

On the martingale problem and Feller and strong Feller properties for weakly coupled Lévy type operators

Fubao Xi^a, Chao Zhu^{b,*}

^a School of Mathematics and Statistics, Beijing Institute of Technology, Beijing 100081, China

^b Department of Mathematical Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, United States

Received 17 February 2016; received in revised form 19 September 2017; accepted 11 February 2018

Available online 21 February 2018

Abstract

This paper considers the martingale problem for a class of weakly coupled Lévy type operators. It is shown that under some mild conditions, the martingale problem is well-posed and uniquely determines a strong Markov process (X, A) . The process (X, A) , called a regime-switching jump diffusion with Lévy type jumps, is further shown to possess Feller and strong Feller properties under non-Lipschitz conditions via the coupling method.

© 2018 Elsevier B.V. All rights reserved.

MSC: 60J25; 60J27; 60J60; 60J75

Keywords: Weakly coupled Lévy type operator; Martingale problem; Feller property; Strong Feller property; Coupling method

1. Introduction

This paper deals with the martingale problem for a weakly coupled Lévy type operator \mathcal{A} defined as follows. Let d and n_0 be two positive integers and set $\mathbb{S} := \{1, 2, \dots, n_0\}$. For all “nice” functions $f: \mathbb{R}^d \times \mathbb{S} \rightarrow \mathbb{R}$, we define

$$\mathcal{A}f(x, k) := \mathcal{L}_k f(x, k) + \mathcal{Q}(x)f(x, k). \quad (1.1)$$

* Corresponding author.

E-mail addresses: xifb@bit.edu.cn (F. Xi), zhu@uwm.edu (C. Zhu).

Here, for each $k \in \mathbb{S}$, \mathcal{L}_k is a Lévy type operator defined as follows:

$$\begin{aligned} \mathcal{L}_k f(x, k) &:= \frac{1}{2} \text{tr}(a(x, k) \nabla^2 f(x, k)) + \langle b(x, k), \nabla f(x, k) \rangle \\ &\quad + \int_{\mathbb{R}^d} (f(x+u, k) - f(x, k) - \langle \nabla f(x, k), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u)) \nu(x, k, du), \end{aligned} \quad (1.2)$$

where for each $(x, k) \in \mathbb{R}^d \times \mathbb{S}$, $a(x, k) = (a_{ij}(x, k)) \in \mathbb{R}^{d \times d}$ is symmetric and nonnegative definite, $b(x, k) = (b_i(x, k)) \in \mathbb{R}^d$, and $\nu(x, k, \cdot)$ is a Lévy kernel such that for each (x, k) , $\nu(x, k, \cdot)$ is a nonnegative σ -finite measure on \mathbb{R}^d satisfying

$$\int_{\mathbb{R}^d} |u|^2 \mathbf{1}_{B(0, \varepsilon_0)}(u) \nu(x, k, du) < \infty \text{ and } \nu(x, k, \mathbb{R}^d \setminus B(0, \varepsilon_0)) < \infty, \quad (1.3)$$

where $\varepsilon_0 > 0$ (one can usually take $\varepsilon_0 = 1$). Here and hereafter, $\nabla f(\cdot, k)$ and $\nabla^2 f(\cdot, k)$ denote respectively the gradient and Hessian matrix of $f(\cdot, k)$, $\langle \cdot, \cdot \rangle$ denotes the inner product in \mathbb{R}^d , $\mathbb{R}_0^d := \mathbb{R}^d \setminus \{0\}$, and $B(0, r) := \{x \in \mathbb{R}^d : |x| < r\}$ for $r > 0$. In (1.1) and throughout the paper, the switching operator $Q(x)$ is defined as follows:

$$Q(x)f(x, k) := \sum_{l \in \mathbb{S}} q_{kl}(x)(f(x, l) - f(x, k)), \quad (1.4)$$

where $Q(x) = (q_{kl}(x))$ is an $n_0 \times n_0$ matrix-valued measurable function on \mathbb{R}^d such that for all $x \in \mathbb{R}^d$ we have $q_{kl}(x) \geq 0$ for $k \neq l$, and for each $k \in \mathbb{S}$, $\sum_{l \in \mathbb{S}} q_{kl}(x) = 0$.

In this paper, we consider the martingale problem for the weakly coupled Lévy type operator \mathcal{A} defined in (1.1) on $\Omega := D([0, \infty), \mathbb{R}^d \times \mathbb{S})$, the space of right continuous functions on $[0, \infty)$ into $\mathbb{R}^d \times \mathbb{S}$ having left limits endowed with the Skorohod topology. Let \mathcal{F}_t be the σ -field generated by the cylindrical sets on $D([0, \infty), \mathbb{R}^d \times \mathbb{S})$ up to time t and set $\mathcal{F} = \bigvee_{t=0}^{\infty} \mathcal{F}_t$. Next, let $C_c^\infty(\mathbb{R}^d \times \mathbb{S})$ denote the family of functions defined on $\mathbb{R}^d \times \mathbb{S}$ such that $f(\cdot, k) \in C_c^\infty(\mathbb{R}^d)$ with $k \in \mathbb{S}$, where $C_c^\infty(\mathbb{R}^d)$ denotes the family of functions defined on \mathbb{R}^d which are infinitely differentiable and have compact supports.

Definition 1.1. For a given $(x, k) \in \mathbb{R}^d \times \mathbb{S}$, we say a probability measure $\mathbb{P}^{(x, k)}$ on $D([0, \infty), \mathbb{R}^d \times \mathbb{S})$ is a solution to the martingale problem for the operator \mathcal{A} starting from (x, k) , if $\mathbb{P}^{(x, k)}((X(0), \Lambda(0)) = (x, k)) = 1$ and for each function $f \in C_c^\infty(\mathbb{R}^d \times \mathbb{S})$,

$$M_t^{(f)} := f(X(t), \Lambda(t)) - f(X(0), \Lambda(0)) - \int_0^t \mathcal{A}f(X(s), \Lambda(s)) ds \quad (1.5)$$

is an $\{\mathcal{F}_t\}$ -martingale with respect to $\mathbb{P}^{(x, k)}$, where (X, Λ) is the coordinate process defined by $(X(t, \omega), \Lambda(t, \omega)) = \omega(t) \in \mathbb{R}^d \times \mathbb{S}$ for all $t \geq 0$ and $\omega \in \Omega$.

Sometimes, we say that the probability measure $\mathbb{P}^{(x, k)}$ is a martingale solution for the operator \mathcal{A} starting from (x, k) . We often call the coordinate process (X, Λ) the *regime-switching jump diffusion with Lévy type jumps*.

Since the seminal work of Stroock and Varadhan [29,30] on martingale problems for second order diffusion operators, the notion of martingale problems have been extensively studied for various processes in the literature. For example, Komatsu [12] and Stroock [28] prove that the martingale problem for a Lévy type operator is well-posed; Bass [1] investigates the martingale problem for pure jump Markov processes; Dawson and Zheng [6] and Feng and Zheng [7] consider the martingale problem for a class of nonlinear master equations for chemical reaction models; Xi [34] and Zheng and Zheng [42] discuss the martingale problem

for Q -processes; Zambotti [41] provides an analytic approach for existence and uniqueness for martingale problems in infinite dimensions; Kurtz [13] presents martingale problems for conditional distributions of Markov processes; Mikulevicius and Rozovskii [18] study martingale problems for stochastic partial differential equations; Perkins [19] investigates the martingale problem for interactive measure-valued branching diffusions; Hoh [9] investigates the martingale problems for pseudo-differential operators; and Bass and Tang [2] are devoted to the martingale problem for stable-like processes.

This paper is motivated by Stroock [28] and considers weakly coupled Lévy type operator \mathcal{A} defined in (1.1). Roughly speaking, in addition to the diffusion term, the drift term, and the jump term spelled out in (1.2) for each $k \in \mathbb{S}$, \mathcal{A} also contains a component $Q(x)$ defined in (1.4), which provides the switching mechanism for the operators \mathcal{L}_k , $k \in \mathbb{S}$. In other words, the operators \mathcal{L}_k , $k \in \mathbb{S}$ are coupled through the operator $Q(x)$ of (1.4). Therefore it is convenient to call the operator \mathcal{A} of (1.1) a weakly coupled Lévy type operator and the coordinate process (X, A) a *regime-switching Lévy type process*. Here we remark that $Q(x) = (q_{kl}(x))$ depends on x . When the Lévy kernel $\nu(x, k, dz)$ is independent of (x, k) , then \mathcal{A} reduces to the infinitesimal generator of a regime-switching jump diffusion process as those considered in [36,38,43]. Thanks to their ability in incorporating both structural changes and jumps of various sizes, regime-switching (jump) diffusion processes have attracted many interests lately. See, for example, [5,16,24,25,33,35–40,43] and references therein for investigations of such processes and their applications in areas such as inventory control, ecosystem modeling, manufacturing and production planning, financial engineering, and risk theory.

However, we notice that in these papers, the jump mechanism is usually assumed to be a finite or a Lévy measure $\nu(dz)$. The study of regime-switching jump diffusions with Lévy type jumps is relatively scarce, which is precisely the focus of this paper. In addition, in lieu of the stochastic differential equation approach in the aforementioned papers, this paper begins with the martingale problem for the weakly coupled Lévy type operator \mathcal{A} of (1.1). We prove that under very mild conditions, the martingale problem for the operator \mathcal{A} is well-posed. That is, we show that for any $(x, k) \in \mathbb{R}^d \times \mathbb{S}$, there is exactly one martingale solution for the operator \mathcal{A} starting from (x, k) . This is achieved in two steps. In the first step, we assume that Q of (1.4) takes a special form (\hat{Q} in (2.1)); consequently \mathcal{A} of (1.1) reduces to $\hat{\mathcal{A}}$ of (2.5). For such a special operator $\hat{\mathcal{A}}$, under Assumption 1.2, we manipulate the Stroock–Varadhan piecing together method (refer to §6.1 of [31]) to construct a martingale solution for the operator $\hat{\mathcal{A}}$ with an arbitrary initial condition $(x, k) \in \mathbb{R}^d \times \mathbb{S}$ and further show that this solution is weakly unique in Theorem 2.2. The second step deals with the general case when $Q(x)$ of (1.4) is x -dependent. For such a case, we utilize the likely ratio martingale M defined in (3.1) to establish the desired existence and uniqueness result in Theorem 3.6. One of the key steps in this approach is to show that the switching times and the jump times are mutually disjoint with probability one; see Proposition 3.5 for details. Such a strategy of using the likelihood ratio martingale was used in the recent paper Xi [36], where the jump component is driven by a finite measure. In this paper, we develop this approach to handle the general weakly coupled Lévy type operator \mathcal{A} .

Having established that the martingale problem for \mathcal{A} is well posed, we then have determined a strong Markov process (X, A) with state space $\mathbb{R}^d \times \mathbb{S}$. The second part of this paper proves that such a process possesses the Feller and strong Feller properties. Here the main tool is the coupling method. For the introduction to coupling method and its applications in various areas of probability and stochastic analysis, we refer to Chen [3], Hairer et al. [8], Lindvall [14], Lindvall and Rogers [15], Priola and Wang [21], Wang [32] and the references therein. In this paper, we first use the coupling method to show that for each $k \in \mathbb{S}$, the process $\tilde{X}^{(k)}$ corresponding to the

Lévy type operator \mathcal{L}_k of (1.2) is Feller under Assumption 4.1, in which the coefficients, and in particular, the Lévy type kernel of the operator \mathcal{L}_k , are *non-Lipschitz* in the x variable. In order to establish the Feller property for the process (X, A) , we kill the Lévy type process $\tilde{X}^{(k)}$ at rate $-q_{kk}$ to obtain the process $X^{(k)}$; see (4.5) for details. A mild condition on the functions $q_{kl}(x)$ (Assumption 4.2) then helps us to derive the Feller property for the killed Lévy type process $X^{(k)}$ in Lemma 4.7. Finally we use a series representation for the resolvent G_α of the process (X, A) and a result in [17] to establish the Feller property for the process (X, A) ; this is spelled out in Theorem 4.4.

Next we use a similar approach to establish the strong Feller property for the process (X, A) in Section 5. More precisely, inspired by Priola and Wang [21], we use a combination of reflection and marching coupling for the operator \mathcal{L}_k to establish the strong Feller property for the processes $\tilde{X}^{(k)}$ and $X^{(k)}$ in Proposition 5.3. Again, we allow the coefficients and the Lévy type kernel of the operator \mathcal{L}_k to be *non-Lipschitz* in the x variable in Proposition 5.3. Then, as in Section 4, the series representation for the resolvent G_α of the process (X, A) and the aforementioned result in [17] lead to the desired strong Feller property for the process (X, A) in Theorem 5.4.

The rest of the paper is arranged as follows. We present the necessary assumptions as well as some preliminary results in Section 1.1. In addition, Section 1.1 presents some martingales associated with the operator \mathcal{A} (Theorem 1.4). These martingales are interesting in their own rights. Moreover, they are useful in the proofs of Section 3. The well-posedness of the martingale problem for \mathcal{A} is divided into two parts: Section 2 treats the special case when \mathcal{A} is given by $\hat{\mathcal{A}}$ of (2.5) and Section 3 deals with the general case. Section 4 is devoted to proving the Feller property for the process (X, A) . Strong Feller property is established in Section 5.

To facilitate later presentations, let us introduce some notations that will be frequently used throughout the paper. Let $D([0, \infty), \mathbb{R}^d)$ (resp., $D([0, \infty), \mathbb{S})$) be the space of right continuous functions on $[0, \infty)$ into \mathbb{R}^d (resp., \mathbb{S}) having left limits endowed with the Skorohod topology, and let \mathcal{G}_t (resp., \mathcal{N}_t) be the σ -field generated by the cylindrical sets on $D([0, \infty), \mathbb{R}^d)$ (resp., $D([0, \infty), \mathbb{S})$) up to time t . Also denote $\mathcal{G} = \bigvee_{t=0}^\infty \mathcal{G}_t$ and $\mathcal{N} = \bigvee_{t=0}^\infty \mathcal{N}_t$. It is easy to see that $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{N}_t$ for any $t \geq 0$ and that $\mathcal{F} = \mathcal{G} \vee \mathcal{N}$. Let $C^2(\mathbb{R}^d \times \mathbb{S})$ be the family of functions defined on $\mathbb{R}^d \times \mathbb{S}$ such that $f(\cdot, k) \in C^2(\mathbb{R}^d)$ for each $k \in \mathbb{S}$ and let $C_b^2(\mathbb{R}^d \times \mathbb{S})$ be the family of bounded functions defined on $\mathbb{R}^d \times \mathbb{S}$ such that $f(\cdot, k) \in C^2(\mathbb{R}^d)$ with bounded first and second order continuous partial derivatives in x for each $k \in \mathbb{S}$. Moreover, we denote by $\mathcal{B}(\mathbb{S})$ the family of all the measurable functions on \mathbb{S} into \mathbb{R} .

1.1. Assumptions and preliminaries

Similar to Definition 1.1, for a given $k \in \mathbb{S}$, we can also define the martingale solution for the Lévy type operator \mathcal{L}_k of (1.2) as follows. For a given $x \in \mathbb{R}^d$, we say a probability measure $\mathbb{P}_k^{(x)}$ on $D([0, \infty), \mathbb{R}^d)$ is a solution to the martingale problem for the operator \mathcal{L}_k starting from x , if $\mathbb{P}_k^{(x)}(X(0) = x) = 1$ and for each function $f \in C_c^\infty(\mathbb{R}^d)$,

$$M_t^{(k)(f)} := f(X(t)) - f(X(0)) - \int_0^t \mathcal{L}_k f(X(s)) ds \quad (1.6)$$

is a $\{\mathcal{G}_t\}$ -martingale with respect to $\mathbb{P}_k^{(x)}$.

For the existence and uniqueness of martingale solution corresponding to the weakly coupled Lévy type operator \mathcal{A} defined in (1.1), we make the following assumption.

Assumption 1.2. Suppose the following conditions hold:

- (i) For each $k \in \mathbb{S}$ and $x \in \mathbb{R}^d$, the Lévy type operator \mathcal{L}_k defined in (1.2) has a unique martingale solution $\mathbb{P}_k^{(x)}$ starting from x ;
- (ii) For each $k \in \mathbb{S}$, the function $q_{kk}(x) \leq 0$ is bounded from below; and
- (iii)

$$\sup_{(x,k) \in \mathbb{R}^d \times \mathbb{S}} \int_{\mathbb{R}_0^d} (1 \wedge |y|^2) v(x, k, dy) < \infty. \quad (1.7)$$

Remark 1.3. The martingale problem for the operator \mathcal{L}_k of (1.2) has been well-studied in the literature. For example, Komatsu [12] and Stroock [28] contain explicit sufficient conditions for the existence and uniqueness of martingale solutions for \mathcal{L}_k .

We will prove in Section 3 that there exists a unique martingale solution for the operator \mathcal{A} defined in (1.1). Throughout the rest of this paper, as standing hypotheses, we assume that Assumption 1.2 holds.

We finish the section with the following theorem, which will be needed in the proof of Theorem 3.6, but also interesting in its own right. Let us introduce a counting measure as follows. For $t \geq 0$ and $\Gamma \in \mathcal{B}(\mathbb{R}_0^d)$ with $0 \notin \bar{\Gamma}$, we let

$$\eta(t, \Gamma) := \sum_{s \leq t} \mathbf{1}_{\Gamma}(\Delta X(s)) = \sum_{s \leq t} \mathbf{1}_{\Gamma}(X(s) - X(s-)); \quad (1.8)$$

it counts the number of jumps for the X component such that $\Delta X(s) \in \Gamma$, $0 \leq s \leq t$.

