Integration by Parts Formula and Applications for SDEs with Lévy Noise

Feng-Yu Wang
School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, Beijing 100875, P. R. China

May 7, 2014

Abstract

By using the Malliavin calculus and finite jump approximations, the Driver-type integration by parts formula is established for the semigroup associated to stochastic differential equations with noises containing a subordinate Brownian motion. As applications, the shift Harnack inequality and heat kernel estimates are derived. The main results are illustrated by SDEs driven by $\alpha$-stable like processes.

AMS subject Classification: 60J75, 47G20, 60G52.
Keywords: Integration by parts formula, shift Harnack inequality, heat kernel, stochastic differential equation.

1 Introduction

A significant application of the Malliavin calculus is to describe the density of a Wiener functional using the integration by parts formula. In 1997, Driver [3] established the following integration by parts formula for the heat semigroup $P_t$ on a compact Riemannian manifold $M$:

$$P_t(\nabla Z f) = \mathbb{E}\{f(X_t)N_i\}, \quad f \in C^1(M), \ Z \in \mathcal{X},$$

where $\mathcal{X}$ is the set of all smooth vector fields on $M$, and $N_i$ is a random variable depending on $Z$ and the curvature tensor. From this formula we are able to characterize the derivative w.r.t. the second variable $y$ of the heat kernel $p_t(x,y)$, see [10] for a recent study on

*Supported in part by NNSFC(11131003), SRFDP, the Fundamental Research Funds for the Central Universities.
integration by parts formulas and applications for stochastic differential equations driven by Wiener processes. The backward coupling method developed in [10] has been also used in [4, 14] for SDEs driven by fractional Brownian motions and SPDEs driven by Wiener processes. The purpose of this paper is to investigate the integration by parts formula and applications for SDEs driven by purely jump Lévy noises, in particular, to derive estimates on the heat kernel and its derivatives for the solutions.

Consider the following stochastic equation on $\mathbb{R}^d$:

\begin{equation}
X_t = X_0 + \int_0^t b_s(X_s)ds + \int_0^t \sigma_s dW_{S(s)} + V_t, \quad t \geq 0,
\end{equation}

where

$\sigma : [0, \infty) \to \mathbb{R}^d \otimes \mathbb{R}^d$, \quad $b : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$

are measurable and locally bounded, $W := (W_t)_{t \geq 0}$, $S := (S(t))_{t \geq 0}$ and $V := (V_t)_{t \geq 0}$ are independent stochastic processes such that

(i) $W$ is the Brownian motion on $\mathbb{R}^d$ with $W_0 = 0$;
(ii) $V$ is a cádlág process on $\mathbb{R}^d$ with $V_0 = 0$;
(iii) $S$ is the subordinator induced by a Bernstein function $B$, i.e. $S$ is a one-dimensional increasing Lévy process with $S(0) = 0$ and Laplace transform

$\mathbb{E} e^{-rS(t)} = e^{-tB(r)}, \quad t, r \geq 0$.

Then $(W_{S(t)})_{t \geq 0}$ is a the Lévy process known as the subordinate Brownian motion (or subordinated process of the Brownian motion) with subordinator $S$, see e.g. [1, 6].

For this equation the Bismut formula and Harnack inequalities have been studied in [15] and [11] by using regularization approximations of $S(t)$, but the study of the integration by parts formula and shift Harnack inequality is not yet done.

To establish the integration by parts formula, we need the following assumptions.

(H1) $b_t \in C^2(\mathbb{R}^d)$ such that for some increasing $K_1, K_2 \in C([0, \infty))$,

$\|\nabla b_t\|_{\infty} := \sup_{x \in \mathbb{R}^d} \|\nabla b_t(x)\| \leq K_1(t), \quad \|\nabla^2 b_t\|_{\infty} := \sup_{x \in \mathbb{R}^d} \|\nabla^2 b_t(x)\| \leq K_2(t), \quad t \geq 0,$

where $\|\cdot\|$ is the operator norm.

(H2) $\sigma_t$ is invertible such that for some increasing $\lambda_1, \lambda_2 \in C([0, \infty))$,

$\|\sigma_t\| \leq \lambda_1(t), \quad \|\sigma_t^{-1}\| \leq \lambda_2(t), \quad t \geq 0.$

By (H1) and (H2), for any $x \in \mathbb{R}^d$ the equation (1.1) with $X_0 = x$ has a unique solution $X_t(x)$. Let $P_t$ be the associated Markov operator, i.e.

$P_tf(x) = \mathbb{E}f(X_t(x)), \quad f \in \mathcal{B}_b(\mathbb{R}^d), t \geq 0, x \in \mathbb{R}^d$.  

As already observed in [10] that comparing with the Bismut formula, the integration by parts formula is usually harder to establish. To strengthen this observation, we explain below that the regularization argument used in [15] for the Bismut formula is no longer valid for the integration by parts formula. For simplicity, let us consider the case that \( V_t = 0, b_t = b \) and \( \sigma_t = \sigma \). As in [15], for any \( \varepsilon > 0 \) let

\[
S_\varepsilon(t) = \frac{1}{\varepsilon} \int_t^{t+\varepsilon} S(s) ds + \varepsilon t, \quad t \geq 0.
\]

Then \( S_\varepsilon(\cdot) \) is differentiable and \( S_\varepsilon \downarrow S \) as \( \varepsilon \downarrow 0 \). Consider the equation (note that we have assumed \( V_t = 0, b_t = b \) and \( \sigma_t = \sigma \))

\[
dX_\varepsilon^t = b(X_\varepsilon^t)dt + \sigma dW_{S_\varepsilon(t)}, \quad X_\varepsilon^0 = X_0.
\]

