WASSERSTEIN-1 DISTANCE BETWEEN SDES DRIVEN BY BROWNIAN MOTION AND STABLE PROCESSES

CHANG-SONG DENG, RENÉ L. SCHILLING, AND LIHU XU

Abstract. We are interested in the following two $\mathbb{R}^d$-valued stochastic differential equations (SDEs):

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma\,\mathrm{d}L_t, \quad X_0 = x,$$
$$\mathrm{d}Y_t = b(Y_t)\,\mathrm{d}t + \sigma\,\mathrm{d}B_t, \quad Y_0 = y,$$

where σ is an invertible $d\times d$ matrix, L_t is a rotationally symmetric α-stable Lévy process, and B_t is a d-dimensional standard Brownian motion. We show that for any $\alpha_0 \in (1,2)$ the Wasserstein-1 distance W_1 satisfies for $\alpha \in [\alpha_0, 2)$

$$W_1\left(\mathrm{law}(X_t^x), \mathrm{law}(Y_t^y)\right) \le C\mathrm{e}^{-Ct}|x-y| + C_{\alpha_0} d\cdot\log(1+d)(2-\alpha)\log\frac{1}{2-\alpha},$$

which implies, in particular,

$$W_1(\mu_\alpha, \mu) \le C_{\alpha_0} d\cdot\log(1+d)(2-\alpha)\log\frac{1}{2-\alpha}, \tag{0.1}$$

where μ_α and μ are the ergodic measures of $(X_t^x)_{t\ge 0}$ and $(Y_t^y)_{t\ge 0}$ respectively.

The term $d\cdot\log(1+d)$ appearing in this estimate seems to be optimal. For the special case of a d-dimensional Ornstein–Uhlenbeck system, we show that $W_1(\mu_\alpha,\mu) \ge C_{\alpha_0,d}(2-\alpha)$; this indicates that the convergence rate with respect to α in (0.1) is optimal up to a logarithmic correction. We conjecture that the sharp rate with respect to α and d is $d\cdot\log(1+d)(2-\alpha)$.

Contents

1. Introduction

We study the following two $\mathbb{R}^d$-valued stochastic differential equations (SDEs):

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma\,\mathrm{d}L_t, \quad X_0 = x, \tag{1.1}$$

$$\mathrm{d}Y_t = b(Y_t)\,\mathrm{d}t + \sigma\,\mathrm{d}B_t, \quad Y_0 = y, \tag{1.2}$$

where σ is an invertible $d \times d$ matrix, B_t is a d-dimensional standard Brownian motion, and L_t is a rotationally symmetric α-stable Lévy process with characteristic function $\mathbb{E}\mathrm{e}^{\mathrm{i}\,\xi L_t} = \mathrm{e}^{-t|\xi|^\alpha/2}$. Under some suitable conditions, we can easily show that both equations have solutions which are ergodic.

SDEs driven by α-stable processes have been intensively studied in recent years. We refer the reader to [5, 17, 32, 16] for gradient estimates, to [26, 27, 25, 10, 31, 23] for structural properties and ergodicity, and to [6, 15, 14] for the existence and uniqueness of solutions and approximation schemes. The aim of this paper is to study the difference between the two ergodic measures in Wasserstein distance.

From Lévy's continuity theorem, see e.g. [11], we know that an α-stable distribution converges to the normal distribution as $\alpha \uparrow 2$. It is natural and important to ask whether this convergence carries over to SDEs driven by Brownian motion and stable processes. As an application, one can justify that heavy tailed financial time series with second moment could be modeled by an SDE driven by Brownian motion [12]. There have been several results in this direction, see for instance [18, 19, 20] and the references therein, but all of these results establish only the convergence without giving a rate. In the present paper we obtain a convergence rate which is optimal up to a logarithmic correction.

Throughout the paper, we make the following two assumptions:

(H1). There exist constants $\theta_0 > 0$ and $K \geq 0$ such that

$$\langle x - y, b(x) - b(y)\rangle \leq -\theta_0 |x-y|^2 + K \quad \text{for all} \quad x, y \in \mathbb{R}^d;$$

(H2). There exist constants $\theta_1, \theta_2, \theta_3 \geq 0$ such that

$$|\nabla_v b(x)| \leq \theta_1 |v|, \quad v, x \in \mathbb{R}^d, \tag{1.3}$$

$$|\nabla_{v_1}\nabla_{v_2} b(x)| \leq \theta_2 |v_1||v_2|, \quad v_1, v_2, x \in \mathbb{R}^d, \tag{1.4}$$

$$|\nabla_{v_1}\nabla_{v_2}\nabla_{v_3} b(x)| \leq \theta_3 |v_1||v_2||v_3|, \quad v_1, v_2, v_3, x \in \mathbb{R}^d; \tag{1.5}$$

where directional derivatives are defined in (1.7), (1.8) and (1.9) below.

It is well-known that if (1.3) holds, then both (1.1) and (1.2) have unique non-explosive (strong) solutions. Whenever we want to emphasize the starting point $X_0 = x$ for a given $x \in \mathbb{R}^d$, we will write X_t^x instead of X_t; we use this also for Y_t^y for a given $y \in \mathbb{R}^d$.

Notation. We denote by $C(\mathbb{R}^d, \mathbb{R})$, $C^k(\mathbb{R}^d, \mathbb{R})$ the sets of continuous and k-times continuously differentiable functions; the subscripts "b" and "c" indicate that the functions and all their derivatives up to order k are bounded, resp., have compact support.

Denote by Lip the set of all Lipschitz functions from $\mathbb{R}^d$ to $\mathbb{R}$. The set of Lipschitz functions with Lipschitz constant 1 is denoted by

$$\mathrm{Lip}(1) = \left\{ h \,:\, |h(x) - h(y)| \leq |x - y| \text{ for all } x, y \in \mathbb{R}^d \right\}.$$

The Wasserstein-1 distance between two probability measures μ_1 and μ_2 is defined as

$$\begin{aligned} W_1(\mu_1,\mu_2) &= \sup_{h\in \mathrm{Lip}(1)} \left\{ \int h(x)\mu_1(\mathrm{d}x) - \int h(x)\mu_2(\mathrm{d}x) \right\} \\ &= \sup_{h\in \mathrm{Lip}(1)} \left\{ \int h(x)\mu_1(\mathrm{d}x) - \int h(x)\mu_2(\mathrm{d}x), \quad |h(x)| \le |x| \right\}, \end{aligned}$$

For $h \in \mathrm{Lip}(1)$, $\|\nabla h\|_\infty$ is defined as $\|\nabla h\|_\infty = \sup_{x\neq y, x\in\mathbb{R}^d, y\in\mathbb{R}^d} \frac{|h(x)-h(y)|}{|x-y|}$, which is the Lipschitz constant.

Throughout this paper, C, C_1, C_2 denote positive constants which may depend on $\theta_0, \theta_1, \theta_2, \theta_3, K, \|\nabla b(0)\|_{\mathrm{HS}}, \|\sigma\|_{\mathrm{HS}}$, but they are always independent of d and α; their value may change, without further notice, from line to line.

From the classical Lyapunov function criterion [22] or Harris' Theorem [24, 13] we know that the solutions to the SDEs (1.1) and (1.2) are ergodic. Denote by μ_α and μ the respective ergodic measures of X_t^x and Y_t^y.

1.1. **Main results.** The following two theorems are the main results of our paper.

Theorem 1.1. *Assume that both* **(H1)** *and* **(H2)** *hold true, and let $\alpha_0 \in (1,2)$ be an arbitrary number. For any $\alpha \in [\alpha_0, 2)$, $x, y \in \mathbb{R}^d$ and $t > 0$, we have*

$$W_1\left(\mathrm{law}(X_t^x), \mathrm{law}(Y_t^y)\right) \le C_1 \mathrm{e}^{-C_2 t}|x-y| + C_{\alpha_0} d \cdot \log(1+d)(2-\alpha)\log\frac{1}{2-\alpha},$$

where C_1, C_2 may depend on $\theta_0, \theta_1, K, \|\sigma\|_{\mathrm{HS}}$, and C_{α_0} may depend on α_0, $\theta_0, \theta_1, \theta_2, \theta_3, K, \|\nabla b(0)\|_{\mathrm{HS}}, \|\sigma\|_{\mathrm{HS}}$. In particular,

$$W_1(\mu_\alpha, \mu) \le C_{\alpha_0} d \cdot \log(1+d)(2-\alpha)\log\frac{1}{2-\alpha}. \tag{1.6}$$

It seems to be difficult to improve the factor $d \cdot \log(1+d)$ in the estimate (1.6). This becomes clear from the proof of Lemma 3.4, since the estimate of the term J_{11} used in this lemma is sharp; further details are given shortly before the statement of Lemma 3.4. For the particular case of a d-dimensional Ornstein–Uhlenbeck system, we show that $W_1(\mu_\alpha, \mu) \ge C_{\alpha_0,d}(2-\alpha)$ in Section 5, from which we see that the convergence rate with respect to α in Theorem 1.1 is optimal up to a logarithmic correction. We conjecture that the sharp rate with respect to α and the dimension d is $d \cdot \log(1+d)(2-\alpha)$.

