Submitted to Bernoulli

Properties of Switching Jump Diffusions: Maximum Principles and Harnack Inequalities

XIAOSHAN CHEN, [1] ZHEN-QING CHEN,[2] KY TRAN,[3] and GEORGE YIN[4]

[1] *School of Mathematical Sciences, South China Normal University, Guangdong, China.*
E-mail: *xschen@m.scnu.edu.cn

[2] *Departments of Mathematics, University of Washington, Seattle, WA 98195, USA.*
E-mail: **zqchen@uw.edu

[3] *Department of Mathematics, College of Education, Hue University, Hue city, Vietnam.*
E-mail: †quankysp@gmail.com

[4] *Department of Mathematics, Wayne State University, Detroit, MI 48202 USA.*
E-mail: gyin@math.wayne.edu

This work examines a class of switching jump diffusion processes. The main effort is devoted to proving the maximum principle and obtaining the Harnack inequalities. Compared with the diffusions and switching diffusions, the associated operators for switching jump diffusions are non-local, resulting in more difficulty in treating such systems. Our study is carried out by taking into consideration of the interplay of stochastic processes and the associated systems of integro-differential equations.

Keywords: jump diffusion, regime switching, maximum principle, Harnack inequality.

1. Introduction

In recent years, many different fields require the handling of dynamic systems in which there is a component representing random environment and other factors that are not given as a solution of the usual differential equations. Such systems have drawn new as well as resurgent attention because of the urgent needs of systems modeling, analysis, and optimization in a wide variety of applications. Not only do the applications arise from the traditional fields of mathematical modeling, but also they have appeared in emerging application areas such as wireless communications, networked systems, autonomous systems, multi-agent systems, flexible manufacturing systems, financial engineering, and biological and ecological systems, among others. Much effort has been devoted to the so-called hybrid systems. Taking randomness into consideration, a class of such systems known as switching diffusions has been investigated thoroughly; see for example, [23, 32] and references therein. Continuing our investigation on regime-switching systems, this paper focuses on a class switching jump diffusion processes. To work on such systems, it is necessary to study a number of fundamental properties. Although we have a good understanding of switching diffusions, switching jump diffusions are more difficult to deal with.

One of the main difficulties is the operator being non-local. When we study switching diffusions, it has been demonstrated that although they are similar to diffusion processes, switching diffusions have some distinct features. With the non-local operator used, the distinctions are even more pronounced. Our primary motivation stems from the study of a family of Markov processes in which continuous dynamics, jump discontinuity, and discrete events coexist. Their interactions reflect the salient features of the underlying systems. Specifically, we focus on regime-switching jump diffusion processes, in which the switching process is not exogenous but depends on the jump diffusions. The distinct features of the systems include the presence of non-local operators, the coupled systems of equations, and the tangled information due to the dependence of the switching process on the jump diffusions.

To elaborate a little more on the systems, similar to [32, Section 1.3, pp. 4-5], we begin with the following description. Consider a two component process (X_t, Λ_t), where $\Lambda_t \in \{1, 2\}$. We call Λ_t the discrete event process with state space $\{1, 2\}$. Imagine that we have two parallel planes. Initially, $\Lambda_0 = 1$. It then sojourns in the state 1 for a random duration. During this period, the diffusion with jump traces out a curve on plane 1 specified by the drift, diffusion, and jump coefficients. Then a random switching takes place at a random time τ_1, and Λ switches to plane 2 and sojourns there for a random duration. During this period, the diffusion with jump traces out a curve on plane 2 with different drift, diffusion, and jump coefficients. What we are interested in is the case that Λ_t itself is not Markov, but only the two-component process (X_t, Λ_t) is a Markov process. Treating such systems, similar to the study of switching diffusions, we may consider a number of questions: Under what conditions, will the processes be recurrent and positive recurrent? Under what conditions, will the process be positive recurrent? Is it true that positive recurrence implies the existence of an ergodic measure. To answer these questions, we need to examine a number of issues of the switching jump diffusions and the associated systems of integro-partial differential equations.

Switching jump diffusions models arise naturally in many applications. To illustrate, consider the following motivational example–an optimal stopping problem. It is an extension of the optimal stopping problem for switching diffusions with diffusion dependent switching in [22]. We assume that the dynamics are described by switching jump diffusions rather than switching diffusions. Consider a two component Markov process (X_t, Λ_t) given by

$$dX_t = b(X_t, \Lambda_t)dt + \sigma(X_t, \Lambda_t)dW(t) + \int_{\mathbb{R}_0} c(X_{t-}, \Lambda_{t-}, z)\widetilde{N}_0(dt, dz),$$

where $b(\cdot)$, $\sigma(\cdot)$, and $c(\cdot)$ are suitable real-valued functions, $\widetilde{N}_0(\cdot)$ is a compensated real-valued Poisson process, $W(\cdot)$ is a real-valued Brownian motion, and $\mathbb{R}_0 = \mathbb{R} - \{0\}$. Because the example is for motivation only, we defer the discussion of the precise setup, formulation, and conditions needed for switching jump diffusions to the next section. We assume that Λ depends on the dynamics of X. Denote the filtration by $\{\mathcal{F}_t\}_{t\geq 0}$ and let $\mathcal{T}$ be the collection of $\mathcal{F}_t$-stopping times. Then the treatment of the optimal stopping problem leads to the consideration of the following value function

$$V(x,i)=\sup_{\widetilde{\tau}\in\mathcal{T}}\mathbb{E}_{x,i}\Big[\int_0^{\widetilde{\tau}}[e^{-\beta t}L(X_t,\Lambda_t)dt+e^{-\beta\widetilde{\tau}}\widetilde{G}(X_{\widetilde{\tau}},\Lambda_{\widetilde{\tau}})]\Big],$$

where $L(\cdot)$ and $\widetilde{G}(\cdot)$ are suitable functions, and $X_0=x$ and $\Lambda_0=i$. As an even more specific example, consider an asset model

$$dX_t=b(X_t,\Lambda_t)dt+\sigma(X_t,\Lambda_t)dW(t)+\int_{\mathbb{R}_0}c(\Lambda_{t-},z)X_{t-}\widetilde{N}_0(dt,dz).$$

Then the risk-neutral price of the perpetual American put option is given by

$$V(x,i)=\sup_{\widetilde{\tau}\in\mathcal{T}}\mathbb{E}_{x,i}[K-X_{\widetilde{\tau}}]^+.$$

One of the motivations for using jump-diffusion type models is that it has been observed empirically that distributions of the returns often have heavier tails than that of normal distributions. In particular, if we take $N(t)$ to be a one-dimensional stationary Poisson process with $\mathbb{E}N(t)=\lambda t$ for some $\lambda>0$, and take the compensated Poisson process to be $\widetilde{N}(t)=N(t)-\lambda t$. The resulted system is used widely in option pricing and mean-variance portfolio selections.

Next, consider a modification of a frequently used system in control theory. Let Γ be a compact subset of $\mathbb{R}^d-\{0\}$ that is the range space of the impulsive jumps. For any subset B in Γ, $N(t,B)$ counts the number of impulses on $[0,t]$ with values in B. Consider

$$\begin{aligned}dX_t&=b(X_t,\Lambda_t)dt+\sigma(X_t,\Lambda_t)dW_t+dJ_t,\\ J_t&=\int_0^t\int_\Gamma c(X_{s-}),\Lambda_{s-}),\gamma)N(ds,d\gamma),\end{aligned}$$

with $X_0=x,\Lambda_0=\Lambda$, together with a transition probability specification of the form

$$\mathbb{P}\{\Lambda_{t+\Delta t}=j|\Lambda_t=i,(X_s,\Lambda_s),s\le t\}=q_{ij}(X_t)\Delta t+o(\Delta t),\quad i\ne j,$$

where b and σ are suitable vector-valued and matrix-valued functions, respectively, and W is a standard vector-valued Brownian motion. Assume that $N(\cdot,\cdot)$ is independent of the Brownian motion $W(\cdot)$ and the switching process $\Lambda(\cdot)$. Alternatively, we can write

$$d\Lambda_t=\int_{\mathbb{R}}h(X_{s-}),\Lambda_{s-}),z)N_1(dt,dz),$$

where $h(x,i,z)=\sum_{j\in\mathcal{M}}(j-i)\mathbf{1}_{\{z\in\Delta_{ij}(x)\}}$ with $\Delta_{ij}(x)$ being the consecutive left closed and right open intervals of the real line, and $\widetilde{N}(t,B)$ being a compensated Poisson measure, which is independent of the Brownian motion $W(t)$, $\lambda\in(0,\infty)$ is known as the jump rate and $\pi(B)$ is the jump measure; $N_1(dt,dz)$ is a Poisson measure with intensity $dt\times m_1(dz)$, and $m_1(dz)$ is the Lebesgue measure on $\mathbb{R}$, $N_1(dt,dz)$ is independent of the Brownian motion $W(t)$ and the Poisson measure $\widetilde{N}(\cdot,\cdot)$. Define a compensated or centered Poisson measure as

$$\widetilde{N}(t,B) = N(t,B) - \lambda t \pi(B), \quad \text{for} \quad B \subset \Gamma,$$

where $0 < \lambda < \infty$ is known as the jump rate and $\pi(\cdot)$ is the jump distribution (a probability measure). In the above, we used the setup similar to [21, p. 37]. With this centered Poisson measure, we can rewrite J_t as

$$J_t = \int_0^t \int_\Gamma g(X_{s-}, \Lambda_{s-}, \gamma) \widetilde{N}(ds, d\gamma) + \lambda \int_0^t \int_\Gamma g(X_{s-}, \Lambda_{s-}, \gamma) \pi(d\gamma) ds.$$

The related jump diffusion models without switching have been used in a wide range of applications in control systems; see [21] and references therein.

We devote our attention to the maximum principle and Harnack inequalities for the jump-diffusion processes with regime-switching in this paper. Apart from being interesting in their own right, they play very important roles in analyzing many properties such as recurrence, positive recurrence, and ergodicity of the underlying systems. There is growing interest in treating switching jump systems; see [31] and many references therein. However, up to date, there seems to be no results on maximum principles and Harnack inequality for jump-diffusion processes with regime switching. As was alluded to in the previous paragraph, the main difficulty is that the operators involved are non-local. Thus, the results obtained for the systems (known as weakly coupled elliptic systems) corresponding to switching diffusions cannot be carried over. Thus new approaches and ideas have to be used.

Looking into the literature, in [15], Evans proved the maximum principle for uniformly elliptic equations. In the classical book [27], Protter and Weinberger treated maximum principle for elliptic equations as well as Harnack inequalities and generalized maximum principle together with a number of other topics. For switching diffusion processes, several papers studied Harnack inequality for the weakly coupled systems of elliptic equations. In [14], Chen and Zhao assumed Hölder continuous coefficients, and established Harnack inequality and full Harnack inequality based on the representations and estimates of the Green function and harmonic measures of the operators in small balls. In [1], Arapostathis, Ghosh, and Marcus assumed only measurability of the coefficients to prove the desired results; their proofs were based on the approach of Krylov [19] for estimating the oscillation of a harmonic function on bounded sets. There have been much interest in treating jump processes and associated non-local operators. In a series of papers, Bass and Kassmann [3], Bass, Kassmann, and Kumagai [4], Bass and Levin [5], Chen and Kumagai [8, 9, 10], Foondun [16], Song and Vondracek [29] examined Harnack inequalities for Markov processes with discontinuous sample paths; see also Chen, Hu, Xie, and Zhang [11] for a related work and a maximum principle. In [6], Caffarelli and Silvestre considered nonlinear integro-differential equations arising from stochastic control problems with pure jump Lévy processes (without a Brownian motion) using a purely analytic approach. Nonlocal version of ABP (Alexandrov-Bakelman-Pucci) estimate, Harnack inequality, and regularity were obtained. Most recently, Harnack inequality for solutions to the Schrödinger operator were dealt with in [2] by Athreya and Ramachandran for jump diffusions on $\mathbb{R}^d$ with $d \geq 3$ whose associate operator is an integro-differential operator includes the pure jump part as well as elliptic part. Their approach is based on the comparability of Green functions and Poisson kernels using conditional gauge function and strong regularity is assumed on the coefficients of the diffusion and jumping components.

In this paper, we focus on stochastic processes that have a switching component in addition to the jump diffusion component. The switching in fact is "jump diffusion dependent"; more precise notion will be given in the formulation section. When the switching component is missing, it reduces to the jump diffusion processes; when the continuous disturbance due to Brownian motion is also missing, it reduces to the case of pure jump processes. If only the jump process is missing, it reduces to the case of switching diffusions. Compared to the case of switching diffusion processes, in lieu of systems of elliptic partial differential equations, we have to deal with systems of integro-differential equations. Using mainly a probabilistic approach, we establish the maximum principles. Because local analysis alone is not adequate, the approach treating Harnack inequality for switching diffusion processes cannot be used in the current case. We adopt the probabilistic approach via Krylov type estimates from [5], which was further extended in [3, 16, 29], to derive the Harnack inequality for the nonnegative solution of the system of integro-differential equations.

We remark that since in this paper we are concerned with regime-switching jump-diffusions, we assume the regime-switching component $m \geq 2$. However, the results and their proofs of this paper hold for the case of $m = 1$ but are much easier as there would be no regime-switching. In particular, as a byproduct we have Harnack inequality for non-negative harmonic functions for Schrodinger operator $\mathcal{L} + q$, where $\mathcal{L}$ is an integro-differential operator of (3.1) and $q \leq 0$ is bounded and measurable. Although in this case our potential q is non-positive and bounded while in [2] the potential q can be a function in a suitable Kato class, our integro-differential operator $\mathcal{L}$ of (3.1) has very general non-local operator component and the diffusion coefficients and the jumping measure are much less regular than that in [2].

The rest of the paper is arranged as follows. Section 2 presents the formulation of the problem. In Section 3, we develop the maximum principle for regime-switching jump-diffusions, using a probabilistic approach that allows us to work under a quite general context. We obtain the Harnack inequality for the regime-switching jump-diffusions processes in Section 4. Finally, the paper is concluded with further remarks.

