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 of obstacle problems on an unbounded domain with finite/infinite variation jumps is in $W^{2,1}_{p,loc}$. As a consequence, the smooth-fit property holds.

1. Introduction

On a probability space $(\Omega, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$, consider a one-dimensional jump diffusion process $X = \{X_t; t \geq 0\}$ whose dynamics is governed by the following stochastic differential equation:

$$dX_t = b(X_{t-}, t)\, dt + \sigma(X_{t-}, t)\, dW_t + dJ_t, \tag{1.1}$$

in which $W = \{W_t; t \geq 0\}$ is a 1-dimensional Wiener process, $J = \{J_t; t \geq 0\}$ is a pure jump Lévy process, independent of the Wiener process, with its Lévy measure denoted by ν. This paper studies 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

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

in which $\mathcal{T}_{t,T}$ is the set of all stopping times valued between t and T. A specific example of such an optimal stopping problem is the American option pricing problem, where X model the logarithm of the stock price process and g represents the pay-off function.

The value function u is expected to satisfy a variational inequality with a nonlocal integral term (see e.g. Chapter 3 of [7]). When the diffusion component of X may vanish, different concepts of solutions were employed to characterize the value function. Pham used the notion of viscosity solution in [21]. Also see [3], [4] for more recent results in this direction. Lamberton and Mikou worked with Lévy processes and showed in [18] that the value function can be understood in the distribution sense.

When the diffusion component in X is nondegenerate, the value function is expected to have higher degree of regularity. Sections 1-3 in Chapter 3 of [7] and [15] analyzed the Cauchy problems for second order partial integro-differential equations and showed the existence and uniqueness of solutions in both Sobolev and Hölder spaces. Also see [19]. The intuition is that the diffusions component dominates the contribution from jumps in determining the regularity of solutions, no matter what the variation of the jumps are. However this intuition is only a folklore theorem for obstacle problems. There are some limited results available whose assumptions on

Key words and phrases. Optimal stopping, variational inequality, Lévy processes, regularity of the value function, smooth fit principle, Sobolev spaces.

The first author is supported in part by the National Science Foundation under an applied mathematics research grant and a Career grant, DMS-0906257 and DMS-0955463, respectively, and in part by the Susan M. Smith Professorship.

obstacles, domains, and the structure of the jumps may not be appropriate for financial applications. For example, Bensoussan and Lions analyzed an obstacle problem for jump diffusions where jumps may have finite/infinite activity with finite/infinite variation; see in Theorem 3.2 in [7] on pp. 234. However, their assumption on the obstacle may not be satisfied by option payoffs. In the mathematical finance literature, when irregular obstacles are considered, the jumps are usually restricted to finite activity or infinite activity with finite variation cases. Zhang studied in [25] an obstacle problem for a jump diffusion with finite active jumps. Also see [20], [24], [5], and [6] for further developments. More recently, Davis et al. in [11], generalizing the results in [16] for the diffusion case, analyzed an impulse control problems for jump diffusions with infinite activity but finite variation jumps. A regularity result which treats obstacle problems with irregular obstacles and infinite variation jumps has been missing in the literature.

In this paper, we consider jump diffusions with finite/infinite activity and finite/infinite variational jumps. We show in Theorem 2.5 that the value function of an obstacle problem solves a variational inequality for almost all points in the domain, and it is an element in $W^{2,1}_{p,loc}$ (see later this section for the definition of this Sobolev space). This regularity result directly implies that the smooth fit property holds and the value function is $C^{2,1}$ inside the continuation region. These results confirm the intuition that the nondegenerate diffusions components dominate any type of Lévy jumps in determining the regularity of the value function for obstacle problems.

The remainder of the paper is organized as follows. After introducing notation at the end of this section, main results are presented in Section 2. Regularity properties of the infinitesimal generator of X are studied in Section 3. Then main results are proved in Section 4.

1.1. **Notation.** For a given open interval $D = (\ell, r)$ with $-\infty < \ell < r < \infty$, let us define the δ-neighborhood of D as $D^\delta := (\ell - \delta, r + \delta)$ for $\delta > 0$. We will also denote $D_s := D \times (0, s)$, $D^\delta_s := D^\delta \times (0, s)$ for any $s > 0$, $E_s := \mathbb{R} \times [0, s]$, and $\overline{\cdot}$ the closure of the indicated set. Let us recall definitions of Sobolev spaces and Hölder spaces in what follows; see pp. 5-7 for further details.

Definition 1.1. $C^{2,1}(D_s)$ denotes the class of continuous functions on D_s with continuous classical time and spatial derivatives up to the second order.

For any positive integer $p \geq 1$, $W^{2,1}_p(D_s)$ is the space of functions $v \in L_p(D_s)$ with generalized derivatives $\partial_t v$, $\partial_x v$, $\partial^2_{xx} v$, and a finite norm $\|v\|_{W^{2,1}_p(D_s)} := \|\partial_t v\|_{L_p(D_s)} + \|\partial_x v\|_{L_p(D_s)} + \|\partial^2_{xx} v\|_{L_p(D_s)}$. The space $W^{2,1}_{p,\,loc}(D_s)$ consists of functions whose $W^{2,1}_p$-norm is finite on any compact subsets of D_s.

For any positive nonintegral real number α, $H^{\alpha,\alpha/2}\left(\overline{D_s}\right)$ is the space of functions v that are continuous in $\overline{D_s}$ with continuous classical derivatives $\partial^r_t \partial^s_x v$ for $2r + s < \alpha$, and have finite norm $\|v\|^{(\alpha)}_{\overline{D_s}} := |v|^{(\alpha)}_x + |v|^{(\alpha/2)}_t + \sum_{2r+s\leq[\alpha]} \|\partial^r_t \partial^s_x v\|^{(0)}$, in which $\|v\|^{(0)} = max_{D_s}|v|$, $|v|^{(\alpha)}_x = \sum_{2r+s=[\alpha]} \sup_{|x-x'|\leq\rho_0} \frac{|\partial^r_t\partial^s_x v(x,t) - \partial^r_t\partial^s_x v(x',t)|}{|x-x'|^{\alpha-[\alpha]}}$, and $|v|^{(\alpha/2)}_t = \sum_{\alpha-2<2r+s<\alpha} \sup_{|t-t'|\leq\rho_0} \frac{|\partial^r_t\partial^s_x v(x,t) - \partial^r_t\partial^s_x v(x,t')|}{|t-t'|^{(\alpha-2r-s)/2}}$, for a constant ρ_0. The space $H^\alpha\left(\overline{\Omega}\right)$ is the Hölder space when only spatial variables are considered.

2. Main results

2.1. **Model.** Let us first specify the jump diffusion X in (1.1). We assume that the drift and volatility of X and the discounting factor r satisfy the following assumption:

Assumption 2.1. Let $a := \frac{1}{2}\sigma^2$. Coefficients $a, b, r \in H^{\ell, \frac{\ell}{2}}(E_T)$ for some $\ell > 1$, $r(x,t) \geq 0$, moreover there exist a strictly positive constant λ such that $a(x,t) \geq \lambda$ for all $(x,t) \in E_T$.

Under above assumption, both b and σ are Lipschitz continuous on E_T. For the pure jump component J in (1.1), we assume that it is a Lévy process with the Lévy measure ν, which is a positive Radon measure on $\mathbb{R}$ with a possible singularity at 0. This measure ν satisfies $\int_{\mathbb{R}} y^2 \wedge 1\, \nu(dy) < \infty$, see [22]. These assumptions on coefficients and the jump component ensure that (1.1) admits a unique strong solution, which we denote by X. This jump diffusion process X is said to have *finite activity*, if ν is a finite measure on $\mathbb{R}$, otherwise it is said to have *infinite activity*.

In Assumption 2.1, X is assumed to have a nondegenerate diffusion component. Meanwhile, it could also have infinitely active jump component. The coexistence of diffusions and infinite activity jumps is motivated by recent studies of Aït-Sahalia and Jacod in [1] and [2].

Among all possible Lévy measures, we consider the following large subclass in this paper:

Assumption 2.2. The Lévy measure satisfies $\int_{|y|>1} |y|\nu(dy) < \infty$. Moreover it has a density, which we denote by ρ, and this density satisfies $\rho(y) \leq \frac{M}{|y|^{1+\alpha}}$ on $|y| \leq 1$, for some constants $M > 0$ and $\alpha \in [0, 2)$.

The interval $|y| \leq 1$ can be replaced by any other neighborhood of 0 in our analysis, here $|y| \leq 1$ is chosen to ease the presentation.

Remark 2.3. Virtually all Lévy processes used in the financial modeling satisfy above assumption. For jump diffusions models, ν is a finite measure as in Merton's and Kou's model. For subordinated Brownian motions, ρ has a power singularity $1/|y|^{1+2\beta}$ at $y = 0$, with $0 \leq \beta < 1$; see (4.25) in [10]. In particular, this class contains Variance Gamma and Normal Inverse Gaussian where $\beta = 0$ or $1/2$ respectively. For generalized tempered stable processes (see Remark 4.1 in [10]), $\rho(y) = \frac{C_-}{|y|^{1+\alpha_-}} e^{-\lambda_-|y|} 1_{\{y<0\}} + \frac{C_+}{|y|^{1+\alpha_+}} e^{-\lambda_+ y} 1_{\{y>0\}}$, with $\alpha_-, \alpha_+ < 2$ and $\lambda_-, \lambda_+ > 0$. In particular, CGMY processes in [9] and regular Lévy processes of exponential type (RLPE) in [8] are special examples of this class.

Having introduced the jump diffusion process X, let us discuss the problem (OS). We assume that the payoff function g satisfies the following assumption. A typical example, where this assumption holds, is the American put option payoff $g(x) = (K - e^x)_+$ for some $K \in \mathbb{R}_+$.

