REGULARITY OF THE OPTIMAL STOPPING PROBLEM FOR LÉVY PROCESSES WITH NON-DEGENERATE DIFFUSIONS

ERHAN BAYRAKTAR AND HAO XING

Abstract. The value function of an optimal stopping problem for a process with Lévy jumps is known to be a generalized solution of a variational inequality. Assuming the diffusion component of the process is nondegenerate and a mild assumption on the singularity of the Lévy measure, this paper shows that the value function is smooth in the continuation region for problems with either finite or infinite variation jumps. Moreover, the smooth-fit property is shown via the global regularity of the value function. This paper confirms the intuition that the nondegenerate diffusion component dictates the regularity of the value function in the optimal stopping problem for jump processes.

1. Introduction

This paper studies the finite horizon optimal stopping problem for an n-dimensional jump diffusion process X. In a filtered complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such a process $X = \{X_t; t \geq 0\}$ is governed by the following stochastic differential equation:

$$dX_t = b(X_{t-}, t)\, dt + \sigma(X_{t-}, t)\, dW_t + d\mathcal{J}_t, \tag{1.1}$$

in which $W = \{W_t; t \geq 0\}$ is the d-dimensional standard Brownian motion under $\mathbb{P}$ and $\mathcal{J} = \{\mathcal{J}_t; t \geq 0\}$ is a pure jump Lévy process independent of the Brownian motion. This jump process $\mathcal{J}$ can be of finite/infinite activity with finite/infinite variation. We denote the Lévy measure of $\mathcal{J}$ as ν (please refer to Section 2 for the definition of $\mathcal{J}$ and its properties).

We investigate the problem of maximizing the discounted terminal reward g by optimally stopping the process X before a fixed time horizon T. The value function of this problem is defined as

$$u(x, t) = \sup_{\tau \in \mathcal{T}_{0,T-t}} \mathbb{E}\left[e^{-r\tau} g(X_\tau) \middle| X_0 = x \right], \tag{1.2}$$

in which $\mathcal{T}_{0,t}$ is the set of all stopping times (with respect to the filtration $(\mathcal{F})_{0 \leq s \leq t}$) valued between 0 and t. A specific example of such an optimal stopping problem is the American option pricing problem, where X models the logarithm of the stock price process and g represents the pay-off function.

This value function satisfies, at least intuitively, a variational inequality with a nonlocal integral term (see e.g. Chapter 3 of [2]). In general, the value function is not expected to be a smooth solution of the variational inequality. Therefore, notions of generalized solutions are needed to characterize the value function. In the literature, different solution concepts were studied. Pham showed in [21] that the value function of the optimal stopping problem for a controlled jump process is a viscosity solution of a variational inequality using the dynamic programming principle. In [18], Lamberton and Mikou proved that the value function associated to the optimal stopping problem on Lévy processes can be understood as the solution in the distribution sense.

Key words and phrases. Optimal Stopping, Variational Inequality, Lévy Process, Nondegenerate Diffusion, Classical Solution, Smooth Fit.

This research is partially supported by the National Science Foundation.

When the jump process X has a nondegenerate diffusion component, intuition tells us that the nondegenerate diffusion component should dominate the jump component, in the sense that the value function can be characterized as a smooth function. This intuition has been confirmed for the partial integro-differential equations associated to the Cauchy problem (e.g. the European option pricing problem) and boundary value problems. For these problems, Sections 1-3 in Chapter 3 of [2] and [11] proved the existence and uniqueness of second order partial integro-differential equations in both Sobolev and Hölder spaces. These regularity properties ensure that the Cauchy problem and boundary value problems have smooth solutions, as long as the diffusion component is nondegenerate.

On the other hand, for variational inequalities associated to the optimal stopping problems with either finite or infinite activity jumps, Bensoussan and Lions showed in Theorem 4.4 of [2] pp. 250 that the solution of a variational inequality on a bounded domain can be characterized as an element in a certain Sobolev space. These types of variational inequalities were also studied in Chapter 6 of [12], where jumps are assumed to be restricted in the bounded domain of the problem. The regularity results for the variational inequality in [2] are not enough to ensure the smooth-fit property to hold. Later, these results were extended to variational inequalities on unbounded domains by [14] and [26], where processes are assumed to be diffusions or jump diffusions with finite activity jumps. Combining with a probabilistic argument, [14] and [26] confirmed the smooth-fit property when there may be finite activity jumps. In addition, assuming jumps have finite activity, Yang, Jiang and Bian proved in [25] that the value function is the unique classical solution of a variational inequality. Following [14] and [21], Pham studied in [20] the free boundary problem associated to the variational inequality. In [1], Bayraktar also investigated the free boundary problem with alternative techniques. In [20], [25] and [1], the smooth-fit property was proved when the jump has finite activity.

In this paper, we study the optimal stopping problem (1.2) which allows infinite activity jumps. Using the regularity theory for parabolic differential equations, we proved that the value function is the unique solution of a variational inequality, on a unbounded domain, in a certain Sobolev space. The smooth-fit property follows directly from our regularity results. Moreover, based on these regularity result, we further show that the value function is smooth inside the continuation region, under a mild assumption on the Lévy measure.

When the jump has infinite activity, the Lévy measure ν has a singularity. This singularity introduces difficulties in the analysis of the value function regularity. When ν does not have such a singularity (the jump is of finite activity), after applying the non-local integral operator, which appears in the infinitesimal operator of X, to the value function, the resulting function is expected to have the same regularity with the value function (see [25]). However, when ν has a singularity, the regularity of the resulting function is reduced compared to the regularity of the value function. This reduction in the regularity gives trouble in defining the resulting function, after applying the integral operator to the value function, in the classical sense. When the jump has finite variation, this resulting function is still well defined in the classical sense, thanks to the a priori regularity of the value function coming from the probabilistic argument in [21]. However, when the jump has infinite variation, the a priori regularity no longer ensures that the resulting function is well defined. We overcome this problem using a fixed point theorem and the verification theorem in [18]. On the other hand, the unbounded jumps also introduce difficulty in estimating the local regularity of the value function. Because of the unbounded jumps, regularity of the value function inside a bounded domain depends on the value function outside this domain (see Lemmas 4.1 and B-1 for more precise explanation). We solve this difficulty via an interior estimate technique in Theorem 5.1.

The rest of the paper is organized as follows. In Section 2, we introduce the variational inequality and recall two notions of generalized solutions studied in [21] and [18]. In Section 3 we discuss the finite variation jump case and analyze the regularity of value function in the continuation region. Section 4 is devoted to study the global regularity when jumps may have infinite variation. The the global regularity (Theorem 4.1) is proved in

Section 5. A key estimate, which is needed to prove Theorem 4.1 is showed in Appendix B. As a corollary of this global regularity result, the smooth-fit property is confirmed. Moreover, based on Theorem 4.1, Theorem 4.2 shows that the value function is $C^{2,1}$ in the continuation region. At last, proofs of several auxiliary lemmas are listed in Appendix A.

2. The optimal stopping problem and the variational inequality

2.1. **A priori regularity of the value function.** Let us first define the pure jump component $\mathcal{J}$ in (1.1). According to the Lévy-Itô decomposition (see e.g. Theorem 19.2 in [22]), $\mathcal{J}$ can be decomposed as

$$\mathcal{J}_t = \mathcal{J}_t^\ell + \lim_{\epsilon \downarrow 0} \mathcal{J}_t^\epsilon, \tag{2.1}$$

in which

$$\mathcal{J}_t^\ell = \int_0^t \int_{|y|>1} y\, \mu(ds, dy), \quad \mathcal{J}_t^\epsilon = \int_0^t \int_{\epsilon \le |y| \le 1} y\, \widetilde{\mu}(ds, dy), \tag{2.2}$$

represent large and small jumps respectively. Here μ is a Poisson random measure on $\mathbb{R}_+ \times (\mathbb{R}^n \setminus \{0\})$. Its mean measure is the Lévy measure ν, which is a positive Radon measure on $\mathbb{R}^n \setminus \{0\}$ with a possible singularity at 0. Even with this possible singularity at 0, the measure ν still satisfies

$$\int_{\mathbb{R}^n} (|y|^2 \wedge 1)\, \nu(dy) < +\infty. \tag{2.3}$$

Here, the norm $|\cdot|$ is the standard Euclidean norm: $|y| \triangleq \left(\sum_{i=1}^n (y^i)^2\right)^{1/2}$. In (2.2), $\tilde{\mu}(ds, dy) = \mu(ds, dy) - ds\, \nu(dy)$ is the compensated Poisson measure. It is also worth noticing that the convergence in the last term of (2.1) is the almost sure convergence. Moreover, the convergence is uniform in t on $[0, T]$.

We assume that the drift and the volatility in (1.1) are bounded and Lipschitz continuous, i.e., there exists a positive constant $L_{b,\sigma}$ such that

$$\begin{aligned} &|b(x,t) - b(y,t)| + |\sigma(x,t) - \sigma(y,t)| \le L_{b,\sigma} |x-y|, \quad \forall x, y \in \mathbb{R}^n, \\ &\text{moreover, } |b(x,t)| \text{ and } |\sigma(x,t)| \text{ are bounded on } \mathbb{R}^n \times [0,T]. \end{aligned} \tag{H1}$$

We name the solution of (1.1), with the initial condition $X_0 = x$, as X^x. Thanks to (H1), X^x has the following norm estimates.

Lemma 2.1. *Let us assume b and σ satisfy (H1). Then there exists a positive constant C such that for any $\tau \in \mathcal{S}_{0,t}$ with $t \le T$ and $x, y \in \mathbb{R}^n$,*

$$\mathbb{E}\, |X_\tau^x - X_\tau^y| \le C\, |x-y|\,. \tag{2.4}$$

Moreover, if the Lévy measure satisfies

$$\int_{|y|>1} |y|\, \nu(dy) < +\infty, \tag{H2}$$

then we have

$$\mathbb{E}\, |X_\tau^x| \le C, \tag{2.5}$$

$$\mathbb{E}\, |X_\tau^x - x| \le C\, t^{1/2}, \tag{2.6}$$

$$\mathbb{E}\left[\sup\nolimits_{0 \le s \le t} |X_s^x - x|\right] \le C\, t^{1/2}. \tag{2.7}$$

Remark 2.1. *Similar estimates were given in Lemma 3.1 of [21] under a slightly stronger assumption on the large jumps:* $\int_{|y|>1} |y|^2 \nu(dy) < +\infty$. *Using the equivalence between the norm* $|y|$ *and the norm* $\sum_{i=1}^n |y^i|$, *one could prove Lemma 2.1 under assumption (H2). We give its proof in Appendix A.*

For the optimal stopping problem (1.2), let us assume the terminal reward $g : \mathbb{R}^n \to \mathbb{R}$ to be a bounded and Lipschitz continuous function, i.e., there exist positive constants K and L such that

$$0 \leq g(x) \leq K \quad \text{and} \tag{H3}$$

$$|g(x) - g(y)| \leq L|x - y|, \quad \forall x, y \in \mathbb{R}^n. \tag{H4}$$

Thanks to (H3), the value function u is uniformly bounded by K. Moreover, the Lipschitz continuity of g in (H4) and norm estimates of X in Lemma 2.1 ensure that the value function u has the following regularity properties, which follow from the same proof of Proposition 3.3 in [21] once its Lemma 3.1 is replaced by our Lemma 2.1.

Lemma 2.2. *Let us assume that* g *satisfies (H3) and (H4). Then there exists a constant* $L_x > 0$ *such that for any* $x_1, x_2 \in \mathbb{R}$, $t \in [0, T]$,

$$|u(x_1, t) - u(x_2, t)| \leq L_x |x_1 - x_2|. \tag{2.8}$$

Moreover, if the Lévy measure satisfies (H2), then there exists a constant $L_t > 0$ *such that for any* $t_1, t_2 \in [0, T]$, $x \in \mathbb{R}$,

$$|u(x, t_1) - u(x, t_2)| \leq L_t \, |t_1 - t_2|^{1/2}. \tag{2.9}$$

The Lipschitz continuity of $u(\cdot, t)$ and semi-Hölder continuity of $u(x, \cdot)$ will be useful to show further regularity properties of u in the next three sections.

For the optimal stopping problem, as usual we define the continuation region $\mathcal{C}$ and the stopping region $\mathcal{D}$ as follows:

$$\mathcal{C} \triangleq \{(x, t) \in \mathbb{R}^n \times [0, T) : u(x, t) > g(x)\} \quad \text{and} \quad \mathcal{D} \triangleq \{(x, t) \in \mathbb{R}^n \times [0, T) : u(x, t) = g(x)\}.$$

2.2. **The variational inequality.** Intuitively, one can expect from the Itô's Lemma for Lévy processes (see e.g. Proposition 8.18 in [5] pp. 279) that the value function u, defined in (1.2), satisfies the following variational inequality:

$$\begin{aligned} &\min \{(-\partial_t - \mathcal{L} + r) \, u(x, t), u(x, t) - g(x)\} = 0, \quad (x, t) \in \mathbb{R}^n \times [0, T), \\ &u(x, T) = g(x), \end{aligned} \tag{2.10}$$

in which the integro-differential operator $\mathcal{L}$, the infinitesimal generator of X, is defined via a bounded test function ϕ as

$$\mathcal{L}\phi(x, t) \triangleq \mathcal{L}_D \phi(x, t) + I\phi(x, t), \quad \text{with} \quad \mathcal{L}_D \phi(x, t) \triangleq \sum_{i,j=1}^n a_{ij}(x, t) \frac{\partial^2 \phi}{\partial x^i \partial x^j} + \sum_{i=1}^n b_i(x, t) \frac{\partial \phi}{\partial x^i}. \tag{2.11}$$

Here $A = (a_{ij})_{n \times n} \triangleq \frac{1}{2} \sigma(x, t) \sigma(x, t)^T$ is a $n \times n$ matrix and the integral term

$$\begin{aligned} I\phi(x, t) &\triangleq \int_{\mathbb{R}^n} \left[\phi(x + y, t) - \phi(x, t) - \sum_{i=1}^n y^i \frac{\partial \phi}{\partial x^i}(x, t) \, 1_{\{|y| \leq 1\}} \right] \nu(dy) \\ &= \int_{\mathbb{R}^n} \left[\phi(x + y, t) - \phi(x, t) - y \cdot \nabla_x \phi(x, t) \, 1_{\{|y| \leq 1\}} \right] \nu(dy). \end{aligned} \tag{2.12}$$

However, one does not know a priori that the value function u is sufficiently regular (i.e., $u \in C^{2,1}(\mathbb{R}^n \times [0,T))$) to justify applying Itô's Lemma. Moreover, the integral term $I\phi(x,t)$ is only well defined in classical sense when ϕ has certain regularity properties. It is sufficient to require that $\phi(\cdot,t) \in C^1(B_\epsilon(x))$, in which $B_\epsilon(x)$ is an open ball in $\mathbb{R}^n$ centered at x with some radius $\epsilon \in (0,1)$, and that $\nabla_x\phi(\cdot,t)$ to be Lipschitz in $B_\epsilon(x)$ uniformly in t, i.e., for $t \in [0,T)$ there exists a positive constant L_B such that

$$|\nabla_x\phi(x_1,t) - \nabla_x\phi(x_2,t)| \leq L_B|x_1 - x_2|, \qquad \text{for } x_1, x_2 \in B_\epsilon(x). \tag{2.13}$$

Indeed, using these regularity properties of ϕ we have that

$$I\phi(x,t) = I_\epsilon\phi(x,t) + I^\epsilon\phi(x,t), \quad \text{where} \tag{2.14}$$

$$I^\epsilon\phi(x,t) = \int_{|y|>\epsilon} [\phi(x+y,t) - \phi(x,t)]\,\nu(dy) - \nabla_x\phi(x,t)\cdot\int_{\epsilon<|y|\leq 1} y\,\nu(dy) \tag{2.15}$$

$$\begin{aligned} I_\epsilon\phi(x,t) &= \int_{|y|\leq\epsilon} [\phi(x+y,t) - \phi(x,t) - y\cdot\nabla_x\phi(x,t)]\,\nu(dy) \\ &= \int_{|y|\leq\epsilon} \sum_{i=1}^n y^i\left(\partial_{x^i}\phi(z_i,t) - \partial_{x^i}\phi(x,t)\right)\nu(dy) \leq \int_{|y|\leq\epsilon} L_B\,|y|^2\nu(dy). \end{aligned} \tag{2.16}$$

In (2.16), z_i are some vectors in $\mathbb{R}^n$ with $|z_i - x| < |y|$ and the second equality follows from the mean value theorem, while the inequality follows from the Cauchy-Schwartz inequality and (2.13). Note that $\epsilon\int_{\epsilon<|y|\leq 1}\nu(dy) \leq \int_{\epsilon<|y|\leq 1}|y|\,\nu(dy) < \int_{\epsilon<|y|\leq 1}\nu(dy)$ and $\int_{\epsilon<|y|\leq 1}\nu(dy) \leq \frac{1}{\epsilon^2}\int_{\epsilon<|y|\leq 1}|y|^2\,\nu(dy) < +\infty$ from (2.3). These inequalities imply that $\int_{\epsilon<|y|\leq 1}|y|\,\nu(dy) < +\infty$. Hence, we have $I\phi(x,t) < +\infty$.

However, given the regularity of u in Lemma 2.2, it is not clear that the value function u has the Lipschitz continuous first derivative to ensure that Iu is well defined in the classical sense in the first place. Yet, the value function u is a solution of (2.10) in certain weak senses. In the literature different notions of generalized solutions were explored. For example, Pham analyzed the value function of an optimal stopping problem of controlled jump diffusion processes in [21] and proved that the value function is a unique viscosity solution of a nonlinear variational inequality. In what follows we will introduce the notions that we will need from [21]. Let us define

$$C_1(\mathbb{R}^n \times [0,T]) \triangleq \left\{\phi \in C^0(\mathbb{R}^n \times [0,T]) : \sup_{(x,t)\in\mathbb{R}^n\times[0,T]} \frac{|\phi(x,t)|}{1+|x|} < +\infty\right\}.$$

We adapt the notion of viscosity solutions used in Definition 2.1 of [21] into our context and give the following definition.

Definition 2.1. *(i) Any $u \in C^0(\mathbb{R}^n \times [0,T])$ is a viscosity supersolution (subsolution) of (2.10) if*

$$\min\{-\partial_t\phi - \mathcal{L}\,\phi + ru, u(x,t) - g(x)\} \geq 0\,(\leq 0), \tag{2.17}$$

for any function $\phi \in C^{2,1}(\mathbb{R}^n \times [0,T]) \cap C_1(\mathbb{R}^n \times [0,T])$ such that $u(x,t) = \phi(x,t)$ and $u(\tilde{x},\tilde{t}) \geq \phi(\tilde{x},\tilde{t})$ ($u(\tilde{x},\tilde{t}) \leq \phi(\tilde{x},\tilde{t})$) for all $(\tilde{x},\tilde{t}) \in \mathbb{R}^n \times [0,T)$.

(ii) u is a viscosity solution of (2.10) if it is both supersolution and subsolution.

Applying the result of [21] to our setting, we obtain the following result.