Our method relies on Duhamel's principle and a comparison of the generators of the solutions of the two SDEs; it may be seen as a continuous version of the probability approximation framework established in [4]. The other key ingredients of our analysis are a time-change technique and Bismut's formula from Malliavin calculus.

1.2. **Preliminaries.** In order to prove the main results, we use the fact that the solutions $(X_t^x)_{t\ge0}$ and $(Y_t^y)_{t\ge0}$ to the SDEs are Markov processes. The operator semigroup induced by the Markov process $(X_t^x)_{t\ge0}$ is given by

$$P_t f(x) = \mathbb{E} f(X_t^x), \quad f \in C_b(\mathbb{R}^d, \mathbb{R}),\ t > 0,$$

The infinitesimal generators $\mathscr{A}^P$ is a closed operator defined on the set of continuous functions vanishing at infinity, $C_\infty = C_\infty(\mathbb{R}^d, \mathbb{R})$,

$$\mathrm{Dom}(\mathscr{A}^P) := \left\{ f \in C_\infty \;:\; \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t} \;\text{ ex. for all } x \text{ and } f \in C_\infty \right\},$$

$$\mathscr{A}^P f(x) := \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t}.$$

It is well-known that $C_c^2(\mathbb{R}^d, \mathbb{R}) \subset \mathrm{Dom}(\mathscr{A}^P)$.

Similarly, we can consider the semigroup $Q_t f(x) = \mathbb{E} f(Y_t^x)$ associated to $(Y_t^y)_{t\geq 0}$ and its infinitesimal generator $\mathscr{A}^Q$.

In applications, we often study a semigroup acting on functions which do not belong to $C_c^\infty(\mathbb{R}^d, \mathbb{R})$ or $C_\infty(\mathbb{R}^d, \mathbb{R})$, and so we need to extend the domains $\mathrm{Dom}(\mathscr{A}^P)$ and $\mathrm{Dom}(\mathscr{A}^Q)$ to a larger function class.

Let $(X_t^x)_{t\geq 0}$ be the solution to (1.1). Because of the Lipschitz property of b, we have

$$\begin{aligned}
\mathbb{E}|X_t^x| &\leq |x| + \int_0^t (|b(0)| + C\mathbb{E}|X_s^x|)\,\mathrm{d}s + \mathbb{E}|\sigma L_t| \\
&\leq |x| + |b(0)|t + Ct^{1/\alpha} + C\int_0^t \mathbb{E}|X_s^x|\,\mathrm{d}s,
\end{aligned}$$

which we may combine with Gronwall's inequality to get

$$\mathbb{E}[|X_t^x|] \leq C_t(1+|x|) \quad \text{and, similarly,} \quad \mathbb{E}[|Y_t^x|] \leq C_t(1+|x|).$$

Thus, it is natural to consider the semigroups $(P_t)_{t\geq 0}$ and $(Q_t)_{t\geq 0}$ on the class of functions with linear growth:

$$C_{\mathrm{lin}}(\mathbb{R}^d, \mathbb{R}) = \left\{ f \in C(\mathbb{R}^d, \mathbb{R}) \;:\; \sup_{x\in\mathbb{R}^d} \frac{|f(x)|}{1+|x|} < \infty \right\}$$

and we define an extension of $\mathscr{A}^P$ as

$$\widetilde{\mathrm{Dom}}(\mathscr{A}^P) := \left\{ f \in C_{\mathrm{lin}}(\mathbb{R}^d, \mathbb{R}) \;:\; \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t} \;\text{ exists for all } x \text{ and } f \in C_{\mathrm{lin}}(\mathbb{R}^d, \mathbb{R}) \right\}$$

$$\mathscr{A}^P f(x) := \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t}.$$

The argument in [32, (1.7)] (this is also used in the proof of (3.2) and (3.3)) still holds for $f \in C_{\mathrm{lin}}$ and shows, in this case, that $|\nabla^k P_t f(x)| \leq C(1+|x|)t^{-k/\alpha}$, $k = 1, 2$. Since $C_{\mathrm{lin}}^2(\mathbb{R}^d, \mathbb{R})$, the space of all twice differentiable functions which grow, together with their deriviatives, at most linearly is contained in $\widetilde{\mathrm{Dom}}(\mathscr{A}^P)$, we conclude that $P_t f \in \widetilde{\mathrm{Dom}}(\mathscr{A}^P)$. In a similar way we can extend $\mathscr{A}^Q$ onto $\widetilde{\mathrm{Dom}}(\mathscr{A}^Q)$. Since we can approximate functions in $C_{\mathrm{lin}}^2(\mathbb{R}^d, \mathbb{R})$ locally uniformly with a sequence from $C_c^2(\mathbb{R}^d, \mathbb{R})$, it is clear that the Kolmogorov equations remain valid in a pointwise sense.

The Kolmogorov backward equations read: For all $f \in C_{\mathrm{lin}}(\mathbb{R}^d, \mathbb{R})$ and $t > 0$,

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t f = \mathscr{A}^P P_t f, \quad \frac{\mathrm{d}}{\mathrm{d}t} Q_t f = \mathscr{A}^Q Q_t f.$$

and the following Kolmogorov forward equations: for all $t > 0$,

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t f = P_t \mathscr{A}^P f, \quad f \in \widetilde{\mathrm{Dom}}(\mathscr{A}^P),$$
$$\frac{\mathrm{d}}{\mathrm{d}t} Q_t f = Q_t \mathscr{A}^Q f, \quad f \in \widetilde{\mathrm{Dom}}(\mathscr{A}^Q).$$

For $f \in C^3(\mathbb{R}^d, \mathbb{R})$ and $v_1, v_2, v_3, x \in \mathbb{R}^d$, the directional derivatives $\nabla_{v_1} f(x)$, $\nabla_{v_2}\nabla_{v_1} f(x)$ and $\nabla_{v_3}\nabla_{v_2}\nabla_{v_1} f(x)$ are defined by

$$\nabla_{v_1} f(x) := \lim_{\epsilon \to 0} \frac{f(x + \epsilon v_1) - f(x)}{\epsilon}, \tag{1.7}$$

$$\nabla_{v_2}\nabla_{v_1} f(x) := \lim_{\epsilon \to 0} \frac{\nabla_{v_1} f(x + \epsilon v_2) - \nabla_{v_1} f(x)}{\epsilon}, \tag{1.8}$$

and

$$\nabla_{v_3}\nabla_{v_2}\nabla_{v_1} f(x) := \lim_{\epsilon \to 0} \frac{\nabla_{v_2}\nabla_{v_1} f(x + \epsilon v_3) - \nabla_{v_2}\nabla_{v_1} f(x)}{\epsilon}. \tag{1.9}$$

The operator norms of $\nabla^2 f(x) \in \mathbb{R}^{d\times d}$ and $\nabla^3 f(x) \in \mathbb{R}^{d\times d\times d}$ are given by

$$\|\nabla^2 f(x)\|_{\mathrm{op}} := \sup\left\{|\nabla_{v_2}\nabla_{v_1} f(x)|\,;\, v_1, v_2 \in \mathbb{R}^d, |v_1| = |v_2| = 1\right\},$$

and

$$\|\nabla^3 f(x)\|_{\mathrm{op}} := \sup\left\{|\nabla_{v_3}\nabla_{v_2}\nabla_{v_1} f(x)|\,;\, v_1, v_2, v_3 \in \mathbb{R}^d, |v_1| = |v_2| = |v_3| = 1\right\}.$$

The Hilbert-Schmidt inner product of two matrices $A, B \in \mathbb{R}^{d\times d}$ is $\langle A, B\rangle_{\mathrm{HS}} = \sum_{i,j=1}^d A_{ij}B_{ij}$, and the Hilbert-Schmidt norm of a matrix $A \in \mathbb{R}^{d\times d}$ is $\|A\|_{\mathrm{HS}} = \sqrt{\sum_{i,j=1}^d A_{ij}^2}$. For $f \in C_b^3(\mathbb{R}^d, \mathbb{R})$, we will use the supremum norms

$$\|\nabla f\|_\infty := \sup_{x\in\mathbb{R}^d} |\nabla f(x)|, \qquad \|\nabla^i f\|_{\mathrm{op},\infty} := \sup_{x\in\mathbb{R}^d} \|\nabla^i f(x)\|_{\mathrm{op}}, \quad (i = 2, 3).$$