Assumption 2.4. The payoff function g is a positive bounded Lipschitz continuous on $\mathbb{R}$, i.e., there exists positive constants K and L such that $0 \leq g(x) \leq K$ for any $x \in \mathbb{R}$ and $|g(x) - g(y)| \leq L|x - y|$ for any $x, y \in \mathbb{R}$. Moreover g satisfies $\partial^2_{xx} g \geq -J$ for some positive constant J in the distributional sense, i.e., $\int_{\mathbb{R}} g(x) \partial^2_{xx}\phi(x)\, dx \geq -J \int_{\mathbb{R}} \phi(x)\, dx$ for any compactly supported smooth function ϕ on $\mathbb{R}$.

For the problem (OS), we define its continuation region $\mathcal{C}$ and stopping region $\mathcal{D}$ as usual:

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

2.2. **Main regularity results.** Intuitively, one can expect from Itô's lemma that the value function u satisfies the following *variational inequality*:

$$\begin{aligned} &\min\{(-\partial_t - \mathcal{L} + r)\, u, u - g\} = 0, && (x,t) \in \mathbb{R} \times [0,T), \\ &u(x,T) = g(x), && x \in \mathbb{R}. \end{aligned} \tag{2.1}$$

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

$$\mathcal{L}\phi := \mathcal{L}_D \phi + I\phi, \tag{2.2}$$

where $\mathcal{L}_D \phi(x,t) := a(x,t)\, \partial^2_{xx} + b(x,t)\, \partial_x$ and the integral term

$$I\phi(x,t) := \int_{\mathbb{R}^n} \left[\phi(x+y,t) - \phi(x,t) - y\partial_x \phi(x,t)\, 1_{\{|y|\leq 1\}}\right] \nu(dy). \tag{2.3}$$

However, one does not know a priori whether u is sufficiently regular so that it solves (2.1) in the classical sense where all previous differential and integral terms are well defined for u. It is not even clear whether Iu is well defined in the classical sense.

When $\phi(\cdot,t)$ is Lipschitz continuous on $\mathbb{R}$ with a Lipschitz continuous derivatibe $\partial_x\phi(\cdot,t)$ in a neighborhood of x, it can be seen that $I\phi(x,t)$ is well defined in the classical sense. Indeed, $I\phi(x,t) = I_\epsilon\phi(x,t) + I^\epsilon\phi(x,t) < \infty$, where

$$
\begin{aligned}
I^\epsilon\phi(x,t) &= \int_{|y|>\epsilon} [\phi(x+y,t)-\phi(x,t)]\,\nu(dy) - \partial_x\phi(x,t)\cdot\int_{\epsilon<|y|\le 1} y\,\nu(dy)\\
&\le C\int_{|y|>\epsilon}|y|\nu(dy) - \partial_x\phi(x,t)\cdot\int_{\epsilon<|y|\le 1} y\,\nu(dy),\\
I_\epsilon\phi(x,t) &= \int_{|y|\le\epsilon}[\phi(x+y,t)-\phi(x,t)-y\partial_x\phi(x,t)]\,\nu(dy) \qquad (2.4)\\
&= \int_{|y|\le\epsilon} y\left(\partial_x\phi(z,t)-\partial_x\phi(x,t)\right)\nu(dy) \le \int_{|y|\le\epsilon}\widetilde{C}\,y^2\nu(dy).
\end{aligned}
$$

Here, C is the Lipschitz constant of $\phi(\cdot,t)$ on $\mathbb{R}$, $z\in\mathbb{R}$ satisfies $|z-x|<|y|$, the second inequality in (2.4) follows from the mean value theorem, and D is the Lipschitz constant of $\partial_x\phi(\cdot,t)$ in a neighborhood of x. However, the value function u, in general, does not have these regularity properties mentioned above. We only know from Lemma 3.1 in [21] that u is Lipschitz continuous in x and $1/2-$Hölder continuous in t. However, we will see that the integral term Iu is well defined in the classical sense, see Lemma 3.2. In fact, more is true as we show in the next theorem, which is the main result of the paper.

Theorem 2.5. *Let Assumptions 2.1, 2.2, and 2.4 hold. Then $u \in W^{2,1}_{p,loc}(\mathbb{R}\times(0,T))$ for any integer $p\in(1,\infty)$. Moreover, u solves* (2.1) *for almost every point in E_T.*

The following corollary is of special interest for the American option problem.

Corollary 2.6. *Under the assumptions of Theorem 2.5,*

(i) *$\partial_x u \in C(\mathbb{R}\times[0,T))$, i.e., the smooth-fit holds;*
(ii) *$u\in C^{2,1}$ in the region where $u>g$.*

Remark 2.7. When jumps of X have finite variation, i.e., $\int_{\mathbb{R}}|y|\wedge 1\,\nu(dy)<\infty$, the proof of the main result is much simpler. This is because, when jumps of X have finite variation, the infinitesimal generator $\mathcal{L}$ can be rewritten such that its integral component has a reduced form. For any test function ϕ that is Lipschitz continuous in its first variable, $\mathcal{L}\phi$ can be decomposed as $\mathcal{L}\phi = \mathcal{L}^f_D\phi + I^f\phi$, in which $\mathcal{L}^f_D\phi = a\,\partial^2_{xx}\phi + [b-\int_{|y|\le 1} y\nu(dy)]\,\partial_x\phi$ and

$$
I^f\phi(x,t) := \int_{\mathbb{R}^n}[\phi(x+y,t)-\phi(x,t)]\;\nu(dy). \qquad (2.5)
$$

The previous integral is clearly well defined. This is because $|I^f\phi(x,t)| \le \int_{\mathbb{R}}|\phi(x+y,t)-\phi(x,t)|\;\nu(dy) \le C\int_{\mathbb{R}}|y|\,\nu(dy) < +\infty$, where C is the Lipschitz constant of ϕ. Moreover, $I^f\phi$ is also Hölder continuous; see Lemma 3.1 below. Since the value function u is known to be Lipschitz continuous in its first variable (see Lemma 3.1 in [21]), Iu is already well defined and Hölder continuous. Therefore, Iu can be treated as a driving term in (2.1). However, this simplification cannot be applied when jumps of X have infinite variation, i.e., $\int_{\mathbb{R}}|y|\wedge 1\nu(dy)=\infty$.

3. Regularity properties of the integro-differential operator

3.1. **The integral operator.** The integral operator I has two basic features. First, ν has a singularity at $y=0$. As a result, I maps functions with certain degree of regularity to functions with less regularity. This is contrast to the case in which ν is a finite measure. In that case $\int_{\mathbb{R}} \phi(x+y,t)\nu(dy)$ is already well defined, for any ϕ with at most linear growth, and this integral has the same regularity as ϕ; see [24]. Second, I is a nonlocal operator. Therefore, regularity of $I\phi$ on a given interval D depends on ϕ outside D. In this subsection, we shall study these two features in detail and analyze regularity of $I\phi$ when ϕ is either a function in certain Hölder or Sobolev spaces.

Consider I as an operator between Hölder spaces. When jumps of X have finite variation, we can work with the reduced integral operator I^f in (2.5). It has the following regularity property.

Lemma 3.1. *Let Assumption 2.2 hold with $0 \le \alpha < 1$. For any ϕ which is Lipschitz continuous in its first variable and 1/2-Hölder continuous in its second variable,*

$$\begin{aligned} &I^f\phi \in H^{1-\gamma,\frac{1-\gamma}{2}}(\overline{D_s}) \quad \forall \gamma \in (0,1), && \text{when } \alpha = 0;\\ &I^f\phi \in H^{1-\alpha,\frac{1-\alpha}{2}}(\overline{D_s}), && \text{when } 0<\alpha<1. \end{aligned}$$

However when jumps of X have infinite variation, the integral term $I^f\phi$ is no longer well defined for Lipschitz continuous functions. Hence we work with $\mathcal{L}$ and its integral part I in the form of (2.2) and (2.3). We will see that if we choose an appropriate test function ϕ, $I\phi$ is still well defined and Hölder continuous in both its variables.

Lemma 3.2. *Let Assumption 2.2 hold with $\alpha \in [1,2)$.*

(i) *Suppose that ϕ satisfies $|\phi(x_1,t_1) - \phi(x_2,t_2)| \le L(|x_1 - x_2| + |t_1 - t_2|^{\frac{1}{2}})$ for some $L > 0$ and any $(x_1,t_1),(x_2,t_2) \in E_s$. If, moreover, $\phi \in H^{\beta,\frac{\beta}{2}}(\overline{D_s^1})$ for some $\beta \in (\alpha,2)$, then $I\phi \in H^{\frac{\beta-\alpha}{2},\frac{\beta-\alpha}{4}}(\overline{D_s})$ and*

$$\|I\phi\|_{\overline{D_s}}^{\left(\frac{\beta-\alpha}{2}\right)} \le C\left(L + \|\phi\|_{\overline{D_s^1}}^{(\beta)}\right), \tag{3.1}$$

for a positive constant C that depends on D, α, and β.

(ii) *If $\phi \in H^{\beta,\frac{\beta}{2}}(E_s)$ for some $\beta \in (\alpha,2)$, then $I\phi \in H^{\frac{\beta-\alpha}{2},\frac{\beta-\alpha}{4}}(E_s)$ and*

$$\|I\phi\|_{E_s}^{\left(\frac{\beta-\alpha}{2}\right)} \le C\,\|\phi\|_{E_s}^{(\beta)}, \tag{3.2}$$

for a positive constant C depending on α and β.

Since the proofs of Lemmas 3.1 and 3.2 are similar, we only present the proof of Lemma 3.2.