Proposition 2.1. *If the Lévy measure ν satisfies (H2), the value function $u(x,t)$ is a viscosity solution of (2.10).*

Proof. Let us first comment that under the assumption (H2), $I\phi(x,t)$ is well defined for $\phi \in C^{2,1}(\mathbb{R}^n \times [0,T]) \cap C_1(\mathbb{R}^n \times [0,T])$. Indeed, for $\phi \in C_1(\mathbb{R}^n \times [0,T])$, we have $|\phi(x+y,t) - \phi(x,t)| \leq C(1+|y|)$ for some C independent of y. Therefore, in (2.15) $\int_{|y|>\epsilon}[\phi(x+y,t) - \phi(x,t)]\,\nu(dy) < +\infty$ as a result of (H2) and the analysis after (2.16).

After replacing Lemma 3.1 of [21] by Lemma 2.1, the statement follows from the same proof of Theorem 3.1 in [21]. □

Remark 2.2. *As a corollary of Theorem 4.1 in [21], u is also the unique viscosity solution in the sense of Definition 2.1. However, this uniqueness result is not necessary for the later development.*

Another notion of generalized solution was studied in [18]. Lamberton and Mikou showed that u is the unique solution of (2.10) in the distribution sense. We will summarize the results of [18] that will be used in the sequel. Let Ω be an open subset of $\mathbb{R}^n \times (0,T)$, and let us denote by $\mathcal{S}(\Omega)$ the set of all C^∞ functions with the compact support in Ω, and by $\mathcal{S}'(\Omega)$ the space of distributions. If $v \in S'(\Omega)$, and it is locally integrable, then the action of the distribution v on the test function ϕ is given by

$$\langle v, \phi \rangle = \int_\Omega v(x,t)\phi(x,t)\, dxdt.$$

Therefore, since the value function u is uniformly bounded, even though it is not clear that u has enough regularity to define $Iu(x,t)$ in classical sense, $Iu(x,t)$ can still be defined as a distribution,

$$\langle Iu, \phi \rangle \triangleq \int_{\mathbb{R}^n \times (0,T)} u(x,t)\, I^*\phi(x,t)\, dxdt, \qquad \text{for } \phi \in \mathcal{S}(\Omega), \tag{2.18}$$

in which the adjoint operator I^* is defined as

$$I^*\phi(x,t) = \int_{\mathbb{R}^n} \left[\phi(x-y,t) - \phi(x,t) + y \cdot \nabla_x \phi(x,t) 1_{\{|y|\leq 1\}}\right] \nu(dy). \tag{2.19}$$

Note that since ϕ is infinitely differentiable with compact support, $I^*\phi$ is well defined in the classical sense thanks to the analysis in (2.15) and (2.16).

Using the theory of the Snell envelope, Lamberton and Mikou proved the following result in Theorem 2.8 of [18].

Proposition 2.2. *The value function $u(x,t)$ is the only continuous and bounded function on $[0,T] \times \mathbb{R}^n$ that satisfies the following conditions:*

(i) $u(x,T) = g(x)$,
(ii) $u \geq g$,
(iii) *the distribution $(\partial_t + \mathcal{L} - r)\, u$ is a nonpositive measure on $\mathbb{R}^n \times (0,T)$, i.e., $(\partial_t + \mathcal{L} - r)\, u \leq 0$ in the distribution sense,*
(iv) *on the open set $\{(x,t) \in \mathbb{R}^n \times (0,T) \,:\, u(x,t) > g(x)\}$, $(\partial_t + \mathcal{L} - r)\, u = 0$.*

Remark 2.3. *In Proposition 2.2, the inequality (equality) $(\partial_t + \mathcal{L} - r)\, u \leq 0\, (= 0)$ is understood in the distribution sense, i.e., for any open set $\Omega \subset \mathbb{R}^n \times (0,T)$ and any nonnegative function $\phi(x,t) \in \mathcal{S}(\Omega)$,*

$$\int_\Omega (\partial_t + \mathcal{L} - r)\, u(x,t)\, \phi(x,t)\, dxdt = \int_\Omega u(x,t)\, (-\partial_t + \mathcal{L}^* - r)\, \phi(x,t)\, dxdt \leq 0\, (= 0), \tag{2.20}$$

where the adjoint operator $\mathcal{L}^*$ is defined as the adjoint operator of the differential part of $\mathcal{L}$ plus the operator I^* in (2.19), i.e.,

$$\mathcal{L}^*\phi(x,t) \triangleq \sum_{i,j=1}^n \frac{\partial^2}{\partial x^i \partial x^j}\, (a_{ij}\phi) - \sum_{i=1}^n \frac{\partial}{\partial x^i}\, (b_i \phi) + I^*\phi(x,t).$$

2.3. **The classical differentiability.** We will apply the regularity theory of parabolic differential equations to analyze the classical differentiability of u in the next three sections. We need foremost make sure that Iu is defined in the classical sense. Throughout this paper, we assume that the Lévy measure ν has a density, which we will denote by $\rho(y)$. Moreover, There exists a positive constants M such that

$$\rho(y) \leq \frac{M}{|y|^{n+\alpha}}, \quad \text{for } |y| \leq 1 \text{ and some constant } \alpha \in [0,2). \tag{H5}$$

Remark 2.4. *The Lévy measures ν, corresponding to Lévy processes widely used in the financial modelling for the single asset case, satisfy (H5) with $n = 1$. In jump diffusions models where ν is a probability measure, if the density $\rho(y)$ is bounded, (H5) is satisfied with sufficiently large M. Examples of this case are Merton's model and Kou's model. On the other hand, if $\rho(y) \in C^0(B_1(0) \setminus \{0\})$ and $\rho(y) \to C/|y|^\beta$ with $0 < \beta < 1$ as $y \to 0$, (H5) is also fulfilled because $\frac{1}{|y|^{1+\alpha}} > \frac{1}{|y|^\beta}$ for any $\alpha \geq 0$ and $|y| \leq 1$.*

Moreover, for Lévy processes that are the Brownian motion subordinated by tempered stable subordinators, it follows from (4.25) in [5] that $\rho(y) \to C/|y|^{1+2\beta}$, with $0 \leq \beta < 1$, as $y \to 0$. Therefore (H5) is satisfied by choosing $\alpha = 2\beta$ and sufficiently large M. In particular, this class of Lévy processes contains Variance Gamma and Normal Inverse Gaussian where $\beta = 0$ or $1/2$ respectively.

Furthermore, for the generalized tempered stable processes (see Remark 4.1 in [5]) whose Lévy measure is

$$\rho(y) = \frac{C_-}{|y|^{1+\alpha_-}} e^{-\lambda_-|x|} 1_{\{x<0\}} + \frac{C_+}{|y|^{1+\alpha_+}} e^{-\lambda_+ x} 1_{\{x>0\}},$$

with $\alpha_-, \alpha_+ < 2$, (H5) is satisfied by choosing $\alpha = \max\{\alpha_-, \alpha_+, 0\}$ and $M = \max\{C_-, C_+\}$. In particular, CGMY processes in [4] are special examples of generalized tempered stable processes. In the similar manner, one can also check that the regular Lévy processes of exponential type (RLPE) in [3] also satisfy (H5).

In order to apply the regularity theory of parabolic differential equations to analyze the regularity of u, let us recall the definition of Sobolev spaces and Hölder spaces on pp. 5 and 7 of [17].

Definition 2.2. *Let Ω be a domain in $\mathbb{R}^n$, $Q_T = \Omega \times (0,T)$ and $\overline{Q_T}$ be the closure of Q_T. $C^{2,1}(Q_T)$ denotes the class of continuous functions on Q_T with continuous classical derivatives on Q_T of the form $\partial_t v$, $\partial_{x^i} v$ and $\partial^2_{x^i x^j} v$ for $i, j \leq n$.*

For any positive integer $p \geq 1$, $W^{2,1}_p(Q_T)$ is the Banach space consisting of the elements of $L_p(Q_T)$ having generalized derivatives of the form $\partial_t v$, $\partial_{x^i} v$ and $\partial^2_{x^i x^j} v$ for $i, j \leq n$. The norm in it is defined as

$$\|v\|_{W^{2,1}_p(Q_T)} = \|\partial_t v\|_{L_p} + \sum_{i=1}^n \|\partial_{x^i} v\|_{L_p} + \sum_{i,j=1}^n \|\partial^2_{x^i x^j} v\|_{L_p},$$

where $\|v\|_{L_p} = \left(\int_0^T \int_\Omega |v(x,t)|^p \, dxdt\right)^{1/p}$. On the other hand, $W^{2,1}_{p,\,loc}(Q_T)$ is the Banach space consisting of functions whose $W^{2,1}_p$-norm is finite on any compact subset of Q_T.

For any positive nonintegral real number α, $H^{\alpha,\alpha/2}\left(\overline{Q_T}\right)$ is the Banach space of functions v that are continuous in $\overline{Q_T}$, together with continuous classical derivatives of the form $\partial^r_t \partial^s_x v$ for $2r + s < \alpha$, and have a finite norm

$$\|v\|^{(\alpha)}_{Q_T} = |v|^{(\alpha)}_x + |v|^{(\alpha/2)}_t + \sum_{2r+s \leq [\alpha]} \|\partial^r_t \partial^s_x v\|^{(0)}, \quad \text{in which}$$

$$
\begin{aligned}
\|v\|^{(0)} &= max_{Q_T}|v|, \quad \partial_x^s v = \partial_{x^{i_1}}^{j_1}\cdots\partial_{x^{i_k}}^{j_k} v, \ with\ j_1+\cdots+j_k = s,\\
|v|_x^{(\alpha)} &= \sum_{2r+s=[\alpha]} <\partial_t^r\partial_x^s v>_x^{(\alpha-[\alpha])}, \quad |v|_t^{(\alpha/2)} = \sum_{\alpha-2<2r+s<\alpha} <\partial_t^r\partial_x^s v>_t^{(\frac{\alpha-2r-s}{2})};\\
<v>_x^{(\beta)} &= \sup_{\substack{(x,t),(x',t)\in\overline{Q_T}\\ |x-x'|\le\rho_0}} \frac{|v(x,t)-v(x',t)|}{|x-x'|^\beta}, \quad 0<\beta<1,\\
<v>_t^{(\beta)} &= \sup_{\substack{(x,t),(x,t')\in\overline{Q_T}\\ |t-t'|\le\rho_0}} \frac{|v(x,t)-v(x,t')|}{|t-t'|^\beta}, \quad 0<\beta<1,
\end{aligned}
$$

where ρ_0 is a positive constant.

On the other hand, $H^\alpha\left(\overline{\Omega}\right)$ is the Banach space whose elements are continuous functions $v(x)$ on $\overline{\Omega}$ that have continuous derivatives up to order $[\alpha]$ and the following norm finite

$$
\|v\|_{\overline{\Omega}}^{(\alpha)} = \sum_{s\le[\alpha]}\|\partial_x^s v\|^{(0)} + \left|\partial_x^{[\alpha]}v\right|^{(\alpha-[\alpha])}, \qquad in\ which \quad |v|^{(\beta)} = \sup_{x,x'\in\overline{\Omega},|x-x'|\le\rho_0}\frac{|v(x)-v(x')|}{|x-x'|^\beta}.
$$

These Hölder norms depend on ρ_0, but for different $\rho_0>0$, the corresponding Hölder norms are equivalent. Hence their dependence on ρ_0 will not be noted in the sequel.

3. Finite variation jumps and regularity in the continuation region

In this section, based on Pham's result in Proposition 2.1, we will analyze the regularity of the value function u when the jump of X has finite variation, i.e.,

$$
\int_{\mathbb{R}^n}|y|\wedge 1\,\nu(dy) < +\infty. \tag{3.1}
$$

It is worth noticing that $\int_{|y|\le 1}|y|\,\nu(dy)<+\infty$ is satisfied when we assume (H5) with $0\le\alpha<1$. As a result, the infinitesimal generator $\mathcal{L}$ can be rewritten as

$$
\begin{aligned}
\mathcal{L}\phi(x,t) &= \mathcal{L}_D^f\phi(x,t) + I^f\phi(x,t), \quad \text{where} \qquad (3.2)\\
\mathcal{L}_D^f\phi(x,t) &= \sum_{i,j=1}^n a_{ij}(x,t)\frac{\partial^2\phi}{\partial x^i\partial x^j} + \sum_{i=1}^n\left[b_i(x,t) - \int_{|y|\le 1} y^i\nu(dy)\right]\frac{\partial\phi}{\partial x^i} \qquad (3.3)\\
I^f\phi(t,x) &\triangleq \int_{\mathbb{R}^n}\left[\phi(x+y,t)-\phi(x,t)\right]\nu(dy). \qquad (3.4)
\end{aligned}
$$

Thanks to this reduced integral form and the Lipschitz continuity of $u(\cdot,t)$ (see Lemma 2.2), $I^f u(x,t)$ is well defined in the class sense. Indeed

$$
I^f u(x,t) \le \int_{\mathbb{R}}|u(x+y,t)-u(x,t)|\,\nu(dy) \le L_x\int_{\mathbb{R}}|y|\,\nu(dy) < +\infty,
$$

as a result of (3.1) and (H2). Moreover, assuming (H5) with $0\le\alpha<1$, we will show that $I^f u(x,t)$ is Hölder continuous in both variables in the following lemma.

Lemma 3.1. *Let Ω be any compact domain in $\mathbb{R}^n$. If the density $\rho(y)$ of the measure ν satisfies (H5) with $0\le\alpha<1$, then $I^f u(x,t)$ is Hölder continuous in both variables on $\Omega\times[0,T]$.*

(i) *For any* $(x_1,t),(x_2,t)\in\Omega\times[0,T]$, *there exist constants* $C_{\Omega,\beta}$ *and* C_Ω *independent of* x_1, x_2 *and* t, *such that*

$$\text{when } \alpha=0: \quad \left|I^f u(x_1,t)-I^f u(x_2,t)\right| \le C_{\Omega,\beta}|x_1-x_2|^{1-\beta}, \quad \text{for any } \beta\in(0,1); \tag{3.5}$$

$$\text{when } 0<\alpha<1: \quad \left|I^f u(x_1,t)-I^f u(x_2,t)\right| \le C_{\Omega}|x_1-x_2|^{1-\alpha}. \tag{3.6}$$

(ii) *For any* $(x,t_1),(x,t_2)\in\Omega\times[0,T]$, *there exist constants* $D_{\Omega,\beta}$ *and* D_Ω *independent of* t_1, t_2 *and* x, *such that*

$$\text{when } \alpha=0: \quad \left|I^f u(x,t_1)-I^f u(x,t_2)\right| \le D_{\Omega,\beta}|t_1-t_2|^{\frac{1-\beta}{2}}, \quad \forall\beta\in(0,1); \tag{3.7}$$

$$\text{when } 0<\alpha<1: \quad \left|I^f u(x,t_1)-I^f u(x,t_2)\right| \le D_{\Omega}|t_1-t_2|^{\frac{1-\alpha}{2}}. \tag{3.8}$$

Proof. This proof is motived by Proposition 2.5 in [23]. We will show the Hölder continuity in x first. Let us break up the integral into two parts:

$$\left|I^f u(x_1,t)-I^f u(x_2,t)\right| \le \int_{\mathbb{R}} |u(x_1+y,t)-u(x_1,t)-u(x_2+y,t)+u(x_2,t)|\,\nu(dy) \le I_1+I_2, \quad \text{in which} \tag{3.9}$$

$$I_1=\int_{|y|\le\epsilon}\left[\,|u(x_1+y,t)-u(x_1,t)|+|u(x_2+y,t)-u(x_2,t)|\,\right]\nu(dy), \tag{3.10}$$

$$I_2=\int_{|y|>\epsilon}\left[\,|u(x_1+y,t)-u(x_2+y,t)|+|u(x_1,t)-u(x_2,t)|\,\right]\nu(dy). \tag{3.11}$$

Here the constant $\epsilon\in(0,1]$ will be determined later. Since $x\to u(t,x)$ is globally Lipschitz (see Lemma 2.2), we have for $i=1,2$

$$|u(x_i+y,t)-u(x_i,t)|\le L_x|y|, \quad |u(x_1+y,t)-u(x_2+y,t)|\le L_x|x_1-x_2| \quad\text{and}\quad |u(x_1,t)-u(x_2,t)|\le L_x|x_1-x_2|.$$

Combining these inequalities with (H5), in which $0\le\alpha<1$, we obtain from (3.10) and (3.11) that

$$I_1 \le \int_{|y|\le\epsilon} 2L_x|y|\,\nu(dy) \le 2L_xM\int_{|y|\le\epsilon}|y|^{1-n-\alpha}dy = 2L_xM|S_1(0)|\int_0^\epsilon r^{-\alpha}dr=\frac{2L_xM|S_1(0)|}{1-\alpha}\epsilon^{1-\alpha}, \tag{3.12}$$

$$\begin{aligned} I_2 &\le \int_{|y|>\epsilon} 2L_x|x_1-x_2|\,\nu(dy)\le 2L_x|x_1-x_2|\int_{|y|>1}\nu(dy)+2L_xM|x_1-x_2|\int_{\epsilon<|y|\le1}|y|^{-n-\alpha}dy \\ &= 2L_x|x_1-x_2|\int_{|y|>1}\nu(dy)+2L_xM|S_1(0)|\,|x_1-x_2|\cdot\begin{cases}\frac{\epsilon^{-\alpha}-1}{\alpha} & \text{if } 0<\alpha<1\\ -\log\epsilon & \text{if } \alpha=0,\end{cases}\end{aligned} \tag{3.13}$$

where $|S_1(0)|$ is the surface area of a unit ball in $\mathbb{R}^n$. Now picking $\epsilon=|x_1-x_2|\wedge 1$ and noticing that $0\le\alpha<1$, we have

$$\epsilon^{1-\alpha}\le|x_1-x_2|^{1-\alpha}, \quad \epsilon^{-\alpha}-1\le|x_1-x_2|^{-\alpha}. \tag{3.14}$$

Moreover, when $\epsilon=|x_1-x_2|<1$,

$$-\log\epsilon=\int_{|x_1-x_2|}^1\frac{1}{z}\,dz\le\int_{|x_1-x_2|}^1\frac{1}{z^{1+\beta}}\,dz=\frac{1}{\beta}\left(|x_1-x_2|^{-\beta}-1\right)\le\frac{1}{\beta}|x_1-x_2|^{-\beta} \quad \forall\beta>0. \tag{3.15}$$

Hence choosing $\epsilon=|x_1-x_2|\wedge1$, we have $-\log\epsilon\le\frac{1}{\beta}|x_1-x_2|^{-\beta}$ for any $\beta>0$. Combining (3.9) and (3.12) - (3.15), we conclude that

$$\text{when } 0<\alpha<1: \quad \left|I^f u(x_1,t)-I^f u(x_2,t)\right|\le\left[\frac{2L_xM|S_1(0)|}{\alpha(1-\alpha)}+2L_xd^\alpha\int_{|y|>1}\nu(dy)\right]|x_1-x_2|^{1-\alpha},$$

$$\text{when } \alpha=0: \quad \left|I^f u(x_1,t)-I^f u(x_2,t)\right|\le\left[2L_xM|S_1(0)|\,d^\beta+\frac{2L_xM|S_1(0)|}{\beta}+2L_xd^\beta\int_{|y|>1}\nu(dy)\right]|x_1-x_2|^{1-\beta},$$

in which $\beta\in(0,1)$ and $d=\max_{x,y\in\Omega}|x-y|$.

