

# REGULARITY OF THE OPTIMAL STOPPING PROBLEM FOR JUMP DIFFUSIONS\*

ERHAN BAYRAKTAR<sup>†</sup> AND HAO XING<sup>‡</sup>

**Abstract.** The value function of an optimal stopping problem for jump diffusions is known to be a generalized solution of a variational inequality. Assuming that 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 this optimal stopping problem on an unbounded domain with finite/infinite variation jumps is in  $W_{p,loc}^{2,1}$  with  $p \in (1, \infty)$ . As a consequence, the smooth-fit property holds.

**Key words.** optimal stopping, variational inequality, Lévy processes, regularity of the value function, smooth-fit principle, Sobolev spaces

**AMS subject classifications.** 60G40, 60J75, 35R35, 45K05

**DOI.** 10.1137/100810915

**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:

$$(1.1) \quad dX_t = b(X_t, t) dt + \sigma(X_t, t) dW_t + \int_{\mathbb{R}} h(X_{t-}, y, t) (N(dt, dy) - 1_{\{|y| \leq 1\}} dt \nu(dy)),$$

in which  $W = \{W_t; t \geq 0\}$  is a one-dimensional Wiener process;  $N$ , independent of the Wiener process, is a Poisson random measure on  $\mathbb{R}_+ \times \mathbb{R}$  with its mean measure  $dt \times \nu(dy)$ ; and  $\nu$  is a Lévy measure on  $\mathbb{R}$ . The coexistence of diffusions and infinite activity jumps is motivated by the recent studies of Aït-Sahalia and Jacod in [1] and [2].

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

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

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  $X$  models the logarithm of the stock price process and  $g$  represents the payoff function.

\*Received by the editors October 7, 2010; accepted for publication (in revised form) March 15, 2012; published electronically May 31, 2012.

<http://www.siam.org/journals/sicon/50-3/81091.html>

<sup>†</sup>Department of Mathematics, University of Michigan, Ann Arbor, MI 48109 (erhan@umich.edu). This author is supported in part by the National Science Foundation under applied mathematics research grants DMS-0906257, DMS-1118673, Career grant DMS-095546. and in part by the Susan M. Smith Professorship.

<sup>‡</sup>Department of Statistics, London School of Economics, London WC2A 2AE, UK (h.xing@lse.ac.uk). This author is supported in part by STICERD at London School of Economics.

The value function  $u$  is expected to satisfy a variational inequality with a non-local integral term (see, e.g., Chapter 3 of [7]). Different concepts of solutions were employed to characterize the value function: Pham used the notion of viscosity solution in [23]. Also see [3], [4] for more recent results in this direction. Lamberton and Mikou worked with Lévy processes and showed in [19] that the value function can be understood in the distributional sense.

When the diffusion component in  $X$  is nondegenerate, the value function is expected to have a higher degree of regularity. Sections 1–3 in Chapter 3 of [7], as well as [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 [20]. The intuition is that the diffusions component dominates the contribution from jumps in determining the regularity of solutions, regardless of whether jumps have finite variation or not. However, this intuition is only a folklore theorem for obstacle problems. There are some limited results available whose assumptions on 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 Theorem 3.2 in [7, p. 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 to infinite activity with finite variation cases. Zhang studied in [27] an obstacle problem for a jump diffusion with finite active jumps. Also see [22], [26], [5], and [6] for further developments. More recently, Davis, Guo, and Wu [11], generalizing the results in [17] for the diffusion case, analyzed an impulse control problem 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 allow for infinite activity, infinite variation 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 that it is an element in  $W_{p,loc}^{2,1}$  with  $p \in (1, \infty)$  (later in this section we define 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. We also develop a nonlocal version of the interior Schauder estimate in Proposition 3.5, which could be useful for studying other integro-differential equations with irregular initial conditions.

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 analyzed in section 3. Then main results are proved in section 4.

**1.1. Notation.** For a given open interval  $D = (\ell, r)$  with  $-\infty \leq \ell < r \leq \infty$ , let us define the  $\delta$ -neighborhood of  $D$  as  $D^\delta := (\ell - \delta, r + \delta)$  for  $\delta > 0$ . We will also denote  $D_s := D \times (0, s)$ ,  $D_s^\delta := D^\delta \times (0, s)$  for any  $s > 0$ ,  $E_s := \mathbb{R} \times [0, s]$ , and by  $\bar{A}$  denote the closure of the indicated set  $A$ . Let us recall definitions of Sobolev spaces and Hölder spaces in what follows; see [18, 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 first and second order, respectively.

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

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

## 2. Main results.

**2.1. Model.** Let us first specify the jump diffusion  $X$  in (1.1). We assume that the drift and the volatility of  $X$ , the discounting factor  $r$ , and the jump size  $h$  satisfy the following set of assumptions.

*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  $\lambda$  such that  $a(x, t) \geq \lambda$  for all  $(x, t) \in E_T$ . The jump size  $h(x, y, t)$  is continuously differentiable in  $x$ , and  $\partial_x h(x, y, t)$  is Hölder continuous in  $(x, t)$ ; moreover, there exists a constant  $C$  such that

$$(2.1) \quad |h(x, y, t)| \leq C|y| \text{ and } |h(x_1, y, t_1) - h(x_2, y, t_2)| \leq C|y|(|x_1 - x_2| + |t_1 - t_2|^{\frac{1}{2}})$$

for  $(x, t), (x_i, t_i) \in \mathbb{R} \times [0, T]$ ,  $i = 1$  or  $2$ , and  $y \in \mathbb{R}$ .

Without loss of generality, we will take  $C$  in (2.1) to be equal to one; otherwise we would rescale the process  $X$ . For the pure jump component in (1.1), we assume that  $\nu$  is a Lévy measure on  $\mathbb{R}$ . See [24] for this terminology. In particular, we require that  $\int_{\mathbb{R}} (y^2 \wedge 1) \nu(dy) < \infty$ . When  $h \equiv y$ , the jump component of (1.1) is a Lévy process. The aforementioned assumptions on coefficients and the jump component ensure that (1.1) admits a unique strong solution (see [16]), which we denote by  $X$ . This jump diffusion process  $X$  is said to have *finite activity* if  $\nu$  is a finite measure on  $\mathbb{R}$ ; otherwise it is said to have *infinite activity*. We say that the jumps of  $X$  have *finite variation* if  $\int_{\mathbb{R}} |y| \nu(dy) < \infty$ ; otherwise we say that they have *infinite variation*.

