A FINITE ELEMENT LIKE SCHEME FOR INTEGRO-PARTIAL DIFFERENTIAL HAMILTON-JACOBI-BELLMAN EQUATIONS

FABIO CAMILLI AND ESPEN R. JAKOBSEN

Abstract. We construct a finite element like scheme for fully non-linear integro-partial differential equations arising in optimal control of jump-processes. Special cases of these equations include optimal portfolio and option pricing equations in Finance. The schemes are monotone and robust. We prove that they converge in very general situations, including degenerate equations, multiple dimensions, relatively low regularity of the data, and for most (if not all) types of jump-models used in Finance. In all cases we provide (probably optimal) error bounds. These bounds apply when grids are unstructured and integral terms are very singular, two features that are new or highly unusual in this setting.

1. Introduction

In this paper we introduce and analyze finite element (FEM) like schemes for nonlocal Hamilton-Jacobi-Bellman equations (HJB equations) of the form

$$\sup_{v\in V}\big\{-\operatorname{tr}[a(x,v)D^2u]-b(x,v)Du+c(x,v)u-f(x,v)-\mathcal{I}^v u(x)\big\}=0 \quad \text{in} \quad \mathbb{R}^N, \tag{1.1}$$

or

$$\sup_{v\in V}\big\{-\operatorname{tr}[a(x,v)D^2u]-b(x,v)Du+c(x,v)u-f(x,v)-\mathcal{J}^v u(x)\big\}=0 \quad \text{in} \quad \mathbb{R}^N, \tag{1.2}$$

where

$$a(x,v)=\frac{1}{2}\sigma(x,v)\sigma^T(x,v), \tag{1.3}$$

$$\mathcal{I}^v u(x)=\int_E [u(x+\eta(x,v,z))-u(x)]\nu(dz), \tag{1.4}$$

$$\mathcal{J}^v u(x)=\int_E [u(x+\eta(x,v,z))-u(x)-1_{|z|<1}\eta(x,v,z)Du(x)]\nu(dz), \tag{1.5}$$

V is a compact metric space and $E=\mathbb{R}^M\setminus\{0\}$. The coefficients σ,η,b,c,f are loosely speaking Lipschitz continuous in x, continuous in v, and Borel measurable in z. The precise assumptions will be given in the next section. These equations are the dynamic programming equations for stochastic control problems involving Levy processes, a class of Markov processes with jumps [11]. The two forms of the equation correspond to different intensities $\nu(dz)$ of small jumps in the corresponding Levy processes. In general these equations are degenerate and fully-nonlinear, and the corresponding solutions are typically only Hölder continuous and have to be understood in the viscosity sense [25, 29, 4].

Equations like (1.1) and (1.2) appear in advanced models of financial markets where price evolution of stocks (and other risky assets) are modeled as (exponential) pure-jump or jump-diffusion processes. Special cases of (1.1) and (1.2) include linear equations used in option pricing problems, obstacle problems used e.g. in pricing of American options, and fully non-linear equations (the full HJB equation) used in optimal portfolio problems. We refer to [19] for the first two cases and

Date: March 24, 2009.

2000 *Mathematics Subject Classification.* Primary 45K05 ; Secondary 65N15, 49L25.

Key words and phrases. Integro-partial differential equation, viscosity solution, numerical scheme, Lévy process, error estimate.

The research was supported by the Research Council of Norway through the project "Integro-PDEs: Numerical methods, Analysis, and Applications to Finance".

to [25, 8] for the last case. It is well-known that the standard Black-Scholes model (a diffusion model) give poor fit to real markets, at least on smaller time scales. E.g. log-returns distributions of stock prices are leptokurtic and have longer and fatter tails than predicted by the Black-Scholes model, see e.g. [19]. To improve upon these shortcomings, many pure-jump and jump-diffusion models have been suggested in the literature over the years, see e.g. [19, 40] for the most popular models. The empirical fact that Levy processes with discontinuous sample paths tend to better model e.g. stock prices, is one main reason for the renewed interest in stochastic optimal control of jump-diffusion processes.

Except in very simple cases equations (1.1) and (1.2) do not have closed form solutions, so numerical methods are needed to obtain (approximate) solutions. In this paper we construct FEM like schemes. Since the equations are fully nonlinear we have no weak/variational formulation, and the usual FEM approach does not work. If the measure ν is finite and the corresponding HJB equation is (1.1), we discretize in two steps. First we obtain a semi-discrete equation ("discrete in time") as the dynamic programming equation of a discrete time control problem approximating the underlying continuous time control problems of (1.1):

$$u_h(x) = \inf_{v\in V}\Big\{hf(x,v) + e^{-hc(x,v)}\Big[\frac{e^{-h\lambda}}{2d}\sum_{m=1}^{d}\Big(u_h(x+hb(x,v)+\sqrt{h}\sigma_m(x,v))+ + u_h(x+hb(x,v)-\sqrt{h}\sigma_m(x,v)\Big) + \frac{1-e^{-h\lambda}}{\lambda}\int_E u_h(x+\eta(x,v,z))\nu(dz)\Big]\Big\}, \tag{1.6}$$

where σ_m is the m-th column of the matrix σ, h is the discretization step and λ is the mass of the measure. The next step is obtain a fully discrete equation by introducing a regular triangulation and look for continuous piecewise linear functions over the chosen triangulation satisfying the semi-discrete equation at every vertex of the triangulation. When ν is not finite we first approximate it by a finite measure $1_{r<|z|<R}\nu(dz)$ (truncation), and then approximate the truncated equation following the above approach. To improve the truncation approximation, we also add small diffusion and/or drift terms to the equation.

We prove that these methods converge and derive (probably optimal) error bounds using the framework developed in [30] and ideas from [17]. We also discuss issues like restricting to a bounded domain, truncating long jumps, and approximating integrals by quadrature. In all cases we provide rigorous error bounds for the various approximations. What remains to do in a computer implementation, is the resolution of the non-linearity. There are various ways to do that, e.g. policy iteration, value iteration, artificial time methods (stationary limits), and so on. We will not address this point in this paper.

FEM schemes like the one we have described above are usually called semi-Lagrangian or control schemes in the literature. They are (usually) monotone and first order accurate, and a comprehensive background and references can be found in [24]. Most results in this field concerns deterministic control problems and first order HJB equations without integral terms. Semi-Lagrangian schemes for second order HJB equations with no integral terms have been considered in [37, 17, 5, 1]. Moreover in [16] such schemes were derived for HJB equations associated to piecewise deterministic processes (compound Poisson processes with drift). One advantage of this type of schemes is that in general they produce also an approximation of the optimal control law in feedback form and the optimal trajectories (this is a key point in the study of a control problem). This advantage can also be shared by some (monotone, low order) finite difference schemes following the construction of Kushner, and we refer to [34] for a discussion of this point. An other advantage is that these schemes are formulated on general grids/triangulations just like FEMs. By contrast this is very cumbersome to achieve with finite difference methods. A third advantage is that the semi-Lagrangian scheme *automatically* handles non-diagonally dominant diffusion matrices a. Such matrices appear in applications and are cumbersome to discretize, we refer to [12] for a finite difference (FDM) approach to this problem.

The construction and analysis of numerical schemes for linear integro-partial differential equations arising as pricing equations in financial markets of jump-diffusion type is currently an active field of research, see e.g. [19, 36, 23, 15] and references therein. By contrast, there are few works on numerical schemes for fully nonlinear degenerate integro-partial differential equations. We mention the discussion about jump-diffusion processes in [34] and the papers [22, 14, 30, 10]. In all cases mainly monotone FDMs are considered and convergence is obtained. In the last two papers the main focus is on convergence rates. The framework of [30] is again based on ideas of Krylov, Barles and Jakobsen [31, 32, 5] in the pure PDE case. We emphasize that the error bounds of this paper apply when grids are unstructured and integral terms are very singular, a feature which is new or highly unusual in this setting.

The rest of this paper is organized as follows. In section 2 we state the assumptions on the data and give well-posedness/regularity results for equations (1.1) and (1.2). We discuss truncation of the Levy measure and reduction to a bounded domain and give error bounds. In section 3 we construct the schemes via the dynamic programming equation (a semi-discretization) and FEM ideas. We prove existence, uniqueness, consistency, and partial error bounds. We also derive a fully discrete scheme of FEM type by piecewise linear reconstruction on a (regular) triangulation of the domain. In section 4 we derive error bounds for the semi-discrete scheme with or without truncation of the Levy measure. Finally in the appendix, we prove the main technical result of this paper, a regularity and continuous dependence result for the semi-discrete scheme.

Notation: By $USC_b(\mathbb{R}^N)$ and $C_b(\mathbb{R}^N)$ we mean the spaces of bounded continuous and upper semicontinuous functions. We will use the following norms

$$|f|_0 = \operatorname{ess\,sup}_{x\in\mathbb{R}^N}|f(x)|, \quad [f]_1 = |Df|_0, \quad \text{and} \quad |f|_1 = |f|_0 + [f]_1.$$

2. Preliminaries

In this section we state the assumptions for equations (1.1) and (1.2) and give well-posedness results. We discuss reduction to a bounded domain and reduction to bounded measure ν with compact support.

2.1. **Assumptions, well-posedness, regularity.** We will use the following assumptions:

(A1) The set V is a compact metric space, the coefficients σ, η, b, c, f are continuous in x and v and Borel measurable in z, and ν is a positive Radon measure on $E := \mathbb{R}^M \setminus \{0\}$.

(A2) There exists $L_1, L_2, \ell \geq 0$ such that for any $v \in V$ and $z \in E$

$$\begin{aligned} &|c(\cdot,v)|_1 + |f(\cdot,v)|_1 \leq L_1,\\ &|\sigma(\cdot,v)|_1 + |b(\cdot,v)|_1 \leq L_2\\ &|\eta(\cdot,v,z)|_1 \leq L_2(|z|1_{|z|<1} + e^{\ell|z|}1_{|z|>1}).\end{aligned}$$

(A3) There exists $c_0 > 0$ such that, for all $x \in \mathbb{R}^N$ and $v \in V$,

$$c(x,v) \geq c_0.$$

(A4) There exists $L_3 \geq 0$, such that for ℓ defined in (A2), any $v \in V$ and $x \in \mathbb{R}^N$,

$$|D_z\eta(x,v,\cdot)|_0 \leq L_3 e^{\ell|z|}.$$

For the measure ν we will use the following integrability assumptions.

$(\mathrm{B1})_k$ The measure ν satisfies for ℓ defined in (A2) and some $k \in \{0,1,2\}$,

$$\int_E (|z|^k 1_{|z|<1} + e^{\ell|z|}1_{|z|>1})\,\nu(dz) < \infty.$$

(B2) The measure ν has a positive density $m : E \to [0,\infty)$ such that $\nu(dz) = m(z)dz$ and for some C, $\varepsilon > 0$, and $n \in \mathbb{N}$

$$|D^k m(z)| \leq C_k e^{-(\ell+\epsilon)|z|} \quad \text{for} \quad k = 0, 1, \ldots, n,$$

where $D^k m$ is the vector of all order k derivatives of m and ℓ is defined in (A2).

(B3) The measure ν has a positive density $m : E \to [0, \infty)$ such that $\nu(dz) = m(z)dz$ and for some $\alpha \in [0, 2)$, C_0, C_1, and $\varepsilon > 0$,

$$|D^k m(z)| \leq C_k \Big(\frac{1}{|z|^{N+\alpha+k}} 1_{|z|<1} + e^{-(\ell+\epsilon)|z|} 1_{|z|>1} \Big) \quad \text{for} \quad k = 0, 1,$$

where $D^0 m = m$, $D^1 m = Dm$ the gradient of m, and ℓ is defined in (A2).

