On the Martingale Problem and Feller and Strong Feller Properties for Weakly Coupled Lévy Type Operators

Fubao Xi*        Chao Zhu†

August 28, 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 \((X, \Lambda)\). The process \((X, \Lambda)\), 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.

Key Words and Phrases. Weakly coupled Lévy type operator, martingale problem, Feller property, strong Feller property, coupling method.

2000 MR Subject Classification. 60J25, 60J27, 60J60, 60J75.

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 \(\mathbb{S} := \{1, 2, \cdots, n_0\}\). For all “nice” functions \(f : \mathbb{R}^d \times \mathbb{S} \rightarrow \mathbb{R}\), we define

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

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

\[
\mathcal{L}_k f(x, k) := \frac{1}{2} \text{tr} (a(x, k) \nabla^2 f(x, k)) + \langle b(x, k), \nabla f(x, k) \rangle
\]

\[
+ \int_{\mathbb{R}^d} (f(x + u, k) - f(x, k) - (\nabla f(x, k), u) 1_{B(0, \varepsilon_0)}(u)) \nu(x, k, du),
\] (1.2)

*School of Mathematics and Statistics, Beijing Institute of Technology, Beijing 100081, China, xifb@bit.edu.cn.
†Department of Mathematical Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, zhu@uwm.edu.
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 \(\sigma\)-finite measure on \(\mathbb{R}_0^d\) satisfying

\[
\int_{\mathbb{R}_0^d} |u|^2 1_{B(0, \varepsilon_0)}(u) \nu(x, k, du) < \infty \text{ and } \nu(x, k, \mathbb{R}^d \setminus B(0, \varepsilon_0)) < \infty,
\]

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:

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

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 \(A\) 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 \(\sigma\)-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 \(A\) starting from \((x, k)\), if \(\mathbb{P}^{(x,k)}((X(0), \Lambda(0)) = (x, k)) = 1\) and for each function \(f \in C_c^\infty(\mathbb{R}^d \times \mathbb{S})\),

\[
M_t^{(f)} := f(X(t), \Lambda(t)) - f(X(0), \Lambda(0)) - \int_0^t Af(X(s), \Lambda(s))ds
\]

is an \(\{\mathcal{F}_t\}\)-martingale with respect to \(\mathbb{P}^{(x,k)}\), where \((X, \Lambda)\) 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 \(A\) starting from \((x, k)\). We often call the coordinate process \((X, \Lambda)\) the *regime-switching jump diffusion with Lévy type jumps*.

Since the seminal work of Stroock and Varadhan (Stroock and Varadhan (1969a,b)) 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 (1973) and Stroock (1975) prove that the martingale problem for a Lévy type operator is well-posed; Bass (1988) investigates the martingale problem for pure jump Markov processes; Dawson and Zheng (1991) and Feng and Zheng (1992) considers the martingale problem;
problem for a class of nonlinear master equations for chemical reaction models; Xi (1998) and Zheng and Zheng (1986) discuss the martingale problem for $Q$-processes; Zambotti (2000) provides an analytic approach for existence and uniqueness for martingale problems in infinite dimensions; Kurtz (1998) presents a martingale problems for conditional distributions of Markov processes; Mikulevicius and Rozovskii (1999) studies martingale problems for stochastic partial differential equations; Perkins (1995) investigates the martingale problem for interactive measure-valued branching diffusions; Hoh (1994) investigates the martingale problems for pseudo-differential operators; and Bass and Tang (2009) is devoted to the martingale problem for stable-like processes.

This paper is motivated by Stroock (1975) and considers weakly coupled Lévy type operator $A$ 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 $k \in S$, $A$ also contains a component $Q(x)$ defined in (1.4), which provides the switching mechanism for the operators $L_k, k \in S$. In other words, the operators $L_k, k \in S$ are coupled through the operator $Q(x)$ of (1.4). Therefore it is convenient to call the operator $A$ of (1.1) a weakly coupled Lévy type operator and the coordinate process $(X, \Lambda)$ a regime-switching Lévy type process. Here we remark that $Q(x) = (q_{kj}(x))$ depends on $x$. When the Lévy kernel $\nu(x, k, dz)$ is independent of $(x, k)$, then $A$ reduces to the infinitesimal generator of a regime-switching jump diffusion process as those considered in Xi (2009), Yin and Xi (2010), Zhu et al. (2015). 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, Cloez and Hairer (2015), Mao and Yuan (2006), Sethi and Zhang (1994), Shao and Xi (2014), Wang (2014), Xi (2008, 2009), Xi and Zhao (2006), Yin and Xi (2010), Yin and Zhang (1998), Yin and Zhu (2010), Zhu et al. (2015) 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, risk theory, etc.

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 $A$ of (1.1). We prove that under very mild conditions, the martingale problem for the operator $A$ is well-posed. That is, we show that for any $(x, k) \in \mathbb{R}^d \times S$, there is exactly one martingale solution for the operator $A$ 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 $A$ of (1.1) reduces to $\hat{A}$ of (2.5). For such a special operator $\hat{A}$, under Assumption 1.2, we manipulate the Stroock-Varadhan piecing together method (refer to §6.1 of Stroock and Varadhan (1979)) to construct a martingale solution for the operator $\hat{A}$ with an arbitrary initial condition $(x, k) \in \mathbb{R}^d \times S$ 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 (2009), 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 \((X, \Lambda)\) 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 (2004), Hairer et al. (2011), Lindvall (2002), Lindvall and Rogers (1986), Priola and Wang (2006), Wang (2010) 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 \((X, \Lambda)\), 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 \( G_\alpha \) of the process \((X, \Lambda)\) and a result in Meyn and Tweedie (1993) to establish the Feller property for the process \((X, \Lambda)\); this is spelled out in Theorem 4.4.

Next we use a similar approach to establish the strong Feller property for the process \((X, \Lambda)\) in Section 5. More precisely, inspired by Priola and Wang (2006), 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 \( G_\alpha \) of the process \((X, \Lambda)\) and the aforementioned result in Meyn and Tweedie (1993) lead to the desired strong Feller property for the process \((X, \Lambda)\) 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 \( \tilde{\mathcal{A}} \) of (2.5) and Section 3 deals with the general case. Section 4 is devoted to proving the Feller property for the process \((X, \Lambda)\). 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 \( \sigma \)-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 \( \mathcal{F}_t = \mathcal{G}_t \cup \mathcal{N}_t \) for any \( t \geq 0 \) and that \( \mathcal{F} = \mathcal{G} \cup \mathcal{N} \). 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 \( C^2_b(\mathbb{R}^d \times \mathbb{S}) \) 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 $k \in S$, we can also define the martingale solution for the Lévy type operator $L_k$ 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 $L_k$ starting from $x$, if $\mathbb{P}_k^{(x)}(X(0) = x) = 1$ and for each function $f \in C_c^\infty(\mathbb{R}^d)$,

\[
M_t^{(k)}(f) := f(X(t)) - f(X(0)) - \int_0^t L_k f(X(s))ds
\]

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

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

**Assumption 1.2.** Suppose the following conditions hold:

(i) For each $k \in S$ and $x \in \mathbb{R}^d$, the Lévy type operator $L_k$ defined in (1.2) has a unique martingale solution $\mathbb{P}_k^{(x)}$ starting from $x$;

(ii) For each $k \in S$, the function $q_{kk}(x) \leq 0$ is bounded from below; and

(iii) \[
\sup_{(x,k)\in \mathbb{R}^d \times S} \int_{\mathbb{R}^d_0} (1 \wedge |y|^2) \nu(x, k, dy) < \infty.
\]

**Remark 1.3.** The martingale problem for the operator $L_k$ of (1.2) has been well-studied in the literature. For example, Komatsu (1973) and Stroock (1975) contain explicit sufficient conditions for the existence and uniqueness of martingale solutions for $L_k$.

We will prove in Section 3 that there exists a unique martingale solution for the operator $A$ 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 $\Gamma \in \mathcal{B}(\mathbb{R}^d_0)$ with $0 \notin \Gamma$, we let

\[
\eta(t, \Gamma) := \sum_{s \leq t} 1_{\Gamma}(\Delta X(s)) = \sum_{s \leq t} 1_{\Gamma}(X(s) - X(s-));
\]

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 $A$ starting from $(x, k) \in \mathbb{R}^d \times S$, then the following assertions are true:
(a) For each $f \in C^2_b(\mathbb{R}^d \times \mathbb{S})$ such that $f$ is uniformly positive,
\[
 f(X(t), \Lambda(t)) \exp\left\{ - \int_0^t A f(X(u), \Lambda(u)) \, du \right\}
\]
is a $\mathbb{P}$-martingale.

(b) For each $\theta \in \mathbb{R}^d$,
\[
 \exp\left\{ \, i \langle \theta, X(t) - X(0) - \int_0^t b(X(u), \Lambda(u)) \, du \rangle \right. \left. + \frac{1}{2} \int_0^t \langle \theta, a(X(u), \Lambda(u)) \theta \rangle \, du \right. \\
 \left. - \int_0^t \int_{\mathbb{R}^d} [e^{\langle \theta, y \rangle} - 1 - i \langle \theta, y \rangle \mathbf{1}_{B(0,\varepsilon_0)}(y)] \nu(X(u), \Lambda(u), dy) \, du \right\}
\]
is a $\mathbb{P}$-martingale, where $i := \sqrt{-1}$.

(c) Let $g$ be a bounded measurable function on $\mathbb{R}^d$ which vanishes in a neighborhood of the origin. Then for any $\theta \in \mathbb{R}^d$,
\[
 \exp\left\{ \, i \langle \theta, X(t) - X(0) - \int_0^t b(X(u), \Lambda(u)) \, du \rangle \right. \left. + \frac{1}{2} \int_0^t \langle \theta, a(X(u), \Lambda(u)) \theta \rangle \, du \right. \\
 \left. + \int_{\mathbb{R}^d} g(y) \eta(t, dy) - \int_0^t \int_{\mathbb{R}^d} [e^{\langle \theta, y \rangle + g(y)} - 1 - i \langle \theta, y \rangle \mathbf{1}_{B(0,\varepsilon_0)}(y)] \nu(X(u), \Lambda(u), dy) \, du \right\}
\]
is a $\mathbb{P}$-martingale.