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

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

Note that the interval  $|y| \leq 1$  can be replaced by any other neighborhood of 0 in our analysis; our choice of this interval is made for notational convenience.

*Remark 2.3.* Virtually all Lévy processes used in financial modeling satisfy the above assumption. For jump diffusions models,  $\nu$  is a finite measure as in Merton's and Kou's models. For normal tempered stable processes,  $\rho$  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, Carr–Geman–Madan–Yor (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 set of assumptions.

**Assumption 2.4.** The payoff function  $g$  is positive, bounded, and Lipschitz continuous on  $\mathbb{R}$ . That is, 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  $\partial_{xx}^2 g \geq -J$  for some positive constant  $J$  in the distributional sense; i.e.,  $\int_{\mathbb{R}} g(x) \partial_{xx}^2 \phi(x) dx \geq -J \int_{\mathbb{R}} \phi(x) dx$  for any compactly supported smooth function  $\phi$  on  $\mathbb{R}$ .

A typical example, where these assumptions hold, is the American put option payoff  $g(x) = (K - e^x)_+$  for some  $K \in \mathbb{R}_+$ .

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

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

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

$$(2.2) \quad \begin{aligned} \min \{(-\partial_t - \mathcal{L} + r)u, u - g\} &= 0, & (x, t) &\in \mathbb{R} \times [0, T), \\ u(x, T) &= g(x), & x &\in \mathbb{R}. \end{aligned}$$

Here, the integro-differential operator  $\mathcal{L}$  is the infinitesimal generator of  $X$ . Its application on a smooth test function  $\phi$  is

$$(2.3) \quad \mathcal{L}\phi := \mathcal{L}_D\phi + I\phi,$$

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

$$(2.4) \quad I\phi(x, t) := \int_{\mathbb{R}} [\phi(x + h(x, y, t), t) - \phi(x, t) - h(x, y, t) \partial_x \phi(x, t) 1_{\{|y| \leq 1\}}] \nu(dy).$$

In what follows we will not write the arguments of  $h$  explicitly but only indicate the argument on which we are focusing in order to keep the notation simple.

In general, one does not know a priori whether  $u$  is sufficiently smooth so that it solves (2.2) in the classical sense. Moreover, 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 derivative  $\partial_x \phi(\cdot, t)$  in a neighborhood of  $x$ , it can be shown 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

$$(2.5) \quad \begin{aligned} I^\epsilon \phi(x, t) &:= \int_{|y| > \epsilon} [\phi(x + h, t) - \phi(x, t)] \nu(dy) - \partial_x \phi(x, t) \int_{\epsilon < |y| \leq 1} h \nu(dy) \\ &\leq C \int_{|y| > \epsilon} |y| \nu(dy) + |\partial_x \phi(x, t)| \int_{\epsilon < |y| \leq 1} |y| \nu(dy), \\ I_\epsilon \phi(x, t) &:= \int_{|y| \leq \epsilon} [\phi(x + h, t) - \phi(x, t) - h \partial_x \phi(x, t)] \nu(dy) \\ &= \int_{|y| \leq \epsilon} h (\partial_x \phi(x + h, t) - \partial_x \phi(x, t)) \nu(dy) \leq C \int_{|y| \leq \epsilon} y^2 \nu(dy). \end{aligned}$$

Here, the first inequality follows from the Lipschitz continuity of  $\phi(\cdot, t)$  and the assumption that  $|h| \leq |y|$ ; the mean value theorem implies the second equality in (2.5) where  $z$  satisfies  $|z - x| < |h|$ ; the last inequality holds due to the Lipschitz continuity of  $\partial_x \phi(\cdot, t)$  in an  $\epsilon$ -neighborhood of  $x$ . However, the value function  $u$ , in general, does not have the regularity properties mentioned above. We only know from Lemma 3.1 in [23] that  $u$  is Lipschitz continuous in  $x$  and  $1/2$ -Hölder continuous in  $t$ . Nevertheless, we will see that the integral term  $Iu$  is well defined in the classical sense in Lemma 3.2 below. 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  $u \in W_{p,loc}^{2,1}(\mathbb{R} \times (0, T))$  for any integer  $p \in (1, \infty)$ . Moreover,  $u$  solves (2.2) 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 so that its integral component has a reduced form. For any test function  $\phi$  that is Lipschitz continuous in its first variable,  $\mathcal{L}\phi$  can be decomposed as  $\mathcal{L}\phi = \mathcal{L}_D^f \phi + I^f \phi$ , in which  $\mathcal{L}_D^f \phi = a \partial_{xx}^2 \phi + [b - \int_{|y| \leq 1} h \nu(dy)] \partial_x \phi$  and

$$(2.6) \quad I^f \phi(x, t) := \int_{\mathbb{R}} [\phi(x + h, t) - \phi(x, t)] \nu(dy).$$

The previous integral is clearly well defined. Indeed,  $|I^f \phi(x, t)| \leq C \int_{\mathbb{R}} |y| \nu(dy) < +\infty$  follows from the Lipschitz continuity of  $\phi(\cdot, t)$  and  $|h| \leq |y|$ . Moreover,  $I^f \phi$  is also Hölder continuous in both of its variables; 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 [23]),  $I^f u$  is already well defined and Hölder continuous. Therefore, in order to study the regularity of  $u$ ,  $I^f u$  can be treated as a driving term in (2.2). 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,  $\nu$  has a singularity at  $y = 0$ . As a result,  $I$  maps functions with a certain degree of regularity to functions with less regularity. This is in contrast to the case in which  $\nu$  is a finite measure. In that case,  $\int_{\mathbb{R}} \phi(x + h, t) \nu(dy)$  is already well defined for any  $\phi$  with at most linear growth, and this integral has the same regularity as  $\phi$ ; see [26]. Second,  $I$  is a nonlocal operator. Therefore, regularity of  $I\phi$  on a given interval  $D$  depends on  $\phi$  outside  $D$ . In this subsection, we shall study these two features in detail and analyze the regularity of  $I\phi$  when  $\phi$  is a function in either certain Hölder spaces 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 (2.6). It has the following regularity property.

**LEMMA 3.1.** *Let Assumption 2.2 hold with  $0 \leq \alpha < 1$  and  $s > 0$ . For any  $\phi$  which is Lipschitz continuous in its first variable and  $1/2$ -Hölder continuous in its*

second variable,

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

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 forms of (2.3) and (2.4). We will see that if we choose an appropriate test function  $\phi$ ,  $I\phi$  is still well defined and Hölder continuous in both its variables. Regularity estimates of the following type have been obtained in [25] and [21].