Assumptions (A1) – (A3) and $(B1)_2$ are standard from a stochastic control theory point of view, as they insure existence and uniqueness of strong solutions of the underlying stochastic differential equations, see [26]. Under assumption $(B1)_2$ or the less general assumption $(B1)_1$ the measure ν may have a (non-integrable) singularity at $z = 0$. If $(B1)_1$ or $(B1)_0$ holds the HJB equation takes the form (1.1). If $(B1)_2$ holds but not $(B1)_1$, then (1.2) gives the correct form of the HJB equation. In this case, the extra term in the integrand of $\mathcal{J}^v$ (compared to $\mathcal{I}^v$) is needed for the integral to converge: If $(B1)_2$ holds, then $\mathcal{J}^v u$ converge for all C^2 functions u is with polynomial growth at infinity. Finally, (B2) and (B3) prescribe densities. Close to $z = 0$, the density of (B3) equals the density of the α-stable processes related to the fractional Laplacian $\Delta^{\alpha/2}$. (B3) implies $(B1)_2$, and if $\alpha \in [0, 1)$, then (B3) also implies $(B1)_1$.

Example 2.1. When the stock returns are modeled as exponential Levy processes the integral terms equals (1.4) or (1.5) with $\eta(x, z) = x(e^z - 1)$ in the one-dimensional uncontrolled case, see [19]. An important extension of this expression satisfying (A2) and (A4) is $\eta(x, v, z) = \bar{\eta}(x, v)\phi(z)$ where $\bar{\eta}$ is bounded and Lipschitz and $|\phi(z)| \leq C(e^{|z|} - 1)$, $|D\phi(z)| \leq Ce^{|z|}$.

Example 2.2. Assumptions (B2) and (B3) is satisfied by all bounded and unbounded Lévy processes used in the literature to model financial markets, see [19] for a nice overview. In the models of Merton and Kou, the Levy measures have bounded densities (no singularity), in the Merton case [38] given by

$$\nu(dz) = \frac{\lambda}{\delta\sqrt{2\pi}} e^{-\frac{|z-\mu|^2}{2\delta^2}} dz \quad \text{for constants } \lambda, \delta, \mu.$$

For these models (B2) holds. The Variance Gamma model has a Levy density with an integrable singularity at $z = 0$, and satisfy (B3) with $\alpha = 0$. The Normal Inverse Gaussian model [40] has a non-integrable density corresponding to $\alpha = 1$ in (B3), and the Levy measure is

$$\nu(dz) = \frac{C}{|z|} e^{Az} K_1(B|z|) \quad \text{for constants } A, B, C,$$

where K_1 is a modified Bessel function of 2nd kind. Finally we mention models using tempered α-stable processes, e.g. the CGMY model [18]. Here $\alpha \in (0, 2)$ in (B3) and the Levy measure is give by

$$\nu(dz) = \frac{c_-}{|z|^{1+\alpha}} e^{-\lambda_-|z|} 1_{z<0} + \frac{c_+}{|z|^{1+\alpha}} e^{-\lambda_+|z|-} 1_{z>0}$$
$$\text{for constants } \alpha \in (0, 2), c_-, c_+, \lambda_-, \lambda_+ \geq 0.$$

It is well known that under the above assumptions the solutions to (1.1) and (1.2) need not be smooth, and that the correct concept of (weak) solutions is that of viscosity solutions. For the definition of viscosity solution in this case we refer to [41, 3, 39, 29, 4]. We state without proof a well-posedness and regularity result for (1.2). The proof of this result is standard, and we refer to [3, 39, 28, 29] for the proofs of similar results.

Theorem 2.1. *Assume (A1) – (A3) and $(B1)_2$ hold.*

(i) There exists a unique viscosity solution $u \in C_b(\mathbb{R}^N)$ of equation (1.2) which is Hölder continuous, i.e., there is a $\delta \in (0, 1]$ such that

$$|u(x) - u(y)| \leq C|x - y|^\delta \quad \textit{for all} \quad x, y \in \mathbb{R}^N.$$

(ii) There exists a constant $c_1 > 0$ depending only on $\sup_v[\sigma(\cdot, v)]_1$, $\sup_v[b(\cdot, v)]_1$, and $\sup_v \int_E [\eta(\cdot, v, z)]_1^2 \nu(dz)$ such that if $c_0 \geq c_1$, then the viscosity solution u of (1.2) is Lipschitz continuous ($\delta = 1$ above).

(iii) Let $u, -v \in USC_b(\mathbb{R}^N)$. If u and v are respectively viscosity sub- and supersolutions of (1.2)*, then $u \le v$ in $\mathbb{R}^N$.*

Remark 2.2. Note that this result also holds for (1.1) under assumptions (A1) – (A3) and either one of $(B1)_1$ or $(B1)_0$. Simply write this equation in the form (1.2) with b replaced by $\bar{b} = b - \int_E \eta\nu(dz)$.

2.2. **Reduction to bounded measure with compact support.** Consider Levy measures ν which are not bounded nor have compact support. We assume that (B3) hold and consider two cases: (i) $\alpha \in [0,1)$ and (ii) $\alpha \in [1,2)$. In the first case we truncate the Levy and compensate by adding a drift and a diffusion term. Let us introduce the "two-scales" truncated Lévy measure

$$\nu_{r,R}(dz) := 1_{r<|z|<R}\, \nu(dz), \tag{2.1}$$

where $r \in (0,1)$ and $R > 1$. Clearly, $\nu_{r,R}(dz)$ is a bounded and compactly supported measure satisfying assumption $(B1)_0$. When $\alpha \in (0,1)$ assumption $(B1)_1$ holds and the HJB equation is (1.1). This equation is approximated by

$$\sup_{v\in V} \left\{ -\operatorname{tr}[\bar{a}(x,v)D^2u] - \bar{b}(x,v)Du + c(x,v)u - f(x,v) - \bar{\mathcal{I}}^v u(x) \right\} = 0 \quad \text{in} \quad \mathbb{R}^N, \tag{2.2}$$

where

$$\bar{a}(x,v) = a(x,v) + \frac{1}{2}\int_{0<|z|<r} \eta(x,v,z)\eta(x,v,z)^T\, \nu(dz), \tag{2.3}$$

$$\bar{b}(x,v) = b(x,v) + \int_{0<|z|<r} \eta(x,v,z)\nu(dz), \tag{2.4}$$

$$\bar{\mathcal{I}}^v u(x) = \int_E [\phi(x+\eta(x,v,z)) - \phi(x)]\; \nu_{r,R}(dz). \tag{2.5}$$

This gives an approximation of $\mathcal{I}^v\phi$ which is "third order" near $z=0$ since

$$\begin{aligned}
&|\bar{\mathcal{I}}^v\phi + (\bar{b}-b)D\phi + \operatorname{tr}[(\bar{a}-a)D^2\phi] - \mathcal{I}^v\phi| \\
&\le K|D^3\phi|_0 \left| \int_{0<|z|<r} |\eta(\cdot,\cdot,z)|^3 \nu(dz) \right|_0 + K|\phi|_0 \int_{|z|>R} \nu(dz).
\end{aligned}$$

In case (ii) we truncate the Levy measure and compensate by adding a diffusion term. When $\alpha \in [1,2)$ assumption $(B1)_1$ is not satisfied and the HJB equation is (1.2). We approximate this equation by

$$\sup_{v\in V} \Big\{ -\operatorname{tr}[\bar{a}(x,v)D^2u] - \tilde{b}(x,v)Du + c(x,v)u - f(x,v) - \bar{\mathcal{I}}^v u(x) \Big\} = 0 \quad \text{in} \quad \mathbb{R}^N, \tag{2.6}$$

where $\bar{a}$ and $\bar{\mathcal{I}}^v u$ is defined in (2.3) and (2.5), and

$$\tilde{b}(x,v) = b(x,v) + \int_E 1_{|z|<1}\eta(x,v,z)\nu_{r,R}(dz). \tag{2.7}$$

Again we obtain an approximation of the integral term ($\mathcal{J}^v\phi$ this time) which is "third order" near $z=0$ since

$$\begin{aligned}
&|\bar{\mathcal{I}}^v\phi + (\tilde{b}-b)D\phi + \operatorname{tr}[(\bar{a}-a)D^2\phi] - \mathcal{J}^v\phi| \\
&\le K|D^3\phi|_0 \left| \int_{0<|z|<r} |\eta(\cdot,\cdot,z)|^3 \nu(dz) \right|_0 + K|\phi|_0 \int_{|z|>R} \nu(dz).
\end{aligned}$$

Equations (2.2) and (2.6) are almost, but not quite, of the same form as (1.1). In [30] it is proved that Theorem 2.1 still holds for solutions of (2.2) and (2.6). The following result gives error bounds for the approximations (2.2), (2.6) of equations (1.1), (1.2).

Lemma 2.3. *Assume (A1) – (A3) and (B3) hold and let $u, \bar{u}, v, \bar{v} \in C^{0,1}(\mathbb{R}^N)$ solve* (1.1), (2.2), (1.2), (2.6) *respectively. Then for $r > 0$ small enough and R large enough we have*

$$|u-\bar{u}|_0 + |v-\bar{v}|_0 \le C_1(r^{1-\alpha/3} + e^{-\ell R}),$$

for some constant C_1 independent of r and R.

The result about $v, \bar{v}$ is proved in [30], while the result about $u, \bar{u}$ is new but the proof is almost identical to the proof of the $v, \bar{v}$ result. We omit it.

Remark 2.4. The compensating drift and diffusion terms are added to improve the convergence of the approximation (of the small jumps). The error is of order $r^{1-\alpha/2}$ without the compensating diffusion (see [30]), while it is of order $r^{1-\alpha}$ without compensating drift and diffusion in the case $\alpha \in [0, 1)$.

Remark 2.5. In probabilistic terms, the explanation for the improved convergence is that the small jumps of a Levy process can be approximated by a Brownian motion, we refer to e.g. [2, 19] for details. Moreover using probabilistic methods (Berry-Essen type estimates), it is possible to prove convergence of order r in some cases [2]. We refer to [20] where such estimates are made explicit for linear parabolic equations with constant coefficients.

Remark 2.6. When $\alpha = 0$ no improvement is obtained by adding drift and diffusion terms to the equation. If ν has a bounded density, no truncation is needed close to $z = 0$.

2.3. **Reduction to a bounded domain.** Reduction to a bounded domain is key step in order to implement a numerical method. Following ideas of [17], we restrict equations (1.1) and (1.2) to bounded domains by truncating the coefficients outside some large ball,

$$B_{\frac{1}{\mu}} = \{x \in \mathbb{R}^n : |x| < \frac{1}{\mu}\} \quad \text{for} \quad \mu > 0.$$

Let $\xi_\mu \in C_c^\infty(\mathbb{R}^N)$ be a cut-off function satisfying $0 \le \xi_\mu \le 1$ and $\xi_\mu(x) = 1$ for $x \in B_{\frac{1}{\mu}}$ and define

$$\sigma_\mu = \xi_\mu(x)\sigma(x,a), \qquad b_\mu = \xi_\mu(x)b(x,a), \qquad \eta_\mu(x,v,z) = \xi_\mu(x)\eta(x,v,z).$$

Since b_μ, σ_μ and η_μ satisfies the same assumptions of σ, b and η for any μ, there exists a unique viscosity solution u_μ of the equation

$$\sup_{v \in V}\left\{-\operatorname{tr}\left[a_\mu(x,v)D^2u\right] - b_\mu(x,v)Du + c(x,v)u - f(x,v) - \mathcal{J}_\mu^v(x)\right\} = 0 \quad \text{in} \quad \mathbb{R}^N \tag{2.8}$$

where $a_\mu(x,v) = \frac{1}{2}\sigma_\mu(x,v)\sigma_\mu^T(x,v)$ and $\mathcal{J}_\mu^v(\cdot)$ is defined as in (1.5) with η_μ in place of η. Since the coefficients b_μ, σ_μ and η_μ are zero outside of $\operatorname{supp}(\xi_\mu)$, the solution u_μ (2.8) is given by

$$u_\mu(x) = \min_{v \in V} \frac{f(x,v)}{c(x,v)} \quad \text{for any} \quad x \notin \operatorname{supp}(\xi_\mu).$$

The next result give a crude bound on the error that this cut-off procedure introduces.