Similarly, in order to show the Hölder continuity in t, we also break up the integral term into two parts:

$$\left|I^f u(x,t_1) - I^f u(x,t_2)\right| \leq \int_{\mathbb{R}} |u(x+y,t_1) - u(x,t_1) - u(x+y,t_2) + u(x,t_2)|\, \nu(dy) \leq I_1 + I_2, \text{ in which} \tag{3.16}$$

$$I_1 = \int_{|y|\leq\epsilon} [\, |u(x+y,t_1) - u(x,t_1)| + |u(x+y,t_2) - u(x,t_2)|\,]\, \nu(dy), \tag{3.17}$$

$$I_2 = \int_{|y|>\epsilon} [\, |u(x+y,t_1) - u(x+y,t_2)| + |u(x,t_1) - u(x,t_2)|\,]\, \nu(dy). \tag{3.18}$$

The constant $\epsilon \in (0,1]$ will be determined later. We can first bound I_1 in (3.17) using (3.12). Then it follows from the semi-Hölder continuity of $t \to u(t,x)$ (see Lemma 2.2) that

$$\begin{aligned} I_2 &\leq \int_{|y|>\epsilon} 2\, L_t |t_1-t_2|^{\frac{1}{2}}\, \nu(dy) = 2\, L_t |t_1-t_2|^{\frac{1}{2}} \int_{\epsilon<|y|\leq 1} \nu(dy) + 2\, L_t |t_1-t_2|^{\frac{1}{2}} \int_{|y|>1} \nu(dy) \\ &\leq 2\, L_t |t_1-t_2|^{\frac{1}{2}} \int_{|y|>1} \nu(dy) + 2\, L_t M |S_1(0)|\, |t_1-t_2|^{\frac{1}{2}} \cdot \begin{cases} \frac{\epsilon^{-\alpha}-1}{\alpha}, & \text{if } 0<\alpha<1 \\ -\log\epsilon, & \text{if } \alpha = 0, \end{cases} \end{aligned} \tag{3.19}$$

in which the second inequality follows from (H5) with $0 \leq \alpha < 1$.

Now picking $\epsilon = |t_1-t_2|^{\frac{1}{2}} \wedge 1$, we have $\epsilon^{1-\alpha} \leq |t_1-t_2|^{\frac{1-\alpha}{2}}$ and $\epsilon^{-\alpha} - 1 \leq |t_1-t_2|^{-\frac{\alpha}{2}}$. A calculation in (3.15) gives us that $-\log\epsilon \leq 2|t_1-t_2|^{-\beta/2}/\beta$ for any $\beta > 0$. Therefore (3.7) and (3.8) follow from combining (3.16), (3.12) and (3.19). □

Having shown that the integral term $I^f u$ is well defined in classical sense and is Hölder continuous on compact domains, we will study the variational inequality (2.10) on a given compact domain inside the continuation region $\mathcal{C}$. Let B be an open ball in $\mathbb{R}^n$ with its closure $\overline{B}$ and $B \times (t_1,t_2) \subset \mathcal{C}$ for some $t_1, t_2 \in [0,T)$. Let us consider the following boundary value problem:

$$\begin{aligned} (-\partial_t - \mathcal{L} + r)\; v(x,t) = 0, &\quad (x,t) \in B \times [t_1,t_2), \\ v(x,t) = u(x,t), &\quad (x,t) \in \mathbb{R}^n \times [t_1,t_2] \setminus B \times [t_1,t_2). \end{aligned} \tag{3.20}$$

Due to Lemma 2.2, the boundary and terminal value u is continuous in $\overline{B} \times [t_1,t_2]$. The viscosity solution of this boundary value problem can be defined as follows (see e.g. Definition 7.4 in [6], Definition 13.1 in [8] or Definition 12.1 in [5]).

Definition 3.1. *(i) Any $v \in C^0(\overline{B} \times [t_1,t_2])$ is a viscosity subsolution of (3.20) if*

$$(-\partial_t - \mathcal{L} + r)\, \phi(x,t) \leq 0, \quad \textit{for } (x,t) \in B \times [t_1,t_2), \tag{3.21}$$

$$\min\{(-\partial_t - \mathcal{L} + r)\, \phi(x,t),\, v(x,t) - u(x,t)\} \leq 0, \quad \textit{for } (x,t) \in \partial B \times [t_1,t_2) \cup \overline{B} \times t_2, \tag{3.22}$$

$$v(x,t) \leq u(x,t), \quad \textit{for } (x,t) \in \mathbb{R}^n \times [t_1,t_2] \setminus \overline{B} \times [t_1,t_2], \tag{3.23}$$

for any function $\phi \in C^{2,1}(\mathbb{R}^n \times [t_1,t_2]) \cap C_1(\mathbb{R}^n \times [t_1,t_2])$ such that $\phi(x,t) = v(x,t)$ and $\phi(\tilde{x},\tilde{t}) \geq v(\tilde{x},\tilde{t})$ for any $(\tilde{x},\tilde{t}) \in \mathbb{R}^n \times [t_1,t_2]$. Any $v \in C^0(\overline{B} \times [t_1,t_2])$ is a viscosity supersolution of (3.20) if

$$(-\partial_t - \mathcal{L} + r)\, \phi(x,t) \geq 0, \quad \textit{for } (x,t) \in B \times [t_1,t_2), \tag{3.24}$$

$$\max\{(-\partial_t - \mathcal{L} + r)\, \phi(x,t),\, v(x,t) - u(x,t)\} \geq 0, \quad \textit{for } (x,t) \in \partial B \times [t_1,t_2) \cup \overline{B} \times t_2, \tag{3.25}$$

$$v(x,t) \geq u(x,t), \quad \textit{for } (x,t) \in \mathbb{R}^n \times [t_1,t_2] \setminus \overline{B} \times [t_1,t_2], \tag{3.26}$$

for any function $\phi \in C^{2,1}(\mathbb{R}^n \times [t_1,t_2]) \cap C_1(\mathbb{R}^n \times [t_1,t_2])$ such that $\phi(x,t) = v(x,t)$ and $\phi(\tilde{x},\tilde{t}) \leq v(\tilde{x},\tilde{t})$ for any $(\tilde{x},\tilde{t}) \in \mathbb{R}^n \times [t_1,t_2]$.

(ii) v is a viscosity solution of (3.20) if it is both a subsolution and a supersolution.

Following Definition 3.1, it is easy to check the following result.

Lemma 3.2. *If the Lévy measure ν satisfies (H2), then $u(x,t)$ is a viscosity solution of (3.20).*

Proof. We will only show that $u(x,t)$ is a viscosity subsolution. That u is a viscosity supersolution can be checked similarly. For any $(x,t) \in \overline{B} \times [t_1, t_2]$, let ϕ be a test function satisfying conditions in Definition 3.1 for subsolutions. Noticing that $u(x,t)$ itself is the boundary and terminal value of (3.20), (3.22) and (3.23) are automatically satisfied. On the other hand, the inequality (3.21) follows from (2.17) and the fact that $u(t,x) \geq g(x)$. □

In Definition 3.1, it is important to note that the test function ϕ is used in evaluating the integral term $I^f \phi(t,x)$. However, thanks to Lemma 3.1, $I^f u$ is well defined in the classical sense. Therefore, we will consider the following parabolic differential equation with an integral driving term

$$
\begin{aligned}
&(-\partial_t - \mathcal{L}_D^f + r)\, v(x,t) = I^f u(x,t), \quad \text{for } (x,t) \in B \times [t_1, t_2), \\
&v(x,t) = u(x,t), \quad \text{for } (x,t) \in \partial B \times [t_1, t_2) \cup \overline{B} \times t_2,
\end{aligned}
\tag{3.27}
$$

where B is the same as in (3.20). The viscosity solution of (3.27) is defined as follows.

Definition 3.2. *Any $v \in C^0(\overline{B} \times [t_1, t_2])$ is a viscosity subsolution of (3.27) if*

$$
(-\partial_t - \mathcal{L}_D^f + r)\, \phi(x,t) \leq I^f u(x,t), \quad \text{for } (x,t) \in B \times [t_1, t_2), \tag{3.28}
$$

$$
\min\Big\{(-\partial_t - \mathcal{L}_D^f + r)\, \phi(t,x) - I^f u(x,t),\, v(x,t) - u(x,t)\Big\} \leq 0, \quad \text{for } (x,t) \in \partial B \times [t_1, t_2)\overline{B} \times t_2 \tag{3.29}
$$

for any function $\phi \in C^{2,1}(\mathbb{R}^n \times [t_1, t_2])$ such that $\phi(x,t) = v(x,t)$ and $\phi(\tilde{x}, \tilde{t}) \geq v(\tilde{x}, \tilde{t})$ for any $(\tilde{x}, \tilde{t}) \in \mathbb{R}^n \times [t_1, t_2]$. The supersolution is defined analogously. As usual, v is a viscosity of (3.27) if it is both a subsolution and a supersolution.

Actually, it turns out the notion of viscosity solutions for (3.20) defined in Definition 3.1 is equivalent to the notion of viscosity solutions for (3.27) defined in Definition 3.2.

Lemma 3.3. *The value function u is a viscosity solution of (3.20) in the sense of Definition 3.1, if and only if u is a viscosity solution of (3.27) in the sense of Definition 3.2.*

Proof. The proof follows from the argument of Lemma 2.1 in [24]. For the completeness of this paper, we will repeat this argument in Appendix A. □

Now we will apply the regularity theory of parabolic differential equation to analyze the regularity of u in the continuation region $\mathcal{C}$. We assume that there exist a positive constant λ such that

$$
\sum_{i,j=1}^{n} a_{ij}(x,t)\, \xi^i \xi^j \geq \lambda |\xi|^2, \quad \forall x, \xi \in \mathbb{R}^n, t \geq 0. \tag{H6}
$$

Additionally, for $i, j \leq n$

(H7) $a_{ij}(x,t), b_i(x,t)$ and $r(x,t)$ are continuously differentiable in both variables on $\mathbb{R}^n \times [0,T]$.

With these two assumptions, now we are ready to state the main theorem of this section.

Theorem 3.1. *Let us assume that the Lévy measure ν satisfies (H2) and (H5) with $0 \leq \alpha < 1$, moreover coefficients of (3.20) satisfy (H6) and (H7). Then the value function u is the unique classical solution, i.e., $u \in C^{2,1}$, of the boundary value problem (3.20). Moreover, $u \in C^{2,1}(\mathcal{C})$.*

Proof. It follows from Lemmas 3.2 and 3.3 that the value function $u(x,t)$ is a viscosity solution of (3.27) in the sense of Definition 3.2. For the boundary value problem (3.27), its boundary and terminal values are continuous on $\partial B \times [t_1, t_2) \cup \overline{B} \times t_2$, as a result of the continuity of u (see Lemma 2.2). On the other hand, the driving term $I^f u(x,t)$ is uniformly Hölder continuous in both variables in $\overline{B} \times [t_1, t_2]$ (see Lemma 3.1). Moreover, thanks to (H7), the coefficients in (3.27) are bounded and Hölder continuous in $\overline{B} \times [t_1, t_2]$. Therefore, combining with the nondegenerate assumption (H6), Theorem 9 in [9] pp. 69 implies that (3.27) has a unique classical solution $u^*(x,t) \in C^{2,1}(B \times (t_1, t_2))$. Since u^* is already a classical solution, u^* is also a viscosity solution of (3.27). Therefore, it follows from the Comparison Theorem for viscosity solutions for parabolic differential equations with the driving term (see e.g. Theorem 7.5 in [6]) that $u(x,t) = u^*(x,t)$ for $(x,t) \in B \times (t_1, t_2)$. This ensures that the value function u is the unique classical solution of (3.20). Since $B \times (t_1, t_2)$ is an arbitrary domain in the continuation region $\mathcal{C}$, we have $u \in C^{2,1}(\mathcal{C})$. □

We have studied the regularity of the value function inside the continuation region when jumps have finite variation. We still want to understand how the value function cross the interface of the continuation region and the stopping region, even when jumps have finite variation. Moreover, we hope to study problems with infinite variation jumps. These analysis depend on the global regularity of the value function, which we shall study in the following section.

4. Infinite variation jumps and the global regularity

4.1. **The integral term.** When the jumps of X have infinite variation, i.e., (3.1) is not satisfied, the integral term cannot be reduced to the form in (3.4). Therefore, throughout this section we need to work with the integro-differential operator $\mathcal{L}$ and its integral part I in the form of (2.11) and (2.12). However, given the regularity properties of the value function u in Lemmas 2.2, it is not clear that u has Lipschitz continuous first derivative to make sure Iu is well defined in the classical sense (see (2.16)). Nevertheless, in the following lemma, we will show that given sufficient regularity properties for the test function ϕ, $I\phi(x,t)$ is Hölder continuous in both variables. Later in this section, we will prove that the value function u does have these regularity properties to guarantee Iu well defined in the classical sense.

Let Ω be a compact domain in $\mathbb{R}^n$, $\Omega^\delta \triangleq \{x \in \mathbb{R}^n : x \in B_\delta(y) \text{ for some } y \in \Omega\}$ for some $\delta > 0$. For $s \in (0, T]$, let us denote $\overline{Q_s} = \overline{\Omega} \times [0,s]$ and $\overline{Q_s^\delta} = \overline{\Omega^\delta} \times [0,s]$. Moreover, we denote $D_s \triangleq \mathbb{R}^n \times [0,s]$.

Lemma 4.1. *Let us assume that the Lévy measure satisfies (H2) and (H5) with $\alpha \in [1,2)$.*

(i) *Let us choose ϕ with finite norms $\max_{\mathbb{R}^n \times [0,s]} |\phi|$ and $\max_{\mathbb{R}^n \times [0,s]} |\nabla_x \phi|$, moreover $|\phi(x,t_1) - \phi(x,t_2)| \leq \widetilde{L}_t |t_1 - t_2|^{1/2}$ for any $x \in \mathbb{R}$ and $t_1, t_2 \in [0,s]$. If $\phi \in H^{\beta, \frac{\beta}{2}}\left(\overline{Q^1_s}\right)$ for some $\beta \in (\alpha, 2)$, then $Iu \in H^{\frac{\beta-\alpha}{2}, \frac{\beta-\alpha}{4}}\left(\overline{Q_s}\right)$. Additionally, there exists a constant $C_\Omega > 0$, depending on Ω, α, β and T, such that*

$$\|I\phi\|^{\left(\frac{\beta-\alpha}{2}\right)}_{Q_s} \leq C_\Omega \left(\max_{\mathbb{R}^n \times [0,s]} |\phi| + \max_{\mathbb{R}^n \times [0,s]} |\nabla_x \phi| + \widetilde{L}_t + \|\phi\|^{(\beta)}_{Q^1_s} \right), \tag{4.1}$$

where the Hölder norm $\|\cdot\|^{(\gamma)}_{Q_s}$ is defined in Definition 2.2.

(ii) *If $\phi \in H^{\beta, \frac{\beta}{2}}(D_s)$ for some $\beta \in (\alpha, 2)$, then $I\phi \in H^{\frac{\beta-\alpha}{2}, \frac{\beta-\alpha}{4}}(D_s)$. Moreover, there exists a constant C, depending on α, β and T, such that*

$$\|I\phi\|^{\left(\frac{\beta-\alpha}{2}\right)}_{D_s} \leq C\, \|\phi\|^{(\beta)}_{D_s}. \tag{4.2}$$

Proof. For the notational simplicity, the constant C denotes a generic constant in different places in the proof.

1. Let us first estimate $\max_{\overline{Q_s}}|I\phi|$. Following (2.12), for $(x,t)\in\overline{Q_s}$, we have

$$
\begin{aligned}
(4.3)\quad |I\phi(x,t)| &\le \int_{|y|\le 1}\left|\phi(x+y,t)-\phi(x,t)-\sum_{i=1}^n y^i\,\partial_{x^i}\phi(x,t)\right|\nu(dy)+\int_{|y|>1}|\phi(x+y,t)-\phi(x,t)|\,\nu(dy)\\
&\le \int_{|y|\le 1}\sum_{i=1}^n\left|y^i\,\partial_{x^i}\phi(z_i,t)-y^i\,\partial_{x^i}\phi(x,t)\right|\nu(dy)+2\max_{\mathbb{R}^n\times[0,s]}|\phi|\int_{|y|>1}\nu(dy)\\
&\le \|\phi\|^{(\beta)}_{Q^1_s}\int_{|y|\le 1}|y|^\beta\nu(dy)+2\max_{\mathbb{R}^n\times[0,s]}|\phi|\int_{|y|>1}\nu(dy)\\
&\le C\left(\max_{\mathbb{R}^n\times[0,s]}|\phi|+\|\phi\|^{(\beta)}_{Q^1_s}\right).
\end{aligned}
$$

In the second inequality of (4.3), z_i are some vectors in $\mathbb{R}^n$ with $|z_i-x|<|y|$. Therefore, when $x\in\Omega$, we have $x+z_i\in\Omega^1$. The third inequality follows from the Hölder continuity of $\partial_{x^i}\phi$ on $\overline{Q^1_s}$, i.e., $\sum_{i=1}^n|\partial_{x^i}\phi(z_i,t)-\partial_{x^i}\phi(x,t)|\le\|\phi\|^{(\beta)}_{Q^1_s}|y|^{\beta-1}$. TO get the last inequality, we applies (H5). Note that $\beta>\alpha$, hence $\int_{|y|\le1}|y|^{-n+\beta-\alpha}dy$ is integrable.