(d) Define $\tilde{\eta}(t, \Gamma) := \eta(t, \Gamma) - \int_0^t \nu(X(u), \Lambda(u), \Gamma) du$. Then for each $\theta \in \mathbb{R}^d$ and any measurable function $g$ on $\mathbb{R}_0^d$ satisfying the condition $|g(y)|^2 \leq C(1 + |y|^2)$ for some positive constant $C$,
\[
 \exp\left\{ \, i \langle \theta, X(t) - X(0) - \int_0^t b(X(u), \Lambda(u)) \, du \rangle \right. \left. + \frac{1}{2} \int_0^t \langle \theta, a(X(u), \Lambda(u)) \theta \rangle \, du \right. \\
 \left. + \int_{\mathbb{R}^d} g(y) \tilde{\eta}(t, dy) \right. - \int_0^t \int_{\mathbb{R}^d} [e^{\langle \theta, y \rangle + g(y)} - 1 - i \langle \theta, y \rangle \mathbf{1}_{B(0,\varepsilon_0)}(y) - g(y)] \nu(X(u), \Lambda(u), dy) \, du \right\}
\]
is a $\mathbb{P}$-martingale. In particular, if $0 \notin \Gamma$, 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 Stroock and Varadhan (1979). For brevity, we shall omit the details here. □
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 $k, l \in S$ with $k \neq l$ and $\hat{q}_{kk} = -(n_0 - 1)$ for all $k \in S$. In other words, we have

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

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

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

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

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

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

We have the following lemma from Zheng and Zheng (1986):

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

Let $\Lambda$ be the coordinate process on $D([0, \infty), S)$ 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 $k \in S$, $Q^{(k)}\{\lim_{n \to \infty} \tau_n = +\infty\} = 1$. Moreover, we have $Q^{(k)}(\tau_1 \geq t) = \exp(-(n_0 - 1)t)$ for all $t \geq 0$ and

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

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

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

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

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

\[
\hat{M}_t^{(f)} := f(X(t), \Lambda(t)) - f(X(0), \Lambda(0)) - \int_0^t \hat{A}f(X(s), \Lambda(s))ds,
\]

where \(f \in C^\infty_c(\mathbb{R}^d \times \mathbb{S})\) and \((X, \Lambda)\) 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 \(\hat{P}(x,k)\) for the operator \(\hat{A}\) starting from \((x, k).\) Our construction of the desired probability measure \(\hat{P}(x,k)\) 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 \(\{P_k^{(x)} : k \in \mathbb{S}, x \in \mathbb{R}^d\}\) and \(\{Q(k) : k \in \mathbb{S}\},\) 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

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

and

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

In view of Lemma 2.4 of Shiga and Tanaka (1985), we know that \(\int_0^t \nu(\Lambda(s); A)ds\) is the compensator of the point process \(n(t, A);\) namely,

\[
\mu(t, A) := n(t, A) - \int_0^t \nu(\Lambda(s); A)ds
\]

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

\[
\hat{Q}f(k) = \sum_{l \in \mathbb{S}} \hat{q}_{kl}(f(l) - f(k)) = \int_{\mathbb{S}} (f(l) - f(k))\nu(k; dl).
\]

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 \(\hat{P}(x,k)\) on \(D([0, \infty), \mathbb{R}^d \times \mathbb{S})\) for the operator \(\hat{A}\) starting from \((x, k).\)

**Proof.** The proof is divided into two steps. The first step establishes the existence of a martingale solution \(\hat{P}\) for the operator \(\hat{A}\) 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 \((\Omega, \mathcal{F})\)
write as follows:

\[
\mathbb{P}^{(1)} = \mathbb{P}^{(x)}_k \times \mathbb{Q}^{(k)}, \quad \text{and for } n \geq 1, \quad \mathbb{P}^{(n+1)} = \mathbb{P}^{(n)} \otimes \tau_n \left( \mathbb{P}^{(X(\tau_n))} \times \mathbb{Q}^{(\Lambda(\tau_n))} \right),
\]

(2.10)

where \( \Omega = D([0, \infty), \mathbb{R}^d \times \mathbb{S}) \). Thanks to Theorem 6.1.2 of Stroock and Varadhan (1979), \( \mathbb{P}^{(n+1)} = \mathbb{P}^{(n)} \) on \( \mathcal{F}_{\tau_n} \).

Let \( f \in C^2_b(\mathbb{R}^d \times \mathbb{S}) \). 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 \( \mathbb{P}^{(x)}_k \) and hence \( \mathbb{P}^{(1)} \). On the other hand, using (2.9), we can write

\[
\int_0^{\tau_1 \wedge t} \hat{Q} f(X(s), \Lambda(s)) ds
\]

\[
= \int_0^{\tau_1 \wedge t} \int_0^{\tau_1 \wedge t} [f(X(s), l) - f(X(s), \Lambda(s))] \nu(\Lambda(s), dl) ds
\]

\[
= - \int_0^{\tau_1 \wedge t} \int_0^{\tau_1 \wedge t} [f(X(s), l) - f(X(s), \Lambda(s))] (n(ds, dl) - \nu(\Lambda(s), dl) ds)
\]

\[
+ \int_0^{\tau_1 \wedge t} \int_0^{\tau_1 \wedge t} [f(X(s), l) - f(X(s), \Lambda(s))] n(ds, dl)
\]

\[
= - \int_0^{\tau_1 \wedge t} \int_0^{\tau_1 \wedge t} [f(X(s), l) - f(X(s), \Lambda(s))] \mu(ds, dl)
\]

\[
+ f(X(\tau_1 \wedge t), \Lambda(\tau_1 \wedge t)) - f(X(\tau_1 \wedge t), \Lambda(\tau_1 \wedge t))
\]

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

\[
\hat{M}^{(f)}_{\tau_1 \wedge t} = f(X(\tau_1 \wedge t), \Lambda(\tau_1 \wedge t)) - f(X(0), \Lambda(0)) - \int_0^{\tau_1 \wedge t} \hat{A} f(X(s), \Lambda(s)) ds
\]

\[
= f(X(\tau_1 \wedge t), \Lambda(0)) - f(X(0), \Lambda(0)) - \int_0^{\tau_1 \wedge t} \mathcal{L}_{\Lambda(0)} f(X(s), \Lambda(0)) ds
\]

\[
+ f(X(\tau_1 \wedge t), \Lambda(\tau_1 \wedge t)) - f(X(\tau_1 \wedge t), \Lambda(0))
\]

\[
+ \int_0^{\tau_1 \wedge t} \mathcal{L}_{\Lambda(0)} f(X(s), \Lambda(0)) ds - \int_0^{\tau_1 \wedge t} \hat{A} f(X(s), \Lambda(s)) ds
\]

\[
= f(X(\tau_1 \wedge t), \Lambda(0)) - f(X(0), \Lambda(0)) - \int_0^{\tau_1 \wedge t} \mathcal{L}_{\Lambda(0)} f(X(s), \Lambda(0)) ds
\]

\[
+ f(X(\tau_1 \wedge t), \Lambda(\tau_1 \wedge t)) - f(X(\tau_1 \wedge t), \Lambda(0)) - \int_0^{\tau_1 \wedge t} \hat{Q} f(X(s), \Lambda(s)) ds
\]

\[
= f(X(\tau_1 \wedge t), \Lambda(0)) - f(X(0), \Lambda(0)) - \int_0^{\tau_1 \wedge t} \mathcal{L}_{\Lambda(0)} f(X(s), \Lambda(0)) ds
\]
Recall that $\mu$ is a martingale measure with respect to $Q^{(k)}$ and hence $P^{(1)}$. Thus it follows that $\hat{M}^{(f)}_{\tau_n\Lambda}$ is a martingale with respect to $P^{(1)}$.

Next,

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

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

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

is a martingale with respect to $P^{(X(\tau_1)) \times Q^{(\Lambda(\tau_1))}}$. Notice that the above displayed equation is equal to $\hat{M}^{(f)}_{\Lambda(\tau_1)} - \hat{M}^{(f)}_{\tau_2 \wedge t}$. Then in view of Theorem 6.1.2 of Stroock and Varadhan (1979), $\hat{M}^{(f)}_{\tau_n\Lambda}$ is a martingale with respect to $P^{(2)}$. In a similar fashion, we can show that $\hat{M}^{(f)}_{\tau_n\Lambda}$ is a martingale with respect to $P^{(n)}$ for any $n \geq 1$.

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

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

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

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

Recall that the probabilities $P^{(n)}$ constructed in (2.10) satisfies $P^{(n+1)} = P^{(n)}$ on $\mathcal{F}_{\tau_n}$. Hence by Tulcea’s extension theorem (see, e.g., (Stroock and Varadhan, 1979, Theorem 1.3.5)), there exists a unique $\hat{P}$ on $(\Omega, \mathcal{F})$ such that $\hat{P}$ equals $P^{(n)}$ on $\mathcal{F}_{\tau_n}$. Thus it follows that $\hat{M}^{(f)}_{\tau_n\Lambda}$ is a martingale with respect to $\hat{P}$ for every $n \geq 1$. In addition, for any $t \geq 0$, we have

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

Thus $\tau_n \to \infty$ a.s. $\hat{P}$ and hence $\hat{M}^{(f)}$ is a martingale with respect to $\hat{P}$. This establishes
that \( \hat{P} \) is the desired martingale solution staring 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 \( \hat{P}(x,k) \).

**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 \( \tilde{P}(x,k) \in \mathcal{P}(\Omega, \mathcal{F}) \) be another solution to the martingale problem associated with \( \hat{A} \) starting from \((x, k)\). We show that \( \tilde{P}(x,k) \) and \( \hat{P}(x,k) \) agree on \( \mathcal{F}_{\tau_1} \).