For further use, set

$$A(d,\alpha) := \frac{\alpha\Gamma(\frac{d+\alpha}{2})}{2^{2-\alpha}\pi^{d/2}\Gamma(1-\frac{\alpha}{2})}, \quad \omega_{d-1} = \frac{2\pi^{d/2}}{\Gamma\left(\frac{d}{2}\right)}. \tag{1.10}$$

2. Gradient estimates, the time-change method, Bismut's formula, and auxiliary lemmas

2.1. Gradient estimates for the SDE (1.1).

We consider the derivative of X_t^x with respect to the initial value $x \in \mathbb{R}^d$. For $v \in \mathbb{R}^d$, the directional derivative $\nabla_v X_t^x$ in direction v is defined by

$$\nabla_v X_t^x = \lim_{\epsilon\to 0} \frac{X_t^{x+\epsilon v} - X_t^x}{\epsilon}, \quad t \geq 0.$$

The above limit exists and satisfies the formally differentiated SDE

$$\frac{\mathrm{d}}{\mathrm{d}t}\nabla_v X_t^x = \nabla b(X_t^x)\nabla_v X_t^x, \quad \nabla_v X_0^x = v,$$

and we have the following representation of the solution

$$\nabla_v X_t^x = \exp\left[\int_0^t \nabla b(\nabla X_r^x)\,\mathrm{d}r\right] v.$$

In a similar way we can define for $v_i \in \mathbb{R}^d$, $i = 1, 2, 3$ the directional derivatives $\nabla_{v_2}\nabla_{v_1} X_t^x$ and $\nabla_{v_3}\nabla_{v_2}\nabla_{v_1} X_t^x$, which satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t}\nabla_{v_2}\nabla_{v_1} X_t^x = \nabla b(X_t^x)\nabla_{v_2}\nabla_{v_1} X_t^x + \nabla^2 b(X_t^x)\nabla_{v_2} X_t^x \nabla_{v_1} X_t^x$$

with $\nabla_{v_2}\nabla_{v_1} X_0^x = 0$, and

(2.1)

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\nabla_{v_3}\nabla_{v_2}\nabla_{v_1} X_t^x &= \nabla b(X_t^x)\nabla_{v_3}\nabla_{v_2}\nabla_{v_1} X_t^x + \nabla^3 b(X_t^x)\nabla_{v_3} X_t^x \nabla_{v_2} X_t^x \nabla_{v_1} X_t^x \\ &\quad + \nabla^2 b(X_t^x)\left(\nabla_{v_3} X_t^x \nabla_{v_2}\nabla_{v_1} X_t^x + \nabla_{v_2} X_t^x \nabla_{v_3}\nabla_{v_1} X_t^x + \nabla_{v_1} X_t^x \nabla_{v_3}\nabla_{v_2} X_t^x\right),\end{aligned}$$

with $\nabla_{v_3}\nabla_{v_2}\nabla_{v_1} X_0^x = 0$.

Lemma 2.1. *Assume* **(H2)**. *Then for all* $v_i \in \mathbb{R}^d$, $i = 1, 2, 3$, *and* $t \in [0, 1]$,

$$|\nabla_{v_1} X_t^x| \le C|v_1|, \tag{2.2}$$

$$|\nabla_{v_2}\nabla_{v_1} X_t^x| \le C|v_1||v_2|, \tag{2.3}$$

$$|\nabla_{v_3}\nabla_{v_2}\nabla_{v_1} X_t^x| \le C|v_1||v_2||v_3|. \tag{2.4}$$

Proof. The estimates (2.2) and (2.3) are taken from [3, Lemma 4.1]. We can prove (2.4) in a similar way: Set $\zeta(t) := \nabla_{v_3}\nabla_{v_2}\nabla_{v_1} X_t^x$. We get from (2.1), (1.3)–(1.5) and (2.2), (2.3) that

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}|\zeta(t)|^2 \\ &= 2\langle \zeta(t), \nabla b(X_t^x)\zeta(t)\rangle + 2\left\langle \zeta(t), \nabla^3 b(X_t^x)\nabla_{v_3} X_t^x \nabla_{v_2} X_t^x \nabla_{v_1} X_t^x\right\rangle \\ &\quad + 2\left\langle \zeta(t), \nabla^2 b(X_t^x)\left(\nabla_{v_3} X_t^x \nabla_{v_2}\nabla_{v_1} X_t^x + \nabla_{v_2} X_t^x \nabla_{v_3}\nabla_{v_1} X_t^x + \nabla_{v_1} X_t^x \nabla_{v_3}\nabla_{v_2} X_t^x\right)\right\rangle \\ &\le 2\theta_1|\zeta(t)|^2 + 2\theta_3|\zeta(t)||\nabla_{v_1} X_t^x||\nabla_{v_2} X_t^x||\nabla_{v_3} X_t^x| \\ &\quad + 2\theta_2|\zeta(t)|\left(|\nabla_{v_3} X_t^x||\nabla_{v_2}\nabla_{v_1} X_t^x| + |\nabla_{v_2} X_t^x||\nabla_{v_3}\nabla_{v_1} X_t^x| + |\nabla_{v_1} X_t^x||\nabla_{v_3}\nabla_{v_2} X_t^x|\right) \\ &\le 2\theta_1|\zeta(t)|^2 + C|\zeta(t)||v_1||v_2||v_3| \\ &\le C|\zeta(t)|^2 + C|v_1|^2|v_2|^2|v_3|^2.\end{aligned}$$

In the last inequality we use the elementary inequality $\pm 2z_1 z_2 \le z_1^2 + z_2^2$. Noting that $\zeta(0) = 0$, we get for $t > 0$

$$|\zeta(t)|^2 \le C|v_1|^2|v_2|^2|v_3|^2 \int_0^t \mathrm{e}^{C(t-r)}\,\mathrm{d}r = \left(\mathrm{e}^{Ct} - 1\right)|v_1|^2|v_2|^2|v_3|^2.$$

This proves (2.4) for all $t \in [0, 1]$. □

2.2. **Bismut's formula.** Let $u \in L^2_{\mathrm{loc}}([0,\infty) \times (\Omega, \mathscr{F}, \mathbb{P}); \mathbb{R}^d)$, i.e. $\mathbb{E}\int_0^t |u(s)|^2\,\mathrm{d}s < \infty$ for all $t > 0$. Let $(W_t)_{t\ge 0}$ be a standard Brownian motion on $\mathbb{R}^d$ and assume that u is adapted to the filtration $(\mathscr{F}_t)_{t\ge 0}$ with $\mathscr{F}_t := \sigma(W_s : 0 \le s \le t)$, and define $U : [0,\infty) \to \mathbb{R}^d$ by

$$U_t := \int_0^t u(s)\,\mathrm{d}s, \quad t \ge 0.$$

For $t > 0$, let $F_t : C([0,t], \mathbb{R}^d) \to \mathbb{R}^m$ be an $\mathscr{F}_t$ measurable map, where $m \in \mathbb{N}$. If the following limit exists

$$D_U F_t(W) = \lim_{\epsilon\to 0} \frac{F_t(W + \epsilon U) - F_t(W)}{\epsilon}$$

in $L^2((\Omega,\mathscr{F},\mathbb{P});\mathbb{R}^m)$, then $F_t(W)$ is said to be *Malliavin differentiable*, and $D_UF_t(W)$ is called the Malliavin derivative of $F_t(W)$ in the direction U.

Let both $F_t(W)$ and $G_t(W)$ be Malliavin differentiable. Then the following product rule holds:

$$D_U\left(\langle F_t(W),G_t(W)\rangle\right)=\langle D_UF_t(W),G_t(W)\rangle+\langle F_t(W),D_UG_t(W)\rangle.$$

If $F_t(W)$ has the following structure,

$$F_t(W)=\int_0^t a(s)\,\mathrm{d}W_s,$$

where $a(s)=(a_1(s),\dots,a_d(s))$ is a $\mathscr{F}_s$-adapted stochastic process such that $\mathbb{E}\int_0^t|a(s)|^2\,\mathrm{d}s<\infty$ for all $t>0$, then

$$D_UF_t(W)=\int_0^t\langle a(s),u(s)\rangle\,\mathrm{d}s+\int_0^t D_Ua(s)\,\mathrm{d}W_s. \tag{2.5}$$

We will use the following integration by parts formula, which is also called *Bismut's formula.* For any Malliavin differentiable $F_t(W)$ such that $F_t(W),D_UF_t(W)\in L^2((\Omega,\mathscr{F},\mathbb{P});\mathbb{R})$, we have

$$\mathbb{E}[D_UF_t(W)]=\mathbb{E}\left[F_t(W)\int_0^t u(s)\,\mathrm{d}W_s\right]. \tag{2.6}$$

Let $\phi\in\mathrm{Lip}(1)$ and let $F_t(W)$ be a d-dimensional Malliavin differentiable functional. Then the following chain rule holds:

$$D_U\phi(F_t(W))=\langle\nabla\phi(F_t(W)),D_UF_t(W)\rangle.$$

2.3. **Time-change method for the SDE** (1.1)**.** It is well-known that a d-dimensional rotationally symmetric α-stable Lévy process Z_t can be represented as subordinated Brownian motion, see for instance [7, 29, 32]. More precisely, let $(S_t)_{t\ge0}$ be an $\frac{\alpha}{2}$-stable subordinator, i.e. S_t is an $\mathbb{R}^+$-valued Lévy process with the following Laplace transform:

$$\mathbb{E}\left[\mathrm{e}^{-rS_t}\right]=\mathrm{e}^{-2^{-1}t(2r)^{\alpha/2}},\quad r>0,\,t\ge0,$$

and let W_t be a d-dimensional standard Brownian motion, which is independent of S_t. The time-changed process $Z_t:=W_{S_t}$ is a d-dimensional rotationally symmetric α-stable Lévy process such that $\mathbb{E}\mathrm{e}^{\mathrm{i}\,\xi Z_t}=\mathrm{e}^{-t|\xi|^\alpha/2}$. We refer the reader to [28] and the references therein for detailed expositions of Lévy and α-stable Lévy processes.