Proof of Lemma 3.2. Statement (ii) is a special case of Statement (i) when the domain is taken to be $\mathbb{R}$, instead of D. In particular, $\|\cdot\|_{E_s}^{(\beta)} \ge L$; see Definition 1.1.

It then suffices to prove Statement (i). This proof is inspired by the proof of Proposition 2.5 in [23]. For notational simplicity C represents a generic constant throughout the rest of proof.

<u>Step 1</u>: *Estimate* $\max_{\overline{D_s}}|I\phi|$. For any $(x,t) \in \overline{D_s}$,

$$\begin{aligned} |I\phi(x,t)| &\le \int_{|y|\le 1} |\phi(x+y,t) - \phi(x,t) - y\,\partial_x\phi(x,t)|\,\nu(dy) + \int_{|y|>1} |\phi(x+y,t) - \phi(x,t)|\,\nu(dy)\\ &\le \int_{|y|\le 1} |y\,\partial_x\phi(z,t) - y\,\partial_x\phi(x,t)|\,\nu(dy) + L\int_{|y|>1}|y|\nu(dy)\\ &\le \|\phi\|_{\overline{D_s^1}}^{(\beta)}\int_{|y|\le 1}|y|^\beta\nu(dy) + L\int_{|y|>1}|y|\nu(dy)\\ &\le C\left(L + \|\phi\|_{\overline{D_s^1}}^{(\beta)}\right), \end{aligned}$$

where the second inequality follows from the mean value theorem with $|z-x| \le |y|$; the third inequality is the result of the Hölder continuity of $\partial_x\phi$ on $\overline{D_s^1}$; the fourth inequality holds thanks to Assumption 2.2.

Step 2: *Show that $I\phi$ is Hölder continuous in x.* For $x_1, x_2 \in D$ and $t \in [0, s]$, we break up $|I\phi(x_1,t) - I\phi(x_2,t)|$ into three parts:

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

Here the constant $\epsilon \leq 1$ will be determined later. Let us estimate each above integral term separately. First, an estimate similar to that in Step 1 shows that $I_1 \leq 2\,\|\phi\|^{(\beta)}_{D^1_s}\int_{|y|\leq\epsilon}|y|^\beta\nu(dy) = C\|\phi\|^{(\beta)}_{D^1_s}\epsilon^{\beta-\alpha}$. Second, the Lipschitz continuity of $x \mapsto \phi(x,t)$ and the Hölder continuity of $x \mapsto \partial_x\phi(x,t)$ on $\overline{D^1_s}$ together imply that

$$
\begin{aligned}
I_2 &\leq \int_{\epsilon<|y|\leq 1}\left[2L|x_1-x_2| + \|\phi\|^{(\beta)}_{D^1_s}|y|\,|x_1-x_2|^{\beta-1}\right]\nu(dy)\\
&\leq CL|x_1-x_2|(\epsilon^{-\alpha}-1) + C\,\|\phi\|^{(\beta)}_{D^1_s}|x_1-x_2|^{\beta-1}\cdot\begin{cases}\epsilon^{1-\alpha}-1 & \text{when } 1<\alpha<2,\\ -\log\epsilon & \text{when } \alpha=1,\end{cases}
\end{aligned}
$$

where the second inequality follows from Assumption 2.2. Third, it is clear from the Lipschitz continuity of ϕ that $I_3 \leq 2L|x_1-x_2|\int_{|y|>1}\nu(dy)$.

Now pick $\epsilon = |x_1-x_2|^{\frac{1}{2}} \wedge 1$. Since $1 \leq \alpha < 2$ and $\beta > \alpha$, we have $\epsilon^{\beta-\alpha} \leq |x_1-x_2|^{\frac{\beta-\alpha}{2}}$, $\epsilon^{-\alpha}-1 \leq |x_1-x_2|^{-\frac{\alpha}{2}}$, $\epsilon^{1-\alpha}-1 \leq |x_1-x_2|^{\frac{1-\alpha}{2}}$ and $-\log\epsilon \leq \frac{2}{\beta-1}|x_1-x_2|^{\frac{1-\beta}{2}}$. All above estimates combined imply that

$$
|I\phi(x_1,t) - I\phi(x_2,t)| \leq C\left(L + \|\phi\|^{(\beta)}_{D^1_s}\right)|x_1-x_2|^{\frac{\beta-\alpha}{2}},
$$

for a constant C independent of x_1, x_2, and t.

Step 3: *Show that $I\phi$ is Hölder continuous in t.* The proof is similar to that in Step 2. First we separate $|I\phi(x,t_1) - I\phi(x,t_2)|$ into three parts as above. Then using that fact that $|\partial_x\phi(x,t_1) - \partial_x\phi(x,t_2)| \leq \|\phi\|^{(\beta)}_{D^1_s}|t_1-t_2|^{\frac{\beta-1}{2}}$ (see Definition 1.1) and choosing $\eta = |t_1-t_2|^{\frac{1}{4}} \wedge 1$, we can obtain

$$
|I\phi(x,t_1) - I\phi(x,t_2)| \leq C\left(L + \|\phi\|^{(\beta)}_{D^1_s}\right)|t_1-t_2|^{\frac{\beta-\alpha}{4}},
$$

for a constant C independent of x, t_1, and t_2. □

When I is considered as an operator between Sobolev spaces, it maps $W^{2,1}_p$−functions to L_p−functions on a smaller domain.

Lemma 3.3. *Let Assumption 2.2 hold. Consider a function $\phi \in W^{2,1}_p(D\times(t_1,t_2))$ such that $|\phi|$ is bounded and $|\partial_x\phi|$ is locally bounded on $\mathbb{R}\times[t_1,t_2]$. Then for any $\eta > 0$,*

$$
\|I\phi\|_{L_p(D\times(t_1,t_2))} \leq C\eta^{2-\alpha}\|\phi\|_{W^{2,1}_p(D^\eta\times(t_1,t_2))} + C\left(\max_{\mathbb{R}\times[t_1,t_2]}|\phi| + \max_{D^1\times[t_1,t_2]}|\partial_x\phi|\right)\cdot\begin{cases}(1+\eta^{1-\alpha}),\ \alpha\neq 1\\ (1-\log\eta),\ \alpha=1\end{cases}, \tag{3.3}
$$

for some constant C depending on D, t_1, and t_2.

Remark 3.4. When X has finite variation jumps, i.e., $0 \leq \alpha < 1$, η can be chosen as zero. Hence L_p−norm of $I\phi$ only depends on $\max_{\mathbb{R}\times[t_1,t_2]}|\phi|$ and $\max_{D^1\times[t_1,t_2]}|\partial_x\phi|$.

Proof. Since $C^{2,1}$ is dense in $W_p^{2,1}$ (c.f. Theorem 5.3.1 in [12] pp.250), ϕ can be considered as a $C^{2,1}$ function. Observing that $\phi(x+y,t)-\phi(x,t)-y\partial_x\phi(x,t) = y^2\int_0^1(1-z)\,\partial^2_{xx}\phi(x+zy,t)\,dz$, the integral $I\phi$ can be separated into the following three parts:

$$
\begin{aligned}
|I\phi(x,t)| \le & \int_{|y|\le\eta} y^2\,\nu(dy)\int_0^1 dz\,\left|\partial^2_{xx}\phi(x+zy,t)\right| \\
& + \int_{\eta<|y|\le 1}\nu(dy)\,|\phi(x+y,t)-\phi(x,t)-y\partial_x\phi(x,t)| + \int_{|y|>1}\nu(dy)\,|\phi(x+y,t)-\phi(x,t)| \\
=: & I_1 + I_2 + I_3.
\end{aligned}
$$

In the rest of proof, the L_p-norm of each above term is estimated respectively. First,

$$
\begin{aligned}
\|I_1(\cdot,t)\|^p_{L_p(D)} &= \int_D dx\left[\int_{|y|\le\eta} y^2\nu(dy)\int_0^1 dz\,\left|\partial^2_{xx}\phi(x+zy,t)\right|\right]^p \\
&\le \int_D dx\int_0^1 dz\left[\int_{|y|\le\eta}\nu(dy)\,|y|^2\left|\partial^2_{xx}\phi(x+zy,t)\right|\right]^p \\
&\le C\int_D dx\int_0^1 dz\left[\int_{|y|\le\eta} dy\,|y|^{1-\alpha}\left|\partial^2_{xx}\phi(x+zy,t)\right|\right]^p \\
&\le C\int_D dx\int_0^1 dz\left(\int_{|y|\le\eta} dy\,|y|^{1-\alpha}\right)^{\frac{p}{q}}\cdot\int_{|y|\le\eta} dy\,|y|^{1-\alpha}\left|\partial^2_{xx}\phi(x+zy,t)\right|^p \\
&\le C\,\eta^{(2-\alpha)p}\left\|\partial^2_{xx}\phi(\cdot,t)\right\|^p_{L_p(D^\eta)},
\end{aligned}
$$

where the first inequality follows from Fubini's Theorem and Jensen's inequality since $p>1$; the second inequality is a result of Assumption 2.2; the third inequality follows from Hölder inequality with $1/p+1/q=1$; the fourth inequality holds since $x+zy\in D^\eta$ for any $|y|\le\eta$ and $z\in[0,1]$. Second, since $x+y\in D^1$ for $x\in D$ and $|y|\le 1$,

$$
\|I_2(\cdot,t)\|_{L_p(D)} \le C\max_{D^1\times[t_1,t_2]}|\partial_x\phi|\cdot\int_{\eta\le|y|\le 1}|y|\nu(dy) \le C\max_{D^1\times[t_1,t_2]}|\partial_x\phi|\cdot\begin{cases}(1+\eta^{1-\alpha}),\ \alpha\ne 1\\ (1-\log\eta),\ \alpha=1\end{cases}
$$

Third, it is clear that $\|I_3\phi(\cdot,t)\|_{L_p(D)}\le C\cdot\max_{\mathbb{R}\times[t_1,t_2]}|\phi|$, since ϕ is bounded.