Recall that \( \hat{P}(x,k) \) agrees with \( P^{(x)} \times Q^{(k)} \) on \( \mathcal{F}_{\tau_1} \) and that \( P^{(x)} \in \mathcal{P}(D([0, \infty); \mathbb{R}^d)) \) 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 \( \delta_k \) contains all functions \( \omega \) in \( D([0, \infty), \mathcal{S}) \) satisfying \( \omega(t) = k \) for all \( 0 \leq t < \tau_1 \) and \( \omega(\tau_1) \in S \setminus \{k\} \). Since \( Q^{(k)}(\delta_k) = 1 \), it follows that

\[
\tilde{P}(x,k)(A) = P^{(x)}_k \times Q^{(k)}(A_1 \times \delta_k) = P^{(x)}_k(A_1).
\]

(2.12)

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

\[
\tilde{M}^{(g)} = g(X(t \wedge \tau_1)) - g(X(0)) - \int_0^{\tau_1 \wedge t} \hat{A}g(X(s))ds
\]

\[
= g(X(t \wedge \tau_1)) - g(X(0)) - \int_0^{\tau_1 \wedge t} \mathcal{L}_k g(X(s))ds
\]

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

\[
\tilde{P}(A_1) := \tilde{P}(x,k)(A_1 \times \delta_k).
\]

(2.13)

Then \( \tilde{M}^{(g)} \) is also a \( \tilde{P} \) martingale and hence \( \tilde{P} \) is a solutions to the martingale problem associated with \( \mathcal{L}_k \) starting from \( x \) up to \( \tau_1 \). Now by the uniqueness of the martingale solution to \( \mathcal{L}_k \) starting from \( x \), we conclude from (2.12) and (2.13) that \( \tilde{P}(x,k)(A) = \tilde{P}(x,k)(A) \) 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 \( \mathcal{F}_{\tau_1} \).

Now suppose that the martingale solution \( \tilde{P}(x,k) \) to \( \hat{A} \) starting from \((x, k)\) is uniquely determined on \( \mathcal{F}_{\tau_n} \). By virtue of Theorem 6.2.1 of Stroock and Varadhan (1979) (also Lemma 5.4.19 of Karatzas and Shreve (1991)), there is a \( \tilde{P}(x,k) \)-null set \( N \in \mathcal{F}_{\tau_n} \) such that

\[
\tilde{P}(X(\tau_n(\omega)), \Lambda(\tau_n(\omega))) := \delta(X(\tau_n(\omega)), \Lambda(\tau_n(\omega)), \omega) \otimes \tau_n(\omega) \tilde{P}_\omega
\]

solves the martingale problem for \( \hat{A} \) starting from \((X(\tau_n(\omega)), \Lambda(\tau_n(\omega))) \) whenever \( \omega \notin N \), where \( \tilde{P}_\omega \) is the regular conditional probability distribution of \( \tilde{P}(x,k) \) given \( \mathcal{F}_{\tau_n} \), whose existence follows from (Karatzas and Shreve, 1991, Theorem 5.3.18). By the argument in the
previous paragraph, \( \hat{P}(x, k) \) is uniquely determined on \( \mathcal{F}_{\tau_n+1} \). Note that by virtue of Theorem 6.1.2 of Stroock and Varadhan (1979),

\[
\hat{P}(x, k) = \hat{P}(x, k) \otimes_{\tau_n(\cdot)} \hat{P}(X(\tau_n(\cdot)), A(\tau_n(\cdot))) ,
\]

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

(i) \( \hat{P}(x, k) \otimes_{\tau_n(\cdot)} \hat{P}(X(\tau_n(\cdot)), A(\tau_n(\cdot))) (A) = \hat{P}(x, k)(A) \), for any \( A \in \mathcal{F}_{\tau_n} \), and

(ii) \( \delta_{\{\tau_n(\omega), A(\tau_n(\omega)), \omega\}} \otimes_{\tau_n(\omega)} \hat{P}_\omega \) is a regular conditional probability distribution of \( \hat{P}(x, k) \otimes_{\tau_n(\cdot)} \hat{P}(X(\tau_n(\cdot)), A(\tau_n(\cdot))) \) given \( \mathcal{F}_{\tau_n} \).

Thus by the induction hypothesis, we conclude that \( \hat{P}(x, k) \) 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 \( \mathbb{P}_n(A) := \hat{P}(x, k)(A) \). Apparently \( \mathbb{P}_n \) satisfies that \( \mathbb{P}_n = \mathbb{P}_{n+1} \) on \( \mathcal{F}_{\tau_n} \) and that for any \( t \geq 0 \), \( \mathbb{P}_n\{\tau_n \leq t\} = \hat{P}(x, k)\{\tau_n \leq t\} \rightarrow 0 \) as \( n \rightarrow \infty \), where we used (2.11). Therefore by Tulcea’s extension theorem (e.g., (Stroock and Varadhan, 1979, Theorem 1.3.5)), the sequence \( \mathbb{P}_n \) has a unique extension \( \hat{\mathbb{P}} \) on \( (\Omega, \mathcal{F}) \) such that \( \hat{\mathbb{P}} = \mathbb{P}_n \) on \( \mathcal{F}_{\tau_n} \). 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:

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

where

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

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 \( (M_t, \mathcal{F}_t, \hat{\mathbb{P}}) \) 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

\[
\prod_{i=0}^{n(t)-1} q_{\Lambda(\tau_i)A(\tau_{i+1})}(X(\tau_{i+1})) = \exp \left\{ \sum_{i=0}^{n(t)-1} \log q_{\Lambda(\tau_i)A(\tau_{i+1})}(X(\tau_{i+1})) \right\}
\]
\[
\begin{align*}
&= \exp\left\{ \int_{[0,t] \times \mathbb{S}} \log q_{A(s-)}(X(s)) n(ds, dl) \right\}, \\
\end{align*}
\]
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
\[
Z(t) := \int_{[0,t] \times \mathbb{S}} \log q_{A(s-)}(X(s)) n(ds, dl) - \int_0^t [q_{A(s)}(X(s)) - n_0 + 1] ds,
\]
Now we apply Itô’s formula for jump processes (see, e.g., (Ikeda and Watanabe, 1989, Theorem II.5.1)) to the process \( M_t \):
\[
M_t(X(\cdot), \Lambda(\cdot)) - 1 = e^{Z(t)} - e^{Z(0)}
= \int_0^t \int_{\mathbb{S}} e^{Z(s)} [q_{A(s-)}(X(s)) - 1] n(ds, dl) - \int_0^t e^{Z(s)} [q_{A(s)}(X(s)) - n_0 + 1] ds. \quad (3.3)
\]
Recall from Section 2 that for any \( s \geq 0 \), \( \hat{\mathbb{P}}\{\Lambda(s) = l, \Lambda(s) \neq \Lambda(s-)\} = \frac{1}{n_0 - 1} \). Thus we have
\[
\hat{\mathbb{E}}[n(t, A)] = \mathbb{E}^\hat{\mathbb{P}} \left[ \sum_{s \leq t} 1_{\{\Lambda(s) \in A, \Lambda(s) \neq \Lambda(s-)\}} \right] = \mathbb{E}^\hat{\mathbb{P}} \left[ \sum_{k \in \mathbb{S}} \sum_{s \leq t} 1_{\{\Lambda(s) \in A, \Lambda(s) \neq \Lambda(s-) \Lambda(s-) = k\}} \right]
= (n_0 - 1) \int_0^t \int_{\mathbb{S}} \frac{1}{n_0 - 1} dl ds = \int_0^t \int_A dl ds,
\]
where \( dl \) is the counting measure on \( \mathbb{S} \). Then it follows that
\[
\int_0^t e^{Z(s)} [q_{A(s)}(X(s)) - n_0 + 1] ds = \int_0^t e^{Z(s)} \sum_{l \neq \Lambda(s-)} [q_{A(s-)(X(s)) - 1} ds
= \int_0^t \int_{\mathbb{S}} e^{Z(s)} [q_{A(s-)}(X(s)) - 1] dl ds.
\]
Putting these observations into (3.3) and using (3.2), we obtain
\[
M_t(X(\cdot), \Lambda(\cdot)) - 1 = \int_{[0,t] \times \mathbb{S}} M_{s-}(X(\cdot), \Lambda(\cdot)) [q_{A(s-)}(X(s)) - 1] \tilde{n}(ds, dl), \quad (3.4)
\]
where \( \tilde{n}(t, A) = n(t, A) - \hat{\mathbb{E}}[n(t, A)] \) is the compensated Poisson random measure with respect to \( \hat{\mathbb{P}} \) 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

\[
q_{kl}^\varepsilon(x) \to q_{kl}(x), \quad \text{and} \quad q_{kk}^\varepsilon(x) \to q_{kk}(x)
\]

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

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

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

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

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

\[
M_t(X(\cdot), \Lambda(\cdot)) - 1 = \int_{[0,t] \times \mathbb{S}} M_{s-}(X(\cdot), \Lambda(\cdot)) \left[ q_{\Lambda(s-)}(X(s)) - 1 \right] \bar{n}(ds, dl). \tag{3.5}
\]

Step 3. From (3.5), we can see that \( M_t(X(\cdot), \Lambda(\cdot)) \) is a martingale with mean 1 under \( \hat{P} \). 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 \( \hat{P} \).