Using the subordination representation, (1.1) can be written in the following form

$$\mathrm{d}X_t=b(X_t)\,\mathrm{d}t+\sigma\,\mathrm{d}W_{S_t},\quad X_0=x. \tag{2.7}$$

Let $\mathbb{W}$ be the space of all continuous functions from $[0,\infty)$ to $\mathbb{R}^d$ vanishing at starting point 0, which is endowed with the locally uniform convergence topology and the Wiener measure $\mu_{\mathbb{W}}$ so that the coordinate process

$$W_t(w):=w_t$$

is a standard d-dimensional Brownian motion. Let $\mathbb{S}$ be the space of all increasing and càdlàg functions from $[0,\infty)$ to $[0,\infty)$, vanishing at starting point 0, which is endowed with the Skorohod metric and the probability measure $\mu_{\mathbb{S}}$ so that for any $\ell\in\mathbb{S}$ the coordinate process

$$S_t(\ell):=\ell_t$$

is an $\frac{\alpha}{2}$-stable subordinator. Consider the following product probability space

$$(\Omega, \mathscr{F}, \mathbb{P}) := (\mathbb{W} \times \mathbb{S}, \mathscr{B}(\mathbb{W}) \times \mathscr{B}(\mathbb{S}), \mu_{\mathbb{W}} \times \mu_{\mathbb{S}}),$$

and define

$$L_t(w, \ell) := w \circ \ell_t.$$

Then $(L_t)_{t\geq 0}$ is an α-stable process on $(\Omega, \mathscr{F}, \mathbb{P})$. We shall use the following two natural filtrations associated with the Lévy process L_t and the Brownian motion W_t:

$$\mathscr{F}_t := \sigma\{L_s(w,\ell) : s \leq t\} \quad \text{and} \quad \mathscr{F}_t^{\mathbb{W}} := \sigma\{W_s(w) : s \leq t\}.$$

In particular, we can regard the solution X_t^x of the SDE (2.7) as an $(\mathscr{F}_t)$-adapted functional on Ω; therefore,

$$\mathbb{E} f\left(X_t^x\right) = \int_{\mathbb{S}} \int_{\mathbb{W}} f\left(X_t^x(w \circ \ell)\right) \mu_{\mathbb{W}}(\mathrm{d}w)\, \mu_{\mathbb{S}}(\mathrm{d}\ell). \tag{2.8}$$

This relation will allow us to freeze the path of the subordinator and do all calculations for a Brownian motion which is time-changed with a deterministic time-change $\ell = (\ell_t)_{t\geq 0} \in \mathbb{S}$. After that, we only have to make sure that our results remain stable under the integration with respect to $\mu_{\mathbb{S}}(\mathrm{d}\ell)$.

For $\ell \in \mathbb{S}$, let $X_t^{\ell;x}$ denote the solution to the SDE

$$\mathrm{d}X_t^{x;\ell} = b(X_t^{x;\ell})\, \mathrm{d}t + \sigma\, \mathrm{d}W_{\ell_t}, \quad X_0^{x;\ell} = x. \tag{2.9}$$

We will now fix a path $\ell \in \mathbb{S}$ and consider the SDE (2.9). Unless stated otherwise, all expectations are taken on the Wiener space $(\mathbb{W}, \mathscr{B}(\mathbb{W}), \mu_{\mathbb{W}})$. First of all, notice that $t \to W_{\ell_t}$ is a Gaussian process with zero mean and independent increments. In particular, W_{ℓ_t} is a càdlàg $\mathscr{F}_{\ell_t}^{\mathbb{W}}$-martingale. Thus, under **(H1)**, **(H2)**, it is known that for each initial value $x \in \mathbb{R}^d$, the SDE (2.9) admits a unique càdlàg $\mathscr{F}_{\ell_t}^{\mathbb{W}}$-adapted solution $X_t^{x;\ell}$.

In order to prove the gradient estimates with respect to the semigroup P_t, we shall use a (deterministic) time-change argument to transform the SDE (2.9) into an SDE driven by standard Brownian motion, and then use Bismut's formula (2.6). We note that Zhang [32] used this technique to obtain a Bismut–Elworthy–Li formula; in the present paper we need to estimate the third order derivative of the semigroup by repeated applications of Bismut's formula, see Lemma 3.1 below. We refer the reader to [2], and the references therein, for a comprehensive treatment on Bismut's formula

For $\epsilon \in (0,1)$, we define

$$\ell_t^\epsilon := \frac{1}{\epsilon} \int_t^{t+\epsilon} \ell_s\, \mathrm{d}s + \epsilon t = \int_0^1 \ell_{\epsilon s + t}\, \mathrm{d}s + \epsilon t.$$

Since $t \mapsto \ell_t$ is increasing and right-continuous, it follows that for each $t \geq 0$,

$$\ell_t^\epsilon \downarrow \ell_t \quad \text{as} \quad \epsilon \downarrow 0. \tag{2.10}$$

Moreover, $t \mapsto \ell_t^\epsilon$ is absolutely continuous and strictly increasing. Let $X_t^{x;\ell^\epsilon}$ be the solution to the SDE

$$\mathrm{d}X_t^{x;\ell^\epsilon} = b(X_t^{x;\ell^\epsilon})\, \mathrm{d}t + \sigma\, \mathrm{d}W_{\ell_t^\epsilon - \ell_0^\epsilon}, \quad X_0^{x;\ell^\epsilon} = x \tag{2.11}$$

and denote by γ^ϵ the inverse function of ℓ^ϵ, i.e.,

$$\ell^\epsilon_{\gamma^\epsilon_t} = t, \quad t \geq \ell^\epsilon_0 \quad \text{and} \quad \gamma^\epsilon_{\ell^\epsilon_t} = t, \quad t \geq 0.$$

By definition, γ^ϵ_t is absolutely continuous on $[\ell^\epsilon_0, \infty)$. Let us now define

$$Z^{x;\ell^\epsilon}_t := X^{x;\ell^\epsilon}_{\gamma^\epsilon_t}, \quad t \geq \ell^\epsilon_0. \tag{2.12}$$

From equation (2.11) and with a change of variables, we see for $t \geq l^\epsilon_0$,

$$Z^{x;\ell^\epsilon}_t = x + \int_0^{\gamma^\epsilon_t} b\left(X^{x;\ell^\epsilon}_s\right) \mathrm{d}s + \sigma W_{t-\ell^\epsilon_0} = x + \int_{\ell^\epsilon_0}^t b\left(Z^{x;\ell^\epsilon}_s\right) \dot{\gamma}^\epsilon_s \,\mathrm{d}s + \sigma W_{t-\ell^\epsilon_0}.$$

Hence, we have for any vector $v \in \mathbb{R}^d$

$$\nabla_v Z^{x;\ell^\epsilon}_t = v + \int_{\ell^\epsilon_0}^t \nabla b\left(Z^{x;\ell^\epsilon}_s\right) \nabla_v Z^{x;\ell^\epsilon}_s \dot{\gamma}^\epsilon_s \,\mathrm{d}s, \tag{2.13}$$

which implies

$$\nabla_v Z^{x;\ell^\epsilon}_t = \exp\left[\int_{\ell^\epsilon_0}^t \nabla b\left(Z^{x;\ell^\epsilon}_s\right) \dot{\gamma}^\epsilon_s \,\mathrm{d}s\right] v.$$

For further use, we use the shorthand

$$J^{x;l^\epsilon}_{s,t} = \exp\left[\int_s^t \nabla b\left(Z^{x;\ell^\epsilon}_s\right) \dot{\gamma}^\epsilon_s \,\mathrm{d}s\right], \quad \ell^\epsilon_0 \leq s \leq t < \infty. \tag{2.14}$$