Now, since $\|I\phi\|_{L_p(D\times(t_1,t_2))} := \left[\int_{t_1}^{t_2}\|I\phi(\cdot,t)\|_{L_p(D)}\,dt\right]^{\frac{1}{p}}$, the statement follows from above L_p-norm estimates on I_k, $k=1,2,3$. □

3.2. **An interior estimate.** The L_p-norm estimate of the integral term in Lemma 3.3 helps to derive the following $W_p^{2,1}-$norm estimate for solutions of the Cauchy problem below. This estimate is a nonlocal version of the parabolic Calderon-Zygmund estimate (c.f. Theorem 9.1 in [17] pp.341).

Proposition 3.5. *Suppose that Assumptions 2.1 and 2.2 are satisfied. Let v be a $W^{2,1}_{p,loc}-$solution of the following Cauchy problem:*

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

where $f\in L_{p,loc}(E_T)$. If v is bounded and $\partial_x v$ is locally bounded on E_T, then for any $s\in(0,T)$, there exist $\delta\in(0,s)$ and C_δ, depending on δ, such that

$$
\|v\|_{W_p^{2,1}(D\times(s,T))} \le C_\delta\left[\max_{E_T}|v| + \max_{D^{\delta/4+1}\times[0,T]}|\partial_x v| + \|f\|_{L^p(D^{\delta/4}\times(\frac{\delta}{2},T))}\right]. \tag{3.4}
$$

Remark 3.6. The main idea of the following proof is to treat Iv as a driving term and utilize the classical Calderon-Zygmund estimate. However, as we have seen in Lemma 3.3, $W_p^{2,1}-$norm of v controls L_p-norm of Iv, which in term bounds the $W_p^{2,1}-$norm of v via the Calderon-Zygmund estimate. Therefore, a careful balance between extending domains and controlling $W_p^{2,1}-$norm of v needs to be employed in the following proof. This is contrast to the case where only finite variation jumps are considered. As we have seen in Remark 3.4, $\max|\partial_x v|$ and $\max|v|$ control the L_p-norm of Iv which bounds the $W_p^{2,1}$-norm of v. Hence, in this case, (3.4) can be obtained directly from the classical Calderon-Zygmund estimate.

Proof. The constant C denotes a generic constant throughout this proof. Domains used in this proof are displayed in Figure 1.

For a constant $\delta \in (0, s)$ which will be determined later, let us choose a cut-off function ζ^δ such that $0 \leq \zeta^\delta \leq 1$, $\zeta^\delta = 1$ inside $D \times (\delta, T)$ and $\zeta^\delta = 0$ outside $D^{\delta/4} \times (\delta/2, T)$. Moreover ζ^δ can be chosen to satisfy

$$\left|\partial_x \zeta^\delta\right| \leq \frac{C}{\delta}, \quad \left|\partial^2_{xx} \zeta^\delta\right| \leq \frac{C}{\delta^2}, \quad \text{and} \quad \left|\partial_t \zeta^\delta\right| \leq \frac{C}{\delta}. \tag{3.5}$$

The function $w := \zeta^\delta\, v$ satisfies

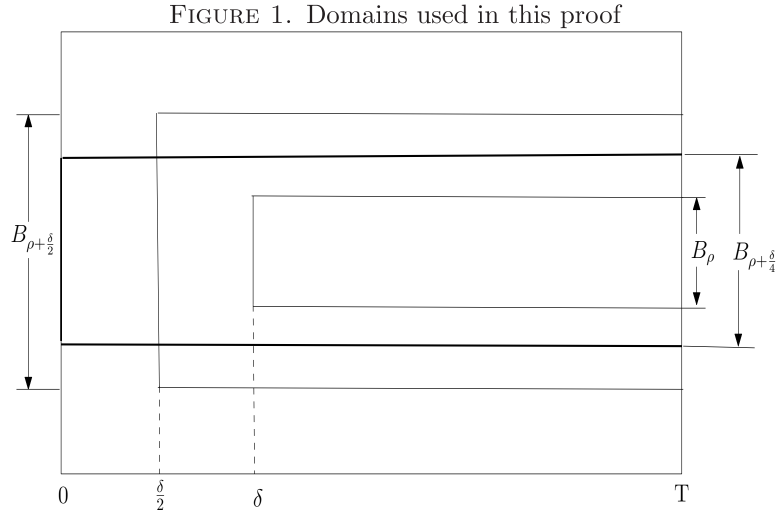
$$\begin{aligned}
&(\partial_t - \mathcal{L}_D + r)\ w = \zeta^\delta\, Iv(x,t) + \zeta^\delta\, f(x,t) + h(x,t), && (x,t) \in D^{\delta/4} \times (0,T),\\
&w(x,t) = 0, && (x,t) \in \partial\overline{D^{\delta/4}} \times [0,T),\\
&w(x,0) = 0, && x \in \overline{D^{\delta/4}},
\end{aligned}$$

in which $h(x,t) := \partial_t \zeta^\delta\, v - a\left(\partial^2_{xx}\zeta^\delta\, v + 2\,\partial_x \zeta^\delta\, \partial_x v\right) - b\,\partial_x \zeta^\delta\, v$. Appealing to Theorem 9.1 in [17] pp.341, we can find a constant C such that

$$\|w\|_{W_p^{2,1}(D^{\delta/4}\times(0,T))} \leq C\left[\left\|\zeta^\delta\, Iv\right\|_{L_p} + \left\|\zeta^\delta\, f\right\|_{L_p} + \|h\|_{L_p}\right], \tag{3.6}$$

where all L_p-norms on the right-hand-side are taken on $D^{\delta/4} \times (0,T)$.

Figure 1. Domains used in this proof



In what follows, we will estimate the terms on the right-hand-side of (3.6) respectively. First, when $\alpha \neq 1$,

$$\begin{aligned}\left\|\zeta^\delta\, Iv\right\|_{L_p(D^{\delta/4}\times(0,T))} &\leq \|Iv\|_{L_p(D^{\delta/4}\times(\frac{\delta}{2},T))}\\ &\leq C\left(\frac{\delta}{4}\right)^{2-\alpha}\|v\|_{W^{2,1}_p(D^{\delta/2}\times(\frac{\delta}{2},T))} + C\left(1+\left(\frac{\delta}{4}\right)^{1-\alpha}\right)\left[\max_{E_T}|v| + \max_{D^{\delta/4+1}\times[0,T]}|\partial_x v|\right],\end{aligned}$$

where the first inequality follows from the choice of ζ^δ; the second inequality follows from Lemma 3.3 where $\eta = \frac{\delta}{4}$, $t_1 = \frac{\delta}{2}$, and $t_2 = T$. When $\alpha = 1$, a similar estimate can be obtained. In that case, the rest of proof is similar to that for $\alpha \neq 1$ case, hence we only present the proof for $\alpha \neq 1$ henceforth. Second, it is clear that $\left\|\zeta^\delta\, f\right\|_{L_p(D^{\delta/4}\times(0,T))} \leq \|f\|_{L_p(D^{\delta/4}\times(\frac{\delta}{2},T))}$. Third, we will estimate the L_p-norm of h. To this end, let us derive a bound for $\|\partial_t\zeta^\delta v\|_{L_p(D^\delta\times(0,T))}$ in what follows. It follows from (3.5) that

$$\left\|\partial_t\zeta^\delta\, v\right\|_{L_p(D^\delta\times(0,T))} \leq C\max_{E_T}|v|\,\delta^{-1}Area\left(D^{\delta/4}\times(\delta/2,T)\setminus D\times(\delta,T)\right)^{\frac{1}{p}} \leq C\max_{E_T}|v|\,\delta^{\frac{1-p}{p}},$$

where $Area(\cdot)$ is the Lebesgue measure. Estimates on other terms of h can be performed similarly to obtain

$$\|h\|_{L_p(D^{\delta/4}\times(0,T))} \leq C\left(\delta^{\frac{1-p}{p}} + \delta^{\frac{1-2p}{p}}\right)\left(\max_{E_T}|v| + \max_{D^{\delta/4}\times[0,T]}|\partial_x v|\right).$$

Using the above estimates on the right-hand-side of (3.6), we obtain that

$$\begin{aligned}\|v\|_{w^{2,1}_p(D\times(\delta,T))} \leq{}& \|w\|_{w^{2,1}_p(D^{\delta/4}\times(0,T))}\\ \leq{}& C\left(\frac{\delta}{4}\right)^{2-\alpha}\|v\|_{W^{2,1}_p(D^{\delta/2}\times(\frac{\delta}{2},T))} + C\left(1+\delta^{1-\alpha}+\delta^{\frac{1-p}{p}}+\delta^{\frac{1-2p}{p}}\right)\left(\max_{E_T}|v| + \max_{D^{\delta/4+1}\times[0,T]}|\partial_x v|\right)\\ &+ \|f\|_{L_p(D^{\delta/4}\times(\frac{\delta}{2},T))}.\end{aligned}$$