Lemma 2.7. *Assume (A1), (A2), (B1)$_2$ hold, that* (1.2) *and* (2.8) *satisfy the dynamic programming principle, and that u, u_μ solve* (1.2), (2.8) *respectively. Then there exists a constant C such that*

$$|u(x) - u_\mu(x)|_0 \le C\mu^2(1 + |x|^2) \quad in \quad B_{\frac{1}{\mu}}.$$

We skip the proof since it is similar to the proof of Proposition 3.1 in [17] when we are equipped with the moment estimates (3.2) and (3.4) of [39]. The dynamic programming principle has recently been extended to the current setting in [11].

Remark 2.8. As in Remark 2.2 we immediately get an analogous result for (1.1) under assumption (B1)$_1$.

Remark 2.9. If the equation is uniformly elliptic or has a non-degenerate singular integral term, then you expect estimates decaying exponentially as $\mu \to \infty$. We refer to [20, 36] for such result in a linear one-dimensional setting.

3. Construction of the scheme

In subsections 3.1 and 3.2 we always assume that (A1) – (A3) and (B1)$_0$ hold: The Levy measure ν is bounded and the HJB equation has the form (1.1). We can always reduce to this case by a truncation. In subsection 3.3 we discuss the general case.

3.1. **Semi-discretization.** We introduce a control problem for which (1.1) is the corresponding dynamic programming or HJB equation. We start by defining the controlled dynamics. Since the measure ν is bounded, we can normalize ν and obtain a probability measure μ as follows

$$\mu(dz) = \frac{1}{\lambda}\nu(dz)$$

where

$$\lambda = \int_E \nu(dz) \tag{3.1}$$

Now we consider a Markov process $X_t^{v_t}$ evolving according the SDE

$$dX_t = b(X_{t^-}, v_{t^-})dt + \sigma(X_{t^-}, v_{t^-})dW_t + \int_{|z|>0} \eta(X_{t^-}, v_{t^-}, z)\bar{\mu}(dz, dt), \tag{3.2}$$

where $\bar{\mu}$ is a Poisson measure corresponding to a compound Poisson process with jump intensity λ and jump distribution μ. The control v_t belongs to $\mathcal{V}$, the set of all progressively measurable processes with values in V. Since we assume that ν and hence $\bar{\mu}$ are bounded, on any finite time interval X_t will only jump finitely many times with probability one. Between two jump times T_i and T_{i+1} the process diffuses according to the SDE

$$dX_t = b(X_{t^-}, v_{t^-})dt + \sigma(X_{t^-}, v_{t^-})dW_t. \tag{3.3}$$

For $v_t \equiv v \in V$ and $a(x, v) = \frac{1}{2}\sigma(x, v)\sigma^T(x, v)$, the infinitesimal generator of the process X_t is

$$L^v\psi(x) = \lim_{t\to 0} \frac{\mathbb{E}_x[\psi(X_t)] - \psi(x)}{t} = \mathrm{tr}\,[a(x, v)D^2\psi] + b(x, v)D\psi + \mathcal{I}^v\psi(x)$$

for $\psi \in C^2(\mathbb{R}^N)$. On the paths of the process X_t we define the discounted cost functional

$$J(x, v_t) = \mathbb{E}_x\Big[\int_0^\infty f(X_t, v_t)e^{-\int_0^t c(X_s, v_s)ds}dt\Big],$$

and we consider the corresponding value function

$$u(x) = \inf_{v_t\in\mathcal{V}} J(x, v_t).$$

In [11] it is proved that u is the unique viscosity solution of equation (1.1).

Following the approach of [16, 17] we construct an approximation scheme for the equation (1.1) by discretizing the associated control problem. We fix a discretization step $h > 0$ and consider two stochastic processes N_n and Z_n, $n \in \mathbb{N}$, taking values in $\mathbb{N}$ and in $\mathbb{R}^N$ and representing the n-th jump time and the corresponding z-jump (size and direction) of the Poisson measure $\bar{\mu}$. We set $N_0 = 0$ and $Z_0 = 0$ and assume that N_n has independent $h\lambda$-exponentially distributed increments, i.e. the probability distribution of N_n is given by

$$\mathbb{P}[N_{n+1} - N_n \geq j | N_0, N_1, \ldots, N_n] = e^{-h\lambda j}, \quad n = 0, 1, 2, \ldots,$$

while the Z_n, $n \in \mathbb{N}$, are i.i.d. random variables with probability density μ. Now we define a discrete time stochastic process X_n approximating the continuous time process X_t.

$$\begin{cases} X_0 = x, \\ X_n = X_{n-1} + hb(X_{n-1}, v_{n-1}) + \sqrt{h}\sum_{m=1}^d \sigma_m(X_{n-1}, v_{n-1})\xi^m_{n-1}, \\ \qquad\qquad\qquad\qquad \text{for } n = N_i + 1, N_i + 2, \ldots, N_{i+1} - 1, \\ X_{N_{i+1}} = X_{N_{i+1}-1} + \eta(X_{N_{i+1}-1}, v_{N_{i+1}-1}, Z_i), \end{cases}$$

for $i = 1, 2, 3, \ldots$, where σ_m denote the m-th column of σ and ξ^m_n, $m = 1, \ldots, d$ are random variables taking values in $\{-1, 0, 1\}$ such that

$$\mathbb{P}[\{\xi^i_n = \pm 1\}] = \frac{1}{2d} \qquad \text{and} \qquad \mathbb{P}[\{\xi^i_n \neq 0\} \cap \{\xi^j_n \neq 0\}] = 0, \quad i \neq j.$$

The discrete control $\{v_n\}$ is a random variable with values in V which is measurable with respect to the σ-algebra generated by $X_1, \ldots, X_n$.

Between jumps the process evolves like a random walk approximating the SDE (3.3), and when the process jumps there is no diffusion/random walk. The generator of the discrete process is

$$L_h^v\psi(x) = \frac{\mathbb{E}_x[\psi(X_1)] - \psi(x)}{h} = e^{-h\lambda}\mathcal{L}_h^v\psi(x) + \frac{1-e^{-h\lambda}}{h}\mathcal{I}_h^v\psi(x), \tag{3.4}$$

for $\psi \in C^0(\mathbb{R}^N)$ and where

$$\mathcal{L}_h^v\psi = \frac{1}{2dh}\sum_{m=1}^{d}\Big[\psi(x+hb(x,v)+\sqrt{h}\sigma_m(x,v)) + \psi(x+hb(x,v)-\sqrt{h}\sigma_m(x,v)) - 2\psi(x)\Big],$$
$$\mathcal{I}_h^v\psi = \frac{1}{\lambda}\mathcal{I}^v\psi,$$

with $\mathcal{I}^v$ as in (1.4). Observe that at this level the space variable is not discretized, therefore the discrete process has the same jump distribution as the continuous process.

On the paths of the discrete process we define the cost functional

$$J_h(x,\{v_n\}) = \mathbb{E}_x[\sum_{n=0}^{\infty} he^{-h\sum_{i=0}^{n-1}c(X_i,v_i)}f(X_n,v_n)], \tag{3.5}$$

(with the convention $\sum_{i=0}^{-1} = 0$), and the corresponding value function for the discrete control problem

$$u_h(x) = \inf_{\{v_n\}} J_h(x,\{v_n\}). \tag{3.6}$$

Now it is easy to see, at least formally [9], that the following dynamic programming principle holds:

$$u_h(x) = \inf_{\{v_i\}}\mathbb{E}_x\Big[\sum_{n=0}^{P} he^{-h\sum_{i=0}^{n-1}c(X_i,v_i)}f(X_n,v_n) + e^{-h\sum_{i=0}^{P}c(X_i,v_i)}u_h(X_{P+1})\Big],$$

for any $P \in \mathbb{N}$. Taking $P=0$ in the above equation and noting that $X_0 = x$, we get

$$\begin{aligned} u_h(x) &= \inf_{v\in V}\mathbb{E}_x[hf(X_0,v) + e^{-hc(X_0,v)}u_h(X_1)] \\ &= \inf_{v\in V}\Big\{hf(x,v) + e^{-hc(x,v)}\Big[\frac{e^{-h\lambda}}{2d}\sum_{m=1}^{d}(u_h(x+hb(x,v)+\sqrt{h}\sigma_m(x,v)) \\ &\quad + u_h(x+hb(x,v)-\sqrt{h}\sigma_m(x,v)) + \frac{1-e^{-h\lambda}}{\lambda}\int_E u_h(x+\eta(x,v,z))\nu(dz)\Big]\Big\}. \end{aligned} \tag{3.7}$$

Rearranging the terms in the previous equation and dividing by $he^{-hc(x,v)}$ we get

$$\sup_{v\in V}\Big\{-L_h^v u_h(x) + \frac{e^{hc(x,v)}-1}{h}u_h(x) - e^{hc(x,v)}f(x,v)\Big\} = 0 \quad \text{in} \quad \mathbb{R}^N, \tag{3.8}$$

where $L_h^v(\cdot)$ is as in (3.4). We will talk about sub- and supersolutions of this equation, meaning that (3.8) holds as an inequality with $\le$ and $\ge$ respectively. For the scheme (3.8) we have the following easy properties:

Proposition 3.1. *Assume (A1) – (A3) and (B1)$_0$.*

(i) If u_h and v_h are bounded sub- and supersolutions of (3.8)*, then $u_h \le v_h$ in $\mathbb{R}^N$.*

(ii) Any solution u_h of (3.8) *is bounded and satisfies for all $h > 0$,*

$$|u_h|_0 \le \frac{\sup_v |f(\cdot,v)|_0}{c_0}, \quad \textit{where } c_0 \textit{ is as in (A3).}$$

(iii) There exists a unique bounded continuous function u_h solving (3.8).

(iv) For $0 < h < 1$ and $\phi \in C^4(\mathbb{R}^N)$ satisfying $|\phi|_0 + \cdots + |D^4\phi|_0 < \infty$,

$$|L^v\psi(x) - L_h^v\psi(x)| \le C_1 h(|D^2\phi|_0 + |D^3\phi|_0 + |D^4\phi|_0) + C_2 h\lambda((1+|\int_E \eta\nu|)|D\phi|_0 + |D^2\phi|_0),$$

where the constants C_1 and C_2 only depend on $\sup_v |\sigma|_0, \sup_v |b|_0$.

Remark 3.2. For the truncation of an unbounded measure to converge as $h \to 0$, we need $\lambda \to \infty$ as $h \to 0$, while the scheme (3.8) converges only if both $h \to 0$ and $\lambda h \to 0$ by *(iv)*. The last condition means that the small jumps must be resolved in the grid.