It is not hard to see that $J^{x;\ell^\epsilon}_{s,t} J^{x;\ell^\epsilon}_{\ell^\epsilon_0,s} = J^{x;\ell^\epsilon}_{\ell^\epsilon_0,t}$ for all $\ell^\epsilon_0 \leq s \leq t < \infty$ and

$$\nabla_v Z^{x;\ell^\epsilon}_t = J^{x;\ell^\epsilon}_{\ell^\epsilon_0,t} v.$$

Lemma 2.2. *Assume* **(H2)**. *For all $v_1, v_2 \in \mathbb{R}^d$ and $t \in [\ell^\epsilon_0, \ell^\epsilon_1]$,*

$$|\nabla_{v_1} Z^{x;\ell^\epsilon}_t| \leq C|v_1|, \tag{2.15}$$

$$|\nabla_{v_2}\nabla_{v_1} Z^{x;\ell^\epsilon}_t| \leq C|v_1||v_2|. \tag{2.16}$$

Proof. From the expression for $\nabla_{v_1} Z^{x;\ell^\epsilon}_t$ and (1.3) we get

$$|\nabla_{v_1} Z^{x;\ell^\epsilon}_t| \leq \exp\left[\theta_1 \int_{\ell^\epsilon_0}^t \dot{\gamma}^\epsilon_s \,\mathrm{d}s\right] |v_1| = \mathrm{e}^{\theta_1 \gamma^\epsilon_t}|v_1|.$$

This implies the first estimate for $t \in [\ell^\epsilon_0, \ell^\epsilon_1]$.

Let $\xi(t) := \nabla_{v_2}\nabla_{v_1} Z^{x;\ell^\epsilon}_t$. Since

$$\frac{\mathrm{d}}{\mathrm{d}t}\xi(t) = \dot{\gamma}^\epsilon_t \nabla b(Z^{x;\ell^\epsilon}_t)\xi(t) + \dot{\gamma}^\epsilon_t \nabla^2 b(Z^{x;\ell^\epsilon}_t)\nabla_{v_2} Z^{x;\ell^\epsilon}_t \nabla_{v_1} Z^{x;\ell^\epsilon}_t,$$

we see from (1.3), (1.4), and the already proved estimate, that for $t \in [\ell^\epsilon_0, \ell^\epsilon_1]$,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}|\xi(t)|^2 &= 2\langle \xi(t), \dot{\gamma}^\epsilon_t \nabla b(Z^{x;\ell^\epsilon}_t)\xi(t)\rangle + 2\langle \xi(t), \dot{\gamma}^\epsilon_t \nabla^2 b(Z^{x;\ell^\epsilon}_t)\nabla_{v_2} Z^{x;\ell^\epsilon}_t \nabla_{v_1} Z^{x;\ell^\epsilon}_t\rangle \\
&\leq 2\theta_1 \dot{\gamma}^\epsilon_t |\xi(t)|^2 + 2\theta_2 \dot{\gamma}^\epsilon_t |\nabla_{v_2} Z^{x;\ell^\epsilon}_t||\nabla_{v_1} Z^{x;\ell^\epsilon}_t||\xi(t)| \\
&\leq 2\theta_1 \dot{\gamma}^\epsilon_t |\xi(t)|^2 + C|v_1||v_2|\dot{\gamma}^\epsilon_t |\xi(t)| \\
&\leq C\dot{\gamma}^\epsilon_t |\xi(t)|^2 + C|v_1|^2|v_2|^2 \dot{\gamma}^\epsilon_t.
\end{aligned}$$

In the last estimate we use $2ab \le a^2+b^2$, taking $a = |v_1||v_2|\sqrt{\dot\gamma_t^\epsilon}$ and $b = \sqrt{\dot\gamma_t^\epsilon}|\xi(t)|$. Noting that $\xi(\ell_0^\epsilon)=0$, this yields that for all $t\in[\ell_0^\epsilon,\ell_1^\epsilon]$,

$$\begin{aligned}
|\xi(t)|^2 &\le C|v_1|^2|v_2|^2\int_{\ell_0^\epsilon}^t \exp\left[C\int_r^t \dot\gamma_s^\epsilon\,\mathrm{d}s\right]\dot\gamma_r^\epsilon\,\mathrm{d}r\\
&= C|v_1|^2|v_2|^2\int_{\ell_0^\epsilon}^t \mathrm{e}^{C(\gamma_t^\epsilon-\gamma_r^\epsilon)}\dot\gamma_r^\epsilon\,\mathrm{d}r\\
&= C|v_1|^2|v_2|^2\int_0^{\gamma_t^\epsilon} \mathrm{e}^{C(\gamma_t^\epsilon-r)}\,\mathrm{d}r\\
&= C|v_1|^2|v_2|^2\int_0^{\gamma_t^\epsilon} \mathrm{e}^{Cr}\,\mathrm{d}r\\
&= |v_1|^2|v_2|^2\left(\mathrm{e}^{C\gamma_t^\epsilon}-1\right)\\
&\le \left(\mathrm{e}^{C}-1\right)|v_1|^2|v_2|^2.
\end{aligned}$$

This completes the proof. □

We will now use the tools from Malliavin calculus described in Section 2.2. Fixing $t\ge\ell_0^\epsilon$ and $x\in\mathbb{R}^d$, the solution $Z_t^{x;\ell^\epsilon}$ is a d-dimensional functional of Brownian motion $(W_{s-\ell_0^\epsilon})_{\ell_0^\epsilon\le s\le t}$. Recall the definition of U in Section 2.2. The Malliavin derivative of $Z_t^{x;\ell^\epsilon}$ along the direction U exists in $L^2\left((\mathbb{W},\mathscr{B}(\mathbb{W}),\mu_{\mathbb{W}});\mathbb{R}^d\right)$ and is given by

$$D_U Z_t^{x;\ell^\epsilon}(W) = \lim_{\delta\to0}\frac{Z_t^{x;\ell^\epsilon}(W+\delta U)-Z_t^{x;\ell^\epsilon}(W)}{\delta}.$$

We drop "W" and write $D_U Z_t^{x;\ell^\epsilon} = D_U Z_t^{x;\ell^\epsilon}(W)$, to keep the notation simple; we have

$$D_U Z_t^{x;\ell^\epsilon} = \int_{\ell_0^\epsilon}^t\left[\nabla b\left(Z_s^{x;\ell^\epsilon}\right)D_U Z_s^{x;\ell^\epsilon}\dot\gamma_s^\epsilon + u(s)\right]\mathrm{d}s,$$

and this integral equation has a unique solution:

$$D_U Z_t^{x;\ell^\epsilon} = \int_{\ell_0^\epsilon}^t J_{s,t}^{x;\ell^\epsilon}u(s)\,\mathrm{d}s,$$

where $J_{s,t}^{x;\ell^\epsilon}$ is defined by (2.14).

For $t>\ell_0^\epsilon$ and $v_1,v_2,v_3,x\in\mathbb{R}^d$, define $u_{t,i},U_{t,i}:[0,\infty)\to\mathbb{R}^d$ by

$$u_{t,i}(s) := \frac{1}{t-\ell_0^\epsilon}\nabla_{v_i}Z_s^{x;\ell^\epsilon}\mathbb{1}_{[\ell_0^\epsilon,t]}(s)\quad\text{and}\quad U_{t,i;s} := \int_0^s u_{t,i}(r)\,\mathrm{d}r$$

for $i=1,2,3$. Clearly,

$$D_{U_{t,i}}Z_s^{x;\ell^\epsilon} = \frac{s-\ell_0^\epsilon}{t-\ell_0^\epsilon}\nabla_{v_i}Z_s^{x;\ell^\epsilon},\quad \ell_0^\epsilon\le s\le t. \tag{2.17}$$

This, together with (2.13), implies that for $s\in[\ell_0^\epsilon,t]$

(2.18)

$$D_{U_{t,2}}\nabla_{v_1}Z_s^{x;\ell^\epsilon} = \int_{\ell_0^\epsilon}^s\left[\nabla^2 b\left(Z_r^{x;\ell^\epsilon}\right)D_{U_{t,2}}Z_r^{x;\ell^\epsilon}\nabla_{v_1}Z_r^{x;\ell^\epsilon} + \nabla b\left(Z_r^{x;\ell^\epsilon}\right)D_{U_{t,2}}\nabla_{v_1}Z_r^{x;\ell^\epsilon}\right]\dot\gamma_r^\epsilon\,\mathrm{d}r.$$

The argument which we used in the proofs of Lemma 2.1 and 2.2 gives the following upper bounds on the Malliavin derivatives.