2. Formulation

Throughout the paper, we use z' to denote the transpose of $z \in \mathbb{R}^{l_1 \times l_2}$ with $l_1, l_2 \geq 1$, and $\mathbb{R}^{d\times 1}$ is simply written as $\mathbb{R}^d$. If $x \in \mathbb{R}^d$, the norm of x is denoted by $|x|$. For $x_0 \in \mathbb{R}^d$ and $r > 0$, $B(x_0, r)$ denotes the open ball in $\mathbb{R}^d$ centered at x_0 with radius $r > 0$. If D is a Borel set in $\mathbb{R}^d$, $\overline{D}$ and $D^c = \mathbb{R}^d \setminus D$ denote the closure and the complement of D, respectively. The space $C^2(D)$ refers to the class of functions whose partial derivatives up to order 2 exist and are continuous in D, and $C_b^2(D)$ is the subspace of $C^2(D)$ consisting of those functions whose partial derivatives up to order 2 are bounded. The indicator function of a set A is denoted $\mathbf{1}_A$. Let $Y_t = (X_t, \Lambda_t)$ be a two component Markov process such that X is an $\mathbb{R}^d$-valued process, and Λ is a switching process taking values in a finite set $\mathcal{M} = \{1, 2, \dots, m\}$. Throughout this paper, $d \geq 1$ and $m \geq 2$. Let $b(\cdot, \cdot) : \mathbb{R}^d \times \mathcal{M} \mapsto \mathbb{R}^d$, $\sigma(\cdot, \cdot) : \mathbb{R}^d \times \mathcal{M} \mapsto \mathbb{R}^d \times \mathbb{R}^d$, and for each $x \in \mathbb{R}^d$, $\pi_i(x, dz)$ is a σ-finite measure on $\mathbb{R}^d$ satisfying

$$\int_{\mathbb{R}^d} (1 \wedge |z|^2) \pi_i(x, dz) < \infty.$$

Let $Q(x) = (q_{ij}(x))$ be an $m \times m$ matrix depending on x such that

$$q_{ij}(x) \geq 0 \quad \text{for } i \neq j, \quad \sum_{j \in \mathcal{M}} q_{ij}(x) \leq 0.$$

Define

$$Q(x) f(x, \cdot)(i) := \sum_{j \in \mathcal{M}} q_{ij}(x) f(x, j).$$

The generator $\mathcal{G}$ of the process (X_t, Λ_t) is given as follows. For a function $f : \mathbb{R}^d \times \mathcal{M} \mapsto \mathbb{R}$ and $f(\cdot, i) \in C^2(\mathbb{R}^d)$ for each $i \in \mathcal{M}$, define

$$\mathcal{G} f(x, i) = \mathcal{L}_i f(x, i) + Q(x) f(x, \cdot)(i), \quad (x, i) \in \mathbb{R}^d \times \mathcal{M}, \tag{2.1}$$

where

$$\begin{aligned} \mathcal{L}_i f(x, i) \quad = \quad & \sum_{k,l=1}^{d} a_{kl}(x, i) \frac{\partial^2 f(x, i)}{\partial x_k \partial x_l} + \sum_{k=1}^{d} b_k(x, i) \frac{\partial f(x, i)}{\partial x_k} \\ & + \int_{\mathbb{R}^d} \big(f(x + z, i) - f(x, i) - \nabla f(x, i) \cdot z \mathbf{1}_{\{|z|<1\}} \big) \pi_i(x, dz), \end{aligned} \tag{2.2}$$

where $a(x, i) := \big(a_{kl}(x, i)\big) = \sigma(x, i)\sigma'(x, i)$, $\nabla f(\cdot, i)$ denotes the gradient of $f(\cdot, i)$.

Let $\Omega = D\big([0, \infty), \mathbb{R}^d \times \mathcal{M}\big)$ denote the space of all right continuous functions mapping $[0, \infty)$ to $\mathbb{R}^d \times \mathcal{M}$, having finite left limits. Define $(X_t, \Lambda_t) = w(t)$ for $w \in \Omega$ and let $\{\mathcal{F}_t\}$ be the right continuous filtration generated by the process (X_t, Λ_t). A probability measure $\mathbb{P}_{x,i}$ on Ω is a solution to the martingale problem for $\big(\mathcal{G}, C_b^2(\mathbb{R}^d)\big)$ started at (x, i) if

(a) $\mathbb{P}_{x,i}(X_0 = x, \Lambda_0 = i) = 1$,
(b) if $f(\cdot, i) \in C_b^2(\mathbb{R}^d)$ for each $i \in \mathcal{M}$, then

$$f(X_t, \Lambda_t) - f(X_0, \Lambda_0) - \int_0^t \mathcal{G} f(X_s, \Lambda_s) ds,$$

is a $\mathbb{P}_{x,i}$ martingale.
If for each (x, i), there is only one such $\mathbb{P}_{x,i}$, we say that the martingale problem for $\big(\mathcal{G}, C_b^2(\mathbb{R}^d)\big)$ is well-posed.

Definition 2.1 Let $U = D \times \mathcal{M}$ with $D \subset \mathbb{R}^d$ being a bounded connected open set. A bounded and Borel measurable function $f : \mathbb{R}^d \times \mathcal{M} \mapsto \mathbb{R}^d$ is said to be $\mathcal{G}$-harmonic in U if for any relatively compact open subset V of U,

$$f(x,i) = \mathbb{E}_{x,i}\big[f\left(X(\tau_V), \Lambda(\tau_V)\right)\big] \quad \text{for all} \quad (x,i) \in V,$$

where $\tau_V = \inf\{t \geq 0 : (X(t), \Lambda(t)) \notin V\}$ is the first exit time from V.

Throughout the paper, we assume conditions (A1)-(A3) hold until further notice.

(A1) The functions $\sigma(\cdot, i)$ and $b(\cdot, i)$ are bounded continuous on $\mathbb{R}^d$, and $q_{ij}(\cdot)$ is bounded Borel measurable on $\mathbb{R}^d$ for every $i, j \in \mathcal{M}$.

(A2) There exists a constant $\kappa_0 \in (0, 1]$ such that

$$\kappa_0|\xi|^2 \leq \xi' a(x,i)\xi \leq \kappa_0^{-1}|\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^d, x \in \mathbb{R}^d, i \in \mathcal{M},$$

and $|b(x,i)| \leq \kappa_0^{-1}$ for all $x \in \mathbb{R}^d$ and $i \in \mathcal{M}$.

(A3) There exists a σ-finite measure $\Pi(dz)$ so that $\pi_i(x, dz) \leq \Pi(dz)$ for every $x \in \mathbb{R}^d$ and $i \in \mathcal{M}$ and

$$\int_{\mathbb{R}^d} \left(1 \wedge |z|^2\right) \Pi(dz) \leq \kappa_1 < \infty.$$

(A4) For any $i \in \mathcal{M}$ and $x \in \mathbb{R}^d$, $\pi_i(x, dz) = \widetilde{\pi}_i(x, z)dz$. Moreover, for any $r \in (0, 1]$, any $x_0 \in \mathbb{R}^d$, any $x, y \in B(x_0, r/2)$ and $z \in B(x_0, r)^c$, we have

$$\widetilde{\pi}_i(x, z - x) \leq \alpha_r \widetilde{\pi}_i(y, z - y),$$

where α_r satisfies $1 \leq \alpha_r \leq \kappa_2 r^{-\beta}$ with κ_2 and β being positive constants.

Remark 2.2 (a) Under Assumptions (A1)-(A3), for each $i \in \mathcal{M}$, the martingale problem for $(\mathcal{L}_i, C_b^2(\mathbb{R}^d))$ is well-posed for every starting point $x \in \mathbb{R}^d$ (see [20, Theorem 5.2]). Then the switched Markov process (X_t, Λ_t) can be constructed from jump diffusions having infinitesimal generators $\mathcal{L}_i$, $1 \leq i \leq m$, as follows. Let X^i be the strong Markov process whose distribution is the unique solution to the martingale problem $(\mathcal{L}_i, C_b^2(\mathbb{R}^d))$. Suppose we start the process at (x_0, i_0), run a subprocess $\widetilde{X}^{i_0}$ of X^{i_0} that got killed with rate $-q_{i_0 i_0}(x)$; that is, via Feynman-Kac transform $\exp\left(\int_0^t q_{i_0 i_0}(X_s^{i_0})ds\right)$. Note that this subprocess $\widetilde{X}^{i_0}$ has infinitesimal generator $\mathcal{L}_{i_0} + q_{i_0 i_0}$. At the lifetime τ_1 of the killed process $\widetilde{X}^{i_0}$, jump to plane $j \neq i_0$ with probability $-q_{i_0 j}(X^{i_0}(\tau_1-))/q_{i_0 i_0}(X^{i_0}(\tau_1-))$ and run an independent copy of a subprocess $\widetilde{X}^j$ of X^j with killing rate $-q_{jj}(x)$ from position $X^{i_0}(\tau_1-)$. Repeat this procedure. The resulting process (X_t, Λ_t) is a strong Markov process with lifetime ζ by [17, 24]. For each $x \in \mathbb{R}^d$, we say that the matrix $Q(x)$ is *Markovian* if $\sum_{j \in \mathcal{M}} q_{ij}(x) = 0$ a.e. on $\mathbb{R}^d$ for every $i \in \mathcal{M}$, and *sub-Markovian* if $\sum_{j \in \mathcal{M}} q_{ij}(x) \leq 0$ a.e. on $\mathbb{R}^d$ for every $i \in \mathcal{M}$. When $Q(x)$ is *Markovian*, the lifetime $\zeta = \infty$, and when $Q(x)$ is just *sub-Markovian*, ζ can be finite. We use the convention that $(X_t, \Lambda_t) = \partial$ for $t \geq \zeta$, where ∂ is a cemetery point, and any function is extended to ∂ by taking value zero there. It is easy to check that the law of (X_t, Λ_t) solves the martingale problem for $\big(\mathcal{G}, C_b^2(\mathbb{R}^d)\big)$ so it is the desired switched jump-diffusion. This way of constructing switched diffusion has been utilized in [13, p.296]. It follows from [30] that law of (X_t, Λ_t) is the unique solution to the martingale problem for $\big(\mathcal{G}, C_b^2(\mathbb{R}^d)\big)$.

(b) Conditions (A1) and (A2) presents the uniform ellipticity of $a(x,i)$ and the uniform boundedness of $b(x,i)$ and $q_{ij}(x)$. The measure $\pi_i(x,dz)$ can be thought of as the intensity of the number of jumps from x to $x+z$ (see [3, 5]). Condition (A4) tells us that $\pi_i(x,dy)$ is absolutely continuous with respect to the Lebesgue measure dx on $\mathbb{R}^d$, and the intensities of jumps from x and y to a point z are comparable if x, y are relatively far from z but relatively close to each other. If $\widetilde{\pi}_i(x,z)$ is such that

$$\frac{c_i^{-1}}{|z|^{d+\beta_i}} \le \widetilde{\pi}_i(x,z) \le \frac{c_i}{|z|^{d+\beta_i}}$$

for some $c_i \ge 1$ and $\beta_i \in (0,2)$, then condition (A4) is satisfied with $1 < \alpha_r < \kappa_2$ independent of $r \in (0,1)$. Condition (A4) is an essential hypothesis in the proof of the Harnack inequality.

Throughout the paper, we use capital letters $C_1, C_2, \dots$ for constants appearing in the statements of the results, and lowercase letters $c_1, c_2, \dots$ for constants appearing in proofs. The numbering of the latter constants afresh in every new proof.

3. Maximum Principle

In this section, we establish maximum principle for the coupled system under conditions (A1)-(A3). We emphasize that we do not assume condition (A4) for the maximum principle. In Subsection 3.1, we prepare three propositions for general diffusions with jumps that will be used several times in the sequel.

3.1. Jump Diffusions and Strict Positivity

Consider

$$\begin{aligned}\mathcal{L}f(x) \quad &= \quad \sum_{k,l=1}^{d} a_{kl}(x)\frac{\partial^2 f(x)}{\partial x_k \partial x_l} + \sum_{k=1}^{d} b_k(x)\frac{\partial f(x)}{\partial x_k} \\ &\quad + \int_{\mathbb{R}^d} \big(f(x+z) - f(x) - \nabla f(x)\cdot z\mathbf{1}_{\{|z|<1\}}\big)\,\pi(x,dz), \qquad (3.1)\end{aligned}$$

where $(a_{kl}(x))$ is a continuous matrix-valued function and $b(x) = (b_1(x),\dots,b_d(x))$ is a $\mathbb{R}^d$-valued function on $\mathbb{R}^d$ such that

$$\lambda^{-1} I_{d\times d} \le (a_{kl}(x)) \le \lambda I_{d\times d} \quad \text{and} \quad \|b\|_\infty \le \lambda \qquad \text{on } \mathbb{R}^d \qquad (3.2)$$

for some $\lambda \ge 1$, and $\pi(x,dz)$ is a σ-finite measure on $\mathbb{R}^d$ satisfying

$$K := \int_{\mathbb{R}^d} \big(1 \wedge |z|^2\big) \sup_{x\in\mathbb{R}^d} \pi(x,dz) < \infty. \qquad (3.3)$$

Here $I_{d\times d}$ denotes the $d\times d$-identity matrix. By [20, Theorem 5.2], there is a unique conservative strong Markov process $X=\{X_t, t\geq 0; \mathbb{P}_x, x\in\mathbb{R}^d\}$ that is the unique solution to the martingale problem $(\mathcal{L}, C_b^2(\mathbb{R}^d))$. Suppose $q\geq 0$ is a bounded function on $\mathbb{R}^d$. One can kill the sample path of X with rate q. For this, let η be an independent exponential random variable with mean 1. Let

$$\zeta=\inf\left\{t>0: \int_0^t q(X_s)ds>\eta\right\}$$

and define $Z_t=X_t$ for $t<\zeta$ and $Z_t=\partial$ for $t\geq\zeta$, where ∂ is a cemetery point. It is easy to see that for any $x\in\mathbb{R}^d$ and $\varphi\geq 0$ on $\mathbb{R}^d$,

$$\mathbb{E}_x[\varphi(Z_t); t<\zeta]=\mathbb{E}_x\left[e_q(t)\varphi(X_t)\right], \quad t\geq 0, \tag{3.4}$$

where

$$e_q(t):=\exp\left(-\int_0^t q(X_s)ds\right).$$

The process Z is called the subprocess of X killed at rate q, and ζ the lifetime of Z. For $A\subset\mathbb{R}^d$, we define its hitting time and exit time of Z by

$$\sigma_A^Z=\inf\{t\geq 0: Z_t\in A\} \quad \text{and} \quad \tau_A^Z=\inf\{t\geq 0: Z_t\notin A\},$$

with the convention that $\inf\emptyset=\infty$. Note that $\tau_A^Z\leq\zeta$. The following two propositions are based on the support theorem for diffusions with jumps in [16].

