Bull. Math. Sci.
https://doi.org/10.1007/s13373-018-0126-0

Maximum principles for nonlocal parabolic Waldenfels operators

Qiao Huang[1] · **Jinqiao Duan**[1,2] · **Jiang-Lun Wu**[3]

Received: 6 January 2018 / Revised: 18 April 2018 / Accepted: 5 May 2018

Abstract As a class of Lévy type Markov generators, nonlocal Waldenfels operators appear naturally in the context of investigating stochastic dynamics under Lévy fluctuations and constructing Markov processes with boundary conditions (in particular the construction with jumps). This work is devoted to prove the weak and strong maximum principles for 'parabolic' equations with nonlocal Waldenfels operators. Applications in stochastic differential equations with α-stable Lévy processes are presented to illustrate the maximum principles.

Keywords Nonlocal operators · Weak and strong maximum principles · Integro-partial differential equations · Waldenfels operators · Fokker–Planck equations · Stochastic differential equations with α-stable Lévy processes

Mathematics Subject Classification 35B50 · 35R09 · 47G20 · 60J75

Communicated by Ari Laptev.

✉ Jinqiao Duan
duan@iit.edu

Qiao Huang
hq932309@hust.edu.cn

Jiang-Lun Wu
j.l.wu@swansea.ac.uk

[1] Center for Mathematical Sciences, Huazhong University of Science and Technology, Wuhan 430074, Hubei, China

[2] Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616, USA

[3] Department of Mathematics, Swansea University, Singleton Park, Swansea SA2 8PP, UK

Published online: 16 May 2018

1 Introduction

The usual maximum principle concerns with second-order differential operators of elliptic or parabolic type. It is a basic property of solutions to boundary value problems for the associated elliptic or parabolic partial differential equations (PDEs) in a bounded domain. See [22,24]) for a general study of maximum principles. Classically, the maximum principle states that the maximum of the solution of a second-order elliptic or parabolic equation in a domain is to be found on the boundary of that domain. In particular, the strong maximum principle says that if the solution achieves its maximum in the interior of the domain, the solution must be a constant, while the weak maximum principle indicates that the maximum is to be found on the boundary but may re-occur in the interior as well. Let us also mention [19] where both weak and strong maximum principle for symmetric Markov generators are discussed via (local) Dirichlet forms. Moreover, a maximum principle for nonlocal operators generated by nonnegative kernels defined on topological groups acting continuously on a Hausdorff space was considered by Coville [7]. The strong maximum principle for semicontinuous viscosity solution of fully nonlinear second-order parabolic integro-differential equations was studied in [5].

A fairly large class of Markov processes on $\mathbb{R}^d$ are governed analytically by their infinitesimal generators, called Lévy type generators or pseudo-differential operators associated with negative definite symbols (cf. e.g. [11]), either via martingale problem (cf. e.g. [14–16,28,29]) or via Dirichlet form (cf. e.g. [9,11,17,18]). From [6,11], these operators are usually integro-differential operators or nonlocal operators, consisting of a combination of second-order elliptic differential operators and integral operators of Lévy type. The nonlocal operator here corresponds to the jump component of a Markov process; in fact, it is an integral with respect to a jump measure.

The well-known Hille-Yosida theorem and the semigroup approach, which can be found in e.g. [12], provide an intrinsic link between Markov processes and partial differential equations, in particular second-order elliptic differential operators, as in the pioneering work of Feller in early 1950s. The monograph [30] (also references therein) explores the functional analytic approach to constructing Markov processes in a prescribed region of $\mathbb{R}^d$, via the elliptic boundary value problems for the associated Lévy-type generators.

Due to the nature of pseudo-differential operators (involving integral operators), the Lévy-type generators are nonlocal operators. This kind of integro-differential operators was initiated by Waldenfels [32] in 1960s. It was elucidated in [30] that a Markov process associated with such an operator as infinitesimal operator could be interpreted with a physical picture: A Markovian particle moves both by jumps and continuously in a certain region of the state space $\mathbb{R}^d$.

The present paper is devoted to the weak and strong maximum principles for the following nonlocal parabolic Waldenfels operator $-\frac{\partial}{\partial t} + L$:

$$\begin{aligned}\Big(-\frac{\partial}{\partial t}+L\Big)u(x,t)&\\ :=-\frac{\partial u}{\partial t}(x,t)+\sum_{j,k=1}^{d}a_{jk}(x,t)\frac{\partial^2 u}{\partial x_j\partial x_k}(x,t)+\sum_{j=1}^{d}b_j(x,t)\frac{\partial u}{\partial x_j}(x,t)+c(x,t)u(x,t)&\\ +\int_{\mathbb{R}^d\setminus\{0\}}\Big[u(x+z,t)-u(x,t)-\sum_{j=1}^{d}z_j\frac{\partial u}{\partial x_j}(x,t)\mathbf{1}_{\{|z|<1\}}\Big]\nu(t,x,dz),&\end{aligned}$$

where the kernel $\{\nu(t,x,\cdot)\mid(x,t)\in\mathbb{R}^d\times[0,\infty)\}$ behaves as the jump measure for the associated Markov process. The operator L is called an elliptic Waldenfels operator. Note that Waldenfels operators L and $-\frac{\partial}{\partial t}+L$ appear in the generator and in the Fokker–Planck equation, respectively, for a stochastic differential equation with Lévy motions [3,8,26,27]. We would like to point out that Waldenfels operators also appear in nonlocal conservation laws [31]. Certain properties for diffusion generators perturbed by the nonlocal Laplacian operator have also been studied recently [1,2].

We will prove the new weak and strong maximum principles for the nonlocal parabolic operator $-\frac{\partial}{\partial t}+L$, and they do not require any "nondegeneracy" conditions. In order to cover the general case with either bounded or unbounded support of the jump measure ν, we will introduce two open sets D and E (with $D\subset E$), where D is the set where the maximum is achieved, and the stochastic process ("Markovian particle") cannot jump from D to the complement of E.

As a preparation for proving these maximum principles, we will prove the maximum principles for nonlocal elliptic Waldenfels operator L. These maximum principles are important for the construction of Markov processes. In [30, Appendix C], weak and strong maximum principles for such elliptic Waldenfels operators were proven, but under stringent conditions, that is, the jump measure has to have bounded support. The results in [5] includes a strong maximum principle for viscosity solutions of certain *nonlinear* nonlocal partial differential equations under a "nondegeneracy" condition.

The rest of this paper is organised as follows. In Sect. 2, we will present our results on maximum principles for elliptic Waldenfels operators. As a corollary, we also obtain the Hopf's Lemma about the sign of the gradient on the boundary. Section 3 is devoted to prove the maximum principles for parabolic Waldenfels operators. Some consequences and examples are presented in Sect. 4. Finally in Sect. 5, we present the proofs of some technical lemmas for the sake of completeness.

2 Maximum principles for elliptic Waldenfels operators

In this section, we consider the weak and strong maximum principles for the elliptic Waldenfels operator L (decomposed into local and nonlocal components)

$$L:=A+K, \tag{1}$$

where A and K are defined as

$$Au(x) := \sum_{j,k=1}^{d} a_{jk}(x)\frac{\partial^2 u}{\partial x_j \partial x_k}(x) + \sum_{j=1}^{d} b_j(x)\frac{\partial u}{\partial x_j}(x) + c(x)u(x),$$

$$Ku(x) := \int_{\mathbb{R}^d\setminus\{0\}} \Big[u(x+z) - u(x) - \sum_{j=1}^{d} z_j \frac{\partial u}{\partial x_j}(x)\mathbf{1}_{\{|z|<1\}}\Big]\nu(x,dz).$$

Note that the coefficients are taken to be independent of time t. Note that the operator K is actually the nonlocal Laplacian operator $-(-\Delta)^{\frac{\alpha}{2}}$, when the jump measure ν is the α-stable type; see [8, Ch. 7].

The elliptic Waldenfels operator L plays an important role [30] in the theory of Markov processes constructed in a given domain of $\mathbb{R}^d$. In that context, the second-order differential operator describes the diffusion part of the associated Markov process and the integral operator of Lévy type corresponds to the jump behavior of the Markov process. Finally, there is an assumption in that context which indicates that a Markovian particle cannot move by jumps from any interior point of certain domain to the outside of closure of the domain. For further remarks and discussions, we refer e.g. to Bony et al. [4] and Taira [30].

To cover more general situations, we introduce two open sets D and E in $\mathbb{R}^d$, with $D \subset E$ and E not necessarily bounded. As usual, we denote the boundary of D by ∂D, its closure by $\overline{D} := D \cup \partial D$ and its complement by $D^c := \mathbb{R}^d \setminus D$.

We make following assumptions:

1. Continuity condition: $a_{jk}, b_j, c \in C(\overline{E})$ $(j,k = 1,\ldots,d)$.
2. Symmetry condition: $a_{jk} = a_{kj}$ $(j,k = 1,\ldots,d)$. Uniform ellipticity condition: there exists a constant $\gamma > 0$ such that

$$\sum_{j,k=1}^{d} a_{jk}(x)\xi_j\xi_k \geq \gamma|\xi|^2, \tag{2}$$

   for all $x \in D$, $\xi \in \mathbb{R}^d$.
3. Lévy measures: The kernel $\{\nu(x,\cdot) \mid x \in \mathbb{R}^d\}$ is a family of Lévy measures, namely, each $\nu(x,\cdot)$ is a Borel measure on $\mathbb{R}^d \setminus \{0\}$ such that

$$\sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d\setminus\{0\}} (1 \wedge |z|^2)\nu(x,dz) < \infty, \tag{3}$$

   and moreover, for fixed $U \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$, the mapping $\mathbb{R}^d \ni x \to \nu(x,U) \in [0,\infty)$ is Borel measurable. Here we further assume that for each $x \in D$ the measure $\nu(x,\cdot)$ is supported in $\overline{E} - x := \{y - x \mid y \in \overline{E}\} = \{z \mid x + z \in \overline{E}\}$, i.e.,

$$\operatorname{supp}\nu(x,\cdot) \subset \overline{E} - x, \quad \forall x \in D. \tag{4}$$

Remark 2.1 The support condition (4) means in probability sense that a Markovian particle cannot move by jumps from a point $x \in D$ to the outside of $\overline{E}$. The motivation for this condition is that the maximizer point will propagate between connected components of the set in which the subsolution achieves maximum. The details will be discussed again in Remark 2.9 below. When the set E is the whole space $\mathbb{R}^d$, $E - x$ is still the whole space, and then there are actually no extra restrictions on the support of each measure $\nu(x, \cdot)$. In the case that $E = D$, the support condition is $\operatorname{supp} \nu(x, \cdot) \subset \overline{D} - x$, and this is related to the assumption in [30] that a Markovian particle cannot move by jumps from a point $x \in D$ to the outside of $\overline{D}$.

For convenience, the notation $\mathbf{a} = (a_{jk})_{j,k=1,\dots,d}$ means $\mathbf{a}$ is a matrix with (j,k)-th entry a_{jk}, and $b = (b_1, \dots, b_d)^T$ is regarded as a row vector. We also recall the gradient operator (for space variable) $\nabla_x = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_d}\right)^T$ and the Hessian operator $\nabla_x^2 = \nabla_x \otimes \nabla_x = \left(\frac{\partial^2}{\partial x_j \partial x_k}\right)_{j,k=1,\dots,d}$, where $\otimes$ means the tensor product. The variables or subscripts will be omitted when there is no ambiguity. Then we can rewrite the operator L as

$$\begin{aligned} Lu &= Au + Ku \\ &= \operatorname{tr}[\mathbf{a}^T(\nabla^2 u)] + b^T \nabla u + cu \\ &\quad + \int_{\mathbb{R}^d \setminus \{0\}} \left[u(\cdot + z) - u - z^T \nabla u \cdot \mathbf{1}_{\{|z|<1\}}\right] \nu(\cdot, dz), \end{aligned} \tag{5}$$

where "tr" denote the trace of a matrix. Both $x^T y$ and $x \cdot y$, for two vectors $x, y \in \mathbb{R}^d$, denote the scalar product. Moreover, we denote the positive and negative part of function u by $u^+ := u \vee 0$ and $u^- := -(u \wedge 0) = (-u) \vee 0$, respectively. Then $u = u^+ - u^-$ and $|u| = u^+ + u^-$.