LEMMA 3.2. *Let Assumption 2.2 hold with  $\alpha \in [1, 2)$  and  $s > 0$ .*

- (i) *Suppose that  $\phi$  satisfies  $|\phi(x_1, t_1) - \phi(x_2, t_2)| \leq L(|x_1 - x_2| + |t_1 - t_2|^{\frac{1}{2}})$  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  $I\phi \in H^{\beta-\alpha-\gamma, \frac{\beta-\alpha-\gamma}{2}}(\overline{D_s})$  and*

$$(3.1) \quad \|I\phi\|_{\overline{D_s}}^{(\beta-\alpha-\gamma)} \leq C \left( L + \|\phi\|_{\overline{D_s^1}}^{(\beta)} \right)$$

*for a positive constant  $C$  depending on  $D$ ,  $\alpha$ , and  $\beta$ .*

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

$$(3.2) \quad \|I\phi\|_{E_s}^{(\beta-\alpha-\gamma)} \leq C \|\phi\|_{E_s}^{(\beta)}$$

*for a positive constant  $C$  depending on  $\alpha$  and  $\beta$ .*

Here  $\gamma = 0$  when  $\alpha \in (1, 2)$ ;  $\gamma$  is an arbitrary number in  $(0, \beta - \alpha)$  when  $\alpha = 1$ .

Since the proofs of Lemmas 3.1 and 3.2 are similar, we present only 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,  $\|\cdot\|_{E_s}^{(\beta)} \geq L$ ; see Definition 1.1. It then suffices to prove statement (i). For notational simplicity,  $C$  represents a generic constant throughout the rest of proof.

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

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

where the second inequality follows from the mean value theorem with  $|z - x| \leq |h| \leq |y| \leq 1$ ; the third inequality is the result of the  $(\beta - 1)$ -Hölder continuity of  $\partial_x \phi$  on  $\overline{D_s^1}$  and  $|h| \leq |y|$ ; 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:

$$|I\phi(x_1, t) - I\phi(x_2, t)| \leq I_1 + I_2 + I_3, \quad \text{in which}$$

$$I_1(x, t) := \int_{|y| \leq \epsilon} [|\phi(x_1 + h(x_1), t) - \phi(x_1, t) - h(x_1) \partial_x \phi(x_1, t)| \\ + |\phi(x_2 + h(x_2), t) - \phi(x_2, t) - h(x_2) \partial_x \phi(x_2, t)|] \nu(dy),$$

$$I_2(x, t) := \int_{\epsilon < |y| \leq 1} [|\phi(x_1 + h(x_1), t) - \phi(x_1, t) - \phi(x_2 + h(x_2), t) + \phi(x_2, t)| \\ + |h(x_1) \partial_x \phi(x_1, t) - h(x_2) \partial_x \phi(x_2, t)|] \nu(dy),$$

$$I_3(x, t) := \int_{|y| > 1} [|\phi(x_1 + h(x_1), t) - \phi(x_2 + h(x_2), t)| + |\phi(x_1, t) - \phi(x_2, t)|] \nu(dy),$$

where variables  $y$  and  $t$  are ignored in  $h$  and 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  $I_1 \leq 2 \|\phi\|_{\frac{\beta}{D_s}}^{(\beta)} \int_{|y| \leq \epsilon} |y|^\beta \nu(dy) = C \|\phi\|_{\frac{\beta}{D_s}}^{(\beta)} \epsilon^{\beta-\alpha}$ . Second, it follows from the Lipschitz continuity of  $\phi(\cdot, t)$ ,  $(\beta - 1)$ -Hölder continuity of  $\partial_x \phi$ , and  $|h| \leq |y|$  that

$$\begin{aligned} & |\phi(x_1 + h(x_1), t) - \phi(x_1, t) - \phi(x_2 + h(x_2), t) + \phi(x_2, t)| \\ & \leq |\phi(x_1 + h(x_1), t) - \phi(x_1, t) - \phi(x_2 + h(x_1), t) \\ & \quad + \phi(x_2, t)| + |\phi(x_2 + h(x_1), t) - \phi(x_2 + h(x_2), t)| \\ & \leq \left| \int_0^{h(x_1)} |\partial_x \phi(x_1 + z, t) - \partial_x \phi(x_2 + z, t)| dz \right| + C|x_1 - x_2| |y| \\ & \leq \|\phi\|_{\frac{\beta}{D_s}}^{(\beta)} |x_1 - x_2|^{\beta-1} |y| + C|x_1 - x_2| |y|. \end{aligned}$$

Similarly,  $|h(x_1) \partial_x \phi(x_1, t) - h(x_2) \partial_x \phi(x_2, t)| \leq \|\phi\|_{\frac{\beta}{D_s}}^{(\beta)} |y| (|x_1 - x_2| + |x_1 - x_2|^{\beta-1})$ . Therefore,

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

where the second inequality follows from Assumption 2.2. Third, it is clear from the Lipschitz continuity of  $\phi$  and (2.1) that  $I_3 \leq C|x_1 - x_2| \int_{|y| > 1} (1 + |y|) \nu(dy)$ .

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

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

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  $|\partial_x \phi(x, t_1) - \partial_x \phi(x, t_2)| \leq \|\phi\|_{\frac{\beta}{D_s}}^{(\beta)} |t_1 - t_2|^{\frac{\beta-1}{2}}$  (see Definition 1.1), together with

semi-Hölder continuity of  $h(x, \cdot)$  in (2.1), and choosing  $\epsilon = |t_1 - t_2|^{\frac{1}{2}} \wedge 1$ , we can obtain

$$|I\phi(x, t_1) - I\phi(x, t_2)| \leq C \left(1 + \|\phi\|_{\frac{(\beta)}{D_s^1}}\right) |t_1 - t_2|^{\frac{\beta - \alpha - \gamma}{2}}$$

for a constant  $C$  independent of  $x$ ,  $t_1$ , and  $t_2$ .  $\square$

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

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

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

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

Remark 3.4. When  $X$  has finite variation jumps, i.e.,  $0 \leq \alpha < 1$ ,  $\eta$  in (3.3) can be chosen as zero. Hence the  $L_p$ -norm of  $I\phi$  depends only on  $\max_{\mathbb{R} \times [t_1, t_2]} |\phi|$  and  $\max_{D^1 \times [t_1, t_2]} |\partial_x \phi|$ .