The proof of the Hölder continuity of $x\to I\phi(x,t)$ and $t\to I\phi(x,t)$ are similar to the proof in Lemmas 3.1. Let us check the Hölder continuity in x first. For any $x_1,\,x_2\in\Omega$ and $t\in[0,s]$, breaking up the integral term into three parts, we obtain

$$
\begin{aligned}
(4.4)\quad &|I\phi(x_1,t)-I\phi(x_2,t)|\le I_1+I_2+I_3,\qquad\text{in which}\\
&I_1(x,t)=\int_{|y|\le\epsilon}\left[|\phi(x_1+y,t)-\phi(x_1,t)-y\cdot\nabla_x\phi(x_1,t)|+|\phi(x_2+y,t)-\phi(x_2,t)-y\cdot\nabla_x\phi(x_2,t)|\right]\nu(dy),\\
&I_2(x,t)=\int_{\epsilon<|y|\le1}\left[|\phi(x_1+y,t)-\phi(x_2+y,t)|+|\phi(x_1,t)-\phi(x_2,t)|+|y|\,|\nabla_x\phi(x_1,t)-\nabla_x\phi(x_2,t)|\right]\nu(dy),\\
&I_3(x,t)=\int_{|y|>1}\left[|\phi(x_1+y,t)-\phi(x_2+y,t)|+|\phi(x_1,t)-\phi(x_2,t)|\right]\nu(dy).
\end{aligned}
$$

Here the constant $\epsilon\le1$ will be determined later. Let us estimate each integral term separately. An estimate similar to (4.3) shows that

$$
(4.5)\qquad I_1\le 2\|\phi\|^{(\beta)}_{Q^1_s}\int_{|y|\le\epsilon}|y|^\beta\nu(dy)\le 2M\|\phi\|^{(\beta)}_{Q^1_s}\int_{|y|\le\epsilon}|y|^{-n+\beta-\alpha}dy=C\|\phi\|^{(\beta)}_{Q^1_s}\,\epsilon^{\beta-\alpha}.
$$

Thanks to the Lipschitz continuity of $x\to\phi(x,t)$ and the Hölder continuity of $x\to\partial_{x_i}\phi(x,t)$, we can estimate I_2 and I_3 as

$$
\begin{aligned}
(4.6)\quad I_2 &\le \int_{\epsilon<|y|\le1}\left[2\max_{\mathbb{R}^n\times[0,s]}|\nabla_x\phi|\,|x_1-x_2|+\|\phi\|^{(\beta)}_{Q^1_s}\,|y|\,|x_1-x_2|^{\beta-1}\right]\nu(dy)\\
&\le M\int_{\epsilon<|y|\le1}\left[2\max_{\mathbb{R}^n\times[0,s]}|\nabla_x\phi|\,|x_1-x_2|+\|\phi\|^{(\beta)}_{Q^1_s}\,|y|\,|x_1-x_2|^{\beta-1}\right]|y|^{-n-\alpha}dy\\
&= C\max_{\mathbb{R}^n\times[0,s]}|\nabla_x\phi|\,|x_1-x_2|(\epsilon^{-\alpha}-1)+C\,\|\phi\|^{(\beta)}_{Q^1_s}|x_1-x_2|^{\beta-1}\cdot\begin{cases}\epsilon^{1-\alpha}-1 & \text{when } 1<\alpha<2,\\ -\log\epsilon & \text{when } \alpha=1.\end{cases},\\
(4.7)\quad I_3 &\le 2\max_{\mathbb{R}^n\times[0,s]}|\nabla_x\phi|\,|x_1-x_2|\int_{|y|>1}\nu(dy).
\end{aligned}
$$

Now pick $\epsilon=|x_1-x_2|^{1/2}\wedge1$. Note that $1\le\alpha<2$, we obtain $\epsilon^{\beta-\alpha}\le|x_1-x_2|^{\frac{\beta-\alpha}{2}}$, $\epsilon^{-\alpha}-1\le|x_1-x_2|^{-\frac{\alpha}{2}}$, $\epsilon^{1-\alpha}-1\le|x_1-x_2|^{\frac{1-\alpha}{2}}$ and $-\log\epsilon\le\frac{1}{\delta}|x_1-x_2|^{-\delta}$ for any $\delta>0$ (see (3.15)). Since $\beta>1$, we will choose $\delta=\frac{\beta-1}{2}$

in the following. Concluding from these inequalities and (4.4) - (4.7), we obtain

$$|I\phi(x_1,t) - I\phi(x_2,t)| \leq C_\Omega \left(\max_{\mathbb{R}^n \times [0,s]} |\nabla_x \phi| + \|\phi\|^{(\beta)}_{Q^1_s} \right) |x_1 - x_2|^{\frac{\beta-\alpha}{2}}, \tag{4.8}$$

where C_Ω is a sufficiently large constant independent of x_1, x_2 and t.

For the Hölder continuity of $t \to I\phi(x,t)$, since $\phi \in H^{\beta,\frac{\beta}{2}}(\overline{Q^1_s})$, it follows from Definition 2.2 that

$$\sum_{i=1}^n |\partial_{x_i}\phi(x,t_1) - \partial_{x_i}\phi(x,t_2)| \leq \|\phi\|^{(\beta)}_{Q_s} |t_1 - t_2|^{\frac{\beta-1}{2}}, \qquad \text{for } x \in \Omega \text{ and } t_1, t_2 \in [0,s].$$

Picking $\epsilon = |x_1 - x_2|^{\frac{1}{4}} \wedge 1$, an estimation similar to Lemma 3.1 gives us

$$|I\phi(x,t_1) - I\phi(x,t_2)| \leq C_\Omega \left(\widetilde{L}_t + \|\phi\|^{(\beta)}_{Q^1_s} \right) |t_1 - t_2|^{\frac{\beta-\alpha}{4}}, \tag{4.9}$$

where C_Ω is a sufficiently large constant independent of x, t_1 and t_2.

Now the first part of the lemma follows from (4.3), (4.8) and (4.9).

2. Noting that $\max_{D_s} |\phi| \leq \|\phi\|^{(\beta)}_{D_s}$ and $\max_{t_1,t_2 \in [0,s]} \frac{|\phi(x,t_1) - \phi(x,t_2)|}{|t_1 - t_2|^{\frac{1}{2}}} \leq s^{\frac{\beta-1}{2}} \|\phi\|^{(\beta)}_{D_s}$ (see Definition 2.2), the second part of the lemma follows from the same argument which we used in the first part of the proof. $\square$

Remark 4.1. *When the Lévy measure ν is a finite measure on $\mathbb{R}^n$, the integral form $\int_{\mathbb{R}^n} \phi(x+y,t)\,\nu(y)$ has the same regularity as $\phi(x,t)$ (see [25]). When the Lévy measure has a singularity, as we have seen in Lemma 4.1, the regularity of $I\,\phi$ decreases compared to the regularity of ϕ. Moreover, as we have seen in (4.1), the Hölder norm of $I\,\phi$ depends on the Hölder norm of ϕ on a slightly larger domain. This extension of domains will introduce a technical difficulty in estimating the Sobolev norm of u. This estimation will be carried out in the following section.*

4.2. **Solutions in the Sobolev sense.** As we have seen in Proposition 2.1, if the Lévy measure ν satisfies (H2), the value function u is the viscosity solution of the variational inequality (2.10). In the following, we will apply the regularity results for partial differential equations to show that u is also a solution of (2.10) in the Sobolev sense.

In this subsection, instead of (H7), we assume that

$$a_{ij}, b_i \text{ and } r \text{ are constants for } i,j \leq n, \text{ and } r \geq 0. \tag{H7'}$$

Moreover, there exist positive constants λ and Λ such that

$$\lambda|\xi|^2 \leq \sum_{i,j=1}^n a_{ij}\,\xi^i\xi^j \leq \Lambda|\xi|^2, \quad \forall \xi \in \mathbb{R}^n. \tag{H6'}$$

Remark 4.2. *Actually, the following two assumptions*

$$\lambda|\xi|^2 \leq \sum_{i,j=1}^n a_{ij}(x,t)\,\xi^i\xi^j \leq \Lambda|\xi|^2, \quad \forall (x,t) \in \mathbb{R}^n \times [0,T] \text{ and } \xi \in \mathbb{R}^n, \text{ and} \tag{H6''}$$

$$a_{ij}(x,t), b(x,t), r(x,t) \in H^{\ell,\frac{\ell}{2}}(\mathbb{R}^n \times [0,T]), \quad \forall \ell \in (0,1) \text{ and } i,j \leq n, \text{ and } r(x,t) \geq 0 \tag{H7''}$$

are sufficient for all results in this section except Lemma 5.5, where the constant coefficient assumption (H7') is necessary.

In order to work with non-smooth payoff functions, we assume that there exists a mollified sequence of g, denoted by $\{g^\epsilon\}_{\epsilon \in (0,\epsilon_0)}$ for some constant $\epsilon_0 < 1$, such that $\lim_{\epsilon \downarrow 0} g^\epsilon(x) = g(x)$ uniformly in compact subsets of $\mathbb{R}^n$ and

$$\text{each } g^\epsilon(x) \in H^{2+\ell}(\mathbb{R}^n) \quad \forall \ell \in (0,1). \tag{H8}$$

Moreover, there exist positive constants K, L and J independent of ϵ such that for all $x \in \mathbb{R}^n$

(H3')
$$0 \le g^\epsilon(x) \le K,$$

(H4')
$$|\nabla g^\epsilon(x)| \le L, \quad \text{and}$$

(H9)
$$\sum_{i,j=1}^{n} \partial^2_{x^i x^j} g^\epsilon(x)\, \xi^i \xi^j \ge -J\,|\xi|^2, \quad \forall \xi, x \in \mathbb{R}^n.$$

Remark 4.3. *Actually, for standard put option payoffs on multiple assets: $g(x) = \left[K - \frac{1}{n}\sum_{i=1}^n e^{x_i}\right]^+$ (the arithmetic average) and $g(x) = \left[K - \exp\left(\frac{1}{n}\sum_{i=1}^n x_i\right)\right]^+$ (the geometric average), mollified sequences can be constructed to satisfy the assumptions (H3'), (H4'), (H8) and (H9). Indeed, we can choose a sequence of functions $H_\epsilon(y) \in C^\infty(\mathbb{R})$ ($\epsilon \in (0, \epsilon_0)$ with ϵ_0 much smaller than K) such that $0 \le H_\epsilon^{'}(y) \le 1$, $H_\epsilon^{''}(y) \ge 0$ and $H_\epsilon(y) = \begin{cases} y, & y \ge \epsilon \\ 0, & y \le -\epsilon \end{cases}$. The mollified sequence $\{g^\epsilon\}_{\epsilon \in (0,\epsilon_0)}$ can be constructed by defining $g^\epsilon(x) = H_\epsilon\left(K - \frac{1}{n}\sum_{i=1}^n e^{x_i}\right)$ or $g^\epsilon(x) = H_\epsilon\left(K - \exp\left(\frac{1}{n}\sum_{i=1}^n x_i\right)\right)$. It is clear that $\lim_{\epsilon\downarrow 0} g^\epsilon(x) = g(x)$ uniformly in $\mathbb{R}$. Note that $H^{'}(y) > 0$ only when $y > -\epsilon$, one can check that (H3'), (H4'), (H8) and (H9) are satisfied for both cases.*

Given these assumptions, we are ready to state main result of this section.

Theorem 4.1. *If (H6'), (H7'), (H3'), (H4'), (H8) and (H9) are satisfied, moreover, the Lévy measure ν satisfies (H2) and (H5) with $\alpha \in [0, 2)$, then $u \in W^{2,1}_p(B_\rho(x_0) \times (0, T-s))$ for any integer $p \in (1, \infty)$, $\rho, s > 0$ and $x_0 \in \mathbb{R}^n$.*

Before we prove this key estimate in Section 5, let us list some corollaries of this result.

Corollary 4.1. *If the assumptions in Theorem 4.1 are satisfied, then for any $\rho, s > 0$ and $x_0 \in \mathbb{R}^n$*

(i) *$u \in H^{\beta, \frac{\beta}{2}}(\overline{B_\rho(x_0)} \times [0, T-s])$ where $\beta = 2 - \frac{n+2}{p} > 0$. In particular, $\nabla_x u \in C(\mathbb{R}^n \times [0, T))$. Therefore the smooth-fit property holds.*

(ii) *if the Lévy measure ν satisfies (H5) with $\alpha \in [1, 2)$, then Iu is well defined in the classical sense in $B_\rho(x_0) \times [0, T)$. Moreover, $I\,u \in H^{\frac{\beta-\alpha}{2}, \frac{\beta-\alpha}{4}}(\overline{B_\rho(x_0)} \times [0, T-s])$ for some $\beta \in (\alpha, 2)$.*

Proof. **(i)** Combining the result in Theorem 4.1 and the Sobolev Inequality (see e.g. Lemma 3.3 in [17] pp. 80), we have $u \in H^{\beta, \frac{\beta}{2}}(\overline{B_\rho(x_0)} \times [0, T-s])$, where $\beta = 2 - \frac{n+2}{p} > 0$. Choosing sufficiently large p such that $\beta > 1$, the continuity of $\nabla_x u$ follows from the definition of Hölder spaces in Definition 2.2 and the arbitrary choice of s.

(ii) It follows from the result in (i) for $\rho+1$ and the estimation (4.3) that Iu is well defined in $B_\rho(x_0) \times [0, T-s]$. Then the first statement of (ii) follows, since the choice of s is arbitrary. Choosing sufficiently large p such that $\beta > \alpha$, the second statement of (ii) follows from Lemma 4.1. □

Thanks to Corollary 4.1 (ii), we can consider the following boundary value problem with the driving term $I\,u$:

(4.10)
$$\begin{aligned} &(-\partial_t - \mathcal{L}_D + r)\, v(x,t) = Iu(x,t), \quad \text{for } (x,t) \in B \times [t_1, t_2), \\ &v(x,t) = u(x,t), \quad \text{for } (x,t) \in \partial B \times [t_1, t_2) \cup \overline{B} \times t_2, \end{aligned}$$

where $B \times (t_1, t_2) \subset \mathcal{C}$ is the bounded domain as in (3.20). The viscosity solution of (4.10) is defined similarly as in Definition 3.2, with operators $\mathcal{L}^f_D$ and I^f replaced by $\mathcal{L}_D$ and I respectively.

Rather than extending Lemma 3.3 to the infinite variation jump case, the following relation between the solutions in the Sobolev sense and the viscosity sense shows that the value function u is a viscosity solution of the boundary value problem (4.10). See Corollary 3 in [19] or Theorem 9.15 (ii) in [15] for its proof.

Lemma 4.2. *If $u \in W^{2,1}_p(B\times(t_1,t_2))$ for $p>n+1$ satisfies (4.10) at almost every point in $B\times(t_1,t_2)$, then u is the viscosity solution of (4.10) in the sense of Definition 3.2.*

Thanks to Corollary 4.1, Lemmas 4.1 and 4.2, the argument in Theorem 3.1 also works for the infinite variation jump case.

Theorem 4.2. *If the Lévy measure ν satisfies (H2) and (H5) with $1\le\alpha<2$, then the value function u is the unique classical solution, i.e., $u\in C^{2,1}$, of the boundary value problem (3.20). Moreover, $u\in C^{2,1}(\mathcal{C})$.*

Proof. Corollary 4.1 (ii) tells us that $Iu(x,t)\in H^{\frac{\beta-\alpha}{2},\frac{\beta-\alpha}{4}}(\overline{B}\times[t_1,t_2])$. As the value function u is shown to be a viscosity solution of (4.10) in Lemma 4.2, the rest proof follows from the same proof for Theorem 3.1. □

5. Proof of Theorem 4.1

Because the jump may have infinite variation, the proof of Theorem 4.1 needs to conquer several technical difficulties. We will carry the proof of Theorem 4.1 in a series of lemmas and point out the difficulties along the way.

Let us first define $v(x,t)=u(x,T-t)$ for $(x,t)\in\mathbb{R}^n\times[0,T]$. It is natural to expect that v solves the following variational inequality

$$\begin{aligned}&\min\left\{(\partial_t-\mathcal{L}_D-I+r)\,v(x,t),v(x,t)-g(x)\right\}=0,\quad (x,t)\in\mathbb{R}\times(0,T],\\ &v(x,0)=g(x).\end{aligned}\tag{5.1}$$

In this section, we will show that v indeed solves (5.1) for almost every point $(x,t)\in\mathbb{R}^n\times[0,T]$. Moreover, its $W^{2,1}_p$-norm is bounded on bounded domains of $\mathbb{R}^n\times[0,T]$. In the following, we will only carry out the proof of Theorem 4.1 for the infinite variation jump case, i.e., the Lévy measure ν satisfies (H5) with $1\le\alpha<2$. Since the integral operator has the reduced form I^f in (3.4) for the finite variation jumps, the proof of $0\le\alpha<1$ case in Theorem 4.1 will be similar and easier.

Motivated by Lemma 3.1 in [10] pp. 24 and [25], we will study the following penalty problem for each $\epsilon\in(0,\epsilon_0)$, where ϵ_0 is chosen before (H8):

$$\begin{aligned}&(\partial_t-\mathcal{L}_D-I+r)\,v^\epsilon(x,t)+p_\epsilon\left(v^\epsilon-g^\epsilon\right)=0,\quad (x,t)\in\mathbb{R}^n\times(0,T],\\ &v^\epsilon(x,0)=g^\epsilon(x),\end{aligned}\tag{5.2}$$

where the mollified sequence $\{g^\epsilon\}_{\epsilon\in(0,\epsilon_0)}$ satisfies (H3'), (H4'), (H8) and (H9). Here the penalty term $p_\epsilon(y)\in C^\infty(\mathbb{R})$ is chosen to satisfy following properties:

(5.3)

$$(i)\, p_\epsilon(y)\le 0,\quad (ii)\, p_\epsilon(y)=0 \text{ if } y\ge\epsilon,\quad (iii)\, p_\epsilon(0)=-n\Lambda J-|b|^{(0)}L-|r|^{(0)}K-J\int_{|y|\le 1}|y|^2\nu(dy)-K\int_{|y|>1}\nu(dy),$$

$$(iv)\, p^{'}_\epsilon(y)\ge 0,\quad (v)\, p^{''}_\epsilon(y)\le 0\quad \text{and}\quad (vi)\, \lim_{\epsilon\downarrow 0}p_\epsilon(y)=\begin{cases}0, & y>0\\ -\infty, & y<0\end{cases}.$$

Here constants Λ, K, L and J come from (H6") (H3'), (H4') and (H9) respectively. Additionally, $|b|^{(0)}=\max_{\mathbb{R}^n\times[0,T]}|b(x,t)|$ and $|r|^{(0)}=\max_{\mathbb{R}^n\times[0,T]}|r(x,t)|$ are finite due to (H7"). Moreover, $p_\epsilon(0)$ is also finite thanks to (2.3). It is also worth noticing that $p_\epsilon(0)$ is independent of ϵ. These properties of p_ϵ will be useful in later development. In particular, (5.3) (iii) will be essential for proofs of Lemma 5.6 and Corollary 5.2.

Let us recall the Schauder Fixed Point Theorem (see e.g. Theorem 2 in [9] pp. 189).

Lemma 5.1. *Let Θ be a closed convex subset of a Banach space and let $\mathcal{T}$ be a continuous operator on Θ such that $\mathcal{T}\Theta$ is contained in Θ and $\mathcal{T}\Theta$ is precompact. Then $\mathcal{T}$ has a fixed point in Θ.*

For each $\epsilon \in (0, \epsilon_0)$, we will show that the penalty problem (5.2) has a classical solution via the Schauder Fixed Point Theorem. Let us recall $D_s = \mathbb{R}^n \times [0, s]$.

Lemma 5.2. *If the Lévy measure ν satisfies (H2) and (H5) with $1 \le \alpha < 2$, then for any $\epsilon \in (0, \epsilon_0)$ and $\beta \in (\alpha, 2)$, (5.2) has a solution $v^\epsilon \in H^{2+\frac{\beta-\alpha}{2}, 1+\frac{\beta-\alpha}{4}}(D_T)$.*

Proof. We will first prove that (5.2) has a solution on a sufficiently small time interval $t \in [0, s]$ via the Schauder Fixed Point Theorem. Then we will extend this solution to the interval $[0, T]$.