Proof. The proof is similar to that of (Xi, 2009, Lemma 4.4) and we shall omit the details here.\( \square \)

Let \( \varepsilon > 0 \) and notice that in view of (1.3), \( \nu(x, k, \mathbb{R}^d \setminus B(0, \varepsilon)) < \infty \) 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 \( n \geq 0 \),

\[
\zeta_n^{(\varepsilon)} := \inf\{ t \geq \zeta_{n-1}^{(\varepsilon)} : |\Delta X(t)| = |X(t) - X(t^-)| \geq \varepsilon \}. \tag{3.6}
\]

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

\[
\hat{P}\{\tau_1 > t\} = \exp\{-(n_0 - 1)t\}, \tag{3.7}
\]

\[
\hat{P}\{\zeta_1^{(\varepsilon)} > t | \mathcal{F}_{\zeta_1^{(\varepsilon)}}\} = \exp\left\{ -\int_0^t \nu(X_\varepsilon(u \wedge \zeta_1^{(\varepsilon)}), \Lambda(u \wedge \zeta_1^{(\varepsilon)}), \mathbb{R}^d \setminus B(0, \varepsilon)) du \right\}. \tag{3.8}
\]

14
Proof. Equation (3.7) follows directly from the construction of \( \hat{\mathbb{P}} \) 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

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

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

\[
\mathbb{E}^{\hat{\mathbb{P}}}[\eta(t \wedge \zeta_1^{(e)}; \Gamma) | \mathcal{F}_{\zeta_1^{(e)}_-}] = \mathbb{E}^{\hat{\mathbb{P}}} \left[ \int_0^{t \wedge \zeta_1^{(e)}} \nu(X(u), \Lambda(u), \Gamma) du | \mathcal{F}_{\zeta_1^{(e)}_-} \right]
\]

\[
= \mathbb{E}^{\hat{\mathbb{P}}} \left[ \int_0^t 1_{\{ \zeta_1^{(e)} > u \}} \nu(X(u \wedge \zeta_1^{(e)}), \Lambda(u \wedge \zeta_1^{(e)}), \Gamma) du | \mathcal{F}_{\zeta_1^{(e)}_-} \right]
\]

\[
= \int_0^t \mathbb{P}^{\hat{\mathbb{P}}} \{ \zeta_1^{(e)} > u | \mathcal{F}_{\zeta_1^{(e)}_-} \} \nu(X(s \wedge \zeta_1^{(e)}), \Lambda(s \wedge \zeta_1^{(e)}), \Gamma) ds. \tag{3.9}
\]

On the other hand, note that

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

Thus we have

\[
\mathbb{P}^{\hat{\mathbb{P}}} \{ \zeta_1^{(e)} > t | \mathcal{F}_{\zeta_1^{(e)}_-} \} = \mathbb{E}^{\hat{\mathbb{P}}} \left[ 1_{\{ \zeta_1^{(e)} > t \}} | \mathcal{F}_{\zeta_1^{(e)}_-} \right]
\]

\[
= \mathbb{E}^{\hat{\mathbb{P}}} \left[ (1 - \eta(t \wedge \zeta_1^{(e)}; \Gamma)) | \mathcal{F}_{\zeta_1^{(e)}_-} \right] = 1 - \mathbb{E}^{\hat{\mathbb{P}}} \left[ \eta(t \wedge \zeta_1^{(e)}; \Gamma) | \mathcal{F}_{\zeta_1^{(e)}_-} \right]. \tag{3.10}
\]

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

\[
1 - \mathbb{P}^{\hat{\mathbb{P}}} \{ \zeta_1^{(e)} > t | \mathcal{F}_{\zeta_1^{(e)}_-} \} = \int_0^t \mathbb{P}^{\hat{\mathbb{P}}} \{ \zeta_1^{(e)} > s | \mathcal{F}_{\zeta_1^{(e)}_-} \} \nu(X(s \wedge \zeta_1^{(e)}), \Lambda(s \wedge \zeta_1^{(e)}), \Gamma) ds. \tag{3.11}
\]

Let us denote \( u(t) := \mathbb{P}^{\hat{\mathbb{P}}} \{ \zeta_1^{(e)} > t | \mathcal{F}_{\zeta_1^{(e)}_-} \} \) and \( v(t) := \nu(X_u(t \wedge \zeta_1^{(e)}), \Lambda(t \wedge \zeta_1^{(e)}), \Gamma) \). Then we can rewrite (3.11) as

\[
u_u(t) + \int_0^t u(s)v(s)ds = 1,
\]

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

\[
\int_0^t u(r)v(r)dr = \int_0^t e^{\int_0^s v(r)dr} v(s)ds = e^{\int_0^t v(r)dr} - 1.
\]
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 \(\hat{\mathbb{P}}\), \(\{\zeta_n^{(\varepsilon)}, n \geq 1\}\) 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 \(m = 0, 1, \ldots, M - 1\), we denote

\[ J^{(\varepsilon)}((m - 1)/M, m/M] := \max\{n \in \mathbb{N} : \zeta_n^{(\varepsilon)} \leq m/M\} - \max\{n \in \mathbb{N} : \zeta_n^{(\varepsilon)} \leq (m - 1)/M\}, \]

\[ S((m - 1)/M, m/M] := \max\{n \in \mathbb{N} : \tau_n \leq m/M\} - \max\{n \in \mathbb{N} : \tau_n \leq (m - 1)/M\}, \]

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 \(K_{\varepsilon}\) 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

\[ \hat{\mathbb{P}} \left\{ \{\zeta_n^{(\varepsilon)} : n \geq 1, \zeta_n^{(\varepsilon)} \leq T\} \cap \{\tau_n : n \geq 1, \tau_n \leq T\} \neq \emptyset \right\} \]

\[ \leq \hat{\mathbb{P}} \left\{ \text{there are one jump and one switch in the interval } \left(\frac{m - 1}{M}, \frac{m}{M}\right) \text{ for some } m \right\} \]

\[ \leq \sum_{m=0}^{M-1} \hat{\mathbb{P}} \left\{ A_m \cap B_m \right\} = \sum_{m=0}^{M-1} \hat{\mathbb{P}}(A_m)\hat{\mathbb{P}}(B_m|A_m) \]

\[ \leq \sum_{m=0}^{M-1} \left(1 - \exp\left\{-\int_{m-1}^m \nu(X_\varepsilon(s), \Lambda(s), \mathbb{R}^d \setminus B(0, \varepsilon))\,ds\right\}\right) \left(1 - \exp\left\{-(n_0 - 1)\frac{1}{M}\right\}\right) \]

\[ \leq \sum_{m=0}^{M-1} \left(1 - \exp\left\{-K_{\varepsilon}\frac{1}{M}\right\}\right) \left(1 - \exp\left\{-(n_0 - 1)\frac{1}{M}\right\}\right) \]

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

\[ \hat{\mathbb{P}} \left\{ \{\zeta_n^{(\varepsilon)} : n \geq 1, \zeta_n^{(\varepsilon)} \leq T\} \cap \{\tau_n : n \geq 1, \tau_n \leq T\} \neq \emptyset \right\} \leq \sum_{m=0}^{M-1} \frac{n_0 - 1}{M} K_{\varepsilon} \frac{1}{M} = \frac{(n_0 - 1)K_{\varepsilon}}{M}, \]

16
which can be arbitrarily small since the denominator $M$ is arbitrary. This implies the desired conclusion and hence completes the proof.

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. Rudin (1976). 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 $\hat{P}$, $\{\zeta_n : n \geq 1\}$ and $\{\tau_n : n \geq 1\}$ are mutually disjoint with probability 1.

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

$$\{\zeta_n : n \geq 1\} \cap \{\tau_n : n \geq 1\} = \bigcup_{m=1}^{\infty} \{\zeta^{(1/m)}_n : n \geq 1\} \cap \{\tau_n : n \geq 1\}.$$ Moreover, for each $m = 1, 2, \ldots$, since (3.12) holds with $\varepsilon = \frac{1}{m}$, Lemma 3.4 implies that

$$\hat{P}\{\{\zeta^{(1/m)}_n : n \geq 1\} \cap \{\tau_n : n \geq 1\} \neq \emptyset\} = 0.$$ Therefore we deduce

$$\hat{P}\{\{\zeta_n : n \geq 1\} \cap \{\tau_n : n \geq 1\} \neq \emptyset\} \leq \sum_{m=1}^{\infty} \hat{P}\{\{\zeta^{(1/m)}_n : n \geq 1\} \cap \{\tau_n : n \geq 1\} \neq \emptyset\} = 0.$$ This completes the proof.

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

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

**Proof.** For each $t \geq 0$ and each $A \in F_t$, define

$$P^{(x,k)}_t(A) = \int_A M_t(X(\cdot), A(\cdot)) \, d\hat{P}^{(x,k)}.$$ Thanks to Lemma 3.1, the family of probability measures $\{P^{(x,k)}_t\}_{t \geq 0}$ is consistent in the sense that if $0 \leq t_1 \leq t_2$ and $A \in F_{t_1}$, then $P^{(x,k)}_{t_2}(A) = P^{(x,k)}_{t_1}(A)$. Thus by Tulcea’s extension theorem (see, e.g., (Stroock and Varadhan, 1979, Theorem 1.3.5)), there exists a unique probability measure $P^{(x,k)}$ on $(\Omega, F)$ which coincides with $P^{(x,k)}_t$ on $F_t$ for all $t \geq 0$. Moreover, we will prove that the $P$ is the desired martingale solution for the operator $A$ starting from $(x, k)$. To do so, analogously to the proof of Lemma 4.2 in Xi (2009), we first
prove that for each function $f \in C_c^\infty(\mathbb{R}^d \times S)$, $(M_t, \mathcal{M}_t, \mathcal{F}_t, \mathcal{P})$ is a martingale, where $M_t^{(f)}$ is defined in (1.5). In fact, using integration by parts, we derive that

$$M_t M_t^{(f)} = \int_0^t M_s^{(f)} dM_s + \int_0^t M_s \, d\mathcal{M}_s^{(f)}$$

$$+ \int_0^t M_s^{-}(dM_s^{(f)} - d\mathcal{M}_s^{(f)}) + \sum_{s \leq t} (M_s - M_{s-}) (M_s^{(f)} - M_{s-}^{(f)}),$$

(3.15)

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

$$\sum_{s \leq t} (M_s - M_{s-}) (M_s^{(f)} - M_{s-}^{(f)})$$

$$= \sum_{s \leq t} (M_s - M_{s-}) [f(X(s), \Lambda(s)) - f(X(s), \Lambda(s-))]$$

$$= \int_{[0,t] \times S} M_{s-} \left( \frac{M_s}{M_{s-}} - 1 \right) [f(X(s), l) - f(X(s), \Lambda(s-))] n(ds, dl)$$

$$= \int_{[0,t] \times S} M_{s-} (q_{\Lambda(s-)}(X(s)) - 1) [f(X(s), l) - f(X(s), \Lambda(s-))] n(ds, dl).$$