Proof. Utilizing truncation and smooth mollification, one can construct a sequence of smooth functions  $(\phi^\epsilon)_{\epsilon>0}$  such that  $\phi^\epsilon$  converges to  $\phi$  in  $W_p^{2,1}$  and  $I\phi^\epsilon$  converges to  $I\phi$  in  $L_p$  as  $\epsilon \rightarrow 0$  (cf. section 5.3 and Appendix C.4 in [12]). Therefore, it suffices to prove the statement for a smooth function  $\phi$ .

To this end, observing that  $\phi(x+h, t) - \phi(x, t) - h\partial_x \phi(x, t) = h^2 \int_0^1 (1-z) \partial_{xx}^2 \phi(x + zh, t) dz$ , the integral  $I\phi$  can be bounded above by three integral terms:

$$\begin{aligned} |I\phi(x, t)| &\leq \int_{|y| \leq \eta} h^2 \nu(dy) \int_0^1 dz |\partial_{xx}^2 \phi(x + zh, t)| \\ &+ \int_{\eta < |y| \leq 1} \nu(dy) |\phi(x + h, t) - \phi(x, t) - h\partial_x \phi(x, t)| \\ &+ \int_{|y| > 1} \nu(dy) |\phi(x + h, t) - \phi(x, t)| \\ &=: I_1 + I_2 + I_3. \end{aligned}$$

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

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



where the first inequality follows from Fubini's theorem and Jensen's inequality since  $p > 1$ , the second inequality is a result of the assumption that  $|h| \leq |y|$  and Assumption 2.2, the third inequality follows from Hölder inequality with  $1/p + 1/q = 1$ , the fourth inequality utilizes Fubini's theorem, and the fifth inequality holds since  $x + zh \in D^\eta$  for any  $|h| \leq |y| \leq \eta$  and  $z \in [0, 1]$ . Second, since  $x + h \in D^1$  for  $x \in D$ , and  $|h| \leq |y| \leq 1$ , it follows that

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

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

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

**3.2. An interior estimate.** The  $L_p$ -norm estimate of the integral term in Lemma 3.3 helps us 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 (cf. Theorem 9.1 in [18, p. 341]).

**PROPOSITION 3.5.** *Suppose that Assumptions 2.1 and 2.2 are satisfied. Let  $v$  be a  $W_{p,loc}^{2,1}$ -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  $C_\delta$ , depending on  $\delta$ , such that

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

**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 for local PDEs. However, as we have seen in Lemma 3.3, the  $W_p^{2,1}$ -norm of  $v$  controls the  $L_p$ -norm of  $Iv$ , which in turn bounds the  $W_p^{2,1}$ -norm of  $v$  via the Calderon–Zygmund estimate. Therefore, a careful balance between extending domains and controlling the  $W_p^{2,1}$ -norm of  $v$  needs to be maintained in the following proof. This is in 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 the case of finite variation jumps, (3.4) can be obtained directly from the classical Calderon–Zygmund estimate for local PDEs.

*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  $\zeta^\delta$  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,  $\zeta^\delta$  can be chosen to satisfy

$$(3.5) \quad |\partial_x \zeta^\delta| \leq \frac{C}{\delta}, \quad |\partial_{xx}^2 \zeta^\delta| \leq \frac{C}{\delta^2}, \quad \text{and} \quad |\partial_t \zeta^\delta| \leq \frac{C}{\delta}.$$

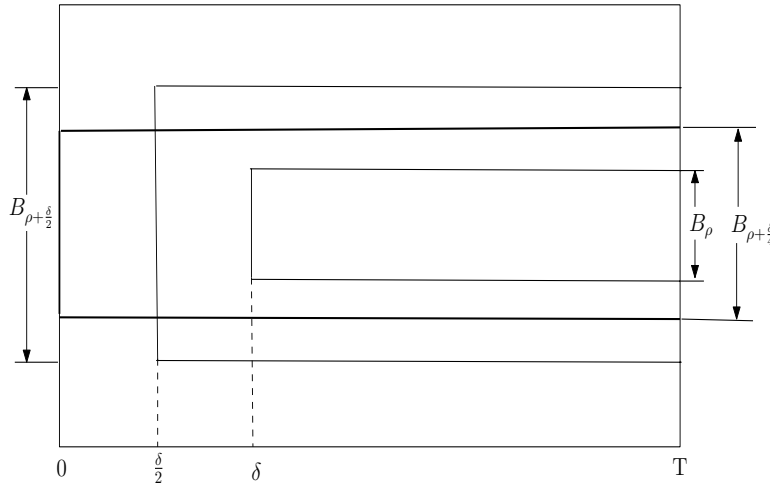


FIG. 1. Domains used in the proof of Proposition 3.5.

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

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

in which  $h := \partial_t \zeta^\delta v - a(\partial_{xx}^2 \zeta^\delta v + 2 \partial_x \zeta^\delta \partial_x v) - b \partial_x \zeta^\delta v$ . Appealing to Theorem 9.1 in [18, p. 341], we can find a constant  $C$  such that

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

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

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

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

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

follows from (3.5) that

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

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

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

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

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

Multiplying  $\delta^2$  on both sides of the previous inequality, we have

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

where

$$K(\delta) = C(\delta^2 + \delta^{3-\alpha} + \delta^{\frac{1+p}{p}} + \delta^{\frac{1}{p}}) \left( \max_{E_T} |v| + \max_{D^{\delta/4+1} \times [0, T]} |\partial_x v| \right) + \delta^2 \|f\|_{L_p(D^{\delta/4} \times (\frac{\delta}{2}, T))}.$$

Denote  $F(\tau) := \tau^2 \|v\|_{W_p^{2,1}(D^{\delta-\tau} \times (\tau, T))}$ . 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  $4C(\frac{\delta}{4})^{2-\alpha} \leq \frac{1}{2}$ , we obtain from the above inequality that

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

Note that  $F(\delta)$  is finite for any  $\delta$ , since the  $W_p^{2,1}$ -norm of  $v$  is finite in any compact domain of  $\mathbb{R} \times (0, T)$ , and  $K(\delta)$  is increasing in  $\delta$ . We then obtain from iterating the previous 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  $W_{p,loc}^{2,1}$ -norms, the previous inequality reads

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

#### 4. Proof of main results.

**4.1. The penalty method.** We use the penalty method (see, e.g., [14] and [26]) to analyze the following variational inequality:

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

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 (2.6); see Remark 2.7. Then all proofs are similar to but easier than those in the infinite variation case.

