |v_* - v_{\tau,h}| \le N(d_1, d, K_4)e^{(M-\lambda)_+T'}(T' + 1)(\tau + h^2).$$

For $\lambda \ge 1 + M$, we use (2.10) again and get on $\mathcal{M}_T$

$$|v - v_{\tau,h}| \le N(d_1, d, K_4)(\tau + h^2).$$

Theorem 2.18 is proved. □

Proof of Theorem 2.13, 2.15 and 2.19. We take $g \equiv 0$ and denote the functions v and $v_{\tau,h}$ from Theorem 2.12, (2.14, and 2.18, respectively) by v^T and $v^T_{\tau,h}$. Obviously, it suffices to prove that for all (t, x)

$$\tilde v(x) = \lim_{T\to\infty} v^T(t, x), \quad \tilde v_h(x) = \lim_{T\to\infty} v^T_{\tau,h}(t, x), \tag{5.19}$$

whenever $\lambda > 0$ and τ is small enough.

The first relation in (5.19) is well known (see, for instance, [9] or [18]). To prove the second, it suffices to prove that for any sequence $T_n \to \infty$ such that $v^{T_n}_{\tau,h}(t, x)$ converges at all points of $\mathcal{M}_\infty$, the limit is independent of t and satisfies (2.11) on the grid

$$G = \{i_1 h\ell_1 + \dots + i_{d_1} h\ell_{d_1} : i_k = 0, \pm 1, \dots, k = 1, \dots, d_1\}.$$

Given the former, the latter is obvious. Also notice that if σ_k, b_k, c and f are independent of t, the translation $t \to t + \tau$ brings any solution of (2.6) on $\mathcal{M}_\infty$ again to a solution. Therefore, it only remains to prove uniqueness of bounded solutions of (2.6) on $\mathcal{M}_\infty$.

Observe that if u_1 and u_2 are two solutions of (2.6) on $\mathcal{M}_\infty$, then they also solve (2.6) on $\mathcal{M}_T$ for any T with terminal condition u_1 and u_2, respectively. By the comparison result

$$|u_1(t, x) - u_2(t, x)| \le e^{-\lambda(T-t)/2} \sup_x |u_1(T, x) - u_2(T, x)|,$$

if τ is small enough. By sending $T \to \infty$ we prove the uniqueness and the theorem. □

6. Proof of Lemma 3.3

Obviously, the first part of Lemma 3.3 follows immediately from the second part. Recall that v is given in (2.3) with $x_s = x_s(t,x)$ defined by (2.4). We fix $(t,x) \in H_T$, take a $\xi \in \mathbb{R}^d$ and set $\xi_s^{(i)} = \xi_s^{(i)}(x,\xi), i = 1, 2, \dots, 2m$ to be the i^{th} order derivative of x_s at point (t,x) in the direction of ξ, i.e.,

$$\xi_s^{(1)} = (x_s(t,x))_{(\xi)}, \quad \xi_s^{(2)} = (x_s(t,x))_{(\xi)(\xi)}, \quad \text{etc.}$$

We know that for example $\xi_s^{(1)}$ and $\xi_s^{(2)}$ satisfy the equations

$$d\xi_s^{(1)} = \sigma_{(\xi_s^{(1)})}(s,x_s)\,dw_s + b_{(\xi_t^{(1)})}(s,x_s)\,ds,$$

$$\begin{aligned} d\xi_s^{(2)} =& \big[\sigma_{(\xi_s^{(2)})}(s,x_s) + \sigma_{(\xi_s^{(1)})(\xi_s^{(1)})}(s,x_s)\big]\,dw_s \\ &+ \big[b_{(\xi_s^{(2)})}(s,x_s) + b_{(\xi_s^{(1)})(\xi_s^{(1)})}(s,x_s)\big]\,ds. \end{aligned}$$