Proof. We work with the scheme in the equivalent form (3.7). Note that the scheme is monotone, it has positive coefficients. If u and v be sub- and supersolutions of (3.8), an easy and standard computation using (3.7) and assumption (A3) shows that

$$u_h(x) - v_h(x) \leq e^{-c_0 h}|(u_h - v_h)^+|_0,$$

and hence $(1 - e^{-c_0 h})|(u_h - v_h)^+|_0 \leq 0$ which proves (i). A similar computation shows (ii) after noting that $\frac{hce^{hc}}{e^{hc}-1} \leq 1$ for $hc \geq 0$. Next denote the right hand side of (3.7) by $T_h u_h$, and note that T_h is contraction in the $|\cdot|_0$ norm,

$$T_h u_h - T_h v_h \leq e^{-hc_0}|u_h - v_h|_0.$$

Existence and uniqueness of a continuous bounded solution to (3.8) follows from Banach's fixed point theorem and this proves (iii). Finally, to prove (iv), note that since $\mathcal{I}^v_h = \lambda^{-1}\mathcal{I}^v$ we may write

$$L^v\psi - L^v_h\psi = (\mathcal{L}^v - \mathcal{L}^v_h)\psi + (\mathcal{I}^v - \mathcal{I}^v)\psi - (1 - e^{-h\lambda})\mathcal{L}^v_h\psi - (1 - h\lambda^{-1}(1 - e^{-h\lambda}))\mathcal{I}^v\psi,$$

where $\mathcal{L}^v\psi(x) = \operatorname{tr}[a(x,v)D^2\psi(x)] + b(x,v)D\psi(x)$. By Taylor expansion, e.g.

$$|(\mathcal{L}^v - \mathcal{L}^v_h)\psi| \leq h|b|_0^2|D^2\psi|_0 + (h^2|b|_0^3 + h|b|_0|\sigma|_0^2)|D^3\psi|_0 + (h^3|b|_0^4 + h^2|b|_0^2|\sigma|_0^2 + h|\sigma|_0^4)|D^4\psi|_0,$$

and the estimates $|1 - e^{-x}|, |1 - x^{-1}(1 - e^{-x})| \leq |x|$ the result follows. □

Next we give an optimal Lipschitz regularity and continuous dependence on coefficients result for the scheme (3.8).

Proposition 3.3. *Let u_h and $\tilde{u}_h$ be solutions of* (3.8) *corresponding to the data $\sigma, b, c, f, \eta, \nu$ and $\tilde{\sigma}, \tilde{b}, \tilde{c}, \tilde{f}, \tilde{\eta}, \nu$ respectively, assume both sets of coefficients satisfy (A1) – (A3) and $(B1)_0$, and that $h\lambda \leq \bar{C}_0$ and $h \in (0,1]$. Then there exist constants $c_1, L, K \geq 0$ (only depending on the data and $\bar{C}_0$) such that if $c_0 \geq c_1$ (c_0 as in (A3)), then for all $h > 0, x, y \in \mathbb{R}^N$,*

$$\begin{aligned}|u_h(x) - \tilde{u}_h(y)| \leq L|x-y| + K \sup_{v \in V}\Big[&|f - \tilde{f}|_0 + |c - \tilde{c}|_0 \\ &+ |b - \tilde{b}|_0 + |\sigma - \tilde{\sigma}|_0 + \Big|\int_E |\eta(\cdot,z) - \tilde{\eta}(\cdot,z)|^2\nu(dz)\Big|_0^{1/2}\Big].\end{aligned}$$

If the coefficients are equal and $u_h = \tilde{u}_h$, then this is a Lipschitz regularity result, while if $x = y$ then this is a continuous dependence on the coefficients result. This result one of the main contributions of this paper, and it will play a key role in the next section where error bounds are derived. It extends similar results of [5] to equations with integral terms. The proof is rather long and technical and we have put the main bulk in the appendix. In the pure diffusion case, the current proof simplifies considerably the arguments of [5].

Proof. The result follows from Theorem A.1 in the Appendix after writing

$$b = \Big[b - \frac{1 - e^{-h\lambda}}{\lambda e^{-\lambda h}}\int_E \eta\nu\Big] + \frac{1 - e^{-h\lambda}}{\lambda e^{-\lambda h}}\int_E \eta\nu,$$

and noting that by the Cauchy-Schwartz inequality,

$$\Big|b - \tilde{b} - \frac{1 - e^{-h\lambda}}{\lambda e^{-\lambda h}}\int_E [\eta - \tilde{\eta}]\nu\Big|_0 \leq |b - \tilde{b}|_0 + (e^{h\lambda} - 1)\lambda^{-1/2}\Big|\int_E |\eta - \tilde{\eta}|^2\nu\Big|_0^{1/2}.$$

□

Remark 3.4. When $c_0 < c_1$ the solution to (3.8) is only Hölder continuous. We will not discuss this case here, but refer instead to [5] for how to obtain results in this case.

3.2. **The fully-discrete scheme.** In this section we introduce a FEM like discretization of (3.8) yielding a fully discrete scheme. For a nice introduction to FEMs we refer to [13]. For $k > 0$ let $\mathcal{T}^k = \{S_j^k\}_{j\in\mathbb{N}}$ be a non-degenerate triangulation of $\mathbb{R}^N$, i.e. a collection of N-simplices S_j^k such that

$$\bigcup_{j\in\mathbb{N}} S_j^k = \mathbb{R}^N, \quad \sup_{j\in\mathbb{N}}(\operatorname{diam} S_j) \le k, \quad \rho k \le \sup_{j\in\mathbb{N}}(\operatorname{diam} B_{S_j^k}),$$

where $\rho \in (0,1)$, diam denotes the diameter of the set, and $B_{S_j^k}$ is the greatest ball contained in S_j^k. We denote by $X^k = \{x_i\}_{i\in\mathbb{N}}$ the corresponding set of the vertices, and introduce the space of continuous piecewise linear functions on $\mathcal{T}^k$,

$$W^k = \{w \in C(\mathbb{R}^N) : Dw(x) \text{ is constant in } S_j^k\}.$$

Every element w in W^k can be expressed as

$$w(x) = \sum_{i\in\mathbb{N}} \beta_i(x) w(x_i),$$

for basis functions (the so-called tent functions) $\beta_i \in W^k$ satisfying $\beta_i(x_j) = \delta_{ij}$ for $i,j \in \mathbb{N}$. It immediately follows that $0 \le \beta_i(x) \le 1$, $\sum_{i\in\mathbb{N}} \beta_i(x) = 1$, β_i has compact support, and at any $x \in \mathbb{R}^N$ at most $N+1$ β_i's are non-zero. The family $\{\beta_j\}_j$ is a partition of unity. On any simplex S_i^k, the β_j's are called the barycentric coordinates of S_i^k.

The fully discrete scheme can then be formulated as follows: Find the function $u \in W^k$ that satisfies (3.8) at every vertex $x_i \in X^k$, or equivalently,

$$u(x_i) = \inf_{v\in V} \Big\{ e^{-hc(x_i,v)} \Big[e^{-\lambda h} \sum_j M_{ij}(v) u(x_j) + (1 - e^{-\lambda h}) \sum_j P_{ij}(v) u(x_j) \Big] + hf(x_i, v) \Big\}, \tag{3.9}$$

for every $x_i \in X^K$. Here the matrices $M(v)$ and $P(v)$ are given by

$$M(v) = \sum_{m=1}^{d} \frac{1}{2d} (M_m^+(v) + M_m^-(v)) \tag{3.10}$$

for $M_{m,ij}^{\pm}(v) = \beta_j\big(x_i + hb(x_i,v) \pm \sqrt{h}\sigma_m(x_i,v)\big)$, and

$$P_{ij}(v) = \frac{1}{\lambda} \int_E \beta_j(x_i + \eta(x_i,v,z)) \nu(dz). \tag{3.11}$$

Note that M is a stochastic matrix, and for any m, only $N+1$ entries of any row of $M_m^{\pm}$ are non-zero. The matrix P is non-zero only if the vertex x_j belong to a simplex which has nonempty intersection with the set $x_i + \eta(x_i, v, \operatorname{supp}(\nu))$ for all $v \in V$.

As a final step we also discretize P_{ij} by (monotone) quadrature

$$Q_{\Delta z}[\phi] := \sum_{j\in\mathbb{N}} \phi(z_j)\omega_j \quad \text{where} \quad z_j \in E, \quad \omega_j \ge 0, \tag{3.12}$$

satisfying the error bound

$$E_{\Delta z}[\phi] := \Big| \int_E \phi(z)dz - Q_{\Delta z}[\phi] \Big| \le \bar{K} \|D\phi\|_{L^1} \Delta z \quad \text{for} \quad \phi \in W^{1,1}(\mathbb{R}^n). \tag{3.13}$$

Here $\Delta z > 0$ is the discretization parameter. An example is the compound midpoint rule where $Q_{\Delta z}[\phi] = \sum_j \phi(z_j)\Delta z^M$ and $\{z_j\}_j$ is a renumeration of the uniform z-grid $(\Delta z\, \mathbb{Z}^M) \cap E$. All sensible tensor product quadratures satisfy the (first order) error bound, and the bound is optimal if the integrand ϕ is not more regular. The monotonicity assumption $\omega_j \ge 0$ is satisfied for compound Gauss and Newton-Cotes types of quadratures in any space dimension and when the order is less than 9 in the Newton-Cotes case. We refer to [30] for examples and a wider discussion of these issues. If we assuming that ν has a density m satisfying (B2), we get the following final scheme: Find $u \in W^k$ such that

$$u(x_i) = \inf_{v\in V} \Big\{ e^{-hc(x_i,v)} \Big[e^{-\lambda h} \sum_j M_{ij}(v) u(x_j) + (1 - e^{-\lambda h}) \sum_j \bar{P}_{ij}(v) u(x_j) \Big] + hf(x_i, v) \Big\}, \tag{3.14}$$

where M_{ij} is defined in (3.10) and

$$\bar{P}_{ij}(v) = \frac{1}{\lambda} Q_{\Delta z}[\beta_j(x_i + \eta(x_i, v, \cdot))m(\cdot)].$$

We have the following existence, uniqueness, and *partial* convergence result for (3.14).