Lemma 2.3. *Assume* **(H2)**. *For all* $v_i \in \mathbb{R}^d$, $i = 1, 2, 3$, *and* $t \in [\ell_0^\epsilon, \ell_1^\epsilon]$,

$$|D_{U_{t,2}} \nabla_{v_1} Z_t^{x;\ell^\epsilon}| \le C|v_1||v_2|, \tag{2.19}$$

$$|D_{U_{t,3}} \nabla_{v_2} \nabla_{v_1} Z_t^{x;\ell^\epsilon}| \le C|v_1||v_2||v_3|, \tag{2.20}$$

$$|D_{U_{t,1}} D_{U_{t,3}} \nabla_{v_2} Z_t^{x;\ell^\epsilon}| \le C|v_1||v_2||v_3|. \tag{2.21}$$

Proof. From (2.18) and (2.17) we see that

$$\frac{\mathrm{d}}{\mathrm{d}t} D_{U_{t,2}} \nabla_{v_1} Z_t^{x;\ell^\epsilon} = \dot{\gamma}_t^\epsilon \nabla b(Z_t^{x;\ell^\epsilon}) D_{U_{t,2}} \nabla_{v_1} Z_t^{x;\ell^\epsilon} + \dot{\gamma}_t^\epsilon \nabla^2 b(Z_t^{x;\ell^\epsilon}) \nabla_{v_2} Z_t^{x;\ell^\epsilon} \nabla_{v_1} Z_t^{x;\ell^\epsilon}.$$

Repeating the argument used in the proof of Lemma 2.2, we get (2.19) for all $t \in [\ell_0^\epsilon, \ell_1^\epsilon]$. The estimates (2.20) and (2.21) can be proved in a similar way. □

3. Proof of Theorem 1.1

For the proof of Theorem 1.1 we need some preparations. First, the following gradient estimates are crucial.

Lemma 3.1. *Assume* **(H2)** *and let* $\alpha_0 \in (1, 2)$ *be an arbitrary fixed number. For all* $h \in \mathrm{Lip}(1)$, $t \in (0, 1]$, *and* $\alpha \in [\alpha_0, 2)$,

$$\|\nabla P_t h\|_\infty \le C, \tag{3.1}$$

$$\|\nabla^2 P_t h\|_{\mathrm{op},\infty} \le C_{\alpha_0} t^{-1/\alpha}, \tag{3.2}$$

$$\|\nabla^3 P_t h\|_{\mathrm{op},\infty} \le C_{\alpha_0} t^{-2/\alpha}. \tag{3.3}$$

Remark 3.2. The gradient estimates (3.1) and (3.2) are essentially taken from [3, Lemma 3.1]; for the reader's convenience and for completeness we sketch the proof in Section 4. As far as we know, (3.3) is new, and we will prove it in the same spirit.

Lemma 3.3. *Let* $A(d, \alpha)$ *and* ω_{d-1} *be as in* (1.10). *For* $\alpha \in (0, 2)$,

$$\left| \frac{A(d, \alpha)\omega_{d-1}}{d(2-\alpha)} - 1 \right| \le C \log(1 + d)(2 - \alpha).$$

Proof. Note that

$$\frac{A(d, \alpha)\omega_{d-1}}{d(2-\alpha)} = \frac{\alpha\Gamma\left(\frac{d+\alpha}{2}\right)}{d(2-\alpha)2^{1-\alpha}\Gamma\left(1-\frac{\alpha}{2}\right)\Gamma\left(\frac{d}{2}\right)} = \frac{\alpha\Gamma\left(\frac{d+\alpha}{2}\right)}{d2^{2-\alpha}\Gamma\left(2-\frac{\alpha}{2}\right)\Gamma\left(\frac{d}{2}\right)}.$$

Set

$$\rho(x) := x\Gamma\left(\frac{d+x}{2}\right) - d2^{2-x}\Gamma\left(2-\frac{x}{2}\right)\Gamma\left(\frac{d}{2}\right), \quad 0 \le x \le 2.$$

It is easy to see that $\rho(2) = 0$ and

$$\begin{aligned} \rho'(x) &= \Gamma\left(\frac{d+x}{2}\right) + \frac{x}{2}\Gamma'\left(\frac{d+x}{2}\right) \\ &\quad + d \cdot \Gamma\left(\frac{d}{2}\right)\left\{\log 2 \cdot 2^{2-x}\Gamma\left(2-\frac{x}{2}\right) + 2^{1-x}\Gamma'\left(2-\frac{x}{2}\right)\right\} \\ &= \Gamma\left(\frac{d+x}{2}\right) + \frac{x}{2}\psi\left(\frac{d+x}{2}\right)\Gamma\left(\frac{d+x}{2}\right) \end{aligned}$$

$$+ d\cdot\Gamma\left(\frac{d}{2}\right)\left\{\log 2\cdot 2^{2-x}\Gamma\left(2-\frac{x}{2}\right)+2^{1-x}\psi\left(2-\frac{x}{2}\right)\Gamma\left(2-\frac{x}{2}\right)\right\},$$

where $\psi(x)=\Gamma'(x)/\Gamma(x)$ is the Psi or Digamma function. Since $\lim_{z\to\infty}\psi(z)/\log z = 1$, cf. [1, 6.3.18, p. 259], we have for all $x\in[0,2]$,

$$|\rho'(x)|\le Cd\cdot\Gamma\left(\frac{d}{2}\right)\log(1+d).$$

From this we conclude that for all $\alpha\in(0,2)$

$$\begin{aligned}\left|\frac{A(d,\alpha)\omega_{d-1}}{d(2-\alpha)}-1\right| &= \frac{|\rho(\alpha)|}{d2^{2-\alpha}\Gamma\left(2-\frac{\alpha}{2}\right)\Gamma\left(\frac{d}{2}\right)}\\ &= \frac{|\rho(2)-\rho(\alpha)|}{d2^{2-\alpha}\Gamma\left(2-\frac{\alpha}{2}\right)\Gamma\left(\frac{d}{2}\right)}\\ &\le C\log(1+d)(2-\alpha).\end{aligned}$$

□

The next lemma is a key step in proving our main results. Note that the terms J_{11}, which appears in the proof below, has a sharp estimate, depending on $d\cdot\log(1+d)$. This seems to indicate that the bound in Theorem 1.1 (1.6) cannot be improved.

Lemma 3.4. *Assume that* **(H2)** *holds and let $\alpha_0\in(1,2)$ be an arbitrary fixed number. For any $\alpha\in[\alpha_0,2)$ and $s\in(0,1]$, we have*

$$\begin{aligned}&\sup_{h\in\mathrm{Lip}(1)}\left|(\mathscr{A}^P-\mathscr{A}^Q)P_s h\right|\\ &\quad\le C_{\alpha_0}d\cdot\log(1+d)\Big[(2-\alpha)s^{-1/\alpha}+\left\{\left[(2-\alpha)s^{-2/\alpha}\right]\wedge s^{-1/\alpha}\right\}\Big].\end{aligned}$$

Proof. Set $f:=P_s h$. Since

$$\mathscr{A}^P f(x)=\langle\nabla f(x),b(x)\rangle+\int_{\mathbb{R}^d\setminus\{0\}}\left[f(x+\sigma z)-f(x)-\langle\nabla f(x),\sigma z\rangle\mathbb{1}_{\{|z|\le1\}}\right]\frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z,$$

and

$$\mathscr{A}^Q f(x)=\langle\nabla f(x),b(x)\rangle+\frac{1}{2}\langle\nabla^2 f(x),\sigma\sigma^\top\rangle_{\mathrm{HS}},$$

it follows that

$$\begin{aligned}&(\mathscr{A}^P-\mathscr{A}^Q)f(x)\\ &=\left\{\int_{|z|\le1}\left[f(x+\sigma z)-f(x)-\langle\nabla f(x),\sigma z\rangle\right]\frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z-\frac{1}{2}\langle\nabla^2 f(x),\sigma\sigma^\top\rangle_{\mathrm{HS}}\right\}\\ &\quad+\int_{|z|>1}\left[f(x+\sigma z)-f(x)\right]\frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z\\ &=:\mathsf{J}_1+\mathsf{J}_2.\end{aligned}$$

With Lemma 3.3 it is easy to check that for any $\alpha\in[\alpha_0,2)$ the following estimate holds:

$$\frac{A(d,\alpha)\omega_{d-1}}{\alpha-1}\le C_{\alpha_0}d\cdot\log(1+d)(2-\alpha).$$

Because of (3.1), $f = P_t h \in \mathrm{Lip}$ if $h \in \mathrm{Lip}(1)$, and so

$$\begin{aligned}|\mathsf{J}_2| &\le \int_{|z|>1} |f(x+\sigma z) - f(x)| \frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z \\ &\le C \int_{|z|>1} |z| \frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z \\ &= C\frac{A(d,\alpha)\omega_{d-1}}{\alpha-1} \\ &\le C_{\alpha_0} d \cdot \log(1+d)(2-\alpha).\end{aligned}$$