In general $\xi_s^{(i)}$ satisfies

$$d\xi_s^{(i)} = \sigma_{(\xi_s^{(i)})}(s,x_s)\,dw_s + b_{(\xi_s^{(i)})}(s,x_s)\,ds + S_1\,dw_s + S_2\,ds, \tag{6.1}$$

where S_1 is the sum of the terms

$$\sigma_{(\xi_s^{(k_1)})(\xi_s^{(k_2)})\dots(\xi_s^{(k_l)})}(s,x_s), \quad \text{for } 1 \le k_j < i,\ \sum_{j=1}^{l} k_j = i.$$

Similarly, S_2 is the sum of the terms

$$b_{(\xi_s^{(k_1)})(\xi_s^{(k_2)})\dots(\xi_s^{(k_l)})}(s,x_s), \quad \text{for } 1 \le k_j < i,\ \sum_{j=1}^{l} k_j = i.$$

Definition 6.1. Given real numbers $A_{i_1,i_2,\dots,i_{2m}}$ defined for $i_1,\dots,i_{2m} = 1,\dots,d$, we say that

$$A = \{A_{i_1,i_2,\dots,i_{2m}}\}_{i_1,\dots,i_{2m}=1}^{d}$$

is strictly positive definite if the following two conditions are satisfied

1) The value of $A_{i_1,i_2,\dots,i_{2m}}$ does not change if we interchange any two indices.
2) For any $x = (x^1, x^2, \dots, x^d) \in \mathbb{R}^d \setminus \{0\}$,

$$A(x) := \sum_{i_1,i_2,\dots,i_{2m}} A_{i_1,i_2,\dots,i_{2m}} x^{i_1} x^{i_2} \dots x^{i_{2m}} > 0.$$

Assumption 6.2. We are given constants $M \ge 0$ and $\delta > 0$ and a strictly positive definite A such that for any $(t,x) \in H_T$ and $\xi \in \mathbb{R}^d$,

$$\begin{aligned} &m(2m-1) \sum_{i_1,\dots,i_{2m}} \sum_{|j|=1}^{d_1} \sigma_{(\xi)}^{i_1,j}(t,x)\sigma_{(\xi)}^{i_2,j}(t,x)\xi^{i_3}\dots\xi^{i_{2m}} A_{i_1,\dots,i_{2m}} \\ &+2m \sum_{i_1,\dots,i_{2m}} b_{(\xi)}^{i_1}(t,x)\xi^{i_2}\xi^{i_3}\dots\xi^{i_{2m}} A_{i_1,\dots,i_{2m}} \le (M-\delta)A(\xi). \end{aligned} \tag{6.2}$$

Denote $\sigma(t,x,y) = \sigma_{(y)}(t,x)$ and $b(t,x,y) = b_{(y)}(t,x)$. The following lemma is proved in [6] and is a generalized version of Lemma 7.2 in [14].

Lemma 6.3. *Let α_s be a $d \times d_1$ matrix-valued and β_s an $\mathbb{R}^d$-valued predictable processes satisfying natural integrability conditions so that the equation*

$$dy_s = [\sigma(s, x_s(t,x), y_s) + \alpha_s]\, dw_s + [b(s, x_s(t,x), y_s) + \beta_s]\, ds, \quad s > t \tag{6.3}$$

makes sense and let y_s be its solution with a nonrandom initial condition $y \in \mathbb{R}^d$. Then under Assumption 6.2, there exists a number $p_0 = p_0(m, \delta, \sigma) > 1$ such that for any stopping time $\tau \le T$ and constant δ_1, δ_2, p,satisfying $0 \le \delta_1 < \delta_2 \le \delta/2$ and $p \in (0, p_0]$, we have

$$\begin{aligned} &E \sup_{t \le s \le \tau} \big[(e^{p(-M+\delta_1)(s-t)}|y_s|^{2mp}\big] \\ &\le N|y|^{2mp} + NE \sup_{t \le s \le \tau} \big[e^{p(-M+\delta_2)(s-t)}(\|\alpha_s\|^{2mp} + |\beta_s|^{2mp})\big], \end{aligned} \tag{6.4}$$

where $N = N(m, d, p, K_1, A, \delta, \delta_1, \delta_2)$.