Proposition 3.1 *There is a positive constant C_1 depending only on λ and K in* (3.2)-(3.3) *and an upper bound on $\|q\|_\infty$ such that for any $R\in(0,1]$, $r\in(0,R/4)$, $x_0\in\mathbb{R}^d$, $x\in B(x_0,3R/2)$ and $y\in B(x_0,2R)$,*

$$\mathbb{P}_y\left(\sigma^Z_{B(x,r)}<\tau^Z_{B(x_0,2R)}\right)\geq C_1 r^6.$$

Proof: Note that $Z_t=X_t$ for $t\in[0,\zeta)$, where ζ is the lifetime of Z. Define

$$\sigma_{B(x,r)}=\inf\{t\geq 0: X_t\in B(x,r)\}, \quad \tau_{B(x_0,2R)}=\inf\{t\geq 0: X_t\notin B(x_0,2R)\}.$$

Define a function $\phi:[0,8]\mapsto\mathbb{R}^d$ as follows

$$\phi(t)=y+\frac{x-y}{|x-y|}t, \quad t\in[0,8].$$

By [16, Theorem 4.2] and [16, Remark 4.3], there exists a constant $c_1>0$ so that

$$\mathbb{P}_y\Big(\sup_{t\leq 8}|X_t-\phi(t)|<r\Big)\geq c_1 r^6, \tag{3.5}$$

for any $x\in B(x_0,3R/2)$ and $r\in(0,R/4)$. Moreover, c_1 depends only on λ and an upper bound on $\|b\|_\infty$ and $\|q\|_\infty$. Since $|\phi'(t)|=1$ and $\phi(|x-y|)=x$, on $\big\{\sup_{t\leq 8}|X_t-\phi(t)|<r\big\}$,

we have $X_{|x-y|} \in B(x,r)$, $X_t \in B(x_0, 2R)$ for $0 \le t \le |x-y|$, and $|X_{8R} - x_0| \ge |X_{8R} - y| - |y - x_0| > 3R$. As a result, $\sigma_{B(x,r)} < |x-y| < \tau_{B(x_0,2R)} < 8R \le 8$ on $\left\{ \sup_{t\le 8} |X_t - \phi(t)| < r \right\}$. Then (3.5) leads to

$$\mathbb{P}_y\Big(\sigma_{B(x,r)} < \tau_{B(x_0,2R)} < 8\Big) \ge c_1 r^6.$$

It follows from (3.4) that

$$\begin{aligned} \mathbb{P}_y\left(\sigma^Z_{B(x,r)} < \tau^Z_{B(x_0,2R)}\right) &\ge \mathbb{P}_y\left(\sigma^Z_{B(x,r)} < \tau^Z_{B(x_0,2R)} < 8 < \zeta\right) \\ &\ge \exp(-6\|q\|_\infty)\mathbb{P}_y\Big(\sigma_{B(x,r)} < \tau_{B(x_0,2R)} < 8\Big) \\ &\ge \exp(-6\|q\|_\infty) c_1 r^6. \end{aligned}$$

This proves the proposition. □

Proposition 3.2 (i) *For any $0 < r \le 1/2$ and $x_0 \in \mathbb{R}^d$, if $A \subset B(x_0, r)$ has positive Lebesgue measure, then $\mathbb{P}_x(\sigma^Z_A < \tau^Z_{B(x_0,2r)}) > 0$ for every $x \in B(x_0, r)$.*

(ii) *Let $\rho \in (0,1)$ be a constant. There exist a nondecreasing function $\Phi : (0,\infty) \mapsto (0,\infty)$ and $r_0 \in (0, 1/2]$, depending only on λ and K in (3.2)-(3.3) and an upper bound on $\|q\|_\infty$, such that for any $x_0 \in \mathbb{R}^d$, any $r \in (0, r_0)$, and any Borel subset A of $B(x_0, r)$ with $|A|/r^d \ge \rho$, we have*

$$\mathbb{P}_x\Big(\sigma^Z_A < \tau^Z_{B(x_0,2r)}\Big) \ge \frac{1}{2}\Phi\left(|A|/r^d\right), \quad x \in B(x_0, r). \tag{3.6}$$

Proof: As in the proof of Proposition 3.1, define

$$\sigma_A = \inf\{t \ge 0 : X_t \in A\} \quad \text{and} \quad \tau_{B(x_0,2r)} = \inf\{t \ge 0 : X_t \notin B(x_0, 2r)\}.$$

By [16, Corollary 4.9], there is a nondecreasing function $\Phi : (0,\infty) \mapsto (0,\infty)$ such that if $A \subset B(x_0, r)$, $|A| > 0$, $r \in (0, 1/2]$ and $x \in B(x_0, r)$, then

$$\mathbb{P}_x(\sigma_A < \tau_{B(x_0,2r)}) \ge \Phi\left(|A|/r^d\right). \tag{3.7}$$

Using test function and Itô's formula, it is easy to derive (see [16, Proposition 3.4(b)] or Proposition 4.4 below) that there is a constant $c_1 > 0$ independent of x_0 and $r \in (0, 1/2]$ so that

$$\mathbb{E}_x \tau_{B(x_0,2r)} \le c_1 r^2 \quad \text{for any } x \in B(x_0, 2r). \tag{3.8}$$

(i) Suppose $0 < r \le 1/2$ and $A \subset B(x_0, r)$ has positive Lebesgue measure. Then by (3.7), $\mathbb{P}_x(\sigma_A < \tau_{B(x_0,2r)}) > 0$. Hence in view of (3.8), we have for every $x \in B(x_0, r)$,

$$\mathbb{P}_x(\sigma^Z_A < \tau^Z_{B(x_0,2r)}) \ge \mathbb{P}_x(\sigma_A < \tau_{B(x_0,2r)} < \zeta) = \mathbb{E}_x\left[e_q(\tau_{B(x_0,2r)})\mathbf{1}_{\{\sigma_A < \tau_{B(x_0,2r)}\}}\right] > 0.$$

(ii) Observe that

$$
\begin{aligned}
\mathbb{P}_x\Big(\sigma_A^Z < \tau^Z_{B(x_0,2r)}\Big) &\geq \mathbb{P}_x\Big(\sigma_A < \tau_{B(x_0,2r)}; \tau_{B(x_0,2r)} < \zeta\Big) \\
&\geq \mathbb{P}_x\Big(\sigma_A < \tau_{B(x_0,2r)}\Big) - \mathbb{P}_x\Big(\tau_{B(x_0,2r)} \geq \zeta\Big)
\end{aligned}
\tag{3.9}
$$

For $A \subset B(x_0, r)$ with $|A| \geq \rho r^d$, we have $\mathbb{P}_x(\sigma_A < \tau_{B(x_0,2r)}) \geq \Phi(\rho)$. On the other hand,

$$
\mathbb{P}_x(\zeta > t) = \mathbb{E}_x\left[\exp\left(-\int_0^t q(X_s)ds\right)\right] \geq \exp(-\|q\|_\infty\, t).
$$

This combined with (3.8) yields that

$$
\begin{aligned}
\mathbb{P}_x(\zeta > \tau_{B(x_0,2r)}) &\geq \mathbb{P}_x(\zeta > r > \tau_{B(x_0,2r)}) \\
&\geq \mathbb{P}_x(\zeta > r) - \mathbb{P}_x(\tau_{B(x_0,2r)} \geq r) \\
&\geq \exp(-\|q\|_\infty r) - \frac{\mathbb{E}_x \tau_{B(x_0,2r)}}{r} \\
&\geq \exp(-\|q\|_\infty r) - c_1 r.
\end{aligned}
$$

Since $\lim_{r\to 0}\left(\exp(-\|q\|_\infty r) - c_1 r\right) = 1$, there is a constant $r_0 \in (0, 1/2]$ such that

$$
\mathbb{P}_x\Big(\tau_{B(x_0,2r)} \geq \zeta\Big) \leq \frac{1}{2}\Phi(\rho) \quad \text{for all } r \in (0, r_0). \tag{3.10}
$$

The desired conclusion follows from (3.9), (3.7), and (3.10). □

For a connected open subset $D \subset \mathbb{R}^d$ and a Borel measurable function $f \geq 0$ on D, define $G_D^q f(x) = \mathbb{E}_x\left[\int_0^{\tau_D^Z} f(Z_s)ds\right]$.

Proposition 3.3 *For $f \geq 0$, either $G_D^q f(x) > 0$ on D or $G_D^q f(x) \equiv 0$ on D. Moreover, if $G_D^q f > 0$ on D if and only if $\{x \in D : f(x) > 0\}$ has positive Lebesgue measure.*

Proof: Suppose that $A := \{x \in D : G_D^q f(x) > 0\}$ has positive Lebesgue measure. We claim that for any $r \in (0, 1]$ and $B(x_0, r) \subset D$ so that $B(x_0, r/2) \cap A$ has positive Lebesgue measure, then $B(x_0, r/2) \subset A$. This is because if $B(x_0, r/2) \cap A$ has positive Lebesgue measure, then there is a compact subset $K \subset B(x_0, r/2) \cap A$ having positive Lebesgue measure. By Proposition 3.2 (i), we have $\mathbb{P}_x(\sigma_K^Z < \tau^Z_{B(x_0,r)}) > 0$ for every $x \in B(x_0, r/2)$. Consequently,

$$
G_D^q f(x) = \mathbb{E}_x \int_0^{\tau_D^Z} f(Z_s)ds \geq \mathbb{E}_x\left[G_D^q f(Z_{\sigma_K}); \sigma_K^Z < \tau^Z_{B(x_0,r)}\right] > 0
$$

for every $x \in B(x_0, r/2)$. This proves the claim. Since $B(x_0, r/2) \subset A$, by a chaining argument, the above reasoning shows that $A = D$ if A has positive Lebesgue measure. Now assume that $G_D^q f = 0$ a.e. on D. Since $G_D^q f(x) = \mathbb{E}_x \int_0^{\tau_D} e_q(s) f(X_s)ds$, we have $G_D f(x) := \mathbb{E}_x \int_0^{\tau_D} f(X_s)ds = 0$ a.e. on D. In particular, $G_D(f \wedge n) = 0$ a.e. on D. By [16, Theorem 2.3], bounded harmonic functions of X is Hölder continuous. By the proof of [4, Proposition 3.3], this together with (3.8) implies that $G_D(f\wedge n)$ is Hölder continuous on D. Therefore we have $G_D(f\wedge n)(x)=0$ for every $x\in D$. Consequently, $G_Df(x)=0$ for every $x\in D$ and so is $G^q_Df(x)$. This proves the first part of the proposition.

For the second part of the proposition, suppose that $f\geq 0$ and $f=0$ a.e. on D. It follows from [25, Corollary 2] that for every $x_0\in\mathbb{R}^d$,

$$\mathbb{E}_x\int_0^{\tau_{B(x_0,1)}}(\mathbf{1}_Df)(X_s)ds=0\quad\text{for every } x\in B(x_0,1). \tag{3.11}$$

We claim that $\mathbb{E}_x\int_0^{\tau_D}f(X_s)ds=0$ for every $x\in D$. For this, we define a sequence of stopping times: $\tau_0:=0$, $\tau_1:=\inf\{t\geq 0:|X_t-X_0|\geq 1\}\wedge\tau_D$, and for $n\geq 2$, $\tau_n:=\inf\{t\geq\tau_{n-1}:|X_t-X_{\tau_{n-1}}|\geq 1\}\wedge\tau_D$. Note that on $\{\lim_{n\to\infty}\tau_n<\tau_D\}$, $\lim_{n\to\infty}X_{\tau_n}=X_{\lim_{n\to\infty}}$ by the left-continuity of X_t. On the other hand, the sequence $\{X_{\tau_n};n\geq 1\}$ diverges on $\{\lim_{n\to\infty}\tau_n<\tau_D\}$ as $|X_{\tau_n}-X_{\tau_{n-1}}|\geq 1$ by the definition of τ_n. This contradiction implies that $\mathbb{P}_x(\lim_{n\to\infty}\tau_n<\tau_D)=0$; in other words, $\lim_{n\to\infty}\tau_n=\tau_D$ $\mathbb{P}_x$-a.s. Consequently, we have by (3.11)

$$\begin{aligned}\mathbb{E}_x\int_0^{\tau_D}f(X_s)ds &= \mathbb{E}_x\sum_{n=1}^\infty\int_{\tau_{n-1}}^{\tau_n}f(X_s)ds\\ &= \sum_{n=1}^\infty\mathbb{E}_x\left[\mathbb{E}_{X_{\tau_{n-1}}}\int_0^{\tau_{B(X_{\tau_{n-1}},1)}\wedge\tau_D}f(X_s)ds;\tau_{n-1}<\tau_D\right]\\ &= 0.\end{aligned}$$

It follows then $G^q_Df(x)=\mathbb{E}_x\int_0^{\tau_D}e_q(X_s)f(X_s)ds=0$ for every $x\in D$. This proves that if $f\geq 0$ and $f=0$ a.e. on D, then $G^q_Df\equiv 0$ on D. Next suppose that $f\geq 0$ is a bounded function on $\mathbb{R}^d$ and $\{x\in D:f(x)>0\}$ has positive Lebesgue measure, we will show that $G^q_D(x)>0$ for every $x\in D$. Let $c_p>0$ be the constant in the Remark following Theorem 3.1 on p.282 of [20]. Using a localization argument if needed, we may assume that $|a_{ij}(x)-a_{ij}(y)|\leq 1/c_p$ for every $x,y\in\mathbb{R}^d$. Let K be a compact subset of D so that $\{x\in K:f(x)>0\}$ has positive Lebesgue measure. Then by Theorem 3.6 and the proof of Theorem 4.2 both in [20], for $\lambda>0$ large, $v(x):=\mathbb{E}_x\int_0^\infty e^{-\lambda t}(\mathbf{1}_Kf)(X_s)ds$ is non-trivial on $\mathbb{R}^d$. We define a sequence of stopping times as follows. Let $S_1:=\sigma_K$, $T_1:=\inf\{t>\sigma_K:X_t\notin D\}$; for $n\geq 2$, define $S_n:=\inf\{t>T_{n-1}:X_t\in K\}$ and $T_n:=\inf\{t>S_n:X_t\notin D\}$. Then

$$v(x)=\sum_{n=1}^\infty\mathbb{E}_x\int_{S_n}^{T_n}e^{-\lambda s}f(X_s)ds=\sum_{n=1}^\infty\mathbb{E}_x\left[e^{-\lambda S_n}G_{D,\lambda}(\mathbf{1}_Kf)(X_{S_n})\right],$$

where $G_{D,\lambda}\varphi(x):=\mathbb{E}_x\int_0^{\tau_D}e^{-\lambda s}\varphi(X_s)ds$. Hence $G_{D,\lambda}(\mathbf{1}_Kf)(x)$ cannot be identically zero on K. By the first part of this proof (by taking $q=\lambda$), we have $G_{D,\lambda}(\mathbf{1}_Kf)(x)>0$ for every $x\in D$. It follows that $G_Df(x)>0$ and so $G^q_Df(x)>0$ for every $x\in D$. □

3.2. Maximum Principle for Switched Markov Processes

Now we return to the setting of switched Markov process (X_t, Λ_t). Let D be a bounded open set in $\mathbb{R}^d$ and $U = D \times \mathcal{M}$. Then $\tau_D = \inf\{t > 0 : X_t \notin D\}$ is the same as $\tau_U := \inf\{t > 0 : Y_t := (X_t, \Lambda_t) \notin U\}$. Suppose u is a $\mathcal{G}$-harmonic function in U. Under some mild assumptions (for example, when u is bounded and continuous up to $\partial D \times \mathcal{M}$), we have

$$u(x,i) = \mathbb{E}_{x,i}[u\,(X_{\tau_D}, \Lambda_{\tau_D})] \qquad \text{for } (x,i) \in U. \tag{3.12}$$

It follows immediately that if $u \geq 0$ on U^c, then $u \geq 0$ in U.