Theorem 1.4. Suppose \mathbb{P} is a solution to the martingale problem associated with \mathcal{A} starting from $(x, k) \in \mathbb{R}^d \times \mathbb{S}$, then the following assertions are true:

- (a) For each $f \in C_b^2(\mathbb{R}^d \times \mathbb{S})$ such that f is uniformly positive,

$$f(X(t), \Lambda(t)) \exp \left\{ - \int_0^t \frac{\mathcal{A}f(X(u), \Lambda(u))}{f(X(u), \Lambda(u))} du \right\}$$

is a \mathbb{P} -martingale.

- (b) For each $\theta \in \mathbb{R}^d$,

$$\begin{aligned} & \exp \left\{ \mathbf{i} \left\langle \theta, X(t) - X(0) - \int_0^t b(X(u), \Lambda(u)) du \right\rangle + \frac{1}{2} \int_0^t \langle \theta, a(X(u), \Lambda(u)) \theta \rangle du \right. \\ & \quad \left. - \int_0^t \int_{\mathbb{R}_0^d} [e^{\mathbf{i} \langle \theta, y \rangle} - 1 - \mathbf{i} \langle \theta, y \rangle \mathbf{1}_{B(0, \varepsilon_0)}(y)] v(X(u), \Lambda(u), dy) du \right\} \end{aligned}$$

is a \mathbb{P} -martingale, where $\mathbf{i} := \sqrt{-1}$.

- (c) Let g be a bounded measurable function on \mathbb{R}^d which vanishes in a neighborhood of the origin. Then for any $\theta \in \mathbb{R}^d$,

$$\begin{aligned} & \exp \left\{ \mathbf{i} \left\langle \theta, X(t) - X(0) - \int_0^t b(X(u), \Lambda(u)) du \right\rangle + \frac{1}{2} \int_0^t \langle \theta, a(X(u), \Lambda(u)) \theta \rangle du \right. \\ & \quad + \int_{\mathbb{R}_0^d} g(y) \eta(t, dy) - \int_0^t \int_{\mathbb{R}_0^d} [e^{\mathbf{i} \langle \theta, y \rangle + g(y)} - 1 - \mathbf{i} \langle \theta, y \rangle \mathbf{1}_{B(0, \varepsilon_0)}(y)] \\ & \quad \times v(X(u), \Lambda(u), dy) du \left. \right\} \end{aligned}$$

is a \mathbb{P} -martingale.

- (d) Define $\tilde{\eta}(t, \Gamma) := \eta(t, \Gamma) - \int_0^t v(X(u), \Lambda(u), \Gamma) du$. Then for each $\theta \in \mathbb{R}^d$ and any measurable function g on \mathbb{R}_0^d satisfying the condition $|g(y)|^2 \leq C(1 \wedge |y|^2)$, for some positive constant C ,

$$\begin{aligned} & \exp \left\{ i \left\langle \theta, X(t) - X(0) - \int_0^t b(X(u), \Lambda(u)) du \right\rangle \right. \\ & \quad + \frac{1}{2} \int_0^t \langle \theta, a(X(u), \Lambda(u)) \theta \rangle du + \int_{\mathbb{R}_0^d} g(y) \tilde{\eta}(t, dy) \\ & \quad \left. - \int_0^t \int_{\mathbb{R}_0^d} [e^{i\langle \theta, y \rangle + g(y)} - 1 - i\langle \theta, y \rangle \mathbf{1}_{B(0, \varepsilon_0)}(y) - g(y)] v(X(u), \Lambda(u), dy) du \right\} \end{aligned}$$

is a \mathbb{P} -martingale. In particular, if $0 \notin \bar{\Gamma}$, then $\tilde{\eta}(t, \Gamma)$ is a \mathbb{P} -martingale.

Proof. This theorem can be established using very similar arguments as those in the proof of Theorem 4.2.1 in [31]. For brevity, we shall omit the details here. \square

2. Martingale solution: special case

We first consider a special Q -matrix $\hat{Q} = (\hat{q}_{kl})$, in which $\hat{q}_{kl} = 1$ for all $k, l \in \mathbb{S}$ with $k \neq l$ and $\hat{q}_{kk} = -(n_0 - 1)$ for all $k \in \mathbb{S}$. In other words, we have

$$\hat{Q} = (\hat{q}_{kl}) = \begin{pmatrix} -(n_0 - 1) & 1 & \cdots & 1 \\ 1 & -(n_0 - 1) & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & -(n_0 - 1) \end{pmatrix}. \quad (2.1)$$

Corresponding to this matrix \hat{Q} , we introduce an operator \hat{Q} on $\mathcal{B}(\mathbb{S})$ as follows:

$$\hat{Q}f(k) = \sum_{l \in \mathbb{S}} \hat{q}_{kl} (f(l) - f(k)), \quad k \in \mathbb{S}. \quad (2.2)$$

For a given $k \in \mathbb{S}$, a probability measure $\mathbb{Q}^{(k)}$ on $D([0, \infty), \mathbb{S})$ is said to be a solution to the martingale problem for the operator \hat{Q} starting from k , if $\mathbb{Q}^{(k)}(\Lambda(0) = k) = 1$ and for each function $f \in \mathcal{B}(\mathbb{S})$,

$$N_t^{(f)} := f(\Lambda(t)) - f(\Lambda(0)) - \int_0^t \hat{Q}f(\Lambda(s)) ds \quad (2.3)$$

is an $\{\mathcal{N}_t\}$ -martingale with respect to $\mathbb{Q}^{(k)}$. Here Λ is the coordinate process $\Lambda(t, \omega) := \omega(t)$ with $\omega \in D([0, \infty), \mathbb{S})$ and $t \geq 0$.

We have the following lemma from [42]:

Lemma 2.1. *For any given $k \in \mathbb{S}$, there exists a unique martingale solution $\mathbb{Q}^{(k)}$ on $D([0, \infty), \mathbb{S})$ for the operator \hat{Q} starting from k .*

Let Λ be the coordinate process on $D([0, \infty), \mathbb{S})$ and let $\{\tau_n\}$ be the sequence of stopping times defined by

$$\tau_0 \equiv 0, \quad \text{and for } n \geq 1, \quad \tau_n := \inf\{t > \tau_{n-1} : \Lambda(t) \neq \Lambda(\tau_{n-1})\}. \quad (2.4)$$

Then it is obvious that for any $k \in \mathbb{S}$, $\mathbb{Q}^{(k)}\{\lim_{n \rightarrow \infty} \tau_n = +\infty\} = 1$. Moreover, we have $\mathbb{Q}^{(k)}(\tau_1 \geq t) = \exp(-(n_0 - 1)t)$ for all $t \geq 0$ and

$$\mathbb{Q}^{(k)}(\Lambda(\tau_1) = l) = 1/(n_0 - 1) \text{ for each } l \in \mathbb{S} \setminus \{k\}.$$

Clearly, the distributions of τ_1 and $\Lambda(\tau_1)$ under $\mathbb{Q}^{(k)}$ are regular.

Now we introduce an operator $\hat{\mathcal{A}}$ on $C_c^2(\mathbb{R}^d \times \mathbb{S})$ as follows:

$$\hat{\mathcal{A}}f(x, k) := \mathcal{L}_k f(x, k) + \hat{\mathcal{Q}}f(x, k), \quad (2.5)$$

where the operators \mathcal{L}_k and $\hat{\mathcal{Q}}$ are defined in (1.2) and (2.2), respectively. Note that $\hat{\mathcal{A}}$ of (2.5) is really a special case of the operator \mathcal{A} defined in (1.1). We can define the martingale solution for the operator $\hat{\mathcal{A}}$ similarly as in Definition 1.1. For convenience of later presentation, let us also denote

$$\hat{M}_t^{(f)} := f(X(t), \Lambda(t)) - f(X(0), \Lambda(0)) - \int_0^t \hat{\mathcal{A}}f(X(s), \Lambda(s))ds, \quad (2.6)$$

where $f \in C_c^\infty(\mathbb{R}^d \times \mathbb{S})$ and (X, Λ) is the coordinate process on $D([0, \infty), \mathbb{R}^d \times \mathbb{S})$.

We will show that for each $(x, k) \in \mathbb{R}^d \times \mathbb{S}$, there exists a unique martingale solution $\hat{\mathbb{P}}^{(x, k)}$ for the operator $\hat{\mathcal{A}}$ starting from (x, k) . Our construction of the desired probability measure $\hat{\mathbb{P}}^{(x, k)}$ on $D([0, \infty), \mathbb{R}^d \times \mathbb{S})$ as well as the proof of uniqueness for such a solution relies heavily on the martingale solutions $\{\mathbb{P}_k^{(x)} : k \in \mathbb{S}, x \in \mathbb{R}^d\}$ and $\{\mathbb{Q}^{(k)} : k \in \mathbb{S}\}$, and the stopping times $\{\tau_n\}$ defined in (2.4).

But first let us introduce a random point process and a family of counting measures on \mathbb{S} as follows. For $t > 0$, $k \in \mathbb{S}$, and $A \subset \mathbb{S}$, set

$$n(t, A) := \sum_{s \leq t} \mathbf{1}_{\{\Lambda(s) \in A, \Lambda(s) \neq \Lambda(s-)\}}, \quad (2.7)$$

and

$$v(k; A) := \sum_{l \in A \setminus \{k\}} \hat{q}_{kl} = \#\{A \setminus \{k\}\}.$$

In view of Lemma 2.4 of [27], we know that $\int_0^t v(\Lambda(s); A)ds$ is the compensator of the point process $n(t, A)$; namely,

$$\mu(t, A) := n(t, A) - \int_0^t v(\Lambda(s); A)ds \quad (2.8)$$

is a martingale measure with respect to $\mathbb{Q}^{(k)}$. Moreover, notice that the operator $\hat{\mathcal{Q}}$ defined in (2.2) can be represented as

$$\hat{\mathcal{Q}}f(k) = \sum_{l \in \mathbb{S}} \hat{q}_{kl}(f(l) - f(k)) = \int_{\mathbb{S}} (f(l) - f(k))v(k; dl). \quad (2.9)$$

Now we present the main result of this section:

Theorem 2.2. *For any given $(x, k) \in \mathbb{R}^d \times \mathbb{S}$, there exists a unique martingale solution $\hat{\mathbb{P}}^{(x, k)}$ on $D([0, \infty), \mathbb{R}^d \times \mathbb{S})$ for the operator $\hat{\mathcal{A}}$ starting from (x, k) .*

Proof. The proof is divided into two steps. The first step establishes the existence of a martingale solution $\hat{\mathbb{P}}$ for the operator $\hat{\mathcal{A}}$ starting from (x, k) while the second step deals with the uniqueness.

Step 1. For any given $(x, k) \in \mathbb{R}^d \times \mathbb{S}$, we define a series of probability measures on (Ω, \mathcal{F}) as follows:

$$\mathbb{P}^{(1)} = \mathbb{P}_k^{(x)} \times \mathbb{Q}^{(k)}, \quad \text{and for } n \geq 1, \quad \mathbb{P}^{(n+1)} = \mathbb{P}^{(n)} \otimes_{\tau_n} (\mathbb{P}_{\Lambda(\tau_n)}^{(X(\tau_n))} \times \mathbb{Q}^{(\Lambda(\tau_n))}), \quad (2.10)$$

where $\Omega = D([0, \infty), \mathbb{R}^d \times \mathbb{S})$. Thanks to Theorem 6.1.2 of [31], $\mathbb{P}^{(n+1)} = \mathbb{P}^{(n)}$ on \mathcal{F}_{τ_n} .

Let $f \in C_c^2(\mathbb{R}^d \times \mathbb{S})$. We have

$$f(X(\tau_1 \wedge t), k) - f(X(0), k) - \int_0^{\tau_1 \wedge t} \mathcal{L}_k f(X(s), k) ds$$

is a martingale with respect to $\mathbb{P}_k^{(x)}$ and hence $\mathbb{P}^{(1)}$. On the other hand, using (2.9), we can write

$$\begin{aligned} & \int_0^{\tau_1 \wedge t} \widehat{Q} f(X(s), \Lambda(s)) ds \\ &= \int_0^{\tau_1 \wedge t} \int_{\mathbb{S}} [f(X(s), l) - f(X(s), \Lambda(s))] v(\Lambda(s), dl) ds \\ &= - \int_0^{\tau_1 \wedge t} \int_{\mathbb{S}} [f(X(s), l) - f(X(s), \Lambda(s))] (n(ds, dl) - v(\Lambda(s), dl) ds) \\ &\quad + \int_0^{\tau_1 \wedge t} \int_{\mathbb{S}} [f(X(s), l) - f(X(s), \Lambda(s))] n(ds, dl) \\ &= - \int_0^{\tau_1 \wedge t} \int_{\mathbb{S}} [f(X(s), l) - f(X(s), \Lambda(s))] \mu(ds, dl) \\ &\quad + f(X(\tau_1 \wedge t), \Lambda(\tau_1 \wedge t)) - f(X(\tau_1 \wedge t), \Lambda(\tau_1 \wedge t-)). \end{aligned}$$

Then using the definitions of the operators \widehat{A} , \mathcal{L}_k and \widehat{Q} , we have

$$\begin{aligned} \widehat{M}_{\tau_1 \wedge t}^{(f)} &= f(X(\tau_1 \wedge t), \Lambda(\tau_1 \wedge t)) - f(X(0), \Lambda(0)) - \int_0^{\tau_1 \wedge t} \widehat{A} f(X(s), \Lambda(s)) ds \\ &= f(X(\tau_1 \wedge t), \Lambda(0)) - f(X(0), \Lambda(0)) - \int_0^{\tau_1 \wedge t} \mathcal{L}_{\Lambda(0)} f(X(s), \Lambda(0)) ds \\ &\quad + f(X(\tau_1 \wedge t), \Lambda(\tau_1 \wedge t)) - f(X(\tau_1 \wedge t), \Lambda(0)) \\ &\quad + \int_0^{\tau_1 \wedge t} \mathcal{L}_{\Lambda(0)} f(X(s), \Lambda(0)) ds - \int_0^{\tau_1 \wedge t} \widehat{A} f(X(s), \Lambda(s)) ds \\ &= f(X(\tau_1 \wedge t), \Lambda(0)) - f(X(0), \Lambda(0)) - \int_0^{\tau_1 \wedge t} \mathcal{L}_{\Lambda(0)} f(X(s), \Lambda(0)) ds \\ &\quad + f(X(\tau_1 \wedge t), \Lambda(\tau_1 \wedge t)) - f(X(\tau_1 \wedge t), \Lambda(0)) - \int_0^{\tau_1 \wedge t} \widehat{Q} f(X(s), \Lambda(s)) ds \\ &= f(X(\tau_1 \wedge t), \Lambda(0)) - f(X(0), \Lambda(0)) - \int_0^{\tau_1 \wedge t} \mathcal{L}_{\Lambda(0)} f(X(s), \Lambda(0)) ds \\ &\quad + \int_0^{\tau_1 \wedge t} \int_{\mathbb{S}} [f(X(s), l) - f(X(s), \Lambda(s))] \mu(ds, dl). \end{aligned}$$

Recall that μ is a martingale measure with respect to $\mathbb{Q}^{(k)}$ and hence $\mathbb{P}^{(1)}$. Thus it follows that $\widehat{M}_{\tau_1 \wedge \cdot}^{(f)}$ is a martingale with respect to $\mathbb{P}^{(1)}$.

Next,

$$f(X(\tau_2 \wedge t), \Lambda(\tau_1)) - f(X(\tau_1), \Lambda(\tau_1)) - \int_{\tau_1}^{\tau_2 \wedge t} L_{\Lambda(\tau_1)} f(X(s), \Lambda(\tau_1)) ds, \quad t \geq \tau_1$$

is a martingale with respect to $\mathbb{P}_{\Lambda(\tau_1)}^{(X(\tau_1))} \times \mathbb{Q}^{(\Lambda(\tau_1))}$. Then a similar argument as above gives that

$$f(X(\tau_2 \wedge t), \Lambda(\tau_2 \wedge t)) - f(X(\tau_1), \Lambda(\tau_1)) - \int_{\tau_1}^{\tau_2 \wedge t} \hat{A}f(X(s), \Lambda(s))ds, \quad t \geq \tau_1$$

is a martingale with respect to $\mathbb{P}_{\Lambda(\tau_1)}^{(X(\tau_1))} \times \mathbb{Q}^{(\Lambda(\tau_1))}$. Notice that the above displayed equation is equal to $\hat{M}_{\tau_2 \wedge t}^{(f)} - \hat{M}_{\tau_1 \wedge t}^{(f)}$. Then in view of Theorem 6.1.2 of [31], $\hat{M}_{\tau_2 \wedge \cdot}^{(f)}$ is a martingale with respect to $\mathbb{P}^{(2)}$. In a similar fashion, we can show that $\hat{M}_{\tau_n \wedge \cdot}^{(f)}$ is a martingale with respect to $\mathbb{P}^{(n)}$ for any $n \geq 1$.