Theorem 3.5. *Assume (A1) – (A4), (B2),* (3.12), (3.13) *hold,* $h\lambda \le \bar{C}_0$*, and* $h \in (0,1]$. *Then there exists a unique bounded solution* $u_{hk\Delta z} \in W^k$ *to* (3.14). *Furthermore, if the solution* u_h *of* (3.8) *belongs to* $C^{0,1}(\mathbb{R}^N)$*, then*

$$|u_h - u_{hk\Delta z}|_0 \le \frac{|u_h|_1}{1 - e^{-c_0 h}} \Big(2k + C\Delta z\Big).$$

Proof. Existence and uniqueness follow from a fixed point argument like in the proof of Proposition 3.1. To prove the error bound, note that since $u_{hk\Delta z}(x) = \sum_{j\in\mathbb{N}} \beta_j(x) u_{hk\Delta z}(x_j)$, $\beta_j \ge 0$, and $\sum \beta_j = 1$,

$$\begin{aligned} |u_h(x) - u_{hk\Delta z}(x)| &\le \sum_{j\in\mathbb{N}} \Big(\beta_j(x)|u_h(x) - u_h(x_j)| + \beta_j(x)|u_h(x_j) - u_{hk\Delta z}(x_j)|\Big) \\ &\le |u_h|_1 k + \sum_{j\in\mathbb{N}} \beta_j(x)|u_h(x_j) - u_{hk\Delta z}(x_j)|. \end{aligned} \tag{3.15}$$

The last term can be estimated by using (3.8) and (3.14). Easy computations show that

$$\begin{aligned} &|u_h(x_j) - u_{hk\Delta z}(x_j)| \\ &\le e^{-hc_0} e^{-\lambda h} |u_{hk\Delta z} - u_h|_0 + e^{-hc_0}(1 - e^{-\lambda h}) \Big[\sum_j \bar{P}_{ij} |u_h(x_j) - u_{hk\Delta z}(x_j)| \\ &+ \underbrace{\Big| \sum_j \bar{P}_{ij} u_h(x_j) - \sum_j P_{ij} u_h(x_j) \Big|}_{A} + \underbrace{\Big| \sum_j P_{ij} u_h(x_j) - \frac{1}{\lambda} \int_E u_h(x_i + \eta(x_i, z, v))\nu(dz) \Big|}_{B} \Big]. \end{aligned}$$

Note that $\sum_j \bar{P}_{ij} = \frac{1}{\lambda} Q_{\Delta z}[1] = 1$ by (3.13) since $\sum \beta_j = 1$. Furthermore,

$$\begin{aligned} A &= \Big| \frac{1}{\lambda} \int_E I_k u_h(x_i + \eta(x_i, v, z)) m(z) dz - Q_{\Delta z}\Big[I_k u_h(x_i + \eta(x_i, v, \cdot)) m(\cdot)\Big] \Big| \\ &\le \frac{\Delta z}{\lambda} \big\| D\big[I_k u_h(x_i + \eta(x_i, v, z)) m(z)\big] \big\|_{L^1} \le \frac{\Delta z}{\lambda} \Big(|u_h|_1 L_3 C_0 + |u_h|_0 C_1\Big) \int_E e^{-\varepsilon|z|} dz, \\ B &= \Big| \frac{1}{\lambda} \int_E (u_h - I_k u_h)(x_i + \eta(x_i, v, z))\nu(dz) \Big| \le k|u_h|_1 \frac{\int_E \nu(dz)}{\lambda} = k|u_h|_1, \end{aligned}$$

where $I_k\phi(x) = \sum_i \beta_i(x)\phi(x_i)$ is piecewise linear interpolation of ϕ on X^k and we have used (B2), (A4), and (3.13). Combining these estimates and (3.15), using the properties of $\beta_i(x)$, and remembering that $h \in (0,1]$ then gives the result. □

Remark 3.6. Since $h \in (0,1]$, $|u_h - u_{hk\Delta z}|_0 \le \frac{C|u_h|_1}{c_0} \frac{k+\Delta z}{h}$, which is consistent with the estimates obtained in [16] in a different setting.

Remark 3.7. If we use the cut-off procedure explained in Section 2.3 (a way of reducing to a bounded domain), then equation (3.14) gives a finite system of equations. In view of the integral term the new effective domain then becomes $B_\mu + \mathrm{supp}(\nu)$ (see Section 2.3).

3.3. **Schemes for unbounded measures** ν**.** We consider general unbounded Levy measures ν under assumption (B3). There are two different cases: (i) $\alpha \in [0,1)$ with HJB equation (1.1), and (ii) $\alpha \in [1,2)$ with HJB equation (1.2). To derive our schemes, we first reduce to a bounded Levy measure $\nu_{r,R}$ as explained in Section 2. The result are the approximate HJB equations (2.2) and (2.6). These equations are then approximated by a slightly modified version of the semi-Lagrangian scheme (3.8) (or equivalently (3.7)) defined in Section 3.1.

We propose the following semi-Lagrangian scheme in case (i)

$$v_h(x) = \inf_{v\in V}\Big\{hf(x,v) + e^{-hc(x,v)}\Big[\frac{e^{-h\lambda_{r,R}}}{4d}\sum_{m=1}^{d}\Big(v_h(x+h\bar{b}(x,v)+\sqrt{h}\bar{\sigma}_{+,m}(x,v)) \tag{3.16}$$
$$+ v_h(x+h\bar{b}(x,v)+\sqrt{h}\bar{\sigma}_{-,m}(x,v)) + v_h(x+h\bar{b}(x,v)-\sqrt{h}\bar{\sigma}_{+,m}(x,v))$$
$$+ v_h(x+h\bar{b}(x,v)-\sqrt{h}\bar{\sigma}_{-,m}(x,v))\Big) + \frac{1-e^{-h\lambda_{r,R}}}{\lambda_{r,R}}\int_E v_h(x+\eta(x,v,z))\nu_{r,R}(dz)\Big]\Big\},$$

and in case (ii)

$$w_h(x) = \inf_{v\in V}\Big\{hf(x,v) + e^{-hc(x,v)}\Big[\frac{e^{-h\lambda_{r,R}}}{4d}\sum_{m=1}^{d}(w_h(x+h\tilde{b}(x,v)+\sqrt{h}\bar{\sigma}_{+,m}(x,v)) \tag{3.17}$$
$$+ w_h(x+h\tilde{b}(x,v)+\sqrt{h}\bar{\sigma}_{-,m}(x,v)) + w_h(x+h\tilde{b}(x,v)-\sqrt{h}\bar{\sigma}_{+,m}(x,v))$$
$$+ w_h(x+h\tilde{b}(x,v)-\sqrt{h}\bar{\sigma}_{-,m}(x,v)) + \frac{1-e^{-h\lambda_{r,R}}}{\lambda_{r,R}}\int_E w_h(x+\eta(x,v,z))\nu_{r,R}(dz)\Big]\Big\},$$

where $\lambda_{r,R} := \int_E \nu_{r,R}(dz)$, $\bar{b}, \tilde{b}, \nu_{r,R}$ are defined in Section 2, and $\bar{\sigma}_{\pm,m}$ is m-th column of

$$\bar{\sigma}_{\pm}(x,v) = \sigma(x,v) \pm \sqrt{\int_{0<|z|<r}\eta(x,v,z)\eta(x,v,z)^T\nu(dz)}, \tag{3.18}$$

where the square root denotes the matrix square root.

Remark 3.8. The additional terms in (3.16) and (3.17) compared with (3.8), enable us to use $\sigma \pm \sqrt{\int \eta\eta^T\nu}$ instead of $\sqrt{\sigma\sigma^T + \int \eta\eta^T\nu}$ as diffusion matrix. The consistency relation Proposition 3.1 (iv) still holds, and if $\eta(x,v,z) = \eta_1(x,v)\eta_2(z)$, then the square root in (3.18) equals $C\eta_1(x,v)$ where C is the precomputable constant matrix $\sqrt{\int_{0<|z|<r}\eta_2(z)\eta_2^T(z)\nu(dz)}$.

Remark 3.9. These schemes are similar to (3.8), and can be derived in a similar way. The conclusions of Propositions 3.1 and 3.3 (when b-terms are replaced by $\bar{b}$- or $\tilde{b}$-terms as defined in Section 2) still hold for solutions of (3.16) and (3.17). We refer to [30] for the technical modifications needed to handle the integral term in the diffusion coefficients.

Remark 3.10. Previously bounded quantities may blow up as $r \to 0$. Indeed by (B3) and (A2) we have for $r \in (0,1)$,

$$\lambda_{r,R} \leq \frac{K}{\alpha r^\alpha} \quad \text{for } \alpha \in (0,2), \qquad \Big|\int_E \eta\nu_{r,R}(dz)\Big| \leq \begin{cases} \frac{L_2K}{(\alpha-1)r^{\alpha-1}} & \text{for } \alpha \in (1,2) \\ L_2K & \text{for } \alpha \in (0,1). \end{cases} \tag{3.19}$$

4. Convergence estimates for the discrete-time problem

In this section we prove a priori error bounds for the convergence of solutions u_h of the semi-discrete scheme (3.8) to the unique viscosity solution u of (1.1). We consider two cases: (i) The measure ν is bounded and $(B1)_0$ holds and (ii) the measure ν is the truncation of an unbounded measure satisfying (B3).

In view of the equi-boundedness, equi-continuity and consistency results of Propositions 3.1 and 3.3, u_h converge locally uniformly to u by the Arzela-Ascoli Theorem, stability and uniqueness results for viscosity solutions (see e.g. [4]), and the consistency result in Proposition 3.1. It is also possible to obtain convergence without equi-continuity (i.e. under weaker assumptions on the coefficients) using so-called half relaxed limits [7]. Such results are given in [14] for some non-local equations, but these results does not cover the HJB equations we consider here.

Now we proceed to obtain a priori estimates for the convergence of u_h to u. To do this we will make use of an abstract result in [30], which we will describe below. Consider the equation

$$F(x,u,Du,D^2u,u(\cdot)) = 0 \qquad x \in \mathbb{R}^N \tag{4.1}$$

where $u(\cdot)$ represents non-local (integral) terms. Let $h > 0$ be a discretization parameter and consider an approximation scheme for (4.1) written in abstract form as

$$S(h, x, u_h(x), [u_h]_x) = 0 \qquad x \in \mathbb{R}^N, \tag{4.2}$$

where $[u_h]_x$ represents a function defined at x via all the possible value of u_h. We need the following set of assumptions.

(C1) *(Monotonicity)* There exists $\bar{c}_0 > 0$ such that for any $h > 0$, $x \in \mathbb{R}^N$, $\zeta \in \mathbb{R}$, $\tau > 0$ and bounded functions u, v such that $u \leq v$ in $\mathbb{R}^N$, then

$$S(h, x, \zeta + \tau, [u + \tau]_x) \geq S(h, x, \zeta, [v]_x) + \bar{c}_0 \tau$$

(C2) *(Regularity)* For any $h > 0$ and any continuous, bounded function ϕ, the function

$$x \mapsto S(h, x, \phi(x), [\phi]_x)$$

is bounded and continuous on $\mathbb{R}^N$ and the function

$$\zeta \mapsto S(h, x, \zeta, [\phi]_x)$$

is uniformly continuous for bounded ζ, uniformly in x.

(C3) *(Consistency)* There exists a function $E(\tilde{K}, h, \epsilon)$ such that for any sequence $\{\phi_\varepsilon\}_\varepsilon$ of smooth functions satisfying

$$|D^\beta \phi_\epsilon(x)| \leq \tilde{K} \epsilon^{1-|\beta|} \quad \text{in } \mathbb{R}^N, \qquad \text{for any } \beta \in \mathbb{N}^N,$$

where $|\beta| = \sum_{i=1}^N \beta_i$, the following inequality holds:

$$|S(h, x, \phi_\epsilon(x), [\phi_\epsilon]_x) - F(x, \phi, D\phi_\epsilon, D^2\phi_\epsilon)| \leq E(\tilde{K}, h, \epsilon) \quad \text{in } \mathbb{R}^N.$$

(C4) *(Convexity)* Let $(\rho_\varepsilon)_{\varepsilon>0}$ be a family of mollifiers (smooth, positive functions with mass 1 and support in $\{|x| < \varepsilon\}$). For any Lipschitz-continuous function ϕ, there exists a constant C such that for any x and h

$$\int_{\mathbb{R}^N} S(h, x, \phi(x - e), [\phi(\cdot - e)]_x)\rho(e) de \geq S(h, x, (\phi * \rho_\varepsilon)(x), [(\phi * \rho_\varepsilon)]_x) - C\varepsilon$$

(C5) *(Commutation with translations)* For any $h > 0$ small enough, $0 < \varepsilon < 1$, $y \in \mathbb{R}^N$, $\zeta \in \mathbb{R}$, continuous bounded function ϕ and $|e| \leq y$, we have

$$S(h, y, \zeta, [\phi]_{y-e}) = S(h, y, \zeta, [\phi(\cdot - e)]_y).$$

(D) For h small enough and $\varepsilon \in [0, 1)$, there is a unique solution u_h^ε of the scheme

$$\max_{|e| \leq \varepsilon} S(h, x + e, u_h^\varepsilon(x), [u_h^\varepsilon]_x) = 0 \quad \text{in } \mathbb{R}^N, \tag{4.3}$$

where $u_h := u_h^0$ also solve (4.2), and a constant C independent of h, ε such that

$$|u_h^\varepsilon|_1 \leq C \qquad \text{and} \qquad |u_h^0 - u_h^\varepsilon|_0 \leq C\varepsilon.$$

We remark that we are using a more general consistency relation here than in [30], and that this extra generality will be needed when we consider unbounded measures ν. The next result is a restatement of Theorem 3.4 in [30] in view of the new consistency relation (C3).