Let us consider the set $\Theta = \left\{ v \in H^{\beta, \frac{\beta}{2}}(D_s) \text{ with its Hölder norm } \|v\|_{D_s}^{(\beta)} \le U_0 \right\}$, where positive constants s and U_0 will be determined later. It is clear that Θ is a bounded, closed and convex set in the Banach space $H^{\beta, \frac{\beta}{2}}(D_s)$. For any $v \in \Theta$, consider the following Cauchy problem for $u - g^\epsilon$:

$$
\begin{aligned}
&(\partial_t - \mathcal{L}_D + r)\,(u - g^\epsilon)(x, t) = Iv(x, t) - p_\epsilon(v - g^\epsilon)(x, t) + (\mathcal{L}_D - r)\, g^\epsilon(x), \quad (x, t) \in \mathbb{R} \times (0, s], \\
&u(x, 0) - g^\epsilon(x) = 0.
\end{aligned} \tag{5.4}
$$

Via the solution u of (5.4), the operator $\mathcal{T}$ can be defined as $u = \mathcal{T}v$. Let us check the conditions for the Schauder Fixed Point Theorem in the sequel.

1. Tv **is well defined.** Note that $v \in H^{\beta, \frac{\beta}{2}}(D_s)$ and $\beta \in (\alpha, 2)$, it follows from Lemma 4.1 (ii) that $Iv \in H^{\frac{\beta-\alpha}{2}, \frac{\beta-\alpha}{4}}(D_s)$ with

$$
\|Iv\|_{D_s}^{(\frac{\beta-\alpha}{2})} \le C\, \|v\|_{D_s}^{(\beta)}, \qquad \text{for some constant } C > 0 \text{ independent of } s. \tag{5.5}
$$

On the other hand, we can check that $p_\epsilon(v - g^\epsilon) \in H^{\frac{\beta-\alpha}{2}, \frac{\beta-\alpha}{4}}(D_s)$. Indeed, $p_\epsilon(v - g^\epsilon)$ is bounded in D_s, since both $v, g^\epsilon \in H^{\beta, \frac{\beta}{2}}(D_s)$ (see (H8)) and $p_\epsilon(y) \in C^0(\mathbb{R})$. Additionally, for any $x_1, x_2 \in \mathbb{R}^n$, $t \in [0, s]$

$$
|p_\epsilon(v - g^\epsilon)(x_1, t) - p_\epsilon(v - g^\epsilon)(x_2, t)| \le \max_{D_s} |p_\epsilon^{'}(v - g^\epsilon)|\; |(v - g^\epsilon)(x_1, t) - (v - g^\epsilon)(x_2, t)| \le \widetilde{C}|x_1 - x_2|.
$$

Here $\max_{D_s} |p_\epsilon^{'}(v - g^\epsilon)|$ is finite, which also follows from the boundness of $v - g^\epsilon$ and $p_\epsilon \in C^1(\mathbb{R})$. The positive constant $\widetilde{C}$ depends on $\max_{D_s} |p_\epsilon^{'}(v - g^\epsilon)|$ and the Hölder norms of v and g^ϵ. Meanwhile, the Hölder continuity of $p_\epsilon(v - g^\epsilon)$ in t can be checked similarly. Furthermore, $(\mathcal{L}_D - r)\, g^\epsilon(x) \in H^{\frac{\beta-\alpha}{2}, \frac{\beta-\alpha}{4}}(D_s)$ as a result of (H8). Therefore, thanks to (H6”) and (H7”), it follows from Theorem 5.1 in [17] pp. 320 that (5.4) has a uniqueness solution $u - g^\epsilon \in H^{2+\frac{\beta-\alpha}{2}, 1+\frac{\beta-\alpha}{4}}(D_s)$. Note that $g^\epsilon \in H^{2+\frac{\beta-\alpha}{2}, 1+\frac{\beta-\alpha}{4}}(D_s)$ (see (H8)), we have $u = Tv \in H^{2+\frac{\beta-\alpha}{2}, 1+\frac{\beta-\alpha}{4}}(D_s)$.

2. $\mathcal{T}\Theta \subset \Theta$. For $u = Tv$, appealing to Lemma 2 in [9] pp. 193, we obtain that there exists a positive constant A_β, depending on β, such that

$$
\begin{aligned}
\|u - g^\epsilon\|_{D_s}^{(\beta)} &\le A_\beta s^\gamma \left[\|Iv\|^{(0)} + \|p_\epsilon(v - g^\epsilon)\|^{(0)} + \|(\mathcal{L}_D - r)\, g^\epsilon\|^{(0)} \right] \\
&\le A_\beta C s^\gamma \|v\|_{D_s}^{(\beta)} + \widetilde{A},
\end{aligned} \tag{5.6}
$$

where $\gamma = \frac{2-\beta}{2}$, C is the constant in (5.5) and $\widetilde{A}$ is a sufficiently large constant dependent on $\|g^\epsilon\|_{\mathbb{R}^n}^{(2+\ell)}$ for some $\ell \in (0, 1)$. Let us take sufficiently small s such that $\tau \triangleq A_\beta C s^\gamma < 1/2$. Moreover, let us take $U_0 = \max\{\frac{2\widetilde{A}}{1-2\tau}, 2\, \|g^\epsilon\|_{D_s}^{(\beta)}\}$. Note that $\|v\|_{D_s}^{(\beta)} \le U_0$, it follows from (5.6) that

$$
\|u\|_{D_s}^{(\beta)} \le \|u - g^\epsilon\|_{D_s}^{(\beta)} + \|g^\epsilon\|_{D_s}^{(\beta)} \le \tau U_0 + \widetilde{A} + \frac{U_0}{2} \le \tau\, U_0 + \frac{1 - 2\tau}{2}\, U_0 + \frac{U_0}{2} = U_0. \tag{5.7}
$$

Therefore, $u = \mathcal{T}v \in \Theta$.

3. $\mathcal{T}\Theta$ is a precompact subset of $H^{\beta,\frac{\beta}{2}}(D_s)$. For any $\eta \in (\beta, 2)$, similar estimate as (5.6) shows that for any $v \in \Theta$, we have $\|\mathcal{T}v\|^{(\eta)}_{D_s} \leq U_1$ for some constant U_1 depending on U_0 and s. On the other hand, argument similar to Theorem 1 in [9] pp.188 shows that bounded subsets of $H^{\eta,\frac{\eta}{2}}(D_s)$ are precompact subsets of $H^{\beta,\frac{\beta}{2}}(D_s)$. Therefore, $\mathcal{T}\Theta$ is a precompact subset in $H^{\beta,\frac{\beta}{2}}(D_s)$.

4. $\mathcal{T}$ is a continuous operator. Let v_n be a sequence in Θ such that $\lim_{n\to\infty} \|v_n - v\|^{(\beta)}_{D_s} = 0$, we will show $\lim_{n\to\infty} \|Tv_n - Tv\|^{(\beta)}_{D_s} = 0$. From (5.4), $w \triangleq Tv_n - Tv$ satisfies the Cauchy problem

$$
\begin{aligned}
&(\partial_t - \mathcal{L}_D + r)\, w(x,t) = I(v_n - v)(x,t) - [p_\epsilon(v_n - g^\epsilon) - p_\epsilon(v - g^\epsilon)]\,, \quad (x,t) \in \mathbb{R}^n \times (0,s] \\
&w(x,0) = 0.
\end{aligned}
$$

It follows again from Lemma 2 in [9] pp. 193 that

$$
\begin{aligned}
\|\mathcal{T}v_n - \mathcal{T}v\|^{(\beta)}_{D_s} = \|w\|^{(\beta)}_{D_s} &\leq A_\beta s^\gamma \left[\|I(v_n - v)\|^{(0)} + \|p_\epsilon(v_n - g^\epsilon) - p_\epsilon(v - g^\epsilon)\|^{(0)}\right] \\
&\leq A_\beta s^\gamma \left[C\|v_n - v\|^{(\beta)}_{D_s} + \max_{D_s, n}\left|p'_\epsilon(v_n - g^\epsilon)\right| \|v_n - v\|^{(0)}\right] \to 0 \quad \text{as } n \to \infty.
\end{aligned}
$$

Concluding from **2. - 4.**, we obtain a fixed point of operator $\mathcal{T}$ in $H^{\beta,\frac{\beta}{2}}(D_s)$ as a result of the Schauder Fixed Point Theorem. We denote this fixed point as v^ϵ. Moreover, it follows from the result in **1.** that $v^\epsilon = \mathcal{T}v^\epsilon \in H^{2+\frac{\beta-\alpha}{2}, 1+\frac{\beta-\alpha}{4}}(D_s)$.

Finally, let us extend v^ϵ to the interval $[0,T]$. Choosing any $\rho \in (0, T-s)$, we replace $g^\epsilon(\cdot)$ by $v^\epsilon(\cdot,\rho)$ in (5.4). Note that the choice of s in **2.** only depend on β and C, but not on ρ. If $\|v^\epsilon(\cdot,\rho)\|^{(2+\frac{\beta-\alpha}{2})}_{\mathbb{R}^n}$ is finite, we can choose a sufficiently large U_0, depending on $\|v^\epsilon(\cdot,\rho)\|^{(2+\frac{\beta-\alpha}{2})}_{\mathbb{R}^n}$, such that (5.7) holds on $[\rho, \rho+s]$, moreover $\|v^\epsilon(\cdot,\rho+s)\|^{(2+\frac{\beta-\alpha}{2})}_{\mathbb{R}^n}$ is finite thanks to the result after **4.**. Noticing that $\|g^\epsilon\|^{(2+\ell)}_{\mathbb{R}^n}$ is finite for any $\ell \in (0,1)$, one can extend the time interval by s each time, until the time interval contains $[0,T]$. Therefore we have the statement of the lemma. □

Remark 5.1. *Because of the regularity decreases after applying the integral operator (see Remark 4.1), it is no longer straight forward to use the "bootstraping scheme" which was used in Theorem 2.1 of [25] to explore the higher regularity of v^ϵ. Instead, we will use a new technique to study the higher regularity of v^ϵ in the proof of Lemma 5.5.*

Thanks to the definition of the Hölder spaces, Lemma 5.2 also tells us that v^ϵ is bounded in D_T. In order to show that v^ϵ is the unique bounded classical solution of the penalty problem (5.2), we need the following Maximum Principle for the parabolic integro-differential operator. The proof of it is provided in Appendix A. (See Lemma 2.1 of [25] for a similar Maximum Principle, where ν is assumed to be a finite measure on $\mathbb{R}$.)

Lemma 5.3. *Let us assume that $a_{ij}(x,t)$, $b_i(x,t)$ and $c(x,t)$ are bounded in $\mathbb{R}^n \times [0,T]$ with $A = (a_{ij})_{n\times n}$ satisfying $\sum_{i,j=1}^n a_{ij}(x,t)\,\xi^i\xi^j > 0$ for any $\xi \in \mathbb{R}^n \setminus \{0\}$, moreover $c(x,t) \geq 0$ and the Lévy measure satisfies (H2). If $v \in C^0([0,T]\times\mathbb{R}^n) \cap C^{2,1}((0,T]\times\mathbb{R}^n)$ satisfies $(\partial_t - \mathcal{L}_\mathcal{D} - I + c(x,t))\, v(x,t) \geq 0$ in $\mathbb{R}\times(0,T]$ and there exists a sufficiently large positive constant m such that $v(x,t) \geq -m$ for $(x,t) \in \mathbb{R}^n \times [0,T]$. Then $v(x,0) \geq 0$ implies that $v(x,t) \geq 0$ for $(x,t) \in \mathbb{R}^n \times [0,T]$.*

As a corollary of this Maximum Principle, the bounded classical solution of the penalty problem (5.2) is unique.

Corollary 5.1. *For each $\epsilon \in (0, \epsilon_0)$, the penalty problem (5.2) has a unique bounded classical solution.*

Proof. Let us assume v_1 and v_2 are two bounded solutions of (5.2). Then $v_1 - v_2$ satisfies

$$
\begin{aligned}
&(\partial_t - \mathcal{L}_D - I + r)\ (v_1 - v_2) + p_\epsilon(v_1 - g^\epsilon) - p_\epsilon(v_2 - g^\epsilon) = 0, \quad (x,t) \in \mathbb{R}^n \times (0,T], \\
&(v_1 - v_2)(x,0) = 0
\end{aligned}
\tag{5.8}
$$

On the other hand, it follows from the mean value theorem that $p_\epsilon(v_1 - g^\epsilon) - p_\epsilon(v_2 - g^\epsilon) = p_\epsilon^{'}(y)(v_1 - v_2)$ for some $y \in \mathbb{R}^n$. Moreover, $p_\epsilon^{'}(y)$ is bounded, say by M, thanks to the fact that $p_\epsilon \in C^1(\mathbb{R})$ and v_1, v_2 and g^ϵ are all bounded. Now applying Lemma 5.3 to the equation (5.8) and choosing $c = r + M \geq 0$ (see (5.3) (iv)), we have $v_1(x,t) \geq v_2(x,t)$ for $(x,t) \in \mathbb{R}^n \times (0,T]$. The other direction of the inequality follows from applying the same argument to $v_2 - v_1$. □

Applying Lemma 5.3, we will analyze some universal properties of v^ϵ for all $\epsilon \in (0, \epsilon_0)$ in the following three lemmas.

Lemma 5.4.

$$0 \leq v^\epsilon(x,t) \leq K + 1, \quad \text{for } (x,t) \in \mathbb{R}^n \times [0,T].$$

Proof. Since the proof is similar to the proof of Lemma 2.2 in [25], we give it in the Appendix A. □

Lemma 5.5.

$$|\partial_{x^k} v^\epsilon(x,t)| \leq L, \quad \text{for } (x,t) \in \mathbb{R}^n \times [0,T], 1 \leq k \leq n.$$

Proof. Intuitively, thanks to the constant coefficient assumption (H7'), it follows from (5.2) that $\partial_{x^k} v^\epsilon$ satisfies

$$\begin{aligned} &(\partial_t - \mathcal{L}_D - I + r)\, w + p_\epsilon^{'}(v^\epsilon - g^\epsilon)(w - \partial_{x^k} g^\epsilon) = 0, \quad (x,t) \in \mathbb{R}^n \times (0,T], \\ &w(x,0) = \partial_{x^k} g^\epsilon(x), \end{aligned} \tag{5.9}$$

where coefficients unchanged compared to (5.2). However, given the result in Lemma 5.2, it is only known that v^ϵ has continuous derivatives of the form $\partial^2_{x^i x^j} v^\epsilon$, $\partial_{x^i} v^\epsilon$ and $\partial_t v^\epsilon$. While it is necessary for v^ϵ to have derivatives of higher orders to ensure $\partial_{x^k} v^\epsilon$ as the classical solution of (5.9). Therefore, we will first prove that $\partial_{x^k} v^\epsilon$ is indeed the classical solution of (5.9).

Let us consider the equation

$$\begin{aligned} &(\partial_t - \mathcal{L}_D - I + r)\, w = -p_\epsilon^{'}(v^\epsilon - g^\epsilon)(\partial_{x^k} v^\epsilon - \partial_{x^k} g^\epsilon), \quad (x,t) \in \mathbb{R}^n \times (0,T], \\ &w(x,0) = \partial_{x^k} g^\epsilon(x). \end{aligned} \tag{5.10}$$

Thanks to Lemma 5.2 and (H8), $-p_\epsilon^{'}(v^\epsilon - g^\epsilon)(\partial_{x^k} v^\epsilon - \partial_{x^k} g^\epsilon)$ is Hölder continuous. Therefore, it follows from Theorem 3.1 in [11] pp. 89 that (5.10) has a unique classical solution. Let us call it w.

For any point $(x,t) \in \mathbb{R}^n \times [0,T]$, we will show that $\partial_{x^k} v^\epsilon(x,t) = w(x,t)$. As a vector in $\mathbb{R}^n$, $x = (x^1, \cdots, x^n)$. Let us also denote $x(z) \triangleq (x^1, \cdots, x^{k-1}, z, x^{k+1}, \cdots, x^n)$. One can check that $v(x,t) \triangleq \int_0^{x^k} w(x(z),t)\, dz + v^\epsilon(x(0),t)$ is a classical solution of the following Cauchy problem

$$\begin{aligned} &(\partial_t - \mathcal{L}_D - I + r)\, v = -p_\epsilon(v^\epsilon - g^\epsilon), \quad (x,t) \in \mathbb{R}^n \times (0,T], \\ &v(x,0) = g^\epsilon(x). \end{aligned} \tag{5.11}$$

Moreover, thanks to estimate (3.6) in Theorem 3.1 of [11] pp. 89, v is a bounded on $\mathbb{R}^n \times [0,T]$. On the other hand, using Lemma 5.3 one can show that (5.11) has a unique bounded classical solution. Therefore, it follows from Corollary 5.1 that $v(x,t) = v^\epsilon(x,t)$ for $(x,t) \in \mathbb{R}^n \times [0,T]$. As a result $\partial_{x^k} v^\epsilon(x,t) = w(x,t)$ and $\partial_{x^k} v^\epsilon$ is a classical solution of (5.9).

The rest of the proof is same as the proof of Lemma 2.4 in [25]. Thanks to Lemma 5.2, $|\partial_{x^k} v^\epsilon|$ is already bounded on $\mathbb{R}^n \times [0,T]$. We will show it is bounded uniformly in ϵ in the following. Let $u = L + \partial_{x^k} v^\epsilon$, $u \in C^0([0,T] \times \mathbb{R}^n) \cap C^{2,1}((0,T] \times \mathbb{R}^n)$ and it satisfies

$$\begin{aligned} &(\partial_t - \mathcal{L}_D - I + r)\, u + p_\epsilon^{'}(v^\epsilon - g^\epsilon)\, u = p_\epsilon^{'}(v^\epsilon - g^\epsilon)(\partial_{x^k} g^\epsilon + L) + r\, L, \\ &u(x,0) = L + \partial_{x^k} g(x). \end{aligned} \tag{5.12}$$

Note (H4') and (5.3) (iv), $u(x,t) \geq 0$ follows from applying Lemma 5.3 to (5.12) by picking $c = r + p_\epsilon^{'}(v^\epsilon - g^\epsilon)$. The proof for the upper bound can be performed similarly by picking $u = L - \partial_{x^k} v^\epsilon$. □

Remark 5.2. *The constant coefficient assumption (H7') makes sure that the coefficient before u in (5.12) is nonnegative in order to apply the Maximal Principle Lemma 5.3.*

Lemma 5.6. *For any $\epsilon \in (0, \epsilon_0)$, $v^\epsilon(x,t) \geq g^\epsilon(x)$ on $\mathbb{R}^n \times [0,T]$.*

Proof. Let us first show that $I g^\epsilon(x)$ is uniformly bounded from below. Compared to [25] where ν is a finite measure, we will bound $I g^\epsilon(x)$ from below in the following way.

$$
\begin{aligned}
I g^\epsilon(x) &= \int_{|y|\leq 1} \left[g^\epsilon(x+y) - g^\epsilon(x) - \sum_{i=1}^n y^i \frac{\partial}{\partial x^i} g^\epsilon(x) \right] \nu(dy) + \int_{|y|>1} \left[g^\epsilon(x+y) - g^\epsilon(x) \right] \nu(dy) \\
&= \int_{|y|\leq 1} \nu(dy) \int_0^1 dz (1-z) \sum_{i,j=1}^n y^i y^j \frac{\partial^2}{\partial x^i \partial x^j} g^\epsilon(x+zy) + \int_{|y|>1} \left[g^\epsilon(x+y) - g^\epsilon(x) \right] \nu(dy) \\
&\geq \int_{|y|\leq 1} \nu(dy) \int_0^1 dz (1-z) \left(-J|y|^2 \right) - K \int_{|y|>1} \nu(dy) \\
&\geq -J \int_{|y|\leq 1} |y|^2 \nu(dy) - K \int_{|y|>1} \nu(dy),
\end{aligned} \tag{5.13}
$$

where the first inequality follows from (H9) and (H3').