In this section, L is the elliptic Waldenfels operator as defined in (1).

2.1 Weak maximum principle for elliptic case

We now prove the weak maximum principle.

Theorem 2.2 (Weak maximum principle for elliptic Waldenfels operators) *Let D be an open and bounded set but not necessarily connected, and E be an open set satisfying $D \subset E$. Assume that $u \in C^2(D) \cap C(\overline{E})$, $Lu \geq 0$ in D, and $\operatorname{supp} \nu(x, \cdot) \subset \overline{E} - x$ for each $x \in D$.*

1. *If $c \equiv 0$ in D, then*

$$\sup_{\overline{E}} u = \sup_{\overline{E} \setminus D} u.$$

2. *If $c \leq 0$ in D, then*

$$\sup_{\overline{E}} u \leq \sup_{\overline{E} \setminus D} u^+.$$

Here the supremum may be infinity.

Proof Assertion 1. We first consider the case with the strict inequality

$$Lu > 0 \quad \text{in } D. \tag{6}$$

Suppose that on the contrary $\sup_{\overline{E}} u > \sup_{\overline{E}\setminus D} u$. Then there exists a point $x^0 \in D$ with $u(x^0) = \sup_{\overline{E}} u$, and

$$u(x^0) = \max_{\overline{D}} u.$$

Thus at the maximizer point x^0, we have

$$\nabla u(x^0) = 0, \tag{7}$$

$$\nabla^2 u(x^0) \le 0, \tag{8}$$

where the last inequality means that the symmetric matrix $\nabla^2 u(x^0)$ is nonpositive definite. In particular, $\frac{\partial^2 u}{\partial x_j^2}(x^0) \le 0$, $j = 1, \ldots, d$. Since the matrix $\mathbf{a} = (a_{jk})$ is symmetric and positive definite at x^0, there exists an orthogonal matrix P such that

$$P[\mathbf{a}(x^0)]P^T = \operatorname{diag}(\lambda_1, \ldots, \lambda_d),$$

where "diag" means the diagonal matrix with diagonal entries $\lambda_j > 0$, $j = 1, \ldots, d$, which are eigenvalues of $\mathbf{a}(x^0)$. Then by changing variables $y - x^0 = P(x - x^0)$, we have

$$\nabla_x u = P^T(\nabla_y u),$$
$$\nabla_x^2 u = P^T(\nabla_y^2 u)P.$$

In light of (8), we find that at point x^0,

$$\begin{aligned} \operatorname{tr}[\mathbf{a}^T(\nabla_x^2 u)] &= \operatorname{tr}[\mathbf{a}^T P^T(\nabla_y^2 u)P] = \operatorname{tr}[P\mathbf{a}^T P^T(\nabla_y^2 u)] \\ &= \operatorname{tr}[(P\mathbf{a}P^T)^T(\nabla_y^2 u)] = \sum_j \lambda_j \frac{\partial^2 u}{\partial y_j^2} \le 0. \end{aligned} \tag{9}$$

Thus, combining (7), (9) and the assumption $c \equiv 0$, together with the fact that u attains a maximum at x^0, we obtain that at x^0,

$$\begin{aligned} Au &= \operatorname{tr}[\mathbf{a}^T(\nabla^2 u)] + b^T \nabla u + cu \le 0, \\ Ku(x^0) &= \int_{\mathbb{R}^d\setminus\{0\}} \left[u(x^0 + z) - u(x^0) - z^T \nabla u(x^0) \cdot \mathbf{1}_{\{|z|<1\}}\right] \nu(x^0, dz) \\ &= \int_{\overline{E}-x} \left[u(x^0 + z) - u(x^0)\right] \nu(x^0, dz) \\ &\le 0. \end{aligned}$$

Hence

$$Lu = Au + Ku \le 0 \quad \text{at } x^0. \tag{10}$$

Therefore, we get a contradiction in light of (6) and (10), which leads to $\sup_{\overline{E}} u = \sup_{\overline{E}\setminus D} u$.

For the general case that $Lu \ge 0$, we introduce a function

$$u^\epsilon(x) := u(x) + \epsilon e^{-\beta x_1}, \quad x \in \overline{E}, \tag{11}$$

where $\beta > 0$ will be selected below and ϵ is a positive parameter. Note that $a_{11} \ge \gamma > 0$, by substituting $z = e_1 = (1, 0, \ldots, 0)$ into condition (2). Then by Taylor expansion and the moment condition (3) of kernel ν, we have

$$\begin{aligned}
Lu^\epsilon &= Lu + \epsilon L(e^{-\beta x_1}) \\
&\ge \epsilon e^{-\beta x_1}\Big[\beta^2 a^{11} - \beta b^1 + \int_{|z|\ge 1} \big(e^{-\beta z_1} - 1\big)\nu(x, dz) \\
&\qquad + \int_{0<|z|<1} \big(e^{-\beta z_1} - 1 + \beta z_1\big)\nu(x, dz)\Big] \\
&\ge \epsilon e^{-\beta x_1}\Big[\beta^2 a^{11} - \beta b^1 - \int_{|z|\ge 1} \nu(x, dz) + \frac{1}{2}\beta^2 \int_{0<|z|<1} z_1^2 e^{-\beta\theta z_1}\nu(x, dz)\Big] \\
&\ge \epsilon e^{-\beta x_1}\Big[\beta^2 a^{11} - \beta b^1 - \int_{|z|\ge 1} \nu(x, dz) + \frac{1}{2}\beta^2 e^{-\beta\theta} \int_{0<|z|<1} z_1^2 \nu(x, dz)\Big] \\
&> 0,
\end{aligned}$$

provided $\beta > 0$ is large enough, where θ is a constant with $0 < \theta < 1$.

Then by the previous conclusion, $\sup_{\overline{E}} u^\epsilon = \sup_{\overline{E}\setminus D} u^\epsilon$. Let $\epsilon \to 0$ to find $\sup_{\overline{E}} u = \sup_{\overline{E}\setminus D} u$ by the continuity. This proves Assertion 1.

Assertion 2. If $u \le 0$ everywhere in D, the second assertion is trivially true. Hence we set $D_+ := \{x \in D \mid u(x) > 0\} \ne \emptyset$. Then

$$(L - c)u \ge -cu \ge 0 \quad \text{in } D_+.$$

The new operator $L - c$ has no zeroth-order term and consequently Assertion 1 implies that

$$\sup_{\overline{E}} u = \sup_{\overline{E}\setminus D_+} u = \Big(\sup_{\overline{E}\setminus D} u\Big) \vee \Big(\sup_{D\setminus D_+} u\Big) = \Big(\sup_{\overline{E}\setminus D} u\Big) \vee 0 = \sup_{\overline{E}\setminus D} u^+.$$

This completes the proof. □

Remark 2.3 From the proof of Assertion 2 in Theorem 2.2, we have the following conclusions.

1. In Assertion 1, if $Lu > 0$ in D, then u can either achieve its (finite) maximum only on $\overline{E} \setminus D$ or be unbounded on $\overline{E}$.

2. In Assertion 2, essentially the following equality holds according to the proof,

$$\sup_{\overline{E}} u^+ = \sup_{\overline{E}\setminus D} u^+,$$

even though the Assertion 1 in Theorem 2.2 cannot be applied directly to u^+ as it is not in $C^2(D)$. Especially if u can take positive values in D, or equivalently, $D_+ \neq \emptyset$, then we have

$$\sup_{\overline{E}} u = \sup_{\overline{E}\setminus D} u^+.$$

Remark 2.4 The proof of Theorem 2.2 still works if the matrix $\mathbf{a} = (a_{jk})$ is only positive semidefinite. Indeed, since the eigenvalues of $\mathbf{a}(x^0)$ are nonnegative ($\lambda_j \geq 0$, $j = 1, \ldots, d$), the inequality (9) still holds.

Remark 2.5 As in Remark 2.1, there are two special cases for Theorem 2.2, that is, $E = \mathbb{R}^d$ or $E = D$. Using the latter as an example, namely, $u \in C^2(D) \cap C(\overline{D})$, $Lu \geq 0$ in D, and $\operatorname{supp} \nu(x, \cdot) \subset \overline{D} - x$ for each $x \in D$, where D is open and bounded but not necessarily connected, then the following conclusions holds:

1. If $c \equiv 0$ in D, then

$$\max_{\overline{D}} u = \max_{\partial D} u.$$

2. If $c \leq 0$ in D, then

$$\max_{\overline{D}} u \leq \max_{\partial D} u^+.$$

Corollary 2.6 *Let D be an open and bounded set but not necessarily connected, and E be an open set satisfying $D \subset E$. Assume that $u \in C^2(D) \cap C(\overline{E})$, and $\operatorname{supp} \nu(x, \cdot) \subset \overline{E} - x$ for each $x \in D$.*

1. *If $c \equiv 0$ and $Lu \leq 0$ both hold in D, then*

$$\inf_{\overline{E}} u = \inf_{\overline{E}\setminus D} u.$$

2. *If $c \leq 0$ and $Lu \leq 0$ both hold in D, then*

$$\inf_{\overline{E}} u \geq -\sup_{\overline{E}\setminus D} u^-.$$

3. *If $c \leq 0$ and $Lu = 0$ both hold in D, then*

$$\sup_{\overline{E}} |u| = \sup_{\overline{E}\setminus D} |u|.$$

In all the three expressions, the supremum and infimum may be infinity.

Proof 1. Apply directly the first assertion of Theorem 2.2 to $-u$.
2. Apply the second assertion of Theorem 2.2 to $-u$.
3. Applying Statement 2 in Remark 2.3 to $-u$, we have

$$\sup_{\overline{E}} u^- = \sup_{\overline{E}\setminus D} u^-.$$

Then it follows that

$$\sup_{\overline{E}} |u| = \left(\sup_{\overline{E}} u^+\right) \vee \left(\sup_{\overline{E}} u^-\right) = \left(\sup_{\overline{E}\setminus D} u^+\right) \vee \left(\sup_{\overline{E}\setminus D} u^-\right) = \sup_{\overline{E}\setminus D} |u|.$$

This completes the proof. □

Going one step further, we suppose E is bounded and then apply Corollary 2.6 to $u - v$, yielding the following corollary which is often used in applications.

Corollary 2.7 *Let D be an open and bounded set but not necessarily connected, and E be an open set satisfying $D \subset E$. Assume that $u, v \in C^2(D) \cap C(\overline{E})$, $c \leq 0$ in D, and $\operatorname{supp} \nu(x, \cdot) \subset \overline{E} - x$ for each $x \in D$.*

1. *(Comparison Principle) If $Lu \leq Lv$ in D and $u \geq v$ on $\overline{E} \setminus D$, then $u \geq v$ in $\overline{E}$.*
2. *(Uniqueness) If $Lu = Lv$ in D and $u = v$ on $\overline{E} \setminus D$, then $u = v$ in $\overline{E}$.*

Proof The two results immediately follow by using the last two assertions of Corollary 2.6 for $u - v$. □

2.2 Strong maximum principle for elliptic case

This section is devoted to the strong maximum principle for the elliptic Waldenfels operator L.