To proceed, we recall the notion of irreducibility of the generator $\mathcal{G}$ or the matrix function $Q(\cdot)$. The operator $\mathcal{G}$ or the matrix function $Q(\cdot)$ is said to be *irreducible on D* if for any $i, j \in \mathcal{M}$, there exist $n = n(i,j) \geq 1$ and $\Lambda_0, \ldots, \Lambda_n \in \mathcal{M}$ with $\Lambda_{k-1} \neq \Lambda_k$ for $1 \leq k \leq n$, $\Lambda_0 = i, \Lambda_n = j$ such that $\{x \in D : q_{\Lambda_{k-1}\Lambda_k}(x) > 0\}$ has positive Lebesgue measure for $k = 1, \ldots, n$.

For each $i \in \mathcal{M}$, denote by X^i the jump diffusion that solves the martingale problem $(\mathcal{L}_i, C_b^2(\mathbb{R}^d))$ and $\widetilde{X}^i$ the subprocess of X^i killed at rate $-q_{ii}(x)$. For a connected open set $D \subset \mathbb{R}^d$, G_D^i denotes the Green operator of $\widetilde{X}^i$ in D.

Theorem 3.4 *Assume that conditions* (A1)-(A3) *hold, that D is a bounded connected open set in $\mathbb{R}^d$, and that Q is irreducible on D. Suppose that u is a $\mathcal{G}$-harmonic function in $U = D \times \mathcal{M}$ given by*

$$u(x,i) = \mathbb{E}_{x,i}[\phi\,(X_{\tau_D}, \Lambda_{\tau_D})\,; \tau_D < \infty] \qquad \textit{for } (x,i) \in U$$

and $\phi \geq 0$ on $D^c \times \mathcal{M}$. Then either $u(x,i) > 0$ for every $(x,i) \in U$ or $u \equiv 0$ on U.

Proof: Clearly $u \geq 0$ on U. Suppose that u is not a.e. zero on U. Without loss of generality, let us assume that $\{x \in D : u(x,1) > 0\}$ has positive Lebesgue measure. Denote by $\tau_1 := \inf\{t > 0 : \Lambda_t \neq \Lambda_0\}$ the first switching time for $Y_t = (X_t, \Lambda_t)$. Let

$$v_i(x) := v(x,i) := \mathbb{E}_{x,i}[\phi(\widetilde{X}^i_{\tau_D}, i)] = \mathbb{E}_{x,i}[\phi(X^i_{\tau_D}, i); \tau_D < \tau_1].$$

Then v_i is a harmonic function of $\mathcal{L}_i + q_{ii}$ in D with $v_i = \phi(\cdot, i)$ on D^c. For $1 \leq i \leq m$, using the strong Markov property τ_1, we have

$$u(x,i) = v_i(x) + \sum_{\substack{j=1 \\ j \neq i}}^{m} G_D^i(q_{ij} u(\cdot, j))(x). \tag{3.13}$$

Under the above assumption, either $\{x \in D : v_1(x) > 0\}$ or $\{x \in D : G_D^1(\sum_{\substack{j=1 \\ j\neq i}}^m q_{ij} u(\cdot, j))(x) > 0\}$ has positive Lebesgue measure. If the latter happens, then by Proposition 3.3, $G_D^1(\sum_{\substack{j=1 \\ j\neq i}}^m q_{ij} u(\cdot, j))(x) > 0$ and hence $u(x,1) > 0$ for every $x \in D$. Note that

$$v_i(x) = \mathbb{E}_x[e_{-q_{ii}}(\tau_D)\phi(X^i(\tau_D), i)] \leq \mathbb{E}_x[\phi(X^i(\tau_D), i)] =: \widetilde{u}_i(x). \tag{3.14}$$

Suppose $|\{x \in D : v_1(x) > 0\}| > 0$. Then so does $A := \{x \in D : \widetilde{u}_1(x) > 0\}$. For any $x_0 \in D$ and $r \in (0,1)$ so that $B(x_0, r) \subset D$ and $B(x_0, r/2) \cap A$ has positive Lebesgue measure, let $K \subset B(x_0, r/2) \cap A$ be a compact set having positive Lebesgue measure. By (3.7), $\mathbb{P}_x(\sigma_K^1 < \tau^1_{B(x_0,r))}) > 0$ for every $x \in B(x_0, r/2)$, where $\sigma_K^1 := \inf\{t \geq 0 : X_t^1 \in K\}$ and $\tau^1_{B(x_0,r)} := \inf\{t \geq 0 : X_t^1 \notin B(x_0, r)\}$. Hence for every $x \in B(x_0, r/2)$, by the strong Markov property of X^i at σ_K^1,

$$\widetilde{u}_1(x) \geq \mathbb{E}_x\left[\widetilde{u}_1(X^i_{\sigma^1_K}); \sigma_K^1 < \tau_D\right] > 0.$$

Consequently, $B(x_0, r/2) \subset A$. By the chaining argument, the same reasoning as above leads to $A = D$; that is, $\widetilde{u}_1(x) > 0$ on D. By the probabilistic representation (3.14) of v_1, we have $v_1(x) > 0$ on D and hence $u(x, 1) > 0$ on D. Thus we have shown that $u(x, 1) > 0$ on D whenever $\{x \in D : u(x, 1) > 0\}$ has positive Lebesgue measure.

For $i \neq 1$, there is a self-avoiding path $i = i_0, \ldots, i_n = i$ so that $\{x \in D : q_{i_{k-1}i_k}(x) > 0\}$ has positive Lebesgue measure for each $k = 1, \ldots, n$. By (3.13) and its iteration, we have

$$\begin{aligned}
u(x,i) &= v_i(x) + \sum_{k=1}^{n} \sum_{\substack{l_1,\ldots,l_k=1 \\ l_1\neq 1, l_2\neq l_1,\ldots,l_k\neq l_{k-1}}}^{m} G_D^i(q_{il_1}(G_D^{l_1} q_{l_1 l_2}(\cdots(G_D^{l_{k-1}} q_{l_{k-1}l_k} v_{l_k}))\cdots)))(x) \\
&\quad + \sum_{\substack{l_1,\ldots,l_n=1 \\ l_1\neq 1, l_2\neq l_1,\ldots,l_n\neq l_{n-1}}}^{m} G_D^i(q_{il_1}(G_D^{l_1} q_{l_1 l_2}(\cdots(G_D^{l_{n-1}} q_{l_{n-1}l_n} u(\cdot, l_n))\cdots)))(x) \\
&\geq G_D^i(q_{ii_1}(G_D^{i_1} q_{i_1 i_2}(\cdots(G_D^{i_{n-1}} q_{i_{n-1}i_n} u(\cdot, i))\cdots)))(x),
\end{aligned}$$

which is strictly positive in D by Proposition 3.3.

Next assume that $u = 0$ a.e. on U. We claim that $u \equiv 0$ on U. In view of (3.13) and Proposition 3.3, it suffices to show that $v_i(x) \equiv 0$ on D for every $i \in \mathcal{M}$. Since

$$v_i(x) = \mathbb{E}_{x,i}\left[\phi(X^i_{\tau_D}, i); \tau_D < \tau_1\right] = \mathbb{E}_{x,i}\left[e^{(i)}_{-q_{ii}}(\tau_D)\phi(X^i_{\tau_D}, i)\right],$$

and $v_i(x) = 0$ a.e. on D, where

$$e^{(i)}_{-q_{ii}}(t) := \exp\left(\int_0^t q_{ii}(X_s^i)ds\right),$$

we have $u_i(x) := \mathbb{E}_{x,i}\left[\phi(X^i_{\tau_D}, i)\right]$ vanishes a.e. on D. The function $u_i(x)$ is harmonic in D with respect to X^i (or equivalently, with respect to the operator $\mathcal{L}_i$ in D). By [16, Theorem 2.3], it is Hölder continuous in D. Hence $u_i(x) = 0$ for every $x \in D$, and so is $v_i(x)$. This proves that $u(x, i) = 0$ for every $x \in D$ and every $i \in \mathcal{M}$. □

Theorem 3.5 (Strong Maximum Principle I) *Assume conditions* (A1)-(A3) *hold, D is a bounded connected open set in $\mathbb{R}^d$, and $Q(x)$ is irreducible on D. Suppose u is a $\mathcal{G}$-harmonic function in $U = D \times \mathcal{M}$ given by*

$$u(x,i) = \mathbb{E}_{x,i}[\phi(X_{\tau_D}, \Lambda_{\tau_D})] \qquad \textit{for } (x,i) \in U$$

for some ϕ with $M := \sup_{(y,j)\in D^c\times\mathcal{M}} \phi(y,j) \in [0,\infty)$. If $(x_0,i_0)\in D\times\mathcal{M}$ and $u(x_0,i_0) = M$, then $u\equiv M$ on $D\times\mathcal{M}$. If in addition $M>0$, then the matrix $Q(x)$ is Markovian.

Proof: (i) First we assume that $Q(x)$ is Markovian in the sense that $\sum_{j\in\mathcal{M}} q_{ij}(x) = 0$ a.e. on $\mathbb{R}^d$ for every $i\in\mathcal{M}$. In this case, by the construction of the switched Markov process (X_t,Λ_t) outlined in Remark 2.2(a), (X_t,Λ_t) has infinite lifetime and so constant 1 is a $\mathcal{G}$-harmonic function on $\mathbb{R}^d$. Hence

$$M - u(x,i) = \mathbb{E}_{x,i}\left[(M-\phi)(X_{\tau_D},\Lambda_{\tau_D}); \tau_D<\infty\right]$$

is a non-negative $\mathcal{G}$-harmonic function in $D\times\mathcal{M}$. (Note that $\tau_D<\infty$ $\mathbb{P}_{x,i}$-a.s. in view of Proposition 4.4 below.) Since $u(x_0,i_0)=M$, we have by Theorem 3.4 that $u(x,i)=M$ for every $(x,i)\in U$.

(ii) We now consider the general case that $Q(x)$ is a sub-Markovian matrix. Define a Markovian matrix $\bar{Q}(x) = (\bar{q}_{ij}(x))$ by taking $\bar{q}_{ij}(x) = q_{ij}(x)$ and $\bar{q}_{ii}(x) = -\sum_{j\in\mathcal{M}\setminus\{i\}} q_{ij}(x)$. Let $(\bar{X}_t,\bar{\Lambda}_t)$ be the conservative switched Markov process corresponding to $\bar{\mathcal{G}}$ as in (2.1) but with $\bar{Q}(x)$ in place of $Q(x)$. The original switched Markov process (X_t,Λ_t) can be viewed as a subprocess of $\left(\bar{X}_t,\bar{\Lambda}_t\right)$ killed at rate $\kappa(x,i) = \bar{q}_{ii}(x) - q_{ii}(x)$; that is, for every $\psi(x,i)\ge 0$ on $\mathbb{R}^d\times\mathcal{M}$,

$$\mathbb{E}_{x,i}[\psi(X_t,\Lambda_t)] = \mathbb{E}_{x,i}\left[\bar{e}_\kappa(t)\psi(\bar{X}_t,\bar{\Lambda}_t)\right],$$

where $\bar{e}_\kappa(t) = \exp(-\int_0^t \kappa(\bar{X}_s,\bar{\Lambda}_s)ds)$. We consider

$$v(x,i) := \mathbb{E}_{x,i}\left[(M-\phi)(X_{\tau_D},\Lambda_{\tau_D}); \tau_D<\infty\right], \tag{3.15}$$

which is a non-negative $\mathcal{G}$-harmonic function in $D\times\mathcal{M}$. We can rewrite $v(x,i)$ on U as

$$v(x,i) = \mathbb{E}_{x,i}[\bar{e}_\kappa(\tau_D)(M-\phi)(\bar{X}_{\tau_D},\bar{\Lambda}_{\tau_D})]. \tag{3.16}$$

Since $u(x_0,i_0) = M\ge 0$,

$$0\le v(x_0,i_0) = M\mathbb{P}_{x_0,i_0}(\tau_D<\infty) - u(x_0,i_0)\le 0,$$

that is, $v(x_0,i_0)=0$. Thus by Theorem 3.4, $v(x,i)\equiv 0$ on $D\times\mathcal{M}$. This implies by (3.16) that $(M-\phi)(\bar{X}_{\tau_D},\bar{\Lambda}_{\tau_D}) = 0$ $\mathbb{P}_{x,i}$-a.s. for every $(x,i)\in D\times\mathcal{M}$. Consequently, we have

$$u(x,i) = M\mathbb{E}_{x,i}\left[\bar{e}_\kappa(\tau_D)\right] \quad \text{for } (x,i)\in D\times\mathcal{M} \tag{3.17}$$

and so

$$\begin{aligned} u(x,i) &= M + M\mathbb{E}_{x,i}[\bar{e}_\kappa(\tau_D)-1)] \\ &= M - \mathbb{E}_{x,i}\left[\int_0^{\tau_D}\kappa(\bar{X}_s,\bar{\Lambda}_s)\exp\left(-\int_s^{\tau_D}\kappa(\bar{X}_r,\bar{\Lambda}_r)dr\right)ds\right] \\ &= M - M\mathbb{E}_{x,i}\left[\int_0^{\tau_D}\kappa(\bar{X}_s,\bar{\Lambda}_s)ds\right]. \end{aligned}$$