On the other hand, thanks to (H6") and (H9), $\sum_{i,j}^n a_{ij}(x,t)\, \partial^2_{x^i x^j} g^\epsilon(x)$ is also bounded from below. Note that $\sum_{i,j}^n a_{ij}(x,t)\, \partial^2_{x^i x^j} g^\epsilon(x) = tr(AH(g^\epsilon))$, where $H(g^\epsilon)$ is the Hessian of g^ϵ, i.e., $H(g^\epsilon)_{ij} = \partial^2_{x^i x^j} g^\epsilon(x)$. It follows from the first inequality in (H6") that A is a positive definite matrix. Then there exists a nonsingular matrix C such that $A = CC^{'}$. Therefore $tr(AH(g^\epsilon)) = tr(CC^{'} H(g^\epsilon)) = tr(C^{'} H(g^\epsilon) C)$. Moreover, (H9) and (H6") give us that

$$
(C\xi)^{'}\, H(g^\epsilon)\, (C\xi) \geq -J \left(\xi^{'} C^{'} C\, \xi \right) = -J \left(\xi^{'} A\, \xi \right) \geq -J\Lambda |\xi|^2, \quad \forall \xi \in \mathbb{R}^n.
$$

Hence $C^{'} H(g^\epsilon)\, C + J\Lambda\, I_n$ is a non-negative definite matrix. As a result, we have $tr\left(C^{'} H(g^\epsilon)\, C \right) + nJ\Lambda = tr\left(C^{'} H(g^\epsilon)\, C + J\Lambda I_n \right) \geq 0$, which implies

$$
\sum_{i,j=1}^n a_{ij}(x,t)\, \frac{\partial^2}{\partial x^i \partial x^j} g^\epsilon(x) = tr\left(AH(g^\epsilon) \right) \geq -nJ\Lambda. \tag{5.14}
$$

Thanks to (5.13) and (5.14), we can bound $(\partial_t - \mathcal{L}_D - I + r)\, g^\epsilon(x)$ from above. Indeed,

$$
\begin{aligned}
& (\partial_t - \mathcal{L}_D - I + r)\, g^\epsilon(x) \\
&\quad = - \sum_{i,j=1}^n a_{ij}(x,t)\, \frac{\partial^2}{\partial x^i \partial x^j} g^\epsilon(x) - \sum_{i=1}^n b_i(x,t)\, \frac{\partial}{\partial x^i} g^\epsilon(x) + r(x,t)\, g^\epsilon(x) - I g^\epsilon(x) \\
&\quad \leq nJ\Lambda + |b|^{(0)} L + |r|^{(0)} K + J \int_{|y|\leq 1} |y|^2 \nu(dy) + K \int_{|y|>1} \nu(dy) \\
&\quad = -p_\epsilon(0),
\end{aligned} \tag{5.15}
$$

where the second equality follows from (5.3) (iii).

Now we will show $v^\epsilon \geq g^\epsilon$ via the Maximum Principle in Lemma 5.3. It follows from (5.15) that

$$
\begin{aligned}
(\partial_t - \mathcal{L}_D - I + r)\, (v^\epsilon - g^\epsilon) &= -p_\epsilon\, (v^\epsilon - g^\epsilon) - (\partial_t - \mathcal{L}_D - I + r)\, g^\epsilon \\
&\geq -p_\epsilon\, (v^\epsilon - g^\epsilon) + p_\epsilon(0).
\end{aligned}
$$

Combining with the mean value theorem, we obtain

$$\left(\partial_t - \mathcal{L}_D - I + r + p_\epsilon^{'}(y)\right)(v^\epsilon - g^\epsilon) \geq 0, \tag{5.16}$$

where $y \in \mathbb{R}^n$ and $p_\epsilon^{'}(y)$ is bounded. Therefore the statement of the lemma follows applying Lemma 5.3 to (5.16) and choosing $c = r + p_\epsilon^{'}(y) \geq 0$. □

As an easy corollary, the penalty terms are uniformly bounded.

Corollary 5.2. *$p_\epsilon\,(v^\epsilon - g^\epsilon)$ is bounded uniformly in $\epsilon \in (0, \epsilon_0)$.*

Proof. Thanks to Lemma 5.6 and (5.3) (i) and (iv), we have $p_\epsilon(0) \leq p_\epsilon\,(v^\epsilon - g^\epsilon) \leq 0$. The statement follows noticing that $p_\epsilon(0)$ (in (5.3) (iii)) is independent of ϵ. □

Thanks to Lemmas 5.2, 5.4, 5.5 and Corollary 5.2, we can apply the following $W_p^{2,1}$-norm estimate for the parabolic integro-differential equation to each solution v^ϵ of the penalty problem.

Since the proof of the following theorem is technical and independent of the penalty problem, we will perform it in the Appendix B.

Theorem 5.1. *Let us assume the Lévy measure satisfies (H5) with $\alpha \in [0, 2)$, if v is a $W_{p,\,loc}^{2,1}$ solution of the following Cauchy problem for some positive integer p,*

$$\begin{aligned} &(\partial_t - \mathcal{L}_D - I + r)\,v = f(x,t), \quad (x,t) \in \mathbb{R}^n \times (0, T], \\ &v(x, 0) = g(x), \end{aligned} \tag{5.17}$$

where the coefficients satisfy (H6”), (H7”) and $f \in L_{p,\,loc}(\mathbb{R}^n \times (0,T))$, moreover $|v|$ is bounded on $\mathbb{R}^n \times [0,T]$ and $|\nabla_x v|$ is bounded on any compact domain of $\mathbb{R}^n \times [0,T]$. Then for any domain $B_\rho(x_0) \times (s,T)$ for any $\rho > 0$, $s \in (0,T)$ and $x_0 \in \mathbb{R}^n$

$$\|v\|_{W_p^{2,1}(B_\rho(x_0)\times(s,T))} \leq C_\delta \left[\max_{\mathbb{R}^n\times[0,T]} |v| + \max_{B_{\rho+\delta/4+1}(x_0)\times[0,T]} |\nabla_x v| + \|f\|_{L^p(B_{\rho+\delta/4}(x_0)\times(\delta/2,T))}\right], \tag{5.18}$$

for some positive constant C_δ and $\delta < s$.

Remark 5.3. *The existence of the $W_p^{2,1}$ solution for (5.17) was ensured by Theorem 3.2 in [2] pp.234. However, the norm estimation was not given there. On the other hand, since the integral operator I is non-local, it is important to study the Cauchy problem (5.17) on the entire domain $\mathbb{R}^n \times [0,T]$. Otherwise, for the Cauchy problem on bounded domains of $\mathbb{R}^n \times [0,T]$ with some boundary conditions, $W_p^{2,1}$ solutions are not expected in general, see [13] for a counterexample.*

A $W_p^{2,1}$-norm estimate, similar to (5.18), for the parabolic integro-differential equation was proved in Theorem 3.5 in [11] pp. 91. However, the estimation in [11] requires the jump restricted in a bounded domain, i.e., if $x \in \Omega$ where Ω is a bounded domain in $\mathbb{R}^n$, the jump size $z(x)$, which is state dependent, can only be chosen such that $x + z(x) \in \Omega$ (see (1.54) in [11] pp. 63). However, this restriction is not satisfied in our case, where the jump size is unbounded and independent of the state variable x.

Applying Theorem 5.1 to each penalty problem (5.2), thanks to Lemmas 5.2, 5.4, 5.5 and Corollary 5.2, we have the following corollary.

Corollary 5.3. *If the Lévy measure satisfies (H2) and (H5) with $\alpha \in [1, 2)$, moreover (H6’), (H7’), (H3’), (H4’), (H8) and (H9) are also satisfied, then for any domain $B_\rho(x_0) \times (s,T)$ for any $\rho > 0$, $s \in (0,T)$ and $x_0 \in \mathbb{R}^n$,*

$\|v^\epsilon\|_{W^{2,1}_p(B_\rho(x_0)\times(s,T))}$ are bounded uniformly in $\epsilon \in (0,\epsilon_0)$ for any integer $p \in (1,\infty)$, i.e., there is a constant C independent of ϵ such that

$$\|v^\epsilon\|_{W^{2,1}_p(B_\rho(x_0)\times(s,T))} \leq C. \tag{5.19}$$

Proof. It follows from Lemma 5.2 that $v^\epsilon \in W^{2,1}_{p,loc}(\mathbb{R}^n\times(0,T))$. Thanks to Lemmas 5.4 and 5.5, both $\max_{\mathbb{R}^n\times[0,T]}|v^\epsilon|$ and $\max_{\mathbb{R}^n\times[0,T]}|\nabla_x v^\epsilon|$ are also bounded uniformly in ϵ. Moreover, picking $f = -p_\epsilon(v^\epsilon - g^\epsilon)$, it follows from Corollary 5.2 that f is also bounded uniformly in ϵ. Concluding from these facts, (5.19) follows (5.18). □

Remark 5.4. *The estimate in Theorem 5.1 is essential for the proof of Corollary 5.3. However, having infinite variation jumps presents two technical difficulties to the proof of Theorem 5.1.*

First, as we shall see in Lemma B-1, once the Lévy measure has a singularity, the L_p-norm of Iv^ϵ depends on the $W^{2,1}_p$-norm of v^ϵ. Therefore, one could not consider Iv^ϵ as a driving term directly and use the classical $W^{2,1}_p$-norm estimate for parabolic differential equations (without the integral term) to bound the $W^{2,1}_p$-norm of v^ϵ by the L_p-norm of $I\,v^\epsilon$. When the Lévy measure is a finite measure as in [25], L_p-norm of Iv^ϵ only depends on L^∞-norm of v^ϵ. Therefore, Lemma 2.6 in [25] follows from the classical $W^{2,1}_p$-norm estimate for parabolic differential equations, i.e., the $W^{2,1}_p$-norm of v^ϵ is bounded by L^∞-norm of v^ϵ.

Second, as we have seen in Remark 4.1 and we shall see it again in Lemma B-1, the regularity of Iv^ϵ actually depends on regularity of v^ϵ on a larger domain. This extension of the domain is another technical difficulty we face in the proof of Theorem 5.1, because the extension of domains implies that $W^{2,1}_p$-norm of v^ϵ on a bounded domains depends on its $W^{2,1}_p$-norm on a slightly larger domain.

To conclude this section, in the following theorem we will find a limit v^* of the sequence $\{v^\epsilon\}_{\epsilon\in(0,\epsilon_0)}$ such that v^* is indeed the value function v defined at the beginning of this section.

Theorem 5.2. *Let us assume that (H6'), (H7'), (H3'), (H4'), (H8) and (H9) are satisfied, moreover, the Lévy measure ν satisfies (H2) and (H5) with $\alpha \in [1,2)$. Then for any $s,\rho > 0$ and $x_0 \in \mathbb{R}^n$, there exists a subsequence $\{\epsilon_k\}_{k\geq 0}$ such that v^{ϵ_k} converges uniformly to the limit v^* uniformly in $\overline{B_\rho(x_0)}\times[s,T]$ as $\epsilon_k \to 0$. Moreover, v^* solves the variational inequality (5.1) for almost every point in $\mathbb{R}^n\times[0,T]$ and $v^* \in W^{2,1}_p(B_\rho(x_0)\times(s,T))$ for any integer $p\in(1,\infty)$.*

Proof. Thanks to Corollary 5.3, there exists a subsequence $\{\epsilon_k\}$ with $\epsilon_k \to 0$ and a function $v^* \in W^{2,1}_p(B_\rho(x_0)\times(s,T))$ such that

$$v^{\epsilon_k} \rightharpoonup v^* \quad \text{in } W^{2,1}_p(B_\rho(x_0)\times(s,T)).$$

Here " $\rightharpoonup$ " represents weak convergence, please refer to Appendix D.4. in [7] pp. 639 for its definition and properties. The rest of the proof is the same as proof of Theorem 3.2 in [25]. It confirms that v^* solves the variational inequality (5.1) for almost every point in $\mathbb{R}^n\times[0,T]$. □

Finally, thanks to the verification result Proposition 2.2, v^* must be the v defined at the beginning of this section. As a result, the $1\leq\alpha<2$ case of Theorem 4.1 follows from Theorem 5.2 after reversing the time.

Appendix A. Proof of several lemmas in Sections 2, 3 and 4

Proof of Lemma 2.1. Throughout this proof, in order to distinguish the Euclidean norm in $\mathbb{R}^n$ from the absolute value in $\mathbb{R}$, we denote the Euclidean norm as $\|\cdot\|$ and the absolute value as $|\cdot|$. Actually, the norm $\|\cdot\|$ is equivalent to the sum of the norms $|\cdot|$ among all components, i.e.,

$$\|y\| \leq \sum_{i=1}^n |y^i| \leq n\,\|y\|, \quad \text{for any } y\in\mathbb{R}^n. \tag{A-1}$$

Thanks to (A-1), (2.4) - (2.7) can be showed under a slighter weaker assumption (H2), compared to the assumption $\int_{|y|>1} |y|^2 \,\nu(dy)$ in Lemma 3.1 of [21]. We will only prove (2.6) and (2.7) in the following.

Following from (1.1) and (2.2), we have for any $\tau \in \mathcal{T}_{0,t}$ that

$$\|X_\tau^x - x\| \le \left\| \int_0^\tau b\left(X_s^x, s\right) ds \right\| + \left\| \int_0^\tau \sigma\left(X_s^x, s\right) dW_s \right\| + \left\| \mathcal{J}_\tau^\ell \right\| + \left\| \lim_{\epsilon \downarrow 0} \mathcal{J}_\tau^\epsilon \right\|. \tag{A-2}$$

Comparing to the proof of Lemma 3.1 in [21], the difference is on the estimation on the large jump term. Therefore, we will focus on $\left\|\mathcal{J}_\tau^\ell\right\|$ in the following.

First it follows from (2.2) and the triangle inequality that

$$\mathbb{E}\left\|\mathcal{J}_\tau^\ell\right\| = \mathbb{E}\left\| \int_0^\tau \int_{\|y\|>1} y\,\mu(ds,dy) \right\| \le \mathbb{E}\left\| \int_0^\tau \int_{\|y\|>1} y\,\widetilde{\mu}(ds,dy) \right\| + \mathbb{E}\left\| \int_0^\tau ds \int_{\|y\|>1} y\,\nu(dy) \right\|. \tag{A-3}$$

Let us estimate the right-hand-side of (A-3) separately. On the one hand, $\int_0^t \int_{\|y\|>1} y\,\widetilde{\mu}(ds,dy)$ is a martingale because of (H2). Hence $\left\|\int_0^t \int_{\|y\|>1} y\,\widetilde{\mu}(ds,dy)\right\|$ is a submartingale (see e.g. Problem 3.7 in [16] pp. 13). Noticing that $\tau \in \mathcal{T}_{0,t}$, it follows from the Optional Sampling Theorem that

$$\mathbb{E}\left\| \int_0^\tau \int_{\|y\|>1} y\,\widetilde{\mu}(ds,dy) \right\| \le \mathbb{E}\left\| \int_0^t \int_{\|y\|>1} y\,\widetilde{\mu}(ds,dy) \right\|. \tag{A-4}$$

Thanks to (A-1), we can estimate the right-hand-side of (A-4) as follows.

$$\begin{aligned}
\mathbb{E}\left\| \int_0^t \int_{\|y\|>1} y\,\widetilde{\mu}(ds,dy) \right\| &\le \mathbb{E}\sum_{i=1}^n \left| \int_0^t \int_{\|y\|>1} y^i\,\widetilde{\mu}(ds,dy) \right| \\
&\le \mathbb{E}\sum_{i=1}^n \left| \int_0^t \int_{\|y\|>1} y^i\,\mu(ds,dy) \right| + \sum_{i=1}^n \int_0^t ds \int_{\|y\|>1} \left|y^i\right|\,\nu(dy) \\
&\le \mathbb{E}\int_0^t \int_{\|y\|\ge 1} \sum_{i=1}^n \left|y^i\right|\,\mu(ds,dy) + \int_0^t ds \int_{\|y\|>1} \sum_{i=1}^n \left|y^i\right|\,\nu(dy) \\
&= 2\int_0^t ds \int_{\|y\|>1} \sum_{i=1}^n \left|y^i\right|\,\nu(dy) \le 2\,n \int_{\|y\|>1} \|y\|\,\nu(dy)\cdot t.
\end{aligned} \tag{A-5}$$

Here the first and fourth inequalities follow from (A-1). Moreover, the third inequality follows since the Poisson random measure μ is a non-negative measure on $\mathbb{R}_+ \times \mathbb{R}^n$ for each $\omega \in \Omega$. On the other hand, the second term on the right-hand-side of (A-3) can be estimated similarly using (A-1).

Concluding from (A-3) - (A-5), we can find a positive constant C such that $\mathbb{E}\left\|\mathcal{J}_\tau^\ell\right\| \le C\,t$ for any $\tau \in \mathcal{T}_{0,t}$. The other three terms on the right-hand-side of (A-2) can be estimated in the same way as in Lemma 3.1 of [21]. In particular, the stochastic integral and the small jump terms are bounded by $C\,t^{1/2}$. Moreover, compared to the estimate (3.3) in [21], the boundness of b and σ ensures that the constant C in (2.6) is independent of x.

For (2.7), we will still focus on the large jump term. Instead of applying the Doob's inequality as in Lemma 3.1 in [21], we will use properties of μ to derive the following estimate.

$$\begin{aligned}\mathbb{E}\left[\sup_{0\leq s\leq t}\left\|\mathcal{J}^\ell_s\right\|\right] &= \mathbb{E}\left[\sup_{0\leq s\leq t}\left\|\int_0^s\int_{\|y\|>1} y\,\mu(du,dy)\right\|\right] \leq \mathbb{E}\left[\sup_{0\leq s\leq t}\sum_{i=1}^n\left|\int_0^s\int_{\|y\|>1} y^i\,\mu(du,dy)\right|\right]\\ &\leq \mathbb{E}\left[\sup_{0\leq s\leq t}\int_0^s\int_{\|y\|>1}\sum_{i=1}^n|y^i|\,\mu(du,dy)\right] \leq \mathbb{E}\left[\int_0^t\int_{\|y\|>1}\sum_{i=1}^n|y^i|\,\mu(du,dy)\right]\\ &= \int_0^t du\int_{\|y\|>1}\sum_{i=1}^n\left|y^i\right|\,\nu(dy) \leq n\int_{\|y\|>1}\|y\|\,\nu(dy)\cdot t.\end{aligned} \tag{A-6}$$

Here the first and fourth inequalities follow from (A-1), the second and the third inequalities hold since μ is a non-negative measure for each $\omega\in\Omega$. The rest proof of (2.7) follows from the same approach used in Lemma 3.1 of [21]. □

Proof of Lemma 3.3. Thanks to Lemma 3.1, the driving term $I^f u$ in (3.27) is well defined in the classical sense and Hölder continuous in both its variables. We will only prove the statement for the subsolution. The statement for the supersolution can be shown in the similar manner.