Theorem 2.8 (Strong maximum principle for elliptic Waldenfels operator) *Let D be an open and connected set but not necessarily bounded, and E be an open set satisfying $D \subset E$. Assume that $u \in C^2(D) \cap C(\overline{E})$, $Lu \geq 0$ in D, and $\operatorname{supp} \nu(x, \cdot) \subset \overline{E} - x$ for each $x \in D$. Moreover, assume that the mapping $x \to \nu(x, \cdot)$ is continuous in D. If one of the following conditions holds:*

1. *$c \equiv 0$ in D and u achieves a (finite) maximum over $\overline{E}$ at an interior point in D;*
2. *$c \leq 0$ in D and u achieves a (finite) nonnegative maximum over $\overline{E}$ at an interior point in D;*
3. *u achieves a zero maximum over $\overline{E}$ at an interior point in D,*

then u is constant on $\overline{D}$.

Before proving this theorem, let us first give some comments on it.

Remark 2.9 The propagation of maximizer point by translation of measure support mentioned in [5,7] is similar in our case. That is, if the assumptions in Theorem 2.8 hold, then u is a constant on the set $\overline{\bigcup_{n=0}^{\infty} \Lambda_n}$, where Λ_n's are defined by induction,

$$\Lambda_0 = x^0, \ \Lambda_{n+1} = \bigcup_{x \in D \cap \Lambda_n} [\operatorname{supp} \nu(x, \cdot) + x].$$

This result depends on the support of every measure $\nu(x, \cdot)$, it can be easily proved by induction and continuity. It is noteworthy that in this scheme, the set D may not be connected, since jumps from one connected component to another might occur when measure supports overlap two or more connected components.

In conclusion, it is the integro-differential term, or jump diffusion term that leads to the propagation of maximizer point between those connected components. Therefore, we need to restrict that the Markovian point can move by jumps only inside the set E, i.e., the support condition (4), to obtain the propagation of maximizer (over E) point.

Remark 2.10 As shown in Remark 2.1, our results on the weak and strong maximum principles formulated in Theorem 2.2 and 2.8, respectively, cover the situations when the support of jump measure is either bounded or unbounded, especially for $E = D$ or $E = \mathbb{R}^d$ in the setting. While Taira [30] only considered the situation for $E = D$. Furthermore, our assumptions are less restrictive than Taira's: In our work, the connectedness is not needed for the weak maximum principle while the boundedness is not necessary for the strong maximum principle. Moreover, the continuity of mapping $x \to \nu(x, \cdot)$ is necessary only in the strong case but not for the weak maximum principle.

Like the weak case, by applying directly Theorem 2.8 to $-u$, one can conclude the strong maximum principle for the converse case $Lu \leq 0$.

Corollary 2.11 *Let D be an open and connected set but not necessarily bounded, and E be an open set satisfying $D \subset E$. Assume that $u \in C^2(D) \cap C(\overline{E})$, $Lu \leq 0$ in D, and $\operatorname{supp} \nu(x, \cdot) \subset \overline{E} - x$ for each $x \in D$. Moreover, assume that the mapping $x \to \nu(x, \cdot)$ is continuous in D. If one of the following conditions holds:*

1. *$c \equiv 0$ in D and u achieves a (finite) minimum over $\overline{E}$ at an interior point in D;*
2. *$c \leq 0$ in D and u achieves a (finite) nonnegative minimum over $\overline{E}$ at an interior point in D;*
3. *u achieves a zero minimum over $\overline{E}$ at an interior point in D,*

then u is constant on $\overline{D}$.

Now we start to prove Theorem 2.8.

Proof of Theorem 2.8 Suppose that $u \not\equiv \max_{\overline{E}} u$ in D. Set $D_< := \{x \in D \mid u(x) < \max_{\overline{E}} u\} \neq \emptyset$. Since D is connected which implies $\partial D_< \cap D \neq \emptyset$, we can always choose a point $x^1 \in D_<$ such that $\operatorname{dist}(x^1, \partial D_< \cap D) < \operatorname{dist}(x^1, \partial D)$. Denote by B the largest ball having x^1 as center with $B \subset D_<$. Then $\overline{B} \subset D$ and there exists some point $x^0 \in \partial B$ with

$$u(x^0) = \max_{\overline{E}} u > u(x), \quad \forall x \in B.$$

Since u achieves its maximum at $x^0 \in D$, we have $\nabla u(x^0) = 0$. We will create a contradiction by proving that

$$\frac{\partial u}{\partial \mathbf{n}}(x^0) > 0, \tag{12}$$

where $\mathbf{n}$ is the unit outer normal vector of B at x^0. Then by this contradiction, u must be constant within D, and the result follows by continuity. Now the rest of the proof is devoted to (12). We divide it into three steps.

Step 1 The closed set $\overline{B}$ is a d-dimension C^2-differential manifold with boundary. Let (U, Φ) be a coordinate chart near x^0, where U is a relatively open neighborhood of x^0 in $\overline{B}$, Φ is a C^2-diffeomorphism to its image from U into the closed upper half plane $\mathbb{H}^d_+ := \{y \in \mathbb{R}^d \mid y_d \geq 0\}$, with inverse Φ^{-1}. Then Φ is an embedding whose rank at x^0 equals to d, equivalently, if we denote by $J\Phi$ the Jacobian matrix of Φ, i.e., $J\Phi := \nabla_x \Phi$, then $J\Phi$ is non-degenerate. As a result, the tangent mapping Φ_* induced by Φ at point x^0 is an isomorphism.

Now we consider the function u restricted in U. We define $\hat{u}(y) := u(\Phi^{-1}(y))$, $y \in \Phi(U)$. Then $\hat{u}$ attains its maximum at $y^0 = \Phi(x^0)$ over $\Phi(U) \subset \mathbb{H}^d_+$. Hence at the maximizer point y^0,

$$\frac{\partial \hat{u}}{\partial y_j} = 0, \quad j = 1, \ldots, d-1. \tag{13}$$

We also denote the image tangent vector of $\frac{\partial}{\partial \mathbf{n}}$ under tangent mapping Φ_* by

$$\frac{\partial}{\partial \hat{\mathbf{n}}} := \Phi_*\Big(\frac{\partial}{\partial \mathbf{n}}\Big).$$

We compute at y^0 (or x^0)

$$\begin{aligned} \frac{\partial \hat{u}}{\partial \hat{\mathbf{n}}} &= \Big\langle \Phi_*\Big(\frac{\partial}{\partial \mathbf{n}}\Big), d\hat{u} \Big\rangle = \Big\langle \frac{\partial}{\partial \mathbf{n}}, \Phi^*(d\hat{u}) \Big\rangle \\ &= \Big\langle \frac{\partial}{\partial \mathbf{n}}, d(\hat{u} \circ \Phi) \Big\rangle = \Big\langle \frac{\partial}{\partial \mathbf{n}}, du \Big\rangle = \frac{\partial u}{\partial \mathbf{n}} = 0, \end{aligned} \tag{14}$$

where Φ^* is denoted as the cotangent mapping induced by Φ at point x^0, $\langle \cdot, \cdot \rangle$ is the dual product between the tangent space and cotangent space at y^0 (or x^0). Now recall that Φ_* is an isomorphism. The tangent vector $\frac{\partial}{\partial \hat{\mathbf{n}}}$ is independent of $\{\frac{\partial}{\partial y_j} \mid j = 1, \ldots, d-1\}$ and consequently by (14),

$$\frac{\partial \hat{u}}{\partial y_d}(y^0) = 0. \tag{15}$$

Combining (13) and (15) together with the fact that $\hat{u}$ attains its maximum at y^0, we have

$$\nabla^2_y \hat{u}(y^0) \leq 0. \tag{16}$$

Combining (13), (15) and (16), we have at x^0,

$$\begin{aligned}
&\nabla_x u = (J\Phi)^T(\nabla_y \hat{u}) = 0,\\
&\nabla_x^2 u = (J\Phi)^T(\nabla_y^2 \hat{u})(J\Phi) + (\nabla_x^2 \Phi)(\nabla_y \hat{u}) = (J\Phi)^T(\nabla_y^2 \hat{u})(J\Phi),
\end{aligned}$$

where we treat $\nabla_x^2\Phi$ as a third-order covariant tensor. Hence at x^0,

$$\begin{aligned}
Au &= \mathrm{tr}[\mathbf{a}^T(\nabla_x^2 u)] + b^T\nabla_x u + cu\\
&= \mathrm{tr}[\mathbf{a}^T(J\Phi)^T(\nabla_y^2\hat{u})(J\Phi)] + cu\\
&= \mathrm{tr}[(J\Phi)\mathbf{a}^T(J\Phi)^T(\nabla_y^2\hat{u})] + cu\\
&= \mathrm{tr}\big[\big((J\Phi)\mathbf{a}(J\Phi)^T\big)^T(\nabla_y^2\hat{u})\big] + cu\\
&=: \mathrm{tr}[\hat{\mathbf{a}}^T(\nabla_y^2\hat{u})] + cu,
\end{aligned} \tag{17}$$

where $\hat{\mathbf{a}} := (J\Phi)\mathbf{a}(J\Phi)^T$. Since $\mathbf{a}(x)$ is symmetric and positive definite and the matrix $J\Phi$ is non-degenerate, we see the matrix $\hat{\mathbf{a}}(x^0)$ is also symmetric and positive definite. Hence, as explained in the proof of Theorem 2.2 and by (16), we have

$$\mathrm{tr}\big[\hat{\mathbf{a}}(x^0)^T\big(\nabla_y^2\hat{u}(y^0)\big)\big] \le 0. \tag{18}$$

Define

$$E_0 := \big\{x \in \overline{E} \mid u(x) = \max_{\overline{E}} u\big\} = \big\{x \in \overline{E}\setminus B \mid u(x) = \max_{\overline{E}} u\big\}. \tag{19}$$

Recall that u attains its maximum over $\overline{E}$ at x^0. Now we have

$$\begin{aligned}
Ku(x^0) &= \int_{\mathbb{R}^d\setminus\{0\}} \big[u(x^0+z) - u(x^0) - z^T\nabla u(x^0)\cdot \mathbf{1}_{\{|z|<1\}}\big]\nu(x^0,dz)\\
&= \int_{\overline{E}} [u(x^0+z) - u(x^0)]\nu(x^0,dz)\\
&= \int_{x^0+z\in\overline{E}\setminus E_0} [u(x^0+z) - u(x^0)]\nu(x^0,dz)\\
&\le 0.
\end{aligned} \tag{20}$$

From (17), (18) and (20), we obtain

$$Lu(x^0) = Ku(x^0) + Au(x^0) \le c(x^0)u(x^0) \le 0.$$

By recalling the assumption on u, we have $Lu(x) \ge 0$ for each $x \in D$, and thus

$$Lu(x^0) = Au(x^0) = Ku(x^0) = 0,$$

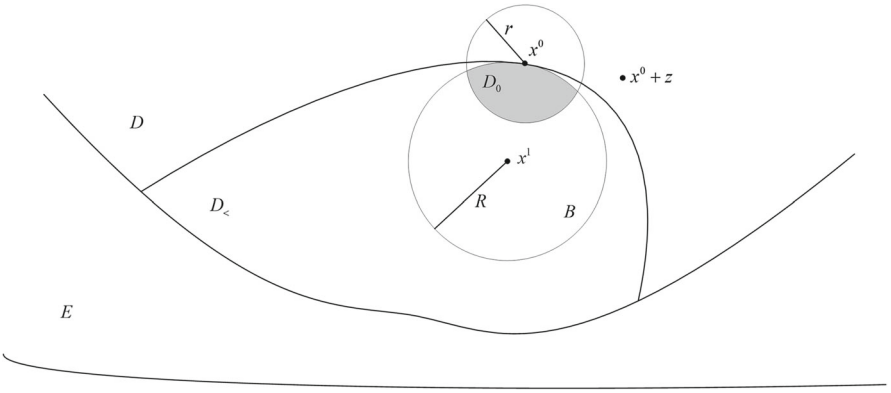

Fig. 1 Sketch for Theorem 2.8

especially,

$$Ku(x^0) = \int_{x^0+z\in\overline{E}\setminus E_0} [u(x^0+z) - u(x^0)]\nu(x^0, dz) = 0.$$