In what follows, we only use Lemma 6.3 for $\tau = T$. Due to the assumption, we have $K_1 \le N_0$. Next we estimate the moments inductively. Since $\xi_t^{(1)}$ satisfies (6.3) with $\alpha = \beta = 0$, by using Lemma 6.3 with $p = 1$, we have

$$E \sup_{t \le s \le T} [e^{(-M+\delta/4)(s-t)}|\xi_s^{(1)}(x,\xi)|^{2m}] \le N|\xi|^{2m}.$$

We have the natural initial conditions $\xi_0^{(i)} = 0$, $\forall i \ge 2$. Since $M, A \ge 0$, condition (6.2) is satisfied if we replace M with $2M$. For $i = 2$, because of (6.1) and (6.4) with $p = 1/2$, we have

$$\begin{aligned} &E \sup_{t \le s \le T} [e^{(-M+\delta/8)(s-t)}|\xi_s^{(2)}(x,\xi)|^m] \\ &= E \sup_{t \le s \le T} [e^{-(1/2)(2M+\delta/4)(s-t)}|\xi_s^{(2)}(x,\xi)|^m] \\ &\le N(d, d_1, m, N_0, A, \delta) K_2^m E \sup_{t \le s \le T} [e^{-(1/2)(2M+\delta/2)(s-t)}|\xi_s^{(1)}(x,\xi)|^{2m}] \\ &\le N(d, d_1, m, N_0, A, \delta)\varepsilon^{-m(2-l)_+}|\xi|^{2m}. \end{aligned}$$

For $i = 3$, we note that in this case α_s is a linear combination of

$$\sigma_{(\xi_s^{(1)})(\xi_s^{(2)})}, \quad \sigma_{(\xi_s^{(1)})(\xi_s^{(1)})(\xi_s^{(1)})},$$

and β_s is a linear combination of

$$b_{(\xi_s^{(1)})(\xi_s^{(2)})}, \quad b_{(\xi_s^{(1)})(\xi_s^{(1)})(\xi_s^{(1)})}.$$

Therefore, upon using (6.4) with $p = 1/3$ we obtain

$$\begin{aligned} &E \sup_{t \le s \le T} [e^{(-M+\delta/16)(s-t)}|\xi_s^{(3)}(x,\xi)|^{2m/3}] \\ &\le N(d, d_1, m, N_0, A, \delta) E \sup_{t \le s \le T} \big[e^{(-M+\delta/8)(s-t)}(|\xi_s^{(1)}|^{2m}\varepsilon^{-2m(3-l)_+/3} \\ &+ |\xi_s^{(1)}|^{2m/3}|\xi_s^{(2)}|^{2m/3}\varepsilon^{-2m(2-l)_+/3})\big] \\ &\le N(d, d_1, m, N_0, A, \delta) E \sup_{t \le s \le T} \big[e^{(-M+\delta/8)(s-t)}(|\xi_s^{(1)}|^{2m}\varepsilon^{-2m(3-l)_+/3} \\ &+ |\xi_s^{(2)}|^m \varepsilon^{m(2-l)_+ - 2m(3-l)_+/3})\big] \le N|\xi|^{2m}\varepsilon^{-2m(1-l/3)_+}. \end{aligned}$$

Using similar arguments, one gets the following estimate for $\xi_s^{(i)}$.

Lemma 6.4. *Under the assumption of Lemma 3.3 (ii), for any ξ in $\mathbb{R}^d$, $(t,x)\in H_T$, we have*

$$E\sup_{t\le s\le\tau} e^{(-M+\delta/2^{1+i})(s-t)}|\xi_s^{(i)}(t,x,\xi)|^{2m/i}\le N|\xi|^{2m}\varepsilon^{-2m(1-l/i)_+},$$

for $i=1,2,\dots,2m$, where $N=N(d,d_1,m,N_0,A,\delta)$. Moreover, due to Hölder's inequality,

$$E\sup_{t\le s\le\tau} e^{(-M+\delta/2^{1+i})(s-t)}|\xi_s^{(i)}(t,x,\xi)|^{q/i}\le N|\xi|^{q}\varepsilon^{-q(1-l/i)_+}, \tag{6.5}$$

for $i=1,2,\dots,2m$, $0\le q\le 2m$, where $N=N(d,d_1,m,N_0,A,\delta)$.