Given u as a subsolution of (3.27), we will show that u is a viscosity subsolution of (3.20). According to Definition 3.1, for any $(x_0,t_0)\in\overline{B}\times[t_1,t_2]$, the test function $\phi(x,t)$ is chosen such that

$$u(x_0,t_0)-\phi(x_0,t_0)=\max_{(x,t)\in\mathbb{R}^n\times[t_1,t_2]}\left[u(x,t)-\phi(x,t)\right].$$

Therefore $u(x_0+y,t_0)-u(x_0,t_0)\leq\phi(x_0+y,t_0)-\phi(x_0,t_0)$ for any $y\in\mathbb{R}^n$. Since ν is a positive measure, we have from (3.4) that

$$I^f u(x_0,t_0)\leq I^f\phi(x_0,t_0). \tag{A-7}$$

Here $\phi(x,t)$ is chosen in $C_1(\mathbb{R}^n\times[t_1,t_2])$ so that $I^f\phi(x_0,t_0)$ is finite under the assumption (H2). Thanks to (A-7), we obtain from (3.28) that

$$(-\partial_t-\mathcal{L}_D+r)\,\phi(x_0,t_0)\leq I^f u(x_0,t_0)\leq I^f\phi(x_0,t_0),\qquad\text{for } (x_0,t_0)\in B\times[t_1,t_2].$$

Moreover, (3.22) and (3.23) are automatically satisfied because $u(x,t)$ itself is the boundary and terminal value (3.27). Therefore according to Definition 3.1, $u(x,t)$ is a subsolution of (3.20).

Conversely, let us assume that $u(x,t)$ is a subsolution of (3.20), for any $(x_0,t_0)\in\overline{B}\times[t_1,t_2]$, given any function $\phi(x,t)\in C^{2,1}(\mathbb{R}^n\times[t_1,t_2])$ such that $\phi(x_0,t_0)=u(x_0,t_0)$ and $\phi(x,t)\geq u(x,t)$ for all $(x,t)\in\mathbb{R}^n\times[t_1,t_2]$, let us construct ϕ^ϵ for $\epsilon\in(0,1)$ as follows.

$$\phi^\epsilon(x,t)\triangleq\phi(x,t)\chi^\epsilon(x)+\widetilde{u}(x,t)\left(1-\chi^\epsilon(x)\right),$$

where χ^ϵ is a smooth function satisfying $0\leq\chi^\epsilon\leq 1$, $\chi^\epsilon(x)=1$ when $x\in B_\epsilon(x_0)$ and $\chi^\epsilon(x)=0$ when $x\in\mathbb{R}^n\setminus B_{2\epsilon}(x_0)$. Moreover, $\widetilde{u}\in C^\infty(\mathbb{R}^n\times[t_1,t_2])$ such that $u\leq\widetilde{u}\leq u+\epsilon^2$ on $\mathbb{R}^n\times[t_1,t_2]$, for example, the usual mollification $\widetilde{u}=u*\zeta^\delta+\epsilon^2$ for sufficiently small δ (Please see [7] pp. 629 for the definition of the mollifier ζ^δ).

Observe that $u(x_0,t_0)=\phi(x_0,t_0)=\phi^\epsilon(x_0,t_0)$ and $u(x,t)-\phi^\epsilon(x,t)=(u-\phi)\chi^\epsilon(x)+(u-\widetilde{u})\left(1-\chi^\epsilon(x)\right)\leq 0$ for $(x,t)\in\mathbb{R}^n\times[t_1,t_2]$. Moreover, $\partial_t\phi^\epsilon(x_0,t_0)=\partial_t\phi(x_0,t_0)$, $\partial_{x_i}\phi^\epsilon(x_0,t_0)=\partial_{x^i}\phi(x_0,t_0)$ and $\partial^2_{x^ix^j}\phi^\epsilon(x_0,t_0)=\partial^2_{x^ix^j}\phi(x_0,t_0)$. Note that $\widetilde{u}$ is uniformly bounded, hence $\phi^\epsilon\in C_1(\mathbb{R}^n\times[t_1,t_2])$, therefore we choose $\phi^\epsilon(x,t)$ as the test function in the Definition 3.1 and obtain from (3.21) that

$$(-\partial_t-\mathcal{L}_D+r)\,\phi(x_0,t_0)-I^f\phi^\epsilon(x_0,t_0)\leq 0, \tag{A-8}$$

where $I^f \phi^\epsilon(x_0,t_0)$ is well defined, because one can show $\phi^\epsilon(x,t_0)$ is globally Lipschitz in x as a result of our choice of χ^ϵ. On the other hand,

$$
\begin{aligned}
&|\phi^\epsilon(x_0+y,t_0) - u(x_0+y,t_0)| \\
&\quad \le |\phi(x_0+y,t_0) - u(x_0+y,t_0)|\, \chi^\epsilon(x_0+y) + |\widetilde{u}(x_0+y,t_0) - u(x_0+y,t_0)|\,(1-\chi^\epsilon(x_0+y)) \\
&\quad \le |\phi(x_0+y,t_0) - u(x_0+y,t_0)|\, 1_{\{|y|\le 2\epsilon\}} + \epsilon^2\, 1_{\{|y|\ge\epsilon\}} \\
&\quad \le [\, |\phi(x_0+y,t_0) - \phi(x_0,t_0)| + |u(x_0+y,t_0) - u(x_0,t_0)|\,]\; 1_{\{|y|\le 2\epsilon\}} + \epsilon^2\, 1_{\{|y|\ge\epsilon\}} \\
&\quad \le (\widetilde{L_x} + L_x)\, |y|\, 1_{\{|y|\le 2\epsilon\}} + \epsilon^2\, 1_{\{|y|\ge\epsilon\}},
\end{aligned}
\tag{A-9}
$$

where $\widetilde{L_x} = \max_{|x-x_0|\le 2\epsilon} \partial_x \phi(t_0,x)$ and L_x is the constant in Lemma 2.2. Due to (A-9), (3.1) and (H2), we have

$$
\begin{aligned}
\left|I^f \phi^\epsilon(x_0,t_0) - I^f u(x_0,t_0)\right| &\le (\widetilde{L_x} + L_x) \int_{|y|\le 2\epsilon} |y|\, \nu(dy) + \int_{|y|\ge\epsilon} \epsilon^2\, \nu(dy) \\
&\le (\widetilde{L_x} + L_x) \int_{|y|\le 2\epsilon} |y|\, \nu(dy) + \epsilon \int_{|y|\ge\epsilon} |y|\, \nu(dy) \to 0 \quad \text{as } \epsilon \downarrow 0.
\end{aligned}
\tag{A-10}
$$

Then the statement that u is a viscosity solution of (3.27) follows from combining (A-8) and (A-10). $\square$

Proof of Lemma 5.3. For any $R_0 > 0$, let us consider the following function

$$
w(x,t) = \frac{m}{f(R_0)}\, [f(|x|) + C_1\, t] + v(x,t),
$$

where $f(R) = \frac{R^2}{1+R}$ and the positive constant C_1 will be determined later. It is clear that $f(R)$ is an increasing function on $(0,+\infty)$ and $\lim_{R\to+\infty} f(R) = +\infty$. On the other hand, $|\partial_{x^i} f(|x|)| \le \frac{|x|\,(2+|x|)}{(1+|x|)^2} < 1$ for any $i \le n$. Moreover, one can also check that $\lim_{|x|\to+\infty} \left|\partial^2_{x^i x^j} f(|x|)\right| = 0$ and $\lim_{|x|\to 0} \left|\partial^2_{x^i x^j} f(|x|)\right| = 2\,\delta_{ij}$ for any $i,j \le n$. Therefore both $\partial_{x^i} f(|x|)$ and $\partial^2_{x^i x^j} f(|x|)$ are bounded on $\mathbb{R}^n$. Thanks to these properties, we can find an upper bound for $|If(|x|)|$ as follows:

$$
\begin{aligned}
\left|If(|x|)\right| &= \left| \int_{\mathbb{R}^n} \left[f(|x+y|) - f(|x|) - \sum_{i=1}^n y^i\, \partial_{x^i} f(|x|)\, 1_{\{|y|\le 1\}} \right] \nu(dy) \right| \\
&\le \int_{|y|\le 1} \nu(dy) \int_0^1 dz\, (1-z) \sum_{i,j=1}^n \left|y^i y^j\right| \left|\partial^2_{x^i x^j} f(|x+zy|)\right| + \int_{|y|>1} \nu(dy)\, |f(|x+y|) - f(|x|)| \\
&\le C \left(\int_{|y|\le 1} |y|^2 \nu(dy) + \int_{|y|>1} |y|\, \nu(dy) \right) < +\infty,
\end{aligned}
\tag{A-11}
$$

for some sufficiently large constant $C > 0$. Here the last inequality in (A-11) follows from (2.3) and (H2).

Now, applying the parabolic integro-differential operator to w, we obtain

$$
\begin{aligned}
(\partial_t - \mathcal{L}_D - I + c)\; w(x,t) &\ge (\partial_t - \mathcal{L}_D - I + c) \left[\frac{m}{f(R_0)}\, (f(|x|) + C_1\, t) \right] \\
&= \frac{m}{f(R_0)} \left[C_1 - \sum_{i,j=1}^n a_{ij}\, \partial^2_{x^i x^j} f(|x|) - \sum_{i=1}^n b_i\, \partial_{x^i} f(|x|) + c\, f(|x|) - If(|x|) \right],
\end{aligned}
$$

where the first inequality follows from the assumption that $(\partial_t - \mathcal{L}_D - I + c)\; v(x,t) \ge 0$. We can choose a sufficiently large constant C_1 independent of R_0 such that

$$
(\partial_t - \mathcal{L}_D - I + c)\; w(x,t) > 0, \qquad \text{for } (x,t) \in \mathbb{R}^n \times [0,T]. \tag{A-12}
$$

This is because $\partial^2_{x^i x^j} f(|x|)$, $\partial_{x^i} f(|x|)$ and coefficients a_{ij}, b_i, c are all bounded, moreover $c \geq 0$ and $|I f(|x|)|$ is bounded thanks to (A-11).

On the other hand, $w(x,0) = \frac{m}{f(R_0)} f(|x|) + v(x,0) \geq 0$ thanks to the assumption $v(x,0) \geq 0$. Moreover, when $|x| = R_0$, $w(x,t) = \frac{m}{f(R_0)} \left(f(R_0) + C_1 t\right) + v(x,t) \geq m + v(x,t) \geq 0$ due to the assumption $v(x,t) \geq -m$. Furthermore, when $|x| > R_0$, we also have $w(x,t) \geq m + v(x,t) \geq 0$ since $f(R)$ is an increasing function. Therefore, we claim that $w(x,t) \geq 0$ for $(x,t) \in B_{R_0} \times (0, T_0]$. Indeed, if there are some points $(x,t) \in B_{R_0} \times (0,T_0]$ such that $w(x,t) < 0$, $w(x,t)$ must take its negative minimum at some point $(x_0, t_0) \in B_{R_0} \times (0, T_0]$. Noticing that $w(x,t) \geq 0$ for $|x| \geq R_0$, we have $w(x_0,t_0) \leq w(x,t)$ for all $(x,t) \in \mathbb{R}^n \times (0,T]$. As a result, we obtain $\partial_t w(x_0,t_0) \leq 0$, $\sum_{i=1}^n b_i \, \partial_{x^i} w(x_0,t_0) = 0$ and $\sum_{i,j=1}^n a_{ij} \, \partial^2_{x^i x^j} w(x_0,t_0) \geq 0$ (see e.g. Lemma 1 in [9] pp. 34). Moreover, $I w(x_0,t_0) \geq 0$, since w achieves its minimum at (x_0,t_0) and $\nabla_x w(x_0,t_0) = 0$. Therefore, we have

$$(\partial_t - \mathcal{L}_D - I + r) \; w(x_0,t_0) \leq 0,$$

which contradicts with (A-12).

Now, for any point $(x,t) \in \mathbb{R}^n \times (0,T]$, taking $R_0 \to +\infty$, we have $v(x,t) \geq 0$ since $\lim_{R_0 \to +\infty} f(R_0) = +\infty$. □

Proof of Lemma 5.4. First, thanks to Lemma 5.2, $|v^\epsilon|$ is bounded on $\mathbb{R}^n \times [0,T]$. In the following, we will show it is bounded uniformly in ϵ. It follows from (5.3) (i) that $(\partial_t - \mathcal{L}_D - I + r)\, v^\epsilon = -p_\epsilon(v^\epsilon - g^\epsilon) \geq 0$. Note that $v^\epsilon(x,0) = g^\epsilon(x) \geq 0$ (see (H3')), the first inequality in the statement follows from Lemma 5.3 directly. On the other hand, defining $u = K + 1 - v^\epsilon$, u satisfies

$$(\partial_t - \mathcal{L}_D - I + r)\, u = r(K+1) + p_\epsilon(v^\epsilon - g^\epsilon), \quad (x,t) \in \mathbb{R}^n \times (0,T]. \tag{A-13}$$

It follows from (H3') and (5.3) (ii) that $p_\epsilon(K + 1 - g^\epsilon) = 0$ with $\epsilon \leq \epsilon_0 \leq 1$. Combining with (A-13) and the mean value theorem, we obtain

$$(\partial_t - \mathcal{L}_D - I + r)\, u + p_\epsilon(K+1-g^\epsilon) - p_\epsilon(v^\epsilon - g^\epsilon) = \left[\partial_t - \mathcal{L}_D - I + r + p^{'}_\epsilon(y)\right] u = r\,(K+1) \geq 0, \tag{A-14}$$

for some $y \in \mathbb{R}$. Note that both $K + 1 - g^\epsilon$ and $v^\epsilon - g^\epsilon$ are bounded, $p^{'}_\epsilon$ is bounded in any bounded domain. Therefore, we have that $r + p^{'}_\epsilon(y)$ is bounded and nonnegative (see (5.3) (iv)). Now apply Lemma 5.3 to u and pick $c = r + p^{'}_\epsilon(y)$, we obtain $u(x,t) = K + 1 - v^\epsilon(x,t) \geq 0$ on $\mathbb{R}^n \times [0,T]$. □

Appendix B. Proof of Theorem 5.1

In this Appendix, for notational simplicity, the constant C denotes a generic constant in different places. Moreover, the center x_0 of the ball $B_\rho(x_0)$ will not be noted in the sequel. For any positive integer p, let us first estimate the L_p-norm of the integral term Iv.

Lemma B-1. *If the assumptions of Theorem 5.1 are satisfied, then for any $\eta > 0$, there exists a positive constant C such that*

(B-1)

$$\|Iv\|_{L_p(B_\rho(x_0) \times (s,T))} \leq C\eta^{2-\alpha} \|v\|_{W^{2,1}_p(B_{\rho+\eta}(x_0)\times(s,T))} + C\left(\max_{\mathbb{R}^n \times [s,T]} |v| + \max_{B_{\rho+1}(x_0)\times[s,T]} |\nabla_x v|\right) \cdot \begin{cases} (1+\eta^{1-\alpha}), \; \alpha \neq 1 \\ (1 - \log \eta), \; \alpha = 1 \end{cases}.$$

Proof. Let us break the integral into three parts.

$$\begin{aligned}
|Iv(x,t)| &= \left|\int_{\mathbb{R}^n}\left[v(x+y,t)-v(x,t)-y\cdot\nabla_x v(x,t)1_{\{|y|\le 1\}}\right]\nu(dy)\right| \\
&\le \int_{|y|\le\eta}\nu(dy)\int_0^1 dz(1-z)\sum_{i,j=1}^n\left|y^iy^j\frac{\partial^2}{\partial x^i\partial x^j}v(x+zy,t)\right| \\
&\quad+\int_{\eta<|y|\le 1}\nu(dy)\,|v(x+y,t)-v(x,t)-y\cdot\nabla_x v(x,t)|+\int_{|y|>1}\nu(dy)\,|v(x+y,t)-v(x,t)| \\
&\le \sum_{i,j=1}^n\int_{|y|\le\eta}|y|^2\nu(dy)\int_0^1 dz\left|\frac{\partial^2}{\partial x^i\partial x^j}v(x+zy,t)\right| \\
&\quad+\int_{\eta<|y|\le 1}\nu(dy)\,|v(x+y,t)-v(x,t)-y\cdot\nabla_x v(x,t)|+\int_{|y|>1}\nu(dy)\,|v(x+y,t)-v(x,t)| \\
&\triangleq \sum_{i,j=1}^n I_{i,j}(x,t)+I_2(x,t)+I_3(x,t).
\end{aligned}$$

In the following, we will estimate the L_p-norm of each term respectively.

(B-2)

$$\begin{aligned}
&\|I_{ij}(\cdot,t)\|^p_{L_p(B_\rho)} \\
&=\int_{B_\rho}dx\left[\int_{|y|\le\eta}|y|^2\nu(dy)\int_0^1 dz\left|\partial^2_{x^ix^j}v(x+zy,t)\right|\right]^p\le\int_{B_\rho}dx\int_0^1 dz\left[\int_{|y|\le\eta}\nu(dy)\,|y|^2\left|\partial^2_{x^ix^j}v(x+zy,t)\right|\right]^p \\
&\le M^p\int_{B_\rho}dx\int_0^1 dz\left[\int_{|y|\le\eta}dy\,|y|^{2-n-\alpha}\left|\partial^2_{x^ix^j}v(x+zy,t)\right|\right]^p \\
&\le M^p\int_{B_\rho}dx\int_0^1 dz\left(\int_{|y|\le\eta}dy\,|y|^{2-n-\alpha}\right)^{\frac{p}{q}}\cdot\left(\int_{|y|\le\eta}dy\,|y|^{2-n-\alpha}\left|\partial^2_{x^ix^j}v(x+zy,t)\right|^p\right) \\
&=M^p\left(|S_1(0)|\frac{\eta^{2-\alpha}}{2-\alpha}\right)^{\frac{p}{q}}\cdot\int_0^1 dz\int_{|y|\le\eta}dy\,|y|^{2-n-\alpha}\int_{B_\rho}dx\left|\partial^2_{x^ix^j}v(x+zy,t)\right|^p \\
&\le M^p\left(|S_1(0)|\frac{\eta^{2-\alpha}}{2-\alpha}\right)^{\frac{p}{q}}\cdot\int_0^1 dz\int_{|y|\le\eta}dy\,|y|^{2-n-\alpha}\left\|\partial^2_{x^ix^j}v(\cdot,t)\right\|^p_{L_p(B_{\rho+\eta})} \\
&=M^p\left(|S_1(0)|\frac{\eta^{2-\alpha}}{2-\alpha}\right)^{p}\cdot\left\|\partial^2_{x^ix^j}v(\cdot,t)\right\|^p_{L_p(B_{\rho+\eta})}.
\end{aligned}$$

Here the first inequality follows from Fubini's Theorem and Jensen's inequality with respect to the Lebesgue measure dz. The assumption (H5) is used in the second inequality. The third inequality follows from Hölder inequality with $1/p+1/q=1$. In the second equality, $|S_1(0)|$ is the surface area of the unit ball in $\mathbb{R}^n$. Note that $x+zy\in B_{\rho+\eta}$ when $x\in B_\rho$, $z\in(0,1)$ and $|y|\le\eta$, the fourth inequality follows.