Let $\tau_1 := \inf\{t > 0 : \bar{\Lambda}_t \neq \bar{\Lambda}_0\}$ and denote by $\bar{G}^i_D$ be the Green function of $\mathcal{L}_i + \bar{q}_{ii}$ in D. Since $u(x_0, i_0) = M$, we have by the strong Markov property and the construction of $(\bar{X}_t, \bar{\Lambda}_t)$ in Remark 2.2(a),

$$\begin{aligned} 0 &= \mathbb{E}_{x_0,i_0}\left[\int_0^{\tau_D} \kappa(\bar{X}_s, \bar{\Lambda}_s)ds\right] \\ &\geq \mathbb{E}_{x_0,i_0}\left[\int_0^{\tau_D\wedge\tau_1} \kappa(\bar{X}_s, i_0)ds\right] = \bar{G}^{i_0}_D(\kappa(\cdot, i_0))(x_0) \geq 0. \end{aligned} \tag{3.18}$$

Thus $\bar{G}^{i_0}_D(\kappa(\cdot, i_0))(x_0) = 0$ and so by Proposition 3.3 we have $\kappa(x, i_0) = 0$ a.e. on D. Observe that

$$\begin{aligned} u(x_0, i_0) &= \mathbb{E}_{x_0,i_0}[\bar{e}_\kappa(\tau_D)\phi(\bar{X}_{\tau_D}, \bar{\Lambda}_{\tau_D})] \\ &= \mathbb{E}_{x_0,i_0}[\phi(\bar{X}_{\tau_D}, \bar{\Lambda}_{\tau_D}); \tau_D < \tau_1] + \mathbb{E}_{x_0,i_0}[\bar{e}_\kappa(\tau_D)\phi(\bar{X}_{\tau_D}, \bar{\Lambda}_{\tau_D}); \tau_1 \leq \tau_D]. \end{aligned}$$

Hence using the strong Markov property, we have

$$\begin{aligned} 0 &= (M - u)(x_0, i_0) = \mathbb{E}_{x_0,i_0}[(M - M\bar{e}_\kappa(\tau_D))] \\ &= \mathbb{E}_{x_0,i_0}[(M - M\bar{e}_\kappa(\tau_D)); \tau_1 \leq \tau_D] \\ &= \mathbb{E}_{x_0,i_0}\left[(M - u)(\bar{X}_{\tau_1}, \bar{\Lambda}_{\tau_1}); \tau_1 \leq \tau_D\right] \\ &= \sum_{j\in\mathcal{M}\setminus\{i_0\}} \mathbb{E}_{x_0,i_0}\left[(M - u)(\bar{X}_{\tau_1-}, j))(q_{i_0j}/\bar{q}_{i_0i_0})(\bar{X}_{\tau_1-}); \tau_1 \leq \tau_D\right] \\ &= \sum_{j\in\mathcal{M}\setminus\{i_0\}} \bar{G}^{i_0}_D(q_{i_0j}(M - u)(x_0). \end{aligned}$$

By Proposition 3.3 again, we have $\sum_{j\in\mathcal{M}\setminus\{i_0\}} q_{i_0j}(M - u) = 0$ a.e. on D. Since $\bar{Q}$ is irreducible on D, for any $j \neq i_0$, there is a self-avoiding path $\{j_0 = i_0, j_1, \ldots, j_n = j\}$ so that $\{x \in D : q_{j_kj_{k+1}}(x) > 0\}$ having positive Lebesgue measure for $k = 0.1, \ldots, n-1$. Thus we have $u(x, j_1) = M$ on $\{x \in D : q_{i_0j_1}(x) > 0\}$. By the argument above, this implies that $\kappa(x, j_1) = 0$ a.e. on D. Continuing as this, we get $\kappa(x, j_k) = 0$ a.e. on D and $\{x \in D : u(x, j_k) = M\}$ has positive Lebesgue measure for $k = 2, \ldots, n$. This proves that $\kappa(x, i) = 0$ a.e. on D for every $i \in \mathcal{M}$ and so $u(x, i) = M$ for every $x \in D$ in view of (3.17). □

Before presenting the next version of strong maximum principle, we first prepare a lemma.

Lemma 3.6 *Assume conditions* (A1)-(A3) *hold, D is a bounded connected open set in $\mathbb{R}^d$, and $Q(x)$ is irreducible on D. For any $\phi \geq 0$ on $D\times\mathcal{M}$, either $\mathbb{E}_{x,i}\int_0^{\tau_D} \phi(X_s, \Lambda_s)ds > 0$ for every $(x, i) \in D \times \mathcal{M}$ or $\mathbb{E}_{x,i}\int_0^{\tau_D} \phi(X_s, \Lambda_s)ds \equiv 0$ on $D \times \mathcal{M}$.*

Proof: Denote by G^i_D the Green function of $\mathcal{L}_i + \bar{q}_{ii}$ in D. Using the strong Markov property at the first switching time $\tau_1 := \inf\{t \geq 0 : \Lambda_t \neq \Lambda_{t-}\}$ in a similar way to that for (3.18), we have for every $(x,i)\in D\times\mathcal{M}$,

$$\begin{aligned} v(x,i) &:= \mathbb{E}_{x,i}\int_0^{\tau_D}\phi(X_s,\Lambda_s)ds \\ &= \mathbb{E}_{x,i}\int_0^{\tau_D\wedge\tau_1}\phi(X_s,\Lambda_s)ds+\mathbb{E}_{x,i}\left[\int_{\tau_1}^{\tau_D\wedge\tau_1}\phi(X_s,\Lambda_s)ds;\tau_1<\tau_D\right] \\ &= G_D^i(\phi(\cdot,i))(x)+\mathbb{E}_{x,i}\left[v(X_{\tau_1},\Lambda_{\tau_1});\tau_1<\tau_D\right] \\ &= G_D^i(\phi(\cdot,i))(x)+\sum_{k\in\mathcal{M}\setminus\{i\}}\mathbb{E}_{x,i}\left[v(X_{\tau_1-},k)(q_{ik}/q_{ii})(X_{\tau_1-});\tau_1<\tau_D\right] \\ &= G_D^i(\phi(\cdot,i))(x)+\sum_{k\in\mathcal{M}\setminus\{i\}}G_D^i\left(q_{ik}v(\cdot,k)\right)(x), \end{aligned} \tag{3.19}$$

where the last identity is due to [28, p.286]; see the proof of [13, Proposition 2.2].

Suppose $v(x_0,i_0)=0$ for some $(x_0,i_0)\in D\times\mathcal{M}$. Then by Proposition 3.3, $v(x,i_0)\equiv 0$ on D. For any $j\in\mathcal{M}\setminus\{i_0\}$, since $Q(x)$ is irreducible on D, there is a self-avoiding path $\{j_0=i_0,j_1,\dots,j_n=j\}$ so that $\{x\in D: q_{j_kj_{k+1}}(x)>0\}$ having positive Lebesgue measure for $k=0,1,\dots,n-1$. It follows from (3.19) and its iteration that

$$0=v(x_0,i_0)\geq G_D^{i_0}(q_{i_0j_1}(G_D^{j_1}q_{j_1j_2}(\cdots(G_D^{j_{n-1}}(q_{j_{n-1}j}v(\cdot,j))\cdots)))(x_0)\geq 0.$$

We conclude from Proposition 3.3 that $q_{j_{n-1}j}(\cdot)v(\cdot,j)=0$ a.e. on D. So there is some $y\in D$ so that $v(y,j)=0$. By (3.19) with (y,j) in place of (x,i) and Proposition 3.3, we have $v(x,j)=0$ for every $x\in D$. □

Theorem 3.7 (Strong Maximum Principle II) *Suppose that conditions* (A1)-(A3) *hold, D is a bounded connected open set in $\mathbb{R}^d$ and $Q(x)$ is irreducible on D. If $f(\cdot,i)\in C^2(D)$, $\sup_{\mathbb{R}^d\times\mathcal{M}}f\geq 0$, and*

$$\mathcal{G}f(x,i)\geq 0 \quad \textit{for } (x,i)\in D\times\mathcal{M},$$

then $f(x,i)$ can not attain its maximum inside $D\times\mathcal{M}$ unless

$$f(x,i)\equiv\sup_{\mathbb{R}^d\times\mathcal{M}}f(y,j) \quad \textit{on } D\times\mathcal{M}.$$

Proof: Suppose f achieves its maximum at some $(x_0,i_0)\in D\times\mathcal{M}$. Let D_1 be any relatively compact connected open subset of D that contains x_0 and that $Q(x)$ is irreducible on D_1. Then by Itô's formula, we have for every $(x,j)\in D_1\times\mathcal{M}$,

$$\begin{aligned} f(x,j) &= \mathbb{E}_{x,j}\left[f(X_{\tau_{D_1}},\Lambda_{\tau_{D_1}})\right]-\mathbb{E}_{x,j}\int_0^{\tau_{D_1}}\mathcal{G}f(X_s,\Lambda_s)ds \\ &\leq \mathbb{E}_{x,j}\left[f(X_{\tau_{D_1}},\Lambda_{\tau_{D_1}}\right]=:h(x,j). \end{aligned} \tag{3.20}$$

Let $M=\sup_{(y,j)\in\mathbb{R}^d\times\mathcal{M}}f(y,j)$, which is non-negative. In view of (3.20),

$$M=\sup_{(y,j)\in D_1^c\times\mathcal{M}}f(y,j)=f(x_0,i_0).$$

Clearly, $h \le M$ and h is $\mathcal{G}$-harmonic in $D_1 \times \mathcal{M}$. We have by (3.19), $h(x_0, i_0) = M$ and

$$\mathbb{E}_{x_0,i_0} \int_0^{\tau_{D_1}} \mathcal{G} f(X_s, \Lambda_s) ds = 0.$$

Theorem 3.5 and Lemma 3.6 tell us that $h \equiv M$ on $D_1 \times \mathcal{M}$ and $\mathbb{E}_{x,i} \int_0^{\tau_{D_1}} \mathcal{G} f(X_s, \Lambda_s) ds = 0$ for every $(x, i) \in D_1 \times \mathcal{M}$. Consequently, $f(x, i) \equiv M$ on $D_1 \times \mathcal{M}$. Letting D_1 increase to D establishes the theorem. □

4. Harnack Inequality

This section is devoted to the Harnack inequality for $\mathcal{G}$-harmonic functions. For simplicity, we introduce some notation as follows. For any $U = D \times \mathcal{M} \subset \mathbb{R}^d \times \mathcal{M}$, recall that

$$\tau_D = \inf\{t \ge 0 : X_t \notin D\}.$$

We define

$$T_D^i := \inf\{t \ge 0 : X_t \in D, \Lambda_t = i\}, \quad i \in \mathcal{M}.$$

Proposition 4.1 *Assume conditions* (A1)-(A3) *hold. There exists a constant* C_2 *not depending on* $x_0 \in \mathbb{R}^d$ *such that for any* $r \in (0, 1)$ *and any* $i \in \mathcal{M}$,

$$\mathbb{P}_{x_0,i} \left(\tau_{B(x_0,r)} \le C_2 r^2\right) \le 1/2. \tag{4.1}$$

Proof: Let $v(\cdot, i) \in C^2(\mathbb{R}^d)$ be a nonnegative function independent of i and

$$v(x, i) = \begin{cases} |x - x_0|^2, & |x - x_0| \le r/2, \\ r^2, & |x - x_0| \ge r \end{cases}$$

such that v is bounded by $c_1 r^2$, and its first and second order derivatives are bounded by $c_1 r$ and c_1, respectively. Since $\mathbb{P}_{x_0,i}$ solves the martingale problem, we have

$$\mathbb{E}_{x_0,i} v(X_{t\wedge\tau_{B(x_0,r)}}, \Lambda_{t\wedge\tau_{B(x_0,r)}}) = v(x_0, i) + \mathbb{E}_{x_0,i} \int_0^{t\wedge\tau_{B(x_0,r)}} \mathcal{G} v(X_s, \Lambda_s) ds.$$

Using the boundedness of the first and second derivatives of $v(\cdot, i)$ and $Q(\cdot)$, we have

$$\int_0^{t\wedge\tau_{B(x_0,r)}} \mathcal{G} v(X_s, \Lambda_s) ds \le c_2 t.$$

It follows that

$$\mathbb{E}_{x_0,i} v(X_{t\wedge\tau_{B(x_0,r)}}, \Lambda_{t\wedge\tau_{B(x_0,r)}}) - v(x_0, i) \le c_2 t.$$

On the other hand, since $v(X_{\tau_{B(x_0,r)}}, \Lambda_{\tau_{B(x_0,r)}}) = r^2$, we obtain

$$\mathbb{E}_{x_0,i} v(X_{t\wedge\tau_{B(x_0,r)}}, \Lambda_{t\wedge\tau_{B(x_0,r)}}) \ge r^2 \mathbb{P}_{x_0,i}(\tau_{B(x_0,r)} \le t).$$

Hence

$$r^2 \mathbb{P}_{x_0,i}(\tau_{B(x_0,r)} \le t) \le c_2 t.$$

Taking $C_2 = \frac{1}{2c_2}$ in the above formula and replacing t by $C_2 r^2$, we obtain (4.1). □

Proposition 4.2 *Assume conditions* (A1)-(A3) *hold. For any constant* $\varepsilon \in (0,1)$*, there exist positive constants* C_3 *and* C_4 *depending only on* ε *such that for any* $(x_0, i) \in \mathbb{R}^d \times \mathcal{M}$ *and any* $r \in (0,1)$*, we have*

(a) $\mathbb{P}_{x,i}\left(\tau_{B(x_0,r)} > C_3 r^2\right) \geq 1/2$ *for* $(x,i) \in B(x_0, (1-\varepsilon)r) \times \mathcal{M}$.
(b) $\mathbb{E}_{x,i}\tau_{B(x_0,r)} \geq C_4 r^2$ *for* $(x,i) \in B(x_0, (1-\varepsilon)r) \times \mathcal{M}$.

PROOF: By Proposition 4.1, there exists a constant c_1 depending only on ε such that for any $(x,i) \in B(x_0, (1-\varepsilon)r) \times \mathcal{M}$, we have

$$\mathbb{P}_{x,i}(\tau_{B(x_0,r)} \leq c_1 r^2) \leq \mathbb{P}_{x,i}(\tau_{B(x,\varepsilon r)} \leq c_1 r^2) \leq 1/2,$$

which implies (a). Hence

$$\mathbb{E}_{x,i}\tau_{B(x_0,r)} \geq c_1 r^2 \mathbb{P}_{x,i}(\tau_{B(x_0,r)} > c_1 r^2) \geq c_1 r^2/2.$$

Then (b) follows. □

For a measure μ on $\mathbb{R}^d$ and $y \in \mathbb{R}^d$, we use $\mu(dx - y)$ to denote the measure $\nu(dx)$ defined by $\nu(A) := \mu(A-y)$ for $A \in \mathcal{B}(\mathbb{R}^d)$, where $A-y := \{x-y : x \in A\}$. We know how the switched Markov processes jumps at the switched times between different plates. The following describes how the switched Markov process (X_s, Λ_s) jumps at non-switching times.