Now we are ready to prove Lemma 3.3 (ii). Firstly, since v satisfies (2.1), we only need to consider the spatial derivatives. It is easy to see (cf. [19]) that for $1\le q\le 2m$ and any unit $\xi\in\mathbb{R}^d$,

$$\big|v_{\underbrace{(\xi)\dots(\xi)}_{q}}(t,x)\big|\le N(d,d_1,m)\big(E\int_t^T(1+s^q)e^{-\varphi_s}I\,ds+(1+T^q)Ee^{-\varphi_T}J\big), \tag{6.6}$$

where

$$\varphi_s=\int_t^s c(s,x_s)\,ds,$$

I is a linear combination of

$$\sup_{H_T}|D_x^i f|\sup_{H_T}|D_x^j c|\prod_{r=1}^{i+j}|\xi_s^{(k_r)}(t,x,\xi)|,\quad \text{for } 1\le i+j\le q,\ \sum_r k_r=q,$$

and J is a linear combination of

$$\sup_{x}|D_x^i g|\sup_{H_T}|D_x^j c|\prod_{r=1}^{i+j}|\xi_s^{(k_r)}(t,x,\xi)|,\quad \text{for } 1\le i+j\le q,\ \sum_r k_r=q,$$

By Hölder's inequality and (6.5),

$$\begin{aligned}
&Ee^{(-M+\delta/2^{1+q})(s-t)}\prod_r|\xi_s^{(k_r)}(t,x,\xi)|\\
&\le\prod_r\big[Ee^{(-M+\delta/2^{1+q})(s-t)}|\xi_s^{(k_r)}(t,x,\xi)|^{q/k_r}\big]^{k_r/q}\\
&\le N\varepsilon^{-\sum_r(k_r-l)_+}.
\end{aligned}$$

Also note that by assumption

$$(\sup_x|D_x^i g|+\sup_{H_T}|D_x^i f|)\sup_{H_T}|D_x^j c|\le N\varepsilon^{-(i-l)_+-(j-l)_+},$$

and

$$\sum_r(k_r-l)_++(i-l)_++(j-l)_+\le(q-l)_+.$$

Thus, the left-hand side of (6.6) is less than or equal to

$$\begin{aligned}
&N(d,d_1,m,A,N_0)\varepsilon^{-(q-l)_+}\Big(\int_t^T(1+s^q)e^{(M-\lambda-\delta/2^{1+q})(s-t)}\,dt\\
&+(1+T^q)e^{(M-\lambda-\delta/2^{1+q})(T-t)}\Big)\\
&\le N(d,d_1,m,\delta,A,N_0)\varepsilon^{-(q-l)_+}e^{(M-\lambda)_+T}.
\end{aligned}$$

This yields Lemma 3.3 (ii) if we make M sufficiently large so that condition (6.2) satisfies for $\delta = 1$ and A being the identity.

7. Discussion of semi-discrete schemes

The following result about semi-discretization allows one to use approximations of the time derivative different from the one in (2.6), in particular, explicit schemes could be used. The semi-discrete approximations for (2.1) are introduced by means of the equation

$$\frac{\partial}{\partial t} u(t,x) + L_h u(t,x) + f(t,x) = 0, \quad (t,x) \in H_T, \tag{7.1}$$

with terminal condition (2.2).

We claim that all the estimates in Theorem 2.12, 2.14, 2.16, and 2.18 still hold if we drop the terms with τ in the right-hand sides. We follow closely the arguments in [13]. The unique solvability of (7.1)-(2.2) in the space of bounded continuous functions is shown by rewriting the problem as

$$u(t,x) = g(x) + \int_t^T \big(L_h u(s,x) + f(s,x)\big)\, ds$$

and using the method of successive approximations.