For I_2 and I_3, noting that $x+y\in B_{\rho+1}$ when $x\in B_\rho$ and $|y|\le 1$, we have

$$\|I_2(\cdot,t)\|_{L_p(B_\rho)}\le C\cdot\max_{B_{\rho+1}\times[s,T]}|\nabla_x v|\cdot\begin{cases}(1+\eta^{1-\alpha}),\ \alpha\ne 1\\ (1-\log\eta),\ \alpha=1\end{cases}\quad\text{and}\tag{B-3}$$

$$\|I_3(\cdot,t)\|_{L_p(B_\rho)}\le C\cdot\max_{\mathbb{R}^n\times[s,T]}|v|\cdot\int_{|y|>1}\nu(dy).\tag{B-4}$$

Combining (B-2) - (B-4), (B-1) follows from noting that $\|Iv\|_{L_p(B_\rho\times(s,T))}\triangleq\left[\int_s^T\|Iv(\cdot,t)\|_{L_p(B_\rho)}\,dt\right]^{1/p}$ and $\|\partial^2_{x^ix^j}v\|_{L_p(B_{\rho+\eta}\times(s,T))}\leq\|v\|_{W^{2,1}_p(B_{\rho+\eta}\times(s,T))}$ (see Definition 2.2). □

In (B-1), when $\alpha\in[0,1)$ (finite variation jumps), the factors of η in both terms on the right-hand-side converge to 0 as $\eta\to 0$. Therefore, the L_p-norm of Iv on the domain $B_\rho(x_0)\times(s,T)$ essentially only depends on $\max_{\mathbb{R}^n\times[s,T]}|v|$ and $\max_{B_{\rho+1}\times[s,T]}|\nabla_x v|$. This can also be confirmed by working with the reduced integral form $I^f v$ in (3.4).

On the contrary, when $\alpha\in[1,2)$ (infinite variation jumps), the factor $1+\eta^{1-\alpha}$ (or $1-\log\eta$) in (B-1) will blow up as $\eta\to 0$ (a similar phenomenon was also observed in Lemma 1.1 of [2] pp.206 for L_p-norm on $\mathbb{R}^n$). Therefore, it is important to note that the L_p-norm of Iv on the domain $B_\rho(x_0)\times(s,T)$ actually depends on $W^{2,1}_p$-norm of v on a larger domain $B_{\rho+\eta}(x_0)\times(s,T)$. Because of the expansion of the domain, instead of using the boundary estimate in Theorem 9.1 in [17] pp. 342, we will use the interior estimate technique in Theorem 10.1 in [17] pp. 351 to prove Theorem 5.1 in the following.

Proof of Theorem 5.1. Let us choose a cut-off function $\zeta^\delta(x,t)$ such that

$$\zeta^\delta(x,t)=\begin{cases}1 & (x,t)\in B_\rho\times(\delta,T)\\ 0 & (x,t)\in\mathbb{R}^n\times(0,T)\setminus B_{\rho+\frac{\delta}{4}}\times(\frac{\delta}{2},T)\end{cases}$$

Here the constant $\delta\in(0,s)$ will be determined later. This cut-off function can be chosen such that

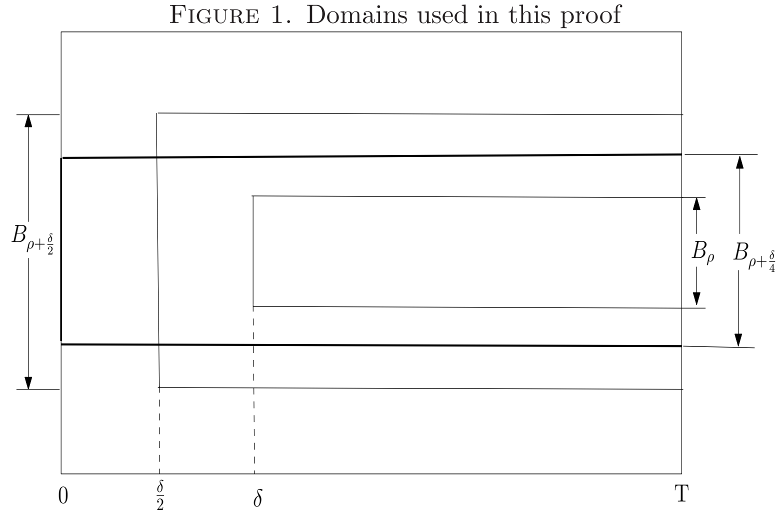
$$\left|\partial_{x^i}\zeta^\delta\right|\leq\frac{C_1}{\delta},\quad\left|\partial^2_{x^ix^j}\zeta^\delta\right|\leq\frac{C_2}{\delta^2}\quad\text{and}\quad\left|\partial_t\zeta^\delta\right|\leq\frac{C_3}{\delta},\tag{B-5}$$

for $i,j\leq n$ and some constants C_1,C_2 and C_3. Please see Figure 1 for the domains used in this proof.

Defining $u(x,t)=\zeta^\delta(x,t)v(x,t)$, it satisfies

$$\begin{aligned}&\left(\partial_t-\mathcal{L}_D+r\right)u(x,t)=\zeta^\delta\cdot Iv(x,t)+\zeta^\delta\cdot f(x,t)+h(x,t),\quad(x,t)\in B_{\rho+\frac{\delta}{4}}\times(0,T),\\&u(x,t)=0,\quad(x,t)\in\partial B_{\rho+\frac{\delta}{4}}\times(0,T),\\&u(x,0)=0,\quad x\in\overline{B_{\rho+\frac{\delta}{4}}},\end{aligned}$$

FIGURE 1. Domains used in this proof



in which $h(x,t) \triangleq \partial_t \zeta^\delta \cdot v - \sum_{i,j=1}^n a_{ij} \left(\partial^2_{x^i x^j} \zeta^\delta \cdot v + 2\, \partial_{x^i} \zeta^\delta \cdot \partial_{x^j} v \right) - \sum_{i=1}^n b_i \cdot \partial_{x^i} \zeta^\delta \cdot v$. Appealing to Theorem 9.1 in [17] pp.341, there exists a constant C such that

$$\begin{aligned} \|u\|_{W_p^{2,1}(B_{\rho+\frac{\delta}{4}}\times(0,T))} \leq & C \Bigg[\left\|\zeta^\delta \cdot Iv\right\|_{L_p} + \left\|\zeta^\delta \cdot f\right\|_{L_p} + \left\|\partial_t \zeta^\delta \cdot v\right\|_{L_p} + \left\| \sum_{i,j=1}^n a_{ij}\, \partial^2_{x^i x^j} \zeta^\delta \cdot v \right\|_{L_p} \\ & + \left\| \sum_{i,j=1}^n 2\, a_{ij}\, \partial_{x^i} \zeta^\delta \cdot \partial_{x^j} v \right\|_{L_p} + \left\| \sum_{i=1}^n b_i \cdot \partial_{x^i} \zeta^\delta \cdot v \right\|_{L_p} \Bigg], \end{aligned} \tag{B-6}$$

in which all L_p-norms on the right-hand-side represent $L_p \left(B_{\rho+\frac{\delta}{4}} \times (0,T) \right)$.

In the following, we will estimate the terms on the right-hand-side of (B-6) respectively.

$$\begin{aligned} \left\|\zeta^\delta \cdot Iv\right\|_{L_p(B_{\rho+\frac{\delta}{4}}\times(0,T))} \leq & \|Iv\|_{L_p(B_{\rho+\frac{\delta}{4}}\times(\frac{\delta}{2},T))} \\ \leq & C \left(\frac{\delta}{4}\right)^{2-\alpha} \|v\|_{W_p^{2,1}(B_{\rho+\frac{\delta}{2}}\times(\frac{\delta}{2},T))} + C \left(1 + \left(\frac{\delta}{4}\right)^{1-\alpha} \right) \left[\max_{\mathbb{R}^n\times[0,T]} |v| + \max_{B_{\rho+\frac{\delta}{4}+1}\times[0,T]} |\nabla_x v| \right]. \end{aligned} \tag{B-7}$$

Here the first inequality follows from the choice of the cut-off function ζ^δ, the second inequality follows from Lemma B-1 for $\alpha \neq 1$ case by picking $\eta = \frac{\delta}{4}$ and $s = \frac{\delta}{2}$. When $\alpha = 1$, we also have an estimate similar to (B-7). On the other hand, we have

$$\left\|\zeta^\delta \cdot f\right\|_{L_p(B_{\rho+\frac{\delta}{4}}\times(0,T))} \leq \|f\|_{L_p(B_{\rho+\frac{\delta}{4}}\times(\frac{\delta}{2},T))}. \tag{B-8}$$

Moreover, we obtain from (B-5) that

$$\begin{aligned} \left\|\partial_t \zeta^\delta \cdot v\right\|_{L_p(B_{\rho+\frac{\delta}{4}}\times(0,T))} \leq & \max_{\mathbb{R}^n\times[0,T]} |v| \cdot \left\|\partial_t \zeta^\delta\right\|_{L_p(B_{\rho+\frac{\delta}{4}}\times(0,T))} \\ \leq & \max_{\mathbb{R}^n\times[0,T]} |v| \left(\int_{B_{\rho+\frac{\delta}{4}}\times(\frac{\delta}{2},T)\setminus B_\rho\times(\delta,T)} dt\, dx\, \frac{C_3^p}{\delta^p} \right)^{\frac{1}{p}} \\ \leq & C \max_{\mathbb{R}^n\times[0,T]} |v| \cdot \delta^{\frac{1-p}{p}}. \end{aligned} \tag{B-9}$$

Similarly, thanks to (H7”), we also have

$$\left\| \textstyle\sum_{i,j=1}^n a_{ij}\, \partial^2_{x^i x^j} \zeta^\delta \cdot v \right\|_{L_p(B_{\rho+\frac{\delta}{4}}\times(0,T))} \leq C \max_{\mathbb{R}^n\times[0,T]} |v| \cdot \delta^{\frac{1-2p}{p}}, \tag{B-10}$$

$$\left\| \textstyle\sum_{i,j=1}^n 2\, a_{ij}\, \partial_{x^i} \zeta^\delta \cdot \partial_{x^j} v \right\|_{L_p(B_{\rho+\frac{\delta}{4}}\times(0,T))} \leq C \max_{B_{\rho+\frac{\delta}{4}}\times[0,T]} |\nabla_x v| \cdot \delta^{\frac{1-p}{p}} \quad \text{and} \tag{B-11}$$

$$\left\| \textstyle\sum_{i=1}^n b_i \cdot \partial_{x^i} \zeta^\delta \cdot v \right\|_{L_p(B_{\rho+\frac{\delta}{4}}\times(0,T))} \leq C \max_{\mathbb{R}^n\times[0,T]} |v| \cdot \delta^{\frac{1-p}{p}}. \tag{B-12}$$

Plugging (B-7) - (B-12) into (B-6) and noticing the choice of the cut-off function ζ^δ, we obtain

(B-13)

$$\begin{aligned}\|v\|_{w_p^{2,1}(B_\rho\times(\delta,T))} \leq \|u\|_{w_p^{2,1}(B_{\rho+\frac{\delta}{4}}\times(0,T))}\\ \leq C\left(\frac{\delta}{4}\right)^{2-\alpha}\|v\|_{W_p^{2,1}(B_{\rho+\frac{\delta}{2}}\times(\frac{\delta}{2},T))} + C\left[1+\delta^{1-\alpha}+\delta^{\frac{1-p}{p}}+\delta^{\frac{1-2p}{p}}\right]\cdot\left[\max_{\mathbb{R}^n\times[0,T]}|v| + \max_{B_{\rho+\frac{\delta}{4}+1}\times[0,T]}|\nabla_x v|\right]\\ + \|f\|_{L_p(B_{\rho+\frac{\delta}{4}}\times(\frac{\delta}{2},T))}.\end{aligned}$$

Multiplying δ^2 on both hand side of (B-13) and defining

$$K(\delta) = C\left[\delta^2+\delta^{3-\alpha}+\delta^{\frac{1+p}{p}}+\delta^{\frac{1}{p}}\right]\cdot\left[\max_{\mathbb{R}^n\times[0,T]}|v| + \max_{B_{\rho+\frac{\delta}{4}+1}\times[0,T]}|\nabla_x v|\right] + \delta^2\|f\|_{L_p(B_{\rho+\frac{\delta}{4}}\times(\frac{\delta}{2},T))},$$

we obtain

$$\delta^2\|v\|_{w_p^{2,1}(B_\rho\times(\delta,T))} \leq 4C\left(\frac{\delta}{4}\right)^{2-\alpha}\cdot\left(\frac{\delta}{2}\right)^2\|v\|_{w_p^{2,1}(B_{\rho+\frac{\delta}{2}}\times(\frac{\delta}{2},T))} + K(\delta). \tag{B-14}$$

Let $F(\tau) \triangleq \tau^2\|v\|_{w_p^{2,1}(B_{\rho+\delta-\tau}\times(\tau,T))}$. The inequality (B-14) gives us the following recursive inequality

$$F(\delta) \leq 4C\left(\frac{\delta}{4}\right)^{2-\alpha}F(\delta/2) + K(\delta). \tag{B-15}$$

Since $\alpha < 2$, we can choose sufficiently small δ such that $4C(\delta/4)^{2-\alpha} \leq \frac{1}{2}$. Therefore, we have from (B-15) that

$$F(\delta) \leq \frac{1}{2}F(\delta/2) + K(\delta). \tag{B-16}$$

On the other hand, thanks to the assumption $v \in W^{2,1}_{p,loc}(\mathbb{R}^n\times(0,T))$, $F(\delta)$ is finite for any $\delta \in (0,\delta_0)$. Iterating the recursive inequality (B-16) gives us

$$F(\delta) \leq \sum_{i=0}^\infty \frac{1}{2^i}K\left(\frac{\delta}{2^i}\right) \leq \sum_{i=0}^\infty \frac{1}{2^i}K(\delta) = 2\,K(\delta),$$

where the second inequality follows from noticing that $K(\delta)$ is increasing in δ. Therefore, it follows from the definitions of $F(\delta)$ and $K(\delta)$ that

$$\begin{aligned}\|v\|_{W_p^{2,1}(B_\rho\times(s,T))} &\leq \|v\|_{W_p^{2,1}(B_\rho\times(\delta,T))}\\ &\leq 2\,C\left[1+\delta^{1-\alpha}+\delta^{\frac{1-p}{p}}+\delta^{\frac{1-2p}{p}}\right]\cdot\left[\max_{\mathbb{R}^n\times[0,T]}|v| + \max_{B_{\rho+\frac{\delta}{4}+1}\times[0,T]}|\nabla_x v|\right] + \|f\|_{L_p(B_{\rho+\frac{\delta}{4}}\times(\frac{\delta}{2},T))}\\ &\leq C_\delta\left[\max_{\mathbb{R}^n\times[0,T]}|v| + \max_{B_{\rho+\frac{\delta}{4}+1}\times[0,T]}|\nabla_x v| + \|f\|_{L_p(B_{\rho+\frac{\delta}{4}}\times(\frac{\delta}{2},T))}\right].\end{aligned}$$

□

References

[1] E. Bayraktar. A proof of the smoothness of the finite time horizon American put option for jump diffusions. *SIAM Journal on Control and Optimization*, 48(2):551–572, 2009.

[2] A. Bensoussan and J.-L. Lions. *Impulse control and quasivariational inequalities*. Gauthier-Villars, Montrouge, 1984.

[3] S. I. Boyarchenko and S. Z. Levendorskiĭ. *Non-Gaussian Merton-Black-Scholes theory*, volume 9 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Inc., River Edge, NJ, 2002.

[4] P. Carr, H. Geman, D. B. Madan, and M. Yor. Stochastic volatility for Lévy processes. *Mathematical Finance*, 13(3):345–382, 2003.

[5] R. Cont and P. Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.

[6] M. Crandall, H. Ishii, and P. L. Lions. User's guide to viscosity solutions of second order partial differential equations. *American Mathematical Society. Bulletin. New Series*, 27(1):1–67, 1992.

[7] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[8] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer, New York, second edition, 2006.

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

[10] A. Friedman. *Variational Principles and Free-boundary Problems*. John Wiley & Sons Inc., New York, 1982.

[11] M. G. Garroni and J.-L. Menaldi. *Green functions for second order parabolic integro-differential problems*, volume 275 of *Pitman Research Notes in Mathematics Series*. Longman Scientific & Technical, Harlow, 1992.

[12] M. G. Garroni and J.-L. Menaldi. *Second order elliptic integro-differential problems*, volume 430 of *Chapman & Hall/CRC Research Notes in Mathematics*. Chapman & Hall/CRC, Boca Raton, FL, 2002.

[13] P. Gimbert and P.-L. Lions. On the regularity of solutions of the Dirichlet problem for elliptic integro-differential operators: a counterexample. *Ricerche di Matematica*, 34(2):283–288, 1985.

[14] P. Jaillet, D. Lamberton, and B. Lapeyre. Variational inequalities and the pricing of American options. *Acta Applicandae Mathematicae*, 21:263–289, 1990.

[15] I. Karatzas. A tutorial introduction to stochastic analysis and its applications, 1998. Available at `http://www.math.columbia.edu/~ik/notes.html`.

[16] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer - Verlag, New York, second edition, 1991.

[17] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Uralchva. *Linear and Quasi-linear Equations of Parabolic Type*. American Mathematical Society, Providence, Rhode Island, 1968.

[18] D. Lamberton and M. Mikou. The critical price for the American put in an exponential Lévy model. *Finance and Stochastics*, 12(4):561–581, 2008.

[19] P. L. Lions. A remark on Bony maximum principle. *Proceedings of the American Mathematical Society*, 88(3):503–508, 1983.

[20] H. Pham. Optimal stopping, free boundary, and American option in a jump-diffusion model. *Applied Mathematics and Optimization*, 35(2):145–164, 1997.

[21] H. Pham. Optimal stopping of controlled jump diffusion processes: a viscosity solution approach. *Journal of Mathematical Systems, Estimation, and Control*, 8(1):1–27, 1998.

[22] K. Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

[23] L. Silvestre. Regularity of the obstacle problem for a fractional power of the laplace operator. *Communications on Pure and Applied Mathematics*, 60(1):67–112, 2006.

[24] H. M. Soner. Optimal control with state-space constraint. II. *SIAM Journal on Control and Optimization*, 24(6):1110–1122, 1986.

[25] C. Yang, L. Jiang, and B. Bian. Free boundary and American options in a jump-diffusion model. *European Journal of Applied Mathematics*, 17(1):95–127, 2006.

[26] X. L. Zhang. *Méthodes Numériques pour le Calcul des Options Américaine dans des Modèles de diffusion avec sauts*. PhD thesis, l'Ecole Nationale des Ponts et Chaussées, Paris, 1994.

(Erhan Bayraktar) Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA; e-mail:erhan@umich.edu

(Hao Xing) Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA; e-mail:haoxing@umich.edu