Proposition 4.3 *Assume conditions* (A1)-(A3) *hold. Suppose* A *and* B *are two bounded open subsets of* $\mathbb{R}^d$ *having a positive distance apart and* $i_0 \in \mathcal{M}$*. Then*

$$\sum_{s \leq t} \mathbf{1}_{\{X_{s-} \in A, X_s \in B, \Lambda_s = i_0\}} - \int_0^t \mathbf{1}_A(X_s)\, \mathbf{1}_{\{i_0\}}(\Lambda_s)\, \pi_{\Lambda_s}(X_s, B - X_s) ds \tag{4.2}$$

is a $\mathbb{P}_{x,i}$*-martingale for each* $(x,i) \in \mathbb{R}^d \times \mathcal{M}$.

PROOF: Let A_1 be a bounded open subset of $\mathbb{R}^d$ so that $\overline{A} \subset A_1 \subset \overline{A}_1 \subset B^c$. Let $v(\cdot, j) \equiv 0$ for all $j \neq i_0$, and $v(\cdot, i_0) \in C_b^2(\mathbb{R}^d)$ so that $v(x, i_0) = 0$ on A_1 and $v(x, i_0) = 1$ on B. Fix $(x,i) \in \mathbb{R}^d \times \mathcal{M}$. Note that

$$M^v(t) := v(X_t, \Lambda_t) - v(X_0, \Lambda_0) - \int_0^t \mathcal{G}v(X_s, \Lambda_s) ds$$

is a $\mathbb{P}_{x,i}$-martingale, so is $\int_0^t \mathbf{1}_A(X_{s-}) dM^v(s)$. Define $\tau_0 = 0$, $\tau_1 = \inf\{t \geq 0 : X_t \in A\}$, $\tau_2 = \inf\{t \geq \tau_1 : X_t \in A_1^c\}$, and for $k \geq 2$,

$$\tau_{2k-1} = \inf\{t \geq \tau_{2(k-1)} : X_t \in A\}, \qquad \tau_{2k} = \inf\{t \geq \tau_{2k-1} : X_t \in A_1^c\}.$$

Note that $v(X_t,\Lambda_t)=0$ for $t\in\cup_{k\geq 1}[\tau_{2k-1},\tau_{2k})$ and $\mathbf{1}_A(X_{t-})=0$ for $t\in\cup_{k\geq 1}[\tau_{2(k-1)},\tau_{2k-1})$. Thus the Riemann sum approximation of stochastic integral yields that

$$\begin{aligned}
&\int_0^t \mathbf{1}_A\left(X_{s-}\right)dM^v(s)\\
&=\sum_{k=1}^{\infty}\mathbf{1}_A(X_{\tau_{2k}\wedge t-})\Big(v(X_{\tau_{2k}\wedge t},\Lambda_{\tau_{2k}\wedge t})-v(X_{\tau_{2k}\wedge t-},\Lambda_{\tau_{2k}\wedge t-})\Big)\\
&\quad-\int_0^t \mathbf{1}_A\left(X_{s-}\right)\mathcal{G}v(X_s,\Lambda_s)ds\\
&=\sum_{s\leq t}\mathbf{1}_A\left(X_{s-}\right)\left[v(X_s,\Lambda_s)-v(X_{s-},\Lambda_{s-})\right]\\
&\quad-\int_0^t \mathbf{1}_A\left(X_s\right)\mathcal{G}v(X_s,\Lambda_s)ds.
\end{aligned}$$

Since $v(y,j)=0$ on $A_1\times\mathcal{M}$, we have

$$\mathcal{G}v(y,j)=\int_{\mathbb{R}^d}v(y+z,j)\pi_j(y,dz)=\int_{\mathbb{R}^d}v(z,j)\pi_j(y,dz-y)$$

for every $(y,j)\in A_1\times\mathcal{M}$. Therefore,

$$\begin{aligned}
&\sum_{s\leq t}\mathbf{1}_A\left(X_{s-}\right)\left[v(X_s,\Lambda_s)-v(X_{s-},\Lambda_{s-})\right]\\
&\qquad-\int_0^t \mathbf{1}_A\left(X_s\right)\int_{\mathbb{R}^d}v(z,\Lambda_s)\pi_{\Lambda_s}(X_s,dz-X_s)ds \ \text{ is a } \mathbb{P}_{x,i}\text{-martingale.}
\end{aligned}$$

Because A and B are a positive distance from each other, the sum on the left of the above formula is in fact a finite sum. With these facts we can pass to the limit to conclude that

$$\begin{aligned}
&\sum_{s\leq t}\mathbf{1}_A\left(X_{s-}\right)\left[\mathbf{1}_{B\times\{i_0\}}\left(X_s,\Lambda_s\right)-\mathbf{1}_{B\times\{i_0\}}\left(X_{s-},\Lambda_{s-}\right)\right]\\
&\qquad-\int_0^t \mathbf{1}_A\left(X_s\right)\int_{\mathbb{R}^d}\mathbf{1}_{B\times\{i_0\}}\left(z,\Lambda_s\right)\pi_{\Lambda_s}(X_s,dz-X_s)ds \ \text{ is a } \mathbb{P}_{x,i}\text{- martingale,}
\end{aligned}$$

which implies

$$\sum_{s\leq t}\mathbf{1}_{\{X_{s-}\in A,X_s\in B,\Lambda_s=i_0\}}-\int_0^t \mathbf{1}_A\left(X_s\right)\mathbf{1}_{\{\Lambda(s)=i_0\}}\pi_{\Lambda_s}(X_s,B-X_s)ds$$

is a $\mathbb{P}_{x,i}$-martingale. □

Proposition 4.4 *There exist $\widetilde{r}_0\in(0,1/2]$ and $C_5>0$, depending only on κ_0 and κ_1 in (A2)- (A3) and an upper bound on $\sum_{k=1}^m\|q_{kk}\|_\infty$, such that for any $x_0\in\mathbb{R}^d$ and any $r\in(0,\widetilde{r}_0)$, we have*

$$\sup_{(x,i)\in B(x_0,r)\times\mathcal{M}}\mathbb{E}_{x,i}\tau_{B(x_0,r)}\leq C_5r^2. \tag{4.3}$$

Proof: Let $u(x) \in C^2(\mathbb{R}^d)$ be a convex function in x with values in $[0,10]$ and increase with respect to $|x|$ such that

$$u(x) = |x|^2, \quad |x| \le 2.$$

Let $\widetilde{r}_0 \in (0,1/2)$ be sufficiently small. For $x_0 \in \mathbb{R}^d$ and $r \in (0, \widetilde{r}_0)$, let $v(x,i) = u(\frac{x-x_0}{r})$. Then for any $(x,i) \in B(x_0,r) \times \mathcal{M}$, since $v(\cdot,\cdot)$ is bounded between 0 and 10 and $Q(\cdot)$ is bounded, there exists $c_1 > 0$ such that

$$Q(x)v(x,\cdot)(i) \ge -c_1. \tag{4.4}$$

Moreover,

$$\begin{aligned}\mathcal{L}^{(c)}v(x,i) &:= \sum_{k,l=1}^{d} a_{kl}(x,i)\frac{\partial^2 v(x,i)}{\partial x_k \partial x_l} + \sum_{k=1}^{d} b_k(x,i)\frac{\partial v(x,i)}{\partial x_k} \\ &= \sum_{k=1}^{d} 2a_{kk}(x,i)r^{-2} + \sum_{k=1}^{d} 2b_k(x,i)(x_k - x_{0,k})r^{-2} \\ &\ge c_2 r^{-2} - c_3 r^{-1} \ge c_4 r^{-2},\end{aligned} \tag{4.5}$$

provided $\widetilde{r}_0$ is small enough. Define

$$\mathcal{L}^{(j)}v(x,i) = \int_{\mathbb{R}^d} \big[f(x+z,i) - f(x,i) - \nabla f(x,i)\cdot z\mathbf{1}_{\{|z|<1\}}\big]\pi_i(x,dz).$$

We break $\mathcal{L}^{(j)}v(x,i)$ into two parts, $|z| \le 1$ and $|z| > 1$, respectively. For the first part, by the convexity of $u(x)$, we deduce

$$\int_{|z|\le 1} [v(x+z,i) - v(x,i) - \nabla v(x,i)\cdot z]\pi_i(x,dz) \ge 0. \tag{4.6}$$

For the second part with $|z| > 1$, since $r < \frac{1}{2}$, we know $x + z \notin B(x_0,r)$ for any $x \in B(x_0,r)$. Then we have $v(x+z,i) \ge 1$ and $v(x,i) \le 1$, it follows that

$$\int_{|z|>1} [v(x+z,i) - v(x,i)]\pi_i(x,dz) \ge 0. \tag{4.7}$$

Since $\mathbb{P}_{x,i}$ solves the martingale problem, together with (4.4), (4.5), (4.6), and (4.7), we deduce that

$$\begin{aligned}&\mathbb{E}_{x,i}v\big(X_{t\wedge\tau_{B(x_0,r)}}, \Lambda_{t\wedge\tau_{B(x_0,r)}}\big) - v(x,i) \\ &\quad= \mathbb{E}_{x,i}\int_0^{t\wedge\tau_{B(x_0,r)}} \mathcal{G}v(X_s,\Lambda_s)ds \ge c_5 r^{-2}\mathbb{E}_{x,i}(t\wedge\tau_{B(x_0,r)}).\end{aligned} \tag{4.8}$$

By the definition of $v(x,i)$,

$$\mathbb{E}_{x,i}v(X_{t\wedge\tau_{B(x_0,r)}}, \Lambda_{t\wedge\tau_{B(x_0,r)}}) - v(x,i) \le 10. \tag{4.9}$$

It follows from (4.8) and (4.9) that

$$c_5 r^{-2}\mathbb{E}_{x,i}(t\wedge\tau_{B(x_0,r)})\le 10.$$

The conclusion follows by letting $t\to\infty$. □

In the remaining of this section, we assume that conditions (A1)-(A4) hold.

Definition 4.5 The generator $\mathcal{G}$ or the matrix function $Q(\cdot)$ is said to be *strictly irreducible* on D if for any $i,j\in\mathcal{M}$ and $i\neq j$, there exists $q_{ij}^0>0$ such that $\inf_{x\in D} q_{ij}(x)\ge q_{ij}^0$.

Proposition 4.6 *Assume conditions* (A1)-(A4) *hold. Let $x_0\in\mathbb{R}^d$ and $r\in(0,\widetilde{r}_0)$, where $\widetilde{r}_0$ is the constant in Proposition 4.4. Suppose that the operator $\mathcal{G}$ is strictly irreducible on $B(x_0,r)$. Let $H:\mathbb{R}^d\times\mathcal{M}\mapsto\mathbb{R}$ be a bounded non-negative function supported in $B(x_0,2r)^c\times\mathcal{M}$. Then there exists a constant $C_6>0$, depending only on κ_0 and κ_1 of (A2)-(A3), an upper bound on $\sum_{k=1}^m\|q_{kk}\|_\infty$ and on $\{q_{ij}^0; i\neq j\in\mathcal{M}\}$ in the definition of strict irreducibility of $\mathcal{G}$, such that for any $x,y\in B(x_0,r/2)$ and any $i\in\mathcal{M}$,*

$$\mathbb{E}_{x,i}H\left(X_{\tau_{B(x_0,r)}},\Lambda_{\tau_{B(x_0,r)}}\right)\le C_6\alpha_{2r}\mathbb{E}_{y,i}H\left(X_{\tau_{B(x_0,r)}},\Lambda_{\tau_{B(x_0,r)}}\right).$$

Here α_r is the constant appeared in condition (A4).

Proof: Denote $B=B(x_0,r)$. Define

$$u(x,i)=\mathbb{E}_{x,i}H\left(X_{\tau_B},\Lambda_{\tau_B}\right)\quad\text{for}\quad (x,i)\in B\times\mathcal{M}.$$

Since $H=0$ on $B(x_0,2r)\times\mathcal{M}$, we have by using the Lévy system formula of $Y=(X,\Lambda)$ given by Proposition 4.3 that

$$\begin{aligned} u(x,i) &= \mathbb{E}_{x,i}\left[H\left(X_{\tau_B},\Lambda_{\tau_B}\right); X_{\tau_B-}\in\overline{B}; X_{\tau_B}\in B(x_0,2r)^c\right]\\ &= \mathbb{E}_{x,i}\left[\int_0^{\tau_B}\int_{B(x_0,2r)^c} H\left(z,\Lambda_s\right)\widetilde{\pi}_{\Lambda_s}\left(X_s,z-X_s\right)dzds\right]. \end{aligned} \tag{4.10}$$

We deduce that

$$\begin{aligned} u(x,i) &\le \mathbb{E}_{x,i}\left[\int_0^{\tau_B}\int_{B(x_0,2r)^c}\sum_{j=1}^m H(z,j)\widetilde{\pi}_j\left(X_s,z-X_s\right)dzds\right]\\ &\le \Big(\sum_{j=1}^m\int_{B(x_0,2r)^c} H(z,j)\sup_{w\in B}\widetilde{\pi}_j(w,z-w)dz\Big)\mathbb{E}_{x,i}\tau_B\\ &\le c_1r^2\sum_{j=1}^m\int_{B(x_0,2r)^c} H(z,j)\sup_{w\in B}\widetilde{\pi}_j(w,z-w)dz\\ &= c_1r^2M, \end{aligned}$$

where the last inequality is a consequence of Proposition 4.4, and

$$M=\sum_{j=1}^m M_j,\quad M_j=\int_{B(x_0,2r)^c} H(z,j)\sup_{w\in B}\widetilde{\pi}_j(w,z-w)dz.$$

Thus,

$$u(x,i) \le c_1 r^2 M \quad \text{for } (x,i) \in B \times \mathcal{M}. \tag{4.11}$$

Let $\tau_1 = \inf\{t > 0 : \Lambda_t \ne \Lambda_0\}$ and denote the Green operator of $\mathcal{L}_i + q_{ii}$ in B by G_B^i. Define

$$h_i(x) = \mathbb{E}_{x,i}[H\left(X_{\tau_B}, \Lambda_{\tau_B}\right); \tau_B < \tau_1], \quad (x,i) \in B \times \mathcal{M}.$$