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

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

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

$$(4.3) \quad \begin{aligned} &\text{(i) } p_\epsilon(y) \leq 0, \quad \text{(ii) } p_\epsilon(y) = 0 \text{ for } y \geq \epsilon, \\ &\text{(iii) } p_\epsilon(0) = -a^{(0)}J - |b|^{(0)}L - r^{(0)}K - J \int_{|y| \leq 1} |y|^2 \nu(dy) - K \int_{|y| > 1} \nu(dy), \\ &\text{(iv) } p'_\epsilon(y) \geq 0, \quad \text{(v) } p''_\epsilon(y) \leq 0, \quad \text{and} \quad \text{(vi) } \lim_{\epsilon \downarrow 0} p_\epsilon(y) = \begin{cases} 0, & y > 0, \\ -\infty, & y < 0, \end{cases} \end{aligned}$$

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,  $p_\epsilon$  can be chosen as a smooth mollification of the function  $\min\{-2p_\epsilon(0)x/\epsilon + p_\epsilon(0), 0\}$ .

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, p. 189]).

**LEMMA 4.1.** *Let  $\Theta$  be a closed convex subset of a Banach space, and let  $\mathcal{T}$  be a continuous operator on  $\Theta$  such that  $\mathcal{T}\Theta$  is contained in  $\Theta$  and  $\mathcal{T}\Theta$  is precompact. Then  $\mathcal{T}$  has a fixed point in  $\Theta$ .*

**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  $v^\epsilon \in H^{2+\beta-\alpha-\gamma, 1+\frac{\beta-\alpha-\gamma}{2}}(E_T)$ . Here  $\gamma = 0$  when  $1 < \alpha < 2$ ;  $\gamma$  is an arbitrary number in  $(0, \beta - \alpha)$  when  $\alpha = 1$ .*

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

Let us consider the set  $\Theta := \{v \in H^{\beta, \frac{\beta}{2}}(E_s)$  with its Hölder norm  $\|v\|_{E_s}^{(\beta)} \leq U_0\}$ , where  $s$  and  $U_0$  will be determined later. It is clear that  $\Theta$  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  $u - g^\epsilon$ :

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

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

*Step 1:  $\mathcal{T}v$  is well defined.* Since  $v \in H^{\beta, \frac{\beta}{2}}(E_s)$  with  $\beta \in (\alpha, 2)$ , Lemma 3.2(ii) implies that  $Iv \in H^{\beta-\alpha-\gamma, \frac{\beta-\alpha-\gamma}{2}}(E_s)$  with  $\|Iv\|_{E_s}^{(\beta-\alpha-\gamma)} \leq C\|v\|_{E_s}^{(\beta)}$ . On the other hand, using properties of  $v$ ,  $g^\epsilon$ , and  $p_\epsilon$ , one can check that  $-p_\epsilon(v - g^\epsilon) + (\mathcal{L}_D - r)g^\epsilon \in H^{\beta-\alpha-\gamma, \frac{\beta-\alpha-\gamma}{2}}(E_s)$ . Therefore, Theorem 5.1 in [18, p. 320] implies that (4.4) has a unique solution  $u - g^\epsilon \in H^{2+\beta-\alpha-\gamma, 1+\frac{\beta-\alpha-\gamma}{2}}(E_s)$ . Hence  $u = \mathcal{T}v \in H^{2+\beta-\alpha-\gamma, 1+\frac{\beta-\alpha-\gamma}{2}}(E_s)$ , since  $g^\epsilon$  is smooth.

*Step 2:  $\mathcal{T}\Theta \subset \Theta$ .* It follows from Lemma 2 in [13, p. 193] that there exists a positive constant  $A_\beta$ , depending on  $\beta$ , such that

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

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

$$(4.6) \quad \|u\|_{E_s}^{(\beta)} \leq \|u - g^\epsilon\|_{E_s}^{(\beta)} + \|g^\epsilon\|_{E_s}^{(\beta)} \leq \tau U_0 + \tilde{A} + \frac{U_0}{2} \leq \tau U_0 + \frac{1-2\tau}{2} U_0 + \frac{U_0}{2} = U_0.$$

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

*Step 3:  $\mathcal{T}\Theta$  is a precompact subset of  $H^{\beta, \frac{\beta}{2}}(E_s)$ .* For any  $\eta \in (\beta, 2)$ , an estimate similar to (4.5) shows that for any  $v \in \Theta$ ,  $\|\mathcal{T}v\|_{E_s}^{(\eta)} \leq U_1$  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, p. 188]), then  $\mathcal{T}\Theta$  is a precompact subset in  $H^{\beta, \frac{\beta}{2}}(E_s)$ .

*Step 4:  $\mathcal{T}$  is a continuous operator.* Let  $v_n$  be a sequence in  $\Theta$  such that  $\lim_{n \rightarrow \infty} \|v_n - v\|_{E_s}^{(\beta)} = 0$ . We will show  $\lim_{n \rightarrow \infty} \|\mathcal{T}v_n - \mathcal{T}v\|_{E_s}^{(\beta)} = 0$ . From (4.4),  $w \triangleq \mathcal{T}v_n - \mathcal{T}v$  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, p. 193] that

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

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

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

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 in our analysis.

**LEMMA 4.3.** *Suppose that  $a > 0$ ,  $a$  and  $b$  are bounded, and the Levy measure  $\nu$  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  $(\partial_t - \mathcal{L}_D - I + c)v(x, t) \geq 0$  and  $v$  is bounded from below on  $E_T$ , then  $v(x, 0) \geq 0$  for  $x \in \mathbb{R}$  implies that  $v \geq 0$  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 function

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

where  $C_1$  will be determined later and  $f: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  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  $\lim_{R \rightarrow +\infty} f(R) = \infty$  and derivatives  $f'$  and  $f''$  are bounded. Then  $If(|x|)$  is bounded on  $\mathbb{R}$ . Indeed, there exists a constant  $C$  such that

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

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

$$\begin{aligned} &(\partial_t - \mathcal{L}_D - I + c)w \\ &= e^{C_0 t} \frac{m}{f(R_0)} \cdot [C_1 + (c + C_0)(f(|x|) + C_1 t) - a \partial_{xx}^2 f(|x|) - b \partial_x f(|x|) - If(|x|)] > C_0 m \end{aligned}$$

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

$$(4.7) \quad (\partial_t - \mathcal{L}_D - I + c + C_0)\tilde{v} > C_0v + C_0m \geq 0 \quad \text{for any } (x, t) \in E_T.$$

On the other hand,  $\tilde{v}(x, 0) = \frac{m}{f(R_0)}f(|x|) + v(x, 0) \geq 0$  due to  $v(x, 0) \geq 0$ , and 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 take 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  $E_T$ , and hence  $I\tilde{v}(x_0, t_0) \geq 0$ ,  $\partial_t \tilde{v}(x_0, t_0) \leq 0$ ,  $\partial_x \tilde{v}(x_0, t_0) = 0$ ,  $\partial_{xx}^2 \tilde{v}(x_0, t_0) \geq 0$ , 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  $\infty$ .  $\square$

This maximum principle implies the uniqueness of classical solutions for the penalty problem (4.2).

