DIFFERENCE-QUADRATURE SCHEMES FOR NONLINEAR DEGENERATE PARABOLIC INTEGRO-PDE

I. H. BISWAS, E. R. JAKOBSEN, AND K. H. KARLSEN

Abstract. We derive and analyze monotone difference-quadrature schemes for Bellman equations of controlled Lévy (jump-diffusion) processes. These equations are fully nonlinear, degenerate parabolic integro-PDEs interpreted in the sense of viscosity solutions. We propose new "direct" discretizations of the non-local part of the equation that give rise to monotone schemes capable of handling singular Lévy measures. Furthermore, we develop a new general theory for deriving error estimates for approximate solutions of integro-PDEs, which thereafter is applied to the proposed difference-quadrature schemes.

Contents

1. Introduction

In this article we derive and analyze numerical schemes for fully nonlinear, degenerate parabolic integro partial differential equations (IPDEs) of Bellman type. To be precise, we consider the initial value problem

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

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

Date: April 12, 2010.
2000 *Mathematics Subject Classification.* Primary 45K05, 65M12; 49L25,65L70.
Key words and phrases. Integro-partial differential equation, viscosity solution, finite difference scheme, error estimate, stochastic optimal control, Lévy process, Bellman equation.
This work was supported by the Research Council of Norway (NFR) through the project "Integro-PDEs: Numerical methods, Analysis, and Applications to Finance". The work of K. H. Karlsen was also supported trough a NFR Outstanding young Investigator Award. This article was written as part of the international research program on Nonlinear Partial Differential Equations at the Centre for Advanced Study at the Norwegian Academy of Science and Letters in Oslo during the academic year 2008–09.

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

$$L^\alpha[\phi](t,x) := \mathrm{tr}\big[a^\alpha(t,x)D^2\phi\big] + b^\alpha(t,x)D\phi,$$
$$J^\alpha[\phi](t,x) := \int_{\mathbb{R}^M\setminus\{0\}} \Big(\phi(t,x+\eta^\alpha(t,x,z)) - \phi - \mathbf{1}_{|z|\le 1}\eta^\alpha(t,x,z)D\phi\Big)\nu(dz),$$

for smooth bounded functions ϕ. Equation (1.1) is convex and non-local. The coefficients $a^\alpha, \eta^\alpha, b^\alpha, c^\alpha, f^\alpha, g$ are given functions taking values respectively in $\mathbb{S}^N$ ($N\times N$ symmetric matrices), $\mathbb{R}^N$, $\mathbb{R}^N$, $\mathbb{R}$, $\mathbb{R}$, and $\mathbb{R}$. The Lévy measure $\nu(dz)$ is a positive, possibly singular, Radon measure on $\mathbb{R}^M\setminus\{0\}$; precise assumptions will be given later.

The non-local term J^α can be a pseudo-differential operator. Specifying $\eta \equiv z$ and $\nu(dz) = \frac{K}{|z|^{N+\gamma}}\,dz$, $\gamma \in (0,2)$, gives rise to the fractional Laplace operator $J = (-\Delta)^{\gamma/2}$. The operator J^α is allowed to degenerate (vanish) since we allow $\eta = 0$ for $z \ne 0$. The second order differential operator L^α is also allowed to degenerate since we only assume that the diffusion matrix a^α is nonnegative definite. Due to these two types of degeneracies there is no (global) smoothing effect on solutions to (1.1) (neither "Laplacian" nor "fractional Laplacian" smoothing). Therefore equation (1.1) will not possess classical solutions in general. In view of the nonlinearity and degeneracy present in (1.1), the natural type of weak solutions are the viscosity solutions [21, 29]. For a precise definition of viscosity solution of (1.1) we refer to [31]. In this paper we will work with Hölder/Lipschitz continuous viscosity solution of (1.1)-(1.2). For other works on viscosity solutions and IPDEs of second order, we refer to [3, 4, 5, 7, 6, 10, 15, 31, 32, 43, 46] and references therein.

Nonlocal equations such as (1.1) appear as the dynamic programming equation associated with optimal control of jump-diffusion processes over a finite time horizon (see [43, 45, 12]). Examples of such control problems include various portfolio optimization problems in mathematical finance where the risky assets are driven by Lévy processes. The linear pricing equations for European and Asian options in Lévy markets are also of the form (1.1) if we take $\mathcal{A}$ to be a singleton. For more information on pricing theory and its relation to IPDEs we refer to [19].

For most nonlinear problems like (1.1)-(1.2), solutions must be computed by a numerical scheme. The construction and analysis of numerical schemes for nonlinear IPDEs is a relatively new area of research. Compared to the PDE case, there are currently only a few works available. Moreover, it is difficult to prove that such schemes converge to the correct (viscosity) solution. In the literature there are two main strategies for the discretization of the non-local term in (1.1). One is indirect in the sense that the Lévy measure is first truncated to obtain a finite measure and then the corresponding finite integral term is approximated by a quadrature rule. Regarding this strategy, we refer to [19, 20] (linear or obstacle problems) and [33, 17] (general nonlinear problems). The other approach is to discretize the integral term directly. Now there are 3 different cases to consider depending on whether (i) $\int_{|z|<1}\nu(dz) < \infty$, (ii) $\int_{|z|<1}|z|\nu(dz) < \infty$, or (iii) $\int_{|z|<1}|z|^2\nu(dz) < \infty$. Case (i) is the simplest one and has been considered by many authors, see, e.g., [38, 19, 14, 25, 2, 33] and references therein. Case (ii) was considered in [1, 41, 28], and case (iii) in [41, 28]. Most of the cited papers restrict their attention to linear, non-degenerate, one-dimensional equations or obstacle problems for such equations.

One of the contributions of this paper is a class of direct approximations of the non-local part of (1.1), giving rise to new monotone schemes that are capable of handling singular Lévy measures and moreover are supported by a theoretical analysis. The proposed schemes are new also in the linear case. As in [1] (cf. also [40] for a related approach), the underlying idea is to perform integration by parts to obtain a bounded "Lévy" measure and an integrand involving derivatives of the solution. In [1], one-dimensional, constant coefficients, linear equations (and obstacle problems) are discretized under the assumption $\int_{|z|<1} |z|\, \nu(dz) < \infty$. Their schemes are high-order and non-monotone, but not supported by rigorous stability and convergence results. In this paper we discretize general nonlinear, multi-dimensional, non-local equations without any additional restrictive integrability condition on the Lévy measure. More precisely, we provide monotone difference-quadrature schemes for (1.1)-(1.2) and prove under weak assumptions that these schemes converge with a rate to the exact viscosity solution of the underlying IPDE. The schemes we put forward and our convergence results apply in much more general situations than those previously treated in the literature.

The second main contribution of this paper is a theory of error estimates for a class of monotone approximations schemes for the initial value problem (1.1)-(1.2). We use this theory to derive error estimates for the proposed numerical schemes. For IPDEs in general and nonlinear IPDEs in particular, there are few error estimates available, see [20, 42] for linear equations and [33, 11, 17] for nonlinear equations.

Error estimates involving viscosity solutions first appeared in 1984 for first order PDEs [22], in 1997/2000 for convex 2nd order PDEs [34, 35], and in 2005/2008 for IPDEs [20, 33]. The results obtained for IPDEs, including those in this paper, are extensions of the results known for convex second order PDEs, which are based on Krylov's method of shaking the coefficients [35]. Krylov's method produces smooth approximate subsolutions of the equation (or scheme) that, via classical comparison and consistency arguments, imply one-sided error estimates. Based on this idea, there are currently two types of error estimates for convex second order PDEs: (i) optimal rates applying to specific schemes and equations (cf., e.g., [36, 30]) and (ii) sub-optimal rates that apply to "any" monotone consistent approximation (cf., e.g., [8, 35]). In particular, type (i) results apply when you have a priori regularity results for the scheme, while type (ii) results do not require this.

In this paper we provide error estimates of type (ii), whereas earlier results for IPDEs are of type (i), see [20, 33, 11, 17]. The problem with type (i) results is the difficulty in establishing the required priori regularity estimates. In the PDE case this can be achieved for particular schemes [36, 30], and attempts to generalize these schemes to the IPDE setting have only been partially successful [11, 17], since the required regularity estimates have been obtained only through unnaturally strong restrictions on the non-local terms. In [17] the Lévy measure is bounded and in [11] the Lévy measure is either bounded or the integral term is independent of x with an (essentially) one-dimensional Lévy measure. Of course, by a truncation procedure only bounded Lévy measures need to be considered [19, 33], but such approximations may not be accurate and the resulting error estimates blow up as the truncation parameter tends to zero. An advantage of the error estimates in the present paper is that they apply without any such restrictions. In particular, we can handle naturally any singular Lévy measures directly in our framework.

To prove our results we extend the approach of [8] to the non-local setting. To this end, we have to invoke a switching system approximation of (1.1) (see Section 6). Switching systems of this generality have not been studied before. In paper [12], we provide well-posedness, regularity, and continuous dependence results for such systems. We also prove that the value function of a combined switching and continuous control problem solve the switching system under consideration.

The remaining part of this paper is organized as follows: First of all, we shall end this introduction by listing some relevant notation. In Section 2 we list a few standing assumptions and provide corresponding well-posedness and regularity results for the IPDE problem (1.1)-(1.2). In Section 3 we present a rather general approximation scheme for this problem, and show that it is consistent, monotone, and convergent. Error estimates for general monotone approximation schemes are stated in Section 4. In Section 5 we present new direct discretizations of the non-local term in (1.1), and prove that these discretizations are consistent, monotone, and also satisfy the requirements introduced in Section 3. The switching system approximation of (1.1) is introduced and analyzed in Section 6. The obtained results are utilized in Section 7 to prove the error estimate stated Section 4. Finally, in Appendix A we give a standard example of a (monotone) discretization of the local PDE part of (1.1) that satisfies the requirements of Section 3.

We now introduce the notation that will be utilized in this paper. By C, K we mean various constants which may change from line to line. The Euclidean norm on any $\mathbb{R}^d$-type space is denoted by $|\cdot|$. For any subset $Q \subset \mathbb{R} \times \mathbb{R}^N$ and for any bounded, possibly vector valued, function on Q, we define the following norms,

$$|w|_0 := \sup_{(t,x)\in Q} |w(t,x)|, \qquad |w|_1 = |w|_0 + \sup_{(t,x)\neq(s,y)} \frac{|w(t,x) - w(t,y)|}{|t-s|^{\frac{1}{2}} + |x-y|}.$$

Note that if w is independent of t, then $|w|_1$ is the Lipschitz (or $W^{1,\infty}$) norm of w. We use $C_b(Q)$ to denote the space of bounded continuous real valued functions on Q. Let $\rho(t,x)$ be a smooth and non-negative function on $\mathbb{R} \times \mathbb{R}^N$ with unit mass and support in $\{0 < t < 1\} \times \{|x| < 1\}$. For any $\epsilon > 0$, we define the mollifier ρ_ϵ by

$$\rho_\epsilon(t,x) := \frac{1}{\epsilon^{N+2}} \rho\Big(\frac{t}{\epsilon^2}, \frac{x}{\epsilon}\Big). \tag{1.3}$$

In this paper we denote by h the vector

$$h = (\Delta t, \Delta x, \Delta z) > 0,$$

and any dependence on Δt, Δx, or Δz will be denoted by subscript h. The grid is denoted by $\mathcal{G}_h$ and is a subset of $\bar{Q}_T$ which need not be uniform or even discrete in general. We also set $\mathcal{G}_h^0 = \mathcal{G}_h \cap \{t = 0\}$ and $\mathcal{G}_h^+ = \mathcal{G}_h \cap \{t > 0\}$.

2. Well-posedness & regularity results for the Bellman equation

In this section we give some relevant well-posedness and regularity results for the Bellman equation (1.1)-(1.2). To this end, we impose the following assumptions:

(**A**.1) The control set $\mathcal{A}$ is a separable metric space. For any $\alpha \in \mathcal{A}$, $a^\alpha = \frac{1}{2}\sigma^\alpha \sigma^{\alpha T}$, and σ^α, b^α, $c^\alpha, f^\alpha, \eta^\alpha$ are continuous in α for all x, t, z.

(**A**.2) There is a positive constant K such that for all $\alpha \in \mathcal{A}$,

$$|g|_1 + |\sigma^\alpha|_1 + |b^\alpha|_1 + |c^\alpha|_1 + |f^\alpha|_1 \leq K.$$

(**A**.3) For every $\alpha \in \mathcal{A}$ and $z \in \mathbb{R}^M$ there is an $\Lambda \geq 0$ such that

$$|e^{-\Lambda|z|}\eta^\alpha(\cdot,\cdot,z)|_1 \leq K(|z| \wedge 1) \quad \text{and} \quad |e^{-\Lambda|\cdot|}\eta^\alpha(t,x,\cdot)|_1 \leq K.$$

(**A**.4) ν is a positive Radon measure on $\mathbb{R}^M \setminus \{0\}$ satisfying

$$\int_{0<|z|\leq 1} |z|^2 \nu(dz) + \int_{|z|\geq 1} e^{(\Lambda+\epsilon)|z|}\nu(dz) \leq K$$

for some $K \geq 0$, $\epsilon > 0$ where Λ is defined in (**A**.3).

Sometimes we need the following stronger assumptions than (**A**.3) and (**A**.4):

(**A**.4') ν is a positive Radon measure having a density $k(z)$ satisfying

$$0 \leq k(z) \leq \frac{e^{-(\Lambda+\varepsilon)|z|}}{|z|^{M+\gamma}} \quad \text{for all} \quad z \in \mathbb{R}^M \setminus \{0\},$$

for some $\gamma \in (0,2)$, $\varepsilon > 0$, where Λ is defined in (**A**.3).