To apply the existing derivative formulas for SDEs driven by the Brownian motion, we take \( Y_\varepsilon^t = X_{S_\varepsilon^{-1}(t)} \) so that this equation reduces to

\[
dY_\varepsilon^t = b(Y_\varepsilon^t)(S_\varepsilon^{-1})'(t)dt + \sigma dW_t, \quad Y_\varepsilon^0 = X_0.
\]

In [15], by using a known Bismut formula for \( Y_\varepsilon^t \) and letting \( \varepsilon \to 0 \), the corresponding formula for \( X_t \) is established. The crucial point for this argument is that the Bismut formula for \( Y_\varepsilon^t \) converges as \( \varepsilon \to 0 \). However, since \( S \) is not differentiable, the existing integration by parts formula of \( Y_\varepsilon^t \) (see e.g. [10, Theorem 5.1] with \( H = \mathbb{R}^d \) and \( A = 0 \))

\[
\mathbb{E}(\nabla_v f)(Y_T) = \frac{1}{T} \mathbb{E}\left\{ f(Y_T^\varepsilon) \int_0^T \langle \sigma^{-1} (v - t(S_\varepsilon^{-1})'(t)) \nabla_v b(Y_\varepsilon^t), dW_t \rangle \right\}
\]

does not converge to any explicit formula as \( \varepsilon \to 0 \), except \( \nabla_v b \) is trivial.

So, to establish the integration by parts formula, we will take a different approximation argument, i.e. the finite jump approximation used in [12] to establish the Bismut formula for SDEs with multiplicative Lévy noises. We have to indicate that in this paper we are not able to establish the integration by parts formula for SDEs with multiplicative Lévy noises. Note that even for SDEs driven by multiplicative Gaussian noises, the existing integration by parts formula using the Malliavin covariant matrix is in general less explicit.

To state our main result, we introduce the \( \mathbb{R}^d \otimes \mathbb{R}^d \)-valued process \( J_t \), which solves the ODE with random coefficients

\[
\frac{d}{dt} J_t = (\nabla b_t)(X_t) J_t, \quad J_0 = I,
\]

where for \( t \geq 0 \) and \( x \in \mathbb{R}^d, (\nabla b_t)(x) \in \mathbb{R}^d \otimes \mathbb{R}^d \) is determined by

\[
(\nabla b_t)(x)v = (\nabla v b_t)(x), \quad v \in \mathbb{R}^d.
\]

By (H1), we have

\[
\|(J_t)^{-1}\| \vee \|J_t\| \leq e^{\int_0^t K_1(s) ds}, \quad t \geq 0.
\]
Theorem 1.1. Assume (H1) and (H2). Let $T > 0$. If

$$(1.4) \quad \mathbb{E} S(T)^{-\frac{1}{2}} < \infty,$$  

then for any $v \in \mathbb{R}^d$ and $f \in C^1_b(\mathbb{R}^d)$,

$$(1.5) \quad P_T(\nabla_v f) = \mathbb{E}\left\{ f(X_T) \frac{M^p_T}{S(T)} \right\},$$

where

$$M^p_T := \left\langle \int_0^T \left( \sigma_t^{-1} J_t \right)^* dW_{S(t)}, J_T^{-1} v \right\rangle$$

$$\quad + \int_0^T dS(t) \int_t^T \text{Tr} \left\{ \sigma_t^{-1} J_t J_s^{-1} \left( \nabla \nabla_{J_t, J_s^{-1}} b_s \right) (X_s) J_s J_t^{-1} \sigma_t \right\} ds$$

is in $L^1(S(T)^{-1}d\mathbb{P})$. Consequently, $P_T$ has a density $p_T(x, y)$ with respect to the Lebesgue measure, which is differentiable in $y$ with

$$\nabla_v \log p_T(x, \cdot)(y) = -\mathbb{E} \left( \frac{M^p_T}{S(T)} \right| X_T(x) = y), \quad v, x \in \mathbb{R}^d.$$

Below we present some consequences of Theorem 1.1 concerning derivative estimates and shift Harnack inequalities. For nonnegative $f \in \mathcal{B}_b(\mathbb{R}^d)$ and $T > 0$, let

$$\text{Ent}_{P_T}(f) = P_T(f \log f) - (P_T f) \log P_T f$$

be the relative entropy of $f$ with respect to $P_T$.

Corollary 1.2. Assume that (H1), (H2) and (1.4) hold. Let

$$\beta(T) = dT \lambda_1(T) \lambda_2(T) K_2(T) e^{3TK_1(T)}, \quad T > 0.$$  

(1) For any $T > 0$ and $v \in \mathbb{R}^d$,

$$\|P_T(\nabla_v f)\|_\infty \leq |v| \cdot \|f\|_\infty \left( \lambda_2(T) e^{TK_1(T)} \mathbb{E} S(T)^{-\frac{1}{2}} + \beta(T) \right), \quad f \in C^1_b(\mathbb{R}^d),$$

$$\int_{\mathbb{R}^d} |\nabla_v p_T(x, \cdot)|(y)dy \leq |v| \left( \lambda_2(T) e^{TK_1(T)} \mathbb{E} S(T)^{-\frac{1}{2}} + \beta(T) \right), \quad x \in \mathbb{R}^d.$$