Next we show that $\lim_{n \rightarrow \infty} \mathbb{P}^{(n)}\{\tau_n \leq t\} = 0$ for any $t \geq 0$. To this end, we consider functions of the form $f(x, k) = g(k)$, where $g \in \mathcal{B}(\mathbb{S})$. Then $M_{\tau_n \wedge \cdot}^{(f)}$ is a $\mathbb{P}^{(n)}$ martingale. But for any $t \geq 0$,

$$\hat{M}_t^{(f)} = N_t^{(g)} = g(\Lambda(t)) - g(\Lambda(0)) - \int_0^t \hat{Q}g(\Lambda(s))ds$$

is a martingale with respect to $\mathbb{Q}^{(k)}$. In particular, $N_{\tau_n \wedge \cdot}^{(g)}$ is a martingale with respect to $\mathbb{Q}^{(k)}$ as well. On the other hand, for any $A \in \mathcal{N}$, we define $\mathbb{Q}(A) := \mathbb{P}^{(n)}\{D([0, \infty), \mathbb{R}^d) \times A\}$. Then $N_{\tau_n \wedge \cdot}^{(g)}$ is a martingale with respect to \mathbb{Q} . By the uniqueness result for the martingale problem for \hat{Q} in Lemma 2.1, we have $\hat{Q} = \mathbb{Q}^{(k)}$. Therefore it follows that

$$\mathbb{P}^{(n)}\{\tau_n \leq t\} = \hat{Q}\{\tau_n \leq t\} = \mathbb{Q}^{(k)}\{\tau_n \leq t\} \rightarrow 0, \text{ as } n \rightarrow \infty.$$

Recall that the probabilities $\mathbb{P}^{(n)}$ constructed in (2.10) satisfy $\mathbb{P}^{(n+1)} = \mathbb{P}^{(n)}$ on \mathcal{F}_{τ_n} . Hence by Tulcea's extension theorem (see, e.g., [31, Theorem 1.3.5]), there exists a unique $\hat{\mathbb{P}}$ on (Ω, \mathcal{F}) such that $\hat{\mathbb{P}}$ equals $\mathbb{P}^{(n)}$ on \mathcal{F}_{τ_n} . Thus it follows that $\hat{M}_{\tau_n \wedge \cdot}^{(f)}$ is a martingale with respect to $\hat{\mathbb{P}}$ for every $n \geq 1$. In addition, for any $t \geq 0$, we have

$$\hat{\mathbb{P}}\{\tau_n \leq t\} = \mathbb{P}^{(n)}\{\tau_n \leq t\} = 0. \quad (2.11)$$

Thus $\tau_n \rightarrow \infty$ a.s. $\hat{\mathbb{P}}$ and hence $\hat{M}_{\cdot}^{(f)}$ is a martingale with respect to $\hat{\mathbb{P}}$. This establishes that $\hat{\mathbb{P}}$ is the desired martingale solution starting from (x, k) to the martingale problem for \hat{A} . When we wish to emphasize the initial data dependence $X(0) = x$ and $\Lambda(0) = k$, we write this martingale solution as $\hat{\mathbb{P}}^{(x,k)}$.

Step 2. Next we show that there is at most one solution to the martingale problem associated with \hat{A} starting from (x, k) . To this purpose, we let $\tilde{\mathbb{P}}^{(x,k)} \in \mathcal{P}(\Omega, \mathcal{F})$ be another solution to the martingale problem associated with \hat{A} starting from (x, k) . We show that $\hat{\mathbb{P}}^{(x,k)}$ and $\tilde{\mathbb{P}}^{(x,k)}$ agree on \mathcal{F}_{τ_1} . Recall that $\hat{\mathbb{P}}^{(x,k)}$ agrees with $\mathbb{P}^{(1)} = \mathbb{P}_k^{(x)} \times \mathbb{Q}^{(k)}$ on \mathcal{F}_{τ_1} and that $\mathbb{P}_k^{(x)} \in \mathcal{P}(D([0, \infty); \mathbb{R}^d))$ is the unique solution to the martingale problem associated with \mathcal{L}_k starting from x . Also notice that any $A \in \mathcal{F}_{\tau_1}$ is necessarily of the form $A_1 \times \delta_k$, where $A_1 \subset D([0, \infty), \mathbb{R}^d)$ and δ_k contains all functions ω in $D([0, \infty), \mathbb{S})$ satisfying $\omega(t) = k$ for all $0 \leq t < \tau_1$ and $\omega(\tau_1) \in \mathbb{S} \setminus \{k\}$. Since $\mathbb{Q}^{(k)}(\delta_k) = 1$, it follows that

$$\hat{\mathbb{P}}^{(x,k)}(A) = \mathbb{P}_k^{(x)} \times \mathbb{Q}^{(k)}(A_1 \times \delta_k) = \mathbb{P}_k^{(x)}(A_1). \quad (2.12)$$

On the other hand, since $\tilde{\mathbb{P}}^{(x,k)}$ is a solution to the martingale problem associated with \hat{A} starting from (x, k) , for any $g \in C_c^2(\mathbb{R}^d)$, $\hat{M}_t^{(g)}$ is a $\tilde{\mathbb{P}}^{(x,k)}$ martingale. In particular,

$$\begin{aligned} \hat{M}_{\tau_1 \wedge t}^{(g)} &= g(X(t \wedge \tau_1)) - g(X(0)) - \int_0^{\tau_1 \wedge t} \hat{A}g(X(s))ds \\ &= g(X(t \wedge \tau_1)) - g(X(0)) - \int_0^{\tau_1 \wedge t} \mathcal{L}_k g(X(s))ds \end{aligned}$$

is a $\tilde{\mathbb{P}}^{(x,k)}$ martingale. Now for any $A_1 \subset D([0, \infty), \mathbb{R}^d)$ with $A_1 \in \mathcal{G}$, we define

$$\tilde{\mathbb{P}}(A_1) := \tilde{\mathbb{P}}^{(x,k)}(A_1 \times \delta_k). \quad (2.13)$$

Then $\hat{M}_{\tau_1 \wedge \cdot}^{(g)}$ is also a $\tilde{\mathbb{P}}$ martingale and hence $\tilde{\mathbb{P}}$ is a solution to the martingale problem associated with \mathcal{L}_k starting from x up to τ_1 . Now by the uniqueness of the martingale solution to \mathcal{L}_k starting from x , we conclude from (2.12) and (2.13) that $\hat{\mathbb{P}}^{(x,k)}(A) = \tilde{\mathbb{P}}^{(x,k)}(A)$ for any $A \in \mathcal{F}_{\tau_1}$. This shows that the martingale solution to \hat{A} starting from (x, k) is uniquely determined on \mathcal{F}_{τ_1} .

Now suppose that the martingale solution $\hat{\mathbb{P}}^{(x,k)}$ to \hat{A} starting from (x, k) is uniquely determined on \mathcal{F}_{τ_n} . By virtue of Theorem 6.2.1 of [31] (also Lemma 5.4.19 of [11]), there is a $\hat{\mathbb{P}}^{(x,k)}$ -null set $N \in \mathcal{F}_{\tau_n}$ such that

$$\hat{\mathbb{P}}(X(\tau_n(\omega)), \Lambda(\tau_n(\omega))) := \delta_{(X(\tau_n(\omega)), \Lambda(\tau_n(\omega)))} \otimes_{\tau_n(\omega)} \hat{\mathbb{P}}_\omega$$

solves the martingale problem for \hat{A} starting from $(X(\tau_n(\omega)), \Lambda(\tau_n(\omega)))$ whenever $\omega \notin N$, where $\hat{\mathbb{P}}_\omega$ is the regular conditional probability distribution of $\hat{\mathbb{P}}^{(x,k)}$ given \mathcal{F}_{τ_n} , whose existence follows from [11, Theorem 5.3.18]. By the argument in the previous paragraph, $\hat{\mathbb{P}}^{(X(\tau_n(\omega)), \Lambda(\tau_n(\omega)))}$ is uniquely determined on $\mathcal{F}_{\tau_{n+1}}$. Note that by virtue of Theorem 6.1.2 of [31],

$$\hat{\mathbb{P}}^{(x,k)} = \hat{\mathbb{P}}^{(x,k)} \otimes_{\tau_n(\cdot)} \hat{\mathbb{P}}^{(X(\tau_n(\cdot)), \Lambda(\tau_n(\cdot)))},$$

In other words, the right-hand side of the above displayed equation satisfies

- (i) $\hat{\mathbb{P}}^{(x,k)} \otimes_{\tau_n(\cdot)} \hat{\mathbb{P}}^{(X(\tau_n(\cdot)), \Lambda(\tau_n(\cdot)))}(A) = \hat{\mathbb{P}}^{(x,k)}(A)$, for any $A \in \mathcal{F}_{\tau_n}$, and
- (ii) $\delta_{(X(\tau_n(\omega)), \Lambda(\tau_n(\omega)))} \otimes_{\tau_n(\omega)} \hat{\mathbb{P}}_\omega$ is a regular conditional probability distribution of $\hat{\mathbb{P}}^{(x,k)} \otimes_{\tau_n(\cdot)}$ given \mathcal{F}_{τ_n} .

Thus by the induction hypothesis, we conclude that $\hat{\mathbb{P}}^{(x,k)}$ is uniquely determined on $\mathcal{F}_{\tau_{n+1}}$.

Now we define for any $n \in \mathbb{N}$ and $A \in \mathcal{F}_{\tau_n}$ that $\mathbb{P}_n(A) := \hat{\mathbb{P}}^{(x,k)}(A)$. Apparently \mathbb{P}_n satisfies that $\mathbb{P}_n = \mathbb{P}_{n+1}$ on \mathcal{F}_{τ_n} and that for any $t \geq 0$, $\mathbb{P}_n\{\tau_n \leq t\} = \hat{\mathbb{P}}^{(x,k)}\{\tau_n \leq t\} \rightarrow 0$ as $n \rightarrow \infty$, where we used (2.11). Therefore by Tulcea's extension theorem (e.g., [31, Theorem 1.3.5]), the sequence \mathbb{P}_n has a unique extension $\hat{\mathbb{P}}$ on (Ω, \mathcal{F}) such that $\hat{\mathbb{P}} = \mathbb{P}_n$ on \mathcal{F}_{τ_n} . The measure $\hat{\mathbb{P}}$ solves the martingale problem for the operator \hat{A} starting from (x, k) . This completes the proof. \square

3. Martingale solution: general case

In this section we construct the martingale solution for the general case. To proceed, for any given $t \geq 0$, we define a function M_t on the sample path space as follows:

$$\begin{aligned} M_t(X(\cdot), \Lambda(\cdot)) &:= \prod_{i=0}^{n(t)-1} q_{\Lambda(\tau_i)\Lambda(\tau_{i+1})}(X(\tau_{i+1})) \\ &\times \exp\left(-\sum_{i=0}^{n(t)} \int_{\tau_i}^{\tau_{i+1} \wedge t} [q_{\Lambda(\tau_i)}(X(s)) - n_0 + 1] ds\right), \end{aligned} \quad (3.1)$$

where

$$q_k(x) = \sum_{l \in \mathbb{S} \setminus \{k\}} q_{kl}(x), \quad n(t) = \max\{i \in \mathbb{N} : \tau_i \leq t\},$$

and $\{\tau_i\}$ is the sequence of stopping times defined in (2.4). In case $n(t) = 0$, we use the convention that $\prod_{i=0}^{-1} a_i := 1$ in (3.1).

Lemma 3.1. We have that $(M_t, \mathcal{F}_t, \widehat{\mathbb{P}})$ is a non-negative martingale with mean one.

Proof. Step 1. We first observe that if $q_{kl}(x) > 0$ for all $k \neq l$ and $x \in \mathbb{R}^d$, then

$$\begin{aligned} \prod_{i=0}^{n(t)-1} q_{\Lambda(\tau_i)\Lambda(\tau_{i+1})}(X(\tau_{i+1})) &= \exp \left\{ \sum_{i=0}^{n(t)-1} \log q_{\Lambda(\tau_i)\Lambda(\tau_{i+1})}(X(\tau_{i+1})) \right\} \\ &= \exp \left\{ \int_{[0,t] \times \mathbb{S}} \log q_{\Lambda(s-)l}(X(s)) n(ds, dl) \right\}, \end{aligned}$$

where $n(t, A)$ is the Poisson random measure defined in (2.7). Then it follows from the definition of M in (3.1) that

$$M_t(X(\cdot), \Lambda(\cdot)) = \exp\{Z(t)\}, \quad (3.2)$$

where

$$Z(t) := \int_{[0,t] \times \mathbb{S}} \log q_{\Lambda(s-)l}(X(s)) n(ds, dl) - \int_0^t [q_{\Lambda(s)}(X(s)) - n_0 + 1] ds,$$

Now we apply Itô's formula for jump processes (see, e.g., [10, Theorem II.5.1]) to the process M_t :

$$\begin{aligned} M_t(X(\cdot), \Lambda(\cdot)) - 1 &= e^{Z(t)} - e^{Z(0)} \\ &= \int_0^t \int_{\mathbb{S}} e^{Z(s-)} [q_{\Lambda(s-)l}(X(s)) - 1] n(ds, dl) - \int_0^t e^{Z(s)} [q_{\Lambda(s)}(X(s)) - n_0 + 1] ds. \end{aligned} \quad (3.3)$$

Recall from Section 2 that for any $s \geq 0$, $\widehat{\mathbb{P}}\{\Lambda(s) = l, \Lambda(s) \neq \Lambda(s-)\} = \frac{1}{n_0 - 1}$. Thus we have

$$\begin{aligned} \mathbb{E}^{\widehat{\mathbb{P}}}[n(t, A)] &= \mathbb{E}^{\widehat{\mathbb{P}}} \left[\sum_{s \leq t} \mathbf{1}_{\{\Lambda(s) \in A, \Lambda(s) \neq \Lambda(s-)\}} \right] = \mathbb{E}^{\widehat{\mathbb{P}}} \left[\sum_{k \in \mathbb{S}} \sum_{s \leq t} \mathbf{1}_{\{\Lambda(s) \in A, \Lambda(s) \neq \Lambda(s-), \Lambda(s-) = k\}} \right] \\ &= (n_0 - 1) \int_0^t \int_A \frac{1}{n_0 - 1} dl ds = \int_0^t \int_A dl ds, \end{aligned}$$

where dl is the counting measure on \mathbb{S} . Then it follows that

$$\begin{aligned} \int_0^t e^{Z(s)} [q_{\Lambda(s)}(X(s)) - n_0 + 1] ds &= \int_0^t e^{Z(s)} \sum_{l \neq \Lambda(s-)} [q_{\Lambda(s-)l}(X(s)) - 1] ds \\ &= \int_0^t \int_{\mathbb{S}} e^{Z(s)} [q_{\Lambda(s-)l}(X(s)) - 1] dl ds. \end{aligned}$$

Putting these observations into (3.3) and using (3.2), we obtain

$$M_t(X(\cdot), \Lambda(\cdot)) - 1 = \int_{[0,t] \times \mathbb{S}} M_{s-}(X(\cdot), \Lambda(\cdot)) [q_{\Lambda(s-)l}(X(s)) - 1] \widetilde{n}(ds, dl), \quad (3.4)$$

where $\widetilde{n}(t, A) = n(t, A) - \mathbb{E}^{\widehat{\mathbb{P}}}[n(t, A)]$ is the compensated Poisson random measure with respect to $\widehat{\mathbb{P}}$ and also a martingale measure on $[0, \infty) \times \mathbb{S}$.

Step 2. In general, there may exist some $i \neq j$ and $x \in \mathbb{R}^d$ so that $q_{ij}(x) = 0$. We define $q_{kl}^\varepsilon(x) := q_{kl}(x) + \varepsilon$ for all $k, l \in \mathbb{S}$ with $k \neq l$ and $x \in \mathbb{R}^d$. Also, we let $q_{kk}^\varepsilon(x) := q_{kk}(x)$

$-(n_0 - 1)\varepsilon$ for all $k \in \mathbb{S}$ and $x \in \mathbb{R}^d$. Then as $\varepsilon \downarrow 0$, we have

$$q_{kl}^\varepsilon(x) \rightarrow q_{kl}(x), \text{ and } q_{kk}^\varepsilon(x) \rightarrow q_{kk}(x)$$

uniformly with respect to $x \in \mathbb{R}^d$ for all $l \neq k \in \mathbb{S}$. Next we define

$$M_t^\varepsilon(X(\cdot), \Lambda(\cdot)) := \exp \left\{ \int_{[0,t] \times \mathbb{S}} \log q_{\Lambda(s-)l}^\varepsilon(X(s)) n(dl, ds) - \int_0^t [q_{\Lambda(s)}^\varepsilon(X(s)) - n_0 + 1] ds \right\}.$$

Thanks to [Assumption 1.2](#) and the bounded convergence theorem, we have $M_t^\varepsilon(X(\cdot), \Lambda(\cdot)) \rightarrow M_t(X(\cdot), \Lambda(\cdot))$ as $\varepsilon \downarrow 0$. Moreover, by (3.4) in Step 1, we have

$$M_t^\varepsilon(X(\cdot), \Lambda(\cdot)) - 1 = \int_0^t \int_{\mathbb{S}} M_{s-}^\varepsilon(X(\cdot), \Lambda(\cdot)) [q_{\Lambda(s-)l}^\varepsilon(X(s)) - 1] \tilde{n}(ds, dl).$$

Now passing to the limit as $\varepsilon \downarrow 0$, we obtain from the bounded convergence theorem that

$$M_t(X(\cdot), \Lambda(\cdot)) - 1 = \int_{[0,t] \times \mathbb{S}} M_{s-}(X(\cdot), \Lambda(\cdot)) [q_{\Lambda(s-)l}(X(s)) - 1] \tilde{n}(ds, dl). \quad (3.5)$$

Step 3. From (3.5), we can see that $M_t(X(\cdot), \Lambda(\cdot))$ is a martingale with mean 1 under $\widehat{\mathbb{P}}$. This completes the proof. \square

Lemma 3.2. For any $T > 0$ and $(x, k) \in \mathbb{R}^d \times \mathbb{S}$, the function $M_T(X(\cdot), \Lambda(\cdot))$ defined in (3.1) is integrable with respect to the measure $\widehat{\mathbb{P}}$.