By the strong Markov property of (X_t, Λ_t), we have

$$u(x,i) = h_i(x) + \sum_{k \ne i} G_B^i\Big(q_{ik}(\cdot)u(\cdot,k)\Big)(x). \tag{4.12}$$

By (4.11) and the fact that $||q_{ik}||_\infty = \sup\limits_{x\in\mathbb{R}^d} |q_{ik}(x)| < \infty$, we arrive at

$$\begin{aligned} \sum_{k\ne i} G_B^i\Big(q_{ik}(\cdot)u(\cdot,k)\Big)(x) &\le \sum_{k\ne i} c_1 r^2 M ||q_{ik}||_\infty \mathbb{E}_{x,i}\tau_B \\ &\le \sum_{k\ne i} c_1 r^2 M ||q_{ik}||_\infty c_1 r^2 = c_2 M r^4. \end{aligned}$$

Combining above estimates, we obtain

$$u(x,i) \le h_i(x) + c_2 M r^4 \quad \text{for} \quad (x,i) \in B \times \mathcal{M}. \tag{4.13}$$

Next, we drive an lower bound for $\mathbb{E}_{x,i}(\tau_B \wedge \tau_1)$ for $(x,i) \in B(x_0, r/2) \times \mathcal{M}$. By Proposition 4.2, there exists $c_3 > 0$ such that

$$\mathbb{P}_{x,i}\left(\tau_B > c_3 r^2\right) \ge \frac{1}{2}.$$

It follows that

$$\begin{aligned} \mathbb{P}_{x,i}(\tau_B \wedge \tau_1 \ge c_3 r^2) &\ge \mathbb{P}_{x,i}\left(\tau_B > c_3 r^2 \text{ and } \tau_1 \ge c_3 r^2\right) \\ &\ge \exp\left(-||q_{kk}||_\infty c_3 r^2\right) \mathbb{P}_{x,i}\left(\tau_B > c_3 r^2\right) \\ &\ge \frac{1}{2}\exp\left(-||q_{kk}||_\infty c_3 \widetilde{r}_0^2\right) =: c_4. \end{aligned}$$

Then we obtain

$$\mathbb{E}_{x,i}\left(\tau_B \wedge \tau_1\right) \ge c_3 c_4 r^2 \quad \text{for } (x,i) \in B(x_0, r/2) \times \mathcal{M}. \tag{4.14}$$

By assumption (A4), for any $z \in B(x_0, 2r)^c$,

$$\sup_{w\in B} \widetilde{\pi}_j\left(w, z-w\right) \le \alpha_{2r} \inf_{w\in B} \widetilde{\pi}_j\left(w, z-w\right).$$

By this inequality and (4.14), we have by using Lévy system formula of $Y = (X, \Lambda)$,

$$\begin{aligned} h_i(x) &= \mathbb{E}_{x,i}\left[\int_0^{\tau_B\wedge\tau_1}\int_{B(x_0,2r)^c} H\left(z,\Lambda_s\right)\widetilde{\pi}_{\Lambda_s}\left(X_s, z-X_s\right)dzds\right] \\ &\ge \mathbb{E}_{x,i}\left[\int_0^{\tau_B\wedge\tau_1}\int_{B(x_0,2r)^c} H(z,i)\inf_{w\in B}\widetilde{\pi}_i\left(w, z-w\right)dzds\right] \\ &\ge \alpha_{2r}^{-1} M_i \mathbb{E}_{x,i}(\tau_B \wedge \tau_1) \\ &\ge c_5 \alpha_{2r}^{-1} M_i r^2 \quad \text{for} \quad x \in B(x_0, r/2). \end{aligned} \tag{4.15}$$

On the other hand,

$$\begin{aligned} h_i(x) &= \mathbb{E}_{x,i}\left[\int_0^{\tau_B\wedge\tau_1}\int_{B(x_0,2r)^c} H\left(z,\Lambda_s\right)\widetilde{\pi}_{\Lambda_s}\left(X_s, z-X_s\right)dzds\right] \\ &\le \mathbb{E}_{x,i}\left[\int_0^{\tau_B\wedge\tau_1}\int_{B(x_0,2r)^c} H(z,i)\sup_{w\in B}\widetilde{\pi}_i\left(w,z-w\right)dzds\right] \\ &\le M_i\mathbb{E}_{x,i}\tau_B \\ &\le c_6 M_i r^2 \quad \text{for} \quad x\in B(x_0,r/2). \end{aligned} \tag{4.16}$$

Note that for $i\neq k$, $\inf_{x\in\mathbb{R}^d} q_{ik}(x)\ge q_{ik}^0>0$. By (4.12) and (4.15), we have

$$\begin{aligned} \sum_{k\neq i} G_B^i\Big(q_{ik}(\cdot)u(\cdot,k)\Big)(x) &\ge \sum_{k\neq i} c_5\alpha_{2r}^{-1}M_k r^2 G_B^i\Big(q_{ik}(\cdot)\mathbf{1}_{B(x_0,3r/4)}(\cdot)\Big)(x) \\ &\ge \sum_{k\neq i} c_5\alpha_{2r}^{-1}M_k r^2 q_{ik}^0\mathbb{E}_{x,i}\left(\tau_{B(x_0,3r/4)}\wedge\tau_1\right) \\ &\ge \sum_{k\neq i} c_5\alpha_{2r}^{-1}M_k r^2 q_{ik}^0 c_7 r^2 \\ &= c_8\alpha_{2r}^{-1}r^4\sum_{k\neq i}M_k \quad \text{for } x\in B(x_0,r/2). \end{aligned} \tag{4.17}$$

In the above, we used the fact that $\mathbb{E}_{x,i}\left(\tau_{B(x_0,3r/4)}\wedge\tau_1\right)\ge c_7r^2$. This can be derived in the same way as that of (4.14). By (4.12), (4.13), (4.15), (4.16), and (4.17), for any $x,y\in B(x_0,r/2)$ and $i\in\mathcal{M}$, we have

$$\begin{aligned} u(y,i) &= h_i(y)+\sum_{k\neq i}G_B^i\Big(q_{ik}(\cdot)u(\cdot,k)\Big)(y) \\ &\ge c_5\alpha_{2r}^{-1}M_ir^2+c_8\alpha_{2r}^{-1}r^4\sum_{k\neq i}M_k \\ &\ge c_9\alpha_{2r}^{-1}\Big(M_ir^2+r^4\sum_{k\neq i}M_k\Big) \\ &\ge c_{10}\alpha_{2r}^{-1}\Big(M_ir^2+r^4M\Big)\ge c_{11}\alpha_{2r}^{-1}u(x,i). \end{aligned}$$

The proof of the proposition is complete. □

Theorem 4.7 *Assume conditions* (A1)-(A4) *hold. Let $D\subset\mathbb{R}^d$ be a bounded connected open set and $\mathcal{K}$ be a compact set in $D\subset\mathbb{R}^d$. Suppose that $\mathcal{G}$ is strictly irreducible on D. Then there exists $C_7>0$ which depends only on $D,\mathcal{K}$ and operator $\mathcal{G}$ such that if $f(\cdot,\cdot)$ is a nonnegative, bounded function in $\mathbb{R}^d\times\mathcal{M}$ that is $\mathcal{G}$-harmonic in $D\times\mathcal{M}$, we have*

$$f(x,i)\le C_7 f(y,j) \quad \textit{for } x,y\in\mathcal{K} \textit{ and } i,j\in\mathcal{M}. \tag{4.18}$$

Proof: We first show that for each fixed ball $B(x_0,4R)\subset D$ with $R<\frac{1}{8}\wedge\widetilde{r}_0$ (where $\widetilde{r}_0$ is given in Proposition 4.4), there exists a constant $C>0$ that depends only on $R\wedge 1$,

κ_0 and κ_1 of (A2)-(A3), an upper bound on $\sum_{k=1}^m \|q_{kk}\|_\infty$ and on $\{q^0_{ij}; i \neq j \in \mathcal{M}\}$ in the definition of strict irreducibility of $\mathcal{G}$ on $B(0,4R)$, such that for any nonnegative, bounded and $\mathcal{G}$-harmonic function $f(\cdot,\cdot)$ in $B(x_0,4R)\times\mathcal{M}$, we have

$$f(x,i) \le Cf(y,j) \quad \text{for } x,y \in B(x_0,R) \text{ and } i,j\in\mathcal{M}. \tag{4.19}$$

By looking at $f+\varepsilon$ and sending ε to 0, we may suppose that f is bounded below by a positive constant. By looking at $af(x,i_0)$ for a suitable constant a if needed, we may assume that $\inf\limits_{(x,i)\in B(x_0,R)\times\mathcal{M}} f(x,i) = 1/2$.

(a) Let us recall several results. Let $r < \widetilde{r}_0 < 1/2$. By Proposition 3.1, there exits a constant $c_1 > 0$ such that for any $x \in B(x_0, 3R/2)$ and any $i\in\mathcal{M}$,

$$\mathbb{P}_{\overline{x}_i,i}\left(T^i_{B(x,r/2)} < \tau_{B(x_0,4R)}\right) \ge c_1 r^6. \tag{4.20}$$

By Proposition 3.2, there exists a nondecreasing function $\Phi : (0,\infty) \mapsto (0,\infty)$ such that if A is a Borel subset of $B(x,r)$ and $|A|/r^d \ge \rho$ for a given ρ, then for any $(y,i) \in B(x,r)\times\mathcal{M}$ and $r \in (0,\widetilde{r}_0)$,

$$\mathbb{P}_{y,i}(T^i_A < \tau_{B(x,2r)}) \ge \frac{1}{2}\Phi\left(|A|/r^d\right). \tag{4.21}$$

By Proposition 4.6 and H being a nonnegative function supported on $B(x,2r)^c$, for any $y,z\in B(x,r/2)$ and $i\in\mathcal{M}$,

$$\mathbb{E}_{y,i}H(X_{\tau_{B(x,r)}},\Lambda_{\tau_{B(x,r)}}) \le c_2\alpha_{2r}\mathbb{E}_{z,i}H(X_{\tau_{B(x,r)}},\Lambda_{\tau_{B(x,r)}}). \tag{4.22}$$

To proceed, we first consider the case that

$$\inf_{x\in B(x_0,2R)} f(x,i) < 1 \quad \text{for each} \quad i\in\mathcal{M}. \tag{4.23}$$

Thus, there exists $\{\overline{x}_i\}_{i\in\mathcal{M}}$ such that

$$\overline{x}_i \in B(x_0,2R) \quad \text{and} \quad f(\overline{x}_i,i) < 1. \tag{4.24}$$

(b) For $n \ge 1$, let

$$r_n = c_3 R/n^2,$$

where c_3 is a positive number such that $\sum_{n=1}^\infty r_n < R/4$ and $r_n \in (0,\widetilde{r}_0)$ for all n, that is,

$$c_3 < \frac{1}{4\sum_{n=1}^\infty 1/n^2} \wedge \frac{\widetilde{r}_0}{R}. \tag{4.25}$$

In particular, it implies $r_n < R/4$. Let ξ, c_4, c_5 be positive constants to be chosen later. Once these constants have been chosen, we can take N_1 large enough so that

$$\xi N_1 \exp(c_4 n) c_5 r_n^{6+\beta} \ge 2\kappa_2 \quad \text{for all} \quad n = 1,2,\ldots \tag{4.26}$$

The constants κ_2 and β are taken from assumption (A4). Such a choice is possible since $c_4 > 0$ and $r_n = c_3 R/n^2$. Suppose that there exists $(x_1, i_1) \in B(x_0, R) \times \mathcal{M}$ with $f(x_1, i_1) = N_1$ for N_1 chosen above. We will show that in this case there exists a sequence $\{(x_k, i_k) : k \geq 1\}$ with

$$\begin{aligned} &(x_{k+1}, i_{k+1}) \in B(x_k, 2r_k) \times \mathcal{M} \subset B(x_0, 3R/2) \times \mathcal{M}, \\ &N_{k+1} := f(x_{k+1}, i_{k+1}) \geq N_1 \exp\left(c_4(k+1)\right). \end{aligned} \tag{4.27}$$

(c) Suppose that we already have $\{(x_k, i_k); 1 \leq k \leq n\}$ so that (4.27) is satisfied for $k = 1, \dots, n-1$. Define

$$A_n = \Big\{ y \in B(x_n, r_n/2) : f(y, i_n) \geq \xi N_n r_n^\beta / \kappa_2 \Big\}.$$

We claim that

$$\frac{|A_n|}{|B(x_n, r_n/2)|} \leq \frac{1}{4}. \tag{4.28}$$

Suppose on the contrary, $\frac{|A_n|}{|B(x_n, r_n/2)|} > 1/4$. Let F be a compact subset of A_n such that $\frac{|F|}{|B(x_n, r_n/2)|} > 1/4$. Then $F \subset B(x_0, 2R)$. By (4.20),

$$\mathbb{P}_{\overline{x}_{i_n}, i_n}\left(T^{i_n}_{B(x_n, r_n/2)} < \tau_{B(x_0, 4R)} \right) \geq c_1 r_n^6,$$

where c_1 is independent of x_n and r_n. By the strong Markov property of (X_t, Λ_t), we have

$$\begin{aligned} \mathbb{P}_{\overline{x}_{i_n}, i_n}\left(T_F^{i_n} < \tau_{B(x_0, 4R)} \right) &\geq \mathbb{E}_{\overline{x}_{i_n}, i_n}\left[\mathbb{P}_{X_{T^{i_n}_{B(x_n, r_n/2)}}, i_n}\left(T_F^{i_n} < \tau_{B(x_n, r_n)} \right); T^{i_n}_{B(x_n, r_n/2)} < \tau_{B(x_0, 4R)} \right] \\ &\geq \frac{1}{2} \Phi\left(\frac{2^d |F|}{r_n^d} \right) \mathbb{P}_{\overline{x}_{i_n}, i_n}\left(T^{i_n}_{B(x_n, r_n/2)} < \tau_{B(x_0, 4R)} \right) \\ &\geq \frac{1}{2} \Phi\left(\frac{\alpha(d)}{4} \right) c_1 r_n^6, \end{aligned}$$

where $\alpha(d)$ is the volume of the unit ball in $\mathbb{R}^d$.