We rewrite J_1 in the following form:

$$\begin{aligned}\mathsf{J}_1 &= \int_{|z|\le 1} \left[f(x+\sigma z) - f(x) - \langle \nabla f(x), \sigma z\rangle\right] \frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z - \frac{1}{2}\langle \nabla^2 f(x), \sigma\sigma^\top\rangle_{\mathrm{HS}} \\ &= \int_{|z|\le 1}\int_0^1 \langle \nabla^2 f(x+r\sigma z), (\sigma z)(\sigma z)^\top\rangle_{\mathrm{HS}}(1-r)\,\mathrm{d}r \frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z - \frac{1}{2}\langle \nabla^2 f(x), \sigma\sigma^\top\rangle_{\mathrm{HS}} \\ &= \left\{\int_{|z|\le 1}\int_0^1 \langle \nabla^2 f(x), (\sigma z)(\sigma z)^\top\rangle_{\mathrm{HS}}(1-r)\,\mathrm{d}r \frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z - \frac{1}{2}\langle \nabla^2 f(x), \sigma\sigma^\top\rangle_{\mathrm{HS}}\right\} \\ &\quad + \int_{|z|\le 1}\int_0^1 \langle \nabla^2 f(x+r\sigma z) - \nabla^2 f(x), (\sigma z)(\sigma z)^\top\rangle_{\mathrm{HS}}(1-r)\,\mathrm{d}r\, \frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z \\ &=: \mathsf{J}_{11} + \mathsf{J}_{12}.\end{aligned}$$

Using the symmetry of the measure $\rho(\mathrm{d}z) = |z|^{-d-\alpha}\mathrm{d}z$, it is clear that $\int_{|z|\le 1} z_i z_j\, \rho(\mathrm{d}z) = \delta_{ij}\frac{1}{d}\int_{|z|\le 1} |z|^2\, \rho(\mathrm{d}z)$, and so we get

$$\begin{aligned}&\int_{|z|\le 1}\int_0^1 \langle \nabla^2 f(x), (\sigma z)(\sigma z)^\top\rangle_{\mathrm{HS}}(1-r)\,\mathrm{d}r \frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z \\ &\quad = \frac{1}{2}\int_{|z|\le 1} \langle \nabla^2 f(x), \sigma(zz^\top)\sigma^\top\rangle_{\mathrm{HS}} \frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z \\ &\quad = \frac{1}{2}\frac{1}{d}\int_{|z|\le 1} \langle \nabla^2 f(x), |z|^2\sigma\sigma^\top\rangle_{\mathrm{HS}} \frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z \\ &\quad = \frac{1}{2d}\langle \nabla^2 f(x), \sigma\sigma^\top\rangle_{\mathrm{HS}} \int_{|z|\le 1} \frac{A(d,\alpha)}{|z|^{d+\alpha-2}}\,\mathrm{d}z \\ &\quad = \frac{A(d,\alpha)\omega_{d-1}}{2d(2-\alpha)}\langle \nabla^2 f(x), \sigma\sigma^\top\rangle_{\mathrm{HS}}.\end{aligned}$$

Combining this with Lemma 3.3 and (3.2) gives

$$\begin{aligned}|\mathsf{J}_{11}| &= \left|\int_{|z|\le 1}\int_0^1 \langle \nabla^2 f(x), (\sigma z)(\sigma z)^\top\rangle_{\mathrm{HS}}(1-r)\,\mathrm{d}r \frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z - \frac{1}{2}\langle \nabla^2 f(x), \sigma\sigma^\top\rangle_{\mathrm{HS}}\right| \\ &= \left|\frac{1}{2}\left[\frac{A(d,\alpha)\omega_{d-1}}{d(2-\alpha)} - 1\right]\langle \nabla^2 f(x), \sigma\sigma^\top\rangle_{\mathrm{HS}}\right| \\ &\le C\log(1+d)(2-\alpha)\left|\langle \nabla^2 f(x), \sigma\sigma^\top\rangle_{\mathrm{HS}}\right| \\ &\le C\log(1+d)(2-\alpha)\|\nabla^2 f\|_{\mathrm{op},\infty} \\ &\le C\log(1+d)(2-\alpha)s^{-1/\alpha}.\end{aligned}$$

Now we turn to the estimate of J_{12}. Since $f = P_s h$ and $h \in \mathrm{Lip}(1)$, we can use (3.2) and (3.3) to see that for all $x \in \mathbb{R}^d$,

$$\begin{aligned}\left\|\nabla^2 f(x+r\sigma z) - \nabla^2 f(x)\right\|_{\mathrm{op}} &\le C\left\{\left(\|\nabla^3 f\|_{\mathrm{op},\infty} r|\sigma z|\right) \wedge \|\nabla^2 f\|_{\mathrm{op},\infty}\right\}\\ &\le C\left\{\left(s^{-2/\alpha} r|z|\right) \wedge s^{-1/\alpha}\right\}.\end{aligned}$$

Using Lemma 3.3, we have for $\alpha \in [\alpha_0, 2)$,

$$\frac{A(d,\alpha)\omega_{d-1}}{3-\alpha} \le C_{\alpha_0} d \cdot \log(1+d)(2-\alpha) \quad \text{and} \quad \frac{A(d,\alpha)\omega_{d-1}}{2-\alpha} \le C_{\alpha_0} d \cdot \log(1+d).$$

Therefore,

$$\begin{aligned}|\mathsf{J}_{12}| &\le \int_{|z|\le 1}\int_0^1 \left|\langle \nabla^2 f(x+r\sigma z) - \nabla^2 f(x), (\sigma z)(\sigma z)^\top\rangle_{\mathrm{HS}}\right| (1-r)\,\mathrm{d}r\, \frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z\\
&\le C\int_{|z|\le 1}\int_0^1 \left\{\left(s^{-2/\alpha} r|z|\right)\wedge s^{-1/\alpha}\right\}|z|^2(1-r)\,\mathrm{d}r\,\frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z\\
&\le C\int_{|z|\le 1}\left\{\left(s^{-2/\alpha}|z|\right)\wedge s^{-1/\alpha}\right\}|z|^2\frac{A(d,\alpha)}{|z|^{d+\alpha}}\,\mathrm{d}z\\
&\le C\left\{\left(s^{-2/\alpha}\int_{|z|\le 1}\frac{A(d,\alpha)}{|z|^{d+\alpha-3}}\,\mathrm{d}z\right)\wedge\left(s^{-1/\alpha}\int_{|z|\le 1}\frac{A(d,\alpha)}{|z|^{d+\alpha-2}}\,\mathrm{d}z\right)\right\}\\
&= C\left\{\left(\frac{A(d,\alpha)\omega_{d-1}}{3-\alpha}\,s^{-2/\alpha}\right)\wedge\left(\frac{A(d,\alpha)\omega_{d-1}}{2-\alpha}\,s^{-1/\alpha}\right)\right\}\\
&\le C_{\alpha_0} d\cdot\log(1+d)\left\{\left[(2-\alpha)s^{-2/\alpha}\right]\wedge s^{-1/\alpha}\right\}.\end{aligned}$$

Combining all estimates, we get for all $x \in \mathbb{R}^d$ and $s \in (0,1]$,

$$\begin{aligned}&\left|(\mathscr{A}^P - \mathscr{A}^Q)f(x)\right|\\
&\quad\le |\mathsf{J}_{11}| + |\mathsf{J}_{12}| + |\mathsf{J}_2|\\
&\quad\le C_{\alpha_0} d\cdot\log(1+d)\left[(2-\alpha)s^{-1/\alpha} + \left\{\left[(2-\alpha)s^{-2/\alpha}\right]\wedge s^{-1/\alpha}\right\} + 2-\alpha\right]\\
&\quad\le C_{\alpha_0} d\cdot\log(1+d)(2-\alpha)s^{-1/\alpha} + C_{\alpha_0} d\cdot\log(1+d)\left\{\left[(2-\alpha)s^{-2/\alpha}\right]\wedge s^{-1/\alpha}\right\},\end{aligned}$$

which implies the claimed estimate. □

Lemma 3.5. *Assume that both* **(H1)** *and* **(H2)** *hold true, and let $\alpha_0 \in (1,2)$ be arbitrary. For all $t > 0$ and $\alpha \in [\alpha_0, 2)$,*

$$\sup_{h\in\mathrm{Lip}(1)}\left|\int_0^{t\wedge 1} Q_{t-s}(\mathscr{A}^P - \mathscr{A}^Q)P_s h\,\mathrm{d}s\right| \le C_{\alpha_0} d\cdot\log(1+d)(2-\alpha)\log\frac{1}{2-\alpha}.$$