Proof. The proof is similar to that of [36, Lemma 4.4] and we shall omit the details here. \square

Let $\varepsilon > 0$ and notice that in view of (1.3), $v(x, k, \mathbb{R}^d \setminus B(0, \varepsilon)) < \infty$ for each $(x, k) \in \mathbb{R}^d \times \mathbb{S}$. Then we can define a sequence of stopping times as follows. Let $\zeta_0^{(\varepsilon)} := 0$ and for $n \geq 0$,

$$\zeta_{n+1}^{(\varepsilon)} := \inf\{t \geq \zeta_n^{(\varepsilon)} : |\Delta X(t)| = |X(t) - X(t-)| \geq \varepsilon\}. \quad (3.6)$$

Lemma 3.3. Let $X_\varepsilon(t) := X(t) - \int_{|y| \geq \varepsilon} y \eta(t, dy)$ for $t \geq 0$ and define $\mathcal{F}_{\zeta_1^{(\varepsilon)}-} := \sigma\{X_\varepsilon(t \wedge \zeta_1^{(\varepsilon)}), \Lambda(t \wedge \zeta_1^{(\varepsilon)}) : t \geq 0\}$. Then we have

$$\widehat{\mathbb{P}}\{\tau_1 > t\} = \exp\{-(n_0 - 1)t\}, \quad (3.7)$$

$$\widehat{\mathbb{P}}\left\{\zeta_1^{(\varepsilon)} > t \mid \mathcal{F}_{\zeta_1^{(\varepsilon)}-}\right\} = \exp\left\{-\int_0^t v(X_\varepsilon(u \wedge \zeta_1^{(\varepsilon)}), \Lambda(u \wedge \zeta_1^{(\varepsilon)}), \mathbb{R}^d \setminus B(0, \varepsilon)) du\right\}. \quad (3.8)$$

Proof. Eq. (3.7) follows directly from the construction of $\widehat{\mathbb{P}}$ in [Theorem 2.2](#). Now we prove (3.8). Let $\Gamma := \mathbb{R}^d \setminus B(0, \varepsilon)$, and recall $\eta(t, \Gamma)$ defined in (1.8). Let us also denote

$$\tilde{\eta}(t, \Gamma) := \eta(t, \Gamma) - \int_0^t v(X(u), \Lambda(u), \Gamma) du.$$

Thanks to [Theorem 1.4](#), $\tilde{\eta}(t, \Gamma)$ is a $\widehat{\mathbb{P}}$ -martingale. Consequently, for any $t \geq 0$, we have

$$\begin{aligned} & \mathbb{E}^{\widehat{\mathbb{P}}} \left[\eta(t \wedge \zeta_1^{(\varepsilon)}; \Gamma) \middle| \mathcal{F}_{\zeta_1^{(\varepsilon)}-} \right] \\ &= \mathbb{E}^{\widehat{\mathbb{P}}} \left[\int_0^{t \wedge \zeta_1^{(\varepsilon)}} v(X(u), \Lambda(u), \Gamma) du \middle| \mathcal{F}_{\zeta_1^{(\varepsilon)}-} \right] \\ &= \mathbb{E}^{\widehat{\mathbb{P}}} \left[\int_0^t \mathbf{1}_{\{\zeta_1^{(\varepsilon)} > u\}} v(X(u \wedge \zeta_1^{(\varepsilon)}), \Lambda(u \wedge \zeta_1^{(\varepsilon)}), \Gamma) du \middle| \mathcal{F}_{\zeta_1^{(\varepsilon)}-} \right] \\ &= \int_0^t \widehat{\mathbb{P}} \left\{ \zeta_1^{(\varepsilon)} > u \middle| \mathcal{F}_{\zeta_1^{(\varepsilon)}-} \right\} v(X_\varepsilon(u \wedge \zeta_1^{(\varepsilon)}), \Lambda(u \wedge \zeta_1^{(\varepsilon)}), \Gamma) du. \end{aligned} \quad (3.9)$$

On the other hand, note that

$$\eta(t \wedge \zeta_1^{(\varepsilon)}; \Gamma) = \begin{cases} 1 & \text{if } \zeta_1^{(\varepsilon)} \leq t, \\ 0 & \text{otherwise.} \end{cases}$$

Thus we have

$$\begin{aligned} \widehat{\mathbb{P}} \left\{ \zeta_1^{(\varepsilon)} > t \middle| \mathcal{F}_{\zeta_1^{(\varepsilon)}-} \right\} &= \mathbb{E}^{\widehat{\mathbb{P}}} \left[\mathbf{1}_{\{\zeta_1^{(\varepsilon)} > t\}} \middle| \mathcal{F}_{\zeta_1^{(\varepsilon)}-} \right] \\ &= \mathbb{E}^{\widehat{\mathbb{P}}} \left[(1 - \eta(t \wedge \zeta_1^{(\varepsilon)}; \Gamma)) \middle| \mathcal{F}_{\zeta_1^{(\varepsilon)}-} \right] \\ &= 1 - \mathbb{E}^{\widehat{\mathbb{P}}} \left[\eta(t \wedge \zeta_1^{(\varepsilon)}; \Gamma) \middle| \mathcal{F}_{\zeta_1^{(\varepsilon)}-} \right]. \end{aligned} \quad (3.10)$$

Combining (3.9) and (3.10), we arrive at

$$\begin{aligned} 1 - \widehat{\mathbb{P}} \left\{ \zeta_1^{(\varepsilon)} > t \middle| \mathcal{F}_{\zeta_1^{(\varepsilon)}-} \right\} &= \int_0^t \widehat{\mathbb{P}} \left\{ \zeta_1^{(\varepsilon)} > s \middle| \mathcal{F}_{\zeta_1^{(\varepsilon)}-} \right\} \\ &\quad \times v(X_\varepsilon(s \wedge \zeta_1^{(\varepsilon)}), \Lambda(s \wedge \zeta_1^{(\varepsilon)}), \Gamma) ds. \end{aligned} \quad (3.11)$$

Let us denote $u(t) := \widehat{\mathbb{P}} \left\{ \zeta_1^{(\varepsilon)} > t \middle| \mathcal{F}_{\zeta_1^{(\varepsilon)}-} \right\}$ and $v(t) := v(X_\varepsilon(t \wedge \zeta_1^{(\varepsilon)}), \Lambda(t \wedge \zeta_1^{(\varepsilon)}), \Gamma)$. Then we can rewrite (3.11) as $u(t) + \int_0^t u(s)v(s)ds = 1$, which, in turn, implies that

$$\frac{d}{dt} \left(e^{\int_0^t v(r)dr} \int_0^t u(r)v(r)dr \right) = e^{\int_0^t v(r)dr} v(t) \left[u(t) + \int_0^t u(r)v(r)dr \right] = e^{\int_0^t v(r)dr} v(t).$$

Then it follows that

$$e^{\int_0^t v(r)dr} \int_0^t u(r)v(r)dr = \int_0^t e^{\int_0^s v(r)dr} v(s)ds = e^{\int_0^t v(r)dr} - 1,$$

and hence

$$u(t) = 1 - \int_0^t u(s)v(s)ds = 1 - \left(1 - e^{-\int_0^t v(r)dr} \right) = e^{-\int_0^t v(r)dr}.$$

This establishes (3.8) and hence completes the proof of the lemma. \square

Lemma 3.4. Let $\varepsilon > 0$ and define the stopping times $\zeta_n^{(\varepsilon)}$ as in (3.6) and recall the sequence of stopping times $\{\tau_n\}$ defined in (2.4). Then under $\widehat{\mathbb{P}}$, $\{\zeta_n^{(\varepsilon)}, n \geq 1\}$ and $\{\tau_n : n \geq 1\}$ are mutually disjoint with probability 1.

Proof. It is enough to show that for any $T > 0$, $\{\zeta_n^{(\varepsilon)} : n \geq 1, \zeta_n^{(\varepsilon)} \leq T\}$ and $\{\tau_n : n \geq 1, \tau_n \leq T\}$ are mutually disjoint with probability 1. To this end, we let $M > 1$ and for $m = 0, 1, \dots, M-1$,

we denote

$$\begin{aligned} J^{(\varepsilon)}((m-1)/M, m/M) &:= \max \{n \in \mathbb{N} : \zeta_n^{(\varepsilon)} \leq m/M\} \\ &\quad - \max \{n \in \mathbb{N} : \zeta_n^{(\varepsilon)} \leq (m-1)/M\}, \\ S((m-1)/M, m/M) &:= \max \{n \in \mathbb{N} : \tau_n \leq m/M\} \\ &\quad - \max \{n \in \mathbb{N} : \tau_n \leq (m-1)/M\}, \end{aligned}$$

and

$$A_m := \{J^{(\varepsilon)}((m-1)/M, m/M) \geq 1\}, \quad B_m := \{S((m-1)/M, m/M) \geq 1\}.$$

Thanks to (1.7), it follows that there exists some positive constant K_ε such that

$$\nu(x, k, \mathbb{R}^d \setminus B(0, \varepsilon)) \leq K_\varepsilon < \infty, \text{ for all } (x, k) \in \mathbb{R}^d \times \mathbb{S}. \quad (3.12)$$

Then we have from (3.7), (3.8), (3.12), and Lemma 3.3 that

$$\begin{aligned} &\widehat{\mathbb{P}}\{\{\zeta_n^{(\varepsilon)} : n \geq 1, \zeta_n^{(\varepsilon)} \leq T\} \cap \{\tau_n : n \geq 1, \tau_n \leq T\} \neq \emptyset\} \\ &\leq \widehat{\mathbb{P}}\left\{\text{there are one jump and one switch in the interval } \left(\frac{m-1}{M}, \frac{m}{M}\right] \text{ for some } m\right\} \\ &\leq \sum_{m=0}^{M-1} \widehat{\mathbb{P}}\{A_m \cap B_m\} = \sum_{m=0}^{M-1} \widehat{\mathbb{P}}(A_m) \widehat{\mathbb{P}}(B_m | A_m) \\ &\leq \sum_{m=0}^{M-1} \left(1 - \exp\left\{-\int_{\frac{m-1}{M}}^{\frac{m}{M}} \nu(X_\varepsilon(s), \Lambda(s), \mathbb{R}^d \setminus B(0, \varepsilon)) ds\right\}\right) \left(1 - \exp\left\{-(n_0 - 1)\frac{1}{M}\right\}\right) \\ &\leq \sum_{m=0}^{M-1} \left(1 - \exp\left\{-K_\varepsilon \frac{1}{M}\right\}\right) \left(1 - \exp\left\{-(n_0 - 1)\frac{1}{M}\right\}\right). \end{aligned}$$

Furthermore, using the elementary inequality $1 - e^{-a} \leq a$ for $a \geq 0$, we obtain

$$\begin{aligned} &\widehat{\mathbb{P}}\{\{\zeta_n^{(\varepsilon)} : n \geq 1, \zeta_n^{(\varepsilon)} \leq T\} \cap \{\tau_n : n \geq 1, \tau_n \leq T\} \neq \emptyset\} \\ &\leq \sum_{m=0}^{M-1} \frac{n_0 - 1}{M} \frac{K_\varepsilon}{M} = \frac{(n_0 - 1)K_\varepsilon}{M}, \end{aligned}$$

which can be arbitrarily small since the denominator M is arbitrary. This implies the desired conclusion and hence completes the proof. \square

Note that since $X \in D([0, \infty), \mathbb{R}^d)$, the set of discontinuity points of X is at most countable for almost all $\omega \in \Omega$, see, e.g. [23]. Therefore we can again define the sequence of jump times for X as follows. Let $\zeta_0 := 0$ and for $n \geq 0$, define $\zeta_{n+1} := \inf\{t \geq \zeta_n : |\Delta X(t)| = |X(t) - X(t-)| > 0\}$.

Proposition 3.5. *Under $\widehat{\mathbb{P}}$, $\{\zeta_n : n \geq 1\}$ and $\{\tau_n : n \geq 1\}$ are mutually disjoint with probability 1.*

Proof. We first notice that $\{\zeta_n : n \geq 1\} = \bigcup_{m=1}^{\infty} \{\zeta_n^{(1/m)} : n \geq 1\}$ and hence

$$\{\zeta_n : n \geq 1\} \cap \{\tau_n : n \geq 1\} = \bigcup_{m=1}^{\infty} \{\zeta_n^{(1/m)} : n \geq 1\} \cap \{\tau_n : n \geq 1\}.$$

Moreover, for each $m = 1, 2, \dots$, since (3.12) holds with $\varepsilon = \frac{1}{m}$, Lemma 3.4 implies that

$$\widehat{\mathbb{P}}\{\{\zeta_n^{(1/m)} : n \geq 1\} \cap \{\tau_n : n \geq 1\} \neq \emptyset\} = 0. \quad (3.13)$$

Therefore we deduce

$$\begin{aligned} \widehat{\mathbb{P}}\{\{\zeta_n : n \geq 1\} \cap \{\tau_n : n \geq 1\} \neq \emptyset\} &\leq \sum_{m=1}^{\infty} \widehat{\mathbb{P}}\{\{\zeta_n^{(1/m)} : n \geq 1\} \cap \{\tau_n : n \geq 1\} \neq \emptyset\} \\ &= 0. \end{aligned}$$

This completes the proof. \square

By virtue of M_t and $\widehat{\mathbb{P}}$, we can construct another probability measure \mathbb{P} on $D([0, \infty), \mathbb{R}^d \times \mathbb{S})$ such that \mathbb{P} is a solution to the martingale problem for the operator \mathcal{A} .

Theorem 3.6. *For any given $(x, k) \in \mathbb{R}^d \times \mathbb{S}$, there exists a unique martingale solution $\mathbb{P}^{(x,k)}$ on $D([0, \infty), \mathbb{R}^d \times \mathbb{S})$ for the operator \mathcal{A} starting from (x, k) .*

Proof. For each $t \geq 0$ and each $A \in \mathcal{F}_t$, define

$$\mathbb{P}_t^{(x,k)}(A) = \int_A M_t(X(\cdot), \Lambda(\cdot)) d\widehat{\mathbb{P}}^{(x,k)}. \quad (3.14)$$

Thanks to Lemma 3.1, the family of probability measures $\{\mathbb{P}_t^{(x,k)}\}_{t \geq 0}$ is consistent in the sense that if $0 \leq t_1 \leq t_2$ and $A \in \mathcal{F}_{t_1}$, then $\mathbb{P}_{t_2}^{(x,k)}(A) = \mathbb{P}_{t_1}^{(x,k)}(A)$. Thus by Tulcea's extension theorem (see, e.g., [31, Theorem 1.3.5]), there exists a unique probability measure $\mathbb{P}^{(x,k)}$ on (Ω, \mathcal{F}) which coincides with $\mathbb{P}_t^{(x,k)}$ on \mathcal{F}_t for all $t \geq 0$. Moreover, we will prove that the \mathbb{P} is the desired martingale solution for the operator \mathcal{A} starting from (x, k) . To do so, analogously to the proof of Lemma 4.2 in [36], we first prove that for each function $f \in C_c^\infty(\mathbb{R}^d \times \mathbb{S})$, $(M_t M_t^{(f)}, \mathcal{F}_t, \widehat{\mathbb{P}})$ is a martingale, where $M_t^{(f)}$ is defined in (1.5). In fact, using integration by parts, we derive that

$$\begin{aligned} M_t M_t^{(f)} &= \int_0^t M_{s-}^{(f)} dM_s + \int_0^t M_{s-} d\widehat{M}_s^{(f)} \\ &\quad + \int_0^t M_{s-} (dM_s^{(f)} - d\widehat{M}_s^{(f)}) + \sum_{s \leq t} (M_s - M_{s-})(M_s^{(f)} - M_{s-}^{(f)}), \end{aligned} \quad (3.15)$$

where $\widehat{M}_t^{(f)}$ is defined in (2.6). Using (1.5), (3.1), and Proposition 3.5, we can compute

$$\begin{aligned} &\sum_{s \leq t} (M_s - M_{s-})(M_s^{(f)} - M_{s-}^{(f)}) \\ &= \sum_{s \leq t} (M_s - M_{s-})[f(X(s), \Lambda(s)) - f(X(s), \Lambda(s-))] \\ &= \int_{[0,t] \times \mathbb{S}} M_{s-} \left(\frac{M_s}{M_{s-}} - 1 \right) [f(X(s), l) - f(X(s), \Lambda(s-))] n(ds, dl) \\ &= \int_{[0,t] \times \mathbb{S}} M_{s-} (q_{\Lambda(s-l)}(X(s)) - 1) [f(X(s), l) - f(X(s), \Lambda(s-))] n(ds, dl). \end{aligned}$$

On the other hand,

$$\begin{aligned} & \int_0^t M_{s-} (dM_s^{(f)} - d\widehat{M}_s^{(f)}) \\ &= - \int_0^t M_{s-} \sum_{l \in \mathbb{S}} (q_{\Lambda(s-)l}(X(s)) - 1) [f(X(s), l) - f(X(s), \Lambda(s-))] ds. \end{aligned}$$

Combining the last two displayed equations, and using the observations concerning the martingale measure $\widetilde{n}(\cdot, \cdot)$ in the proof of [Lemma 3.1](#), we obtain

$$\begin{aligned} & \int_0^t M_{s-} (dM_s^{(f)} - d\widehat{M}_s^{(f)}) + \sum_{s \leq t} (M_s - M_{s-}) (M_s^{(f)} - M_{s-}^{(f)}) \\ &= \int_{[0,t] \times \mathbb{S}} M_{s-} (q_{\Lambda(s-)l}(X(s)) - 1) [f(X(s), l) - f(X(s), \Lambda(s-))] \widetilde{n}(ds, dl). \end{aligned}$$

Then upon plugging the above equation into [\(3.15\)](#), it follows that

$$\begin{aligned} M_t M_t^{(f)} &= \int_0^t M_{s-}^{(f)} dM_s + \int_0^t M_{s-} d\widehat{M}_s^{(f)} \\ &+ \int_{[0,t] \times \mathbb{S}} M_{s-} (q_{\Lambda(s-)l}(X(s)) - 1) \\ &\times [f(X(s), l) - f(X(s), \Lambda(s-))] \widetilde{n}(ds, dl). \end{aligned} \quad (3.16)$$

We have shown respectively in [Theorem 2.2](#) and [Lemma 3.1](#) that $\widehat{M}^{(f)}$ and M are martingales under the measure $\widehat{\mathbb{P}}^{(x,k)}$. Also recall from the proof of [Lemma 3.1](#) that $\widetilde{n}(\cdot, \cdot)$ is a martingale measure on $[0, \infty) \times \mathbb{S}$ under $\widehat{\mathbb{P}}^{(x,k)}$. Thus in view of [\(3.16\)](#), we conclude immediately that $M_t M_t^{(f)}$ is a martingale under $\widehat{\mathbb{P}}^{(x,k)}$.