We take $c_5 = \Phi\left(\alpha(d)/4\right) c_1/2$. By the definition of $\mathcal{G}$-harmonicity and the above estimates, we obtain

$$\begin{aligned} 1 > f(\overline{x}_{i_n}, i_n) &\geq \mathbb{E}_{\overline{x}_{i_n}, i_n}[f(X_{T_F^{i_n} \wedge \tau_{B(x_0, 4R)}}, \Lambda_{T_F^{i_n} \wedge \tau_{B(x_0, 4R)}}); T_F^{i_n} < \tau_{B(x_0, 4R)}] \\ &\geq \frac{\xi N_n r_n^\beta}{\kappa_2} \mathbb{P}_{\overline{x}_{i_n}, i_n}(T_F^{i_n} < \tau_{B(x_0, 4R)}) \\ &\geq \frac{\xi N_n r_n^{\beta+6} c_5}{\kappa_2} \\ &\geq 2, \end{aligned} \tag{4.29}$$

which is a contradiction. Note that the last inequality follows from $N_n \geq N_1 \exp(c_4 n)$ and our choice of N_1 given by (4.26). Thus, (4.28) is valid. Therefore, there is a compact subset $\widetilde{F}$ of $B(x_n, r_n/2) \setminus A_n$ such that $|\widetilde{F}| \geq \frac{1}{2}|B(x_n, r_n/2)|$. By the definition of $\widetilde{F}$ and A_n,

$$f(x, i_n) < \frac{\xi N_n r_n^\beta}{\kappa_2} \quad \text{for} \quad x \in \widetilde{F}.$$

Denote $\tau_{r_n} := \tau_{B(x_n, r_n)}$, $p_n := \mathbb{P}_{x_n, i_n}(T_{\widetilde{F}}^{i_n} < \tau_{r_n})$ and $M_n := \sup_{(y,j)\in B(x_n, 2r_n)\times\mathcal{M}} f(y, j)$. Since $|\widetilde{F}| \geq \frac{1}{2}|B(x_n, r_n/2)|$, using (4.21), we obtain

$$p_n \geq \frac{1}{2}\Phi\left(\frac{\alpha(d)}{2^{d+1}}\right) := c_6 \quad \text{for} \quad n = 1, 2, \ldots \tag{4.30}$$

By the definition of $\mathcal{G}$-harmonic function and the right continuity of the sample paths of (X_t, Λ_t), we have

$$\begin{aligned} N_n &= f(x_n, i_n) = \mathbb{E}_{x_n, i_n}[f(X_{T_{\widetilde{F}}^{i_n}}, \Lambda_{T_{\widetilde{F}}^{i_n}}) : T_{\widetilde{F}}^{i_n} < \tau_{r_n}] \\ &\quad + \mathbb{E}_{x_n, i_n}[f(X_{\tau_{r_n}}, \Lambda_{\tau_{r_n}}) : X(\tau_{r_n}) \in B(x_n, 2r_n), \tau_{r_n} < T_{\widetilde{F}}^{i_n}] \\ &\quad + \mathbb{E}_{x_n, i_n}[f(X_{\tau_{r_n}}, \Lambda_{\tau_{r_n}}) : X(\tau_{r_n}) \notin B(x_n, 2r_n), \tau_{r_n} < T_{\widetilde{F}}^{i_n}] \\ &\leq \frac{\xi N_n r_n^\beta}{\kappa_2} + M_n(1 - p_n) \\ &\quad + \mathbb{E}_{x_n, i_n}[f(X_{\tau_{r_n}}, \Lambda_{\tau_{r_n}}) : X_{\tau_{r_n}} \notin B(x_n, 2r_n), \tau_{r_n} < T_{\widetilde{F}}^{i_n}]. \end{aligned} \tag{4.31}$$

Take a point $y_n \in \widetilde{F}$. Then $f(y_n, i_n) < \dfrac{\xi N_n r_n^\beta}{\kappa_2}$. We then deduce from (4.22) that

$$\begin{aligned} \frac{\xi N_n r_n^\beta}{\kappa_2} &> f(y_n, i_n) \\ &\geq \mathbb{E}_{y_n, i_n}[f(X_{\tau_{r_n}}, \Lambda_{\tau_{r_n}}) : X_{\tau_{r_n}} \notin B(x_n, 2r_n)] \\ &\geq \frac{1}{c_2 \alpha_{2r_n}} \mathbb{E}_{x_n, i_n}[f(X_{\tau_{r_n}}, \Lambda_{\tau_{r_n}}) : X_{\tau_{r_n}} \notin B(x_n, 2r_n)]. \end{aligned}$$

It follows that

$$\begin{aligned} \mathbb{E}_{x_n, i_n}[f(X_{\tau_{r_n}}, \Lambda_{\tau_{r_n}}) : X_{\tau_{r_n}} \notin B(x_n, 2r_n)] &\leq \frac{\xi N_n r_n^\beta c_2 \alpha_{2r_n}}{\kappa_2} \\ &\leq \frac{\xi c_2}{2^\beta} N_n, \end{aligned}$$

where the last inequality is obtained by noting that $\alpha_{2r_n} \leq \kappa_2 (2r_n)^{-\beta}$. Hence by (4.31),

$$N_n \leq \Big(\frac{\xi}{\kappa_2} + \frac{\xi c_2}{2^\beta}\Big) N_n + M_n(1 - p_n). \tag{4.32}$$

Denote $\eta = 1 - \Big(\dfrac{\xi}{\kappa_2} + \dfrac{\xi c_2}{2^\beta}\Big)$. Let $\xi > 0$ be sufficiently small such that $\dfrac{\eta}{1 - c_6} > 3/2$. By (4.30) and (4.32), $M_n/N_n > 3/2$. Using the definition of M_n, there is $(x_{n+1}, i_{n+1}) \in$

$B(x_n, 2r_n) \times \mathcal{M}$ so that

$$N_{n+1} := f(x_{n+1}, i_{n+1}) \ \geq 3N_n/2.$$

We take $c_4 = \ln(3/2)$. Then (4.27) holds for $k = n$. By induction, we have constructed a sequence of points $\{(x_k, i_k)\}$ such that (4.27) holds for all $k \geq 1$. It can be seen that $N_k \to \infty$ as $k \to \infty$, a contradiction to the assumption that f is bounded. Thus, for a positive constant N_1 sufficiently large such that (4.26) holds, we have

$$f(x,i) < N_1 \quad \text{for all} \quad (x,i) \in B(x_0, R) \times \mathcal{M}.$$

Since $\inf_{(x,i)\in B(x_0,R)\times\mathcal{M}} f(x,i) = 1/2$, we arrive at

$$f(x,i) \leq 2N_1 f(y,j), \quad x, y \in B(x_0, R) \text{ and } i, j \in \mathcal{M}. \tag{4.33}$$

For any compact set $\mathcal{K} \subset D$, we use a standard finite ball covering argument. Since $\mathcal{K}$ is compact, there exists a finite number of points $z_k \in \mathcal{K}$, $k = 1, 2, \ldots, n$ such that

$$\mathcal{K} \subset \bigcup_{k=1}^{n} B(z_i, R) \subset D,$$

and $|z_k - z_{k-1}| < R/2$. Let $x, y \in \mathcal{K}$ and $i, j \in \mathcal{M}$. Applying Harnack inequality (4.33) at most $n+1$ times, we obtain $f(x,i) \leq (2N_1)^{n+1} f(y,j)$.

Now we suppose that (4.23) is invalid. Then there exists $i \in \mathcal{M}$ such that $f(x,i) \geq 1$ for all $x \in B(x_0, 2R)$. Set

$$K_i := \inf_{x\in B(x_0,2R)} f(x,i), \quad K := \sup_{i\in\mathcal{M}} K_i, \quad g(x,i) := f(x,i)/3K.$$

It follows that (4.23) holds with g in place of f. Moreover, if $i_0 \in \mathcal{M}$ and $K = K_{i_0}$, then

$$g(x, i_0) \geq 1/3 \quad \text{for} \quad x \in B(x_0, 2R).$$

By the same argument as in Theorem 3.4, there is a constant $c_7 > 0$ such that

$$\begin{aligned} g(x,i) \ &\geq G^i_{B(x_0,2R)}\Big(q_{ii_0}(\cdot) g(\cdot, i_0)\Big)(x) \\ &\geq c_7, \end{aligned}$$

for all $(x,i) \in B(x_0, R) \times \mathcal{M}$, where $G^i_{B(x_0,2R)}$ is the Green operator of $\widetilde{X}^i$ in $B(x_0, 2R)$. Note also that $\inf_{x\in D} q_{ii_0}(x) \geq q^0_{ii_0} > 0$. The Harnack inequalities for g, and for f can be established similarly as in the previous case. □

5. Further Remarks

This paper has been devoted to switching jump diffusions. Important properties such as maximum principle and Harnack inequality have been obtained. The utility and applications of these results will be given in a subsequent paper [7] for obtaining recurrence and ergodicity of switching jump diffusions. The ergodicity can be used in a wide variety of control and optimization problems with average cost per unit time objective functions (see also various variants of the long-run average cost problems in [18]), in which the instantaneous measures are replaced by the corresponding ergodic measures.

We note that the references [8, 9, 10, 12, 26] are devoted to regularity for the parabolic functions of non-local operators (on each of the parallel plane). The results obtained in this paper should be useful when one considers regularity of the coupled systems or switched jump-diffusions.

Acknowledgements

The research of X. Chen was partially supported by NNSF of China (No.11601163) and NSF of Guangdong Province of China (No. 2016A030313448); the research of Z.-Q. Chen, was partially supported by NSF grant DMS-1206276; the research of K. Tran, was partially supported by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant 101.03-2017.23; the research of G. Yin was partially supported by the National Science Foundation under grant DMS-1710827.

References

[1] Arapostathis, A., Ghosh, M.K., and Marcus, S.I. (1999), Harnack's inequality for cooperative weakly coupled elliptic systems, *Comm. Partial Differential Equations*, **24**, 1555–1571.

[2] Athreya, S. and Ramachandran, K. (2016) Harnack inequality for non-local Schrödinger operators, to appear in *Potential Analysis.*

[3] Bass, R.F. and Kassmann, M. (2005) Harnack inequalities for non-local operators of variable order, *Trans. Amer. Math. Soc.*, **357**, 837–850.

[4] Bass, R.F., Kassmann, M. and Kumagai, T. (2010) Symmetric jump processes: Localization, heat kernels and convergence, *Ann. Inst. Henri Poincaré Probab. Stat.*, **46**, 59–71.

[5] Bass, R.F. and Levin, D.A. (2002) Harnack inequalities for jump processes, *Potential Anal.*, **17**, 375–388.

[6] Caffarelli, L. and Silvestre, L. (2009) Regularity theory for fully nonlinear integro-differential equations, *Comm. Pure Appl. Math.* **62**, 597–638.

[7] Chen, X., Chen, Z.-Q., Tran, K. and Yin, G. (2017) Recurrence and ergodicity for a class of regime-switching jump diffusions, to appear in *Appl. Math. Optim.*

[8] Chen, Z.-Q. and Kumagai, T. (2003) Heat kernel estimates for stable-like processes on d-sets. *Stochastic Process Appl.* **108**, 27-62.

[9] Chen, Z.-Q. and Kumagai, T. (2008) Heat kernel estimates for jump processes of mixed types on metric measure spaces. *Probab. Theory Relat. Fields* **140**, 277-317.

[10] Chen, Z.-Q. and Kumagai, T. (2010) A priori Hölder estimate, parabolic harnack principle and heat kernel estimates for diffusions with jumps, *Revista Matematica Iberoamericana*, **26**, 551–589.

[11] Chen, Z.-Q., Hu, E., Xie, L. and Zhang, X. (2016) Heat kernels for non-symmetric diffusions operators with jumps, *J. Differential Equations* **263** (2017), 6576-6634.

[12] Chen, Z.-Q., Wang, H. and Xiong, J. (2012) Interacting superprocesses with discontinuous spatial motion, *Forum Math.*, **24**, 1183–1223.

[13] Chen, Z.-Q. and Zhao, Z. (1996) Potential theory for elliptic systems, *Ann. Probab.*, **24**, 293–319.

[14] ——, (1997) Harnack principle for weakly coupled elliptic systems, *J. Differential Equations*, **139**, 261–282.

[15] Evans, L.C. (2010) *Partial differential equations*, Amer. Math. Soc., Providence, RI, Second Ed.

[16] Foondun, M. (2009) Harmonic functions for a class of integro-differential operators, *Potential Anal.*, **31**, 21–44.

[17] Ikeda, N., Nagasawa, M. and Watanabe, S. (1966) A construction of Markov process by piecing out, *Proc. Japan Academy*, **42**, 370–375.

[18] Jasso-Fuentes, H., Yin, G., (2013) *Advanced Criteria for Controlled Markov-Modulated Diffusions in an Infinite Horizon: Overtaking, Bias, and Blackwell Optimality*, Science Press, Beijing, China.

[19] Krylov, N.V. (1987) *Nonlinear Elliptic and Parabolic Equations of the Second Order*, D. Reidel Publishing Co., Dordrecht, Translated from the Russian by P.L. Buzytsky.

[20] Komatsu, T. (1973) Markov processes associated with certain integro-differential, *Osaka J. Math.*, **10**, 271-303.

[21] Kushner, H. (1990) *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, Birkhäuser, Boston.

[22] Liu, R. (2016) Optimal stopping of switching diffusions with state dependent switching rates, *Stochastics*, **88**, 586-605.

[23] Mao, X. and Yuan, C. (2006) *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, London, UK.

[24] Meyer, P. (1975) Renaissance, recollements, mélanges, relentissement de processus de markov, *Ann. Inst. Fourier*, **25**, 465–497.

[25] Mikulevicius, R. and Pragarauskas, H. (1988) On Hölder continuity of solutions of certain integro-differential equations. *Ann. Acad. Scien. Fenn. Ser. A. I. Math.*, **13**, 231-238.

[26] Negoro, A. and Tsuchiya, M. (1989) Stochastic processes and semigroups associate

with degenerate Lévy generating operators, *Stochastics Stochastic Reports*, **26**, 29–61.

[27] Protter, M.H. and Weinberger, H.F. (1967) *Maximum Principles in Differential Equations*, Prentice-Hall, Inc., Englewood Cliffs, N.J.

[28] Sharpe, M. (1986) *General Theory of Markov Processes*, Academic, New York, 1986.

[29] Song, R. and Vondraček, Z. (2005) Harnack inequality for some discontinuous Markov processes with a diffusion part, *Glas. Mat. Ser. III*, **40(60)**, 177–187.

[30] Wang, J. (2014) Martingale problems for switched processes, *Mathematische Nachrichten*, **287**, 1186–1201.

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

[32] Yin, G. and Zhu, C. (2010) *Hybrid Switching Diffusions: Properties and Applications*, Springer, New York.