(**A**.5) Assume that (**A**.3) holds and let γ as in (**A**.4'). There is a constant K such that for every $\alpha \in \mathcal{A}$ and $z \in \mathbb{R}^M$

$$|D_z^k \eta^\alpha(\cdot,\cdot,z)|_0 + |D_x^l \eta^\alpha(\cdot,\cdot,z)|_0 \le K e^{\Lambda|z|},$$

for all

$$\begin{array}{ll} k = l = 1 & \text{when } \gamma = 0, \\ k, l \in \{1,2\} & \text{when } \gamma \in (0,1), \\ k \in \{1,2,3,4\},\ l \in \{1,2\} & \text{when } \gamma \in [1,2). \end{array}$$

Assumptions (**A**.1)–(**A**.4) are standard and general. The assumptions on the non-local term are motivated by applications in finance. Almost all Lévy models in finance are covered by these assumptions. It is easy to modify the results in this paper so that they apply to IPDEs under different assumptions on the Lévy measures, e.g., to IPDEs of fractional Laplace type where there is no exponential decay of the Lévy measure at infinity. Finally, assumption (A.5) is not strictly speaking needed in this paper. We use it in some results because it simplifies some of our error estimates.

Under these assumptions the following results hold:

Proposition 2.1. *Assume (**A**.1)–(**A**.4).*

(a) There exists a unique bounded viscosity solution u of the initial value problem (1.1)–(1.2) *satisfying $|u|_1 < \infty$.*

(b) If u_1 and u_2 are respectively viscosity sub and supersolutions of (1.1) *satisfying $u_1(0,\cdot) \le u_2(0,\cdot)$, then $u_1 \le u_2$.*

The precise definition of viscosity solutions for the non-local problem (1.1)–(1.2) and the proof of Proposition 2.1 can be found in [31], for example.

3. Difference-Quadrature schemes for the Bellman equation

Now we explain how to discretize (1.1)–(1.2) by convergent monotone schemes on a uniform grid (for simplicity). We start with the spatial part and approximate L^α and J^α by monotone and consistent (see (3.3)–(3.6) below) approximations L_h^α and J_h^α on the spatial grid $\Delta x\, \mathbb{Z}^N$. We refer to Remark 3.1 (items a, b) for how to realize such approximations, see also Section 5 and Appendix A. The result is a system of ODEs in $\Delta x\, \mathbb{Z}^N \times (0,T)$:

$$u_t + \sup_{\alpha \in \mathcal{A}} \Big\{ - L_h^\alpha[u](t,x) + c^\alpha(t,x)u - f^\alpha(t,x) - J_h^\alpha[u](t,x) \Big\} = 0.$$

Then we discretize in the time variable using two separate θ-methods, one for the differential part and one for the integral part. For $\vartheta, \theta \in [0,1]$, the fully discrete scheme reads

$$\begin{aligned} U_\beta^n = U_\beta^{n-1} - \Delta t \sup_{\alpha \in \mathcal{A}} \Big\{ & -\theta L_h^\alpha[U]_\beta^n - (1-\theta) L_h^\alpha[U]_\beta^{n-1} + c_\beta^{\alpha,n-1} U_\beta^{n-1} \\ & - f_\beta^{\alpha,n-1} - \vartheta J_h^\alpha[U]_\beta^n - (1-\vartheta) J_h^\alpha[U]_\beta^{n-1} \Big\} \quad \text{in} \quad \mathcal{G}_h^+, \end{aligned} \tag{3.1}$$

$$U_\beta^0 = g(x_\beta) \quad \text{in} \quad \mathcal{G}_h^0, \tag{3.2}$$

where $\mathcal{G}_h = \Delta x \mathbb{Z}^N \times \Delta t\{0,1,2,\dots,\frac{T}{\Delta t}\}$ and $U_\beta^n = U(t_n, x_\beta)$, $f_\beta^{\alpha,n} = f^\alpha(t_n, x_\beta)$, etc., for $t_n = n\Delta t$ ($n \in \mathbb{N}_0$) and $x_\beta = \beta \Delta x$ ($\beta \in \mathbb{Z}^N$).

The approximations L_h and J_h are assumed to be consistent, meaning that

$$|L^\alpha[\phi] - L_h^\alpha[\phi]| \le K_L(|D^2\phi|_0 \Delta x + |D^4\phi|_0 \Delta x^2), \tag{3.3}$$

$$|J^\alpha[\phi] - J_h^\alpha[\phi]| \le K_J \Delta x \begin{cases} |D^2\phi|_0 & \text{when} \quad \gamma = [0,1), \\ (|D^2\phi|_0 + |D^4\phi|_0) & \text{when} \quad \gamma = [1,2), \end{cases} \tag{3.4}$$

for smooth bounded functions ϕ and where $\gamma \in (0,2)$ is defined in (A.4'). They are also assumed to be monotone in the sense that they can be written as

$$L_h^\alpha[\phi](t_n, x_{\bar\beta}) = \sum_{\beta\in\mathbb{Z}^N} l_{h,\beta,\bar\beta}^{\alpha,n}\big[\phi(t_n,x_\beta) - \phi(t_n,x_{\bar\beta})\big] \quad \text{with} \quad l_{h,\beta,\bar\beta}^{\alpha,n} \geq 0, \tag{3.5}$$

$$J_h^\alpha[\phi](t_n, x_{\bar\beta}) = \sum_{\beta\in\mathbb{Z}^N} j_{h,\beta,\bar\beta}^{\alpha,n}\big[\phi(t_n,x_\beta) - \phi(t_n,x_{\bar\beta})\big] \quad \text{with} \quad j_{h,\beta,\bar\beta}^{\alpha,n} \geq 0, \tag{3.6}$$

for any $\bar\beta \in \mathbb{Z}^N$ and $n \in \mathbb{N}_0$. We also assume without loss of generality that $j_{\cdot,\beta,\beta}^{\cdot,\cdot} = 0 = l_{\cdot,\beta,\beta}^{\cdot,\cdot}$ for all $\beta \in \mathbb{Z}^N$. The sum (3.5) is always finite, while the sum (3.6) is finite if the Lévy measure ν is compactly supported. With $\gamma \in [0,2)$ defined in (A.4') and $\Delta x < 1$, we also have that

$$\bar l_{\bar\beta}^{\alpha,n} := \sum_{\beta\in\mathbb{Z}^N} l_{h,\beta,\bar\beta}^{\alpha,n} \leq K_l \sup_\alpha \Big\{|a^\alpha|_0 \Delta x^{-2} + |b^\alpha|_0 \Delta x^{-1}\Big\}, \tag{3.7}$$

$$\bar j_{\bar\beta}^{\alpha,n} := \sum_{\beta\in\mathbb{Z}^N} j_{h,\beta,\bar\beta}^{\alpha,n} \leq K_j \Delta x^{-1}. \tag{3.8}$$

From (3.3) and (3.4) it immediately follows that the scheme (3.1) is a consistent approximation of (1.1), with the truncation error bounded by

$$\begin{aligned} \frac{1}{2}|\phi_{tt}|_0\Delta t + \sup_{\alpha,n}\Big\{&|L^\alpha[\phi]^n - L_h^\alpha[\phi]^n|_0 + |J^\alpha[\phi]^n - J_h^\alpha[\phi]^n|_0 \\ &+ (1-\theta)|L^\alpha[\phi]^{n-1} - L^\alpha[\phi]^n|_0 + (1-\vartheta)|J^\alpha[\phi]^{n-1} - J^\alpha[\phi]^n|_0\Big\}, \end{aligned} \tag{3.9}$$

for smooth functions ϕ. The last two terms are again bounded by

$$\Delta t \sup_\alpha \Big\{|L^\alpha[\phi_t]|_0 + |J^\alpha[\phi_t]|_0\Big\} \leq K\Delta t\Big\{|\partial_t D\phi|_0 + |\partial_t D^2\phi|_0\Big\}. \tag{3.10}$$

Under a CFL condition, the scheme (3.1) is also monotone, meaning that there are numbers $b_{\beta,\bar\beta}^{m,k}(\alpha) \geq 0$ such that it can be written as

$$\sup_\alpha \Big\{b_{\bar\beta,\bar\beta}^{n,n}(\alpha)U_{\bar\beta}^n - \sum_{\beta\neq\bar\beta} b_{\bar\beta,\beta}^{n,n}(\alpha)U_\beta^n - \sum_\beta b_{\bar\beta,\beta}^{n,n-1}(\alpha)U_\beta^{n-1} - \Delta t f_{\bar\beta}^{n-1,\alpha}\Big\} = 0, \tag{3.11}$$

for all $(x_{\bar\beta}, t_n) \in \mathcal{G}_h^+$. From (3.5) and (3.6), we see that

$$b_{\bar\beta,\bar\beta}^{n,m}(\alpha) = \begin{cases} 1 + \Delta t\theta\, \bar l_{\bar\beta}^{\alpha,m} + \Delta t\vartheta\, \bar j_{\bar\beta}^{\alpha,m} & \text{when } m = n, \\ 1 - \Delta t\Big[(1-\theta)\bar l_{\bar\beta}^{\alpha,m} + (1-\vartheta)\bar j_{\bar\beta}^{\alpha,m} - c_{\bar\beta}^{\alpha,m}\Big] & \text{when } m = n-1, \end{cases}$$

$$b_{\bar\beta,\beta}^{n,m}(\alpha) = \begin{cases} \Delta t\theta l_{h,\bar\beta,\beta}^{\alpha,m} + \Delta t\vartheta j_{h,\bar\beta,\beta}^{\alpha,m} & \text{when } m = n, \\ \Delta t(1-\theta) l_{h,\bar\beta,\beta}^{\alpha,m} + \Delta t(1-\vartheta) j_{h,\bar\beta,\beta}^{\alpha,m} & \text{when } m = n-1, \end{cases}$$

where $\bar\beta \neq \beta$ and other choices of m give zero. These coefficients are positive provided the following CFL condition holds:

$$\Delta t\Big[(1-\theta)\bar l_\beta^{\alpha,m} + (1-\vartheta)\bar j_\beta^{\alpha,m} - c_\beta^{\alpha,m}\Big] \leq 1 \quad \text{for all} \quad \alpha, \beta, m, \tag{3.12}$$

or alternatively if $a^\alpha \not\equiv 0$, $c^\alpha \geq 0$, $\Delta x < 1$, and (A4') hold, by (3.7) and (3.8),

$$\Delta t\Big[(1-\theta)K_l C\Delta x^{-2} + (1-\vartheta)K_j\Delta x^{-1}\Big] \leq 1.$$

Existence, uniqueness, and convergence results for the above approximation scheme are collected in the next theorem, while error estimates are postponed to Theorem 4.3 in Section 4.

Theorem 3.1. *Assume (**A**.1)–(**A**.3), (A.4'),* (3.3)–(3.8)*, and* (3.12)*.*
(a) There exists a unique bounded solution U_h of (3.1)–(3.2)*.*
(b) The scheme is L^∞-stable, i.e., $|U_h| \le e^{\sup_\alpha |c^\alpha|_0 t_n}\big[|g|_0 + t_n \sup_\alpha |f^\alpha|_0\big]$.
(c) U_h converge uniformly to the viscosity solution u of (1.1)–(1.2) *as $h \to 0$.*

Proof. The existence and uniqueness of bounded solutions follow by an induction argument. Consider $t = t_n$ and assume U^{n-1} is a given bounded function. For $\varepsilon > 0$ we define the operator $T : U^n \to U^n$ by

$$TU^n_\beta = U^n_\beta - \varepsilon \cdot (\text{left hand side of (3.11)}) \qquad \text{for all} \qquad \beta \in \mathbb{Z}^M.$$

Note that the fixed point equation $U^n = TU^n$ is equivalent to equation (3.1). Moreover, for sufficiently small ε, T is a contraction operator on the Banach space of bounded functions on $\Delta x\, \mathbb{Z}^N$ 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 $U^0 = g|_{\mathcal{G}_h^0}$ is bounded.

To see that T is a contraction we use the definition and sign of the b-coefficients:

$$\begin{aligned}
&TU^n_\beta - T\tilde{U}^n_\beta \\
&\le \sup_\alpha \Big\{ \Big[1 - \varepsilon[1 + \Delta t(\theta \bar{l}^{\alpha,n}_\beta + \vartheta \bar{j}^{\alpha,n}_\beta)]\Big](U^n_\beta - \tilde{U}^n_\beta) + \varepsilon \Delta t(\theta \bar{l}^{\alpha,n}_\beta + \vartheta \bar{j}^{\alpha,n}_\beta)|U^n_\cdot - \tilde{U}^n_\cdot|_0 \Big\} \\
&\le (1-\varepsilon)|U^n_\cdot - \tilde{U}^n_\cdot|_0,
\end{aligned}$$

provided $1 - \varepsilon(1 + \Delta t(\theta \bar{l}^{\alpha,n}_\beta + \vartheta \bar{j}^{\alpha,n}_\beta))] \ge 0$ for all α, β, n. Taking the supremum over all β and interchanging the role of U and $\tilde{U}$ proves that T is a contraction.