Theorem 4.1. *Assume (A1) – (A3), (B1)$_2$, (C1) – (C5), and (D) hold, and let u and u_h be solutions of respectively* (4.1) *and* (4.2) *satisfying $\tilde{K} := |u|_1 \vee |u_h|_1 < \infty$. Then there exists a constant C depending only on L_1, L_2, c_0 from (A2) and (A3) such that*

$$|u - u_h| \leq C \min_{\epsilon > 0} \left(\epsilon + E_1(\tilde{K}, h, \epsilon)\right) \quad in \quad \mathbb{R}^N.$$

Remark 4.2. To prove (D) for the scheme (3.8), we will need to assume also (B1)$_0$, $h\lambda \leq \bar{C}_0$, $h \leq 1$, and $c_0 \geq c_1$ for both c_1 defined in Theorem 2.1 and Proposition 3.3. Under these assumptions we also have $\tilde{K} := |u|_1 \vee |u_h|_1 < \infty$.

We will apply this abstract result to derive error bounds for the scheme (3.8). We rewrite the scheme in the form (4.2) with

$$
\begin{aligned}
&S(h,x,r,[\psi]_x) = \\
&\sup_{v\in V}\Big\{\frac{-e^{-\lambda h}}{2dh}\sum_{m=1}^{d}\Big[[\psi]_x(hb(x,v)+\sqrt{h}\sigma_m(x,v))-2r+[\psi]_x(hb(x,v)-\sqrt{h}\sigma_m(x,v))\Big] \\
&-\frac{1-e^{-\lambda h}}{\lambda h}\int_E[\psi(x+\eta(x,v,z))-r]+\frac{e^{hc(x,v)}-1}{h}r-e^{hc(x,v)}f(x,v)\Big\},
\end{aligned}
\tag{4.4}
$$

and $[\psi]_x(z)=\psi(x+z)$.

Lemma 4.3. *Assume (A1) – (A3) and (B1)$_0$, $h\lambda\leq\bar{C}_0$, $h\leq 1$. Then the scheme* (3.8) *(and equivalently* (4.4)*) satisfies assumptions (C1)–(C5) with*

$$E(\tilde{K},h,\varepsilon)=C_1h(\tilde{K}\varepsilon^{-1}+\tilde{K}\varepsilon^{-2}+\tilde{K}\varepsilon^{-3})+C_2h\lambda((1+|\int_E\eta\nu|)\tilde{K}+\tilde{K}\varepsilon^{-1}),$$

where the constant C_1 and C_2 only depend on $\sup_v|\sigma|_0,\sup_v|b|_0$.

If in addition $c_0\geq c_1$ for both c_1's in Theorem 2.1 and Proposition 3.3, then assumption (D) also holds with constants C only depending on the data and $\bar{C}_0$.

Proof. It is straightforward to verify (C1) with $\overline{c}_0=c_0$ where c_0 is defined in (A3). (C2) follows from the assumptions on the coefficients, while (C3) follows from Proposition 3.1 (iv). By a straight forward computation, it follows that (C4) holds with $C=0$. We refer to [30] for similar computations. Finally, (C5) holds since $[\phi]_{x-e}=[\phi(\cdot-e)]_x$ for any continuous function ϕ.

To prove (D), observe that (4.3) can be rewritten in the form (4.4) (by defining a new control $\bar{v}=(v,e)$). The coefficients of this new equation still satisfy (A1) - (A3). Therefore (D) follows after an application of Propositions 3.1 and 3.3. □

Now we are in a position to state the error bounds. First we consider the bounded case, i.e (B1)$_0$ holds. In this case the equation is (1.1) which is approximated by the scheme (3.8). Note that the integral operator has not yet been discretized.

Theorem 4.4 (Bounded measure)**.** *Assume (A1) – (A3), (B1)$_0$, $h\leq 1$, and $c_0\geq c_1$ for both c_1's in Theorem 2.1 and Proposition 3.3. Let u be the solution of* (1.1) *and u_h be the solution of* (3.8) *(or equivalently* (4.4)*).*

(a) (General IPDEs) Then $|u-u_h|_0\leq Ch^{1/4}$.

(b) (1st order IPDEs) If in addition $\sigma\equiv 0$, then $|u-u_h|_0\leq Ch^{1/2}$.

In both cases the constant C depends only on the coefficients, c_1, and λ.

Proof. Part (a) is an easy consequence of Theorem 4.1 and Lemma 4.3. Part (b) follows in a similar manner after noting that the consistency relation corresponding to Proposition 3.1 *(iv)* now becomes

$$|L^v\psi(x)-L_h^v\psi(x)|\leq C_1h|D^2\phi|_0+C_2h\lambda((1+|\int_E\eta\nu|)|D\phi|_0+|D^2\phi|_0).$$

□

Remark 4.5. The convergence rate obtained in Theorem 4.4 is the same as in the pure PDE case, see [27, 5]. In the first order case and when the Levy measure is bounded, convergence rate 1/2 for semi-Lagrangian schemes like (3.8) have previously be obtained in [16]. However the integral term in [16] has a different form compared to the one we consider here.

Remark 4.6. The scheme (3.8) uses a first order accurate approximation of 2nd derivatives as can be seen from the consistency relation Proposition 3.1 *(iv)*. This leads to lower rates of convergence than for some monotone FDMs that use 2nd order accurate approximations of 2nd derivatives, see [33, 30, 10]. There the rate is 1/2, while for a more general class of FDMs the rate is at least 1/5, see [6] for the pure PDE case.

We proceed to the case of general unbounded Levy measure ν under assumption (B3). There are two different cases: (i) $\alpha \in [0,1)$ with HJB equation (1.1), and (ii) $\alpha \in [1,2)$ with HJB equation (1.2). In case (i), $\alpha \in (0,1)$, as a consequence of Lemma 4.3, Theorems 4.1 and 2.3 we have the following convergence result.

Theorem 4.7 (Unbounded measure I). *Assume (A1) – (A3), (B3) with* $\alpha \in (0,1)$, $h \le 1$, *and* $c_0 \ge c_1$ *for both* c_1*'s in Theorem 2.1 and Proposition 3.3. Let* u *be the solution of* (1.1) *and* u_h *be the solution of* (3.16).

Then the best rate is obtained choosing $r = h^{\frac{3}{6+\alpha}}$, *and in this case*

$$|u - u_h|_0 \le C(h^{1/4} + h^{\frac{3-\alpha}{6+\alpha}}) \le Ch^{1/4},$$

where the constant C *depend only on the coefficients,* c_1, *and quantities from (B3)/(B1)*$_2$.

Proof. When $\alpha \in (0,1)$, Lemma 4.3 still holds for the scheme (3.16), and in view of (3.19) we have the following form of E,

$$E(\tilde{K}, h, r, \varepsilon) = C_1 \tilde{K} h(\varepsilon^{-1} + \varepsilon^{-2} + \varepsilon^{-3}) + C_2 \tilde{K} h r^{-\alpha}(1 + \varepsilon^{-1}),$$

where the constant C_1 and C_2 are independent of h, r, ε. Let u_r denote the solution of (2.2). Theorem 4.1 and ("term-wise") minimization in ε, lead to the bound

$$|u_r - u_h|_0 \le C(h^{1/4} + r^{-\alpha} h + r^{-\frac{\alpha}{2}} h^{\frac{1}{2}}),$$

where the constant C depend only on the coefficients, c_1, and quantities from (B3)/(B1)$_2$. In view of Theorem 2.3 and the optimal choice of r, $r = h^{\frac{3}{6+\alpha}}$, the result follows. □

Remark 4.8. Since $\frac{1}{4} < \frac{3-\alpha}{6+\alpha} < \frac{1}{2}$ for $\alpha \in (0,1)$, there is no reduction of rate due to truncation of the measure ν.

The case $\alpha \in (1,2)$ is more difficult, since now also $\int_E \eta \nu_{r,R}$ and hence $\tilde{b}$ in (3.17) blows up as $r \to 0$. As a consequence Theorem 4.1 can no longer be used directly. The convergence result is the following:

Theorem 4.9 (Unbounded measure II). *Assume (A1) – (A3), (B3) with* $\alpha \in (1,2)$, $h \le 1$, *and* $c_0 \ge c_1$ *for both* c_1*'s in Theorem 2.1 and Proposition 3.3. Let* u *be the solution of* (1.2) *and* u_h *be the solution of* (3.17).

Then the best rate is obtained choosing $r = h^{\frac{3}{3+5\alpha}}$, *and in this case*

$$|u - u_h|_0 \le Ch^{\frac{3-\alpha}{3+5\alpha}},$$

where the constant C *depend only on the coefficients,* c_1, *and quantities from (B3)/(B1)*$_2$.

Remark 4.10. This result is consistent with Theorem 4.7 since $\lim_{\alpha \to 1^+} \frac{3-\alpha}{3+5\alpha} = \frac{1}{4}$. For $\alpha \in (1,2)$ the rate degrades as α increases, and $\lim_{\alpha \to 2^-} \frac{3-\alpha}{3+5\alpha} = \frac{1}{13}$.

Outline of proof. From Proposition 3.3 we have uniform in r Lipschitz continuity, but the continuous dependence estimates will be proportional to $\int_E |z| \nu_{r,R}$ through the $\tilde{b}$-term. Because of this we must redo the arguments leading to Theorem 4.1, and the result will be an estimate of the form

$$|u_r - u_h| \le C \min_{\epsilon > 0} \left(\epsilon \int_E |z| \nu_{r,R} + E_1(\tilde{K}, h, r, \epsilon) \right) \quad \text{in} \quad \mathbb{R}^N,$$

when u_r solve (2.6). We omit the details, since the argument is same as the one used to prove Theorem 4.1 in [30]. Because of the blow up in $\int_E \eta \nu_{r,R}$, we also need a much more precise consistency relation than given in Proposition 3.1 (iv) tracking all $\tilde{b}$ dependence. From the proof of Proposition 3.1, it is easy to see that it will have the following form

$$\begin{aligned} E(\tilde{K}, h, \varepsilon) = C_1 \tilde{K} \Big[h |\tilde{b}|_0^2 \varepsilon^{-1} + (h^2 |\tilde{b}|_0^3 + h |\tilde{b}|_0) \varepsilon^{-2} + (h + h^3 |\tilde{b}|_0^4 + h^2 |\tilde{b}|_0^2) \varepsilon^{-3} \Big] \\ + C_2 \tilde{K} \lambda_r h \Big[|\tilde{b}|_0 + \varepsilon^{-1} \Big] + C_3 \tilde{K} \lambda_r h \Big[\int_E \eta \nu_{r,R} \Big], \end{aligned}$$

where the constants C_1, C_2, C_3 only depend on $\sup_v |\bar\sigma|_0$. In view of (3.19), $\lambda_r = Cr^{-\alpha}$ and $|\tilde b|_0 + \int_E |z|\nu_{r,R} \le C(1+r^{\alpha-1})$. The rest of the proof is a long computation consisting of choosing optimal ε and r as in the proof of Theorem 4.7. We omit the details only remarking that the worst term in E turns out to be the $h\varepsilon^{-3}$-term. □