Hence, we conclude

$$\nu(x^0, (\overline{E} \setminus E_0) - x^0) = 0. \tag{21}$$

Step 2 We set $B = B(x^1, R)$ with $R = |x^0 - x^1|$. See Fig. 1. Define

$$v(x) := e^{-\beta|x-x^1|^2} - e^{-\beta R^2}, \quad x \in \overline{E},$$

for $\beta > 0$ as selected below. Then

$$\begin{aligned} Av &= e^{-\beta|x-x^1|^2}\Big\{\mathrm{tr}\big[\mathbf{a}^T\big(4\beta^2(x-x^1)\otimes(x-x^1) - 2\beta I\big)\big] \\ &\qquad - 2\beta b^T(x-x^1) + c\big(1 - e^{-\beta(R^2-|x-x^1|^2)}\big)\Big\} \\ &= e^{-\beta|x-x^1|^2}\big[4\beta^2(x-x^1)^T\mathbf{a}^T(x-x^1) - 2\beta\mathrm{tr}(\mathbf{a}) \\ &\qquad - 2\beta b^T(x-x^1) + c\big(1 - e^{-\beta(R^2-|x-x^1|^2)}\big)\big] \\ &\geq e^{-\beta|x-x^1|^2}\big[4\gamma\beta^2|x-x^1|^2 - 2\beta\mathrm{tr}(\mathbf{a}) - 2\beta|b||x-x^1| \\ &\qquad + c\big(1 - e^{-\beta(R^2-|x-x^1|^2)}\big)\big]. \end{aligned} \tag{22}$$

Consider next the open set $D_0 := B(x^1, R)\cap B(x^0, r)$ (see Fig. 1) with some $r \in (0, R)$ which will be chosen later. When β is large enough, we have

$$Av \geq e^{-\beta R^2}\big[4\gamma\beta^2(R-r)^2 - 2\beta\mathrm{tr}(\mathbf{a}) - 2\beta|b|R\big] > C_1\beta^2 - C_2\beta, \tag{23}$$

for $x \in D_0$, where C_1, C_2 are two positive constants.

Moreover, by recalling (21), we have

$$\begin{aligned} K v(x^0) &= \int_{\mathbb{R}^d \setminus \{0\}} \big[e^{-\beta|x^0-x^1+z|^2} - e^{-\beta R^2} \\ &\qquad + 2\beta z^T (x^0 - x^1) e^{-\beta R^2} \mathbf{1}_{\{|z|<1\}}\big] \nu(x^0, dz) \\ &= \int_{\substack{x^0+z\in E_0 \\ |z|\geq 1}} \big[e^{-\beta|x^0-x^1+z|^2} - e^{-\beta R^2}\big] \nu(x^0, dz) \\ &\quad + \int_{\substack{x^0+z\in E_0 \\ 0<|z|<1}} \big[e^{-\beta|x^0-x^1+z|^2} - e^{-\beta R^2} \\ &\qquad + 2\beta z^T (x^0 - x^1) e^{-\beta R^2}\big] \nu(x^0, dz) \\ &=: I + II. \end{aligned} \tag{24}$$

For the term I, it is clear that $E_0 \cap \overline{B} = \{x^0\}$ and consequently

$$|x^0 + z - x^1| > |x^0 - x^1| = R$$

for point z satisfying $x^0 + z \in \overline{E} \setminus E_0$. Thus for sufficiently large β, we have

$$-C_3 < e^{-\beta|x^0-x^1+z|^2} - e^{-\beta R^2} < 0$$

with a constant $C_3 > 0$. Hence,

$$I > -C_3 \int_{\substack{x^0+z\in E_0 \\ |z|\geq 1}} \nu(x^0, dz) \gtrsim -C_3. \tag{25}$$

For the term II, using the Taylor expansion, and for $x^0 + z \in E_0$ and β large enough,

$$\begin{aligned} &e^{-\beta|x^0-x^1+z|^2} - e^{-\beta R^2} + 2\beta z^T (x^0 - x^1) e^{-\beta R^2} \\ &= \frac{1}{2}\big[4\beta^2 e^{-\beta|x^0-x^1+\theta z|^2} |z^T (x^0 - x^1 + \theta z)|^2 - 2\beta e^{-\beta|x^0-x^1+\theta z|^2} |z|^2\big] \\ &\geq -\beta e^{-\beta|x^0-x^1+\theta z|^2} |z|^2 \\ &\geq -C_4 \beta |z|^2, \end{aligned}$$

with some $\theta \in (0, 1)$ and a constant $C_4 > 0$. Hence,

$$II > -C_4\beta \int_{\substack{x^0+z\in E_0 \\ 0<|z|<1}} |z|^2 \nu(x, dz) \gtrsim -C_4\beta. \tag{26}$$

Thus, combining the results of (23), (24), (25) and (26), we find that

$$\begin{aligned} Lv(x^0) &= Av(x^0) + Kv(x^0) = Av(x^0) + I + II \\ &\gtrsim C_1\beta^2 - (C_2 + C_4)\beta - C_3 \\ &> 0, \end{aligned}$$

provided $\beta > 0$ is fixed large enough. Since $Lv(x)$ is continuous in $x \in D$ in light of the continuity of $v(x, \cdot)$, we have

$$Lv(x) \geq 0, \tag{27}$$

for $x \in D_0$, provided r is small enough.

Step 3 Define

$$u^\epsilon(x) = u(x) + \epsilon v(x) - u(x^0), \quad x \in \overline{E},$$

for a constant $\epsilon > 0$. We can choose ϵ so small that

$$u^\epsilon(x) \leq 0, \quad x \in \overline{E} \setminus D_0,$$

since $v(x) \leq 0$ for $x \in \overline{E} \setminus B$, and $u(x) < u(x^0)$ for $x \in B \setminus D_0$ by recalling $u(x^0) > u(x)$ for all $x \in D_<$.

For the first two cases, $c \equiv 0$ in D, or $c \leq 0$ in D also $u(x^0) \geq 0$, from (27) and the fact that $Lu \geq 0$ in D, we see that

$$Lu^\epsilon \geq -cu(x^0) \geq 0 \quad \text{in } D_0.$$

In view of the weak maximum principle of elliptic Waldenfels operator, Theorem 2.2, we know that $u^\epsilon \leq 0$ in $\overline{E}$. Note that $u^\epsilon(x^0) = 0$. Thus we have,

$$0 = \frac{\partial u^\epsilon}{\partial \mathbf{n}}(x^0) = \frac{\partial u}{\partial \mathbf{n}}(x^0) + \epsilon \frac{\partial v}{\partial \mathbf{n}}(x^0).$$

Consequently,

$$\frac{\partial u}{\partial \mathbf{n}}(x^0) = -\epsilon \frac{\partial v}{\partial \mathbf{n}}(x^0) = -\epsilon \nabla v(x^0) \cdot \frac{(x^0 - x^1)}{R} = 2\epsilon\beta R e^{-\beta R^2} > 0,$$

as required.

For the third case that $u(x^0) = 0$, obviously $u \leq 0$ in D. We find

$$(L - c^+)u = Lu - c^+u \geq Lu \geq 0 \quad \text{in } D.$$

Notice that the zeroth-order coefficient of operator $L - c^+$ is $c - c^+$, which is non-positive in D. Hence we apply the result of the second case by replacing L and c respectively with $L - c^+$ and $c - c^+$ to get the same result for this case.

We have thus completed the proof. □

Some comments will be helpful for understanding the long proof of Theorem 2.8.

Remark 2.12 In Theorem 2.8, we restrict the set D to be connected to ensure $\partial D_< \cap D \neq \emptyset$. More generally, if D is not connected, one may merely replace D with the connected component of D which contains the maximizer point, and we thus conclude that u is constant in this connected component.

Recalling Remark 2.13, we could see that the diffusion term gives rise to the propagation of maximizer point in the corresponding connected component. This is why we need the set D to be connected.

Remark 2.13 We can see from (23), (25) and (26) that, it is the second-order differential term $\mathrm{tr}[\mathbf{a}^T(\nabla^2)]$, namely, the diffusion term that plays a leading role in Step 2 in the proof of Theorem 2.8.

Remark 2.14 Theorem 2.8 still holds if the matrix $\mathbf{a}(x) = (a_{jk}(x))_{j,k=1,\ldots,d}$ is only positive semidefinite and the unit outer normal vector $\mathbf{n}$ is not in the nullspace of $\mathbf{a}(x^0)$.

In fact, recall that $\hat{\mathbf{a}} = (J\Phi)\mathbf{a}(J\Phi)^T$ is also semidefinite as the Jacobian matrix $J\Phi$ is invertible. Due to the reason mentioned in Remark 2.4, we confirm that (18) still holds. Moreover, noting that there exists a positive constant γ such that $\mathbf{n}^T a(x^0)\mathbf{n} \geq \gamma > 0$ with $\mathbf{n}$ not in the nullspace of $\mathbf{a}(x^0)$, and consequently $(x^0 - x^1)^T a(x^0)(x^0 - x^1) \geq \gamma |x^0 - x^1|^2$. By continuity we can choose r so small that for all $x \in D_0 = B(x^1, R) \cap B(x^0, r)$,

$$(x - x^1)^T a(x)(x - x^1) \geq \gamma_1 |x - x^1|^2,$$

with a positive constant γ_1. Hence (22) holds with γ_1 in placing of γ and (23) also holds for some other constants C_1, C_2.

By a similar way to prove (12), we can easily obtain the following version of Hopf's boundary point lemma, which is a generalization of [30, Lemma C.3].

Proposition 2.15 (Hopf's boundary point lemma for elliptic Waldenfels operators) *Let D be an open set (not necessarily bounded or connected) with boundary ∂D being C^2. Assume that $u \in C^2(\overline{D})$, $Lu \geq 0$ in D, and $\mathrm{supp}\, \nu(x, \cdot) \subset \overline{D} - x$ for each $x \in D$, and furthermore the mapping $x \to \nu(x, \cdot)$ is continuous in D. Suppose that u achieves its (finite) maximum over $\overline{D}$ at point $x^0 \in \partial D$ such that $u(x^0) > u(x)$ for all $x \in D$, and that one of the following conditions holds:*

1. *$c \equiv 0$ in D;*
2. *$c \leq 0$ in D and $u(x^0) \geq 0$;*
3. *$u(x^0) = 0$.*

Then the outer normal derivative is positive: $\frac{\partial u}{\partial \mathbf{n}}(x^0) > 0$.

In fact, if we let $E = D$ and replace B by D in Step 1 in the proof of Theorem 2.8, also replace $D_<$ by D in Step 3, then the three-step argument also works in the context of Proposition 2.15 and the result follows.

3 Maximum principles for parabolic Waldenfels operators

We assume that D, E are two open sets in $\mathbb{R}^d$ and $D \subset E$, where E is not necessarily bounded. Set $D_T := D \times (0, T]$ and $E_T := E \times (0, T]$ for arbitrarily fixed $T > 0$.