Much the same argument, utilizing (3.11), establishes that U_h is bounded by a constant independent of h:

$$|U^n|_0 \le (1 + \Delta t \sup_\alpha |c^\alpha|_0)^n \Big[|g|_0 + n\Delta t \sup_\alpha |f^\alpha|_0\Big] \le e^{\sup_\alpha |c^\alpha|_0 t_n}\Big[|g|_0 + t_n \sup_\alpha |f^\alpha|_0\Big].$$

In view of this bound, the convergence of U_h to the solution u of (1.1)–(1.2) follows by adapting the Barles-Souganidis argument [9] to the present non-local context. Alternatively, convergence follows from Theorem 4.3 if we also assume (A.5). □

Remark 3.1.

a. We will derive a new direct approximation J_h^α in Section 5 which satisfies the above assumptions, i.e., (3.4) and (3.6). The traditional approximation satisfying (3.6) and (a variant of) (3.4) is based on an indirect approach in which $\nu(dz)$ is replaced by the bounded measure $1_{|z|>\kappa}\nu(dz)$ and a small, suitably chosen viscosity term is added to the equation, cf., e.g., [20, 19, 33, 17] for details.

b. For the approximation L_h^α there are several known choices that satisfy (3.3) and (3.5). For example the scheme of Bonnans and Zidani [13] (see also Krylov [36]) or the standard schemes of Kushner [37, 39, 38]. We also mention that semi-Lagrangian type approximations L_h^α satisfy (3.5) and a (variant of) (3.3), cf., e.g., [16, 23, 24] and references therein. In Appendix A we show explicitly that one of the schemes of Kushner fall into our framework if a^α is diagonally dominant. If a^α is not diagonally dominant, then semi-Lagrangian and Bonnans-Zidani approximations still work but the convergence rate may be reduced. We refer the interested reader to [24] for thorough a discussion of these issues.

c. For the differential part, the choices $\theta = 0,\ 1$, and $1/2$ give explicit, implicit, and Crank-Nicholson discretizations. When $\vartheta > 0$, the integral term is evaluated fully implicitly. The implicit approach is not used much in practice since it leads to linear systems of equations with full matrices.

d. The CFL condition is needed for all the θ-schemes to be monotone except the purely implicit one. The CFL condition is determined by the local diffusion term, and if this term is not present (i.e., if $a \equiv 0$), then the CFL condition improves for all θ-schemes with explicit part. Also, if θ is close to 1 then the constants in the CFL condition (3.12) are very small, and hence the condition is less restrictive. Finally we mention that semi-Lagrangian diffusion approximations may have a CFL condition like $\Delta t \leq C\Delta x$ but at the cost of a lower rate of convergence [16, 24].

e. By parabolic regularity, "$D^2 \sim \partial_t$" and (3.10) is similar to $\Delta t|\phi_{tt}|_0$. When $\theta = 1/2 = \vartheta$ the scheme (3.1) (Crank-Nicholson!) is second order in time $\mathcal{O}(\Delta t^2)$ and (3.9) is no longer optimal.

f. When $\gamma = 0$ the leading error term in $J_h[u]$ (see (3.4)) comes from the difference approximation of the term $Du\int \eta\nu$. This approximation also gives rise to the term Δx^{-1} in (3.8).

4. Error estimates for general monotone approximations

In this section we present error estimates for general monotone approximation schemes for IPDEs. As a corollary we obtain an error estimate for the scheme (3.1)–(3.2) defined in Section 3. These results, which extend those in [8] to the non-local IPDE context, can be applied to "any" Lévy-type integro-differential operator. Earlier results apply to either linear problems, specific schemes, or restricted types of Lévy operators, see [42, 20, 33, 11]. In particular, previous error estimates do not apply to the approximation scheme (3.1) with the J_h^α defined in Section 5.

Let us write (1.1) as $u_t + F[u] = 0$ where $F[u] := F(t, x, u, Du, D^2u, u(t,\cdot))$ denotes the sup part of (1.1). We write approximations of $u_t + F[u] = 0$ as

$$S(h, t, x, u_h(t, x), [u_h]_{t,x}) = 0 \quad \text{in} \quad \mathcal{G}_h^+, \tag{4.1}$$

$$u_h(0, x) = g_h(x) \quad \text{in} \quad \mathcal{G}_h^0, \tag{4.2}$$

where S is the approximation of (1.1) defined on the mesh $\mathcal{G}_h \subset \bar{Q}_T$ with "mesh" parameter $h = (\Delta t, \Delta x, \Delta z)$ (time, space, quadrature parameters). The solution of (4.1) is termed u_h and by $[u_h]_{t,x}$ we denote a function defined at (t, x) in terms of the values taken by u_h evaluated at points other than (t, x). Note that the grid does not have to be uniform or even discrete.

We assume that (4.1) satisfies the following set of (very weak) assumptions:

(S1) **(Monotonicity)** There exist $\lambda, \mu \geq 0$, $h_0 > 0$ such that, if $|h| \leq h_0$, $u \leq v$ are functions in $C_b(\mathcal{G}_h)$ and $\phi(t) = e^{\mu t}(a + bt) + c$ for $a, b, c \geq 0$, then

$$S(h, t, x, r + \phi(t), [u + \phi]_{t,x}) \geq S(h, t, x, r, [v]_{t,x}) + \frac{b}{2} - \lambda c \quad \text{in } \mathcal{G}_h^+.$$

(S2) **(Regularity)** For each h and $\phi \in C_b(\mathcal{G}_h)$, the mapping

$$(t, x) \mapsto S\big(h, t, x, \phi(t, x), [\phi]_{t,x}\big)$$

is bounded and continuous in $\mathcal{G}_h^+$ and the function $r \mapsto S(h, t, x, r, [\phi]_{t,x})$ is uniformly continuous for bounded r, uniformly in t, x.

(S3) **(i)** **(Sub-consistency)** There exists a function $E_1(\tilde{K}, h, \epsilon)$ such that, for any sequence $\{\phi_\epsilon\}_\epsilon$ of smooth bounded functions satisfying

$$|\partial_t^{\beta_0} D^{\beta'} \phi_\epsilon| \leq \tilde{K}\epsilon^{1-2\beta_0-|\beta'|} \quad \text{in } \bar{Q}_T, \quad \text{for any } \beta_0 \in \mathbb{N}, \beta' \in \mathbb{N}^N,$$

where $|\beta'| = \sum_{i=1}^N \beta_i'$, the following inequality holds in $\mathcal{G}_h^+$:

$$S\big(h, t, x, \phi_\epsilon(t, x), [\phi_\epsilon]_{t,x}\big) \leq \phi_{\epsilon t} + F(t, x, \phi_\epsilon, D\phi_\epsilon, D^2\phi, \phi_\epsilon(t,\cdot)) + E_1(\tilde{K}, h, \epsilon).$$

(S3) **(ii)** **(Super-consistency)** There exists a function $E_2(\tilde{K}, h, \epsilon)$ such that, for any sequence $\{\phi_\epsilon\}_\epsilon$ of smooth bounded functions satisfying

$$|\partial_t^{\beta_0} D^{\beta'} \phi_\epsilon| \leq \tilde{K}\epsilon^{1-2\beta_0-|\beta'|} \quad \text{in } \bar{Q}_T, \quad \text{for any } \beta_0 \in \mathbb{N}, \beta' \in \mathbb{N}^N,$$

the following inequality holds in $\mathcal{G}_h^+$:

$$S\big(h,t,x,\phi_\epsilon(t,x),[\phi_\epsilon]_{t,x}\big) \geq \partial_t\phi_\epsilon + F(t,x,\phi_\epsilon,D\phi_\epsilon,D^2\phi,\phi_\epsilon(t,\cdot)) - E_2(\tilde{K},h,\epsilon).$$

Remark 4.1. In (S3), we typically take $\phi_\epsilon = w_\epsilon * \rho_\epsilon$ for some sequence $(w_\epsilon)_\epsilon$ of uniformly bounded and Lipschitz continuous functions, and ρ_ϵ is the mollifier defined in Section 1.

Remark 4.2. Assumption (S1) implies monotonicity in $[u]$ (take $\phi = 0$), and parabolicity of the scheme (4.1) (take $u = v$). This last point is easier to understand from the following more restrictive assumption:

(S1') (Monotonicity) There exist $\lambda \geq 0, \bar{K} > 0$ such that if $u \leq v; u,v \in C_b(\mathcal{G}_h)$ and $\phi : [0,T] \to \mathbb{R}$ smooth, then

$$\begin{aligned} &S(h,t,x,r+\phi(t),[u+\phi]_{t,x}) \\ &\geq S(h,t,x,r,[v]_{t,x}) + \phi'(t) - \bar{K}\Delta t|\phi''(t)|_0 - \lambda\phi^+(t). \end{aligned}$$

It is easy to see that (S1') implies (S1), cf. [8].

The main consequence of (S1) and (S2) is the following comparison principle satisfied by scheme (4.1) (for a proof, cf. [8]):

Lemma 4.1. *Assume (S1), (S2), $g_1, g_2 \in C_b(\mathcal{G}_h)$, and $u, v \in C_b(\mathcal{G}_h)$ satisfy*

$$S(h,t,x,u(t,x),[u]_{t,x}) \leq g_1 \quad \text{and} \quad S(h,t,x,v(t,x),[v]_{t,x}) \geq g_2 \quad \text{in} \quad \mathcal{G}_h^+.$$

Then, for λ and μ as in (S1),

$$u - v \leq e^{\mu t}|(u(0,\cdot)-v(0,\cdot))^+|_0 + 2te^t|(g_1-g_2)^+|_0.$$

The following theorem is our first main result.

Theorem 4.2 (Error Estimate)**.** *Assume (**A**.1)–(**A**.4), (S1), (S2) hold, and that the approximation scheme* (4.1)–(4.2) *has a unique solution $u_h \in C_b(\mathcal{G}_h)$, for each sufficiently small h. Let u be the exact solution of* (1.1)–(1.2).

a) **(Upper Bound)** *If (S3)(i) holds, then there exists a constant C, depending only on μ, K in (S1) and (**A**.2), such that*

$$u - u_h \leq e^{\mu t}|(g-g_h)^+|_0 + C\min_{\epsilon>0}\big(\epsilon + E_1(|u|_1,h,\epsilon)\big) \quad \text{in } \mathcal{G}_h.$$

b) **(Lower Bound)** *If (S3)(ii) holds, then there exists a constant C, depending only on μ, K in (S1) and (**A**.2), such that*

$$u - u_h \geq -e^{\mu t}|(g-g_h)^-|_0 - C\min_{\epsilon>0}\big(\epsilon^{\frac{1}{3}} + E_2(|u|_1,h,\epsilon)\big) \quad \text{in } \mathcal{G}_h.$$

We prove this theorem in Section 7.

Remark 4.3. Theorem 4.2 applies to all Lévy type non-local operators. Note that the lower bound is worse than the upper bound, and may not be optimal. In certain special cases it is possible to prove better bounds, however until now such results have only been obtained in the non-degenerate linear case [42, 20] or under very strong restrictions on the non-local term [11, 17]. More information on such non-symmetric error bounds can be found in [8].

Remark 4.4. For a finite difference-quadrature type discretization of (1.1), the truncation error would typically look like

$$\begin{aligned} &|\phi_t + F(t,x,\phi,D\phi,D^2\phi,\phi(t,\cdot)) - S(h,t,x,\phi(t,x),[\phi]_{t,x})| \\ &\leq K\sum_{\beta_0}|\partial_t^{\beta_0^0}D^{\beta_0'}\phi|_0\Delta t^{k_{\beta_0}} + K\sum_{\beta_1}|\partial_t^{\beta_1^0}D^{\beta_1'}\phi|_0\Delta x^{k_{\beta_1}} + K\sum_{\beta_2}|\partial_t^{\beta_2^0}D^{\beta_2'}\phi|_0\Delta z^{k_{\beta_2}}, \end{aligned}$$

where $\beta_0 = (\beta_0^0, \beta_0')$, $\beta_1 = (\beta_1^0, \beta_1')$, $\beta_2 = (\beta_2^0, \beta_2')$ are multi-indices and $k_{\beta_0}, k_{\beta_2}, k_{\beta_2}$ are real numbers. In this case, the function E in (S3) is obtained by taking $\phi := \phi_\epsilon$ in the above inequality:

$$E_1 = E_2 = \tilde{K}K \sum_{\beta_0,\beta_1,\beta_2} \Big[\epsilon^{1-2\beta_0^0-|\beta_0'|}\Delta t^{k_{\beta_0}} + \epsilon^{1-2\beta_1^0-|\beta_1'|}\Delta x^{k_{\beta_1}} + \epsilon^{1-2\beta_2^0-|\beta_2'|}\Delta z^{k_{\beta_2}}\Big].$$

An optimization with respect to ε yields the final convergence rate. Observe that the obtained rate reflects a potential lack of smoothness of the solution.

We shall now use Theorem 4.2 to prove error estimates for the finite difference-quadrature scheme (3.1).

Theorem 4.3. *Assume (**A**.1)–(**A**.3), (A.4'), (A.5),* (3.3)–(3.8), (3.12) *hold, and that u and U_h are the solutions respectively of* (1.1)–(1.2) *and* (3.1)–(3.2).

There are constants $K_L, K_J \geq 0$, $\delta > 0$ such that if $\Delta x \in (0,\delta)$ and Δt satisfies the CFL condition (3.12)*, then in $\mathcal{G}_h$,*

$$-K(\Delta t^{1/10} + \Delta x^{1/5}) \leq u - U_h \leq K(\Delta t^{1/4} + \Delta x^{1/2}) \quad \textit{for} \quad \gamma \in [0,1),$$

$$-K(\Delta t^{1/10} + \Delta x^{1/10}) \leq u - U_h \leq K(\Delta t^{1/4} + \Delta x^{1/4}) \quad \textit{for} \quad \gamma \in [1,2).$$

See Remark 3.1 (items a, b) for a discussion on approximations J_h^α and L_h^α that satisfy the assumptions of Theorem 4.3.