Appendix A. Lipschitz regularity and continuous dependence

In this section we obtain a combined Lipschitz regularity and continuous dependence result for solutions to the equation

$$u_h(x) = \inf_{v\in V}\Big\{hf(x,v) + e^{-hc(x,v)}\Big[\frac{e^{-h\lambda}}{2d}\sum_{m=1}^{d}\Big(u_h(x+h\bar b(x,v)+\sqrt{h}\sigma_m(x,v))+ \\ + u_h(x+h\bar b(x,v)-\sqrt{h}\sigma_m(x,v)\Big) + \frac{1-e^{-h\lambda}}{\lambda}\int_E u_h(x+\eta(x,v,z))\nu(dz)\Big]\Big\}, \tag{A.1}$$

where $\lambda = \nu(E)$ and

$$\bar b = b + \frac{1-e^{-h\lambda}}{\lambda e^{-\lambda h}}\int_E \eta\nu.$$

Theorem A.1. *Let u_h and $\tilde u_h$ be solutions of* (A.1) *corresponding to the data $\sigma, \bar b, c, f, \eta, \nu$ and $\tilde\sigma, \bar{\tilde b}, \tilde c, \tilde f, \tilde\eta, \nu$ respectively, and assume both sets of coefficients satisfy (A1) – (A3) and (B1)$_0$, and that $h\lambda \le C$ and $h \le 1$. Then there exist constants $c_1, L, K \ge 0$ (only depending on the data and C) such that if $c_0 \ge c_1$, then for all $h > 0, x, y \in \mathbb{R}^N$,*

$$|u_h(x) - \tilde u_h(y)| \le L|x-y| + K\sup_{v\in\mathcal V}\big[|f-\tilde f|_0 + |c-\tilde c|_0 \\ + |b-\tilde b|_0 + |\sigma-\tilde\sigma|_0 + |\int|\eta(\cdot,z)-\tilde\eta(\cdot,z)|^2\nu(dz)|_0^{1/2}\big].$$

Remark A.2. The precise dependence of the constants c_1, L, K is given in the proof below.

Remark A.3. This result extends the corresponding results of [5] to non-local HJB equations, with general (singular) Levy measures. Moreover, the proof below simplifies the corresponding proofs of Barles and Jakobsen [5] because we do not use the somewhat unnatural "doubling schemes" as in [5]. Instead we give a more direct proof.

Proof. We will use doubling of variables techniques similar to those used to prove corresponding results for equation (1.1). We define

$$\phi(x,y) = m + \alpha M + \frac{L}{2}(\alpha^{-1} + \alpha|x-y|^2) + \varepsilon(|x|^2+|y|^2),$$
$$\psi(x,y) = u_h(x) - \tilde u_h(y) - \phi(x,y),$$

for $\alpha, \varepsilon > 0$, $m, M, L \ge 0$, and then we let $(\bar x, \bar y)$ be a point such that $\psi(\bar x,\bar y) = \sup_{x,y}\psi(x,y)$. We will prove that $\psi(\bar x,\bar y) \le o(1)$ as $\varepsilon \to 0$ for a suitable constant L, and for

$$m = \Big(\frac{e^{c_0 h}-1}{h}\Big)^{-1}\sup_{v\in\mathcal V} e^{|c|_0\vee|\tilde c|_0 h}\Big[|f-\tilde f|_0 + (|u_h|_0\wedge|\tilde u_h|_0 + h|f|_0\wedge|\tilde f|_0)|c-\tilde c|_0\Big], \tag{A.2}$$

$$M = K\Big(\frac{e^{c_0 h}-1}{h}\Big)^{-1}\sup_{v\in\mathcal V}\big[|\sigma-\tilde\sigma|_0^2 + |b-\tilde b|_0^2 + |\int|\eta(\cdot,z)-\tilde\eta(\cdot,z)|^2\nu(dz)|_0\big]. \tag{A.3}$$

This implies Theorem A.1 after sending $\varepsilon \to 0$. To see this, fix x, y and note that for any $\alpha > 0$,

$$u_h(x) - \tilde u_h(y) - m - \alpha M - \frac{L}{2}(\alpha^{-1}+\alpha|x-y|^2) \le \psi(\bar x,\bar y) + \varepsilon(|x|^2+|y|^2) \le o(1) \quad \text{as } \varepsilon\to 0.$$

In this inequality we send $\varepsilon \to 0$ and choose $\alpha^{-1} = |x-y| \vee (2M/L)^{1/2}$ to obtain

$$u_h(x) - \tilde u_h(y) \le m + 2(LM)^{1/2} + L|x-y|,$$

and hence Theorem A.1 follows since x, y were arbitrary.

For simplicity we will not be explicit about the form of the ε-terms appearing in the computations below. The role of these terms is only to guaranty that the maximum is attained at a point $(\bar{x}, \bar{y})$, and they vanish in the final estimate when $\varepsilon \to 0$. We refer to the proof of Theorem 3.4 in [5] for details concerning the ε-terms.

We proceed by contradiction assuming that $\psi(\bar{x}, \bar{y}) > o(1)$ as $\varepsilon \to 0$. Note that by the definition of ψ this implies that $u_h(\bar{x}) - u_h(\bar{y}) > 0$. Furthermore, observe that since $(\bar{x}, \bar{y})$ is a maximum point,

$$\begin{aligned}\psi(\bar{x}+b+a, \bar{y}+\bar{b}+\bar{a}) + \psi(\bar{x}+b-a, \bar{y}+\bar{b}-\bar{a}) &\le 2\psi(\bar{x}, \bar{y}),\\ \psi(\bar{x}+\zeta, \bar{y}+\bar{\zeta}) &\le \psi(\bar{x}, \bar{y}),\end{aligned}$$

for every $a, b, \zeta, \bar{a}, \bar{b}, \bar{\zeta} \in \mathbb{R}^N$, and hence by the definition of ψ,

$$\begin{aligned} I_1 &:= [u_h(\bar{x}+b+a) - 2u_h(\bar{x}) + u_h(\bar{x}+b-a)]\\ &\quad - [\tilde{u}_h(\bar{y}+\bar{b}+\bar{a}) - 2\tilde{u}_h(\bar{y}) + \tilde{u}_h(\bar{y}+\bar{b}-\bar{a})]\\ &\le \phi(\bar{x}+b+a, \bar{y}+\bar{b}+\bar{a}) - 2\phi(\bar{x}, \bar{y}) + \phi(\bar{x}+b-a, \bar{y}+\bar{b}-\bar{a}),\\ I_2 &:= [u_h(\bar{x}+\zeta) - u_h(\bar{x})] - [\tilde{u}_h(\bar{y}+\bar{\zeta}) - \tilde{u}_h(\bar{y})] \le \phi(\bar{x}+\zeta, \bar{y}+\bar{\zeta}) - \phi(\bar{x}, \bar{y}).\end{aligned}$$

Finally, by the definition of ϕ and the simple identities

$$\begin{aligned}&|\bar{x}+b \pm a - (\bar{y}+\bar{b} \pm \bar{a})|^2 = |\bar{x} \pm a - (\bar{y} \pm \bar{a})|^2 + 2(\bar{x} \pm a - (\bar{y} \pm \bar{a})) \cdot (b - \bar{b}) + |b - \bar{b}|^2,\\ &|\bar{x} - \bar{y} + (a - \bar{a})|^2 - 2|\bar{x} - \bar{y}|^2 + |\bar{x} - \bar{y} - (a - \bar{a})|^2 = 2|a - \bar{a}|^2,\end{aligned}$$

we are lead to

$$I_1 \le 2\frac{L}{2}\alpha|a-\bar{a}|^2 + 2\frac{L}{2}\alpha|b-\bar{b}|^2 + 4\frac{L}{2}\alpha(\bar{x}-\bar{y}, b-\bar{b}) + o(1) \qquad \text{as } \varepsilon \to 0, \tag{A.4}$$

$$I_2 \le 2\frac{L}{2}\alpha(\bar{x}-\bar{y}, \zeta-\bar{\zeta}) + \frac{L}{2}\alpha|\zeta-\bar{\zeta}|^2 + o(1) \qquad \text{as } \varepsilon \to 0. \tag{A.5}$$

These two inequalities are crucial for the rest of the proof.

Now we divide (A.1) by he^{-hc} and rewrite it as

$$\begin{aligned}0 = \sup_{v \in V}\Big\{&\frac{e^{hc}-1}{h}u_h(x) - e^{hc}f(x,v)\\ &- \frac{e^{-h\lambda}}{2dh}\sum_{m=1}^{d}\Big(u_h(x+h\bar{b}(x,v)+\sqrt{h}\sigma_m(x,v)) - 2u_h(x) + u_h(x+h\bar{b}(x,v)-\sqrt{h}\sigma_m(x,v)\Big)\\ &- \frac{1-e^{-h\lambda}}{\lambda h}\int_E\Big(u_h(x+\eta(x,v,z)) - u_h(x)\Big)\nu(dz)\Big\}.\end{aligned}$$

Upon subtracting the equations (in this new form) for $\tilde{u}_h$ and u_h, we get

$$\begin{aligned}0 \le \sup_{v \in V}\Big\{&\frac{e^{\tilde{c}h}-1}{h}\tilde{u}_h(\bar{y}) - \frac{e^{ch}-1}{h}u_h(\bar{x})\\ &+ [e^{ch}f(\bar{x},v) - e^{\tilde{c}h}\tilde{f}(\bar{y},v)] + \frac{e^{-h\lambda}}{2dh}\sum_{m=1}^{d}I_{1,m} + \frac{1-e^{-h\lambda}}{\lambda h}\int_E I_2\nu(dz)\Big\},\end{aligned}$$

where $I_{1,m}$ corresponds to I_1 above with the choice $a = \sqrt{h}\sigma_m(\bar{x}, v)$, $b = h\bar{b}(\bar{x}, v)$, $\bar{a} = \sqrt{h}\tilde{\sigma}_m(\bar{y}, v)$, and $\bar{b} = h\bar{\tilde{b}}(\bar{y}, v)$, and for I_2 we have taken $\zeta = \eta(\bar{x}, v, z)$ and $\bar{\zeta} = \eta(\bar{y}, v, z)$. By (A.4) and (A.5) we then have

$$
\begin{aligned}
0 \le \sup_{v\in V} \Big\{ & \frac{e^{\tilde{c}h}-1}{h}\tilde{u}_h(\bar{y}) - \frac{e^{ch}-1}{h}u_h(\bar{x}) \\
& + [e^{ch}f(\bar{x},v) - e^{\tilde{c}h}\tilde{f}(\bar{y},v)] + \frac{e^{-h\lambda}}{2h}\frac{L}{2}\alpha\Big[\frac{1}{d}\sum_{m=1}^{d} 2h|\sigma_m(\bar{x},v)-\tilde{\sigma}_m(\bar{y},v)|^2 \\
& + 2h^2|\bar{b}(\bar{x},v)-\bar{\tilde{b}}(\bar{y},v)|^2 + 4h(\bar{x}-\bar{y},\bar{b}(\bar{x},v)-\bar{\tilde{b}}(\bar{y},v))\Big] \\
& + e^{-hc}\frac{1-e^{-h\lambda}}{\lambda h}\frac{L}{2}\alpha\int_E \Big[|\eta(\bar{x},v,z)-\tilde{\eta}(\bar{y},v,z)|^2 + 2\big(\bar{x}-\bar{y},\eta(\bar{x},v,z)-\tilde{\eta}(\bar{y},v,z)\big)\Big]\nu(dz)\Big\} \\
& + o(1) \quad \text{as } \varepsilon \to 0.
\end{aligned}
$$