**COROLLARY 4.4.** *Under the assumptions of Lemma 4.2,  $v^\epsilon$  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 (4.3)(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.  $\square$

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

**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  $v^\epsilon$  is bounded on  $E_T$  for each  $\epsilon \in (0, 1)$ . In this proof, we will show that the bounds are uniform in  $\epsilon$ . First, it follows from (4.3)(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 the first inequality in the statement follows directly from Lemma 4.3. 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)(ii) and  $g^\epsilon \leq K$ , we have  $p_\epsilon(K + 1 - g^\epsilon) = 0$ . Hence,

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



where the first equality follows from the mean value theorem. Now applying Lemma 4.3 to the above equation with  $c = r + p'(y) \geq 0$  (see (4.3)(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.  $\square$

LEMMA 4.6. *Let Assumptions 2.1, 2.2, with  $1 \leq \alpha < 2$ , and 2.4 hold. Then for any  $\epsilon \in (0, 1)$ ,*

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

in which  $C$  depends on  $T$  and  $L$ .

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

$$\begin{aligned} (4.9) \quad & \left[ \partial_t - a \partial_{xx}^2 - (b + \partial_x a) \partial_x - \hat{I} + \left( r - \partial_x b - \int_{|y|>1} \partial_x h \nu(dy) \right) \right] w \\ & + v^\epsilon \partial_x r + p'_\epsilon (v^\epsilon - g^\epsilon) (w - (g^\epsilon)') = 0, \\ & (x, t) \in \mathbb{R} \times (0, T], \\ w(x, 0) &= (g^\epsilon)'(x), \quad x \in \mathbb{R}. \end{aligned}$$

Here  $\hat{I}\phi := I\phi + \int_{\mathbb{R}} [\phi(x+h, t) - \phi(x, t)] \partial_x h \nu(dy)$ , where the second integral is well defined for Lipschitz bounded function  $\phi$  because  $|\partial_x h| \leq |y|$  from (2.1) and  $\int_{|y|>1} |y| \nu(dy) < \infty$  from Assumption 2.2. We will show that  $\partial_x v^\epsilon$  is indeed a classical solution of (4.9). To this end, let us consider the equation

$$\begin{aligned} & \left[ \partial_t - a \partial_{xx}^2 - (b + \partial_x a) \partial_x - \hat{I} + \left( r - \partial_x b - \int_{|y|>1} \partial_x h \nu(dy) \right) \right] w \\ & = -\partial_x r v^\epsilon - p'_\epsilon (v^\epsilon - g^\epsilon) (\partial_x v^\epsilon - (g^\epsilon)'), \\ & (x, t) \in \mathbb{R} \times (0, T], \\ w(x, 0) &= (g^\epsilon)'(x), \quad x \in \mathbb{R}. \end{aligned}$$

Using Assumption 2.1 and Lemma 4.2, one can check that the driving term  $-\partial_x r v^\epsilon - p'_\epsilon (v^\epsilon - g^\epsilon) (\partial_x v^\epsilon - (g^\epsilon)')$  and all coefficients of the previous equation are Hölder continuous. It then follows from Theorem 3.1 in [15, p. 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 straightforward to check that  $v$  is a classical solution of the following equation:

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

Since  $g^\epsilon$  and  $v^\epsilon$  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, p. 89] implies that  $v$  is a bounded solution of the last equation. However, Corollary 4.4 already shows that  $v^\epsilon$  is the unique bounded solution of the last equation, therefore  $v = v^\epsilon$ , and 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  $\epsilon$ . Consider  $\tilde{v} = e^{\gamma t} L + \partial_x v^\epsilon$ , where  $L$  is given by Assumption 2.4 and  $\gamma > 0$  will be determined later. The function

$\tilde{v}$  satisfies the following equation:

$$(4.10) \quad \left[ \partial_t - a \partial_{xx}^2 - (b + \partial_x a) \partial_x - \hat{I} + r - \partial_x b - \int_{|y|>1} \partial_x h \nu(dy) + p'_\epsilon(v^\epsilon - g^\epsilon) \right] \tilde{v} \\ = \left( \gamma + r - \partial_x b - \int_{|y|>1} \partial_x h \nu(dy) \right) e^{\gamma t} L - \partial_x r v^\epsilon + p'_\epsilon(v^\epsilon - g^\epsilon) (e^{\gamma t} L + (g^\epsilon)'), \\ (x, t) \in \mathbb{R} \times (0, T],$$

$$\tilde{v}(x, 0) = e^{\gamma t} L + (g^\epsilon)'(x), \quad x \in \mathbb{R}.$$

Recall  $\partial_x b$  and  $\partial_x r$  are bounded from Assumption 2.1. Observe that  $\int_{|y|>1} \partial_x h \nu(dy) < \infty$  because  $|\partial_x h| \leq |y|$  and  $\int_{|y|>1} |y| \nu(dy) < \infty$ . Moreover,  $v^\epsilon$  is bounded uniformly in  $\epsilon$  thanks to Lemma 4.5. Therefore, one can find a sufficiently large  $\gamma$ , independent of  $\epsilon$ , such that  $(\gamma + r - \partial_x b - \int_{|y|>1} \partial_x h \nu(dy)) e^{\gamma t} L - \partial_x r v^\epsilon > 0$ . On the other hand,  $p'_\epsilon(v^\epsilon - g^\epsilon)(e^{\gamma t} L + (g^\epsilon)')$  is also positive due to (4.3)(iv) and  $|(g^\epsilon)'| \leq L$ . As a result, the right-hand side of (4.10) is positive. Now since  $r - \partial_x b - \int_{|y|>1} \partial_x h \nu(dy) + p'_\epsilon(v^\epsilon - g^\epsilon)$  is bounded from below, and moreover, the Lévy measure  $(1 + \partial_x h) \nu(dy)$  associated with  $\hat{I}$  satisfies  $\int_{|y|>1} |y| (1 + |\partial_x h|) \nu(dy) \leq \int_{|y|>1} (|y| + |y|^2) \nu(dy) < \infty$  (see (2.1) and Assumption 2.2), we then have from Lemma 4.3 with  $I = \hat{I}$  that  $\tilde{v} \geq 0$  on  $E_T$ . Hence  $\partial_x v^\epsilon \geq -e^{\gamma T} L$  on  $E_T$  for some positive  $\gamma$  independent of  $\epsilon$ . The upper bound can be shown similarly by working with  $\tilde{v} = e^{\gamma t} L - \partial_x v^\epsilon$ .  $\square$