Proof of Theorem 4.3. Let us write the scheme (3.1) in abstract form (4.1). To this end, set $[u]_{t,x}(s,y) = u(t+s, x+y)$ and divide (3.11) by Δt to see that (3.1) takes the form (4.1) with

$$\begin{aligned} S(h, t_n, x_\beta, r, [u]_{t_n,x_\beta}) = \sup_{\alpha\in\mathcal{A}} \Bigg\{ &\frac{b_{\beta,\beta}^{n,n}(\alpha)}{\Delta t} r - \sum_{\bar\beta\neq\beta} \frac{b_{\beta,\bar\beta}^{n,n}(\alpha)}{\Delta t}[u]_{t_n,x_\beta}(0, x_{\bar\beta} - x_\beta) \\ &- \sum_{\bar\beta} \frac{b_{\beta,\bar\beta}^{n,n-1}(\alpha)}{\Delta t}[u]_{t_n,x_\beta}(-\Delta t, x_{\bar\beta} - x_\beta)\Bigg\}. \end{aligned}$$

By its definition (3.1), monotonicity (3.11), and consistency (3.9), this scheme obviously satisfies assumptions (S1) – (S3) if the CFL condition (3.12) holds. In particular, from (3.9) and (3.3), (3.4), (3.10), we find that

$$E_1(\tilde K, h, \varepsilon) = E_2(\tilde K, h, \varepsilon) = \begin{cases} C\tilde K(\Delta t\varepsilon^{-3} + \Delta x \varepsilon^{-1} + \Delta x^2\varepsilon^{-3}), & \gamma \in [0,1) \\ C\tilde K(\Delta t\varepsilon^{-3} + \Delta x \varepsilon^{-1} + (\Delta x^2 + \Delta x)\varepsilon^{-3}), & \gamma \in [1,2). \end{cases}$$

The result then follows from Theorem 4.2 and a minimization with respect to ε. □

Remark 4.5. The error estimate is independent of γ and robust in the sense that it applies to non-smooth solutions. Compared with analogous results for pure PDEs, see Theorem 4.1 in [8], the lower error bound in Theorem 4.3 remains unchanged whereas the upper bound is unchanged for $\gamma \in [0,1)$ but lower when $\gamma \in [1,2)$.

5. New monotone approximations of the non-local term

In this section we derive new direct approximations $J_h^\alpha[u]$ of the non-local integro term $J^\alpha[u]$ appearing in (1.1) that are monotone (3.6), consistent (3.4), and satisfy assumption (3.8). Hence they satisfy the assumptions of the convergence theory and error estimates of Sections 3 and 4. As in [1] (cf. also [40]), the idea is to perform integration by parts to reduce the singularity of the measure.

We consider 3 cases separately: (i) $\int_{|z|<1}\nu(dz)<\infty$, (ii) $\int_{|z|<1}|z|\nu(dz)<\infty$, and (iii) $\int_{|z|<1}|z|^2\nu(dz)<\infty$. Note that in cases (i) and (ii) we can write the non-local operator in the form

$$J^\alpha[\phi](t,x)=I^\alpha[\phi](t,x)-\bar b^\alpha(x)D\phi, \tag{5.1}$$

where

$$I^\alpha[\phi](t,x):=\int_{|z|>0}\Big(\phi(t,x+\eta^\alpha(t,x,z))-\phi\Big)\nu(dz),$$
$$\bar b^\alpha(x):=\int_{0<|z|<1}\eta^\alpha(t,x,z)\nu(dz),$$

for smooth bounded functions ϕ. The reason is that $I^\alpha[\phi]$ and $\bar b^\alpha(x)$ are well-defined under assumptions (A.2), (A.3), (A.4) if either (i) or (ii) holds. Furthermore, $\bar b^\alpha(x)$ will be bounded and x-Lipschitz. The term $\bar b^\alpha D\phi$ will be approximated by quadrature and upwind finite differences as in Appendix A leading to a first order method. We skip the standard details and focus on the non-local term $I^\alpha[\phi]$.

To simplify the presentation a bit, we will only consider the Cartesian x-grid $\{x_\beta\}_\beta=\Delta x\,\mathbb{Z}^N$, but it is possible to consider unstructured non-degenerate families of grids. On our grid we define a positive and 2nd order interpolation operator i_h, i.e., an operator satisfying

$$i_h\phi(x)=\sum_{\beta\in\mathbb{Z}^N}w_\beta(x)\phi(x_\beta)\quad\text{with}\quad w_\beta(x)\geq 0, \tag{5.2}$$

$$|E_I[\phi](x)|:=|\phi(x)-i_h\phi(x)|\leq K_I\Delta x^2|D^2\phi|_0, \tag{5.3}$$

for all $x\in\mathbb{R}^N$ and where $w_\beta(x)\geq 0$ are basis functions satisfying $w_\beta(x_{\bar\beta})=\delta_{\beta,\bar\beta}$ and $\sum_\beta w_\beta\equiv 1$. Linear and multi-linear interpolation satisfy these assumptions. Note that higher order interpolation is not monotone in general.

We will also need the following monotone difference operators:

$$\delta^\pm_{r,h}\phi(r,y)=\pm\frac{1}{\Delta x}\big\{\phi(r\pm\Delta x,y)-\phi(r,y)\big\}, \tag{5.4}$$

$$\Delta_{rr,k}\phi(r,y)=\frac{1}{k^2}\big\{\phi(r+k,y)-2\phi(r,y)+\phi(r-k,y)\big\}, \tag{5.5}$$

for functions $\phi(r,y)$ on $\mathbb{R}\times\mathbb{R}^K$ for some $K\in\mathbb{N}$. For smooth ϕ we have

$$|\delta^\pm_{r,h}\phi-\partial_r\phi|\leq\frac12|\phi_{rr}|_0\Delta x,\qquad|\Delta_{rr,k}\phi-\partial_r^2\phi|\leq\frac{1}{12}|\partial_r^4\phi|_0|k|^2.$$

5.1. **Finite Lévy measures.** Assuming $\int_{|z|<1}\nu(dz)<\infty$, we approximate the term $I^\alpha[\phi]$ defined in (5.1) by

$$I^\alpha_h[\phi](t,x)=Q_h\Big[(i_h\phi)(t,x+\eta^\alpha(t,x,z))-\phi(t,x)\Big],$$

where Q_h denotes a positive quadrature rule on the z-grid $\{z_\beta\}_\beta\subset\mathbb{R}^M$ with maximal grid spacing Δz, satisfying

$$Q_h[\phi]=\sum_{\beta\in\mathbb{Z}^M}\omega_\beta\phi(z_\beta)\quad\text{with}\quad\omega_\beta\geq 0,$$
$$|E_Q[\phi]|:=|\textstyle\int\phi(z)\nu(dz)-Q_h[\phi]|\leq K_Q\Delta z^{k_Q}|D^{k_Q}\phi|_0\textstyle\int\nu(dz),$$

for smooth bounded functions ϕ, where $K_Q\geq 0$ and $k_Q\in\mathbb{N}$. Many quadrature methods satisfies these requirements, e.g., compound Newton-Cotes methods of order less than 9 and Gauss methods of arbitrary order. Note that the z-grid does not have to be a Cartesian grid. This method is at most 2nd order accurate because

$$I^\alpha[\phi]=I^\alpha_h[\phi]+E_I[\phi(\cdot,\cdot+\eta^\alpha)]\textstyle\int\nu(dz)+E_Q[\phi(\cdot,\cdot+\eta^\alpha)-\phi],$$

and it is monotone by construction, satisfying (3.6) and (3.8). The $\mathcal{O}(\Delta x^{-1})$ term in (3.8) comes from the discretization of the $\bar{b}^\alpha$ term in (5.1).

5.2. **Unbounded Lévy measures I.** Now we assume that $\int_{|z|<1} |z|\nu(dz) < \infty$, or more precisely that (A.4') holds with $\gamma < 1$. We consider the one-dimensional and multi-dimensional cases separately.

5.2.1. *One-dimensional case* ($M = 1$). Now $I^\alpha[\phi]$ in (5.1) takes the form

$$I^\alpha[\phi](t,x) = \int_{\mathbb{R}\setminus\{0\}} \big[\phi(t, x + \eta^\alpha(t,x,z)) - \phi(t,x)\big]k(z)dz.$$

We approximate this term by

$$I_h^\alpha[\phi](t,x) = \sum_{n=0}^{\infty} \Big[\delta_{z,h}^{+}(i_h\phi)(t, x + \eta^\alpha(t,x,z_n))k_{h,n}^{+} - \delta_{z,h}^{-}(i_h\phi)(t, x + \eta^\alpha(t,x,z_{-n}))k_{h,n}^{-}\Big], \tag{5.6}$$

where $z_n = n\,\Delta x$, $\delta_{z,h}^{\pm}$ is defined in (5.4), the x-interpolation i_h satisfies (5.2) and (5.3). Moreover,

$$\begin{cases} k_{h,n}^{+} := \int_{z_n}^{z_{n+1}} \hat{k}(z)dz, \\ k_{h,n}^{-} := \int_{z_{-(n+1)}}^{z_{-n}} \hat{k}(z)dz \end{cases} \quad \text{and} \quad \hat{k}(z) := \begin{cases} \int_{-\infty}^{z} k(\zeta)\,d\zeta, & \text{if } \ z < 0, \\ \int_{z}^{\infty} k(\zeta)\,d\zeta, & \text{if } \ z > 0. \end{cases}$$

By (A.4') ($M = 1$ and $\gamma < 1$), $0 \le \int_{\mathbb{R}} \hat{k}(z)dz < \infty$.

To derive this approximation, the key idea is to perform integration by parts:

$$\begin{aligned} I^\alpha[\phi](t,x) &= \Big(\int_{-\infty}^{0} + \int_{0}^{\infty}\Big)\big(\phi(t, x + \eta^\alpha(t,x,z)) - \phi(t,x)\big)k(z)dz \\ &= \int_0^\infty \frac{\partial}{\partial z}\big(\phi(t, x+\eta^\alpha(t,x,z))\big)\hat{k}(z)dz - \int_{-\infty}^{0} \frac{\partial}{\partial z}\big(\phi(t, x+\eta^\alpha(t,x,z))\big)\hat{k}(z)dz, \end{aligned}$$

for bounded C^1 functions ϕ. Write $I^\alpha[\phi] = I^{\alpha,+}[\phi] + I^{\alpha,-}[\phi]$, and use quadrature, finite differencing, and interpolation to proceed as follows:

$$\begin{aligned} I^{\alpha,+}[\phi](t,x) &:= \int_0^\infty \partial_z\big[\phi(t, x + \eta^\alpha(t,x,z))\big]\hat{k}(z)dz \\ &\simeq \sum_{n=0}^{\infty} \partial_z\big[\phi(t, x+\eta^\alpha(t,x,z))\big]\big|_{z=z_n} k_{h,n}^{+} \\ &\simeq \sum_{n=0}^{\infty} \frac{\phi(t, x + \eta^\alpha(t,x,z_n + \Delta x)) - \phi(t, x+\eta^\alpha(t,x,z_n))}{\Delta x} k_{h,n}^{+} \\ &\simeq \sum_{n=0}^{\infty} \frac{(i_h\phi)(t, x + \eta^\alpha(t,x,z_n + \Delta x)) - (i_h\phi)(t, x+\eta^\alpha(t,x,z_n))}{\Delta x} k_{h,n}^{+}. \end{aligned}$$

In a similar way we can discretize $I^{\alpha,-}[\phi]$ and (5.6) follows.

The approximation just proposed is consistent since

$$I^\alpha[\phi](t,x) = I_h^\alpha[\phi](t,x) + E_Q + E_{\text{FDM}} + E_I,$$

where E_Q, E_{FDM}, and E_I denote respectively the error contributions from the approximation of the integral (1st order), the difference approximation (up-winding, 1st order), and the 2nd order interpolation. These terms can be estimated as follows:

$$|E_Q| \le \Delta x\,|\partial_z^2 \phi(\cdot + \eta^\alpha)|_0 \int_{\mathbb{R}} \hat{k}(z)dz,$$

$$|E_{\mathrm{FDM}}| \leq \frac{1}{2}\Delta x\,|\partial_z^2\phi(\cdot+\eta^\alpha)|_0 \int_{\mathbb{R}} \hat{k}(z)dz,$$
$$|E_I| \leq 2\Delta x\,|D_x^2\phi(\cdot+\eta^\alpha)|_0 \int_{\mathbb{R}} \hat{k}(z)dz.$$

The discretization (5.6) is also monotone satisfying (3.6) and also (3.8) (when I_h^α replaces J_h^α). To see this, note that $i_h\phi(x_{\bar\beta}) = \phi(x_{\bar\beta})$ and that by (A.2) $\eta(t,x,0)=0$. Hence we can reorganize the sum defining $I_h^{\alpha,+}$ and write

$$\begin{aligned}
&I_h^{\alpha,+}[\phi](t,x_{\bar\beta})\\
&\quad= -\frac{1}{\Delta x}k_{h,0}^+\phi(t,x_{\bar\beta}) + \frac{1}{\Delta x}\sum_{n=1}^{\infty}(k_{h,n-1}^+ - k_{h,n}^+)(i_h\phi)(t,x+\eta^\alpha(t,x_{\bar\beta},z_n))\\
&\quad= \frac{1}{\Delta x}\sum_{n=1}^{\infty}(k_{h,n-1}^+ - k_{h,n}^+)\big[(i_h\phi)(t,x_{\bar\beta}+\eta^\alpha(t,x_{\bar\beta},z_n)) - \phi(t,x_{\bar\beta})\big].
\end{aligned}$$

In a similar way

$$I_h^{\alpha,-}[\phi](t,x_{\bar\beta}) = \frac{1}{\Delta x}\sum_{n=1}^{\infty}(k_{h,n-1}^- - k_{h,n}^-)\big[(i_h\phi)(t,x_{\bar\beta}+\eta^\alpha(t,x_{\bar\beta},z_{-n})) - \phi(t,x_{\bar\beta})\big].$$

Since $\hat{k}$ is increasing on $(0,\infty)$ and decreasing on $(-\infty,0)$,

$$k_{h,n-1}^\pm > k_{h,n}^\pm,$$

and hence by (5.2) and $\sum_\beta w_\beta \equiv 1$, (3.6) and (3.8) hold with

$$\begin{aligned}
j_{h,\beta,\bar\beta}^{\alpha,n} &= \frac{1}{\Delta x}\sum_{l\in\mathbb{Z}\setminus\{0\}} w_\beta(x_{\bar\beta}+\eta^\alpha(t_n,x_{\bar\beta},z_l))(k_{h,|l|-1}^{\mathrm{sign}(l)} - k_{h,|l|}^{\mathrm{sign}(l)}) \geq 0,\\
\bar{j}_{\bar\beta}^{\alpha,n} &= \frac{1}{\Delta x}\sum_{l\in\mathbb{Z}\setminus\{0\}}(k_{h,|l|-1}^{\mathrm{sign}(l)} - k_{h,|l|}^{\mathrm{sign}(l)})\sum_\beta w_\beta(x_{\bar\beta}+\eta^\alpha(t_n,x_{\bar\beta},z_l)) = \frac{k_{h,0}^+ + k_{h,0}^-}{\Delta x},
\end{aligned}$$

and $k_{h,0}^\pm = \mathcal{O}(\Delta x^{1-\gamma})$. The leading $\mathcal{O}(\Delta x^{-1})$ term in (3.8) comes from discretizing the $\bar{b}^\alpha$ term in (5.1).