Since $\bar{b} = b + \frac{1-e^{-h\lambda}}{\lambda e^{-\lambda h}}\int_E \eta\nu$ and $\bar{\tilde{b}}$ is defined similarly we see that the $(\bar{x}-\bar{y},\eta-\tilde{\eta})$-terms cancel in the above inequality. Since $\nu(E)=\lambda$, Jensen's inequality implies that

$$
\lambda^2\Big|\int_E (\eta(\bar{x},v,z)-\tilde{\eta}(\bar{y},v,z))\frac{\nu(dz)}{\lambda}\Big|^2 \le \lambda^2\int_E |\eta(\bar{x},v,z)-\tilde{\eta}(\bar{y},v,z)|^2\frac{\nu(dz)}{\lambda}.
$$

Also note that since $u_h(\bar{x})-\tilde{u}_h(\bar{y})>0$ and $c\ge c_0>0$,

$$
\frac{e^{\tilde{c}h}-1}{h}\tilde{u}_h(\bar{y}) - \frac{e^{ch}-1}{h}u_h(\bar{x}) \le -\frac{e^{c_0h}-1}{h}[u_h(\bar{x})-\tilde{u}_h(\bar{y})] + \frac{|e^{\tilde{c}h}-e^{ch}|}{h}|u_h(\bar{x})|\wedge|\tilde{u}_h(\bar{y})|.
$$

Therefore after cancellations, Jensen's inequality, and the inequality $\frac{1-e^{-h\lambda}}{\lambda e^{-\lambda h}} \le he^{h\lambda}$, we get

$$
\begin{aligned}
& \frac{e^{c_0h}-1}{h}[u_h(\bar{x})-\tilde{u}_h(\bar{y})] \\
& \le \sup_{v\in V}\Big\{ e^{|c|_0\vee|\tilde{c}|_0h}\Big[|f(\bar{x},v)-\tilde{f}(\bar{y},v)| + (|u_h|_0\wedge|\tilde{u}_h|_0 + h|f|_0\wedge|\tilde{f}|_0)|c(\bar{x},v)-\tilde{c}(\bar{y},v)|\Big] \\
& \quad + \frac{e^{-h\lambda}}{2h}\frac{L}{2}\alpha\Big[\frac{1}{d}\sum_{m=1}^{d}2h|\sigma_m(\bar{x},v)-\tilde{\sigma}_m(\bar{y},v)|^2 \\
& \quad + 4h^2|b(\bar{x},v)-\tilde{b}(\bar{y},v)|^2 + 4h(\bar{x}-\bar{y},b(\bar{x},v)-\tilde{b}(\bar{y},v))\Big] \\
& \quad + \frac{1-e^{-h\lambda}}{\lambda h}\frac{L}{2}\alpha\int_E\Big[4h^2\lambda he^{\lambda h}\int_E|\eta(\bar{x},v,z)-\tilde{\eta}(\bar{y},v,z)|^2\nu(dz) \\
& \quad + |\eta(\bar{x},v,z)-\tilde{\eta}(\bar{y},v,z)|^2\Big]\nu(dz)\Big\} + o(1) \quad \text{as } \varepsilon\to 0.
\end{aligned}
$$

Now since $\lambda h \leq C$ and $h < 1$, it follows from simple computations that

$$\begin{aligned}
&\frac{e^{c_0 h}-1}{h}[\psi(\bar{x},\bar{y})+\phi(\bar{x},\bar{y})] = \frac{e^{c_0 h}-1}{h}[u_h(\bar{x})-\tilde{u}_h(\bar{y})] \\
&\leq \sup_{v\in V}\Big\{\underbrace{e^{|c|_0\vee|\tilde{c}|_0 h}\Big[|f-\tilde{f}|_0+(|u_h|_0\wedge|\tilde{u}_h|_0+h|f|_0\wedge|\tilde{f}|_0)|c-\tilde{c}|_0\Big]}_{\frac{e^{c_0 h}-1}{h}m} \\
&\quad+\frac{L}{2}\alpha\frac{e^{-h\lambda}}{2}\Big[\frac{1}{d}\sum_{m=1}^{d}4|\sigma_m-\tilde{\sigma}_m|_0^2+16h|b-\tilde{b}|_0^2\Big] \\
&\quad+\frac{L}{2}\alpha 5Ce^{C}|\int_E|\eta(\cdot,v,z)-\tilde{\eta}(\cdot,v,z)|^2\nu(dz)|_0\Big\} \\
&\quad+\sup_{v\in V}\Big\{\underbrace{e^{|c|_0\vee|\tilde{c}|_0 h}\Big[L_f+(|u_h|_0\wedge|\tilde{u}_h|_0+h|f|_0\wedge|\tilde{f}|_0)L_c\Big]}_{\overline{L}}\frac{1}{2}(\alpha^{-1}+\alpha|\bar{x}-\bar{y}|^2) \\
&\quad+\frac{L}{2}\alpha|\bar{x}-\bar{y}|^2\underbrace{\Big(\frac{e^{-h\lambda}}{2}\Big[\frac{1}{d}\sum_{m=1}^{d}4L_\sigma^2+16hL_b+8\Big]+5Ce^{C}L_\eta^2\int_E|z|^2\nu(dz)\Big)}_{L_0}\Big\} \\
&\quad+o(1)\quad\text{as }\varepsilon\to 0.
\end{aligned}$$

Let $\overline{L}, L_0$ be defined as in the inequality above. If c_0 is so big that

$$\frac{e^{c_0 h}-1}{h}-L_0>0,$$

and we choose m, M as in (A.2) and (A.3) for K big enough, and

$$L=\frac{\overline{L}}{\frac{e^{c_0 h}-1}{h}-L_0},$$

then $\psi(\bar{x},\bar{y})\leq o(1)$ as $\varepsilon\to 0$ and the proof is complete. $\square$

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, 93–117

[2] S. Asmussen and J Rosinski. Approximations of small jumps of Lévy processes with a view towards simulation. *J. Appl. Probab.* 38(2): 482–493, 2001.

[3] 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.

[4] G. Barles and C. Imbert. Second-Order Elliptic Integro-Differential Equations: Viscosity Solutions' Theory Revisited. *Ann. Inst. H. Poincare Anal. Non Lineaire*, 25(3): 567-585, 2008.

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

[6] G. Barles and E. R. Jakobsen. Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.* 43(2):540-558, 2005.

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

[8] F. E. Benth, K. H. Karlsen, and K. Reikvam. Portfolio optimization in a Lévy market with intertemporal substitution and transaction costs. *Stoch. Stoch. Rep.*, 74(3-4):517–569, 2002.

[9] D.P.Bertsekas and S.Shreve. *Stochastic optimal control: the discrete time case.* Academic Press, New York, 1978.

[10] I. H. Biswas, E. R. Jakobsen, and K. H. Karlsen. Error estimates for finite difference-quadrature schemes for fully nonlinear degenerate parabolic integro-PDEs. *J. Hyperbolic Differ. Equ.*, 5(1): 187-219, 2008.

[11] 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. Submitted, 2008.

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

[13] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods.* Springer-Verlag, New York, 2002.

[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 (4):607–646, 2004.
[15] M. Briani, R. Natalini, and G. Russo. Implicit-explicit numerical schemes for jump-diffusion processes. *CALCOLO*, 44(1): 33–57, 2007.
[16] F. Camilli Approximation of integro-differential equations associated with piecewise deterministic process. *Optimal Control Appl. Methods* 18(6): 423–444, 1997.
[17] F. Camilli and M. Falcone, An approximation scheme for the optimal control of diffusion processes, RAIRO Modél. Math. Anal. Numér. 29:97–122, 1995.
[18] P. Carr, H. Geman, D. B. Madan, and M. Yor. Stochastic volatility for Lévy processes. *Math. Finance* 13 (3):345–382, 2003.
[19] R. Cont and P. Tankov. *Financial modelling with jump processes.* Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.
[20] R. Cont and E. Voltchkova. Finite difference methods for option pricing in jump-diffusion and exponential Levy models. *SIAM J. Numer. Anal.* 43(4): 1596-1626, 2005.
[21] M. G. Crandall and P.-L. Lions Two approximations of solutions of Hamilton-Jacobi equations. *Math. Comp.* 43(167): 1–19, 1984.
[22] S. Elganjoui and K. H. Karlsen. A markov chain approximation scheme for a singular investment-consumption problem with Lévy driven stock prices. available at the URL `http://www.math.uio.no/~kennethk/`, 2002.
[23] Y. d'Halluin, P. A. Forsyth, and G. Labahn. A penalty method for American options with jump-diffusion processes. *Numer. Math.*, 97(2):321–352, 2004.
[24] M. Falcone Appendix A in M. Bardi and I. Capuzzo Dolcetta *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman equations*, Birkhäuser, Boston, 1997.
[25] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions.* Second edition. Springer, New York, 2006.
[26] I. I. Gihman and A. V. Skorohod *Stochastic differential equations. Translated from the Russian by Kenneth Wickwire.* Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 72. Springer-Verlag, New York-Heidelberg, 1972.
[27] 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. (M3AS)* 13(5): 613-644, 2003.
[28] E. R. Jakobsen and K. H. Karlsen. Continuous dependence estimates for viscosity solutions of integro-PDEs. *J. Differential Equations*, 212(2):278–318, 2005.
[29] E. R. Jakobsen and K. H. Karlsen. A "maximum principle for semicontinuous functions" applicable to integro–partial differential equations. *NoDEA Nonlinear Differential Equations Appl.* 13:137-165, 2006.
[30] 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.
[31] N. V. Krylov. On the rate of convergence of finite-difference approximations for Bellman's equations. *Algebra i Analiz*, 9(3):245–256, 1997.
[32] 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.
[33] 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.
[34] H. J. Kushner and P. Dupuis. *Numerical methods for stochastic control problems in continuous time. Second edition.* Applications of Mathematics 24. Stochastic Modelling and Applied Probability. Springer-Verlag, New York, 2001.
[35] P.-L. Lions Existence results for first-order Hamilton-Jacobi equations. *Ricerche Mat.* 32(1):3–23, 1983.
[36] A.-M. Matache, T. von Petersdorff, and C. Schwab. Fast deterministic pricing of options on Lévy driven assets. *M2AN Math. Model. Numer. Anal.*, 38(1):37–71, 2004.
[37] J. L. Menaldi Some estimates for finite difference approximations. *SIAM J. Control Optim.* 27: 579–607, 1989.
[38] R. Merton. Option prices when the underlying stock returns are discontinuous. *J. Financial Economics,* 3:125–144, 1976.
[39] H. Pham. Optimal stopping of controlled jump-diffusion processes: a viscosity solution approach. *J. Math. Systems Estim. Control*, 8(1), 1998.
[40] T. H. Rydberg. The normal inverse Gaussian Lévy process: simulation and approximation. Heavy tails and highly volatile phenomena. *Comm. Statist. Stochastic Models* 13(4):887–910,1997.
[41] A. Sayah. Équations d'Hamilton-Jacobi du premier ordre avec termes intégro-différentiels. I. Unicité des solutions de viscosité. *Comm. Partial Differential Equations*, 16(6-7):1057–1074, 1991.

Dip. di Matematica Pura e Applicata, Univ. dell'Aquila, loc. Monteluco di Roio, 67040 l'Aquila, ITALY
E-mail address: `camilli@ing.univaq.it`

Department of Mathematical Sciences, Norwegian University of Science and Technology, 7491 Trondheim, NORWAY
E-mail address: `erj@math.ntnu.no`