As in [14–16,28,29], we define a time dependent elliptic Waldenfels operator

$$L := A + K, \tag{28}$$

where A and K are defined as, respectively

$$Au(x,t) := \sum_{j,k=1}^{d} a_{jk}(x,t)\frac{\partial^2 u}{\partial x_j \partial x_k}(x,t) + \sum_{j=1}^{d} b_j(x,t)\frac{\partial u}{\partial x_j}(x,t) + c(x,t)u(x,t),$$

$$Ku(x,t) := \int_{\mathbb{R}^d\setminus\{0\}} \Big[u(x+z,t) - u(x,t) - \sum_{j=1}^{d} z_j \frac{\partial u}{\partial x_j}(x,t)\mathbf{1}_{\{|z|<1\}}\Big]\nu(t,x,dz).$$

We make the following assumptions:

1. Continuity condition: $a_{jk}, b_j, c \in C(\overline{E_T})$ $(j,k = 1,\ldots,d)$.
2. Symmetry condition: $a_{jk} = a_{kj}$ $(j,k = 1,\ldots,d)$. Uniform ellipticity condition: There exists a constant $\gamma > 0$ such that

$$\sum_{j,k=1}^{d} a_{jk}(x,t)\xi_j\xi_k \geq \gamma|\xi|^2,$$

   for all $(x,t) \in D_T$, $\xi \in \mathbb{R}^d$.
3. Lévy measures: The kernel $\{\nu(t,x,\cdot) \mid (x,t) \in \mathbb{R}^d \times [0,T]\}$ is a family of Lévy measures, namely, each $\nu(t,x,\cdot)$ is a Borel measure on $\mathbb{R}^d \setminus \{0\}$ such that for all $(x,t) \in \mathbb{R}^d \times [0,T]$,

$$\int_{\mathbb{R}^d\setminus\{0\}} (1 \wedge |z|^2)\nu(t,x,dz) < \infty, \tag{29}$$

   and moreover, for fixed $U \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$, the mapping $\mathbb{R}^d \times [0,T] \ni (x,t) \to \nu(t,x,U) \in [0,\infty)$ is Borel measurable. Here we further assume that for each $(x,t) \in D_T$, the measure $\nu(t,x,\cdot)$ is supported in $\overline{E} - x := \{y - x \mid y \in \overline{E}\} = \{z \mid x + z \in \overline{E}\}$. That is,

$$\operatorname{supp}\nu(t,x,\cdot) \subset \overline{E} - x, \quad \forall (x,t) \in D_T. \tag{30}$$

The Markov process associated with such a generator L can be determined as a solution to the martingale problem induced by L (see, e.g., [29]). However, it is not clear if the Markov process determined by the martingale problem is linked to a stochastic differential equation with certain boundary conditions.

Now we consider the parabolic Waldenfels operator

$$-\frac{\partial}{\partial t}+L,$$

with L being defined in (28), and we are concerned with the maximum principles for such a parabolic operator.

3.1 Weak maximum principle for parabolic case

We are in the position to present both weak and strong maximum principles for parabolic Waldenfels operator $-\frac{\partial}{\partial t}+L$. First we prove the weak one.

Theorem 3.1 (Weak maximum principle for parabolic Waldenfels operators) *Let D be an open and bounded set but not necessarily connected, and E be an open set satisfying $D \subset E$. Assume that $u \in C^{2,1}(D_T) \cap C(\overline{E_T})$, $-\frac{\partial u}{\partial t}+Lu \geq 0$ in D_T, and $supp\, \nu(t,x,\cdot) \subset \overline{E}-x$ for each $(x,t) \in D_T$.*

1. *If $c \equiv 0$ in D_T, then*

$$\sup_{\overline{E_T}} u = \sup_{\overline{E_T}\setminus D_T} u.$$

2. *If $c \leq 0$ in D_T, then*

$$\sup_{\overline{E_T}} u \leq \sup_{\overline{E_T}\setminus D_T} u^+.$$

Here the supremum may be infinity.

Proof Assertion 1. We prove this by contradiction. Suppose that the strict inequality holds, i.e.,

$$-\frac{\partial u}{\partial t}+Lu>0 \quad \text{in } D_T, \tag{31}$$

but there exists a point $(x^0,t^0) \in D_T$ such that

$$u(x^0,t^0)=\max_{\overline{E_T}} u.$$

On one hand, as explained in the proof of Theorem 2.2, we note that $Lu \leq 0$ at point (x^0,t^0). On the other hand, if $0<t^0<T$, then $(x^0,t^0) \in (D_T)^\circ$ and consequently

$$\frac{\partial u}{\partial t}=0 \quad \text{at } (x^0,t^0);$$

if $t^0=T$, then $(x^0,t^0) \in \partial(D_T)$ and consequently

$$\frac{\partial u}{\partial t} \geq 0 \quad \text{at } (x^0,t^0).$$

Thus we always have $-\frac{\partial u}{\partial t} + Lu \leq 0$ at point (x^0, t^0), a contradiction to (31).

In the general case that $-\frac{\partial u}{\partial t} + Lu \geq 0$ holds in D_T, define

$$u^\epsilon(x,t) := u(x,t) - \epsilon t \quad \text{in } \overline{E_T}, \tag{32}$$

with a positive parameter ϵ. Then

$$-\frac{\partial u^\epsilon}{\partial t} + Lu^\epsilon = -\frac{\partial u}{\partial t} + Lu + \epsilon > 0,$$

and hence $\sup_{\overline{E_T}} u^\epsilon = \sup_{\overline{E_T}\setminus D_T} u^\epsilon$. Now Assertion 1 follows by setting $\epsilon \to 0$.

Assertion 2. If u is nonpositive throughout D, Assertion 2 is trivially true. Hence we may assume on the contrary that u achieves a positive maximum at a point $(x^0, t^0) \in D_T$ over $\overline{E_T}$.

We first consider the case with strict inequality $-\frac{\partial u}{\partial t} + Lu > 0$ in D_T. Since $u(x^0, t^0) > 0$ and $c \leq 0$, we derive the contradiction to Assertion 1,

$$-\frac{\partial u}{\partial t} + Lu \leq 0 \quad \text{at } (x^0, t^0).$$

More generally, if $-\frac{\partial u}{\partial t} + Lu \geq 0$ in D_T, then set as before $u^\epsilon(x,t) := u(x,t) - \epsilon t$ with $\epsilon > 0$, which leads to

$$-\frac{\partial u^\epsilon}{\partial t} + Lu^\epsilon = -\frac{\partial u}{\partial t} + Lu + \epsilon - c \cdot \epsilon t \geq \epsilon - c \cdot \epsilon t > 0.$$

Moreover, if u achieves a positive maximum at a point $(x^0, t^0) \in D_T$ over $\overline{E_T}$, then by the continuity, u^ϵ also achieves a positive maximum at a point $(x^0, t^0) \in D_T$ over $\overline{E_T}$, provided that ϵ is small enough. However, as in the previous proof, we obtain a contradiction.

This completes the proof. □

Remark 3.2 As in the first statement of Remark 2.3, we also conclude that in Assertion 1 of Theorem 3.1 if strictly $-\frac{\partial u}{\partial t} + Lu > 0$ in D, then u can either achieve its (finite) maximum only on $\overline{E_T} \setminus D_T$ or be unbounded on $\overline{E_T}$.

Remark 3.3 We cannot prove Assertion 2 of Theorem 3.1 in the same way as the corresponding assertion in Theorem 2.2. In fact, if we introduce similarly the set $D_T^+ := \{(x,t) \in D_T \mid u(x,t) > 0\}$, it will never be the form of $U \times (0, T]$ for some $U \subset D$. Hence we may not take advantage of the first assertion of Theorem 3.1. Consequently, the similar judgment with Assertion 2 of Remark 2.3, which lies on the proof of Assertion 2 in Theorem 2.2, cannot be established here.

Remark 3.4 From Theorem 2.2 and Remark 2.4, we have already known that, for $u \in C^{2,2}(D_T) \cap C(\overline{E_T})$, the supremum (or respectively, positive supremum) is achieved on $\overline{E_T} \setminus (D_T)^\circ$. The alert reader could notice that we may have appeared to be cheating here, as we should also verify that the kernel ν still satisfy the third assumption in the

definition of elliptic Waldenfels operator (1) when regarding it as a kernel in $\mathbb{R}^{d+1}$. In fact, the modified kernel $\hat{\nu}((x,t),dzds) := \nu(t,x,dz)\delta_0(ds)$ does satisfy the moment condition (3), which is enough for us even though $\hat{\nu}$ is not supported inside $\mathbb{R}^{d+1}\setminus\{0\}$. See the proof of Lemma 3.11 for details.

Moreover, if $-\frac{\partial u}{\partial t} + Lu > 0$ in D_T, we see that the maximum (or respectively, positive maximum) cannot be achieved on the upper boundary $D\times\{T\}$ by the same argument as in the proof of Theorem 3.1. Hence, it is clear that Theorem 3.1 holds for $u \in C^{2,2}(D_T)\cap C(\overline{E_T})$, which is a natural consequence of Theorem 2.2.

However, the result for $u \in C^{2,1}(D_T)\cap C(\overline{E_T})$ cannot be obtained in this way. We only need the first-order differentiability in t, benefiting from the form of the operator $-\frac{\partial}{\partial t} + L$. This is evident in the different forms of u^ϵ in (11) and (32).

Remark 3.5 There are two special cases for the weak maximum principle Theorem 3.1 for the parabolic operator $-\frac{\partial}{\partial t} + L$. That is, $E = \mathbb{R}^d$ or $E = D$. Take the latter as an example. Let $u \in C^{2,1}(D_T)\cap C(\overline{D_T})$, $-\frac{\partial u}{\partial t} + Lu \geq 0$ in D_T, and $\operatorname{supp}\nu(t,x,\cdot)\subset \overline{D} - x$ for each $(x,t)\in D_T$, where D is open and bounded but not necessarily connected.

1. If $c\equiv 0$ in D_T, then

$$\max_{\overline{D_T}} u = \max_{\Gamma_T} u.$$

2. If $c\leq 0$ in D_T, then

$$\max_{\overline{D_T}} u \leq \max_{\Gamma_T} u^+.$$

Here Γ_T is the parabolic boundary of D_T, i.e., $\Gamma_T := \overline{D_T}\setminus D_T = (\partial D\times[0,T])\cup(D\times\{0\})$.

There are some consequences of the weak maximum principle for a parabolic Waldenfels operator. We only highlight the following results.

Corollary 3.6 *Let D be an open and bounded set but not necessarily connected, and E be an open set satisfying $D\subset E$. Assume that $u\in C^{2,1}(D_T)\cap C(\overline{E_T})$, and $\operatorname{supp}\nu(t,x,\cdot)\subset\overline{E}-x$ for each $(x,t)\in D_T$.*

1. *If $c\equiv 0$ and $-\frac{\partial u}{\partial t} + Lu \leq 0$ both hold in D_T, then*

$$\inf_{\overline{E_T}} u = \inf_{\overline{E_T}\setminus D_T} u.$$

2. *If $c\leq 0$ and $-\frac{\partial u}{\partial t} + Lu \leq 0$ both hold in D_T, then*

$$\inf_{\overline{E_T}} u \geq -\sup_{\overline{E_T}\setminus D_T} u^-.$$

3. *If* $c \le 0$ *and* $-\frac{\partial u}{\partial t} + Lu = 0$ *both hold in* D_T, *then*

$$\sup_{\overline{E_T}} |u| = \sup_{\overline{E_T} \setminus D_T} |u|.$$

Here the supremum and infimum may be infinity.

Corollary 3.7 *Let* D *be an open and bounded set but not necessarily connected, and* E *be an open set satisfying* $D \subset E$. *Assume that* $u, v \in C^{2,1}(D_T) \cap C(\overline{E_T})$, $c \le 0$ *in* D_T, *and* $\operatorname{supp} v(t, x, \cdot) \subset \overline{E} - x$ *for each* $(x, t) \in D_T$. *There is no sign condition on* c.