5.2.2. *Multi-dimensional case* ($M>1$). In this case we write $I_h^\alpha[\phi]$ of (5.1) in polar coordinates and propose the following approximation:

$$I_h^\alpha[\phi](t,x) = \int_{|y|=1}\sum_{n=0}^{\infty}\delta_{r,h}^+\big[i_h\phi(t,x+\eta^\alpha(t,x,r_ny))\big]k_{h,n}(y)dS_y, \tag{5.7}$$

where $r_n = n\,\Delta x$, dS_y is the surface measure on the unit sphere in $\mathbb{R}^M$, $\delta_{z,h}^\pm$ is defined in (5.4), the x-interpolation i_h satisfies (5.2) and (5.3). Moreover,

$$k_{h,n}(y) = \int_{r_n}^{r_{n+1}}\hat{k}(r,y)dr \qquad\text{and}\qquad \hat{k}(r,y) = \int_r^\infty k(sy)s^{M-1}ds.$$

By assumption (A.4') with $\gamma\in(0,1)$, $0\leq\int_0^\infty \hat{k}(r,y)dr \leq C < \infty$ for all $|y|=1$.

To derive this approximation we use polar coordinates and integrate by parts in the radial direction. Let ϕ be a bounded C^1 function, and set

$$G^\alpha(t,x,z) := \phi(t,x+\eta^\alpha(t,x,z)) - \phi(t,x).$$

Then

$$\begin{aligned}
I^\alpha[\phi](t,x) &= \int_{\mathbb{R}^M\setminus\{0\}} G^\alpha(t,x,z)k(z)dz\\
&= \int_{|y|=1}\Big[\int_0^\infty G^\alpha(t,x,ry)r^{M-1}k(ry)dr\Big]dS_y
\end{aligned}$$

$$= \int_{|y|=1} \Big[\int_0^\infty \frac{\partial}{\partial r} G^\alpha(t,x,ry)\hat{k}(r,y)dr \Big] dS_y,$$

and (5.7) follows by discretizing the inner integral as in Section 5.2.1.

This is a consistent first order approximation of $I^\alpha[\phi]$ since

$$I^\alpha[\phi](t,x) = I_h^\alpha[\phi](t,x) + E_Q + E_{\mathrm{FDM}} + E_I,$$

where E_Q, E_{FDM}, E_I have the same meaning as in Section 5.2.1, and these terms can be estimated as follows:

$$\begin{aligned}
|E_Q| &\le \Delta x\, |D_z^2 \phi(\cdot + \eta^\alpha)|_0 M_{\hat{k}},\\
|E_{\mathrm{FDM}}| &\le \frac{1}{2}\Delta x\, |D_z^2 \phi(\cdot + \eta^\alpha)|_0 M_{\hat{k}},\\
|E_I| &\le \Delta x\, |D_x^2 \phi(\cdot + \eta^\alpha)|_0 M_{\hat{k}},
\end{aligned}$$

where $M_{\hat{k}} = \int_{|y|=1} \int_0^\infty \hat{k}(r,y)dr\, dS_y$.

The approximation $I_h^\alpha[\phi]$ is also monotone, and satisfies (3.6) and (3.8). This follows as in Section 5.2.1, since $I_h^\alpha[\phi](t,x_{\bar\beta})$ can be written as

$$\frac{1}{\Delta x} \int_{|y|=1} \sum_{n=1}^\infty \big[k_{h,n-1}(y) - k_{h,n}(y)\big] \Big[(i_h\phi)(t, x_{\bar\beta} + \eta^\alpha(t,x_{\bar\beta}, r_n y)) - \phi(t,x_{\bar\beta})\Big] dS_y,$$

where for fixed y, $k_{h,n}(y)$ is a decreasing function in n since $\hat{k}(r,y)$ decreasing in r. Moreover, $\bar{l}_\beta^{\alpha,n}$ has a term like $\frac{1}{\Delta x} k_{h,0}(y) = \mathcal{O}(\Delta x^{-\gamma})$ plus the leading $\mathcal{O}(\Delta x^{-1})$ term which comes from the discretization of the $\bar{b}^\alpha$ term in (5.1).

5.3. **Unbounded Lévy measures II.** We assume that $\int_{|z|<1} |z|^2 \nu(z) dx < \infty$, or more precisely that (A.4') hold with $\gamma \in [1,2)$. In this case the decomposition (5.1) is not valid. Again, we consider the one-dimensional and multi-dimensional cases separately.

5.3.1. *One-dimensional Lévy process* ($M = 1$). Now the nonlocal operator takes the form

$$J^\alpha[\phi](t,x) = \int_{\mathbb{R}\setminus\{0\}} \big[\phi(t, x+\eta^\alpha(t,x,z)) - \phi(t,x) - \eta^\alpha(t,x,z)D\phi\big] k(z)dz,$$

or, after two integrations by parts (more details are given below),

$$J^\alpha[\phi](t,x) = J^{\alpha,+}[\phi](t,x) + J^{\alpha,-}[\phi](t,x) - \tilde{b}^\alpha(t,x)D\phi, \tag{5.8}$$

where $\tilde{b}^\alpha(t,x) = \int_{-\infty}^\infty \partial_z^2 \eta^\alpha(t,x,z)\tilde{k}(z)dz$,

$$J^{\alpha,\pm}[\phi] = \pm \int_0^{\pm\infty} \partial_z^2 \big[\phi(t, x+\eta^\alpha(t,x,z))\big]\tilde{k}(z)dz, \tag{5.9}$$

$$\tilde{k}(z) = \begin{cases} \int_{-\infty}^z \int_{-\infty}^w k(r)dr\, dw, & \text{for } z<0\\ \int_z^\infty \int_w^\infty k(r)dr\, dw, & \text{for } z>0. \end{cases}$$

By (A.4') ($M=1$, $\gamma<2$), $0 \le \tilde{k}(z) \le C|z|^{1-\gamma} e^{-(\Lambda+\varepsilon)|z|}$ and $\tilde{k}$ is integrable.

Note that $\tilde{b}^\alpha$ is bounded and x-Lipschitz, and that $\tilde{b}^\alpha D\phi$ can be discretized using quadratures and finite differences as in Appendix A. This leads to a first order monotone (upwind) approximation – we skip the standard details.

We propose the following approximation of $J^{\alpha,\pm}[\phi]$:

$$J_h^{\alpha,\pm}[\phi](t,x) = \sum_{n=0}^\infty \Delta_{zz,\Delta z}\big[i_h\phi(t, x+\eta^\alpha(t,x,z_n))\big]\tilde{k}_{h,n}^\pm, \tag{5.10}$$

where $z_n = n\Delta x$ (not $n\Delta z$!), $\Delta_{zz,\Delta z}$ is defined in (5.5), the x-interpolation i_h satisfies (5.2) and (5.3). Moreover,

$$\tilde{k}^+_{h,n} = \int_{z_n}^{z_{n+1}} \tilde{k}(z)dz \quad \text{and} \quad \tilde{k}^-_{h,n} = \int_{z_{-n-1}}^{z_{-n}} \tilde{k}(z)dz.$$

The approximation (5.10) can be derived from (5.9) using quadrature, finite differencing, and interpolation.

To obtain (5.9) and (5.8), we integrate by parts twice:

$$\begin{aligned}
&\int_0^\infty \big[\phi(t,x+\eta^\alpha(t,x,z)) - \phi(t,x) - \eta^\alpha(t,x,z)D\phi\big]k(z)dz \\
&= \Big[\big[\phi(t,x+\eta^\alpha(t,x,z)) - \phi(t,x) - \eta^\alpha(t,x,z)D\phi\big]\big(-\int_z^\infty k(w)dw\big)\Big]_{z=0}^{z=\infty} \\
&\qquad + \int_0^\infty \partial_z\big[\phi(t,x+\eta^\alpha(t,x,z)) - \eta^\alpha(t,x,z)D\phi\big]\big(\int_z^\infty k(w)dw\big)dz \\
&= 0 + \Big[\partial_z\big[\phi(t,x+\eta^\alpha(t,x,z)) - \eta^\alpha(t,x,z)D\phi\big]\big(-\tilde{k}\big)(z)\Big]_0^\infty \\
&\quad + \int_0^\infty \partial_z^2\big[\phi(t,x+\eta^\alpha(t,x,z)) - \eta^\alpha(t,x,z)D\phi\big]\tilde{k}(z)dz \\
&= 0 + 0 + \int_0^\infty \partial_z^2\big[\phi(t,x+\eta^\alpha(t,x,z))\big]\tilde{k}(z)dz - D\phi\int_0^\infty \partial_z^2\eta^\alpha(t,x,z)\tilde{k}(z)dz.
\end{aligned}$$

In view of this result and similar computations for the integral on $(-\infty,0)$, (5.8) follows. These computations are rigorous if $\phi(t,x+\eta), \partial_z\phi(t,x+\eta), \partial_z^2\phi(t,x+\eta)$ and $\eta, \partial_z\eta, \partial_z^2\eta$ are z-integrable and bounded by $e^{\Lambda|z|}$ at infinity.

The approximation is consistent and has the error expansion

$$J^{\alpha,\pm}[\phi](t,x) = J_h^{\alpha,\pm}[\phi](t,x) + E_Q^\pm + E_{\mathrm{FDM}}^\pm + E_I^\pm,$$

where E_Q, E_{FDM}, E_I have the same meaning as in Section 5.2.1, and these terms can be estimated as follows:

$$\begin{aligned}
|E_Q^\pm| &\le \Delta x\,|\partial_z^3\phi(\cdot+\eta^\alpha)|_0 \int_{\mathbb{R}} \tilde{k}(z)dz, \\
|E_{\mathrm{FDM}}^\pm| &\le \frac{1}{24}\Delta z^2\,|\partial_z^4\phi(\cdot+\eta^\alpha)|_0 \int_{\mathbb{R}} \tilde{k}(z)dz, \\
|E_I^\pm| &\le 4\frac{\Delta x^2}{\Delta z^2}\,|D_x^2\phi(\cdot+\eta^\alpha)|_0 \int_{\mathbb{R}} \tilde{k}(z)dz.
\end{aligned}$$

The proposed approximation is first order accurate if $\Delta z = \Delta x^{1/2}$, it is monotone satisfying (3.6), and (3.8) holds if $\Delta z = \Delta x^{1/2}$. These properties follow as in Section 5.2.1, since $J_h^{\alpha,\pm}[\phi](t,x_{\bar\beta})$ can be written as

$$\begin{aligned}
&\frac{1}{\Delta z^2}\tilde{k}^\pm_{h,0}\big[(i_h\phi)(t,x+\eta^\alpha(t,x_{\bar\beta},z_{\mp 1})) - \phi(t,x_{\bar\beta})\big] \\
&+ \frac{1}{\Delta z^2}\sum_{n=1}^\infty (\tilde{k}^\pm_{h,n+1} - 2\tilde{k}^\pm_{h,n} + \tilde{k}^\pm_{h,n-1})\big[(i_h\phi)(t,x+\eta^\alpha(t,x_{\bar\beta},z_{\pm n})) - \phi(t,x_{\bar\beta})\big],
\end{aligned}$$

and, by convexity of $\tilde{k}(z)$ on $(0,\infty)$ and $(-\infty,0)$,

$$\tilde{k}^\pm_{h,n+1} - 2\tilde{k}^\pm_{h,n} + \tilde{k}^\pm_{h,n-1} \ge 0 \qquad \text{for} \qquad n \ge 1.$$

Moreover, $\bar{j}_\beta^{\alpha,n}$ equals $\frac{2}{\Delta z^2}(\tilde{k}^+_{h,0} - \tilde{k}^+_{h,1} + \tilde{k}^-_{h,0} - \tilde{k}^-_{h,1}) = \mathcal{O}(\Delta x^{2-\gamma}/\Delta z^2)$ plus a $\mathcal{O}(\Delta x^{-1})$ term from the discretization of the $\tilde{b}^\alpha$-term in (5.8). When $\Delta z = \Delta x^{1/2}$ the leading term is the $\mathcal{O}(\Delta x^{-1})$ term.