(2) For any $p > 1$, there exists a constant $C(p) \geq 1$ such that for any $T > 0$,

$$|P_T(\nabla f)| \leq C(p) |P_T| f^p \frac{1}{p} \left( \lambda_2(T) e^{TK_1(T)} \left( \mathbb{E} S(T)^{-\frac{p}{2(p-1)}} \right)^{\frac{p-1}{p}} + \beta(T) \right), \quad f \in C^1_b(\mathbb{R}^d),$$

$$\int_{\mathbb{R}^d} |\nabla \log p_T(x, \cdot)|^{-\frac{1}{p}}(y)p_T(x, y)dy \leq C(p) \left( \lambda_2(T) e^{TK_1(T)} \left( \mathbb{E} S(T)^{-\frac{p}{2(p-1)}} \right)^{\frac{p-1}{p}} + \beta(T) \right), \quad x \in \mathbb{R}^d.$$
Corollary 1.3. Assume (H1) and (H2). Let \( p > 1, T > 0 \). If

\[
\Gamma_{T,p}(r) := \mathbb{E} \exp \left[ \frac{p^2 \lambda_2(T)^2 e^{2TK_1(T)} |v|}{2(p-1)^2 S(T)} \right] < \infty, \quad r \geq 0,
\]

then the shift Harnack inequality

\[
(P_T f)^p(x) \leq \exp \left[ \frac{p(\log p) \beta(T) |v|}{p-1} + p - 1 \right] \log \Gamma_{T,p}(|v|) P_T(f^p(v + \cdot))(x)
\]

holds for all \( v, x \in \mathbb{R}^d \) and positive \( f \in \mathcal{B}_b(\mathbb{R}^d) \). Consequently,

\[
\sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} p_T(x, y) \frac{\rho}{\rho - 1} dy \leq \left( \int_{\mathbb{R}^d} \exp \left[ - \frac{p(\log p) \beta(T) |v|}{p-1} - p - 1 \right] \log \Gamma_{T,p}(|v|) dv \right)^{\frac{1}{\rho - 1}}.
\]

To illustrate the above results, we consider below the SDE driven by \( \alpha \)-stable like noises.

Corollary 1.4. Assume (H1) and (H2). Let \( B(r) \geq cr^\alpha \) for \( r \geq r_0 \), where \( \alpha \in (0, 2) \) and \( c, r_0 > 0 \) are constants.

1. For any \( p > 1 \) there exists a constant \( C(p) > 0 \) such that

\[
|P_T(\nabla f)| \leq \frac{C(p)(P_T|f|^p)^{\frac{\rho}{\rho - 1}}}{1 \wedge T^{\frac{\rho}{\rho - 1}}}, \quad T > 0, f \in C^1_b(\mathbb{R}^d),
\]

\[
\sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} |\nabla \log p_T(x, \cdot)|^{\frac{\rho}{\rho - 1}}(y) p_T(x, y) dy \leq \frac{C(p)}{1 \wedge T^{\frac{\rho}{\rho - 1}}}, \quad T > 0.
\]

2. Let \( \alpha \in (1, 2) \). Then there exists a constant \( C > 0 \) such that for any \( p > 1, \delta > 0, v \in \mathbb{R}^d \) and \( f \in C^1(\mathbb{R}^d) \),

\[
|P_T(\nabla_v f)| \leq \delta \text{Ent}_{P_T}(f) + (P_T f) \left( \beta(T) |v| + \frac{C|v|^2}{\delta^2 (1 \wedge T)^{\frac{\alpha}{\alpha - 1}}} + \frac{C|v|^{\frac{\alpha}{\alpha - 1}}}{\{\delta^\alpha (1 \wedge T)\}^{\frac{1}{\alpha - 1}}} \right),
\]

\[
\sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} \exp \left[ \frac{\nabla_v \log p_T(x, \cdot)(y)}{\delta} \right] p_T(x, y) dy \leq \exp \left[ \beta(T) |v| + \frac{C|v|^2}{\delta^2 (1 \wedge T)^{\frac{\alpha}{\alpha - 1}}} + \frac{C|v|^{\frac{\alpha}{\alpha - 1}}}{\{\delta^\alpha (1 \wedge T)\}^{\frac{1}{\alpha - 1}}} \right].
\]
(3) Let $\alpha \in (1, 2)$. Then there exists a constant $C > 0$ such that for any $p > 1, T > 0, v \in \mathbb{R}^d$ and positive $f \in \mathcal{B}_b(\mathbb{R}^d)$,
\[
(P_T f)^p \leq \exp \left[ \frac{C(p \log p)}{p - 1} |v| + \frac{Cp|v|^2}{(p - 1)(1 \wedge T)^{\alpha}} + \frac{Cp^{\frac{1}{\alpha - 1}} |v|^{\frac{\alpha}{\alpha - 1}}}{(p - 1)(1 \wedge T)^{\alpha - 1}} \right] P_T(f^p(v \cdot)),
\]
\[
\sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} p_T(x, y) \frac{e^{-p \alpha}}{y^p} dy \leq \frac{1}{(1 \wedge T)^{\alpha p - 1}} \exp \left[ \frac{Cp \log p}{(p - 1)^2} + \frac{Cp^{\frac{1}{\alpha - 1}}}{(p - 1)^{\alpha - 1}} \right].
\]

The remainder of the paper is organized as follows. In Section 2, we fix a path $\ell$ of $S$ with finite jumps, and establish the integration by parts formula for the corresponding equation, i.e. the equation (1.1) with $\ell$ in place of $S$. In Section 3, we use this integration by parts formula to prove the above results by using finite jump approximations.

2 Integration by parts formula for the equation with finite jump

In this section, we let $\ell$ be a càdlàg and increasing function on $[0, \infty)$ with $\ell(0) = 0$ such that the set $\{ t \in [0, T]: \Delta \ell(t) := \ell(t) - \ell(t-) > 0 \}$ is finite. We call $\ell$ a path of $S$ with finite many jumps on $[0, T]$. Let $X_t^\ell$ solve the equation
\[
X_t^\ell = X_0^\ell + \int_0^t b_s(X_s^\ell) ds + \int_0^t \sigma_s dW_{\ell(s)} + V_t, \quad t \geq 0,
\]
and let $P_t^\ell$ be the associated Markov operator; i.e.
\[
P_t^\ell f(x) := \mathbb{E} f(X_t^\ell(x)),
\]
where $X_t^\ell(x)$ solves (2.1) for $X_0^\ell = x$. Moreover, let $J_t^\ell$ solve the ODE with random coefficients
\[
\frac{d}{dt} J_t^\ell = (\nabla b_t)(X_t^\ell) J_t^\ell, \quad J_0^\ell = I.
\]
By (H1), we have
\[
\| (J_t^\ell)^{-1} \| \| J_t^\ell \| \leq e^{t K_1(s)} ds, \quad t \geq 0.
\]
The main result in this section is the following.