On the other hand,

$$\int_0^t M_{s-} (dM_s^{(f)} - d\mathcal{M}_s^{(f)})$$

$$= - \int_0^t M_{s-} \sum_{l \in S} (q_{\Lambda(s-)}(X(s)) - 1) [f(X(s), l) - f(X(s), \Lambda(s-))] ds.$$

Combining the last two displayed equations, and using the observations concerning the martingale measure $\mathcal{M}(\cdot, \cdot)$ in the proof of Lemma 3.1, we obtain

$$\int_0^t M_{s-} (dM_s^{(f)} - d\mathcal{M}_s^{(f)}) + \sum_{s \leq t} (M_s - M_{s-}) (M_s^{(f)} - M_{s-}^{(f)})$$

$$= \int_{[0,t] \times S} M_{s-} (q_{\Lambda(s-)}(X(s)) - 1) [f(X(s), l) - f(X(s), \Lambda(s-))] \mathcal{M}(ds, dl).$$

Then upon plugging the above equation into (3.15), it follows that

$$M_t M_t^{(f)} = \int_0^t M_s^{(f)} dM_s + \int_0^t M_s \, d\mathcal{M}_s^{(f)}$$

$$+ \int_{[0,t] \times S} M_{s-} (q_{\Lambda(s-)}(X(s)) - 1) [f(X(s), l) - f(X(s), \Lambda(s-))] \mathcal{M}(ds, dl).$$

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

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

\[
\int_A M_t^{(f)} d\hat{P}^{(x,k)} = \int_A M_t M_t^{(f)} d\hat{P}^{(x,k)} = \int_A M_s M_s^{(f)} d\hat{P}^{(x,k)} = \int_A M_s^{(f)} d\hat{P}^{(x,k)},
\]

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

It remains to show that any martingale solution \( \hat{P} \) for the operator \( \mathcal{A} \) starting from \((x, k)\) must agree with \( \hat{P}^{(x,k)} \) and therefore establishing the desired uniqueness. From Wang (2014), for any martingale solution \( \hat{P}^{(x,k)} \) to the operator \( \mathcal{A} \), we have

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

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 \((x, k) \in \mathbb{R}^d \times S\) is well-posed. Thus the process \((X, \Lambda)\) 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 \( \Omega = D([0, \infty), \mathbb{R}^d \times S) \) under which the coordinate process \((X(t), \Lambda(t))\) satisfies \( \mathbb{P}\{(X(0), \Lambda(0)) = (x, k)\} = 1 \) and that for any \( f \in C^c_c(\mathbb{R}^d \times S) \), 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 \( \rho : [0, \infty) \rightarrow [0, \infty) \) satisfying \( \rho(r) > 0 \) for \( r > 0 \) and

\[
\int_{0^+} \frac{dr}{\rho(r)} = \infty, \tag{4.1}
\]
such that for all $k \in S$ and $x, z \in \mathbb{R}^d$,
\[
\|\sigma(x, k) - \sigma(z, k)\|^2 + 2 \langle x - z, b(x, k) - b(z, k) \rangle \leq H|x - z|\rho(|x - z|),
\] (4.2)
and
\[
\int_{\mathbb{R}_0^d} |u| \|\nu(x, k, \cdot) - \nu(z, k, \cdot)\|(du) \leq H\rho(|x - z|),
\] (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
\[
|q_{kl}(x) - q_{kl}(z)| \leq H|x - z|
\] (4.4)
for all $x, z \in \mathbb{R}^d$ and $k \neq l \in S$, 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 (1979) for the 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 $(X, \Lambda)$ 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 $L_k$ 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 $(X, \Lambda)$ can be represented by a series of the resolvents of the killed processes $X^{(k)}$, $k \in S$. This representation further helps us to establish (4.28). Finally we use (4.28) together with Proposition 6.1.1 in Meyn and Tweedie (1993) to derive the Feller property for the process $(X, \Lambda)$.

**Remark 4.5.** The recent paper Wang (2010) 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 Wang (2010). By contrast, our goal is to establish the Feller property for the two-component process $(X, \Lambda)$ 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 below 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 Wang (2010).

Recall that for each $k \in S$ and $x \in \mathbb{R}^d$, Assumption 1.2 guarantees that the operator $L_k$ 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 $(-q_{kk})$:

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

(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 $\partial$ 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 $k \in S$, 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 $L_k$. 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 $h(x,z) \in C^2_0(\mathbb{R}^d \times \mathbb{R}^d)$, set

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

(4.6)

which is a coupling of the diffusion part in the generator $L_k$ defined in (1.2) (refer to Chen and Li (1989)). Next, for $h(x,z) \in C^2_0(\mathbb{R}^d \times \mathbb{R}^d)$, set

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

(4.7)

where $(\nu(x,k,\cdot) - \nu(z,k,\cdot))^+ = \sup\{\nu(x,k,A) - \nu(z,k,A) : A \in \mathcal{B}(\mathbb{R}^d)\}$ and $(\nu(z,k,\cdot) - \nu(x,k,\cdot))^+$.
\(\nu(x, k, \cdot)\) 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:

\[
\tilde{\mathcal{L}}_k h(x, z) = \tilde{\Omega}_d(k) h(x, z) + \tilde{\Omega}_j(k) h(x, z),
\]

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

\[
W(P_1, P_2) = \inf \int |x - z| \tilde{P}(dx, dz),
\]

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 \(x \to z\). By virtue of Theorem 5.6 in Chen (2004), it suffices to prove that

\[
W(\tilde{P}^{(k)}(t, x, \cdot), \tilde{P}^{(k)}(t, z, \cdot)) \to 0 \quad \text{as} \quad x \to z. \tag{4.9}
\]

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 Chen and Li (1989), 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 \(\rho\), we can find a strictly decreasing sequence \(\{a_n\} \subset (0, 1]\) with \(a_0 = 1, \lim_{n \to \infty} a_n = 0\) 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 \(\rho_n\) 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

\[
\psi_n(r) := \int_0^{|r|} \int_0^y \rho_n(u) du dy, \quad r \in \mathbb{R}, n \geq 1. \tag{4.10}
\]
We can immediately verify that $\psi_n$ is even and continuously differentiable, with $|\psi_n'(r)| \leq 1$ and $\lim_{n \to \infty} \psi_n(r) = |r|$ 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 $n \in \mathcal{N}$, $\psi_n$, $\psi_n'$ and $\psi_n''$ all vanish on the interval $(-a_n, a_n)$.

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

$$A(x, z, k) = a(x, k) + a(z, k) - 2\sigma(x, k)\sigma(z, k)^T,$$

$$\hat{B}(x, z, k) = \langle x - z, b(x, k) - b(z, k) \rangle,$$

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 Chen and Li (1989), we can verify directly that

$$2\tilde{\Omega}_d(k)\psi_n(|x - z|) = \psi_n''(|x - z|)|\overline{A}(x, z, k)|$$

$$+ \frac{\psi_n'(|x - z|)}{|x - z|} \left[ \text{tr}(A(x, z, k)) - \overline{A}(x, z, k) + 2\hat{B}(x, z, k) \right].$$

(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

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

On the other hand, using (4.2) again,

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

Thanks to the construction of $\psi_n$, 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

$$\tilde{\Omega}_d(k)\psi_n(|x - z|) \leq \frac{1}{2}\psi_n''(|x - z|)H|x - z|\rho(|x - z|) + \frac{1}{2}\psi_n'(|x - z|)H\rho(|x - z|)$$

$$\leq \frac{H}{n}|x - z|I_{(a_n,a_{n-1})}(|x - z|) + \frac{1}{2}H\rho(|x - z|)$$

$$\leq \frac{Ha_{n-1}}{n} + \frac{1}{2}H\rho(|x - z|).$$

(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
\[
\int \left( \psi_n(|x + u - z|) - \psi_n(|x - z|) - \langle \nabla x \psi_n(|x - z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u) \right) \left( \nu(x, k, du) - \nu(z, k, du) \right)^+ \\
\leq 2 \int |u| \left( \nu(x, k, du) - \nu(z, k, du) \right)^+.
\]

Similarly, we have
\[
\int \left( \psi_n(|x - z - u|) - \psi_n(|x - z|) - \langle \nabla z \psi_n(|x - z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u) \right) \left( \nu(x, k, du) - \nu(z, k, du) \right)^+ \\
\leq 2 \int |u| \left( \nu(x, k, du) - \nu(z, k, du) \right)^+.
\]

Note that \( \nabla x \psi_n(|x - z|) = -\nabla z \psi_n(|x - z|) \). Thus
\[
\int \left[ \psi_n(|x + u - z - u|) - \psi_n(|x - z|) - \langle \nabla x \psi_n(|x - z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u) \right] \\
- \langle \nabla z \psi_n(|x - z|), u \rangle \mathbf{1}_{B(0, \varepsilon_0)}(u) \right) \left( \nu(x, k, du) \wedge \nu(z, k, du) \right) = 0.
\]

Then, using the definition of \( \tilde{\Omega}_j(k) \) in (4.7), we obtain
\[
\tilde{\Omega}_j(k) \psi_n(|x - z|) \\
\leq 2 \int |u| \left( \nu(x, k, du) - \nu(z, k, du) \right)^+ + 2 \int |u| \left( \nu(z, k, du) - \nu(x, k, du) \right)^+ \\
\leq 2 \int |u| \left( \nu(x, k, \cdot) - \nu(z, k, \cdot) \right)(du) \leq 2H \rho(|x - z|),
\]

where the last inequality follows from (4.3).