We now prove that for each function $f \in C_c^\infty(\mathbb{R}^d \times \mathbb{S})$, $(M_t^{(f)}, \mathcal{F}_t, \mathbb{P}^{(x,k)})$ is a martingale. Indeed, for any given $0 \leq s < t$ and any given $A \in \mathcal{F}_s$, we have

$$\int_A M_t^{(f)} d\mathbb{P}^{(x,k)} = \int_A M_t M_t^{(f)} d\widehat{\mathbb{P}}^{(x,k)} = \int_A M_s M_s^{(f)} d\widehat{\mathbb{P}}^{(x,k)} = \int_A M_s^{(f)} d\mathbb{P}^{(x,k)},$$

where the second equality follows from the martingale property of $(M_t M_t^{(f)}, \mathcal{F}_t, \widehat{\mathbb{P}}^{(x,k)})$, while the first and third equalities hold true since $\mathbb{P}^{(x,k)}$ coincides with the probability measure $\mathbb{P}_t^{(x,k)}$ given in [\(3.14\)](#). This shows that $\mathbb{P}^{(x,k)}$ is a martingale solution for the operator \mathcal{A} starting from (x, k) .

It remains to show that any martingale solution $\widetilde{\mathbb{P}}$ for the operator \mathcal{A} starting from (x, k) must agree with $\mathbb{P}^{(x,k)}$ and therefore establishing the desired uniqueness. From [\[33\]](#), for any martingale solution $\mathbb{P}^{(x,k)}$ to the operator \mathcal{A} , we have

$$\mathbb{P}^{(x,k)}(\Lambda(\tau_1) \in \mathbb{S} \setminus \{k\} | \mathcal{F}_{\tau_1-}) = - \sum_{l \in \mathbb{S} \setminus \{k\}} \frac{q_{kl}}{q_{kk}} (X(\tau_1-)) = 1.$$

Then the uniqueness can be established by using a similar argument as that in the proof of [Theorem 2.2](#). \square

Remark 3.7. Thanks to [Theorem 3.6](#), the martingale problem for the operator \mathcal{A} defined in [\(1.1\)](#) with any initial condition $(x, k) \in \mathbb{R}^d \times \mathbb{S}$ is well-posed. Thus the process (X, Λ) is strong Markov.

4. Feller property

We proved in [Theorem 3.6](#) that the martingale problem for the operator \mathcal{A} defined in (1.1) is well-posed. Consequently for any (x, k) , there exists a unique probability measure \mathbb{P} on $\Omega = D([0, \infty), \mathbb{R}^d \times \mathbb{S})$ under which the coordinate process $(X(t), \Lambda(t))$ satisfies $\mathbb{P}\{(X(0), \Lambda(0)) = (x, k)\} = 1$ and that for any $f \in C_c^\infty(\mathbb{R}^d \times \mathbb{S})$, the process M_t^f defined in (1.5) is an $\{\mathcal{F}_t\}$ -martingale. In this section, we will prove that in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the process $(X(t), \Lambda(t))$ possesses the Feller property under the following conditions.

Assumption 4.1. Assume that there exist a positive constant H and a nondecreasing and concave function $\rho : [0, \infty) \mapsto [0, \infty)$ satisfying $\rho(r) > 0$ for $r > 0$ and

$$\int_{0+} \frac{dr}{\rho(r)} = \infty, \quad (4.1)$$

such that for all $k \in \mathbb{S}$ and $x, z \in \mathbb{R}^d$,

$$\|\sigma(x, k) - \sigma(z, k)\|^2 + 2\langle x - z, b(x, k) - b(z, k) \rangle \leq H|x - z|\rho(|x - z|), \quad (4.2)$$

and

$$\int_{\mathbb{R}_0^d} |u| \|v(x, k, \cdot) - v(z, k, \cdot)\| (du) \leq H\rho(|x - z|), \quad (4.3)$$

where $\sigma(x, k) \in \mathbb{R}^{d \times d}$ satisfies $\sigma(x, k)\sigma(x, k)^T = a(x, k)$, and $\|\cdot\|$ denotes the Hilbert–Schmidt norm for matrices or the total variation norm for signed measures. Here and below, T denotes the transpose of a vector or matrix.

Assumption 4.2. Assume that

$$|q_{kl}(x) - q_{kl}(z)| \leq H|x - z| \quad (4.4)$$

for all $x, z \in \mathbb{R}^d$ and $k \neq l \in \mathbb{S}$, where constant $H > 0$ is the same as that in [Assumption 4.1](#) without loss of generality.

Remark 4.3. For existence of a square root $\sigma(x, k)$ of $a(x, k)$ such as in [Assumption 4.1](#) and the equivalence of different choices of the square root, we refer the reader to Stroock and Varadhan [31] for details. Some common functions satisfying the conditions in [Assumption 4.1](#) include $\rho(r) = r$ and concave and increasing functions such as $\rho(r) = r \log(1/r)$, $\rho(r) = r \log(\log(1/r))$, and $\rho(r) = r \log(1/r) \log(\log(1/r))$ for $r \in (0, \delta)$ with $\delta > 0$ small enough.

The main result of this section is:

Theorem 4.4. Suppose that [Assumptions 1.2, 4.1 and 4.2](#) hold. Then the process (X, Λ) has Feller property.

Let us first briefly describe our strategy toward the proof of [Theorem 4.4](#). We first use the coupling method to show in [Lemma 4.6](#) that the Lévy type process $\tilde{X}^{(k)}$ corresponding to the operator \mathcal{L}_k of (1.2) has Feller property under [Assumptions 1.2\(i\) and 4.1](#). [Lemma 4.7](#) further establishes the Feller property for the killed Lévy type process $X^{(k)}$ under [Assumption 4.2](#). Next we show in [Lemma 4.9](#) that the resolvent of (X, Λ) can be represented by a series of the resolvents of the killed processes $X^{(k)}$, $k \in \mathbb{S}$. This representation further helps us to establish (4.28). Finally we use (4.28) together with Proposition 6.1.1 in [17] to derive the Feller property for the process (X, Λ) .

Remark 4.5. (i) The recent paper Wang [32] also establishes the Feller property for the Lévy type process $\tilde{X}^{(k)}$ under a different set of conditions. In particular, the Lévy type kernel is assumed to have a certain representation in [32]. By contrast, our goal is to establish the Feller property for the two-component process (X, Λ) under [Assumptions 1.2, 4.1](#) and [4.2](#). This is achieved by establishing the Feller property for the Lévy type process $\tilde{X}^{(k)}$ as well as the killed Lévy type process $X^{(k)}$ under these assumptions. It is worth pointing out that [Lemma 4.6](#) indicates that [Assumptions 1.2\(i\)](#) and [4.1](#) are sufficient conditions for the Feller property for the Lévy type process $\tilde{X}^{(k)}$. These assumptions, in particular, [Assumption 4.1](#), seem more direct and easier to verify in some sense compared with those in [32].

Recall that for each $k \in \mathbb{S}$ and $x \in \mathbb{R}^d$, [Assumption 1.2](#) guarantees that the operator \mathcal{L}_k of [\(1.2\)](#) uniquely determines a Lévy type process $\tilde{X}^{(k)(x)}$ with initial condition $\tilde{X}^{(k)(x)}(0) = x$. Next we kill the process $\tilde{X}^{(k)(x)}$ at rate $(-q_{kk})$:

$$\begin{aligned} \mathbb{E}_k[f(X^{(k)(x)}(t))] &= \mathbb{E}_k[f(\tilde{X}^{(k)(x)}(t)); t < \tau] \\ &= \mathbb{E}_k\left[\exp\left\{\int_0^t q_{kk}(\tilde{X}^{(k)(x)}(s))ds\right\} f(\tilde{X}^{(k)(x)}(t))\right], \end{aligned} \quad (4.5)$$

where $\tau := \inf\{t \geq 0 : \Lambda(t) \neq \Lambda(0)\}$. Equivalently, the killed Lévy type $X^{(k)(x)}$ can be defined as $X^{(k)(x)}(t) = \tilde{X}^{(k)(x)}(t)$ if $t < \tau$ and $X^{(k)(x)}(t) = \partial$ if $t \geq \tau$, where ∂ is a cemetery point added to \mathbb{R}^d . Moreover, we denote the transition probability families of the Lévy type process $\tilde{X}^{(k)}$ and the killed Lévy type process $X^{(k)}$ by $\{\tilde{P}^{(k)}(t, x, A) : t \geq 0, x \in \mathbb{R}^d, A \in \mathcal{B}(\mathbb{R}^d)\}$ and $\{P^{(k)}(t, x, A) : t \geq 0, x \in \mathbb{R}^d, A \in \mathcal{B}(\mathbb{R}^d)\}$, respectively.

For an arbitrarily fixed $k \in \mathbb{S}$, we now construct a coupling of the Lévy type process $\tilde{X}^{(k)}$. To this end, we need only to construct a coupling for its generator \mathcal{L}_k . For $x, z \in \mathbb{R}^d$, set

$$a(x, z, k) = \begin{pmatrix} a(x, k) & \sigma(x, k)\sigma(z, k)^T \\ \sigma(z, k)\sigma(x, k)^T & a(z, k) \end{pmatrix}, \quad b(x, z, k) = \begin{pmatrix} b(x, k) \\ b(z, k) \end{pmatrix}.$$

Obviously, $a(x, z, k)$ is nonnegative definite for all $x, z \in \mathbb{R}^d$. For $h(x, z) \in C_0^2(\mathbb{R}^d \times \mathbb{R}^d)$, set

$$\tilde{\Omega}_d(k)h(x, z) = \frac{1}{2}\text{tr}(a(x, z, k)\nabla^2 h(x, z)) + \langle b(x, z, k), \nabla h(x, z) \rangle, \quad (4.6)$$

which is a coupling of the diffusion part in the generator \mathcal{L}_k defined in [\(1.2\)](#) (refer to Chen and Li [4]). Next, for $h(x, z) \in C_0^2(\mathbb{R}^d \times \mathbb{R}^d)$, set

$$\begin{aligned} &\tilde{\Omega}_j(k)h(x, z) \\ &= \int [h(x+u, z) - h(x, z) - \langle \nabla_x h(x, z), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u)](v(x, k, du) - v(z, k, du))^+ \\ &\quad + \int [h(x, z+u) - h(x, z) - \langle \nabla_z h(x, z), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u)] \\ &\quad \times (v(z, k, du) - v(x, k, du))^+ \\ &\quad + \int [h(x+u, z+u) - h(x, z) - \langle \nabla_x h(x, z), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u) \\ &\quad \quad - \langle \nabla_z h(x, z), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u)](v(x, k, (\cdot)) \wedge v(z, k, (\cdot)))(du), \end{aligned} \quad (4.7)$$

where $(v(x, k, \cdot) - v(z, k, \cdot))^+ = \sup\{v(x, k, A) - v(z, k, A) : A \in \mathcal{B}(\mathbb{R}_0^d)\}$ and $(v(z, k, \cdot) - v(x, k, \cdot))^+$ is defined in a similar fashion. Note that the operator $\tilde{\Omega}_j(k)$ defined in [\(4.7\)](#) is a coupling of the jump part in the generator \mathcal{L}_k defined in [\(1.2\)](#). Finally, combining the two

couplings together, we get a coupling $\tilde{\mathcal{L}}_k$ of the generator \mathcal{L}_k as follows:

$$\tilde{\mathcal{L}}_k h(x, z) = \tilde{\mathcal{Q}}_d(k)h(x, z) + \tilde{\mathcal{Q}}_j(k)h(x, z), \quad (4.8)$$

for $h(x, z) \in C_0^2(\mathbb{R}^d \times \mathbb{R}^d)$.

To proceed, we now introduce the Wasserstein metric between two probability measures as follows. For two probability measures P_1 and P_2 on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, define

$$W(P_1, P_2) = \inf_{\tilde{P}} \int |x - z| \tilde{P}(dx, dz),$$

where \tilde{P} varies over all coupling probability measures with marginals P_1 and P_2 ; that is,

$$\tilde{P}(A \times \mathbb{R}^d) = P_1(A), \text{ and } \tilde{P}(\mathbb{R}^d \times A) = P_2(A), \text{ for any } A \in \mathcal{B}(\mathbb{R}^d).$$

Lemma 4.6. Suppose that Assumptions 1.2(i) and 4.1 hold. For each $k \in \mathbb{S}$, the Lévy type process $\tilde{X}^{(k)}$ generated by the Lévy type operator \mathcal{L}_k defined in (1.2) has Feller property.

Proof. For an arbitrarily fixed $k \in \mathbb{S}$, we need only to prove that for any $t > 0$, $x, z \in \mathbb{R}^d$, $\tilde{P}^{(k)}(t, x, \cdot)$ converges weakly to $\tilde{P}^{(k)}(t, z, \cdot)$ as $x \rightarrow z$. By virtue of Theorem 5.6 in [3], it suffices to prove that

$$W(\tilde{P}^{(k)}(t, x, \cdot), \tilde{P}^{(k)}(t, z, \cdot)) \rightarrow 0 \quad \text{as } x \rightarrow z. \quad (4.9)$$

We use the coupling $\tilde{\mathcal{L}}_k$ constructed in (4.8) to establish (4.9). Let $(\tilde{X}^{(k)}, \tilde{Z}^{(k)})$ denote the coupling process corresponding to the coupling generator $\tilde{\mathcal{L}}_k$. Also let \mathbb{P}_k denote the distribution of $(\tilde{X}^{(k)}, \tilde{Z}^{(k)})$ and \mathbb{E}_k the corresponding expectation with a slight abuse of notation. By Assumption 1.2 we readily know that the coupling process $(\tilde{X}^{(k)}, \tilde{Z}^{(k)})$ is non-explosive. Similarly to the proof of Theorem 2.3 in [4], set

$$T_R := \inf\{t \geq 0 : |\tilde{X}^{(k)}(t)|^2 + |\tilde{Z}^{(k)}(t)|^2 > R\}.$$

Thanks to the assumptions imposed on the function ρ , we can find a strictly decreasing sequence $\{a_n\} \subset (0, 1]$ with $a_0 = 1$, $\lim_{n \rightarrow \infty} a_n = 0$ and $\int_{a_n}^{a_{n-1}} \rho^{-1}(r) dr = n$ for every $n \geq 1$. For each $n \geq 1$, there exists a continuous function ρ_n on \mathbb{R} with support in (a_n, a_{n-1}) so that $0 \leq \rho_n(r) \leq 2n^{-1} \rho^{-1}(r)$ holds for every $r > 0$, and $\int_{a_n}^{a_{n-1}} \rho_n(r) dr = 1$.

Now consider the sequence of functions

$$\psi_n(r) := \int_0^{|r|} \int_0^y \rho_n(u) du dy, \quad r \in \mathbb{R}, n \geq 1. \quad (4.10)$$

We can immediately verify that ψ_n is even and continuously differentiable, with $|\psi'_n(r)| \leq 1$ and $\lim_{n \rightarrow \infty} \psi_n(r) = |r|$ for $r \in \mathbb{R}$. Furthermore, for each $r > 0$, the sequence $\{\psi_n(r)\}_{n \geq 1}$ is nondecreasing. Note also that for each $n \in \mathcal{N}$, ψ_n , ψ'_n and ψ''_n all vanish on the interval $(-a_n, a_n)$.