**Theorem 2.1.** Let $\ell$ be a path of $S$ with finite many jumps on $[0, T]$. Then
\[
P_T^\ell(\nabla_v f) = \mathbb{E} \{ f(X_T^\ell) M_T^\ell, v \}, \quad v \in \mathbb{R}^d, f \in C_b^1(\mathbb{R}^d),
\]
where
\[
M_T^\ell := \left\langle \int_0^T \left( \sigma_t^{-1} J_t^\ell \right)^* dW_{\ell(t)}, (J_t^\ell)^{-1} \right\rangle
\]
\[
+ \int_0^T \int_0^T \text{Tr} \left\{ \sigma_t^{-1} J_t^\ell (J_s^\ell)^{-1} (\nabla \nabla J_t^\ell)^{-1} v b_s) (X_s^\ell) J_t^\ell (J_t^\ell)^{-1} \sigma_t \right\} ds.
\]
Proof. We shall use the integration by parts formula in the Malliavin calculus, see, for instance [8, 9]. For the Brownian motion \((W_t)_{t \in [0, \ell(T)]}\), let \((D, \mathcal{D}(D))\) be the Malliavin gradient, and let \((D^*, \mathcal{D}(D^*))\) be its adjoint operator (i.e. the Malliavin divergence). Let
\[
h(t) = \sum_{i=1}^t (t \wedge \ell(t_i) - \ell(t_{i-1}))^{+} \sigma_{t_i}^{-1} J_t^i (J_T^i)^{-1} v, \quad t \in [0, \ell(T)].
\]
From (H1) it is easy to see that \(J^i_t\) and \((J^i_t)^{-1}\) are Malliavin differentiable for every \(t \in [0, \ell(T)]\), such that \(h \in \mathcal{D}(D^*)\). Since \((V_t)_{t \geq 0}\) is independent of \((W_t)_{t \geq 0}\), we have \(D_h V_t = 0\), so that (2.1) yields
\[
d_h X^i_t = (\nabla_{D_h X^i_t} b_t)(X^i_t) dt + \sigma_t dh_{t \ell(t)}, \quad D_h X^0_0 = 0.
\]
Then
\[
D_h X^i_T = J^i_T \int_0^T (J^i_t)^{-1} \sigma_t dh_{t \ell(t)} = J^i_T \sum_{i=1}^n (J^i_t)^{-1} \sigma_t \sigma_t^{-1} J^i_t (J_T^i)^{-1} \Delta \ell(t_i) = \ell(T) v.
\]
Therefore,
\[
P_T^d(\nabla_v f) = \mathbb{E}(\nabla_v f)(X^i_T) = \frac{1}{\ell(T)} \mathbb{E}(\nabla_{D_h X^i_T} f)(X^i_T)
\]
\[
= \frac{1}{\ell(T)} \mathbb{E}\{D_h f(X^i_T)\} = \frac{1}{\ell(T)} \mathbb{E}\{f(X^i_T) D^*(h)\}.
\]
To calculate \(D^*(h)\), let
\[
h_{ik}(t) = (t \wedge \ell(t_i) - \ell(t_{i-1}))^{+} e_k, \quad F_{ik} = \langle \sigma_t^{-1} J_t^i (J_T^i)^{-1} v, e_k \rangle
\]
for \(1 \leq i \leq n, 1 \leq k \leq d, t \in [0, \ell(T)]\), where \(\{e_k\}_{k=1}^d\) is the canonical orthonormal basis on \(\mathbb{R}^d\). Then
\[
h(t) = \sum_{k=1}^d \sum_{i=1}^n F_{ik} h_{ik}(t), \quad t \in [0, \ell(T)].
\]
Noting that \(h_{ik}\) is deterministic with \(\int_0^{\ell(T)} |h_{ik}'(t)|^2 dt < \infty\), we have
\[
D^*(h_{ik}) = \int_0^{\ell(T)} \langle h_{ik}'(t), dW_t \rangle = \langle e_k, W_{t(t_i)} - W_{t(t_{i-1})} \rangle.
\]
Thus, using the formula \(D^*(F_{ik} h_{ik}) = F_{ik} D^*(h_{ik}) - D_{h_{ik}} F_{ik}\), we obtain
\[
D^*(h) = \sum_{k=1}^d \sum_{i=1}^n \left\{ F_{ik} D^*(h_{ik}) - D_{h_{ik}} F_{ik} \right\}
\]
\[
= \sum_{k=1}^d \sum_{i=1}^n \left\{ F_{ik} \langle e_k, W_{t(t_i)} - W_{t(t_{i-1})} \rangle - \langle \sigma_t^{-1} D_{h_{ik}} (J_t^i (J_T^i)^{-1}) v, e_k \rangle \right\}
\]
\[
= \left( \int_0^T (\sigma_t^{-1} J_t^i)\ast dW_{t \ell(t)}, (J_T^i)^{-1} v \right) - \sum_{k=1}^d \sum_{i=1}^n \langle \sigma_t^{-1} D_{h_{ik}} (J_t^i (J_T^i)^{-1}) v, e_k \rangle.
\]
Since \( dh_{ik}(t) \) is supported on \( (\ell(t_{i-1}), \ell(t_i)) \) but \( J^f_{t_i} \) is measurable with respect to \( \mathcal{F}_{\ell(t_{i-1})} := \sigma\{ W_t : t \leq \ell(t_{i-1}) \} \), we have \( D_{h_{ik}} J^f_{t_i} = 0 \). So,

\[
(2.7) \quad D_{h_{ik}} (J^f_{t_i}(J^f_T)^{-1}) = J^f_{t_i} D_{h_{ik}} (J^f_T)^{-1} = -J^f_{t_i} (J^f_T)^{-1} (D_{h_{ik}} J^f_{t_i})(J^f_T)^{-1}.
\]