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  $Ig^\epsilon$  is uniformly bounded from below. Indeed,

$$Ig^\epsilon(x) = \int_{|y|\leq 1} \nu(dy) \int_0^1 dz (1-z) h^2 \partial_{xx}^2 g^\epsilon(x+zh) + \int_{|y|>1} [g^\epsilon(x+h) - g^\epsilon(x)] \nu(dy) \\ \geq -J \int_{|y|\leq 1} |y|^2 \nu(dy) - K \int_{|y|>1} \nu(dy),$$

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

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

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

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

The previous inequality and the mean value theorem combined imply that

$$(\partial_t - \mathcal{L}_D - I + r + p'_\epsilon(y))(v^\epsilon - g^\epsilon) \geq 0$$

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

**COROLLARY 4.8.** *Let the 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)(i), (iv) together 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  $\epsilon$ ; see (4.3)(iii).  $\square$

## 4.2. Proofs 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  $v^* \in W_p^{2,1}(B \times (s, T))$  for any integer  $p \in (1, \infty)$ , compact domain  $B \subset \mathbb{R}$ , and  $s \in [0, T)$ . Second, we confirm that  $u^*(x, t) := v^*(x, T - t)$  is the value function for the problem (OS).

*Step 1.* First, it follows from Lemma 4.2 that  $\partial_t v^\epsilon$ ,  $\partial_x v^\epsilon$ , and  $\partial_{xx}^2 v^\epsilon$  are continuous, and hence locally bounded on  $\mathbb{R} \times (0, T)$ . Therefore  $v^\epsilon \in W_{p,loc}^{2,1}(\mathbb{R} \times (0, T))$  for each  $\epsilon \in (0, 1)$ . Second, Lemmas 4.5 and 4.6 show that  $v^\epsilon$  and  $\partial_x v^\epsilon$  are bounded on  $E_T$ , uniformly in  $\epsilon$ . Moreover, the penalty term  $p_\epsilon(v^\epsilon - g^\epsilon)$  is also bounded uniformly in  $\epsilon$  due to Corollary 4.8. Therefore these boundedness properties and Proposition 3.5 with  $f = -p_\epsilon(v^\epsilon - g^\epsilon)$  together imply that

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

Thanks to the weak compactness of the Sobolev space  $W_p^{2,1}$ ,  $1 < p < \infty$ , we can then find a subsequence  $(\epsilon_k)_{k \geq 0}$  converging to zero and a function  $v^*$  such that  $v^{\epsilon_k} \rightharpoonup v^* \in W_p^{2,1}(B \times (s, T))$ . Here “ $\rightharpoonup$ ” represents the weak convergence; cf. Appendix D.4 in [12, p. 639]. In fact this convergence can be shown to be pointwise and uniform in the index. Indeed, (4.11) and the Sobolev embedding theorem (cf. Lemma 3.3 in [18, p. 80]) combined imply that

$$\|v^\epsilon\|_{B \times [s, T]}^{(\beta)} \leq C, \quad \text{where } \beta = 2 - \frac{3}{p} \text{ and } C \text{ is some constant independent of } \epsilon.$$

Choosing  $p > 1$  so that  $\beta > 0$  and using the previous uniform estimate along with the Arzelà–Ascoli theorem, we then find a further subsequence of  $(\epsilon_k)_{k \geq 0}$ , which is still denoted by  $(\epsilon_k)_{k \geq 0}$ , such that  $(v^{\epsilon_k})_{k \geq 0}$  converge to  $v^*$  uniformly on  $B \times [s, T]$ . Since each  $v^{\epsilon_k}$  is continuous,  $v^*$  is also continuous on  $B \times [s, T]$ .

Let us show that  $v^*$  solves (4.1). On the one hand, since  $p_\epsilon(v^{\epsilon_k} - g^{\epsilon_k}) \leq 0$ , we have  $(\partial_t - \mathcal{L}_D - I + r)v^{\epsilon_k} \geq 0$  for each  $\epsilon_k$ . Hence

$$\int (\partial_t - \mathcal{L}_D - I + r)v^* \phi \, dx dt = \lim_{\epsilon_k \rightarrow 0} \int (\partial_t - \mathcal{L}_D - I + r)v^{\epsilon_k} \phi \, dx dt \geq 0$$

for any compactly supported smooth function  $\phi$ . Here the identity above follows from applying the dual operator of  $\partial_t - \mathcal{L}_D - I + r$  to  $\phi$  and utilizing the dominated convergence theorem. The previous inequality then yields  $(\partial_t - \mathcal{L}_D - I + r)v^* \geq 0$  on  $B \times [s, T]$  in the distributional sense, which implies the same inequality on  $\mathbb{R} \times (0, T]$  in the distributional sense, since the choices of  $B$  and  $s$  are arbitrary. On the other hand, Lemma 4.7 shows that  $v^{\epsilon_k} \geq g^{\epsilon_k}$ . Then  $v^* \geq g$  after sending  $\epsilon_k \rightarrow 0$ . Therefore, we obtain  $\min\{(\partial_t - \mathcal{L}_D - I + r)v^*, v^* - g\} \geq 0$  on  $\mathbb{R} \times (0, T]$  in the distributional sense. 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. Utilizing the uniform

convergence of  $(v^{\epsilon_k})_{k \geq 0}$  and  $(g^{\epsilon_k})_{k \geq 0}$ , we can then find sufficiently small  $\epsilon_k$  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  $\epsilon_k \rightarrow 0$ , we conclude that  $(\partial_t - \mathcal{L}_D - I + r)v^* = 0$  in the distributional sense when  $v^* > g$ . Finally, since  $v^* \in W_{p,loc}^{2,1}$ ,  $v^*$  also solves (4.1) at 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 [23]. 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  $(\tilde{x}, \tilde{t}) \in \mathbb{R} \times (0, T)$ . The function  $v$  is said to be a viscosity solution of (4.1) if it is both a supersolution and a subsolution.