For any $x, z \in \mathbb{R}^d$, set

$$\begin{aligned} A(x, z, k) &= a(x, k) + a(z, k) - 2\sigma(x, k)\sigma(z, k)^T, \\ \widehat{B}(x, z, k) &= \langle x - z, b(x, k) - b(z, k) \rangle, \end{aligned}$$

and

$$\overline{A}(x, z, k) = \langle x - z, A(x, z, k)(x - z) \rangle / |x - z|^2.$$

Then as in the proof of Theorem 3.1 in [4], we can verify directly that

$$2\tilde{\Omega}_d(k)\psi_n(|x-z|) = \psi_n''(|x-z|)\bar{A}(x, z, k) + \frac{\psi_n'(|x-z|)}{|x-z|} [\text{tr}(A(x, z, k)) - \bar{A}(x, z, k) + 2\widehat{B}(x, z, k)]. \quad (4.11)$$

Note that $\text{tr}(A(x, z, k)) = \|\sigma(x, k) - \sigma(z, k)\|^2$ and hence we obtain from (4.2) that

$$\text{tr}A(x, z, k) + 2\widehat{B}(x, z, k) \leq H|x-z|\rho(|x-z|).$$

On the other hand, using (4.2) again,

$$\begin{aligned} \bar{A}(x, z, k) &= \frac{\langle x-z, (\sigma(x, k) - \sigma(z, k))(\sigma(x, k) - \sigma(z, k))^T(x-z) \rangle}{|x-z|^2} \\ &\leq H|x-z|\rho(|x-z|). \end{aligned}$$

Thanks to the construction of ψ_n , we have $0 \leq \psi_n'(r) \leq 1$ and $\psi_n''(r) = \rho_n(r) \leq \frac{2}{n\rho(r)}I_{(a_n, a_{n-1})}(r)$ for all $r \geq 0$. Putting the above estimates into (4.11), it then follows that

$$\begin{aligned} \tilde{\Omega}_d(k)\psi_n(|x-z|) &\leq \frac{1}{2}\psi_n''(|x-z|)H|x-z|\rho(|x-z|) + \frac{1}{2}\psi_n'(|x-z|)H\rho(|x-z|) \\ &\leq \frac{H}{n}|x-z|I_{(a_n, a_{n-1})}(|x-z|) + \frac{1}{2}H\rho(|x-z|) \\ &\leq \frac{Ha_{n-1}}{n} + \frac{1}{2}H\rho(|x-z|). \end{aligned} \quad (4.12)$$

By virtue of the mean value theorem and the fact that $|\psi_n'| \leq 1$, we have

$$\psi_n(|x+u-z|) - \psi_n(|x-z|) \leq ||x+u-z| - |x-z|| \leq |u|,$$

and

$$|\langle \nabla_x \psi_n(|x-z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u)| \leq |u|.$$

Then it follows that

$$\begin{aligned} &\int (\psi_n(|x+u-z|) - \psi_n(|x-z|) - \langle \nabla_x \psi_n(|x-z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u)) \\ &\quad \times (v(x, k, du) - v(z, k, du))^+ \\ &\leq 2 \int |u| (v(x, k, du) - v(z, k, du))^+. \end{aligned}$$

Similarly, we have

$$\begin{aligned} &\int (\psi_n(|x-z-u|) - \psi_n(|x-z|) - \langle \nabla_z \psi_n(|x-z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u)) \\ &\quad \times (v(z, k, du) - v(x, k, du))^+ \\ &\leq 2 \int |u| (v(x, k, du) - v(z, k, du))^+. \end{aligned}$$

Note that $\nabla_x \psi_n(|x-z|) = -\nabla_z \psi_n(|x-z|)$. Thus

$$\begin{aligned} &\int [\psi_n(|x+u-z-u|) - \psi_n(|x-z|) - \langle \nabla_x \psi_n(|x-z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u) \\ &\quad - \langle \nabla_z \psi_n(|x-z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u)] (v(x, k, du) \wedge v(z, k, du)) = 0. \end{aligned}$$

Then, using the definition of $\tilde{\Omega}_j(k)$ in (4.7), we obtain

$$\begin{aligned} & \tilde{\Omega}_j(k)\psi_n(|x-z|) \\ & \leq 2 \int |u| (v(x, k, du) - v(z, k, du))^+ + 2 \int |u| (v(z, k, du) - v(x, k, du))^+ \\ & \leq 2 \int |u| \|v(x, k, \cdot) - v(z, k, \cdot)\| (du) \leq 2H\rho(|x-z|), \end{aligned} \quad (4.13)$$

where the last inequality follows from (4.3).

A combination of (4.12) and (4.13) yields

$$\begin{aligned} \tilde{\mathcal{L}}_k \psi_n(|x-z|) &= \tilde{\Omega}_d(k)\psi_n(|x-z|) + \tilde{\Omega}_j(k)\psi_n(|x-z|) \\ &\leq \frac{Ha_{n-1}}{n} + 3H\rho(|x-z|), \quad \forall x, z \in \mathbb{R}^d. \end{aligned}$$

Now we apply Itô's formula to the process $\psi_n(|\tilde{X}^{(k)(x)}(\cdot) - \tilde{Z}^{(k)(z)}(\cdot)|)$ to obtain

$$\begin{aligned} & \mathbb{E}_k[\psi_n(|\tilde{X}^{(k)(x)}(t \wedge T_R) - \tilde{Z}^{(k)(z)}(t \wedge T_R)|)] \\ &= \psi_n(|x-z|) + \mathbb{E}_k\left[\int_0^{t \wedge T_R} \tilde{\mathcal{L}}_k \psi_n(|\tilde{X}^{(k)(x)}(s) - \tilde{Z}^{(k)(z)}(s)|) ds\right] \\ &\leq \psi_n(|x-z|) + \frac{Ha_{n-1}t}{n} + 3H\mathbb{E}_k\left[\int_0^{t \wedge T_R} \rho(|\tilde{X}^{(k)(x)}(s) - \tilde{Z}^{(k)(z)}(s)|) ds\right]. \end{aligned} \quad (4.14)$$

Recall that $\psi_n(|x|) \uparrow |x|$ and $a_n \rightarrow 0$ as $n \rightarrow \infty$. Therefore, passing to the limit as $n \rightarrow \infty$ on both sides of (4.14), it follows from the monotone convergence theorem that

$$\begin{aligned} & \mathbb{E}_k[|\tilde{X}^{(k)(x)}(t \wedge T_R) - \tilde{Z}^{(k)(z)}(t \wedge T_R)|] \\ &\leq |x-z| + 3H\mathbb{E}_k\left[\int_0^{t \wedge T_R} \rho(|\tilde{X}^{(k)(x)}(s) - \tilde{Z}^{(k)(z)}(s)|) ds\right]. \end{aligned}$$

Furthermore, passing to the limit as $R \rightarrow \infty$, we have by Fatou's lemma and the monotone convergence theorem that

$$\begin{aligned} \mathbb{E}_k[|\tilde{X}^{(k)(x)}(t) - \tilde{Z}^{(k)(z)}(t)|] &\leq |x-z| + 3H\mathbb{E}_k\left[\int_0^t \rho(|\tilde{X}^{(k)(x)}(s) - \tilde{Z}^{(k)(z)}(s)|) ds\right] \\ &\leq |x-z| + 3H\mathbb{E}_k\left[\int_0^t \rho(|\tilde{X}^{(k)(x)}(s) - \tilde{Z}^{(k)(z)}(s)|) ds\right] \\ &\leq |x-z| + 3H \int_0^t \rho(\mathbb{E}_k[|\tilde{X}^{(k)(x)}(s) - \tilde{Z}^{(k)(z)}(s)|]) ds, \end{aligned} \quad (4.15)$$

where the last inequality follows from Fubini's theorem and Jensen's inequality. Denote $u(t) := \mathbb{E}_k[|\tilde{X}^{(k)(x)}(t) - \tilde{Z}^{(k)(z)}(t)|]$ and $v(t) := |x-z| + 3H \int_0^t \rho(u(s)) ds$. Then by (4.15), we have $0 \leq u(t) \leq v(t)$. Define $G(r) := \int_1^r \frac{ds}{\rho(s)}$ for $r > 0$. Then G is nondecreasing and satisfies $\lim_{r \downarrow 0} G(r) = -\infty$ thanks to (4.1). In addition, we have

$$\begin{aligned} G(u(t)) &\leq G(v(t)) = G(|x-z|) + \int_0^t G'(v(s))v'(s) ds \\ &= G(|x-z|) + 3H \int_0^t \frac{\rho(u(s))}{\rho(v(s))} ds \leq G(|x-z|) + 3Ht, \end{aligned}$$

where the last inequality follows from the assumption that ρ is nondecreasing. Let also $G^{-1}(r) := \inf\{s \geq 0 : G(s) > r\}$ for $r \in \mathbb{R}$. Then G^{-1} is nondecreasing and satisfies $\lim_{r \rightarrow -\infty} G^{-1}(r) = 0$.

Furthermore, we have

$$0 \leq u(t) = \mathbb{E}_k[\tilde{X}^{(k)(x)}(t) - \tilde{Z}^{(k)(z)}(t)] \leq G^{-1}(G(|x - z|) + 3Ht). \quad (4.16)$$

In particular, when $|x - z| \rightarrow 0$, we see that the right most expression of (4.16) converges to 0 and so does $u(t)$. This implies (4.9) and hence completes the proof. \square

Lemma 4.7. Suppose that Assumptions 1.2, 4.1 and 4.2 hold. For each $k \in \mathbb{S}$, the killed Lévy type process $X^{(k)}$ introduced in (4.5) has Feller property.

Proof. For an arbitrarily fixed $k \in \mathbb{S}$, we need only to prove that for any given $t > 0$ and $f \in C_b(\mathbb{R}^d)$,

$$\begin{aligned} & |\mathbb{E}_k[f(X^{(k)(x)}(t))] - \mathbb{E}_k[f(X^{(k)(z)}(t))]| \\ &= \left| \mathbb{E}_k \left[f(\tilde{X}^{(k)(x)}(t)) \exp \left\{ \int_0^t q_{kk}(\tilde{X}^{(k)(x)}(s)) ds \right\} \right] \right. \\ & \quad \left. - \mathbb{E}_k \left[f(\tilde{X}^{(k)(z)}(t)) \exp \left\{ \int_0^t q_{kk}(\tilde{X}^{(k)(z)}(s)) ds \right\} \right] \right| \end{aligned} \quad (4.17)$$

tends to zero as $|x - z| \rightarrow 0$. Using the coupling process $(\tilde{X}^{(k)}, \tilde{Z}^{(k)})$ generated by the coupling generator $\tilde{\mathcal{L}}_k$ as in the proof of Lemma 4.6, we obtain that for any given $\varepsilon > 0$, the right-hand side of equality (4.17) equals

$$\begin{aligned} & \left| \mathbb{E}_k \left[f(\tilde{X}^{(k)(x)}(t)) \exp \left\{ \int_0^t q_{kk}(\tilde{X}^{(k)(x)}(s)) ds \right\} \right] \right. \\ & \quad \left. - \mathbb{E}_k \left[f(\tilde{Z}^{(k)(z)}(t)) \exp \left\{ \int_0^t q_{kk}(\tilde{Z}^{(k)(z)}(s)) ds \right\} \right] \right| \\ & \leq \mathbb{E}_k \left[\left| f(\tilde{X}^{(k)(x)}(t)) \exp \left\{ \int_0^t q_{kk}(\tilde{X}^{(k)(x)}(s)) ds \right\} \right. \right. \\ & \quad \left. \left. - f(\tilde{Z}^{(k)(z)}(t)) \exp \left\{ \int_0^t q_{kk}(\tilde{Z}^{(k)(z)}(s)) ds \right\} \right| \right] \\ & \leq \|f\| \mathbb{E}_k \left[\left| \exp \left\{ \int_0^t q_{kk}(\tilde{X}^{(k)(x)}(s)) ds \right\} - \exp \left\{ \int_0^t q_{kk}(\tilde{Z}^{(k)(z)}(s)) ds \right\} \right| \right] \\ & \quad + 2\|f\| \mathbb{E}_k \left[\exp \left\{ \int_0^t q_{kk}(\tilde{Z}^{(k)(z)}(s)) ds \right\} \mathbf{1}_{\{|f(\tilde{X}^{(k)(x)}(t)) - f(\tilde{Z}^{(k)(z)}(t))| \geq \varepsilon\}} \right] \\ & \quad + \varepsilon \mathbb{E}_k \left[\exp \left\{ \int_0^t q_{kk}(\tilde{Z}^{(k)(z)}(s)) ds \right\} \mathbf{1}_{\{|f(\tilde{X}^{(k)(x)}(t)) - f(\tilde{Z}^{(k)(z)}(t))| < \varepsilon\}} \right] \\ & := (4.18.I) + (4.18.II) + (4.18.III), \end{aligned} \quad (4.18)$$

where $\|f\|$ denotes the uniform (or supremum) norm of the function f . Noting that $q_{kk} \leq 0$ and the elementary inequality $|e^{-a} - e^{-b}| \leq |a - b|$ for $a, b > 0$, we obtain from (4.4) and (4.16) that

$$\begin{aligned} (4.18.I) & \leq \|f\| \mathbb{E}_k \left[\left| \int_0^t q_{kk}(\tilde{X}^{(k)(x)}(s)) ds - \int_0^t q_{kk}(\tilde{Z}^{(k)(z)}(s)) ds \right| \right] \\ & \leq (n_0 - 1)H\|f\| \int_0^t \mathbb{E}_k \left[|\tilde{X}^{(k)(x)}(s) - \tilde{Z}^{(k)(z)}(s)| \right] ds \\ & \leq (n_0 - 1)H\|f\| \int_0^t G^{-1}(G(|x - z|) + 3Hs) ds, \end{aligned} \quad (4.19)$$

where G and G^{-1} are the functions defined in the proof of Lemma 4.6. Since both G and G^{-1} are nondecreasing, for all $s \in [0, t]$ and $x, z \in \mathbb{R}^d$ with $|x - z| \leq 1$, we have $0 \leq G^{-1}(G(|x - z|) + 3Hs) \leq G^{-1}(G(1) + 3Ht) = G^{-1}(3Ht)$, which is integrable on the interval $[0, t]$. Thus it follows from the dominated convergence theorem and (4.16) that (4.18).I $\rightarrow 0$ as $|x - z| \rightarrow 0$. Moreover, in view of (4.16), we have that $\tilde{X}^{(k)(x)}(t)$ converges to $\tilde{Z}^{(k)(z)}(t)$ in probability \mathbb{P}_k as $|x - z| \rightarrow 0$. Thus, from the continuity of f , we obtain that $f(\tilde{X}^{(k)(x)}(t))$ also converges to $f(\tilde{Z}^{(k)(z)}(t))$ in probability \mathbb{P}_k as $|x - z| \rightarrow 0$. Combining this with $q_{kk} \leq 0$, we derive that

$$(4.18).II \leq 2\|f\| \mathbb{P}_k(|f(\tilde{X}^{(k)(x)}(t)) - f(\tilde{Z}^{(k)(z)}(t))| \geq \varepsilon) \rightarrow 0 \quad (4.20)$$

as $|x - z| \rightarrow 0$. Using the fact that $q_{kk} \leq 0$ again, we see that (4.18).III does not exceed ε ; which can be arbitrarily small. Combining this, (4.18), (4.19) and (4.20) together, we conclude that the right-hand side of equality (4.17) tends to zero as $|x - z| \rightarrow 0$. The proof is complete. \square

Lemma 4.8. Let Z be the subprocess of \tilde{Z} killed at the rate q with lifetime ζ , that is,

$$\mathbb{E}[f(Z^{(x)}(t))] = \mathbb{E}[f(\tilde{Z}^{(x)}(t)); t < \zeta] = \mathbb{E}\left[\exp\left\{-\int_0^t q(\tilde{Z}^{(x)}(s))ds\right\} f(\tilde{Z}^{(x)}(t))\right], \quad (4.21)$$

where \tilde{Z} is a right continuous strong Markov process, $q \geq 0$ on \mathbb{R}^d , and $f \in \mathcal{B}_b(\mathbb{R}^d)$. Then for any nonnegative function ϕ on \mathbb{R}^d and constant $\alpha > 0$, we have

$$\mathbb{E}[e^{-\alpha\zeta} \phi(Z^{(x)}(\zeta-))] = G_\alpha^Z(q\phi)(x), \quad (4.22)$$

where $\{G_\alpha^Z, \alpha > 0\}$ denotes the resolvent for the killed process Z .

Proof. By the definition of the resolvent and (4.21), we get

$$\begin{aligned} G_\alpha^Z(q\phi)(x) &= \mathbb{E}\left[\int_0^\infty e^{-\alpha t} (q\phi)(Z^{(x)}(t))dt\right] \\ &= \mathbb{E}\left[\int_0^\infty e^{-\alpha t} (q\phi)(\tilde{Z}^{(x)}(t)) \exp\left\{-\int_0^t q(\tilde{Z}^{(x)}(s))ds\right\} dt\right], \end{aligned}$$

which by page 286 in [26] (putting $m_t = \exp\{-\int_0^t q(\tilde{Z}(s))ds\} \mathbf{1}_{(t < \zeta)}$ there) equals the left-hand side in (4.22). \square

For each $k \in \mathbb{S}$, let $\{G_\alpha^{(k)}, \alpha > 0\}$ be the resolvent for the generator $\mathcal{L}_k + q_{kk}$. Let us also denote by $\{G_\alpha, \alpha > 0\}$ the resolvent for the generator \mathcal{A} defined in (1.1). Let

$$\begin{aligned} G_\alpha^0 &= \begin{pmatrix} G_\alpha^{(1)} & 0 & \cdots & 0 \\ 0 & G_\alpha^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & G_\alpha^{(n_0)} \end{pmatrix} \text{ and} \\ Q^0(x) &= Q(x) - \begin{pmatrix} q_{11}(x) & 0 & \cdots & 0 \\ 0 & q_{22}(x) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & q_{n_0 n_0}(x) \end{pmatrix}. \end{aligned}$$

Lemma 4.9. Suppose that [Assumption 1.2](#) holds. There exists a constant $\alpha_1 > 0$ such that for any $\alpha \geq \alpha_1$ and any $f(\cdot, k) \in \mathcal{B}_b(\mathbb{R}^d)$ with $k \in \mathbb{S}$,

$$G_\alpha f = G_\alpha^0 f + \sum_{m=1}^{\infty} G_\alpha^0 (\mathcal{Q}^0 G_\alpha^0)^m f. \quad (4.23)$$

Proof. Let $f(x, k) \geq 0$ on $\mathbb{R}^d \times \mathbb{S}$. Applying the strong Markov property at the first switching time τ and recalling the construction of (X, Λ) , we obtain

$$\begin{aligned} G_\alpha f(x, k) &= \mathbb{E}_{x,k} \left[\int_0^\infty e^{-\alpha t} f(X(t), \Lambda(t)) dt \right] \\ &= \mathbb{E}_{x,k} \left[\int_0^\tau e^{-\alpha t} f(X(t), k) dt \right] + \mathbb{E}_{x,k} \left[\int_\tau^\infty e^{-\alpha t} f(X(t), \Lambda(t)) dt \right] \\ &= G_\alpha^{(k)} f(x, k) + \mathbb{E}_{x,k} \left[e^{-\alpha \tau} G_\alpha f(X(\tau), \Lambda(\tau)) \right] \\ &= G_\alpha^{(k)} f(x, k) + \sum_{l \in \mathbb{S} \setminus \{k\}} \mathbb{E}_{x,k} \left[e^{-\alpha \tau} \left(-\frac{q_{kl}}{q_{kk}} \right) (X(\tau-)) G_\alpha f(X(\tau-), l) \right] \\ &= G_\alpha^{(k)} f(x, k) + \sum_{l \in \mathbb{S} \setminus \{k\}} G_\alpha^{(k)} (q_{kl} G_\alpha f(\cdot, l))(x), \end{aligned}$$

where the last equality follows from (4.22) in [Lemma 4.8](#). Hence we have

$$G_\alpha f(x, k) = G_\alpha^{(k)} f(\cdot, k)(x) + G_\alpha^{(k)} \left(\sum_{l \in \mathbb{S} \setminus \{k\}} q_{kl} G_\alpha f(\cdot, l) \right)(x). \quad (4.24)$$