Multiplying δ^2 on both hand sides of the previous inequality, we obtain

$$\delta^2\,\|v\|_{w^{2,1}_p(D\times(\delta,T))} \leq 4C\left(\frac{\delta}{4}\right)^{2-\alpha}\left(\frac{\delta}{2}\right)^2\|v\|_{w^{2,1}_p(D^{\delta/2}\times(\frac{\delta}{2},T))} + K(\delta),$$

where $K(\delta) = C\left(\delta^2 + \delta^{3-\alpha} + \delta^{\frac{1+p}{p}} + \delta^{\frac{1}{p}}\right)\left(\max_{E_T}|v| + \max_{D^{\delta/4+1}\times[0,T]}|\partial_x v|\right) + \delta^2\,\|f\|_{L_p(D^{\delta/4}\times(\frac{\delta}{2},T))}$. Denote $F(\tau) := \tau^2\,\|v\|_{w^{2,1}_p(D^{\delta-\tau}\times(\tau,T))}$. The previous inequality gives the following recursive inequality

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

Now choosing a sufficiently small $\delta \in (0,s)$ such that $4C\left(\delta/4\right)^{2-\alpha} \leq \frac{1}{2}$, we obtain from the above inequality that

$$F(\delta) \leq \frac{1}{2}\,F\left(\frac{\delta}{2}\right) + K(\delta).$$

$F(\delta)$ is finite for any δ, since the $W^{2,1}_p-$norm of v is finite in any compact domain of $\mathbb{R}\times(0,T)$, and $K(\delta)$ is increasing in δ. We then obtain from iterating the previous recursive inequality that

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

In terms of $W^{2,1}_{p,loc}-$norms, the previous inequality reads

$$\begin{aligned}\|v\|_{W^{2,1}_p(D\times(s,T))} &\leq 2\,C\left[1+\delta^{1-\alpha}+\delta^{\frac{1-p}{p}}+\delta^{\frac{1-2p}{p}}\right]\left[\max_{E_T}|v| + \max_{D^{\delta/4+1}\times[0,T]}|\partial_x v|\right] + 2\,\|f\|_{L_p(D^{\delta/4}\times(\frac{\delta}{2},T))}\\ &\leq C_\delta\left[\max_{E_T}|v| + \max_{D^{\delta/4+1}\times[0,T]}|\partial_x v| + \|f\|_{L_p(D^{\delta/4}\times(\frac{\delta}{2},T))}\right].\end{aligned}$$

□

4. Proof of main results

4.1. **The penalty method.** We use the penalty method (see e.g. [14] and [24]) to analyze the following variational inequality.

$$
\begin{aligned}
&\min\{(\partial_t - \mathcal{L}_D - I + r)\, v, v - g\} = 0, && (x,t) \in \mathbb{R} \times (0,T], \\
&v(x,0) = g(x), && x \in \mathbb{R}.
\end{aligned} \tag{4.1}
$$

The nonlocal integral term introduces several technical difficulties in applying the penalty method. In this section, we will focus on the case where X has infinite variation jumps, i.e., Assumption 2.2 holds with $1 \leq \alpha < 2$. When X has finite variation jumps, i.e., $0 \leq \alpha < 1$, the integral operator has the reduced form I^f in (2.5). Then all proofs are similar but easier than those in infinite variation case.

For each $\epsilon \in (0,1)$, consider the following penalty problem:

$$
\begin{aligned}
&(\partial_t - \mathcal{L}_D - I + r)\, v^\epsilon + p_\epsilon(v^\epsilon - g^\epsilon) = 0, && (x,t) \in \mathbb{R} \times (0,T], \\
&v^\epsilon(x,0) = g^\epsilon(x), && x \in \mathbb{R},
\end{aligned} \tag{4.2}
$$

Here $\{g^\epsilon\}_{\epsilon \in (0,1)}$ is a mollified sequence of g such that $\partial^2_{xx} g^\epsilon(x) \geq -J$, $0 \leq g \leq K$, and $|(g^\epsilon)'(x)| \leq L$ for any $x \in \mathbb{R}$; see [14] pp.27 for its construction. These constants J, K, and L, appearing in Assumption 2.4, are independent of ϵ. The penalty term $p_\epsilon(y) \in C^\infty(\mathbb{R})$ is chosen to satisfy following properties:

(4.3)

$$
(i)\, p_\epsilon(y) \leq 0, \quad (ii)\, p_\epsilon(y) = 0 \text{ for } y \geq \epsilon, \quad (iii)\, p_\epsilon(0) = -a^{(0)}J - |b|^{(0)}L - r^{(0)}K - J\int_{|y|\leq 1} |y|^2 \nu(dy) - K\int_{|y|>1} \nu(dy),
$$

$$
(iv)\, p_\epsilon'(y) \geq 0, \quad (v)\, p_\epsilon''(y) \leq 0, \quad \text{and} \quad (vi)\, \lim_{\epsilon \downarrow 0} p_\epsilon(y) = \begin{cases} 0, & y > 0 \\ -\infty, & y < 0 \end{cases},
$$

where $a^{(0)} = \max_{E_T} a$, $|b|^{(0)} = \max_{E_T} |b|$, and $r^{(0)} = \max_{E_T} r$ are finite thanks to Assumption 2.1. Indeed, p_ϵ can be chosen as a smooth mollification of the function $\min\{-\frac{2p_\epsilon(0)}{\epsilon}x + p_\epsilon(0), 0\}$.

Now we show that each penalty problem (4.2) has a classical solution. To this end, let us first recall the Schauder fixed point theorem (see Theorem 2 in [13] pp. 189).

Lemma 4.1. *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.*

Lemma 4.2. *Let Assumptions 2.1, 2.2 with $1 \leq \alpha < 2$, and 2.4 hold. Then for any $\epsilon \in (0,1)$ and $\beta \in (\alpha, 2)$, (4.2) has a solution $v^\epsilon \in H^{2+\frac{\beta-\alpha}{2}, 1+\frac{\beta-\alpha}{4}}(E_T)$.*

Proof. We will first prove the existence on a sufficiently small time interval $[0,s]$ via the Schauder fixed point theorem, then extend this solution to the interval $[0,T]$.

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

$$
\begin{aligned}
&(\partial_t - \mathcal{L}_D + r)(u - g^\epsilon) = Iv - p_\epsilon(v - g^\epsilon) + (\mathcal{L}_D - r)\, g^\epsilon, && (x,t) \in \mathbb{R} \times (0,s], \\
&u(x,0) - g^\epsilon(x) = 0, && x \in \mathbb{R}.
\end{aligned} \tag{4.4}
$$

We define an operator $\mathcal{T}$ via $u = \mathcal{T}v$ using the solution u of (4.4). Let us check the conditions for the Schauder fixed point theorem are satisfied in the following four steps:

<u>*Step 1*</u>*: Tv is well defined.* Since $v \in H^{\beta, \frac{\beta}{2}}(E_s)$ with $\beta \in (\alpha, 2)$, Lemma 3.2 part (ii) implies that $Iv \in H^{\frac{\beta-\alpha}{2}, \frac{\beta-\alpha}{4}}(E_s)$ with $\|Iv\|_{E_s}^{(\frac{\beta-\alpha}{2})} \le C\,\|v\|_{E_s}^{(\beta)}$. On the other hand, using properties of v, g^ϵ and p_ϵ, one can check that $-p_\epsilon(v-g^\epsilon) + (\mathcal{L}_D - r)g^\epsilon \in H^{\frac{\beta-\alpha}{2}, \frac{\beta-\alpha}{4}}(E_s)$. Therefore, Theorem 5.1 in [17] pp. 320 implies that (4.4) has a unique solution $u - g^\epsilon \in H^{2+\frac{\beta-\alpha}{2}, 1+\frac{\beta-\alpha}{4}}(E_s)$. Hence $u = Tv \in H^{2+\frac{\beta-\alpha}{2}, 1+\frac{\beta-\alpha}{4}}(E_s)$, since g^ϵ is smooth.

<u>*Step 2*</u>. $\mathcal{T}\Theta \subset \Theta$. It follows from Lemma 2 in [13] pp. 193 that there exists a positive constant A_β, depending on β, such that

$$
\begin{aligned}
\|u-g^\epsilon\|_{E_s}^{(\beta)} &\le A_\beta s^\gamma \left[\|Iv\|_{E_s}^{(0)} + \|p_\epsilon(v-g^\epsilon)\|_{E_s}^{(0)} + \|(\mathcal{L}_D - r)\, g^\epsilon\|_{E_s}^{(0)}\right] \\
&\le A_\beta C s^\gamma \|v\|_{E_s}^{(\beta)} + \widetilde{A},
\end{aligned}
\tag{4.5}
$$

where $\gamma = \frac{2-\beta}{2}$, C is the constant in Step 1, and $\widetilde{A}$ is a sufficiently large constant. Let s be such that $\tau := A_\beta C s^\gamma < 1/2$ and let $U_0 := \max\{\frac{2\widetilde{A}}{1-2\tau}, 2\,\|g^\epsilon\|_{E_s}^{(\beta)}\}$. Since $\|v\|_{E_s}^{(\beta)} \le U_0$, it then follows from (4.5) that

$$
\|u\|_{E_s}^{(\beta)} \le \|u-g^\epsilon\|_{E_s}^{(\beta)} + \|g^\epsilon\|_{E_s}^{(\beta)} \le \tau U_0 + \widetilde{A} + \frac{U_0}{2} \le \tau\, U_0 + \frac{1-2\tau}{2}\, U_0 + \frac{U_0}{2} = U_0. \tag{4.6}
$$

This confirms that $u = \mathcal{T}v \in \Theta$.

<u>*Step 3*</u>. *$\mathcal{T}\Theta$ is a precompact subset of $H^{\beta,\frac{\beta}{2}}(E_s)$.* For any $\eta \in (\beta, 2)$, a similar estimate as (4.5) shows that for any $v \in \Theta$, $\|Tv\|_{E_s}^{(\eta)} \le U_1$ for some constant U_1 depending on U_0 and s. Since bounded subsets of $H^{\eta,\frac{\eta}{2}}(E_s)$ are precompact subsets of $H^{\beta,\frac{\beta}{2}}(E_s)$ (see Theorem 1 in [13] pp.188), then $\mathcal{T}\Theta$ is a precompact subset in $H^{\beta,\frac{\beta}{2}}(E_s)$.