Next, as in [13] on the basis of the comparison results and Theorem 4.1 one shows that for $(t,x),(s,y) \in \bar{H}_T$ we have

$$|v_{\tau,h}(t,x) - v_{\tau,h}(t,y)| \le N|x-y|,$$
$$|v_{\tau,h}(t,x) - v_{\tau,h}(s,x)| \le N(|t-s|^{1/2} + \tau^{1/2})$$

with N independent of τ, h, t, x, s, y. It follows easily that one can find a sequence $\tau_n \downarrow 0$ such that $v_{\tau_n,h}(t,x)$ converges at each point of $\mathbb{R}^d$ uniformly in $t \in [0,T]$. Call u the limit of one of subsequences and introduce

$$\kappa_n(t) = i\tau_n \quad \text{for } i\tau_n \le t < (i+1)\tau_n, \quad i = 0,1,\dots$$

Then for any smooth $\psi(t)$ vanishing at $t = T$ and $t = 0$,

$$\begin{aligned} &\int_0^T [\psi(L_h v_{\tau_n,h} + f)](\kappa_n(t),x)\, dt \\ &= \int_0^T v_{\tau_n,h}(\kappa_n(t),x)\tau_n^{-1}\big(\psi(\kappa_n(t),x) - \psi(\kappa_n(t)-\tau_n,x)\big)\, dt. \end{aligned}$$

Since the integrands converge uniformly on $[0,T]$ to their natural limits, we conclude that u satisfies (7.1) in the weak sense. On the other hand, u is also a continuous function and $u(T,x) = g(x)$. It follows that u satisfies (7.1). Now our assertion follows directly from Theorem 2.12, 2.14, 2.16, and 2.18.

8. Discussion of equations in cylinders

Some methods of this article can also be applied to equations in cylinders like $Q = [0,T) \times D$, where D is a domain in $\mathbb{R}^d$. It is natural to consider (2.1) and (2.6) in Q with terminal condition $u(T,x) = g(x)$ in D and require v and v_h be zero in $[0,T] \times (\mathbb{R}^d \setminus D)$. We assume that $g = 0$ on ∂D and that there is a sufficiently smooth function ψ on $\mathbb{R}^d$ such that $\psi > 0$ in D, $\psi = 0$ on ∂D, $1 \le |\psi_x| \le K_0$ on

∂D, and $L\psi < -1$ in Q. Then, for sufficiently small h, by the smoothness of ψ, we also have $L_h\psi \le -1/2$ in Q. Due to Lemma 2.5 and Remark 2.6,

$$|v(t,x)| \le K_0\psi(t,x), \quad |v_{\tau,h}(t,x)| \le 2K_0(\psi(t,x)+h),$$

for any $(t,x) \in Q$. These estimates give us necessary control of solutions near the boundary of Q. Now instead of Theorem 4.3 and 3.4, we have the following results, which is deduced from Theorem 4.1 and 3.1 respectively.

Theorem 8.1. *Let Q_1 be a finite set in $\bar{\mathcal{M}}_T$, and suppose a_k, b_k, c, f satisfy the same assumption as in Theorem 4.3. Let u be a function on $\bar{\mathcal{M}}_T$ satisfying* (2.6) *in $Q_1 \cap H_T$ and let $\hat{u}$ be a function on $\bar{\mathcal{M}}_T$ satisfying* (2.6) *in $Q_1 \cap H_T$ with $\hat{a}_k$, $\hat{b}_k$, $\hat{c}$, $\hat{f}$ in place of a_k, b_k, c, f respectively. Assume that u and $\bar{u}$ are bounded on $\bar{\mathcal{M}}_T$ and*

$$|u(T,\cdot)|, |\hat{u}(T,\cdot)| \le K_1.$$

Assume that there is an $\varepsilon > 0$ such that

$$\sup_{\mathcal{M}_T,k} \big(|\sigma_k - \hat{\sigma}_k| + |b_k - \hat{b}_k| + |c-\hat{c}| + |f-\hat{f}|\big) \le K_1\varepsilon.$$