1. *(Comparison Principle) If* $-\frac{\partial u}{\partial t} + Lu \le -\frac{\partial v}{\partial t} + Lv$ *in* D_T *and* $u \ge v$ *on* $\overline{E_T} \setminus D_T$, *then* $u \ge v$ *in* $\overline{E_T}$.
2. *(Uniqueness) If* $-\frac{\partial u}{\partial t} + Lu = -\frac{\partial v}{\partial t} + Lv$ *in* D_T *and* $u = v$ *on* $\overline{E_T} \setminus D_T$, *then* $u = v$ *in* $\overline{E_T}$.

Proof In the case that $c \le 0$ in D_T, the two conclusions are trivially followed by applying Corollary 3.6 to $u - v$.

For general case without any assumption on the sign of c, we only need to prove that if $-\frac{\partial u}{\partial t} + Lu \le 0$ in D_T and $u \ge 0$ on $\overline{E_T} \setminus D_T$, then $u \ge 0$ in $\overline{E_T}$. Define $u^\beta := ue^{-\beta t}$. Then $u \ge 0$ is equivalent to $u^\beta \ge 0$. We calculate

$$-\frac{\partial u}{\partial t} + Lu = e^{\beta t}\Big(-\frac{\partial u^\beta}{\partial t} + Lu^\beta - \beta u^\beta \Big).$$

Hence $-\frac{\partial u}{\partial t} + Lu \le 0$ is equivalent to $-\frac{\partial u^\beta}{\partial t} + (L - \beta)u^\beta \le 0$. Choose a sufficiently large β, we can ensure $c - \beta$, the zeroth-order coefficient of operator $L - \beta$, to be nonpositive in D_T. By the preceding statements, we know $u^\beta \ge 0$ in $\overline{E_T}$, equivalently, $u \ge 0$ in $\overline{E_T}$. □

Corollary 3.8 *Let* D *be an open and bounded set but not necessarily connected, and* E *be an open set satisfying* $D \subset E$. *Assume that* $u, v \in C^{2,1}(D_T) \cap C(\overline{E_T})$, $-\frac{\partial u}{\partial t} + Lu = 0$ *in* D_T, *and* $\operatorname{supp} v(t, x, \cdot) \subset \overline{E} - x$ *for each* $(x, t) \in D_T$. *If* $\max_{\overline{D_T}} c \le \beta < 0$, *then*

$$\max_{\overline{E_T}} |u| \le e^{\beta T} \max_{\overline{E_T} \setminus D_T} |u|.$$

Proof We consider the function $u^\beta := ue^{-\beta t}$ like in the previous corollary. By the same argument we know that $-\frac{\partial u}{\partial t} + Lu = 0$ is equivalent to $-\frac{\partial u^\beta}{\partial t} + (L - \beta)u^\beta = 0$. The zeroth-order coefficient of operator $L - \beta$, i.e., $c - \beta$, is nonpositive in D_T. Therefore, Assertion 3 of Corollary 3.6 implies that

$$e^{-\beta T} \max_{\overline{E_T}} |u| \le \max_{\overline{E_T}} |u^\beta| = \max_{\overline{E_T} \setminus D_T} |u^\beta| \le \max_{\overline{E_T} \setminus D_T} |u|.$$

Our result follows. □

3.2 Strong maximum principle for parabolic case

We now turn to the strong maximum principle for the parabolic Waldenfels operator $-\frac{\partial}{\partial t} + L$.

Theorem 3.9 (Strong maximum principle for parabolic Waldenfels operators) *Let D be an open and connected set but not necessarily bounded, and E be an open set satisfying $D \subset E$. Assume that $u \in C^{2,2}(D_T) \cap C(\overline{E_T})$, $-\frac{\partial u}{\partial t} + Lu \geq 0$ in D_T, and $\operatorname{supp} \nu(t, x, \cdot) \subset \overline{E} - x$ for each $(x, t) \in D_T$. Moreover, assume that the mapping $(x, t) \to \nu(t, x, \cdot)$ is continuous in D_T. If one of the following conditions holds:*

1. *$c \equiv 0$ in D_T and u achieves a (finite) maximum over $\overline{E_T}$ at a point $(x^0, t^0) \in D_T$;*
2. *$c \leq 0$ in D_T and u achieves a (finite) nonnegative maximum over $\overline{E_T}$ at a point $(x^0, t^0) \in D_T$;*
3. *u achieves a zero maximum over $\overline{E_T}$ at a point $(x^0, t^0) \in D_T$,*

then u is constant on $\overline{D_{t^0}}$, where $D_{t^0} = D \times (0, t^0]$.

A result of strong maximum principle for viscosity solutions of certain *nonlinear* nonlocal parabolic operators proved in [5] required a "nondegeneracy" condition, which is crucial in that context. But our strong maximum principle for *linear* nonlocal parabolic operator in Theorem 3.9 does not need this or any other conditions like this.

The converse case that $-\frac{\partial u}{\partial t} + Lu \leq 0$ in D_T is immediate.

Corollary 3.10 *Let D be an open and connected set but not necessarily bounded, and E be an open set satisfying $D \subset E$. Assume that $u \in C^{2,2}(D_T) \cap C(\overline{E_T})$, $-\frac{\partial u}{\partial t} + Lu \leq 0$ in D_T, and $\operatorname{supp} \nu(t, x, \cdot) \subset \overline{E} - x$ for each $(x, t) \in D_T$. Moreover, assume that the mapping $(x, t) \to \nu(t, x, \cdot)$ is continuous in D_T. If one of the following conditions holds:*

1. *$c \equiv 0$ in D_T and u achieves a (finite) minimum over $\overline{E_T}$ at a point $(x^0, t^0) \in D_T$;*
2. *$c \leq 0$ in D_T and u achieves a (finite) nonpositive minimum over $\overline{E_T}$ at a point $(x^0, t^0) \in D_T$;*
3. *u achieves a zero minimum over $\overline{E_T}$ at a point $(x^0, t^0) \in D_T$,*

then u is constant on $\overline{D_{t^0}}$, where $D_{t^0} = D \times (0, t^0]$.

To prove the strong maximum principle, we will consider the horizontal propagation of maximizer point in space by the similar arguments in elliptic case, and further obtain the vertical propagation of maximizer point locally in time by the weak maximum principle in elliptic case.

Denote $M := \max_{\overline{E_T}} u < \infty$ for convenience. Under the assumptions in Theorem 3.9, that is, $u \in C^{2,2}(D_T) \cap C(\overline{E_T})$, $-\frac{\partial u}{\partial t} + Lu \geq 0$ in D_T, and $u(x^0, t^0) = M$ with point $(x^0, t^0) \in D_T$, $\operatorname{supp} \nu(t, x, \cdot) \subset \overline{E} - x$ for each $(x, t) \in D_T$, and the mapping $(x, t) \to \nu(t, x, \cdot)$ is continuous in D_T. Furthermore, one of the following assumptions holds:

Assumption 1 $c \equiv 0$ in D_T.

Assumption 2 $c \leq 0$ in D_T and $M \geq 0$.

Assumption 3 $M = 0$.

The following lemma follows from Theorem 2.8 and Remark 2.14.

Lemma 3.11 *Let $B \subset \mathbb{R}^{d+1}$ be a open ball with $\overline{B} \subset D_T$. Assume that there exists a point $(x^0, t^0) \in \partial B$ such that $u(x^0, t^0) = M$ and $u(x, t) < M$ for each point $(x, t) \in B$. Then t^0 is either the smallest or the largest value over all the time coordinates of points in $\overline{B}$.*

Proof If $t^0 = T$, the theorem is trivial. Hence we assume $t^0 < T$. Equivalently, (x^0, t^0) is an interior point of D_T.

We regard the parabolic Waldenfels operator $-\frac{\partial}{\partial t} + L$ as a degenerate elliptic Waldenfels operator by writing

$$\begin{aligned}
&\Big(-\frac{\partial}{\partial t} + L\Big)u(x,t) \\
&\quad = -\frac{\partial u}{\partial t}(x,t) + \mathrm{tr}[\mathbf{a}^T(\nabla^2 u)](x,t) + b^T \nabla u(x,t) + cu(x,t) \\
&\qquad + \int_{\mathbb{R}^d \setminus \{0\}} \Big[u(x+z,t) - u(x,t) - z^T \nabla u(x,t) \cdot \mathbf{1}_{\{|z|<1\}}\Big] \nu(t,x,dz) \\
&\quad = \mathrm{tr}\left[\begin{pmatrix} \mathbf{a} & 0 \\ 0 & 0 \end{pmatrix}^T (\nabla^2_{x,t} u)\right](x,t) + (b^T, -1) \cdot \nabla_{x,t} u(x,t) + cu(x,t) \\
&\qquad + \int_{\mathbb{R}^{d+1}} \Big[u(x+z,t+s) - u(x,t) \\
&\qquad\quad - (z^T, s) \cdot \nabla_{x,t} u(x,t) \cdot \mathbf{1}_{\{|z|^2+|s|^2 \le 1\}}\Big] \nu(t,x,dz)\delta_0(ds).
\end{aligned}$$

Thus, we can replace the matrix $\mathbf{a}$ in the elliptic Waldenfels operator (5) by $\hat{\mathbf{a}} = \begin{pmatrix} \mathbf{a} & 0 \\ 0 & 0 \end{pmatrix}$, vector b by $\hat{b} = \begin{pmatrix} b \\ -1 \end{pmatrix}$, and the kernel $\nu(x, dz)$ by $\hat{\nu}((x,t), dzds) := \nu(t, x, dz)\delta_0(ds)$.

Now we verify that the kernel $\hat{\nu}$, defined on $\mathbb{R}^{d+1}$, satisfies the moment condition (3), although its support is not contained in $\mathbb{R}^{d+1} \setminus \{\mathbf{0}\}$. By recalling (29), condition (3) for $\hat{\nu}$ immediately follows as we see that

$$\begin{aligned}
&\int_{\mathbb{R}^{d+1}} \Big[1 \wedge (|z|^2 + |s|^2)\Big] \hat{\nu}(x,t,dzds) \\
&\quad = \int_{\mathbb{R}^{d+1}} \Big[1 \wedge (|z|^2 + |s|^2)\Big] \nu(t,x,dz)\delta_0(ds) \\
&\quad = \int_{\mathbb{R}^d \setminus \{0\}} (1 \wedge |z|^2) \nu(t,x,dz) \\
&\quad < \infty.
\end{aligned} \tag{33}$$

Since (x^0, t^0) is the maximizer point over $\overline{E_T}$ of u, we may replace the set E_0 in (19) by

$$\hat{E}_0 := \{(x,t) \in \overline{E} \times \{t^0\} \mid u(x,t) = M\}.$$

Due to the fact that the support of measure $\hat{\nu}((x^0, t^0), \cdot)$ is contained in $\overline{E} \times \{0\}$, we can further derive (21) for $\hat{\nu}$ and $\hat{E}_0$. It turns out that (24) and (25), (26) also hold in this situation.

Combining Remark 2.14 and the preceding arguments , we conclude that, as in Theorem 2.8, $\frac{\partial u}{\partial \mathbf{n}} > 0$ holds in the case that the unit outer normal vector $\mathbf{n}$ of (x^0, t^0) over ∂B is not in the nullspace of $\mathbf{a}(x^0, t^0)$, which is exactly the space $N := \{(x, t) \mid x = 0\}$. But this leads to a contradiction: since u attains a maximum at the interior point (x^0, t^0), we have $\frac{\partial u}{\partial \mathbf{n}} = 0$. Therefore, (x^0, t^0) must be a pole of ball B, whose unit outer normal vector is just in N. This completes the proof. □

The next lemma shows that for every $t \in (0, T)$, we have either $u(x, t) < M$ or $u(x, t) = M$ for all $x \in D$. This means that the non-maximizer point (or the maximizer point) may propagation horizontally in space. The proof can be found in [24], we do not present it here and the main points of the proof can be found in Sect. 5.