<u>*Step 4*</u>. *$\mathcal{T}$ is a continuous operator.* Let v_n be a sequence in Θ such that $\lim_{n\to\infty} \|v_n - v\|_{E_s}^{(\beta)} = 0$, we will show $\lim_{n\to\infty} \|Tv_n - Tv\|_{E_s}^{(\beta)} = 0$. From (4.4), $w \triangleq Tv_n - Tv$ satisfies the Cauchy problem

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

It follows again from Lemma 2 in [13] pp. 193 that

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

Now all conditions of the Schauder fixed point theorem are checked, hence $\mathcal{T}$ has a fixed point in $H^{\beta,\frac{\beta}{2}}(E_s)$, which is denoted by v^ϵ. Moreover, it follows from results in Step 1 that $v^\epsilon = \mathcal{T}v^\epsilon \in H^{2+\frac{\beta-\alpha}{2}, 1+\frac{\beta-\alpha}{4}}(E_s)$.

Finally, let us extend v^ϵ to the interval $[0,T]$. We can replace $g^\epsilon(\cdot)$ by $v^\epsilon(\cdot, s)$ in (4.4), since $\|v^\epsilon(\cdot,s)\|_{\mathbb{R}}^{(2+\frac{\beta-\alpha}{2})}$ is finite thanks to the result after Step 4 and because the choice of s in Step 2 only depends on β and C. If we choose a sufficiently large U_0, depending on $\|v^\epsilon(\cdot,s)\|_{\mathbb{R}}^{(2+\frac{\beta-\alpha}{2})}$, such that (4.6) holds on $[s,2s]$, then $\|v^\epsilon(\cdot,2s)\|_{\mathbb{R}}^{(2+\frac{\beta-\alpha}{2})}$ is finite thanks to the argument after Step 4. Now one can repeat this procedure to extend the time interval by s each time, until it contains $[0,T]$. □

After the existence of classical solutions for (4.2) is established, we will study properties of the sequence $(v^\epsilon)_{\epsilon\in(0,1)}$ in the rest of this subsection. The following maximum principle is a handy tool for our analysis.

Lemma 4.3. *Suppose that $a > 0$, a and b are bounded and the Levy measure ν satisfies $\int_{|y|>1} |y|\nu(dy) < \infty$. Assume also that we are given a function c bounded from below on E_T. If $v \in C^0(E_T) \cap C^{2,1}(E_T)$ satisfies $(\partial_t - \mathcal{L}_D - I + c)\, v(x,t) \ge 0$ and v is bounded from below on E_T, then $v(x,0) \ge 0$ for $x \in \mathbb{R}$ implies that $v \ge 0$ on E_T.*

Proof. Let $v \geq -m$ and $c \geq -C_0$ on E_T for some positive constants m and C_0. For any positive R_0, consider the following function:

$$w(x,t) := \frac{m}{f(R_0)} \left(f(|x|) + C_1 t\right) e^{C_0 t}, \quad (x,t) \in E_T,$$

where C_1 will be determined later and $f : \mathbb{R}_+ \to \mathbb{R}_+$ is an increasing C^2 function such that $f = 0$ in a neighborhood of 0 and $f(R) = \frac{R^2}{1+R}$ for sufficiently large R. It is clear that $\lim_{R\to+\infty} f(R) = \infty$ and derivatives $f^{'}$ and $f^{''}$ are bounded. Then $If(|x|)$ is bounded on $\mathbb{R}$. Indeed, there exists a constant C such that

$$\begin{aligned}
\left|If(|x|)\right| &\leq \int_{|y|\leq 1} \nu(dy) \int_0^1 dz\,(1-z)\,y^2 \left|\partial^2_{xx} f(|x+zy|)\right| + \int_{|y|>1} \nu(dy)\,|f(|x+y|) - f(|x|)| \\
&\leq C\left(\int_{|y|\leq 1} y^2 \nu(dy) + \int_{|y|>1} |y|\,\nu(dy)\right) < +\infty.
\end{aligned}$$

Combining above estimate with $c + C_0 \geq 0$, one can find a sufficient large constant C_1 such that

$$(\partial_t - \mathcal{L}_D - I + c)w = e^{C_0 t}\frac{m}{f(R_0)}\left[C_1 + (c+C_0)(f(|x|) + C_1 t) - a\,\partial^2_{xx} f(|x|) - b\,\partial_x f(|x|) - If(|x|)\right] > C_0 m, \quad \text{on } E_T.$$

Now define $\tilde{v} := v + w$. The previous estimate gives

$$(\partial_t - \mathcal{L}_D - I + c + C_0)\tilde{v} > C_0 v + C_0 m \geq 0, \quad \text{for any } (x,t) \in E_T. \tag{4.7}$$

On the other hand, $\tilde{v}(x,0) = \frac{m}{f(R_0)} f(|x|) + v(x,0) \geq 0$ due to $v(x,0) \geq 0$, moreover $\tilde{v}(x,t) \geq m + v(x,t) \geq 0$ for $|x| \geq R_0$ because f is increasing and $v \geq -m$. Therefore, we claim that $\tilde{v} \geq 0$ for $(x,t) \in [-R_0, R_0] \times [0,T]$. Indeed, if there exists $(x,t) \in [-R_0,R_0] \times (0,T]$ such that $\tilde{v}(x,t) < 0$, $\tilde{v}$ must takes its negative minimum at some point $(x_0,t_0) \in [-R_0,R_0] \times (0,T]$. Note that this is also a global minimum for $\tilde{v}$ on E_T, hence $I\tilde{v}(x_0,t_0) \geq 0$, $\partial_t \tilde{v}(x_0,t_0) \leq 0$, $\partial_x \tilde{v}(x_0,t_0) = 0$, $\partial^2_{xx}\tilde{v}(x_0,t_0) \geq 0$, and $(c+C_0)\tilde{v}(x_0,t_0) \leq 0$. As a result, $(\partial_t - \mathcal{L}_D - I + c + C_0)\tilde{v}(x_0,t_0) \leq 0$, which contradicts with (4.7). Now for fixed point (x,t), the statement follows from sending the constant R_0 in $\tilde{v}$ to ∞. □

This maximum principle implies the uniqueness of classical solutions for the penalty problem.

Corollary 4.4. *Under assumptions of Lemma 4.2, v^ϵ is the unique bounded classical solution of* (4.2).

Proof. Lemma 4.2 and the definition of Hölder spaces combined imply that $v_1 = v^\epsilon$ is a bounded classical solution. Now suppose there exists another solution v_2, then $v_1 - v_2$ satisfies

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

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 $y \in \mathbb{R}$, where $p^{'}_\epsilon(y) \geq 0$ thanks to (4.2) part (iv). Now it follows from Lemma 4.3, with $c = r + p'_\epsilon(y)$ that $v_1 \geq v_2$ on $\mathbb{R} \times (0,T]$. The same argument applied to $v_2 - v_1$ gives the reverse inequality. □

Utilizing the maximum principle, we will analyze properties of the sequence $(v^\epsilon)_{\epsilon\in(0,1)}$ in the following result.

Lemma 4.5. *Let Assumptions 2.1, 2.2 with $1 \leq \alpha < 2$, and 2.4 hold. Then for any $\epsilon \in (0,1)$,*

$$0 \leq v^\epsilon \leq K + 1 \quad \text{on } E_T.$$

Proof. It follows from Lemma 4.2 that v^ϵ is bounded on E_T for each $\epsilon \in (0,1)$. In this proof, we will show that the bounds are uniform in ϵ. First, it follows from (4.3) part (i) that $(\partial_t - \mathcal{L}_D - I + r)\, v^\epsilon = -p_\epsilon(v^\epsilon - g^\epsilon) \geq 0$. Moreover, $v^\epsilon(x,0) = g^\epsilon(x) \geq 0$ for $x \in \mathbb{R}$. Then first inequality in the statement follows from Lemma 4.3 directly. Second, consider $w = K + 1 - v^\epsilon$, it satisfies

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

Combining (4.3) part (ii) and $g^\epsilon \leq K$, we have $p_\epsilon(K + 1 - g^\epsilon) = 0$. Hence,

$$(\partial_t - \mathcal{L}_D - I + r)\, w + p_\epsilon(K+1-g^\epsilon) - p_\epsilon(v^\epsilon - g^\epsilon) = \left[\partial_t - \mathcal{L}_D - I + r + p^{'}_\epsilon(y)\right] w = r\,(K+1) \geq 0, \tag{4.8}$$

where the first equality follows from the mean value theorem. Now applying Lemma 4.3 to above equation with $c = r + p'(y) \geq 0$ (see (4.3)) part (iv), we obtain $w(x,t) = K + 1 - v^\epsilon(x,t) \geq 0$ on E_T for any $\epsilon \in (0,1)$, which confirms the second inequality in the statement of the lemma. □

Lemma 4.6. *Let Assumptions 2.1, 2.2 with $1 \leq \alpha < 2$, and 2.4 hold. Then for any $\epsilon \in (0,1)$, there exists some positive γ independent of ϵ such that*

$$|\partial_x v^\epsilon| \leq C \quad \text{on } E_T,$$

in which depends on T and L.