Suppose that there exist constants $N_0, c_0 \ge 0$, $\gamma > 0$ such that (4.2) *holds. Then there exists a constant N depending only on N_0, K_1, γ, d, and d_1 such that in Q_1*

$$|u - \hat{u}| \le N\varepsilon e^{c_0(T+\tau)} I,$$

where

$$I = 1 + \sup_{Q_1}(|u| + |\hat{u}|) + \sup_{\partial_0 Q_1}\big(\max_k |\delta_{h,l_k} u| + \max_k |\delta_{h,l_k}\hat{u}| + \varepsilon^{-1}|u-\hat{u}|\big).$$

Theorem 8.2. *Let Q_2 be a bounded set in $\bar{H}_T$, and suppose a_k, b_k, c, f satisfy the same assumption as in Theorem 3.4. Let v be a function on $\bar{H}_T$ satisfying* (2.1) *in $Q_2 \cap H_T$ and let $\hat{v}$ be a function on $\bar{H}_T$ satisfying* (2.1) *in $Q_2 \cap H_T$ with $\hat{a}_k$, $\hat{b}_k$, $\hat{c}$, $\hat{f}$ in place of a_k, b_k, c, f respectively. Assume that v and $\bar{v}$ are bounded on $\bar{H}_T$ and*

$$|v(T,\cdot)|, |\hat{v}(T,\cdot)| \le K_1.$$

Assume that there is an $\varepsilon > 0$ such that

$$\sup_{H_T,k} \big(|\sigma_k - \hat{\sigma}_k| + |b_k - \hat{b}_k| + |c-\hat{c}| + |f-\hat{f}|\big) < K_1\varepsilon.$$

Then there are constants N and M depending only on K_1, d, and d_1, such that on $\bar{H}_T$

$$|v - \hat{v}| \le N\varepsilon e^{(M-\lambda)_+ T} I$$

where

$$I = 1 + \sup_{Q_2}(|v| + |\hat{v}|) + \sup_{\partial Q_2}\big(\max_k |D_{l_k} v| + \max_k |D_{l_k}\hat{v}| + \varepsilon^{-1}|v-\hat{v}|\big).$$

However, in what concerns the rate of convergence, these results do not allow us to carry over the methods of the present article to equations in cylinders. The point is that, no matter how smooth the coefficients and the domain are, the true solutions may be just discontinuous. Interestingly enough, it seems that for bounded domains one has to apply the theory of controlled diffusion processes (= the theory of fully nonlinear PDEs) in order to deal with the rate of convergence for *linear* equations. We plan to show this in a subsequent article.

References

[1] G. Barles and E. R. Jakobsen, *On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations*, Math. Modelling and Numerical Analysis, ESAIM, M2AM, Vol. 36, No. 1 (2002), 33-54.

[2] G. Barles and E. R. Jakobsen, *Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations*, SIAM J. Numer. Anal., Vol. 43 (2005), No. 2, 540-558.

[3] L. Bers, F. John, and M. Schechter, "Partial differential equations", Interscience Publishers, New York-London-Sydney, 1964.

[4] M. Bondesson, *Interior a priori estimates in discrete L_p norms for solutions of parabolic and elliptic difference equations*, Ann. Mat. Pura Appl. (4), Vol. 95 (1973), 1–43.

[5] R. Bürger and K. H. Karlsen, *On some upwind difference schemes for the phenomenological sedimentation-consolidation model*, J. Engineering Math., Vol., 41 (2001), 145-166.

[6] Hongjie Dong, *About smoothness of solutions of the heat equations in closed smooth space-time domains*, Comm. Pure Appl. Math., Vol. 58 (2005) No. 6, 799-820.

[7] Hongjie Dong, N. V. Krylov, *On the rate of convergence of finite-difference approximations for Bellman's equation with constant coefficients*, Algebra i Analis (St. Petersburg Math. J.) Vol. 17 (2005), No 2.

[8] W. Fleming and M. Soner, "Controlled Markov processes and viscosity solutions", Springer Verlag, 1993.

[9] A. Friedman, "Partial differential equations of parabolic type", Prentice-Hall, Inc., Englewood Cliffs, N. J., 1964.