A combination of (4.12) and (4.13) yields
\[
\tilde{\mathcal{L}}_k \psi_n(|x - z|) = \tilde{\Omega}_d(k) \psi_n(|x - z|) + \tilde{\Omega}_j(k) \psi(|x - z|) \leq \frac{H \sigma_{n-1}}{n} + 3H \rho(|x - z|), \forall x, z \in \mathbb{R}^d.
\]

Now we apply Itô’s formula to the process \( \psi_n(|\tilde{X}^{(k)}(\cdot) - \tilde{Z}^{(k)}(\cdot)|) \) to obtain
\[
\mathbb{E}_k \left[ \psi_n(|\tilde{X}^{(k)}(t \wedge T_R) - \tilde{Z}^{(k)}(t \wedge T_R)|) \right] \\
= \psi_n(|x - z|) + \mathbb{E}_k \left[ \int_0^{t \wedge T_R} \tilde{\mathcal{L}}_k \psi_n(|\tilde{X}^{(k)}(s) - \tilde{Z}^{(k)}(s)|) ds \right] \\
\leq \psi_n(|x - z|) + \frac{H \sigma_{n-1} t}{n} + 3H \mathbb{E}_k \left[ \int_0^{t \wedge T_R} \rho(|\tilde{X}^{(k)}(s) - \tilde{Z}^{(k)}(s)|) ds \right].
\]

24
Recall that $\psi_n(|x|) \uparrow |x|$ and $a_n \to 0$ as $n \to \infty$. Therefore, passing to the limit as $n \to \infty$ on both sides of (4.14), it follows from the monotone convergence theorem that

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

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

$$
\mathbb{E}_k[|\tilde{X}^{(k)}(x)(t) - \tilde{Z}^{(k)}(z)(t)|] \leq |x - z| + 3HE_k\left[\int_0^t \rho(|\tilde{X}^{(k)}(x)(s) - \tilde{Z}^{(k)}(z)(s)|)ds\right]
$$

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 $G(r) := \int_0^r \frac{ds}{\rho(s)}$ for $r > 0$. Then $G$ is nondecreasing and satisfies $\lim_{r \uparrow 0} G(r) = -\infty$ thanks to (4.1). In addition, we have

$$
G(u(t)) \leq G(v(t)) = G(|x - z|) + \int_0^t G'(v(s))v'(s)ds
$$

$$
= G(|x - z|) + 3H \int_0^t \frac{\rho(u(s))}{\rho(v(s))}ds \leq G(|x - z|) + 3Ht,
$$

where the last inequality follows from the assumption that $\rho$ 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 $\lim_{r \to -\infty} G^{-1}(r) = 0$. Furthermore, we have

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

In particular, when $|x - z| \to 0$, 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. \hfill \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
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 
\((4.18.I) \to 0\) as \(|x - z| \to 0\). Moreover, in view of (4.16), we have that \(\tilde{X}^{(k)}(t)\) converges to \(\tilde{Z}^{(k)}(t)\) in probability \(\mathbb{P}_k\) as \(|x - z| \to 0\). Thus, from the continuity of \(f\), we obtain that 
\(f(\tilde{X}^{(k)}(t))\) also converges to \(f(\tilde{Z}^{(k)}(t))\) in probability \(\mathbb{P}_k\) as \(|x - z| \to 0\). Combining this 
with \(q_{kk} \leq 0\), we derive that 

\[
(4.18.II) \leq 2\|f\|\mathbb{P}_k\left(\left| f(\tilde{X}^{(k)}(t)) - f(\tilde{Z}^{(k)}(t)) \right| \geq \varepsilon \right) \to 0 \quad (4.20)
\]
as \(|x - z| \to 0\). Using the fact that \(q_{kk} \leq 0\) again, we see that \((4.18.III)\) does not exceed 
\(\varepsilon\); 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 \(|x - z| \to 0\). The proof 
is complete. \(\square\)

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

\[
\mathbb{E}[f(Z^{(x)}(t))] = \mathbb{E}[f(\tilde{Z}^{(x)}(t)); t < \zeta] = \mathbb{E}\left[\exp\left\{-\int_{0}^{t} q(\tilde{Z}^{(x)}(s))ds\right\} f(\tilde{Z}^{(x)}(t))\right],
\]

where \(\tilde{Z}\) is a right continuous strong Markov process, \(q \geq 0\) on \(\mathbb{R}^d\), and \(f \in \mathcal{B}_0(\mathbb{R}^d)\). Then 
for any nonnegative function \(\phi\) on \(\mathbb{R}^d\) and constant \(\alpha > 0\), we have 

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

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 

\[
G_{\alpha}^Z(q\phi)(x) = \mathbb{E}\left[\int_{0}^{\infty} e^{-\alpha t}(q\phi)(Z^{(x)}(t))dt\right] = \mathbb{E}\left[\int_{0}^{\infty} e^{-\alpha t}(q\phi)(\tilde{Z}^{(x)}(t)) \exp\left\{-\int_{0}^{t} q(\tilde{Z}^{(x)}(s))ds\right\} dt\right],
\]

which by page 286 in Sharpe (1988) (putting \(m_t = \exp\{-\int_{0}^{t} q(\tilde{Z}(s))ds\}\) \(1_{(t<\zeta)}\) there) equals 
the left-hand side in (4.22). \(\square\)

For each \(k \in S\), 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 

\[
G_{\alpha}^0 = \begin{pmatrix}
G_{\alpha}^{(1)} & 0 & \cdots & 0 \\
0 & G_{\alpha}^{(2)} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & G_{\alpha}^{(n_0)}
\end{pmatrix}
\]

and 

\[
Q^0(x) = Q(x) - \begin{pmatrix}
q_{11}(x) & 0 & \cdots & 0 \\
0 & q_{22}(x) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & q_{n_0 n_0}(x)
\end{pmatrix}.
\]
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 $k \in \mathbb{S}$,

$$G_\alpha f = G_\alpha^0 f + \sum_{m=1}^{\infty} G_\alpha^0 (Q^\alpha_0 G^\alpha_0)^m f.$$ (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 $\tau$ and recalling the construction of $(X, \Lambda)$, we obtain

$$G_\alpha f(x, k) = \mathbb{E}_{x,k} \left[ \int_0^\infty e^{-\alpha t} f(X(t), \Lambda(t)) dt \right]$$

$$= \mathbb{E}_{x,k} \left[ \int_0^\tau e^{-\alpha t} f(X(t), k) dt \right] + \mathbb{E}_{x,k} \left[ \int_\tau^\infty e^{-\alpha t} f(X(t), \Lambda(t)) dt \right]$$

$$= G_\alpha^{(k)} f(x, k) + \mathbb{E}_{x,k} \left[ e^{-\alpha \tau} G_\alpha f(X(\tau), \Lambda(\tau)) \right]$$

$$= G_\alpha^{(k)} f(x, k) + \sum_{l \in \mathbb{S} \setminus \{k\}} \mathbb{E}_{x,k} \left[ e^{-\alpha \tau} \left( -\frac{q_{kl}}{q_{kk}} \right) (X(\tau) -) G_\alpha f(X(\tau), l) \right]$$

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

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

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

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

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

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

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

where

$$\psi_0^{(k)} = G_\alpha^{(k)} f(\cdot, k),$$

$$\psi_1^{(k)} = G_\alpha^{(k)} \left( \sum_{l \in \mathbb{S} \setminus \{k\}} q_{kl} G_\alpha^{(l)} f(\cdot, l) \right) = G_\alpha^{(k)} \left( \sum_{l \in \mathbb{S} \setminus \{k\}} q_{kl} \psi_0^{(l)} \right).$$
and for $i \geq 1$,
\[
\psi_i^{(k)} = G_\alpha^{(k)} \left( \sum_{l \in S \setminus \{k\}} q_{kl} \psi_{i-1}^{(l)} \right).
\]

By Assumption 1.2 we know that $+\infty > H := \max\{\|q_{kk}\| : k \in S\} \geq \max\{\|q_{kl}\| : k \neq l \in S\}$, 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,
\[
\|\psi_1^{(k)}\| \leq \sum_{l \in S \setminus \{k\}} \|G_\alpha^{(k)}(q_{kl} \psi_0^{(l)})\| \leq \frac{H}{\alpha} \sum_{l \in S \setminus \{k\}} \|\psi_0^{(l)}\|.
\]
Thus, we get that
\[
\sum_{k \in S} \|\psi_1^{(k)}\| \leq \frac{(n_0 - 1)H}{\alpha} \sum_{k \in S} \|\psi_0^{(k)}\| \leq \frac{1}{2} \sum_{k \in S} \|\psi_0^{(k)}\|
\]
when $\alpha \geq \alpha_1 := 2(n_0 - 1)H$. A similar argument yields that for $i \geq 1$,
\[
\sum_{k \in S} \|\psi_i^{(k)}\| \leq \frac{1}{2} \sum_{k \in S} \|\psi_{i-1}^{(k)}\| \leq \frac{1}{2} \sum_{k \in S} \|\psi_0^{(k)}\| \quad (4.26)
\]
and
\[
\|R_m^{(k)}(\cdot)\| \leq \frac{1}{2m} \sum_{k \in S} \|G_\alpha f(\cdot, k)\| \quad (4.27)
\]
when $\alpha \geq \alpha_1$. Combining (4.26) and (4.27) with (4.25) and letting $m \uparrow \infty$, we conclude that for each $k \in S$, $G_\alpha f(\cdot, k) = \sum_{i=0}^{\infty} \psi_i^{(k)}$, 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 $(X, \Lambda)$ and those of the killed Lévy type processes $X^{(k)}$, $k \in S$. 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 $(X, \Lambda)$ by \(\{P(t, (x, k), A) : t \geq 0, (x, k) \in \mathbb{R}^d \times S, A \in \mathcal{B}(\mathbb{R}^d \times S)\}\). We first prove that for any given $t > 0$, $x \in \mathbb{R}^d$, $k, l \in S$ and $A \in \mathcal{B}(\mathbb{R}^d)$,
\[
P(t, (x, k), A \times \{l\})
= \delta_{kl} P^{(k)}(t, x, A) + \sum_{m=1}^{+\infty} \int \cdots \int \sum_{0 < t_1 < t_2 < \cdots < t_m < t, t_{i+1} \neq t_i, t_m = t} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} P^{(l_0)}(t_1, x, dy_1) q_{l_0 l_1}(y_1) \times P^{(l_1)}(t_2 - t_1, y_1, dy_2) \cdots q_{l_{m-1} l_m}(y_m) P^{(l_m)}(t - t_m, y_m, A) dt_1 dt_2 \cdots dt_m, \quad (4.28)
\]
where $\delta_{kl}$ 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 $\mathbb{R}^d \times S$ such that $f(\cdot,k)$ is Lipschitz continuous for each $k \in S$, we define
\[
P_t f(x,k) : = \mathbb{E}_{x,k}[f(X(t),\Lambda(t))] = \sum_{l \in S} \int_{\mathbb{R}^d} f(y,l)P(t,(x,k),dy \times \{l\}), \tag{4.29}
\]
and
\[
\tilde{P}_t f(x,k) : = \sum_{l \in S} \int_{\mathbb{R}^d} f(y,l)\tilde{P}(t,(x,k),dy \times \{l\})
\]
\[
= \sum_{l \in S} \left[ \int_{\mathbb{R}^d} \delta_{kl}f(y,l)P^{(k)}(t,x,dy) + \sum_{m=1}^{+\infty} \int_{0< t_1 < t_2 < \cdots < t_m < t} \cdots \cdots \right.
\]
\[
\sum_{l_0,t_1,t_2,\ldots,t_m \in S \atop l_i \neq t_{i+1}, t_0 = k, t_m = t} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} P^{(l_0)}(t_1,x,dy_1)q_{0l_1}(y_1)P^{(l_1)}(t_2-t_1,y_1,dy_2) \cdots \cdots \)
\]
\[
x q_{l_{m-1}l_m}(y_m)P^{(l_m)}(t-t_m,y_m,dy)dt_1dt_2 \cdots dt_m \right], \tag{4.30}
\]