Noting that (2.2) yields

\[
\sum_{k} \int_{t_i}^{t_f} \langle \sigma^{-1} \text{d} h_{ik}(s) - 1_{\{t_i \leq t\}}(\ell(t_i) - \ell(t_{i-1})) J^f_{t_i}(J^f_T)^{-1} \sigma_{t_i e_k}, e_k \rangle \text{d} t.
\]

we obtain

\[
(2.8) \quad D_{h_{ik}} J^f_T = J^f_{t_i} \int_0^T (J^f_T)^{-1} (\nabla_{D_{h_{ik}}} X^f_t) J^f_t \text{d} t.
\]

Moreover, it follows from (2.4) that

\[
D_{h_{ik}} X^f_t = J^f_{t_i} \int_0^T (J^f_T)^{-1} (\nabla_{D_{h_{ik}}} X^f_t) J^f_t \text{d} t.
\]

Combining this with (2.8), we arrive at

\[
D_{h_{ik}} J^f_T = (\Delta \ell(t_i)) J^f_T \int_0^T (J^f_T)^{-1} (\nabla_{D_{h_{ik}}} X^f_t) J^f_t \text{d} t.
\]

Substituting this into (2.7) we obtain

\[
\sum_{k=1}^{d} \sum_{i=1}^{n} \langle \sigma_{t_i}^{-1} D_{h_{ik}} (J^f_{t_i}(J^f_T)^{-1}), e_k \rangle = - \sum_{k=1}^{d} \sum_{i=1}^{n} (\Delta \ell(t_i)) \sigma_{t_i}^{-1} J^f_{t_i} \int_0^T (J^f_T)^{-1} (\nabla_{J^f_{t_i}(J^f_T)^{-1}} \sigma_{t_i e_k} \nabla b_{s}) (X^f_s) J^f_T (J^f_T)^{-1} v, e_k \rangle \text{d} s
\]

\[
= - \sum_{k=1}^{d} \int_0^T \text{d} \ell(t) \int_0^T (J^f_T)^{-1} (\nabla_{J^f_{t_i}(J^f_T)^{-1}} \sigma_{t_i e_k} \nabla b_{s}) (X^f_s) J^f_T (J^f_T)^{-1} e_k \rangle \text{d} s
\]

\[
= - \sum_{k=1}^{d} \int_0^T \text{d} \ell(t) \int_0^T \sum_{k=1}^{d} \langle \sigma_{t_i}^{-1} J^f_{t_i}(J^f_T)^{-1} (\nabla_{J^f_{t_i}(J^f_T)^{-1}} \sigma_{t_i e_k} \nabla b_{s}) (X^f_s) J^f_T (J^f_T)^{-1} \sigma_{t_i e_k}, e_k \rangle \text{d} s
\]

\[
= - \int_0^T \text{d} \ell(t) \int_0^T \sum_{k=1}^{d} \langle \sigma_{t_i}^{-1} J^f_{t_i}(J^f_T)^{-1} (\nabla_{J^f_{t_i}(J^f_T)^{-1}} \nabla b_{s}) (X^f_s) J^f_T (J^f_T)^{-1} \sigma_{t_i e_k}, e_k \rangle \text{d} s
\]

Therefore, we derive from (2.6) that \( D^*(h) = M^f_T \) and hence, the proof is finished by (2.5).
Remark 2.1. It is indicated by Xicheng Zhang to the author that this result can also be proved using Remark 2.1 in [13], which says that

\[ P_T^\ell(\nabla_v f) = \mathbb{E}\left( f(X_T^\ell) \sum_{i,k=1}^d \left[ D^* (h_k) \left\{ (\nabla X_T^\ell)^{-1} \right\}_{ki} - D_{h_k} \left\{ (\nabla X_T^\ell)^{-1} \right\}_{ki} \right] v_i \right), \]

where \( h_k \in \mathcal{D}(D^*) \) is such that \( D_{h_k} X_T^\ell = \nabla_{e_k} X_T^\ell \).

Noting that \( \nabla X_T^\ell = J_T^\ell \), we can then prove Theorem 2.1 by calculating \( D^* (h_k) \) and \( D_{h_k} T_T^\ell \) as above.

3 Proofs

Proof of Theorem 1.1. According to [10, Theorem 2.4(1)], the second assertion follows from the first. So, it suffices to prove the desired integration by parts formula. For any path \( \ell \) of \( S \),

\[ \ell_\varepsilon(t) = \sum_{s \leq t} \Delta \ell(s) 1_{\{\Delta \ell(s) \geq \varepsilon\}}, \quad t \geq 0. \]