5.3.2. *Multi-dimensional Lévy process* ($M > 1$). Writing $J^\alpha[\phi]$ in polar coordinates and performing two integrations by parts in the radial direction leads to

$$J^\alpha[\phi](t,x) = \tilde{J}^\alpha[\phi](t,x) - \tilde{b}^\alpha(t,x)D\phi, \tag{5.11}$$

where $\tilde{b}^\alpha(t,x) = \int_{|y|=1}\int_0^\infty \partial_r^2\big[\eta^\alpha(t,x,ry)\big]\tilde{k}(r,y)dr\,dS_y$ and

$$\tilde{J}^\alpha[\phi](t,x) = \int_{|y|=1}\int_0^\infty \partial_r^2\big[\phi(t,x+\eta^\alpha(t,x,ry))\big]\tilde{k}(r,y)dr\,dS_y,$$
$$\tilde{k}(s) = \int_s^\infty\int_w^\infty r^{M-1}k(ry)dr\,dw.$$

By (A.4') ($\gamma < 2$), $\tilde{k}(r,y) \le Cr^{1-\gamma}e^{-(\Lambda+\varepsilon)r}$ and thus $\tilde{k}$ is r-integrable uniformly in y. Note that $\tilde{b}^\alpha$ is bounded and x-Lipschitz, and that $\tilde{b}^\alpha D\phi$ can be discretized using quadrature and finite differencing as in Appendix A. This leads to a first order monotone (upwind) approximation – we skip the standard details.

We propose the following approximation of $\tilde{J}^\alpha[\phi]$:

$$\tilde{J}_h^\alpha[\phi](t,x) = \int_{|y|=1}\sum_{n=0}^\infty \Delta_{rr,\Delta z}\big[(i_h\phi)(t,x+\eta^\alpha(t,x,r_ny))\big]\tilde{k}_{h,n}(y)\,dS_y, \tag{5.12}$$

where $r_n = n\Delta x$ (not $n\Delta z$!), $\Delta_{rr,\Delta z}$ is defined in (5.5), the x-interpolation i_h satisfies (5.2) and (5.3), and

$$\tilde{k}_{h,n}(y) = \int_{r_n}^{r_{n+1}}\tilde{k}(r,y)dz.$$

The approximation (5.12) follows from (5.11) by quadrature, finite differencing, and interpolation, and the derivation of (5.11) is rigorous provided the functions $\phi(t,x+\eta), D_z\phi(t,x+\eta), D_z^2\phi(t,x+\eta)$ and $\eta, D_z\eta, D_z^2\eta$ are z-integrable and bounded by $e^{\Lambda|z|}$ at infinity.

The approximation is consistent and has the error expansion

$$\tilde{J}^\alpha[\phi](t,x) = \tilde{J}_h^\alpha[\phi](t,x) + E_Q + E_{\mathrm{FDM}} + E_I,$$

where E_Q, E_{FDM}, E_I have the same meaning as in Section 5.3.1, and can be estimated as follows:

$$|E_Q| \le \Delta x\,|D_z^3\phi(\cdot+\eta^\alpha)|_0 M_{\tilde{k}},$$
$$|E_{\mathrm{FDM}}| \le \frac{1}{24}\Delta z^2\,|D_z^4\phi(\cdot+\eta^\alpha)|_0 M_{\tilde{k}},$$
$$|E_I| \le 4\frac{\Delta x^2}{\Delta z^2}\,|D_x^2\phi(\cdot+\eta^\alpha)|_0 M_{\tilde{k}},$$

where $M_{\tilde{k}} := \int_{|y|=1}\int_0^\infty \tilde{k}(r,y)dr\,dS_y$. Whenever $\Delta z = \Delta x^{1/2}$, this is a first order approximation. Moreover, the approximation is monotone satisfying (3.6) and, whenever $\Delta z = \Delta x^{1/2}$, it also satisfies (3.8). This follows as in Section 5.2.1 since $\tilde{J}_h^\alpha[\phi](t,x_{\bar\beta})$ can be written as an integral over $\{|y| = 1\}$ with integrand

$$\frac{1}{\Delta z^2}\big[(i_h\phi)(t,x_{\bar\beta}+\eta^\alpha(t,x_{\bar\beta},r_{-1}y)) - \phi(t,x_{\bar\beta})\big]\tilde{k}_{h,0}(y)$$
$$+\frac{1}{\Delta z^2}\sum_{n=1}^\infty\big(\tilde{k}_{h,n+1}(y) - 2\tilde{k}_{h,n}(y) + \tilde{k}_{h,n-1}(y)\big)$$
$$\times\Big[(i_h\phi)(t,x_{\bar\beta}+\eta^\alpha(t,x_{\bar\beta},r_ny)) - \phi(t,x_{\bar\beta})\Big].$$

Furthermore, for each fixed y, $\tilde{k}(r,y)$ is convex on $(0,\infty)$ and thus

$$\tilde{k}_{h,n+1}(y) - 2\tilde{k}_{h,n}(y) + \tilde{k}_{h,n-1}(y) \ge 0 \qquad \text{for } n \ge 1 \text{ and } |y| = 1.$$

Remark 5.1.

a. (Order of schemes) In general, our discretizations of the non-local term in (1.1) are at most first order accurate. In the case $\gamma \in [1,2)$, a first order rate is obtained by choosing $\Delta z = \Delta x^{1/2}$. Higher order discretizations can be derived using higher order quadrature and interpolation rules, but the resulting discretizations are not monotone in general. On the other hand, if $\eta \equiv z$, then interpolation is not needed and consequently the monotone discretizations of Section 5.3 are 2nd order accurate.

b. (Remaining discretizations) To obtain fully discrete schemes it remains to discretize the various terms involving $\tilde{b}^\alpha D\phi$, for example by quadrature and finite differencing, cf. Appendix A. In applications, the densities $\hat{k}$ and $\tilde{k}$ can often be explicitly calculated, e.g., using incomplete gamma functions as in [1]. Otherwise these quantities also have to be computed by quadrature. Furthermore, regarding Sections 5.2.2 and 5.3.2, it also remains to discretize the surface integral in y. This discretization does not pose any problems, neither numerically nor in the analysis, as long as positive quadratures are used. The details are left to the reader.

c. (Increasing efficiency) From a practical point of view in terms computational efficiency, quadratures should be implemented using FFT. This is standard and we refer to, e.g., [25] for the details.

d. (Generalization I) The above approximations (with obvious modifications) also apply to integral terms of the type

$$J^\alpha[\phi](t,x) = \sum_{i=1}^{M} \int_{\mathbb{R}\setminus\{0\}} \big(\phi(t,x+\eta_i^\alpha(t,x,z)) - \phi(t,x) - \eta_i^\alpha(t,x,z)D\phi(t,x)\big)k_i(z)dz.$$

Such terms appear in M-dimensional Lévy models based on M independent Poisson random measures coming from one-dimensional Lévy processes. This is a rich class of models with many applications. For more information and analysis of such models we refer to the book [44].

e. (Generalization II) With obvious modifications, our approximations also apply to linear and nonlinear equations involving the fractional Laplace operator

$$(-\Delta)^\alpha u(x) = c_\alpha \int_{|z|>0} \frac{u(x+z)-u(x)-zDu(x)}{|z|^{N+2\alpha}}dz, \qquad \alpha \in (0,1),$$

where $x,z \in \mathbb{R}^N$ and c_α is a constant, in which case the Lévy measure takes the form $\nu(dz) = |z|^{-N-2\alpha}dz$. This measure satisfies (**A**.3) except for the "exponential decay at infinity" requirement. It is straightforward to recast the entire theory to allow for a fractional Laplace setting where assumption (A.3) is replaced by

$$\int_{|z|>0} |z|^2 \wedge 1\, \nu(dz) < \infty.$$

6. Error estimates for a switching system approximation

In this section we derive error estimates for a switching system approximation of (1.1)–(1.2). This result, which has independent interest, plays a crucial role in the proof of Theorem 4.2 in Section 7.

The switching system will be written as

$$F_i(t,x,v,\partial_t v^i, Dv^i, D^2v^i, u^i(t,\cdot)) = 0 \quad \text{in} \quad Q_T, \quad i \in \{1,2,.....,m\}, \tag{6.1}$$

$$v(0,x) = (g(x),\dots,g(x)) \quad \text{in} \quad \mathbb{R}^N, \tag{6.2}$$

where $v = (v_1,\dots,v_m)$ is in $\mathbb{R}^m$ and for sets $\mathcal{A}_i$ such that $\cup_i \mathcal{A}_i = \mathcal{A}$,

$$F_i(t,x,r,p_t,p_x,X,\phi(\cdot))$$

$$= \max\big\{p_t + \sup_{\alpha\in\mathcal{A}_i}\big[\mathcal{L}^\alpha(t,x,r_i,p_x,X) - J^\alpha[\phi](t,x)\big]; r_i - \mathcal{M}^i r\big\},$$
$$\mathcal{L}^\alpha(t,x,r,p,X) := -\mathrm{tr}\big(a^\alpha(t,x)X\big) - b^\alpha(t,x)\cdot p + c^\alpha(t,x)r - f^\alpha(t,x),$$
$$\mathcal{M}^i r = \min_{j\neq i}\{r_j + k\}, \quad k > 0,$$

for $\big(t,x,r,p_t,p_x,X\big) \in \mathbb{R}\times\mathbb{R}^N\times\mathbb{R}^m\times\mathbb{R}\times\mathbb{R}^N\times\mathbb{S}^N$ and any smooth bounded function ϕ. The operator $J^\alpha[\phi]$ is defined below (1.1).

In the pure PDE case, such approximations have been studied in, e.g., [18, 27, 8]. Here we extend the error estimates of [8] to non-local Bellman equations. In a complimentary article [12], we develop a viscosity solution theory covering switching systems like (6.1). We refer to that paper for the precise definition of viscosity solutions and proofs of the associated results utilized herein. If assumptions (**A**.1) – (**A**.4) of Section 3 hold, then we have the following well posedness result [12]:

Proposition 6.1. *Assume that conditions (**A**.1)–(**A**.4) hold. There exists a unique viscosity solution v of* (6.1)–(6.2)*, satisfying $|v|_1 \le C$ for a constant C depending only on T and K from (**A**.1)–(**A**.3). Furthermore, if w_1, w_2 are respectively viscosity sub and supersolutions of* (6.1) *satisfying $w_1(0,\cdot) \le w_2(0,\cdot)$, then $w_1 \le w_2$.*

Before we continue, we need the following remark.

Remark 6.1. The functions $\sigma^\alpha, b^\alpha, c^\alpha, f^\alpha, \eta^\alpha$ are only defined for times $t\in[0,T]$. But they can be easily extended to times $[-r, T+r]$ for any $r>0$ in such a way that (**A**.1) – (**A**.3) still hold. In view of Proposition 6.1 we can then solve the initial value problem up to time $T+r$ or, by using a translation in time, we may start from time $-r$. We will use these facts several times below.

By equi-continuity and the Arzela-Ascoli theorem it easily follows that the each component of the solution of (6.1)–(6.2) converges locally uniformly to the solution of (1.1)–(1.2) as $k\to 0$. To derive an error estimate we use Krylov's method of shaking the coefficients coupled with an idea of P.-L. Lions as in [8]. We need the following auxiliary system

$$F_i^\epsilon(t,x,v^\epsilon,\partial_t v_i^\epsilon, Dv_i^\epsilon, D^2 v_i^\epsilon, v_i^\epsilon(t,\cdot)) = 0 \quad \text{in} \quad Q_{T+\epsilon^2}, \quad i\in\{1,\dots,m\}, \tag{6.3}$$
$$v^\epsilon(0,x) = (g(x),\dots,g(x)) \quad \text{in} \quad \mathbb{R}^N,$$

where $v^\epsilon = (v_1^\epsilon,\dots,v_m^\epsilon)$ and

$$F_i^\epsilon(t,x,r,p_t,p_x,X,\phi(\cdot)) = \max\big\{p_t + \sup_{\alpha\in\mathcal{A}_i; |e|\le\epsilon; 0\le s\le\epsilon^2}\big(\mathcal{L}^\alpha(t+s,x+e,r_i,p_x,X) - J^\alpha(t+s,x+e)\phi\big); r_i - \mathcal{M}^i r\big\}.$$

The operators $\mathcal{L}$, J, and $\mathcal{M}$ are as previously defined.

Note that we have used the extension of the data mentioned in Remark 6.1. By regularity and continuous dependence results from [12] we have

Proposition 6.2. *Assume that (**A**.1)–(**A**.4) hold. There exists a unique viscosity solution $v^\epsilon : \bar{Q}_{T+\epsilon^2}\to\mathbb{R}$ of* (6.3) *satisfying*

$$|v^\epsilon|_1 + \frac{1}{\epsilon}|v^\epsilon - v|_0 \le C,$$

where C depends T and K. Furthermore, if w_1 and w_2 are respectively sub and supersolutions of (6.3) *satisfying $w_1(0,\cdot)\le w_2(0,\cdot)$, then $w_1\le w_2$.*

We are now in a position to prove the following main result of this section:

Theorem 6.3. *Assume that (**A**.1)–(**A**.4) hold. If u and v are respectively viscosity solutions of* (1.1)–(1.2) *and* (6.1)–(6.2)*, then for sufficiently small k,*

$$0 \le v_i - u \le Ck^{\frac{1}{3}}, \qquad i\in\{1,\dots,m\},$$

where C depends only on K and T.

Proof. Since $w = (u, \dots, u)$ is a viscosity subsolution of (6.1), the first inequality $u \le v_i$ follows from the comparison principle.