Proof. Formally differentiating (4.2) with respect to x gives the following equation:

$$\begin{aligned} &\left[\partial_t - a\partial^2_{xx} - (b + \partial_x a)\partial_x - I + (r - \partial_x b)\right] w + \partial_x r\, v^\epsilon + p^{'}_\epsilon\,(v^\epsilon - g^\epsilon)\left(w - (g^\epsilon)^{'}\right) = 0, \quad && (x,t) \in \mathbb{R} \times (0,T], \\ &w(x,0) = (g^\epsilon)^{'}(x), && x \in \mathbb{R}. \end{aligned} \tag{4.9}$$

We will show that $\partial_x v^\epsilon$ is indeed a classical solution of (4.9). To this end, let us consider the equation

$$\begin{aligned} &\left[\partial_t - a\partial^2_{xx} - (b + \partial_x a)\partial_x - I + (r - \partial_x b)\right] w = -\partial_x r\, v^\epsilon - p^{'}_\epsilon\,(v^\epsilon - g^\epsilon)\left(\partial_x v^\epsilon - (g^\epsilon)^{'}\right), \quad && (x,t) \in \mathbb{R} \times (0,T], \\ &w(x,0) = (g^\epsilon)^{'}(x), && x \in \mathbb{R}. \end{aligned}$$

Using Assumption 2.1 and Lemma 4.2, one can check that $-\partial_x r\, v^\epsilon - p^{'}_\epsilon\,(v^\epsilon - g^\epsilon)\left(\partial_x v^\epsilon - (g^\epsilon)^{'}\right)$ is Hölder continuous. It then follows from Theorem 3.1 in [15] on pp. 89 that the last equation has a classical solution, say w. Define $v(x,t) := \int_0^x w(z,t)dz + v^\epsilon(0,t)$. It is straight forward to check that v is a classical solution of the following equation

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

Since g^ϵ and v^ϵ are both bounded, then $-p_\epsilon(v^\epsilon - g^\epsilon)$ is also bounded. As a result estimate (3.6) in Theorem 3.1 of [15] on pp. 89 implies that v is bounded solution of the last equation. However, Corollary 4.4 already shows that v^ϵ is the unique bounded solution of the last solution, therefore $v = v^\epsilon$, hence $\partial_x v^\epsilon = w$ on E_T and $\partial_x v^\epsilon$ is a classical solution of (4.9).

Now we shall show $\partial_x v^\epsilon$ is bounded uniformly in ϵ. Consider $\tilde{v} = e^{\gamma t} L + \partial_x v^\epsilon$, where $\gamma > 0$ will be determined later. $\tilde{v}$ satisfies the following equation

$$\begin{aligned} &\left[\partial_t - a\partial^2_{xx} - (b + \partial_x a)\partial_x - I + r - \partial_x b + p^{'}_\epsilon(v^\epsilon - g^\epsilon)\right] \tilde{v} \\ &\qquad = (\gamma + r - \partial_x b)e^{\gamma t} L - \partial_x r\, v^\epsilon + p^{'}_\epsilon(v^\epsilon - g^\epsilon)\left(e^{\gamma t} L + (g^\epsilon)^{'}\right), \quad && (x,t) \in \mathbb{R} \times (0,T], \\ &\tilde{v}(x,0) = e^{\gamma t} L + (g^\epsilon)^{'}(x), && x \in \mathbb{R}. \end{aligned} \tag{4.10}$$

Recall that $\partial_x b$ and $\partial_x r$ are bounded from Assumption 2.1 and that v^ϵ is bounded uniformly in ϵ thanks to Lemma 4.5. Therefore, one can find a sufficiently large γ, independent of ϵ, such that $(\gamma + r - \partial_x b)e^{\gamma t} L - \partial_x r\, v^\epsilon > 0$.

Moreover, $p_\epsilon^{'}(v^\epsilon - g^\epsilon)\left(e^{\gamma t}L + (g^\epsilon)^{'}\right)$ is also positive due to (4.3) part (iv) and $|(g^\epsilon)^{'}| \leq L$. As a result, the right-hand-side of (4.10) is positive. Now since $r - \partial_x b + p_\epsilon^{'}(v^\epsilon - g^\epsilon)$ is bounded from below, we have from Lemma 4.3 that $\tilde{v} \geq 0$ on E_T, hence $\partial_x v^\epsilon \geq -e^{\gamma T}L$ on E_T, for some positive γ independent of ϵ. The upper bound can be shown similarly by working with $\tilde{v} = e^{\gamma t} - \partial_x v^\epsilon$. ☐

Lemma 4.7. *Let Assumptions 2.1, 2.2 with $1 \leq \alpha < 2$, and 2.4 hold. Then for any $\epsilon \in (0,1)$,*

$$v^\epsilon \geq g^\epsilon \quad \text{on } E_T.$$

Proof. Let us first show that Ig^ϵ is uniformly bounded from below. Indeed,

$$\begin{aligned} Ig^\epsilon(x) &= \int_{|y|\leq 1} \nu(dy) \int_0^1 dz(1-z)\, y^2 \partial^2_{xx} g^\epsilon(x+zy) + \int_{|y|>1} \left[g^\epsilon(x+y) - g^\epsilon(x)\right] \nu(dy) \\ &\geq -J \int_{|y|\leq 1} y^2 \nu(dy) - K \int_{|y|>1} \nu(dy), \end{aligned}$$

where the inequality follows from $\partial^2_{xx} g^\epsilon \geq -J$ and $0 \leq g^\epsilon \leq K$. As a result, $(\partial_t - \mathcal{L}_D - I + r)\, g^\epsilon$ from above. This is because

$$\begin{aligned} (\partial_t - \mathcal{L}_D - I + r)\, g^\epsilon(x) &= -a(x,t)\, \partial^2_{xx} g^\epsilon(x) - b(x,t)\, \partial_x g^\epsilon(x) + r(x,t)\, g^\epsilon(x) - Ig^\epsilon(x) \\ &\leq a^{(0)} J + |b|^{(0)} L + r^{(0)} K + J \int_{|y|\leq 1} |y|^2 \nu(dy) + K \int_{|y|>1} \nu(dy) \\ &= -p_\epsilon(0), \end{aligned}$$

where the second equality follows from (4.3) part (iii). Therefore,

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

The previous inequality and the mean value theorem combined imply that

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

for some $y \in \mathbb{R}$. Hence the statement of the lemma follows applying Lemma 4.3 to the previous inequality and choosing $c = r + p_\epsilon^{'}(y) \geq 0$. ☐

Corollary 4.8. *Let assumptions of Lemma 4.7 hold. Then $p_\epsilon\,(v^\epsilon - g^\epsilon)$ is bounded uniformly in $\epsilon \in (0,1)$.*

Proof. Lemma 4.7 and (4.3) parts (i), (iv) imply that $p_\epsilon(0) \leq p_\epsilon\,(v^\epsilon - g^\epsilon) \leq 0$. Then the statement follows since $p_\epsilon(0)$ is independent of ϵ; see (4.2) part (iii). ☐

4.2. **Proof of Theorem 2.5 and Corollary 2.6.**

Proof of Theorem 2.5. The proof consists of two steps. First, we show that there exists a function v^* which solves (4.1) and $v^* \in W^{2,1}_p(D \times (s,T))$ for any integer $p \in (1,\infty)$. Second, we confirm that $u(x,t) = v(x, T-t)$ is the value function for the problem (OS).

Step 1: First, $v^\epsilon \in W^{2,1}_{p,loc}(\mathbb{R} \times (0,T))$ for each $\epsilon \in (0,1)$, since Lemma 4.2 shows that $\partial_t v^\epsilon$, $\partial_x v^\epsilon$, and $\partial^2_{xx} v^\epsilon$ are continuous, hence locally bounded, on $\mathbb{R} \times (0,T)$. Second, Lemmas 4.5 and 4.6 show that v^ϵ and $\partial_x v^\epsilon$ are bounded on E_T, uniformly in ϵ. Moreover, the penalty term $p_\epsilon(v^\epsilon - g^\epsilon)$ is also bounded uniformly in ϵ due to Corollary 4.8. Then applying Proposition 3.5 with $f = -p_\epsilon(v^\epsilon - g^\epsilon)$, we obtain that

$$\|v^\epsilon\|_{W^{2,1}_p(D\times(s,T))} \leq C, \quad \text{for some constant } C \text{ independent of } \epsilon. \tag{4.11}$$

Combining with the fact that the Sobolev space $W_p^{2,1}$, $1 < p < \infty$, is weakly compact, we can find a subsequence $(\epsilon_k)_{k\geq 0}$, with $\epsilon_k \to 0$, and a function v^*, such that $v^{\epsilon_k} \rightharpoonup v^*$ in $W_p^{2,1}(D \times (s,T))$. Here " $\rightharpoonup$ " represents the weak convergence; c.f. Appendix D.4. in [12] pp. 639. Furthermore, the convergence is uniform for a further subsequence. Indeed, it follows from (4.11) and the Sobolev embedding theorem (c.f. Lemma 3.3 in [17] pp. 80) that

$$\|v^\epsilon\|^{(\beta)}_{\overline{D}\times[s,T]} \leq C, \quad \text{for some constant } C \text{ independent of } \epsilon.$$

Here $\beta = 2 - \frac{3}{p}$. We choose $p > 1$ so that $\beta > 0$. Using the previous uniform estimate and the Arzelà-Ascoli theorem, we can find a further subsequence of $(\epsilon_k)_{k\geq 0}$, which we still denote by $(\epsilon_k)_{k\geq 0}$, such that $(v^{\epsilon_k})_{k\geq 0}$ converge to v^* uniformly on $\overline{D} \times [s,T]$.