Then \( \ell_\varepsilon \) has finite many jumps on \([0, T]\). Moreover, \( d\ell_\varepsilon(t) \to d\ell(t) \) on \([0, T] \) strongly as \( \varepsilon \to 0 \). Note that by (2.3), (H1) and (H2),

\[ \left\| \sigma_1^{-1} J_{\ell_\varepsilon}^t (J_{\ell_\varepsilon}^t)^{-1} \right\| + \int_t^T \left\| \sigma_1^{-1} J_{\ell_\varepsilon}^s (J_{\ell_\varepsilon}^s)^{-1} (\nabla J_{\ell_\varepsilon}^s (J_{\ell_\varepsilon}^s)^{-1} - \sigma_1 e_k \nabla b_s)(X_{\ell_\varepsilon}^s) J_{\ell_\varepsilon}^s (J_{\ell_\varepsilon}^s)^{-1} \right\| ds \]

is bounded in \((t, \varepsilon) \in [0, T] \times [0, 1]\), and by [12, Lemma 3.1]

\[ \lim_{\varepsilon \to 0} \mathbb{E} \sup_{t \in [0, T]} |X_{\ell_\varepsilon}^t - X_{\ell}^t|^2 = 0, \]

which together with (H1) implies

\[ \lim_{\varepsilon \to 0} \mathbb{E} \sup_{t \in [0, T]} (\left\| J_{\ell_\varepsilon}^t - J_{\ell}^t \right\|^2 + \left\| (J_{\ell_\varepsilon}^t)^{-1} - (J_{\ell}^t)^{-1} \right\|^2) = 0. \]

Combining these with (2.3), (H1) and (H2), we conclude that

\[ \lim_{\varepsilon \to 0} P_T^\ell(\nabla_v f) = \lim_{\varepsilon \to 0} \mathbb{E}(\nabla_v f)(X_T^\ell) = P_T^\ell(\nabla_v f), \]

\[ \lim_{\varepsilon \to 0} \mathbb{E}\{ f(X_T^\ell) M_T^{\ell, v} \} = \mathbb{E}\{ f(X_T) M_T^{\ell, v} \}, \quad f \in C_b^1(\mathbb{R}^d). \]

Therefore, first applying Theorem 2.1 to \( \ell_\varepsilon \) in place of \( \ell \) then letting \( \varepsilon \downarrow 0 \), we obtain

\[ P_T^\ell(\nabla_v f) = \frac{1}{\ell(T)} \mathbb{E}\{ f(X_T^\ell) M_T^{\ell, v} \} \]
for all sample path \( \ell \) of \( S \) with \( \ell(T) > 0 \). Since \( \mathbb{E}_S(T)^{-\frac{1}{2}} < \infty \) implies \( S(T) > 0 \), and noting that \( X_T = X_T^S, M_T^v = M_T^{S,v} \), we obtain

\[
P_T^{S}(\nabla_v f) = \frac{1}{S(T)} \mathbb{E}_T \{ f(X_T) M_T^v \},
\]

where \( \mathbb{E}^S \) is the conditional expectation given \( S \). Moreover, it follows from (2.3), (H1), (H2), and \( \mathbb{E}S(T)^{-\frac{1}{2}} < \infty \) that

\[
\mathbb{E} \left| \frac{M_T^v}{S(T)} \right| \leq \mathbb{E} \left| \frac{1}{S(T)} \mathbb{E}_T^S M_T^v \right| \\
\leq \mathbb{E} \left[ \frac{1}{S(T)} \left( \mathbb{E}^S \int_0^T |\sigma_t^{-1} J_t(J_T)^{-1} v|^2 dS(t) \right)^{1/2} \right. \\
\left. + \sum_{k=1}^d \frac{1}{S(T)} \int_0^T dS(t) \int_t^T \|\sigma_t^{-1} J_t^{-1}\| \cdot \left| (\nabla_{J_t} J_t^{-1} v)_s \right| \mathbb{E} \left( |(\nabla_{J_t} J_t^{-1} v)_s| \right) ds \right] \\
\leq |v| \left( \lambda_2(T) e^{TK_1(T)} \mathbb{E} S(T)^{-\frac{1}{2}} + dT \lambda_1(T) \lambda_2(T) K_2(T) e^{TK_1(T)} \right) < \infty.
\]

Then \( M_T^v \in L^1(S(T)^{-1}d\mathbb{P}) \) and (3.1) yields that

\[
P_T(\nabla_v f) = \mathbb{E} P_T^S(\nabla_v f) = \mathbb{E} \left( \frac{1}{S(T)} f(X_T) M_T^v \right).
\]

This completes the proof. \( \Box \)

**Proof of Corollary 1.2.** Assertion (1) follows immediately from (3.2), Theorem 1.1 and [10, Theorem 2.4(1)] with \( H(r) = r \).

Next, by (1.3), (H1), (H2) and the Burkholder inequality [15, Theorem 2.3] (see also [12, Lemma 2.1]), for any \( p > 1 \) there exists a constant \( C(p) \geq 1 \) such that

\[
\left( \mathbb{E} \left| \frac{M_T^v}{S(T)} \right|^{\frac{p}{p-1}} \right)^{\frac{p-1}{p}} \leq \beta(T)|v| + C(p) \left( \frac{\mathbb{E} \left( \int_0^T |\sigma_t^{-1} J_t(J_T)^{-1} v|^2 dS(t) \right)^{\frac{p}{2(p-1)}}}{S(T)^{\frac{p}{2-1}}} \right)^{\frac{p-1}{p}} \\
\leq \beta(T)|v| + C(p)|v| \lambda_2(T) e^{TK_1(T)} \mathbb{E} S(T)^{-\frac{1}{2}} \left( \frac{S(T)^{\frac{p}{2-1}}}{S(T)^{\frac{p}{2-1}}} \right)^{\frac{p-1}{p}}.
\]

Then assertion (2) follows from [10, Theorem 2.4(1)] with \( H(r) = r^{\frac{p}{p-1}} \) and the fact that

\[
|P_T(\nabla_v f)| = \left| \mathbb{E} \left\{ f(X_T^t) M_T^v \right\} \right| \leq (P_T|f|^p)^{\frac{1}{p}} \left( \mathbb{E} \left| \frac{M_T^v}{S(T)} \right|^{\frac{p}{p-1}} \right)^{\frac{p-1}{p}}, \quad v \in \mathbb{R}^d.
\]