[10] D. W. Holtby, *Higher-order estimates for fully nonlinear difference equations. II*, Proc. Edinb. Math. Soc. (2) 44 (2001), No. 1, 87–102.

[11] E. R. Jakobsen, *On the rate of convergence of approximation schemes for Bellman equations associated with optimal stopping time problems*, Math. Models Methods Appl. Sci., Vol. 13 (2003), No. 5, 613-644.

[12] K. H. Karlsen and N. H. Risebro, *Convergence of finite difference schemes for viscous and inviscid conservation laws with rough coefficients*, M2AN Math. Model. Numer. Anal., Vol. 35 (2001), No. 2, 239–269.

[13] N. V. Krylov, *On the rate of convergence of finite-difference approximations for Bellman's equation with Lipschitz coefficients*, to appear in Appl Math Optim.

[14] N. V. Krylov, *Adapting some idea from stochastic control theory to studying the heat equation in closed smooth domain*, Appl Math Optim 46:231-261 (2002).

[15] N. V. Krylov, *On the rate of convergence of finite–difference approximations for Bellman's equations with variable coefficients*, Probab. Theory Relat. Fields, Vol. 117, No. 1 (2000), 1-16.

[16] N. V. Krylov, *Approximating value functions for controlled degenerate diffusion processes by using piece-wise constant policies*, Electronic Journal of Probability, Vol. 4 (1999), paper No. 2, 1-19,
http://www.math.washington.edu/ ejpecp/EjpVol4/paper2.abs.html

[17] N. V. Krylov, *On the rate of convergence of finite-difference approximations for Bellman's equation*, St. Petersburg Math. J., Vol. 9, No. 3 (1997), 245-256.

[18] N. V. Krylov, "Controlled diffusion processes", Nauka, Moscow, 1977; English transl., Springer, 1980.

[19] N. V. Krylov, *Introduction to the theory of diffusion processes*, American Mathematical Society, Providence, RI, 1995.

[20] H. J. Kuo and N. S. Trudinger, *Evolving monotone difference operators on general space-time meshes*, Duke Math. J., Vol. 91 (1998), No. 3, 587-607.

[21] H. J. Kuo and N. S. Trudinger, *Schauder estimates for fully nonlinear elliptic difference operators*, Proc. Roy. Soc. Edinburgh Sect. A 132 (2002), No. 6, 1395–1406.

[22] O. A. Ladyzhenskaya, "Kraevye zadachi matematicheskoj fiziki", Nauka, Moscow, 1973 in Russian; English translation: "The boundary value problems of mathematical physics", Applied Math. Sciences, Vol. 49, Springer-Verlag, New York etc, 1985.

[23] S. Larsson and V. Thomée, "Partial differential equations with numerical methods", Texts in Applied Mathematics, 45. Springer-Verlag, Berlin, 2003.

[24] R. D. Richtmyer and K. W. Morton, "Difference methods for initial-value problems", Interscience Publishers, New York, 1967.

[25] D. C. Shreve, *Interior estimates in l^p for elliptic difference operators*, SIAM J. Numer. Anal., Vol. 10 (1973), No. 1. 69-80.
[26] R. Stevenson, *Discrete Sobolev spaces and regularity of elliptic difference schemes*, Mathematical Modelling and Numerical Analysis, Vol. 25 (1991), No. 5, 607-640.
[27] O. B. Widlund, *On the rate of convergence for parabolic difference schemes. II*, Comm. Pure Appl. Math., Vol. 23 (1970), 79-96.
[28] Jianfeng Zhang, *Rate of convergence of finite difference approximations for degenerate ODEs*, to appear in Math. Computation.

Hongjie Dong
Department of Mathematics, University of Chicago, 5734 S. University Avenue, Chicago, Illinois 60637, USA
E-mail address: `hjdong@math.uchicago.edu`

Nicolai V. Krylov
127 Vincent Hall, University of Minnesota, Minneapolis, MN 55455, USA
E-mail address: `krylov@math.umn.edu`