Lemma 3.12 *Assume that $u(x^0, t^0) < M$ with $x^0 \in D$ and $t^0 \in (0, T)$. Then $u(x, t^0) < M$ for every $x \in D$.*

Remark 3.13 In Lemma 3.12, we restrict the set D to be connected to make sure that the point (x^1, t^0) can be chosen. More generally, if D is not connected, we may replace D in the previous proof with the connected component of D which contains the maximizer point. We thus conclude that for every fixed $t \in (0, T)$, either $u(x, t) < M$ or $u(x, t) = M$ holds in each connected component of D. Then as in Remark 2.12, we see that the diffusion term gives rise to the horizontal propagation of maximizer point in the corresponding connected component.

As in Remark 2.9, or from [5,7], we also see that the horizontal propagation of maximizer point by translation of measure support. Namely, if $u(x^0, t^0) = M$ with $x^0 \in D$ and $t^0 \in (0, T)$, then $u \equiv M$ on the set $\overline{\bigcup_{n=0}^{\infty} \Lambda_n}$, where $\tilde{\Lambda}_n$'s are defined by induction,

$$\tilde{\Lambda}_0 = x^0, \tilde{\Lambda}_{n+1} = \bigcup_{x \in D \cap \tilde{\Lambda}_n} [\operatorname{supp} \nu(t^0, x, \cdot) + x].$$

Thus in this scheme, the jump diffusion term leads to the horizontal propagation of maximizer point between those connected components, since jumps from one connected component to another might occur when measure supports overlap two or more connected components.

Furthermore, we present the final lemma we need. It means the maximizer point may propagate vertically in time in a local sense. The proof can also be found in Sect. 5

Lemma 3.14 *Assume that $u < M$ in $D \times (t^0, t^1)$, with $0 \le t^0 < t^1 \le T$. Then $u < M$ in $D \times \{t^1\}$.*

Finally we can prove Theorem 3.9.

Proof of Theorem 3.9 Set $D_T^{<} := \{(x, t) \in D_T \mid u(x, t) < M\}$. Then $D_T^{<}$ is a relatively open subset of D_T. From Lemma 3.12, we know that for each fixed $t \in$

$(0, T)$, either $u(x, t) < M$ or $u(x, t) = M$ holds for all $x \in D$. Therefore, $D_T^{\leq}$ must be of the form $D \times I$, for some $I \subset (0, T]$ relatively open in $(0, T]$.

For fixed $s \in I$ and $s \neq T$, define $\tau(s) := \sup\{t \mid (s, t) \subset I\}$, where the set in supremum is never empty as I is relatively open in $(0, T]$. From Lemma 3.14, we see $\tau(s) \in I$. Then $(r, \tau(s)]$ is the connected component in I containing s, for some $r < s$. Consequently $\tau(s) = T$, since I is relatively open. Thus we summarize that for each $s \in I$, $[s, T] \subset I$, which is trivial when $s = T$. Hence, the relatively open set I only has two options: either $[0, T]$ or $(s, T]$ for some $s \in [0, T)$. In light of the fact $u(x^0, t^0) = M$, or equivalently $t^0 \in I$, we conclude that I must be of the form $(s, T]$ for some $s \in [t^0, T)$, as required. This finishes the proof of Theorem 3.9. □

4 Examples

We will give some examples in this section. These examples are all concerned with symmetric α-stable Lévy noise which are not covered in Taira's framework in [30], since the jump measure is of unbounded support.

Example 4.1 (**Mean exit time**) Consider a stochastic system in $\mathbb{R}^d$:

$$dX_t = b(X_t)dt + dW_t + dL_t^{\alpha},$$

where W_t is a standard Wiener process, and L_t^{α} is a Lévy process with jump measure $\nu_\alpha(dz) = c_{\alpha,d}\frac{dz}{|z|^{d+\alpha}}$, for $\alpha \in (0, 2)$ and $c_{\alpha,d}$ a positive constant depending on α and d, together with zero drift and zero diffusion. The generator for this system is the following elliptic Waldenfels operator

$$\begin{aligned} Lu(x) =& \frac{1}{2}\Delta u(x) + b(x) \cdot \nabla u(x) \\ &+ \int_{\mathbb{R}^d \setminus \{0\}} \left[u(x + z) - u(x) - \sum_{j=1}^{d} z_j \frac{\partial u}{\partial x_j}(x)\mathbf{1}_{\{|z|<1\}}\right]\nu_\alpha(dz). \end{aligned}$$

Let D be a domain in $\mathbb{R}^d$. The mean exit time for X_t, starting at x, exits firstly from D is denoted by $\tau(x)$. By Dynkin formula for such jump diffusion process [3,13], as shown in [8,21,25], we know that τ satisfies the following equation,

$$\begin{cases} L\tau = -1 & \text{in } D, \\ \tau = 0 & \text{in } D^c. \end{cases}$$

By the strong maximum principle 2.8, or precisely Corollary 2.11 with the special case $E = \mathbb{R}^d$, we conclude that the mean exit time τ cannot take zero value inside D, unless it is constant (inside the domain D).

Example 4.2 (**Escape probability**) Similarly, let U be a subset of D^c. The likelihood that X_t, starting at x, exits firstly from D by landing in the target set U is called the

escape probability from D to U, denoted by $p(x)$. As shown in [20,23], the escape probability p satisfies the following equation,

$$\begin{cases} Lp = 0 & \text{in } D, \\ p = 1 & \text{in } U, \\ p = 0 & \text{in } D^c \setminus U. \end{cases}$$

By Theorem 2.8 and Corollary 2.11 with $E = \mathbb{R}^d$, we conclude that p cannot take values of zero or one at any point inside D.

Example 4.3 (Fokker–Planck equation) Consider a stochastic system in $\mathbb{R}^d$:

$$\begin{cases} dX_t = b(X_t)dt + dW_t + dL_t^\alpha, \\ X_0 = x^0, \end{cases} \tag{34}$$

where W_t is a standard Wiener process and L_t^α is a Lévy process with pure jump measure $\nu_\alpha(dz) = c_{\alpha,d}\frac{dz}{|z|^{d+\alpha}}$, for $\alpha \in (0,2)$ and $c_{\alpha,d}$ a positive constant depending on α and d. The Fokker–Planck equation for the probability density of the solution, as shown in [8,10], is

$$\begin{aligned} \frac{\partial p}{\partial t} &= \frac{1}{2}\Delta p - b(x)\cdot\nabla p - (\nabla\cdot b)p \\ &\quad + \int_{\mathbb{R}^d\setminus\{0\}} \Big[p(x+z,t) - p(x,t) - \sum_{j=1}^{d} z_j \frac{\partial p}{\partial x_j}(x,t)\mathbf{1}_{\{|z|<1\}}\Big]\nu_\alpha(dz). \end{aligned} \tag{35}$$

Let D be a domain in $\mathbb{R}^d$. In this case, the coefficient of zeroth-order term is $c = -\nabla\cdot b$. We apply the strong maximum principle in Theorem 3.9 and Corollary 3.10 with $E = \mathbb{R}^d$. If $\nabla\cdot b \equiv 0$, which means the deterministic vector field of stochastic system (34) is divergence-free, then the probability density p cannot attain its maximum (or minimum) over $\mathbb{R}^d \times [0,\infty)$ in $D \times [0,\infty)$, unless it is constant at all time before the maximizer (or minimizer) point. Moreover, if $\nabla\cdot b \geq 0$, then p cannot attain its maximum or zero minimum over $\mathbb{R}^d \times [0,\infty)$ in $D \times [0,\infty)$ (note that p only takes nonnegative values), unless it is constant at all time before this point as well.

Acknowledgements This work was partly supported by the National Science Foundation Grant DMS-1620449. We appreciate the helpful comments and suggestions of the anonymous reviewer.

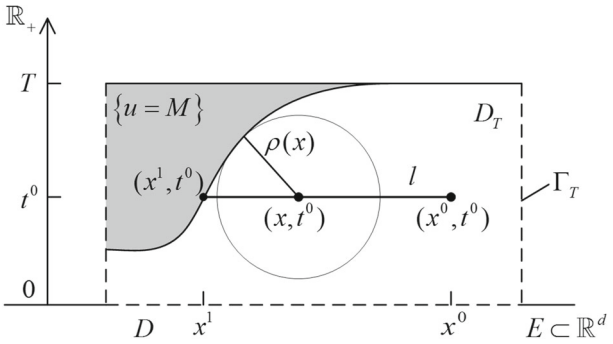

Fig. 2 Sketch for Lemma 3.12

5 Appendix: Proofs for lemmas

The proofs of some technical lemmas will be presented here for the sake of completeness.

5.1 Proof of Lemma 3.12

Assume on the contrary that $D \times \{t^0\}$ contains some points at which $u = M$. Choose the point (x^1, t^0) nearest to (x^0, t^0) at the line $\{(x, t^0) \mid x \in D\}$, such that $u(x^1, t^0) = M$, which is possible since D is connected and the set $\{(x, t) \in D_T \mid u(x, t) = M\}$ (or simply, $\{u = M\}$) is relatively closed in D_T. Denote by l the line segment connecting (x^0, t^0) with (x^1, t^0), and set

$$\delta = \min\{|x^1 - x^0|, \operatorname{dist}(l, \partial D)\}.$$

For $x \in l_0 := \{x \mid (x, t^0) \in l, 0 < |x - x^1| < \delta\}$, define

$$\rho(x) = \operatorname{dist}((x, t^0), (D_T)^\circ \cap \{u = M\}).$$

Obviously, $0 < \rho(x) \leq |x - x^1|$ in l_0. See Fig. 2.

Consider next the open ball $B := B((x, t^0), \rho(x)) \subset \mathbb{R}^{d+1}$ with center (x, t^0) and radius $\rho(x)$. Then $\overline{B} \subset D_T$, $u < M$ in B and ∂B contains points where $u = M$. Thus by Lemma 3.11 we conclude that either $u(x, t^0 + \rho(x)) = M$ or $u(x, t^0 - \rho(x)) = M$. By the Pythagorean theorem, we assert

$$\rho(x + \epsilon e)^2 \leq \rho(x)^2 + \epsilon^2,$$

for sufficiently small $|\epsilon| > 0$, where e is the unit vector along l. In the same way, $\rho(x)^2 \leq \rho(x + \epsilon e)^2 + \epsilon^2$. We note that

$$\lim_{\epsilon \to 0} \frac{\sqrt{\rho(x)^2 - \epsilon^2} - \rho(x)}{\epsilon} \leq \lim_{\epsilon \to 0} \frac{\rho(x + \epsilon e) - \rho(x)}{\epsilon} \leq \lim_{\epsilon \to 0} \frac{\sqrt{\rho(x)^2 + \epsilon^2} - \rho(x)}{\epsilon},$$

which implies that

$$\left.\frac{d\sqrt{\rho(x)^2-\epsilon^2}}{d\epsilon}\right|_{\epsilon=0} \le \left.\frac{d\rho(x+\epsilon e)}{d\epsilon}\right|_{\epsilon=0} \le \left.\frac{d\sqrt{\rho(x)^2+\epsilon^2}}{d\epsilon}\right|_{\epsilon=0}.$$

This leads to

$$\left.\frac{d\rho(x+\epsilon e)}{d\epsilon}\right|_{\epsilon=0} = 0,$$

for each $x \in l_0$. Consequently, $\rho(x)$ is a constant in l_0, which is a contradiction since $\rho(x) > 0$ in l_0 and $\rho(x) \to 0$ as $x \to x^1$. The proof is complete.