Since the process $(X,\Lambda)$ 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 $l \in S$ and each $m = 0, 1, \ldots$, 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 $\alpha_1$ 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 (Chen, 2004, Theorem 1.38)) to conclude that that $e^{\alpha_1 t}P_t f(x,k) = e^{\alpha_1 t}\tilde{P}_t f(x,k)$. That is,
\[
\sum_{l \in S} \int f(y,l)P(t,(x,k),dy \times \{l\}) = \sum_{l \in S} \int f(y,l)\tilde{P}(t,(x,k),dy \times \{l\}) \tag{4.31}
\]

Now we prove (4.28) by the Monotone Class Theorem (see, e.g., Theorem 1.35 in Chen (2004)). 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 Chen (2004)). 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 ((Chen, 2004, Definition 1.34)), \( L \) is an \( \mathcal{L} \)-system. Moreover, let \( C \) denote the set of all the open sets in \( \mathbb{R}^d \times \mathbb{S} \). Note that \( C \) is a \( \pi \)-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 (Chen, 2004, 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 \( 1_{A \times \{l\}}(x,k) \), which implies that (4.28) holds.

Finally, we use (4.28) to prove the Feller property for \((X, \Lambda)\). 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 Meyn and Tweedie (1993) 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, \((X, \Lambda)\) has the Feller property (see Proposition 6.1.1 in Meyn and Tweedie (1993) 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 \( \delta_0, H \) and a nonnegative function \( \vartheta \) defined on \([0, \delta_0]\) satisfying \( \lim_{r \to 0} \vartheta(r) = 0 \) such that

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

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

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 Peszat and Zabczyk (1995), Priola and Wang (2006), Qiao (2014) to obtain the strong Feller property.

Proposition 5.3. Under Assumptions 4.1, 4.2, and 5.1, for each \( k \in S \), 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 (2006). Fix an arbitrary \( k \in S \) 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

\[
\hat{a}(x, z, k) = \begin{pmatrix} a(x, k) & c(x, z, k) \\ c(x, z, k)^T & a(z, k) \end{pmatrix}, \quad b(x, z, k) = \begin{pmatrix} b(x, k) \\ b(z, k) \end{pmatrix}
\]

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

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

and

\[
\tilde{\mathcal{L}}_k h(x, z) := \tilde{\Omega}_d(k)h(x, z) + \tilde{\Omega}_j(k)h(x, z), \tag{5.3}
\]

where \( h \in C^2_0(\mathbb{R}^d \times \mathbb{R}^d) \) and \( \tilde{\Omega}_j(k) \) 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

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

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

\[
\tilde{\Omega}_d(k)F(|x - z|) = \frac{1}{2}F''(|x - z|)\overline{A}(x, z, k)
\]

\[
+ \frac{F'(|x - z|)}{2|x - z|} [\text{tr}(A(x, z, k)) - \overline{A}(x, z, k) + 2B(x, z, k)]
\]

\[
\leq 2\lambda_0 F''(|x - z|) + H F'(|x - z|) \vartheta(|x - z|)
\]

32
\[
\frac{\partial}{\partial x} = -\frac{4\lambda_0}{(1 + |x - z|^2)^3} + \frac{H}{(1 + |x - z|)^2}\vartheta(|x - z|), \tag{5.5}
\]
for all \(x, z \in \mathbb{R}^d\) with \(|x - z| \leq \delta_0\).

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

\[
F'(|x + u - z|) - F'(|x - z|) - \langle \nabla_x F(|x - z|), u \rangle 1_{B(0, \epsilon_0)}(u)
\]
\[
\leq F'(|x - z|)(|x + u - z| - |x - z|) - \frac{F'(|x - z|)}{|x - z|}(x - z, u) 1_{B(0, \epsilon_0)}(u)
\]
\[
\leq \frac{2|u|}{(1 + |x - z|)^2}.
\]

Hence it follows that

\[
\int (F(|x + u - z|) - F(|x - z|) - \langle \nabla_x F(|x - z|), u \rangle 1_{B(0, \epsilon_0)}(u)) \left(\nu(x, k, du) - \nu(z, k, du)\right)^+ 
\]
\[
\leq \frac{2}{(1 + |x - z|)^2} \int |u| \left(\nu(x, k, du) - \nu(z, k, du)\right)^+
\]

In the same manner, we have

\[
\int (F(|x - (z + u)|) - F(|x - z|) - \langle \nabla_x F(|x - z|), u \rangle 1_{B(0, \epsilon_0)}(u)) \left(\nu(x, k, du) - \nu(z, k, du)\right)^+
\]
\[
\leq \frac{2}{(1 + |x - z|)^2} \int |u| \left(\nu(z, k, du) - \nu(x, k, du)\right)^+
\]

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

\[
\int (F(|x + u - z - u|) - F(|x - z|) - \langle \nabla_z F(|x - z|), u \rangle 1_{B(0, \epsilon_0)}(u)
\]
\[
- \langle \nabla_z F(|x - z|), u \rangle 1_{B(0, \epsilon_0)}(u) \left(\nu(x, k, du) \land \nu(z, k, du)\right) = 0.
\]

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

\[
\tilde{\Omega}_j(k)F(|x - z|) 
\]
\[
\leq \frac{2}{(1 + |x - z|)^2} \left[ \int |u| \left(\nu(x, k, du) - \nu(z, k, du)\right)^+ + \int |u| \left(\nu(z, k, du) - \nu(x, k, du)\right)^+ \right]
\]
\[
\leq \frac{2}{(1 + |x - z|)^2} \int |u| \left(\nu(x, k, du) - \nu(z, k, du)\right)^+
\]
\[
\leq \frac{2}{(1 + |x - z|)^2} \int |u| \left(\nu(x, k, du) - \nu(z, k, du)\right)^+ (du)
\]
\[
\leq \frac{H}{(1 + |x - z|)^2} \vartheta(|x - z|), \tag{5.6}
\]

33
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 $|x - z| \leq \delta_0$,

$$
\tilde{\mathcal{L}}_k F(|x - z|) \leq \frac{-4\lambda_0}{(1 + |x - z|)^3} + \frac{2H}{(1 + |x - z|)^2} \vartheta(|x - z|)
$$

$$
\leq \frac{-4\lambda_0}{(1 + \delta_0)^3} + 2H \vartheta(|x - z|).
$$

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

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

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 $\tilde{\mathcal{L}}_k$ and denote by $T$ the coupling time. For $n, N \in \mathbb{N}$ and the $\delta$ 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

$$
0 \leq F(\delta) \mathbb{P}_k \{T_n \land \sigma_N > S_\delta\}
$$

$$
\leq \mathbb{E}_k[F(|\tilde{X}^{(k)}(x)(T_n \land S_\delta \land \sigma_N) - \tilde{Z}^{(k)}(z)(T_n \land S_\delta \land \sigma_N)|)]
$$

$$
= \mathbb{E}_k\left[\int_0^{T_n \land S_\delta \land \sigma_N} \tilde{\mathcal{L}}_k F(|\tilde{X}^{(k)}(x) - \tilde{Z}^{(k)}(z)|) ds\right]
$$

$$
\leq \mathbb{E}_k[T_n \land S_\delta \land \sigma_N],
$$

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

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

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

$$
F(\delta) \mathbb{P}_k \{T > S_\delta\} + \kappa \mathbb{E}[T \land S_\delta] \leq F(|x - z|).
$$
Then for any $t > 0$ and $0 < |x - z| < \delta$, 
\[
P_k \{ T > t \} = P_k \{ T > t, S_\delta > t \} + P_k \{ T > t, S_\delta \leq t \} \\
\leq P_k \{ T \land S_\delta > t \} + P_k \{ T > S_\delta \} \\
\leq \frac{1}{t} E_k [T \land S_\delta] + P_k \{ T > S_\delta \} \\
\leq \left( \frac{1}{t^{\kappa}} + \frac{1}{F(\delta)} \right) F(|x - z|).
\]