The second inequality will be obtained in the following. Since v^ϵ is the viscosity solution of (6.3), it follows that

$$\partial_t v_i^\epsilon + \sup_{\alpha\in\mathcal{A}_i} \big(\mathcal{L}^\alpha(t+s, x+e, v_i^\epsilon(t,x), Dv_i^\epsilon, D^2 v_i^\epsilon) - J^\alpha(t+s, x+e) v_i^\epsilon\big) \le 0$$

in $Q_{T+\epsilon^2}$ in the viscosity sense, $i = 1, \dots, m$. After a change of variable, we conclude that for every $0 \le s \le \epsilon^2$ and $|e| \le \epsilon$, $v^\epsilon(t-s, x-e)$ is a viscosity subsolution of the uncoupled system

$$\partial_t w_i^\epsilon + \sup_{\alpha\in\mathcal{A}_i} \big(\mathcal{L}^\alpha(t, x, w_i^\epsilon, Dw_i^\epsilon, D^2 w_i^\epsilon) - J^\alpha(t,x) w_i^\epsilon\big) = 0 \quad \text{in} \quad Q_T^\epsilon, \tag{6.4}$$

where $Q_T^\epsilon := (\epsilon^2, T) \times \mathbb{R}^N$. Now set $v_\epsilon := v^\epsilon \star \rho_\epsilon$, where ρ_ϵ is the mollifier defined in (1.3). A Riemann-sum approximation shows that this function is the limit of convex combinations of viscosity subsolutions $v(t-s, x-e)$ of the convex system (6.4). Hence v_ϵ is also a viscosity subsolution of (6.4) (see the appendix of [33] for more details). On the other hand, since v^ϵ is a continuous subsolution of (6.3),

$$v_i^\epsilon \le \min_{j\ne i} v_j^\epsilon + k \quad \text{in } Q_{T+\epsilon^2},\ i \in \{1, \dots, m\}.$$

It follows that $\max_i v_i^\epsilon(t,x) - \min v_j^\epsilon(t,x) \le k$ in $Q_{T+\epsilon^2}$, and therefore

$$|v_i^\epsilon - v_j^\epsilon|_0 \le k, \qquad i, j \in \{1, \dots, m\}.$$

Then, by the definition and properties of v_ϵ,

$$|\partial_t v_{\epsilon i} - \partial_t v_{\epsilon j}|_0 \le C\frac{k}{\epsilon^2} \qquad \text{and} \qquad |D^n v_{\epsilon i} - D^n v_{\epsilon j}|_0 \le C\frac{k}{\epsilon^n},$$

for $n \in \mathbb{N}$, $i, j \in \{1, \dots, m\}$, where C depends only on ρ, T, and K. For $\epsilon < 1$, it follows that

$$\begin{aligned} |\partial_t v_{\epsilon j} &+ \sup_{\alpha\in\mathcal{A}_i} \big(\mathcal{L}^\alpha(t, x, v_{\epsilon j}(t,x), Dv_{\epsilon j}, D^2 v_{\epsilon j}) - J^\alpha(t,x) v_{\epsilon j}\big) \\ &- \partial_t v_{\epsilon i} - \sup_{\alpha\in\mathcal{A}_i} \big(\mathcal{L}^\alpha(t, x, v_{\epsilon i}(t,x), Dv_{\epsilon i}, D^2 v_{\epsilon i}) - J^\alpha(t,x) v_{\epsilon i}\big)| \le C\frac{k}{\epsilon^2} \end{aligned}$$

and, since v_ϵ is subsolution of (6.4),

$$\partial_t v_{\epsilon i} + \sup_{\alpha\in\mathcal{A}} \big(\mathcal{L}^\alpha(t, x, v_{\epsilon i}(t,x), Dv_{\epsilon i}, D^2 v_{\epsilon i}) - J^\alpha(t,x) v_{\epsilon i}\big) \le C\frac{k}{\epsilon^2} \quad \text{in } Q_T^\epsilon,$$

where the constant C depends on ρ, T, and K. From this inequality it is easy to see that $v_{\epsilon i} - te^{Kt} C\frac{k}{\epsilon^2}$ is a subsolution of (1.1) restricted to Q_T^ϵ. Hence, by the comparison principle,

$$v_{\epsilon i} - u \le e^{Kt}\Big(|v_{\epsilon i}(\epsilon^2, \cdot) - u(\epsilon^2, \cdot)|_0 + Ct\frac{k}{\epsilon^2}\Big) \quad \text{in } Q_T^\epsilon,\ i \in \{1, \dots, m\}.$$

By regularity in time, $|u(t,\cdot) - v_i(t,\cdot)|_0 \le (|u|_1 + |v_i|_1)\epsilon$, and by Proposition 6.2 and properties of mollification we conclude that

$$v_i - u \le v_i - v_{\epsilon i} + v_{\epsilon i} - u \le C(\epsilon + \frac{k}{\epsilon^2}) \quad \text{in } Q_T,\ i \in \{1, \dots, m\}.$$

Now the theorem follows by minimizing with respect to ϵ. $\square$

7. The Proof of Theorem 4.2

To prove Theorem 4.2 we will use different arguments for the upper and lower bounds. The upper bound, part (a), is the "easy" part, and it is essentially a reformulation of the general upper bound established in [33]. We skip the details, and prove only part (b) which is a new result.

Proof of Theorem 4.2 (b). Without loss of generality we will assume that $\mathcal{A}$ is finite:

$$\mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_m\}.$$

The proof of this statement is similar to the one given in [8] in the pure PDE case and relies on assumption ($\mathbf{A}$.1). Now we follow [8] and use a switching system approximation to construct approximate supersolutions of (1.1) which are pointwise minima of smooth functions and approximates the viscosity solution of (1.1)–(1.2). Consider

$$\begin{aligned} F_i^\epsilon\big(t, x, v^\epsilon, \partial_t v_i^\epsilon, Dv_i^\epsilon, D^2 v_i^\epsilon, v_i^\epsilon(t,\cdot)\big) = 0 \quad \text{in } Q_{T+2\epsilon^2} \\ v^\epsilon(0,x) = v_0(x) \quad \text{in} \quad \mathbb{R}^N, \end{aligned} \tag{7.1}$$

where $v^\epsilon = (v_1^\epsilon, \dots, v_m^\epsilon)$, $v_0 = (g, \dots, g)$, and

$$\begin{aligned} &F_i^\epsilon\big(t, x, r, p_t, p_x, X, \phi(t,\cdot)\big) \\ &= \max\big\{p_t + \min_{0\le s\le \epsilon^2, |e|\le\epsilon} \big(\mathcal{L}^{\alpha_i}\big(t+s-\epsilon^2, x+e, r, p_t, p_x, X\big) \\ &\qquad - J^{\alpha_i}(t+s-\epsilon^2, x+e)\phi\big); r_i - \mathcal{M}^i r\big\}, \end{aligned}$$

where $\mathcal{L}, J$ and $\mathcal{M}$ are defined below (6.1) in Section 6. This new problem is well-posed and each component of the solution of this switching system will converge to the viscosity solution of (1.1) as $k, \epsilon \to 0$:

Lemma 7.1. *Assume that conditions ($\mathbf{A}$.1)–($\mathbf{A}$.4) hold. There exists a unique solution v^ϵ of* (7.1) *satisfying*

$$|v^\epsilon|_1 \le \bar{K}, \quad \max_{i,j} |v_i^\epsilon - v_j^\epsilon|_0 \le k, \quad \text{and for } k \text{ small, } \max_i |u - v_i^\epsilon| \le C(\epsilon + k^{\frac{1}{3}}),$$

where $\bar{K}, C$ only depend on T and K from ($\mathbf{A}$.2)–($\mathbf{A}$.4).

Proof. From [12] we have the existence and uniqueness of a viscosity solution, and moreover the uniform bounds

$$|v^\epsilon|_1 \le \bar{K} \quad \text{and} \quad |v^\epsilon - v^0|_0 \le C\epsilon,$$

where v^0 is the unique viscosity solution of (7.1) corresponding to $\epsilon = 0$. The last inequality in the lemma now follows since $|u - v_i^0|_0 \le Ck^{\frac{1}{3}}$ by Theorem 6.3. To second inequality follows since arguing as in the proof of Theorem 6.3 leads to $0 \le \max_i v_i^\epsilon - \min_j v_j^\epsilon \le k \quad \text{in} \quad Q_{T+2\epsilon^2}$. □

Next we time-shift and mollify v^ϵ. For $i = 1, \dots, m$, set

$$\bar{v}_i^\epsilon(t,x) := v_i^\epsilon(t+\epsilon^2, x), \qquad v_{\epsilon i}(t,x) := \rho_\epsilon * \bar{v}_i^\epsilon(t,x),$$

where ρ is defined in (1.3). Note that $\mathrm{supp}(\rho_\varepsilon) \subset (0,\varepsilon^2) \times B(0,\varepsilon)$ and that the functions v^ϵ, $\bar{v}^\epsilon$, v_ϵ are well-defined respectively on $Q_{T+2\varepsilon^2}$, $(-\epsilon^2, T+\epsilon^2] \times \mathbb{R}^N$, $Q_{T+\epsilon^2}$. By Lemma 7.1 and properties of mollifiers,

$$\begin{aligned} &|\bar{v}^\epsilon|_1 \le \bar{K}, \quad |\bar{v}^\epsilon - v^\epsilon|_0 \le \bar{K}\epsilon, \\ &\max_{i,j} |v_{\epsilon i} - v_{\epsilon j}| \le C(k+\epsilon) \quad \text{in} \quad Q_{T+\epsilon^2}, \\ &\max_i |u - v_{\epsilon,i}| \le C(\epsilon + k^{\frac{1}{3}}) \quad \text{in} \quad Q_T, \end{aligned} \tag{7.2}$$

where C depends only on ρ and K, T from (**A**.2)–(**A**.4). A supersolution of (1.1) can now be produced by setting

$$w := \min_i v_{\epsilon i}.$$

Lemma 7.2. *Assume that conditions (**A**.1)—(**A**.4) hold and* $\epsilon \le (8\sup_i [v_i^\epsilon]_1)^{-1} k$. *For every* $(t,x) \in Q_T$*, if* $j := \mathrm{argmin}_i v_{\epsilon i}(t,x)$*,*

$$\partial_t v_{\epsilon j} + \mathcal{L}^{\alpha_j}\big(t, x, v_{\epsilon j}(t,x), Dv_{\epsilon j}(t,x), D^2 v_{\epsilon j}(t,x)\big) - J^{\alpha_j}(t,x) v_{\epsilon j} \ge 0. \tag{7.3}$$

We postpone the proof of this lemma. From this lemma it follows that w is an approximate supersolution to the scheme (4.1) when $\epsilon \le (8\sup_i [v_i^\epsilon]_1)^{-1} k$:

$$S\big(h, t, x, w(t,x), [w]_{t,x}\big) \ge -E_2(\bar{K}, h, \epsilon) \quad \text{in} \quad \mathcal{G}_h^+, \tag{7.4}$$

where $\bar{K}$ comes from Lemma 7.1. To see this, let $(t,x) \in Q_T$ and set $j := \mathrm{argmin}_i v_{\epsilon i}(t,x)$. At (t,x), $w(t,x) = v_{\epsilon j}(t,x)$ and $w \le v_{\epsilon j}$ in $\mathcal{G}_h$. Hence (S1) implies that

$$S\big(h, t, x, w(t,x), [w]_{t,x}\big) \ge S\big(h, t, x, v_{\epsilon j}(t,x), [v_{\epsilon j}]_{t,x}\big).$$

By consistency (S3)(ii) we have

$$\begin{aligned}
&S\big(h, t, x, v_{\epsilon j}(t,x), [v_{\epsilon j}]_{t,x}\big)\\
&\ge \partial_t v_{\epsilon j} + F\big(t, x, v_{\epsilon j}(t,x), Dv_{\epsilon j}, D^2 v_{\epsilon j}, v_{\epsilon j}(t,\cdot)\big) - E_2(\bar{K}, h, \epsilon),\\
&\ge \partial_t v_{\epsilon j} + \mathcal{L}^{\alpha_j}\big(t, x, v_{\epsilon j}(t,x), Dv_{\epsilon j}(t,x), D^2 v_{\epsilon j}(t,x)\big) - J^{\alpha_j}(t,x) v_{\epsilon j} - E_2(\bar{K}, h, \epsilon),
\end{aligned}$$

and (7.4) then follows from Lemma 7.2.

To derive the lower bound on the error $u_h - u$, we take $\epsilon = (8\sup_i [v_i^\epsilon]_1)^{-1} k$ and use (7.4) and comparison Lemma 4.1 to get

$$u_h - w \le e^{\mu t} |(g_h - w(0,\cdot))^+| + 2te^{\mu t} E_2(\bar{K}, h, \epsilon) \quad \text{in } \mathcal{G}_h.$$

By (7.2), $|w - u| \le C(\epsilon + k + k^{\frac{1}{3}})$, and hence

$$u_h - u \le e^{\mu t} |(g_h - w(0,\cdot))^+| + 2te^{\mu t} E_2(\bar{K}, h, \epsilon) + C(\epsilon + k + k^{\frac{1}{3}}) \quad \text{in} \quad \mathcal{G}_h,$$

possibly with a new constant C. Since $\epsilon = Ck$, the proof is complete by minimizing the right hand side with respect to ϵ. □

Proof of Lemma 7.2. We begin by fixing $(t,x) \in Q_T$ and set $j = \mathrm{argmin}_i\, v_{\epsilon i}(t,x)$. Then

$$v_{\epsilon j}(t,x) - \mathcal{M}^j v_\epsilon(t,x) = \max_{i \ne j} \big\{v_{\epsilon j}(t,x) - v_{\epsilon i} - k\big\} \le -k.$$

Therefore, by the Hölder continuity of $\bar{v}^\epsilon$ and basic properties of mollifiers,

$$\bar{v}_j^\epsilon(t,x) - \mathcal{M}^j \bar{v}^\epsilon(t,x) \le -k + 2\max_i [v_i^\epsilon]_1 2\epsilon$$

and

$$\bar{v}_j^\epsilon(s,y) - \mathcal{M}^j \bar{v}^\epsilon(s,y) \le -k + 2\max_i [v_i^\epsilon]_1 (2\epsilon + |x-y| + |t-s|^{\frac{1}{2}}),$$

for all $(s,y) \in Q_T$. Consequently, if $|x-y| < \epsilon$, $|t-s| < \epsilon^2$, and $\epsilon \le (8\sup_i [v_i^\epsilon]_1)^{-1} k$, then

$$\bar{v}_j^\epsilon(s,y) - \mathcal{M}^j \bar{v}^\epsilon(s,y) < 0. \tag{7.5}$$

To continue we need the following remark. Let $u^1, \dots, u^k$ be functions satisfying (7.5) at (t,x), then any linear combination $u^\lambda = \sum_{i=1}^k \lambda_i u^i$ with $\lambda_i \ge 0$ and $\sum_{i=1}^k \lambda_i = 1$, also satisfies (7.5) at (t,x). If, in addition, $u^1, \dots, u^k$ are supersolutions of

$$\max\big\{\partial_t u_j + \mathcal{L}^{\alpha_j}(t, x, u_j, Du_j, D^2 u_j) - J^{\alpha_j}[u_j](t,x); u_j - \mathcal{M}^j u\big\} = 0, \tag{7.6}$$

then in view of (7.5) they are also supersolutions of the linear equation

$$\partial_t u_j + \mathcal{L}^{\alpha_j}(t,x,u_j,Du_j,D^2u_j) - J^{\alpha_j}[u_j](t,x) = 0 \tag{7.7}$$

at (t,x). An easy adaptation of the proof of Lemma 6.3 in [33] then shows that u^λ is also a viscosity supersolution of (7.7) at (t,x).