Finally, by Theorem 1.1 and the Young inequality (see [2, Lemma 2.4]), if \( f \in C_b^1(\mathbb{R}^d) \) is nonnegative, then

\[
|P_T(\nabla_v f)| = \left| \mathbb{E} \left\{ f(X_T^t) M_T^v \right\} \right| \leq \delta \text{Ent}_{P_T}(f) + \delta (P_T f) \log \mathbb{E} \exp \left[ \frac{M_T^v}{\delta S(T)} \right], \quad \delta > 0.
\]
Obviously, by (1.3), (H1) and (H2),
\[
\frac{M_T}{S(T)} \leq \beta(T)|v| + \frac{1}{S(T)} \int_0^T \langle \sigma_t^{-1} J_t J_t^{-1} v, dW_s(t) \rangle,
\]
\[
\mathbb{E}^S \exp \left[ \frac{1}{\delta S(T)} \int_0^T \langle \sigma_t^{-1} J_t J_t^{-1} v, dW_s(t) \rangle \right] \leq \exp \left[ \frac{\lambda_2(T)^2 |v|^2 e^{2TK_1(T)}}{2\delta^2 S(T)} \right], \quad \delta > 0.
\]
Then
\[
\log \mathbb{E} \exp \left[ \frac{M_T}{\delta S(T)} \right] \leq \frac{\beta(T)|v|}{\delta} + \log \mathbb{E} \exp \left[ \frac{\lambda_2(T)^2 |v|^2 e^{2TK_1(T)}}{2\delta^2 S(T)} \right], \quad \delta > 0.
\]
By combining this with (3.3) and [10, Theorem 2.4(1)] for \( H(r) = e^{r/\delta} \), we prove (3). \( \square \)

Proof of Corollary 1.3. By [10, Theorem 2.5(2)], the second assertion follows from the first. So, we only need to prove the required shift Harnack inequality (1.6) for \( v \neq 0 \). By Corollary 1.2(3), we have
\[
|P_T(\nabla_v f)| \leq \delta \text{Ent}_{P_T}(f) + (P_T f) \left( \beta(T)|v| + \delta \log \mathbb{E} \exp \left[ \frac{\lambda_2(T)^2 |v|^2 e^{2TK_1(T)}}{2\delta^2 S(T)} \right] \right), \quad \delta > 0.
\]
So, letting
\[
\beta_v(\delta) = \beta(T)|v| + \delta \log \mathbb{E} \exp \left[ \frac{\lambda_2(T)^2 |v|^2 e^{2TK_1(T)}}{2\delta^2 S(T)} \right], \quad \delta > 0,
\]
we obtain from [10, Proposition 2.3] that
\[
(P_T f)^p \leq (P_T f^p (v + \cdot)) \exp \left[ \int_0^1 \frac{p}{1 + (p - 1)s} \beta_v \left( \frac{p - 1}{1 + (p - 1)s} \right) ds \right].
\]
By the Jensen inequality, for \( \delta = \frac{p - 1}{1 + (p - 1)s} \), we have
\[
\mathbb{E} \exp \left[ \frac{\lambda_2(T)^2 |v|^2 e^{2TK_1(T)}}{2\delta^2 S(T)} \right] \leq \left( \mathbb{E} \exp \left[ \frac{p^2 \lambda_2(T)^2 |v|^2 e^{2TK_1(T)}}{2(p - 1)^2 S(T)} \right] \right)^{\frac{(1 + (p - 1)s)^2}{p^2}}\Gamma_{T,p}(|v|)^{\frac{1}{p}}.
\]
Thus,
\[
\int_0^1 \frac{p}{1 + (p - 1)s} \beta_v \left( \frac{p - 1}{1 + (p - 1)s} \right) ds \leq \beta(T)|v| \int_0^1 \frac{p}{1 + (p - 1)s} ds + \frac{p - 1}{p} \log \Gamma_{T,p}(|v|) ds
\]
\[
= \frac{p \log p}{p - 1} \beta(T)|v| + \frac{p - 1}{p} \log \Gamma_{T,p}(|v|).
\]
Then the proof is finished by combining this with (3.4). \( \square \)
Proof of Corollary 1.4. Since assertions in Corollaries 1.2 and 1.3 are uniform in \( V \), we may apply them for any deterministic path of \( V \) in place of the process \( V \), so that these two Corollaries remain true for \( P_T^V \) in place of \( P_T \), where

\[
P_T^V f(x) = \mathbb{E}^V(f(X_T(x)) := \mathbb{E}(f(X_T(x))|V).
\]

Next, we observe that by the Markov property it suffices to prove the assertions for \( P_T^V \) in place of \( P_T \) with \( T \in (0,1) \). In fact, for \( T > 1 \) let

\[
P_{T,T}^V f(x) = \mathbb{E}^V f(X_{1,T}(x)), \ f \in \mathcal{B}_b(\mathbb{R}^d), x \in \mathbb{R}^d,
\]

where \((X_{1,t}(x))_{t \geq 1}\) solves the equation

\[
X_{1,t}(x) = x + \int_1^t b_s(X_{1,s}(s))ds + \int_1^t \sigma_s dW(s) + V_t - V_1, \ t \geq 1.
\]

Then by the Markov property of \( X_t \) under \( \mathbb{E}^V \), we obtain,

\[
P_T^V f = P_{1,T}^V(P_1^V f), \ f \in \mathcal{B}_b(\mathbb{R}^d).
\]

Combining this with the assertions for \( T = 1 \) and using the Jensen inequality, we prove the assertions for \( T > 1 \). For instance, if for \( p > 1 \) one has

\[
|P_1^V(\nabla f)| \leq C(p)(P_1^V |f|^p)^{\frac{1}{p}}
\]

then for any \( T > 1 \),

\[
|P_T(\nabla f)| = |\mathbb{E}P_{1,T}^V P_1^V(\nabla f)| \leq \mathbb{E}P_{1,T}^V |P_1^V(\nabla f)|
\]

\[
\leq C(p)\mathbb{E}P_{1,T}^V (P_T^V |f|^p)^{\frac{1}{p}} \leq C(p)(P_T^V |f|^p)^{\frac{1}{p}} = \frac{C(p)(P_T^V |f|^p)^{\frac{1}{p}}}{(1 \wedge T)^{\frac{1}{p}}}.
\]

Below we prove assertions (1)-(3) for \( T \in (0,1) \) respectively.