Of course, we know that the second term on the right hand side of (4.24) equals

$$G_\alpha^{(k)} \left(\sum_{l \in \mathbb{S} \setminus \{k\}} q_{kl} G_\alpha^{(l)} f(\cdot, l) \right)(x) = G_\alpha^{(k)} \left(\sum_{l \in \mathbb{S} \setminus \{k\}} q_{kl} G_\alpha^{(l)} \left(\sum_{l_1 \in \mathbb{S} \setminus \{l\}} q_{ll_1} G_\alpha f(\cdot, l_1) \right) \right)(x).$$

Hence, we further obtain that for any fixed $k \in \mathbb{S}$ and any integer $m \geq 1$,

$$G_\alpha f(x, k) = \sum_{i=0}^m \psi_i^{(k)}(x) + R_m^{(k)}(x), \quad (4.25)$$

where

$$\begin{aligned} \psi_0^{(k)} &= G_\alpha^{(k)} f(\cdot, k), \\ \psi_1^{(k)} &= G_\alpha^{(k)} \left(\sum_{l \in \mathbb{S} \setminus \{k\}} q_{kl} G_\alpha^{(l)} f(\cdot, l) \right) = G_\alpha^{(k)} \left(\sum_{l \in \mathbb{S} \setminus \{k\}} q_{kl} \psi_0^{(l)} \right), \end{aligned}$$

and for $i \geq 1$,

$$\psi_i^{(k)} = G_\alpha^{(k)} \left(\sum_{l \in \mathbb{S} \setminus \{k\}} q_{kl} \psi_{i-1}^{(l)} \right).$$

By [Assumption 1.2](#) we know that $+\infty > H := \max\{\|q_{kk}\| : k \in \mathbb{S}\} \geq \max\{\|q_{kl}\| : k \neq l \in \mathbb{S}\}$, where $\|q_{kl}\|$ denotes the uniform (or supremum) norm of the function q_{kl} as before and constant H is the same as that in [Assumption 4.1](#). Therefore,

$$\|\psi_1^{(k)}\| \leq \sum_{l \in \mathbb{S} \setminus \{k\}} \|G_\alpha^{(k)}(q_{kl} \psi_0^{(l)})\| \leq \frac{H}{\alpha} \sum_{l \in \mathbb{S} \setminus \{k\}} \|\psi_0^{(l)}\|.$$

Thus, we get that

$$\sum_{k \in \mathbb{S}} \|\psi_1^{(k)}\| \leq \frac{(n_0 - 1)H}{\alpha} \sum_{k \in \mathbb{S}} \|\psi_0^{(k)}\| \leq \frac{1}{2} \sum_{k \in \mathbb{S}} \|\psi_0^{(k)}\|$$

when $\alpha \geq \alpha_1 := 2(n_0 - 1)H$. A similar argument yields that for $i \geq 1$,

$$\sum_{k \in \mathbb{S}} \|\psi_i^{(k)}\| \leq \frac{1}{2} \sum_{k \in \mathbb{S}} \|\psi_{i-1}^{(k)}\| \leq \frac{1}{2^i} \sum_{k \in \mathbb{S}} \|\psi_0^{(k)}\| \quad (4.26)$$

and

$$\|R_m^{(k)}(\cdot)\| \leq \frac{1}{2^m} \sum_{k \in \mathbb{S}} \|G_\alpha f(\cdot, k)\| \quad (4.27)$$

when $\alpha \geq \alpha_1$. Combining (4.26) and (4.27) with (4.25) and letting $m \uparrow \infty$, we conclude that for each $k \in \mathbb{S}$, $G_\alpha f(\cdot, k) = \sum_{i=0}^{\infty} \psi_i^{(k)}$, which clearly implies (4.23). The lemma is proved. \square

Lemma 4.9 and in particular (4.23) establishes the relationship between the resolvent of (X, Λ) and those of the killed Lévy type processes $X^{(k)}$, $k \in \mathbb{S}$. Now we are in the position to give the proof of **Theorem 4.4**.

Proof of Theorem 4.4. Denote the transition probability family of Markov process (X, Λ) by $\{P(t, (x, k), A) : t \geq 0, (x, k) \in \mathbb{R}^d \times \mathbb{S}, A \in \mathcal{B}(\mathbb{R}^d \times \mathbb{S})\}$. We first prove that for any given $t > 0$, $x \in \mathbb{R}^d$, $k, l \in \mathbb{S}$ and $A \in \mathcal{B}(\mathbb{R}^d)$,

$$\begin{aligned} &P(t, (x, k), A \times \{l\}) \\ &= \delta_{kl} P^{(k)}(t, x, A) + \sum_{m=1}^{+\infty} \int_{0 < t_1 < t_2 < \dots < t_m < t} \dots \int \sum_{\substack{l_0, l_1, l_2, \dots, l_m \in \mathbb{S} \\ l_i \neq l_{i+1}, l_0 = k, l_m = l}} \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} P^{(l_0)} \\ &\quad \times (t_1, x, dy_1) q_{l_0 l_1}(y_1) \\ &\quad \times P^{(l_1)}(t_2 - t_1, y_1, dy_2) \dots q_{l_{m-1} l_m}(y_m) P^{(l_m)}(t - t_m, y_m, A) dt_1 dt_2 \dots dt_m, \end{aligned} \quad (4.28)$$

where δ_{kl} is the Kronecker symbol in k, l , which equals 1 if $k = l$ and 0 if $k \neq l$. To prove (4.28), denote its the right-hand side by $\tilde{P}(t, (x, k), A \times \{l\})$ for brevity. For any bounded function $f(x, k)$ defined on $\mathbb{R}^d \times \mathbb{S}$ such that $f(\cdot, k)$ is Lipschitz continuous for each $k \in \mathbb{S}$, we define

$$P_t f(x, k) := \mathbb{E}_{x, k}[f(X(t), \Lambda(t))] = \sum_{l \in \mathbb{S}} \int_{\mathbb{R}^d} f(y, l) P(t, (x, k), dy \times \{l\}), \quad (4.29)$$

and

$$\begin{aligned} \tilde{P}_t f(x, k) &:= \sum_{l \in \mathbb{S}} \int_{\mathbb{R}^d} f(y, l) \tilde{P}(t, (x, k), dy \times \{l\}) \\ &= \sum_{l \in \mathbb{S}} \left[\int_{\mathbb{R}^d} \delta_{kl} f(y, l) P^{(k)}(t, x, dy) + \sum_{m=1}^{+\infty} \int_{0 < t_1 < t_2 < \dots < t_m < t} \dots \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} P^{(l_0)}(t_1, x, dy_1) q_{l_0 l_1}(y_1) P^{(l_1)}(t_2 - t_1, y_1, dy_2) \dots \right. \\ &\quad \sum_{\substack{l_0, l_1, l_2, \dots, l_m \in \mathbb{S} \\ l_i \neq l_{i+1}, l_0 = k, l_m = l}} \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} P^{(l_0)}(t_1, x, dy_1) q_{l_0 l_1}(y_1) P^{(l_1)}(t_2 - t_1, y_1, dy_2) \dots \\ &\quad \left. \times q_{l_{m-1} l_m}(y_m) P^{(l_m)}(t - t_m, y_m, dy) dt_1 dt_2 \dots dt_m \right]. \end{aligned} \quad (4.30)$$

Since the process (X, Λ) has right continuous sample paths, it follows from the continuity of f and the bounded convergence theorem that the function $t \mapsto P_t f(x, k)$ is right continuous. Similarly for every $l \in \mathbb{S}$ and each $m = 0, 1, \dots$, every term on the right-hand side of (4.30) is a right-continuous function in t . Moreover, using Assumption 1.2 and the boundedness of the function f , we can see that the series on the right-hand side of (4.30) is absolutely convergent. Therefore it follows that the function $t \mapsto \tilde{P}_t f(x, k)$ is also right continuous.

On the other hand, using Lemma 4.9 and in particular (4.23), for any $\alpha > 0$, we have

$$\int_0^\infty e^{-(\alpha+\alpha_1)t} e^{\alpha_1 t} P_t f(x, k) dt = \int_0^\infty e^{-(\alpha+\alpha_1)t} e^{\alpha_1 t} \tilde{P}_t f(x, k) dt,$$

where α_1 is as in the statement of Lemma 4.9. Since both $P_t f(x, k)$ and $\tilde{P}_t f(x, k)$ are right continuous in t , we can apply the uniqueness theorem of Laplace transform (refer to [3, Theorem 1.38]) to conclude that $e^{\alpha_1 t} P_t f(x, k) = e^{\alpha_1 t} \tilde{P}_t f(x, k)$. That is,

$$\sum_{l \in \mathbb{S}} \int f(y, l) P(t, (x, k), dy \times \{l\}) = \sum_{l \in \mathbb{S}} \int f(y, l) \tilde{P}(t, (x, k), dy \times \{l\}). \quad (4.31)$$

Now we prove (4.28) by the Monotone Class Theorem (see, e.g., Theorem 1.35 in [3]). Denote by L the family of bounded and Borel measurable functions defined on $\mathbb{R}^d \times \mathbb{S}$ such that (4.31) holds. From the above argument, we know that L contains all bounded and Lipschitz continuous functions on $\mathbb{R}^d \times \mathbb{S}$. Next we show that L is a so-called \mathcal{L} -system (c.f. Definition 1.34 in Section 1.5 of [3]). Firstly, L obviously contains the constant function 1. Secondly, for c_1 and c_2 in \mathbb{R} and f_1 and f_2 in L , we clearly have $c_1 f_1 + c_2 f_2$ in L . Thirdly, if $f_n \in L$ with $0 \leq f_n \uparrow f$, then $f \in L$ by the monotone convergence theorem. Hence, according to the definition of \mathcal{L} -system [3, Definition 1.34], L is an \mathcal{L} -system. Moreover, let \mathcal{C} denote the set of all the open sets in $\mathbb{R}^d \times \mathbb{S}$. Note that \mathcal{C} is a π -system and recall that L contains the set of all bounded Lipschitz continuous functions defined on $\mathbb{R}^d \times \mathbb{S}$. Therefore, by virtue of the monotone class theorem (refer to [3, Theorem 1.35]), the family L contains the set of all bounded measurable functions defined on $\mathbb{R}^d \times \mathbb{S}$. In particular, for any given $A \in \mathcal{B}(\mathbb{R}^d)$ and $l \in \mathbb{S}$, the family L contains the function $\mathbf{1}_{A \times \{l\}}(x, k)$, which implies that (4.28) holds.

Finally, we use (4.28) to prove the Feller property for (X, Λ) . By Lemma 4.7, we know that for every $k \in \mathbb{S}$, $X^{(k)}$ has the Feller property. Therefore, in view of Proposition 6.1.1 in [17] and Assumption 1.2, we derive that $P^{(k)}(t, x, A)$ and every term in the series on the right-hand side of (4.28) are lower semicontinuous with respect to x whenever A is an open set in $\mathcal{B}(\mathbb{R}^d)$. This then implies that the left-hand side of (4.28) is lower semicontinuous with respect to (x, k) for every $l \in \mathbb{S}$ whenever A is an open set in $\mathcal{B}(\mathbb{R}^d)$ by noting that \mathbb{S} is a finite set and has discrete metric. Consequently, (X, Λ) has the Feller property (see Proposition 6.1.1 in [17] again). The theorem is proved. \square

5. Strong Feller property

In this section, we study the strong Feller property for the coordinate process $(X(t), \Lambda(t))$ in the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ as specified in Section 4. We first make the following assumption.

Assumption 5.1. There exists a $\lambda_0 > 0$ such that $\langle \xi, a(x, k)\xi \rangle \geq \lambda_0 |\xi|^2$ for all $x, \xi \in \mathbb{R}^d$ and $k \in \mathbb{S}$. Denote by $\sigma_{\lambda_0}(x, k)$ the unique symmetric nonnegative definite matrix-valued function such that $\sigma_{\lambda_0}^2(x, k) = a(x, k) - \lambda_0 I$. In addition, there exist positive constants δ_0, H and a

nonnegative function ϑ defined on $[0, \delta_0]$ satisfying $\lim_{r \rightarrow 0} \vartheta(r) = 0$ such that

$$2\langle x - z, b(x, k) - b(z, k) \rangle + |\sigma_{\lambda_0}(x, k) - \sigma_{\lambda_0}(z, k)|^2 \leq 2H|x - z|\vartheta(|x - z|), \quad (5.1)$$

$$\int_{\mathbb{R}_0^d} |u| \|v(x, k, \cdot) - v(z, k, \cdot)\| (du) \leq \frac{H}{2} \vartheta(|x - z|) \quad (5.2)$$

for all $x, z \in \mathbb{R}^d$ with $|x - z| \leq \delta_0$ and all $k \in \mathbb{S}$.

Remark 5.2. The uniform ellipticity condition for the diffusion matrix $a(x, k)$ in [Assumption 5.1](#) is quite standard in the literature. Indeed, similar assumptions were used in [\[20–22\]](#) to obtain the strong Feller property.

Proposition 5.3. Under [Assumptions 4.1, 4.2, and 5.1](#), for each $k \in \mathbb{S}$, both the Lévy type process $\tilde{X}^{(k)}$ and the killed Lévy type process $X^{(k)}$ are strong Feller.

Proof. The proof is motivated by Priola and Wang [\[21\]](#). Fix an arbitrary $k \in \mathbb{S}$ throughout the proof. Let $\sigma_{\lambda_0}(x, k)$ be as in [Assumption 5.1](#) and put $c(x, z, k) := \lambda_0(I - 2(x - z)(x - z)^T / |x - z|^2) + \sigma_{\lambda_0}(x, k)\sigma_{\lambda_0}(z, k)^T$ for all $x, z \in \mathbb{R}^d$. For $x, z \in \mathbb{R}^d$, set

$$\widehat{a}(x, z, k) = \begin{pmatrix} a(x, k) & c(x, z, k) \\ c(x, z, k)^T & a(z, k) \end{pmatrix}, \quad b(x, z, k) = \begin{pmatrix} b(x, k) \\ b(z, k) \end{pmatrix}.$$

We can verify directly that $\widehat{a}(x, z, k)$ is symmetric and nonnegative definite for all $x, z \in \mathbb{R}^d$. Then we define

$$\widehat{\Omega}_d(k)h(x, z) := \frac{1}{2} \text{tr}(\widehat{a}(x, z, k) \nabla^2 h(x, z)) + \langle b(x, z, k), \nabla h(x, z) \rangle,$$

and

$$\widehat{\mathcal{L}}_k h(x, z) := \widehat{\Omega}_d(k)h(x, z) + \widetilde{\Omega}_j(k)h(x, z), \quad (5.3)$$

where $h \in C_0^2(\mathbb{R}^d \times \mathbb{R}^d)$ and $\widetilde{\Omega}_j(k)$ is defined in [\(4.7\)](#). Let

$$A(x, z, k) = a(x, k) + a(z, k) - 2c(x, z, k),$$

$$\overline{A}(x, z, k) = \frac{1}{|x - z|^2} \langle x - z, A(x, z, k)(x - z) \rangle,$$

$$B(x, z, k) = \langle x - z, b(x, k) - b(z, k) \rangle.$$

Straightforward computations lead to

$$\text{tr}(A(x, z, k)) = \|\sigma(x, k) - \sigma(z, k)\|^2 + 4\lambda_0 \text{ and } \overline{A}(x, z, k) \geq 4\lambda_0. \quad (5.4)$$

Consider the function $F(r) := \frac{r}{1+r}$, $r \geq 0$. Then $F'(r) = \frac{1}{(1+r)^2} > 0$ and $F''(r) = \frac{-2}{(1+r)^3} < 0$ for all $r \geq 0$. Consequently it follows from [\(5.1\)](#) and [\(5.4\)](#) that

$$\begin{aligned} \widetilde{\Omega}_d(k)F(|x - z|) &= \frac{1}{2} F''(|x - z|) \overline{A}(x, z, k) \\ &\quad + \frac{F'(|x - z|)}{2|x - z|} [\text{tr}(A(x, z, k)) - \overline{A}(x, z, k) + 2B(x, z, k)] \\ &\leq 2\lambda_0 F''(|x - z|) + H F'(|x - z|) \vartheta(|x - z|) \\ &= \frac{-4\lambda_0}{(1 + |x - z|)^3} + \frac{H}{(1 + |x - z|)^2} \vartheta(|x - z|), \end{aligned} \quad (5.5)$$

for all $x, z \in \mathbb{R}^d$ with $|x - z| \leq \delta_0$.