A change of variables reveals that $\{\bar{v}^\epsilon(\cdot - s, \cdot - e)\}_{s,e}$, $0 \le s < \epsilon^2$, $|e| < \epsilon$, is a family of supersolutions of (7.6) with $[\bar{v}^\epsilon_j - \mathcal{M}^j\bar{v}^\epsilon](t-s,x-e) < 0$. We note that by approximating the function $v_{\epsilon j}$ by a Riemann sum, we see that it is the limit of convex combinations of $\bar{v}^\epsilon(\cdot - s, \cdot - e)$. In view of the above remark, these convex combinations are supersolutions of (7.6), and hence by the stability result for viscosity supersolutions, so is the limit $v_{\epsilon j}$. Finally, since this function is smooth it is also a classical supersolution of (7.6) and hence (7.3) holds. □

Appendix A. An example of a monotone discretization of L^α

Let $\{e_i\}_{i=1}^N$ be the standard basis of $\mathbb{R}^N$ and a_{ij} the ij-th element of the matrix a. The following discretization of L^α in (1.1) can be found in [38] (see also [37, 39]):

$$L_h^\alpha\phi := \sum_{i=1}^N \Big[a_{ii}^\alpha\Delta_{ii} + \sum_{i\neq j}\big(a_{ij}^{\alpha+}\Delta_{ij}^+ - a_{ij}^{\alpha-}\Delta_{ij}^-\big) + b_i^{\alpha+}\delta_i^+ - b_i^{\alpha-}\delta_i^-\Big]\phi,$$

where $b^+ = \max\{b,0\}$, $b^- = (-b)^+$, and

$$\begin{aligned}
\delta_i^\pm\phi(x) &= \pm\frac{1}{\Delta x}\big\{\phi(x\pm e_i\Delta x) - \phi(x)\big\},\\
\Delta_{ii}\phi(x) &= \frac{1}{\Delta x^2}\big\{\phi(x+e_i\Delta x) - 2\phi(x) + \phi(x-e_i\Delta x)\big\},\\
\Delta_{ij}^+\phi(x) &= \frac{1}{2\Delta x^2}\big\{2\phi(x) + \phi(x+e_i\Delta x + e_j\Delta x) + \phi(x-e_i\Delta x - e_j\Delta x)\big\}\\
&\quad - \frac{1}{2\Delta x^2}\big\{\phi(x+e_i\Delta x) + \phi(x-e_i\Delta x) + \phi(x+e_j\Delta x) + \phi(x-e_j\Delta x)\big\}\\
\Delta_{ij}^-\phi(x) &= \frac{1}{2\Delta x^2}\big\{2\phi(x) + \phi(x+e_i\Delta x - e_j\Delta x) + \phi(x-e_i\Delta x + e_j\Delta x)\big\}\\
&\quad - \frac{1}{2\Delta x^2}\big\{\phi(x+e_i\Delta x) + \phi(x-e_i\Delta x) + \phi(x+e_j\Delta x) + \phi(x-e_j\Delta x)\big\}.
\end{aligned}$$

By Taylor expansion it is easy to check that the truncation error is given by (3.3). Moreover L_h can be written in the form (3.5) with

$$l_{h,\beta,\beta\pm e_i}^{\alpha,n} = \frac{1}{\Delta x^2}\Big[a_{ii}^\alpha(t,x) - \frac{1}{2}\sum_{j\neq i}|a_{ij}^\alpha(t,x)|\Big] + \frac{b_i^{\alpha\pm}(t,x)}{\Delta x},$$

$$l_{h,\beta,\beta+e_i\pm e_j}^{\alpha,n} = \frac{a_{ii}^{\alpha\pm}(t,x)}{\Delta x^2}, \qquad l_{h,\beta,\beta-e_i\pm e_j}^{\alpha,n} = \frac{a_{ii}^{\alpha\pm}(t,x)}{\Delta x^2}, \qquad i\neq j,$$

and $l_{h,\beta,\bar\beta}^{\alpha,n} = 0$ otherwise. This approximation is monotone if $l_{h,\beta,\bar\beta}^{\alpha,n} \ge 0$ for all $\alpha,\beta,\bar\beta,n$ and $\Delta x > 0$, which happens to be the case, e.g., if a is diagonally dominant:

$$a_{ii}^\alpha(t,x) - \sum_{j\neq i}|a_{ij}^\alpha(t,x)| \ge 0 \quad \text{in } Q_T, \text{ for each } \alpha\in\mathcal{A}.$$

We refer to Remark 3.1 b for a discussion of other approximations that monotone even when a is not diagonally dominant.

References

[1] A. Almendral and C. W. Oosterlee Accurate evaluation of European and American options under the CGMY process. *SIAM J. Sci. Comput.* 29 (2007), no. 1.

[2] A. Almendral and C. W. Oosterlee On American options under the variance gamma process. Appl. Math. Finance 14 (2007), no. 2.

[3] O. Alvarez and A. Tourin. Viscosity solutions of nonlinear integro-differential equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 13(3):293–317, 1996.

[4] A. L. Amadori. Nonlinear integro-differential evolution problems arising in option pricing: a viscosity solutions approach. *Differential Integral Equations*, 16(7):787–811, 2003.

[5] G. Barles, R. Buckdahn, and E. Pardoux. Backward stochastic differential equations and integral-partial differential equations. *Stochastics Stochastics Rep.*, 60(1-2):57–83, 1997.

[6] G. Barles, E. Chasseigne and C. Imbert. On the Dirichlet Problem for Second-Order Elliptic Integro-Differential Equations *Indiana Univ. Math. J.* 57 (2008), no. 1, 213–246.

[7] G. Barles and C. Imbert. Second-Order Elliptic Integro-Differential Equations: Viscosity Solutions' Theory Revisited. *Ann. Inst. H. Poincar Anal. Non Linaire* 25 (2008), 567–585.

[8] 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.

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

[10] F. E. Benth, K. H. Karlsen, and K. Reikvam. Optimal portfolio selection with consumption and nonlinear integrodifferential equations with gradient constraint: a viscosity solution approach. *Finance and Stochastic*, 5:275–303, 2001.

[11] I. H. Biswas, E. R. Jakobsen, and K. H. Karlsen. Error estimate for finite difference schemes for parabolic integro-PDEs associated with controlled jump-diffusions. *J. Hyperbolic Differ. Equ.* 5(1): 187-219, 2008.

[12] I. H. Biswas, E. R. Jakobsen, and K. H. Karlsen. Viscosity solutions for a system of integro-PDEs and connections to optimal switching and control of jump-diffusion processes. *Applied Mathematics and Optimisation.*, DOI 10.1007/s00245-009-9095-8.

[13] J. F. Bonnans and H. Zidani. Consistency of generalized finite difference schemes for the stochastic HJB equation. *SIAM J. Numer. Anal.* 41 (2003), no. 3, 1008–1021.

[14] M. Briani, C. La Chioma, and R. Natalini. Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in financial theory. *Numer. Math.* 98 (2004), no. 4.

[15] L. Caffarelli and L. Silvestre. Regularity theory for fully nonlinear integro-differential equations. *Comm. Pure Appl. Math.* 62 (2009), no. 5, 597–638.

[16] 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.

[17] F. Camilli and E. R. Jakobsen. A Finite Element like Scheme for Integro-Partial Differential Hamilton-Jacobi-Bellman Equations. *SIAM J. Numer. Anal.* 47 (2009), no. 4, 2407–2431.

[18] I. Capuzzo Dolcetta and L.C. Evans Optimal switching for ordinary differential equations *SIAM J. Cont. Optim*, 22(365-389), 1984.

[19] R. Cont and P. Tankov. *Finacial modeling with jump processes.* Chapman & Hall/CRC, Boca Raton, FL, 2004.

[20] R. Cont and E. Voltchkova. A finite difference scheme for option pricing in jump diffusion and exponential Lvy models. *SIAM J. Numer. Anal.* 43 (2005), no. 4, 1596–1626.

[21] 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.

[22] M. G. Crandall and P.-L. Lions. Two approximations of solutions of Hamilton-Jacobi equations. *Math. Comp.* 43 (1984), no. 167.

[23] 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.

[24] K. Debrabant and E. Jakobsen. Semi-Lagrangian schemes for linear and fully nonlinear diffusion equations. Preprint 2009: http://arxiv.org/abs/0910.1046.

[25] Y. d'Halluin, P. A. Forsyth, and K. R. Vetzal Robust numerical methods for contingent claims under jump diffusion processes. *IMA J. Numer. Anal.* 25 (2005), no. 1.

[26] 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.

[27] L.C. Evans and A. Friedman. Optimal stochastic switching and the dirichlet problem for the Bellman equation. *Trans. Amer. Math.Soc.*, 253: 365-389, 1979.

[28] W. Farkas, N. Reich, and C. Schwab. Anisotropic stable Lévy copula processes—analytical and numerical aspects. *Math. Models Methods Appl. Sci.* 17 (2007), no. 9.

[29] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions.* Springer-Verlag, New York, 1993.
[30] 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.*, 13(5):613–644, 2003.
[31] E. R. Jakobsen and K. H. Karlsen. Continuous dependence estimates for viscosity solutions of integro-PDEs. *J. Differential Equations*, 212(2):278–318, 2005.
[32] E. R. Jakobsen and K. H. Karlsen. A "maximum principle for semicontinuous functions" applicable to integro-partial differential equations. *NoDEA Nonlinear Differential Equations Appl.*, 49(3,4): 249-273, 2006.
[33] E. R. Jakobsen, K. H. Karlsen, and C. La Chioma. Error estimates for approximate solutions to Bellman equations associated with controlled jump-diffusions. *Numer. Math.*, 110(2): 221-255, 2008.
[34] N. V. Krylov. On the rate of convergence of finite-difference approximations for Bellman's equations. *Algebra i Analiz*, 9(3):245–256, 1997.
[35] N. V. Krylov. On the rate of convergence of finite-difference approximations for Bellman's equations with variable coefficients. *Probab. Theory Related Fields*, 117(1):1–16, 2000.
[36] N. V. Krylov. The rate of convergence of finite-difference approximations for Bellman equations with Lipschitz coefficients. *Appl. Math. Optim.*, 52(3):365–399, 2005.
[37] H. J. Kushner and C. F. Yu. Approximation, existence, and Numerical Procedures for optimal stochastic controls. *J. Math. Anal. Appl.* 45(1974), 563-587.
[38] H. J. Kushner and P. G. Dupuis. *Numerical methods for stochastic control problems in continuous time*. Springer-Verlag, New York, 1992.
[39] P. L. Lions and B. Mercier. Approximation numérique équations de Hamilton-Jacobi-Bellman. *RAIRO Analyse numérique* 14(1980), 369-393.
[40] V. E. Lynch, B. A. Carreras, D. del Castillo-Negrete, K. M. Ferreira-Mejias, and H. R. Hicks. Numerical methods for the solution of partial differential equations of fractional order. *J. Comput. Phys.*, 192(2):406–421, 2003.
[41] A. M. Matache, P. A. Nitsche and C. Schwab. Wavelet Galerkin pricing of American options on Lévy driven assets. *Quant. Finance* 5 (2005), no. 4, 403–424.
[42] A. M. Matache, T. P. Whiler and C. Schwab. Linear complexity solution of parabolic integro-differential equations. *Numer. Math.* 104 (2006), no. 1, 69–102.
[43] R. Mikulyavichyus and G. Pragarauskas. Nonlinear potentials of the Cauchy-Dirichlet problem for the Bellman integro-differential equation. *Liet. Mat. Rink.*, 36(2):178–218, 1996.
[44] B. Øksendal and A. Sulem. *Applied Stochastic Control of Jump Diffusions.* Springer, 2005.
[45] H. Pham. Optimal stopping of controlled jump diffusion processes: a viscosity solution approach. *J. Math. Systems Estim. Control*, 8(1), 1998.
[46] A. Sayah. Équations d'Hamilton-Jacobi du premier ordre avec termes intégro-différentiels. II. *Comm. Partial Differential Equations*, 16(6-7):1075–1093, 1991.

(Imran H. Biswas)
Centre for Applicable Mathematics, Tata Institute of Fundamental Research, Bangalore 560065, India
E-mail address: `imran@math.tifrbng.res.in`

(Espen R. Jakobsen)
Norwegian University of Science and Technology, NO–7491, Trondheim, Norway
E-mail address: `erj@math.ntnu.no`

(Kenneth Hvistendahl Karlsen)
Centre of Mathematics for Applications, University of Oslo, P.O. Box 1053, Blindern, NO–0316 Oslo, Norway
E-mail address: `kennethk@math.uio.no`
URL: `folk.uio.no/kennethk`