Let us show that v^* solves (4.1). On the one hand, since $p_\epsilon(v^\epsilon - g^\epsilon) \leq 0$, we have $(\partial_t - \mathcal{L}_D - I + r)\, v^\epsilon \geq 0$ for any $\epsilon \in (0,1)$. Since $(v^{\epsilon_k})_{k\geq 0}$ converges uniformly to v^*, we obtain that $(\partial_t - \mathcal{L}_D - I + r)\, v^* \geq 0$ on $\overline{D} \times [s,T]$ in the distribution sense. Since the choices of D and s are arbitrary, we have $(\partial_t - \mathcal{L}_D - I + r)\, v^* \geq 0$ on $\mathbb{R} \times (0,T]$ in the distributional sense. On the other hand, Lemma 4.7 shows that $v^\epsilon \geq g^\epsilon$. Then $v^* \geq g$ after sending $\epsilon \to 0$. Therefore, we obtain $\min\{(\partial_t - \mathcal{L}_D - I + r)\, v^*, v^* - g\} \geq 0$ on $\mathbb{R} \times (0,T]$. It then remains to show $(\partial_t - \mathcal{L}_D - I + r)\, v^* = 0$ when $v^* > g$. To this end, take any (x,t) such that $v^*(x,t) > g(x)$. Since both v^* and g are continuous, one can find a sufficiently small $\delta > 0$ and a small neighborhood of (x,t), such that $v^*(\tilde{x},\tilde{t}) \geq g(\tilde{x}) + 2\delta$ for any $(\tilde{x},\tilde{t})$ inside this neighborhood. Since the convergence of $(v^{\epsilon_k})_{k\geq 0}$ and $(g^{\epsilon_k})_{k\geq 0}$ is uniform, we can find sufficiently small ϵ_k such that $v^{\epsilon_k}(\tilde{x},\tilde{t}) \geq g^{\epsilon_k}(\tilde{x}) + \delta$ in the aforementioned neighborhood. Hence $p_{\epsilon_k}(v^{\epsilon_k} - g^{\epsilon_k})(x,t) = 0$, due to (4.3)-(ii), which induces $(\partial_t - \mathcal{L}_D - I + r)\, v^{\epsilon_k}(x,t) = 0$. After sending $\epsilon_k \to 0$, we conclude that $(\partial_t - \mathcal{L}_D - I + r)\, v^* = 0$ when $v^* > g$. Since $v^* \in W^{2,1}_{p,loc}$, v^* also solves (4.1) for almost every point in E_T.

Step 2: Let us first show that v^* is a viscosity solution of (4.1). We will use the definition of viscosity solutions in [21]. Denote by $C_1(E_T)$ the class of functions which have at most linear growth, i.e., $|\phi(x,t)| \leq C(1 + |x|)$ for some C and any $(x,t) \in E_T$. Then viscosity solutions of (4.1) are defined as follows.

Any $v \in C(E_T)$ is a viscosity supersolution (subsolution) of (4.1) if

$$\begin{aligned} &\min\{\partial_t\phi - \mathcal{L}_D\phi - I\phi + rv, v - g\} \geq 0\, (\leq 0), && (x,t) \in \mathbb{R} \times (0,T], \\ &v(x,0) \geq g(x)\, (\leq g(x)), && x \in \mathbb{R}, \end{aligned}$$

for any function $\phi \in C^{2,1}(\mathbb{R} \times (0,T)) \cap C_1(E_T)$ such that $v(x,t) = \phi(x,t)$ and $v(\tilde{x},\tilde{t}) \geq \phi(\tilde{x},\tilde{t})$ $(v(\tilde{x},\tilde{t}) \leq \phi(\tilde{x},\tilde{t}))$ for any other point $(x,t) \in \mathbb{R} \times (0.T)$. v is a viscosity solution of (4.1) if it is both supersolution and subsolution.

Let us show that v^* is a viscosity subsolution of (4.1). Fix $(x,t) \in \mathbb{R} \times (0,T]$, consider $v^*(x,t) > g(x)$, otherwise $\min\{\partial_t\phi - \mathcal{L}_D\phi - I\phi + rv, v(x,t) - g(x)\} \leq 0$ is automatically satisfied. Without loss of generality we can assume that (x,t) is the strict maximum of $v^* - \phi$ in a neighborhood $B(x,t;\delta)$, otherwise the test function can be modified appropriately. On the other hand, since $(v^{\epsilon_k})_{k\geq 0}$ converges to v^* uniformly in compact domains, we can find sufficiently small ϵ_k such that $v^{\epsilon_k} - \phi$ attains its maximum over $B(x,t;\delta)$ at $(x_k,t_k) \in B(x,t;\delta)$. Moreover, $(x_k,t_k) \to (x,t)$ as $\epsilon_k \to 0$. Since v^{ϵ_k} is a classical solution of (4.2) (see Lemma 4.2), it is also a viscosity solution. Hence $(\partial_t - \mathcal{L}_D - I + r)\phi(x_k,t_k) + p_{\epsilon_k}(v^{\epsilon_k}(x_k,t_k) - g^\epsilon(x_k)) \leq 0$. Now, since $v^*(x,t) > g(x)$ and $v^{\epsilon_k}(x_k,t_k) - g(x_k)$ converges to $v^*(x,t) - g(x)$, we obtain $\lim_{\epsilon_k\to 0} p_{\epsilon_k}(v^{\epsilon_k}(x_k,t_k) - g^\epsilon(x_k)) = 0$. As a result, $(\partial_t - \mathcal{L}_D - I + r)\phi(x,t) \leq 0$ by sending $\epsilon_k \to 0$. This confirms that v^* is a viscosity subsolution of (4.1).

For the supersolution property, since $v^* \geq g$, it suffices to show that $(\partial_t - \mathcal{L}_D - I + r)\phi(x,t) \geq 0$ for any test function ϕ. Then the rest proof follows from the arguments we used for the subsolution property.

Define $u^*(x,t) = v^*(x,T-t)$. It is clear that u^* is a viscosity solution of (2.1). Then the statement follows from Theorem 4.1 in [21], which states that u is the unique viscosity solution of (2.1). □

Proof of Corollary 2.6. (i) Combining Theorem 2.5 and the Sobolev embedding theorem (c.f. Lemma 3.3 in [17] pp. 80), we have $u \in H^{\beta,\frac{\beta}{2}}(D \times [0, T-s])$, where $\beta = 2 - \frac{3}{p}$ and $s < T$. Choosing $p > 3$ so that $\beta > 1$, the continuity of $\partial_x u$ follows from Definition 1.1.

(ii) Let us first show that Iu is well defined and Hölder continuous. Since $u \in H^{\beta,\frac{\beta}{2}}(D \times [0, T-s])$ (which follows due to (i)), choosing sufficiently large p so that $\beta > \alpha$, $Iu \in H^{\frac{\beta-\alpha}{2},\frac{\beta-\alpha}{4}}(\overline{D_{T-s}})$ by Lemma 3.2 part (i). Now, for $B \subset \mathbb{R}$ and $t_1, t_2 \in [0, T)$ such that $B \times (t_1, t_2) \subset \mathcal{C}$, consider the following boundary value problem:

$$
\begin{aligned}
&(-\partial_t - \mathcal{L}_D + r)\, v = Iu, \quad (x,t) \in B \times [t_1, t_2), \\
&v(x,t) = u(x,t), \qquad\qquad (x,t) \in \partial B \times [t_1, t_2) \cup \overline{B} \times t_2.
\end{aligned} \tag{4.12}
$$

It is straight forward to show that u is the unique viscosity solution for the previous problem using the fact that u is the unique viscosity solution for (2.1). On the other hand, since the boundary and terminal values of (4.12) are continuous and the driving term Iu is Hölder continuous, it follows from Theorem 9 in [13] pp. 69 that (4.12) has a classical solution $u^* \in C^{2,1}(B \times (t_1, t_2))$. Hence $u = u^*$ on $B \times (t_1, t_2)$, since u^* is also a viscosity solution. Therefore, $u \in C^{2,1}(B \times (t_1, t_2))$. The statement now follows, since $B \times (t_1, t_2)$ is an arbitrary subset of $\mathcal{C}$. □

References

[1] Y. Aït-Sahalia and J. Jacod, *Estimating the degree of activity of jumps in high frequency data*, The Annals of Statistics, 37 (2009), pp. 2202–2244.

[2] ———, *Is Brownian motion necessary to model high frequency data?*, To appear in the Annals of Statistics, (2010).

[3] A. L. Amadori, *Obstacle problem for nonlinear integro-differential equations arising in option pricing*, Ric. Mat., 56 (2007), pp. 1–17.

[4] G. Barles and C. Imbert, *Second-order elliptic integro-differential equations: viscosity solutions' theory revisited*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 25 (2008), pp. 567–585.

[5] 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 (2009), pp. 551–572.

[6] E. Bayraktar and H. Xing, *Analysis of the optimal exercise boundary of American options for jump diffusions*, SIAM Journal on Mathematical Analysis, 41 (2009), pp. 825–860.

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

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

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

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

[11] M. Davis, X. Guo, and G. Wu, *Impulse control of multidimensional jump diffusions*, Available at http://arxiv.org/abs/0912.3297v1, (December 2009).

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

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

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

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

[16] X. Guo and G. Wu, *Smooth fit principle for impulse control of multidimensional diffusion processes*, SIAM J. Control Optim., 48 (2009), pp. 594–617.

[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 (2008), pp. 561–581.

[19] S. Lenhart, *Integro-differential operators associated with diffusion processes with jumps*, Appl. Math. Optim., 9 (1982/83), pp. 177–191.

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

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

[22] K. Sato, *Lévy processes and infinitely divisible distributions*, vol. 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 (2006), pp. 67–112.

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

[25] 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 Statistics, London School of Economics, London, WC2A 2AE, UK; e-mail:xinghao82@gmail.com