5.2 Proof of Lemma 3.14

Assume that there exists a point $x^1 \in D$ such that $u(x^1, t^1) = M$. Let $\hat{B} := B((x^1, t^1), r) \subset \mathbb{R}^{d+1}$ be such a small ball that $\hat{B} \cap (E \times (t^0, t^1))$ is contained in D_T. Define

$$v(x,t) := e^{-|x-x^1|^2-\beta(t-t^1)} - 1 \quad \text{for } (x,t) \in \overline{E} \times [t^0, t^1],$$

where β is a positive constant as selected below. Then for $(x,t) \in \hat{B} \cap (E \times (t^0, t^1))$,

$$\frac{\partial v}{\partial t} = -\beta e^{-|x-x^1|^2-\beta(t-t^1)},$$

and also

$$\begin{aligned} Av &= e^{-|x-x^1|^2-\beta(t-t^1)}\Big\{\mathrm{tr}\big[\mathbf{a}^T\big(4(x-x^1)\otimes(x-x^1)-2I\big)\big] - 2b^T(x-x^1) \\ &\qquad + c\big(1-e^{-\beta(t^1-t)+|x-x^1|^2}\big)\Big\} \\ &= e^{-|x-x^1|^2-\beta(t-t^1)}\big[4(x-x^1)^T\mathbf{a}^T(x-x^1) - 2\mathrm{tr}(\mathbf{a}) - 2b^T(x-x^1) \\ &\qquad + c\big(1-e^{-\beta(t^1-t)+|x-x^1|^2}\big)\big] \\ &\ge e^{-|x-x^1|^2-\beta(t-t^1)}\big[4\gamma|x-x^1|^2 - 2\mathrm{tr}(\mathbf{a}) - 2|b||x-x^1| - |c|\big], \end{aligned}$$

for $\beta > 0$ large enough. We set

$$\big|4\gamma|x-x^1|^2 - 2\mathrm{tr}(\mathbf{a}) - 2|b||x-x^1| - |c|\big| \le C_1 < \infty,$$

for some positive constant C_1, independent of β. Moreover, we have

$$\begin{aligned} Kv(x,t) = e^{-|x-x^1|^2-\beta(t-t^1)} \int_{\mathbb{R}^d\setminus\{0\}} \big[e^{-|x+z-x^1|^2+|x-x^1|^2} - 1 \\ + 2z^T(x-x^1)\mathbf{1}_{\{|z|<1\}}\big]\nu(t,x,dz). \end{aligned}$$

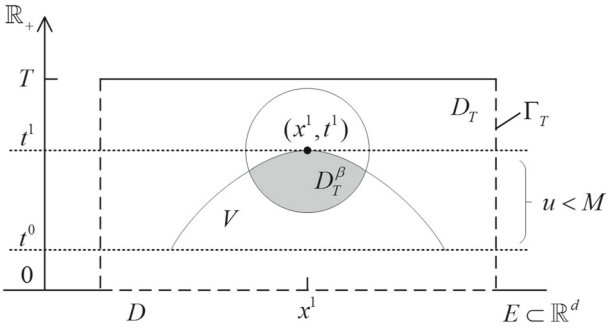

Fig. 3 Sketch for Lemma 3.14

Using Taylor expansions and the fact that each $\nu(t, x, \cdot)$ is Lévy measure, we know that

$$\left| \int_{\mathbb{R}^d \setminus \{0\}} \left[e^{-|x+z-x^1|^2+|x-x^1|^2} - 1 + 2z^T (x - x^1) \mathbf{1}_{\{|z|<1\}} \right] \nu(t, x, dz) \right| \leq C_2 < \infty,$$

where C_2 is a positive constant independent of β. Hence, in $\hat{B} \cap (E \times (t^0, t^1))$

$$\begin{aligned} -\frac{\partial v}{\partial t} + Lv &\geq -\frac{\partial v}{\partial t} + Av - |Kv| \\ &\geq e^{-|x-x^1|^2-\beta(t-t^1)}(\beta - C_1 - C_2) \\ &\geq 0, \end{aligned} \tag{36}$$

provided β is sufficiently large.

Next limit our attention within the domain

$$D_T^\beta := \{(x, t) \in \hat{B} \mid |x - x^1|^2 + \beta(t - t^1) < 0\},$$

and define

$$u^\epsilon = u + \epsilon v - M \quad \text{on } \overline{E} \times [t^0, t^1],$$

for a constant $\epsilon > 0$. Then choosing ϵ small enough, we have

$$u^\epsilon \leq 0 \quad \text{on } (\overline{E} \times [t^0, t^1]) \setminus D_T^\beta,$$

since $v \leq 0$ on $\overline{E} \times [t^0, t^1] \setminus V$, and $u < M$ on $V \setminus D_T^\beta$, where $V := \{(x, t) \in E \times (t^0, t^1) \mid |x - x^1|^2 + \beta(t - t^1) < 0\} = (E \times (t^0, t^1)) \cap \{v > 0\}$. See Fig. 3.

Under Assumptions 1 and 2, $c \equiv 0$ in D_T, or $c \leq 0$ in D_T also $M \geq 0$, from (36) and the assumption $-\frac{\partial u}{\partial t} + Lu \geq 0$ in D_T, we compute

$$-\frac{\partial u^\epsilon}{\partial t} + Lu^\epsilon \geq -cM \geq 0 \quad \text{in } D_T^\beta.$$

Then Remark 3.4, or more precisely, the weak maximum principle for elliptic case in Theorem 2.2, implies that $u^\epsilon \leq 0$ throughout $\overline{E} \times [t^0, t^1]$. But $u^\epsilon(x^1, t^1) = 0$, and thus at the point (x^1, t^1),

$$0 \leq \frac{\partial u^\epsilon}{\partial t} = \frac{\partial u}{\partial t} + \epsilon \frac{\partial v}{\partial t} = \frac{\partial u}{\partial t} - \epsilon\beta,$$

that is, $\frac{\partial u}{\partial t}(x^1, t^1) > 0$. However, as explained already in Theorem 2.2 or Theorem 3.1, we know that $Lu \leq 0$ at point (x^1, t^1). Hence

$$-\frac{\partial u}{\partial t} + Lu < 0 \quad \text{at } (x^1, t^1).$$

This is a contradiction.

Under Assumption 3, we may use the same argument as in the end of the proof of Theorem 2.8, that is, replacing L and c respectively with $L - c^+$ and $c - c^+$ and applying the acquired result for Assumption 2, the same result for this case follows.

References

1. Albeverio, S., Rüdiger, B., Wu, J.-L.: Invariant measures and symmetry property of Lévy type operators. Potential Anal. **13**(2), 147–168 (2000)
2. Albeverio, S., Rüdiger, B., Wu, J.-L.: Analytic and probabilistic aspects of Lévy processes and fields in quantum theory. In: Barndorff-Nielsen, O., Mikosch, T., Resnick, S. (eds.) Lévy Processes: Theory and Applications, pp. 187–224. Birkhäuser, Boston (2001)
3. Applebaum, D.: Lévy Processes and Stochastic Calculus, 2nd edn. Cambridge University Press, Cambridge (2009)
4. Bony, J.-M., Courrège, P., Priouret, P.: Semi-groupes de feller sur une variété à bord compacte et problèmes aux limites intégro-différentiels du second ordre donnant lieu au principe du maximum. Annales de l'institut Fourier **18**(2), 369–521 (1968)
5. Ciomaga, A.: On the strong maximum principle for second order nonlinear parabolic integro-differential equations. Adv. Differ. Equ. **17**(7/8), 635–671 (2012)
6. Courrège, P.: Sur la forme intégro-différentielle des opérateurs de C_k^∞ dans C satisfaisant au principe du maximum. Séminaire Brelot-Choquet-Deny. Théorie du Potentiel **10**(1), 1–38 (1965)
7. Coville, J.: Remarks on the strong maximum principle for nonlocal operators. Electron. J. Differ. Equ. **2008**(66), 1–10 (2008)
8. Duan, J.: An Introduction to Stochastic Dynamics. Cambridge University Press, New York (2015)
9. Fukushima, M., Oshima, Y., Takeda, M.: Dirichlet Forms and Symmetric Markov Processes, vol. 19 of De Gruyter Studies in Mathematics. Walter de Gruyter, Berlin (2010)
10. Garroni, M.G., Menaldi, J.L.: Green Functions for Second Order Parabolic Integro-Differential Problems, vol. 275 of Chapman& Hall/CRC Research Notes in Mathematics Series. Chapman & Hall/CRC, London (1992)
11. Jacob, N.: Pseudo differential operators & Markov Processes: Markov Processes and Applications, vol. 1, 2, 3. Imperial College Press, London (2001, 2002, 2005)
12. Kallenberg, O.: Foundations of Modern Probability. Probability and Its Applications, 2nd edn. Springer, New York (2006)
13. Kolokoltsov, V.N.: Symmetric stable laws and stable-like jump-diffusions. Proc. Lond. Math. Soc. **80**(03), 725–768 (2000)
14. Kolokoltsov, V.N.: Markov Processes, Semigroups, and Generators, vol. 38 of De Gruyter Studies in Mathematics. Walter de Gruyter, Berlin (2011)
15. Komatsu, T.: Markov processes associated with certain integro-differential. Osaka J. Math. **10**, 271–303 (1973)

16. Komatsu, T.: Pseudo-differential operators and Markov processes. J. Math. Soc. Jpn. **36**(3), 387–418 (1984)
17. Komatsu, T.: Continuity estimates for solutions of parabolic equations associated with jump type Dirichlet forms. Osaka J. Math. **25**(3), 697–728 (1988)
18. Komatsu, T.: Uniform estimates for fundamental solutions associated with non-local Dirichlet forms. Osaka J. Math. **32**(4), 833–860 (1995)
19. Kuwae, K.: Maximum principles for subharmonic functions via local semi-Dirichlet forms. Can. J. Math. **60**(4), 822–874 (2008)
20. Liao, M.: The Dirichlet problem of a discontinuous Markov process. Acta Math. Sin. **5**(1), 9–15 (1989)
21. Øksendal, B.K., Sulem, A.: Applied Stochastic Control of Jump Diffusions, 2nd edn. Springer, Berlin (2007)
22. Protter, M.H., Weinberger, H.F.: Maximum Principles in Differential Equations. Springer, New York (1984)
23. Qiao, H., Kan, X., Duan, J.: Escape probability for stochastic dynamical systems with jumps. Proc. Math. Stat. **34**, 195–216 (2013)
24. Renardy, M., Rogers, R.C.: An Introduction to Partial Differential Equations, vol. 13 of Texts in Applied Mathematics, 2nd edn. Springer, New York (2004)
25. Sato, K.-I.: Lévy Processes and Infinitely Divisible Distributions, 2nd edn. Cambridge University Press, Cambridge (1999)
26. Schertzer, D., Larchevêque, M., Duan, J., Yanovsky, V.V., Lovejoy, S.: Fractional Fokker–Planck equation for nonlinear stochastic differential equations driven by non-Gaussian Lévy stable noises. J. Math. Phys. **42**(1), 200–212 (2001)
27. Shlesinger, M.F., Zaslavsky, G.M., Frisch, U.: Lévy Flights and Related Topics in Physics, vol. 450 of Lecture Notes in Physics. Springer, Berlin (1995)
28. Stroock, D.W.: Diffusion processes associated with Lévy generators. Probab. Theory Relat. Fields **32**(3), 209–244 (1975)
29. Stroock, D.W.: Markov Processes from K. Itô's Perspective, vol. 1555 of Annals of Mathematics Studies. Princeton University Press, Princeton (2003)
30. Taira, K.: Semigroups, Boundary Value Problems and Markov Processes. Springer Monographs in Mathematics. Springer, Berlin (2004)
31. Truman, A., Wu, J.-L.: On a stochastic nonlinear equation arising from 1D integro-differential scalar conservation laws. J. Funct. Anal. **238**(2), 612–635 (2006)
32. Waldenfels, W.v.: Positive Halbgruppen auf einem n-dimensionalen Torus. Archiv der Mathematik **15**(1), 191–203 (1964)