1. Since \( \beta(T) + \lambda_2(T)e^{TK_1(T)} \) is bounded for \( T \in (0,1) \), and by [12, (ii) in the proof of Theorem 1.1]

\[
\left( \mathbb{E}S(T)^{-\frac{p}{2(p-1)}} \right)^{\frac{p-1}{p}} \leq \frac{C}{T^{\frac{1}{\alpha}}}, \ T \in (0,1]
\]

holds for some constant \( C > 0 \), the desired assertion follows from Corollary 1.2(2).

2. Let \( \alpha \in (1,2) \), and let \( S_\alpha \) be the subordinator induced by the Bernstein function \( r \mapsto r^{\frac{\alpha}{2}} \). Then as shown in [11, Proof of Corollary 1.2] that

\[
\mathbb{E} \frac{1}{S(T)^k} \leq c_0 \mathbb{E} \frac{1}{S_\alpha(T)^k}, \ k \geq 1, T \in (0,1]
\]

holds for some constant \( c_0 \geq 1 \). Combining this with the third display from below in the proof of [7, Theorem 1.1] for \( \kappa = 1 \), i.e. (note the \( \alpha \) therein is \( \alpha/2 \) here)

\[
\mathbb{E} e^{\lambda/S(t)} \leq 1 + \left( \exp \left[ c_1 \frac{\lambda^{\frac{\alpha}{t^{\frac{\alpha}{\alpha-1}}}}}{t^{\frac{\alpha}{\alpha-1}}} \right] - 1 \right)^{2(\alpha-1)/\alpha} \leq \exp \left[ \frac{c_2 \lambda}{t^{\frac{\alpha}{2}}} + \frac{c_2 \lambda^{\frac{\alpha}{t^{\frac{\alpha}{\alpha-1}}}}}{t^{\frac{\alpha}{\alpha-1}}} \right], \ \lambda, t \geq 0
\]

for some constants $c_1, c_2 > 0$, we obtain
\[
\mathbb{E}e^{\lambda/S(T)} \leq 1 + c_0(\mathbb{E}e^{\lambda/S_\alpha(T)} - 1) \leq \mathbb{E}e^{c_0 \lambda/S_\alpha(T)}
\]
(3.5)
\[
\leq \exp \left[ \frac{c_3 \lambda}{t^{\frac{\alpha}{\alpha-1}}} \right], \quad T \in (0, 1], \lambda \geq 0
\]
for some constant $c_3 > 0$. By Corollary 1.2(3) and (3.5), we prove the desired assertion.

(3) By (3.5), there exists a constant $c_4 > 0$ such that
\[
\Gamma_{T,p}(r) \leq \exp \left[ \frac{c_4 p^2 r^2}{(p-1)^2 T^{\frac{\alpha}{\alpha-1}}} \right] + \frac{c_4 (pr)^{\frac{\alpha}{\alpha-1}}}{(p-1)\alpha T^{\frac{\alpha}{\alpha-1}}}, \quad r \geq 0, T \in (0, 1].
\]
(3.6)
Then there exists a constant $c_5 > 0$ such that
\[
\frac{p(\log p)\beta(T)|v|}{p-1} + \frac{p-1}{p} \log \Gamma_{T,p}(|v|) \leq \frac{c_5 (p \log p)|v|}{p-1} + \frac{c_5 p^{\frac{1}{\alpha-1}}|v|^{\frac{\alpha}{\alpha-1}}}{(p-1)\alpha T^{\frac{\alpha}{\alpha-1}}}, \quad T \in (0, 1], v \in \mathbb{R}^d.
\]
By Corollary 1.3, this implies the first inequality in (3) for some constant $C > 0$. Finally, the second inequality in (3) follows since (3.6) and Corollary 1.3 imply
\[
\sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} p_{T}(x, y)\frac{p}{p-1}dy \leq \left( \int_{\mathbb{R}^d} \exp \left[ - \frac{C(p \log p)|v|}{p-1} - \frac{Cp|v|^2}{(p-1)T^{\frac{\alpha}{\alpha-1}}} - \frac{Cp^{\frac{1}{\alpha-1}}|v|^{\frac{\alpha}{\alpha-1}}}{[(p-1)T]^{\frac{1}{\alpha-1}}} \right] dv \right)^{\frac{1}{p-1}}
\]
\[
\leq \left( \int_{|v| \leq T^{\frac{1}{p-1}}} \exp \left[ \frac{Cp(1+\log p)}{(p-1)^2} + \frac{Cp^{\frac{1}{\alpha-1}}}{(p-1)^{\frac{1}{\alpha-1}}} \right] dv \right)^{\frac{1}{p-1}}
\]
\[
\leq \frac{1}{T^{\frac{d}{\alpha(p-1)}}} \exp \left[ C' p \log p \frac{1}{(p-1)^2} + \frac{C'p^{\frac{1}{\alpha-1}}}{(p-1)^{\frac{1}{\alpha-1}}} \right]
\]
for some constant $C' \geq C$.

Acknowledgement. The author would like to thank professor Xicheng Zhang and the referee for helpful comments.

References