This implies the strong Feller property for the Lévy type process \( \tilde{X}^{(k)} \) immediately. Indeed, for any \( f \in B_b(\mathbb{R}^d) \), $t > 0$, and $0 < |x - z| < \delta$, we have 
\[
|E_k[f(\tilde{X}^{(k)}(x)(t))] - E_k[f(\tilde{X}^{(k)}(z)(t))]| \\
\leq E_k[|f(\tilde{X}^{(k)}(x)(t)) - f(\tilde{X}^{(k)}(z)(t))|] \\
= E_k[|f(\tilde{X}^{(k)}(x)(t)) - f(\tilde{X}^{(k)}(z)(t))|I_{\{T > t\}}] \\
\leq 2\|f\|_\infty P_k \{ T > t \} \to 0, \text{ as } |x - z| \to 0.
\]

Finally, as in the proof of Lemma 4.7, for any $f \in B_b(\mathbb{R}^d)$, $t > 0$, and $0 < |x - z| < \delta$, we can write 
\[
|E_k[f(X^{(k)}(x)(t))] - E_k[f(X^{(k)}(z)(t))]| \\
\leq E_k \left[ f(X^{(k)}(x)(t))e_{i_0}^t q_{kk}(\tilde{X}^{(k)}(x)(s))ds - f(\tilde{Z}^{(k)}(z)(t))e_{i_0}^t q_{kk}(\tilde{Z}^{(k)}(z)(s))ds \right] \\
\leq E_k \left[ f(X^{(k)}(x)(t)) - f(\tilde{Z}^{(k)}(z)(t)) \right] e_{i_0}^t q_{kk}(\tilde{X}^{(k)}(x)(s))ds \\
+ E_k \left[ f(\tilde{Z}^{(k)}(z)(t)) \right] e_{i_0}^t q_{kk}(\tilde{X}^{(k)}(x)(s))ds - e_{i_0}^t q_{kk}(\tilde{Z}^{(k)}(z)(s))ds \right] \\
\leq 2\|f\|_\infty P_k \{ T > t \} + \|f\|_\infty E_k \left[ \int_0^t q_{kk}(\tilde{X}^{(k)}(x)(s))ds - \int_0^t q_{kk}(\tilde{Z}^{(k)}(z)(s))ds \right] \\
\leq 2\|f\|_\infty \left( \frac{1}{t^{\kappa}} + \frac{1}{F(\delta)} \right) F(|x - z|) + H(n_0 - 1)\|f\|_\infty E_k \left[ \int_0^t |\tilde{X}^{(k)}(x)(s) - \tilde{Z}^{(k)}(z)(s)|ds \right] \\
\leq 2\|f\|_\infty \left( \frac{1}{t^{\kappa}} + \frac{1}{F(\delta)} \right) F(|x - z|) + H(n_0 - 1)\|f\|_\infty \int_0^t G^{-1}(G(|x - z|) + 3H)ds, 
\]

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 $\int_0^t G^{-1}(G(|x - z|) + 3H)ds \to 0$ as $|x - z| \to 0$. Thus it follows that $|E_k[f(X^{(k)}(x)(t))] - E_k[f(X^{(k)}(z)(t))]| \to 0$ as $|x - z| \to 0$. On the other hand, for any $|x - z| \geq \delta$, we have 
\[
|E_k[f(X^{(k)}(x)(t))] - E_k[f(X^{(k)}(z)(t))]| \leq 2\delta^{-1}\|f\|_\infty |x - z|.
\]
Therefore we obtain the desired strong Feller property for the killed Lévy process $X^{(k)}$. This completes the proof. □

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 $(X, \Lambda)$ possesses the strong Feller property.

**Acknowledgements**

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**

Bass, R. F. (1988). Uniqueness in law for pure jump Markov processes. *Probab. Theory Related Fields*, 79(2):271–287.

Bass, R. F. and Tang, H. (2009). The martingale problem for a class of stable-like processes. *Stochastic Process. Appl.*, 119(4):1144–1167.

Chen, M.-F. (2004). *From Markov chains to non-equilibrium particle systems*. World Scientific Publishing Co. Inc., River Edge, NJ, second edition.

Chen, M. F. and Li, S. F. (1989). Coupling methods for multidimensional diffusion processes. *Ann. Probab.*, 17(1):151–177.

Cloez, B. and Hairer, M. (2015). Exponential ergodicity for Markov processes with random switching. *Bernoulli*, 21(1):505–536.

Dawson, D. A. and Zheng, X. (1991). Law of large numbers and central limit theorem for unbounded jump mean-field models. *Adv. in Appl. Math.*, 12(3):293–326.

Feng, S. and Zheng, X. (1992). Solutions of a class of nonlinear master equations. *Stochastic Process. Appl.*, 43(1):65–84.

Hairer, M., Mattingly, J. C., and Scheutzow, M. (2011). Asymptotic coupling and a general form of Harris’ theorem with applications to stochastic delay equations. *Probab. Theory Related Fields*, 149(1-2):223–259.

Hoh, W. (1994). The martingale problem for a class of pseudo-differential operators. *Math. Ann.*, 300(1):121–147.

Ikeda, N. and Watanabe, S. (1989). *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, second edition.

Karatzas, I. and Shreve, S. E. (1991). *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition.
Komatsu, T. (1973). Markov processes associated with certain integro-differential operators. *Osaka J. Math.*, 10:271–303.

Kurtz, T. G. (1998). Martingale problems for conditional distributions of Markov processes. *Electron. J. Probab.*, 3:no. 9, 29 pp. (electronic).

Lindvall, T. (2002). *Lectures on the coupling method*. Dover Publications Inc., Mineola, NY. Corrected reprint of the 1992 original.

Lindvall, T. and Rogers, L. C. G. (1986). Coupling of multidimensional diffusions by reflection. *Ann. Probab.*, 14(3):860–872.

Mao, X. and Yuan, C. (2006). *Stochastic differential equations with Markovian switching*. Imperial College Press, London.

Meyn, S. P. and Tweedie, R. L. (1993). *Markov chains and stochastic stability*. Communications and Control Engineering Series. Springer-Verlag London, Ltd., London.

Mikulevicius, R. and Rozovskii, B. L. (1999). Martingale problems for stochastic PDE’s. In *Stochastic partial differential equations: six perspectives*, volume 64 of *Math. Surveys Monogr.*., pages 243–325. Amer. Math. Soc., Providence, RI.

Perkins, E. (1995). On the martingale problem for interactive measure-valued branching diffusions. *Mem. Amer. Math. Soc.*, 115(549):vi+89.

Peszat, S. and Zabczyk, J. (1995). Strong Feller property and irreducibility for diffusions on Hilbert spaces. *Ann. Probab.*, 23(1):157–172.

Priola, E. and Wang, F.-Y. (2006). Gradient estimates for diffusion semigroups with singular coefficients. *J. Funct. Anal.*, 236(1):244–264.

Qiao, H. (2014). Exponential ergodicity for SDEs with jumps and non-Lipschitz coefficients. *J. Theoret. Probab.*, 27(1):137–152.

Rudin, W. (1976). *Principles of mathematical analysis*. McGraw-Hill Book Co., New York-Auckland-Düsseldorf, third edition. International Series in Pure and Applied Mathematics.

Sethi, S. P. and Zhang, Q. (1994). *Hierarchical decision making in stochastic manufacturing systems*. Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA.

Shao, J. and Xi, F. (2014). Stability and recurrence of regime-switching diffusion processes. *SIAM J. Control Optim.*, 52(6):3496–3516.

Sharpe, M. (1988). *General theory of Markov processes*, volume 133 of *Pure and Applied Mathematics*. Academic Press, Inc., Boston, MA.

Shiga, T. and Tanaka, H. (1985). Central limit theorem for a system of Markovian particles with mean field interactions. *Z. Wahrsch. Verw. Gebiete*, 69(3):439–459.

Stroock, D. W. (1975). Diffusion processes associated with Lévy generators. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 32(3):209–244.

Stroock, D. W. and Varadhan, S. R. S. (1969a). Diffusion processes with continuous coefficients. I. *Comm. Pure Appl. Math.*, 22:345–400.

Stroock, D. W. and Varadhan, S. R. S. (1969b). Diffusion processes with continuous coefficients. II. *Comm. Pure Appl. Math.*, 22:479–530.
Stroock, D. W. and Varadhan, S. R. S. (1979). *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin.

Wang, J. (2010). Regularity of semigroups generated by Lévy type operators via coupling. *Stochastic Process. Appl.*, 120(9):1680–1700.

Wang, J.-M. (2014). Martingale problems for switched processes. *Math. Nachr.*, 287(10):1186–1201.

Xi, F. (1998). A note on the jump numbers of \( Q \)-processes. *J. Math. (Wuhan)*, 18(2):187–190.

Xi, F. (2008). Feller property and exponential ergodicity of diffusion processes with state-dependent switching. *Sci. China Ser. A*, 51(3):329–342.

Xi, F. (2009). Asymptotic properties of jump-diffusion processes with state-dependent switching. *Stochastic Process. Appl.*, 119(7):2198–2221.

Xi, F. and Zhao, L. (2006). On the stability of diffusion processes with state-dependent switching. *Sci. China Ser. A*, 49(9):1258–1274.

Yin, G. and Xi, F. (2010). Stability of regime-switching jump diffusions. *SIAM J. Control Optim.*, 48(7):4525–4549.

Yin, G. G. and Zhang, Q. (1998). *Continuous-Time Markov Chains and Applications: A Singular Perturbation Approach*, volume 37 of *Applications of Mathematics (New York)*. Springer-Verlag, New York.

Yin, G. G. and Zhu, C. (2010). *Hybrid Switching Diffusions: Properties and Applications*, volume 63 of *Stochastic Modelling and Applied Probability*. Springer, New York.

Zambotti, L. (2000). An analytic approach to existence and uniqueness for martingale problems in infinite dimensions. *Probab. Theory Related Fields*, 118(2):147–168.

Zheng, J. L. and Zheng, X. G. (1986). A martingale method for \( Q \)-processes. *Kexue Tongbao (Chinese)*, 31(17):1296–1298.

Zhu, C., Yin, G., and Baran, N. A. (2015). Feynman-Kac formulas for regime-switching jump diffusions and their applications. *Stochastics*, 87(6):1000–1032.