Next we estimate $\tilde{\mathcal{Q}}_j(k)F(|x - z|)$. To this end, we note that since F is concave, it follows that for any $x, z \in \mathbb{R}^d$ and $u \in \mathbb{R}_0^d$, we have

$$\begin{aligned} & F(|x + u - z|) - F(|x - z|) - \langle \nabla_x F(|x - z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u) \\ & \leq F'(|x - z|)(|x + u - z| - |x - z|) - \frac{F'(|x - z|)}{|x - z|} \langle x - z, u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u) \\ & \leq \frac{2|u|}{(1 + |x - z|)^2}. \end{aligned}$$

Hence it follows that

$$\begin{aligned} & \int [F(|x + u - z|) - F(|x - z|) - \langle \nabla_x F(|x - z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u)] \\ & \quad \times (v(x, k, du) - v(z, k, du))^+ \\ & \leq \frac{2}{(1 + |x - z|)^2} \int |u| (v(x, k, du) - v(z, k, du))^+. \end{aligned}$$

In the same manner, we have

$$\begin{aligned} & \int [F(|x - (z + u)|) - F(|x - z|) - \langle \nabla_z F(|x - z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u)] \\ & \quad \times (v(x, k, du) - v(z, k, du))^+ \\ & \leq \frac{2}{(1 + |x - z|)^2} \int |u| (v(z, k, du) - v(x, k, du))^+. \end{aligned}$$

On the other hand, since $\nabla_x F(|x - z|) = -\nabla_z F(|x - z|)$, we have

$$\begin{aligned} & \int [F(|x + u - z - u|) - F(|x - z|) - \langle \nabla_x F(|x - z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u) \\ & \quad - \langle \nabla_z F(|x - z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u)] (v(x, k, du) \wedge v(z, k, du)) = 0. \end{aligned}$$

Then, using the definition of $\tilde{\mathcal{Q}}_j(k)$ in (4.7) and condition (5.2), we obtain

$$\begin{aligned} & \tilde{\mathcal{Q}}_j(k)F(|x - z|) \\ & \leq \frac{2}{(1 + |x - z|)^2} \left[\int |u| (v(x, k, du) - v(z, k, du))^+ \right. \\ & \quad \left. + \int |u| (v(z, k, du) - v(x, k, du))^+ \right] \\ & \leq \frac{2}{(1 + |x - z|)^2} \int |u| \|v(x, k, \cdot) - v(z, k, \cdot)\| (du) \\ & \leq \frac{H}{(1 + |x - z|)^2} \vartheta(|x - z|), \end{aligned} \tag{5.6}$$

for all $x, z \in \mathbb{R}^d$ with $|x - z| \leq \delta_0$.

Plugging (5.5) and (5.6) into (5.3), we obtain that for all $x, z \in \mathbb{R}^d$ with $|x - z| \leq \delta_0$,

$$\begin{aligned} \hat{\mathcal{L}}_k F(|x - z|) & \leq \frac{-4\lambda_0}{(1 + |x - z|)^3} + \frac{2H}{(1 + |x - z|)^2} \vartheta(|x - z|) \\ & \leq \frac{-4\lambda_0}{(1 + \delta_0)^3} + 2H\vartheta(|x - z|). \end{aligned}$$

Furthermore, since $\lambda_0 > 0$ and $\lim_{r \downarrow 0} \vartheta(r) = 0$, it follows that there exist positive constants κ and δ ($0 < \delta < \delta_0$), we have

$$\widehat{\mathcal{L}}_k F(|x - z|) \leq -\kappa, \text{ for all } 0 < |x - z| \leq \delta. \quad (5.7)$$

Given $x \neq z$ with $\delta > |x - z| > \frac{1}{m_0}$, where $m_0 \in \mathbb{N}$. Let $(\tilde{X}^{(k)(x)}, \tilde{Z}^{(k)(z)})$ be the coupling process corresponding to the operator $\widehat{\mathcal{L}}_k$ and denote by T the coupling time. For $n, N \in \mathbb{N}$ and the δ in (5.7), define

$$T_n := \inf \left\{ t \geq 0 : |\tilde{X}^{(k)(x)}(t) - \tilde{Z}^{(k)(z)}(t)| < \frac{1}{n} \right\},$$

$$\sigma_N := \inf \{ t \geq 0 : |\tilde{X}^{(k)(x)}(t)| + |\tilde{Z}^{(k)(z)}(t)| > N \},$$

and

$$S_\delta := \inf \{ t \geq 0 : |\tilde{X}^{(k)(x)}(t) - \tilde{Z}^{(k)(z)}(t)| > \delta \}.$$

We have

$$\begin{aligned} 0 &\leq F(\delta) \mathbb{P}_k \{ T_n \wedge \sigma_N > S_\delta \} \\ &\leq \mathbb{E}_k [F(|\tilde{X}^{(k)(x)}(T_n \wedge S_\delta \wedge \sigma_N) - \tilde{Z}^{(k)(z)}(T_n \wedge S_\delta \wedge \sigma_N)|)] \\ &= F(|x - z|) + \mathbb{E}_k \left[\int_0^{T_n \wedge S_\delta \wedge \sigma_N} \widehat{\mathcal{L}}_k F(|\tilde{X}^{(k)(x)} - \tilde{Z}^{(k)(z)}|) ds \right] \\ &\leq F(|x - z|) - \kappa \mathbb{E}_k [T_n \wedge S_\delta \wedge \sigma_N], \end{aligned}$$

where the last inequality follows from (5.7). Then it follows that

$$F(\delta) \mathbb{P}_k \{ T_n \wedge \sigma_N > S_\delta \} + \kappa \mathbb{E}_k [T_n \wedge S_\delta \wedge \sigma_N] \leq F(|x - z|).$$

Since $T_n \rightarrow T$ a.s. as $n \rightarrow \infty$ and $\sigma_N \rightarrow \infty$ a.s. as $N \rightarrow \infty$, we have

$$F(\delta) \mathbb{P}_k \{ T > S_\delta \} + \kappa \mathbb{E}_k [T \wedge S_\delta] \leq F(|x - z|).$$

Then for any $t > 0$ and $0 < |x - z| < \delta$,

$$\begin{aligned} \mathbb{P}_k \{ T > t \} &= \mathbb{P}_k \{ T > t, S_\delta > t \} + \mathbb{P}_k \{ T > t, S_\delta \leq t \} \\ &\leq \mathbb{P}_k \{ T \wedge S_\delta > t \} + \mathbb{P}_k \{ T > S_\delta \} \\ &\leq \frac{1}{t} \mathbb{E}_k [T \wedge S_\delta] + \mathbb{P}_k \{ T > S_\delta \} \\ &\leq \left(\frac{1}{t\kappa} + \frac{1}{F(\delta)} \right) F(|x - z|). \end{aligned}$$

This implies the strong Feller property for the Lévy type process $\tilde{X}^{(k)}$ immediately. Indeed, for any $f \in \mathcal{B}_b(\mathbb{R}^d)$, $t > 0$, and $0 < |x - z| < \delta$, we have

$$\begin{aligned} |\mathbb{E}_k [f(\tilde{X}^{(k)(x)}(t))] - \mathbb{E}_k [f(\tilde{X}^{(k)(z)}(t))]| &\leq \mathbb{E}_k [|f(\tilde{X}^{(k)(x)}(t)) - f(\tilde{X}^{(k)(z)}(t))|] \\ &= \mathbb{E}_k [|f(\tilde{X}^{(k)(x)}(t)) - f(\tilde{X}^{(k)(z)}(t))| I_{\{T > t\}}] \\ &\leq 2 \|f\|_\infty \mathbb{P}_k \{ T > t \} \rightarrow 0, \text{ as } |x - z| \rightarrow 0. \end{aligned}$$

Finally, as in the proof of [Lemma 4.7](#), for any $f \in \mathcal{B}_b(\mathbb{R}^d)$, $t > 0$, and $0 < |x - z| < \delta$, we can write

$$\begin{aligned}
 & \left| \mathbb{E}_k[f(X^{(k)(x)}(t))] - \mathbb{E}_k[f(X^{(k)(z)}(t))] \right| \\
 & \leq \mathbb{E}_k \left[\left| f(\tilde{X}^{(k)(x)}(t)) e^{\int_0^t q_{kk}(\tilde{X}^{(k)(x)}(s)) ds} - f(\tilde{Z}^{(k)(z)}(t)) e^{\int_0^t q_{kk}(\tilde{Z}^{(k)(z)}(s)) ds} \right| \right] \\
 & \leq \mathbb{E}_k \left[\left| f(\tilde{X}^{(k)(x)}(t)) - f(\tilde{Z}^{(k)(z)}(t)) \right| e^{\int_0^t q_{kk}(\tilde{X}^{(k)(x)}(s)) ds} \right] \\
 & \quad + \mathbb{E}_k \left[\left| f(\tilde{Z}^{(k)(z)}(t)) \right| e^{\int_0^t q_{kk}(\tilde{X}^{(k)(x)}(s)) ds} - e^{\int_0^t q_{kk}(\tilde{Z}^{(k)(z)}(s)) ds} \right] \\
 & \leq 2\|f\|_\infty \mathbb{P}_k\{T > t\} + \|f\|_\infty \mathbb{E}_k \left[\left| \int_0^t q_{kk}(\tilde{X}^{(k)(x)}(s)) ds - \int_0^t q_{kk}(\tilde{Z}^{(k)(z)}(s)) ds \right| \right] \\
 & \leq 2\|f\|_\infty \left(\frac{1}{t\kappa} + \frac{1}{F(\delta)} \right) F(|x - z|) + H(n_0 - 1)\|f\|_\infty \\
 & \quad \times \mathbb{E}_k \left[\int_0^t |\tilde{X}^{(k)(x)}(s) - \tilde{Z}^{(k)(z)}(s)| ds \right] \\
 & \leq 2\|f\|_\infty \left(\frac{1}{t\kappa} + \frac{1}{F(\delta)} \right) F(|x - z|) + H(n_0 - 1)\|f\|_\infty \int_0^t G^{-1}(G(|x - z|) + 3Hs) ds,
 \end{aligned}$$

where the second last inequality above follows from [Assumption 4.2](#) and the last inequality follows from (4.16). Note that $F(\cdot)$ is continuous with $F(0) = 0$. In addition, recall that we argued in the proof of [Lemma 4.7](#) that $\int_0^t G^{-1}(G(|x - z|) + 3Hs) ds \rightarrow 0$ as $|x - z| \rightarrow 0$. Thus it follows that $|\mathbb{E}_k[f(X^{(k)(x)}(t))] - \mathbb{E}_k[f(X^{(k)(z)}(t))]| \rightarrow 0$ as $|x - z| \rightarrow 0$. On the other hand, for any $|x - z| \geq \delta$, we have

$$|\mathbb{E}_k[f(X^{(k)(x)}(t))] - \mathbb{E}_k[f(X^{(k)(z)}(t))]| \leq 2\delta^{-1}\|f\|_\infty|x - z|.$$

Therefore we obtain the desired strong Feller property for the killed Lévy process $X^{(k)}$. This completes the proof. \square

With [Proposition 5.3](#) at our hands, we can use exactly the same arguments as those in the proof of [Theorem 4.4](#) to establish the following theorem.

Theorem 5.4. *Under the conditions of [Proposition 5.3](#), the process (X, Λ) possesses the strong Feller property.*

Acknowledgments

The authors would like to thank the anonymous reviewers for their useful comments and suggestions. The research was supported in part by the National Natural Science Foundation of China under Grant No. 11671034, the Beijing Natural Science Foundation under Grant No. 1172001, the Simons foundation collaboration under Grant No. 523736, and a grant from the Research Growth Initiative of UW-Milwaukee.

References

- [1] R.F. Bass, Uniqueness in law for pure jump Markov processes, *Probab. Theory Related Fields* 79 (2) (1988) 271–287.
- [2] R.F. Bass, H. Tang, The martingale problem for a class of stable-like processes, *Stochastic Process. Appl.* 119 (4) (2009) 1144–1167.

- [3] M.-F. Chen, *From Markov Chains to Non-Equilibrium Particle Systems*, second ed., World Scientific Publishing Co. Inc., River Edge, NJ, 2004.
- [4] M.F. Chen, S.F. Li, Coupling methods for multidimensional diffusion processes, *Ann. Probab.* 17 (1) (1989) 151–177.
- [5] B. Cloez, M. Hairer, Exponential ergodicity for Markov processes with random switching, *Bernoulli* 21 (1) (2015) 505–536.
- [6] D.A. Dawson, X. Zheng, Law of large numbers and central limit theorem for unbounded jump mean-field models, *Adv. Appl. Math.* 12 (3) (1991) 293–326.
- [7] S. Feng, X. Zheng, Solutions of a class of nonlinear master equations, *Stochastic Process. Appl.* 43 (1) (1992) 65–84.
- [8] M. Hairer, J.C. Mattingly, M. Scheutzow, Asymptotic coupling and a general form of Harris’ theorem with applications to stochastic delay equations, *Probab. Theory Related Fields* 149 (1–2) (2011) 223–259.
- [9] W. Hoh, The martingale problem for a class of pseudo-differential operators, *Math. Ann.* 300 (1) (1994) 121–147.
- [10] N. Ikeda, S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, second ed., in: North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Kodansha, Ltd., Amsterdam, Tokyo, 1989.
- [11] I. Karatzas, S.E. Shreve, *Brownian Motion and Stochastic Calculus*, second ed., in: Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991.
- [12] T. Komatsu, Markov processes associated with certain integro-differential operators, *Osaka J. Math.* 10 (1973) 271–303.
- [13] T.G. Kurtz, Martingale problems for conditional distributions of Markov processes, *Electron. J. Probab.* 3 (9) (1998) 29 (electronic).
- [14] T. Lindvall, *Lectures on the Coupling Method*, Dover Publications Inc., Mineola NY, 2002. Corrected reprint of the 1992 original.
- [15] T. Lindvall, L.C.G. Rogers, Coupling of multidimensional diffusions by reflection, *Ann. Probab.* 14 (3) (1986) 860–872.
- [16] X. Mao, C. Yuan, *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, London, 2006.
- [17] S.P. Meyn, R.L. Tweedie, *Markov Chains and Stochastic Stability*, in: Communications and Control Engineering Series, Springer-Verlag London, Ltd., London, 1993.
- [18] R. Mikulevicius, B.L. Rozovskii, Martingale problems for stochastic PDE’s, in: *Stochastic Partial Differential Equations: Six Perspectives*, in: Math. Surveys Monogr., vol. 64, Amer. Math. Soc., Providence, RI, 1999, pp. 243–325.
- [19] E. Perkins, On the martingale problem for interactive measure-valued branching diffusions, *Mem. Amer. Math. Soc.* 115 (549) (1995) vi+89.
- [20] S. Peszat, J. Zabczyk, Strong Feller property and irreducibility for diffusions on Hilbert spaces, *Ann. Probab.* 23 (1) (1995) 157–172.
- [21] E. Priola, F.-Y. Wang, Gradient estimates for diffusion semigroups with singular coefficients, *J. Funct. Anal.* 236 (1) (2006) 244–264.
- [22] H. Qiao, Exponential ergodicity for SDEs with jumps and non-Lipschitz coefficients, *J. Theoret. Probab.* 27 (1) (2014) 137–152.
- [23] W. Rudin, *Principles of Mathematical Analysis*, third ed., in: International Series in Pure and Applied Mathematics, McGraw-Hill Book Co., New York-Auckland-Düsseldorf, 1976.
- [24] S.P. Sethi, Q. Zhang, *Hierarchical Decision Making in Stochastic Manufacturing Systems*, in: Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA, 1994.
- [25] J. Shao, F. Xi, Stability and recurrence of regime-switching diffusion processes, *SIAM J. Control Optim.* 52 (6) (2014) 3496–3516.
- [26] M. Sharpe, *General Theory of Markov Processes*, in: Pure and Applied Mathematics, vol. 133, Academic Press, Inc., Boston, MA, 1988.
- [27] T. Shiga, H. Tanaka, Central limit theorem for a system of Markovian particles with mean field interactions, *Z. Wahrscheinlichkeitstheor. Verwandte Geb.* 69 (3) (1985) 439–459.
- [28] D.W. Stroock, Diffusion processes associated with Lévy generators, *Z. Wahrscheinlichkeitstheor. Verwandte Geb.* 32 (3) (1975) 209–244.
- [29] D.W. Stroock, S.R.S. Varadhan, Diffusion processes with continuous coefficients. I, *Comm. Pure Appl. Math.* 22 (1969) 345–400.
- [30] D.W. Stroock, S.R.S. Varadhan, Diffusion processes with continuous coefficients. II, *Comm. Pure Appl. Math.* 22 (1969) 479–530.

- [31] D.W. Stroock, S.R.S. Varadhan, Multidimensional Diffusion Processes, in: Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 233, Springer-Verlag, Berlin, 1979.
- [32] J. Wang, Regularity of semigroups generated by Lévy type operators via coupling, *Stochastic Process. Appl.* 120 (9) (2010) 1680–1700.
- [33] J.-M. Wang, Martingale problems for switched processes, *Math. Nachr.* 287 (10) (2014) 1186–1201.
- [34] F. Xi, A note on the jump numbers of Q -processes, *J. Math. (Wuhan)* 18 (2) (1998) 187–190.
- [35] F. Xi, Feller property and exponential ergodicity of diffusion processes with state-dependent switching, *Sci. China Ser. A* 51 (3) (2008) 329–342.
- [36] F. Xi, Asymptotic properties of jump-diffusion processes with state-dependent switching, *Stochastic Process. Appl.* 119 (7) (2009) 2198–2221.
- [37] F. Xi, L. Zhao, On the stability of diffusion processes with state-dependent switching, *Sci. China Ser. A* 49 (9) (2006) 1258–1274.
- [38] G. Yin, F. Xi, Stability of regime-switching jump diffusions, *SIAM J. Control Optim.* 48 (7) (2010) 4525–4549.
- [39] G.G. Yin, Q. Zhang, Continuous-Time Markov Chains and Applications: A Singular Perturbation Approach, in: Applications of Mathematics (New York), vol. 37, Springer-Verlag, New York, 1998.
- [40] G.G. Yin, C. Zhu, Hybrid Switching Diffusions: Properties and Applications, in: Stochastic Modelling and Applied Probability, vol. 63, Springer, New York, 2010.
- [41] L. Zambotti, An analytic approach to existence and uniqueness for martingale problems in infinite dimensions, *Probab. Theory Related Fields* 118 (2) (2000) 147–168.
- [42] J.L. Zheng, X.G. Zheng, A martingale method for Q -processes, *Kexue Tongbao (Chinese)* 31 (17) (1986) 1296–1298.
- [43] C. Zhu, G. Yin, N.A. Baran, Feynman-Kac formulas for regime-switching jump diffusions and their applications, *Stochastics* 87 (6) (2015) 1000–1032.