Let us show that  $v^*$  is a viscosity subsolution of (4.1). Fix  $(x, t) \in \mathbb{R} \times (0, T]$  and 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  $\epsilon_k$  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,  $(x_k, t_k) \rightarrow (x, t)$  as  $\epsilon_k \rightarrow 0$ . Since  $v^{\epsilon_k}$  is a classical solution of (4.2) (see Lemma 4.2), it is also a viscosity solution. Hence  $(\partial_t - \mathcal{L}_D - I + r)\phi(x_k, t_k) + p_{\epsilon_k}(v^{\epsilon_k}(x_k, t_k) - g^{\epsilon_k}(x_k)) \leq 0$ . Now, since  $v^*(x, t) > g(x)$  and  $v^{\epsilon_k}(x_k, t_k) - g(x_k)$  converge to  $v^*(x, t) - g(x)$ , we obtain  $\lim_{\epsilon_k \rightarrow 0} p_{\epsilon_k}(v^{\epsilon_k}(x_k, t_k) - g^{\epsilon_k}(x_k)) = 0$ . As a result,  $(\partial_t - \mathcal{L}_D - I + r)\phi(x, t) \leq 0$  by sending  $\epsilon_k \rightarrow 0$ . 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  $\phi$ . This follows from similar arguments which we used for the subsolution property in the previous paragraph.

Define  $u^*(x, t) = v^*(x, T - t)$ . It is clear that  $u^*$  is a viscosity solution of (2.2). Then the statement follows from Theorem 4.1 in [23], which states that the value function  $u$  is the unique viscosity solution of (2.2) when the Lévy measure satisfies  $\int_{|y|>1} |y|^2 \nu(dy) < \infty$ .  $\square$

*Proof of Corollary 2.6.* (i) Combining Theorem 2.5 and the Sobolev embedding theorem (cf. Lemma 3.3 in [18, p. 80]), we have  $u \in H^{\beta, \frac{\beta}{2}}(D \times [0, T - s])$ , where  $\beta = 2 - \frac{2}{p}$  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  $\beta > \alpha$ ,  $Iu \in H^{\beta-\alpha-\gamma, \frac{\beta-\alpha-\gamma}{2}}(\overline{D_{T-s}})$  by Lemma 3.2(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:

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

It is straightforward to show that  $u$  is the unique viscosity solution for the previous problem using the fact that  $u$  is the unique viscosity solution for (2.2). 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, p. 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  $\mathcal{C}$ .  $\square$

## REFERENCES

- [1] Y. AÏT-SAHALIA AND J. JACOD, *Estimating the degree of activity of jumps in high frequency data*, Ann. Statist., 37 (2009), pp. 2202–2244.
- [2] Y. AÏT-SAHALIA AND J. JACOD, *Is Brownian motion necessary to model high frequency data?*, Ann. Statist., 38 (2010), pp. 3093–3128.
- [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 J. Control Optim., 48 (2009), pp. 551–572.
- [6] E. BAYRAKTAR AND H. XING, *Analysis of the optimal exercise boundary of American options for jump diffusions*, SIAM J. Math. Anal., 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. LEVENDORSKIÏ, *Non-Gaussian Merton-Black-Scholes Theory*, Adv. Ser. Statist. Sci. Appl. Probab. 9, World Scientific, River Edge, NJ, 2002.
- [9] P. CARR, H. GEMAN, D. B. MADAN, AND M. YOR, *Stochastic volatility for Lévy processes*, Math. Finance, 13 (2003), pp. 345–382.
- [10] R. CONT AND P. TANKOV, *Financial Modelling with Jump Processes*, Chapman & Hall/CRC Financ. Math. Ser., Chapman & Hall/CRC, Boca Raton, FL, 2004.
- [11] M. H. A. DAVIS, X. GUO, AND G. WU, *Impulse control of multidimensional jump diffusions*, SIAM J. Control Optim., 48 (2010), pp. 5276–5293.
- [12] L. C. EVANS, *Partial Differential Equations*, Grad. Stud. Math. 19, American Mathematical Society, Providence, RI, 1998.
- [13] A. FRIEDMAN, *Partial Differential Equations of Parabolic Type*, Prentice-Hall, Englewood Cliffs, NJ, 1964.
- [14] A. FRIEDMAN, *Variational Principles and Free-Boundary Problems*, John Wiley, New York, 1982.
- [15] M. G. GARRONI AND J.-L. MENALDI, *Green Functions for Second Order Parabolic Integro-Differential Problems*, Pitman Res. Notes Math. Ser. 275, Longman Scientific & Technical, Harlow, UK, 1992.
- [16] I. GIHMAN AND A. V. SKOROHOD, *Stochastic Differential Equations*, Springer Verlag, New York, Heidelberg, 1972.
- [17] X. GUO AND G. WU, *Smooth fit principle for impulse control of multidimensional diffusion processes*, SIAM J. Control Optim., 48 (2009), pp. 594–617.
- [18] O. A. LADYŽENSKAJA, V. A. SOLONNIKOV, AND N. N. URALCHVA, *Linear and Quasi-Linear Equations of Parabolic Type*, American Mathematical Society, Providence, RI, 1968.
- [19] D. LAMBERTON AND M. MIKOU, *The critical price for the American put in an exponential Lévy model*, Finance Stoch., 12 (2008), pp. 561–581.
- [20] S. LENHART, *Integro-differential operators associated with diffusion processes with jumps*, Appl. Math. Optim., 9 (1982/83), pp. 177–191.
- [21] R. MIKULEVIČIUS AND C. ZHANG, *On the rate of convergence of weak Euler approximation for nondegenerate SDEs driven by Lévy processes*, Stochastic Process. Appl., 121 (2011), pp. 1720–1748.
- [22] H. PHAM, *Optimal stopping, free boundary, and American option in a jump-diffusion model*, Appl. Math. Optim., 35 (1997), pp. 145–164.
- [23] H. PHAM, *Optimal stopping of controlled jump diffusion processes: A viscosity solution approach*, J. Math. Systems Estim. Control, 8 (1998), pp. 1–27.

- [24] K. SATO, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge Stud. Adv. Math. 68, Cambridge University Press, Cambridge, UK, 1999.
- [25] L. SILVESTRE, *Regularity of the obstacle problem for a fractional power of the Laplace operator*, Comm. Pure Appl. Math., 60 (2006), pp. 67–112.
- [26] C. YANG, L. JIANG, AND B. BIAN, *Free boundary and American options in a jump-diffusion model*, European J. Appl. Math., 17 (2006), pp. 95–127.
- [27] X. L. ZHANG, *Méthodes Numériques pour le Calcul des Options Américaine dans des Modèles de diffusion avec sauts*, Ph.D. thesis, l'Ecole Nationale des Ponts et Chaussées, Paris, 1994.