Proof. By Lemma 3.4,

$$\begin{aligned}&\left|\int_0^{t\wedge 1} Q_{t-s}(\mathscr{A}^P - \mathscr{A}^Q)P_s h\,\mathrm{d}s\right| \le \int_0^{t\wedge 1}\left|Q_{t-s}(\mathscr{A}^P - \mathscr{A}^Q)P_s h\right|\mathrm{d}s\\
&\le C_{\alpha_0} d\cdot\log(1+d)\int_0^1\left\{(2-\alpha)s^{-1/\alpha} + \left[(2-\alpha)s^{-2/\alpha}\right]\wedge s^{-1/\alpha}\right\}\,\mathrm{d}s\\
&= C_{\alpha_0} d\cdot\log(1+d)\left[\int_0^1 (2-\alpha)s^{-1/\alpha}\,\mathrm{d}s + \int_0^{(2-\alpha)^\alpha} s^{-1/\alpha}\,\mathrm{d}s + \int_{(2-\alpha)^\alpha}^1 (2-\alpha)s^{-2/\alpha}\,\mathrm{d}s\right]\\
&\le C_{\alpha_0} d\cdot\log(1+d)\left\{(2-\alpha) + (2-\alpha)^{\alpha-1} + \left[(2-\alpha)^{\alpha-2} - 1\right]\right\}\end{aligned}$$

$$\le C_{\alpha_0} d \cdot \log(1+d)(2-\alpha)\log\frac{1}{2-\alpha},$$

where the last inequality follows from

$$\lim_{\alpha\uparrow 2}\frac{(2-\alpha)^{\alpha-1}}{(2-\alpha)\log\frac{1}{2-\alpha}}=0 \quad\text{and}\quad \lim_{\alpha\uparrow 2}\frac{(2-\alpha)^{\alpha-2}-1}{(2-\alpha)\log\frac{1}{2-\alpha}}=1.$$

This completes the proof. □

Finally, the following result can be found in [3, Proposition 2.2].

Lemma 3.6. *Assume that both* **(H1)** *and* **(H2)** *hold. Then there exist constants* $C_i = C_i(\theta_0, \theta_1, K, \|\sigma\|_{\mathrm{HS}}) > 0$, $i = 1, 2$, *such that for all* $x, y \in \mathbb{R}^d$ *and* $t > 0$,

$$W_1\left(\mathrm{law}(X_t^x), \mathrm{law}(X_t^y)\right) \le C_1 \mathrm{e}^{-C_2 t}|x-y|.$$

After these preparations we can now proceed with the proof of Theorem 1.1.

Proof of Theorem 1.1. (1) Note that

$$W_1\left(\mathrm{law}(X_t^x), \mathrm{law}(Y_t^y)\right) \le W_1\left(\mathrm{law}(X_t^x), \mathrm{law}(X_t^y)\right) + W_1\left(\mathrm{law}(X_t^y), \mathrm{law}(Y_t^y)\right).$$

In view of Lemma 3.6, it suffices to prove that

$$W_1\left(\mathrm{law}(X_t^y), \mathrm{law}(Y_t^y)\right) \le C_{\alpha_0} d \cdot \log(1+d)(2-\alpha)\log\frac{1}{2-\alpha}. \tag{3.4}$$

Recall that

$$W_1\left(\mathrm{law}(X_t^y), \mathrm{law}(Y_t^y)\right) = \sup_{h\in \mathrm{Lip}(1)} |P_t h(y) - Q_t h(y)|$$

and use that

$$P_t h - Q_t h = \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} Q_{t-s} P_s h \,\mathrm{d}s = \int_0^t Q_{t-s}(\mathscr{A}^P - \mathscr{A}^Q) P_s h \,\mathrm{d}s,$$

to get

$$W_1\left(\mathrm{law}(X_t^y), \mathrm{law}(Y_t^y)\right) = \sup_{h\in \mathrm{Lip}(1)} \left| \int_0^t Q_{t-s}(\mathscr{A}^P - \mathscr{A}^Q) P_s h(y) \,\mathrm{d}s \right|. \tag{3.5}$$

If $t \in (0, 1]$, the desired estimate (3.4) follows immediately from (3.5) and Lemma 3.5. We will now consider the case $t > 1$. By (3.5),

$$\begin{aligned} W_1\left(\mathrm{law}(X_t^y), \mathrm{law}(Y_t^y)\right) &\le \sup_{h\in \mathrm{Lip}(1)} \left| \int_0^1 Q_{t-s}(\mathscr{A}^P - \mathscr{A}^Q) P_s h(y) \,\mathrm{d}s \right| \\ &\quad + \sup_{h\in \mathrm{Lip}(1)} \left| \int_1^t Q_{t-s}(\mathscr{A}^P - \mathscr{A}^Q) P_s h(y) \,\mathrm{d}s \right| \\ &=: \mathsf{I}_1 + \mathsf{I}_2. \end{aligned}$$

Lemma 3.5 shows that for all $y \in \mathbb{R}^d$,

$$\mathsf{I}_1 \le C_{\alpha_0} d \cdot \log(1+d)(2-\alpha)\log\frac{1}{2-\alpha}.$$

Lemma 3.6 yields for $h \in \mathrm{Lip}(1)$ and $s > 1$,

$$|P_{s-1}h(x) - P_{s-1}h(y)| \le \mathbb{E}\left|h(X_{s-1}^x) - h(X_{s-1}^y)\right| \le C_1 \mathrm{e}^{-C_2(s-1)}|x-y|, \quad x, y \in \mathbb{R}^d.$$

Combining this with the semigroup property and Lemma 3.4 with $s = 1$, gives for all $y \in \mathbb{R}^d$,

$$\begin{aligned}
\mathsf{I}_2 &= \sup_{h\in \mathrm{Lip}(1)} \left| \int_1^t Q_{t-s}(\mathscr{A}^P - \mathscr{A}^Q) P_1(P_{s-1}h)(y)\,\mathrm{d}s \right| \\
&\le \sup_{g\in \mathrm{Lip}(1)} \left| \int_1^t C_1 \mathrm{e}^{-C_2(s-1)} Q_{t-s}(\mathscr{A}^P - \mathscr{A}^Q) P_1 g(y)\,\mathrm{d}s \right| \\
&\le C_{\alpha_0} d \cdot \log(1+d)(2-\alpha) \int_1^t \mathrm{e}^{-C_2(s-1)}\,\mathrm{d}s \\
&\le C_{\alpha_0} d \cdot \log(1+d)(2-\alpha) \int_0^\infty \mathrm{e}^{-C_2 s}\,\mathrm{d}s \\
&\le C_{\alpha_0} d \cdot \log(1+d)(2-\alpha).
\end{aligned}$$

Combining these two estimates implies (3.4) for $t > 1$. This proves the first assertion.

(2) From the classical ergodic theory for Markov processes, see e.g. [21], it follows that for all $x \in \mathbb{R}^d$

$$\lim_{t\to\infty} W_1\left(\mathrm{law}(X_t^x), \mu_\alpha\right) = 0,$$

see e.g. [3, (1.10)] for further details. Using a similar argument, it is not hard to verify that

$$\lim_{t\to\infty} W_1\left(\mathrm{law}(Y_t^y), \mu\right) = 0, \quad y \in \mathbb{R}^d.$$

Since

$$W_1(\mu_\alpha, \mu) \le W_1\left(\mu_\alpha, \mathrm{law}(X_t^x)\right) + W_1\left(\mathrm{law}(X_t^x), \mathrm{law}(Y_t^y)\right) + W_1\left(\mathrm{law}(Y_t^y), \mu\right),$$

the second assertion follows immediately from the first one with $t \to \infty$. □

4. Proof of Lemma 3.1

We will frequently use the following mollifier: Let g_δ be the density of the d-dimensional normal distribution $N(0, \delta^2 I_d)$, $\delta > 0$. We define for every $h \in \mathrm{Lip}(1)$

$$h_\delta(x) := \int_{\mathbb{R}^d} g_\delta(y) h(x-y)\,\mathrm{d}y.$$

It is easy to see that h_δ is smooth, $\lim_{\delta\downarrow 0} h_\delta(x) = h(x)$ for all $x \in \mathbb{R}^d$, and

$$\|\nabla h_\delta\|_\infty \le \|\nabla h\|_\infty \le 1.$$

4.1. **Proof of** (3.1)**.** The dominated convergence theorem and (2.2) imply for any $v \in \mathbb{R}^d$ and $t \in [0, 1]$

$$|\nabla_v P_t h_\delta(x)| = |\mathbb{E}[\nabla h_\delta(X_t^x) \nabla_v X_t^x]| \le \|\nabla h_\delta\|_\infty \mathbb{E}|\nabla_v X_t^x| \le C|v|.$$

Since it holds from the dominated convergence theorem that

$$\lim_{\delta\downarrow 0} \nabla_v P_t h_\delta(x) = \nabla_v P_t h(x),$$

the desired estimate follows by letting $\delta \downarrow 0$.

4.2. **Proof of** (3.2)**.** We use the dominated convergence theorem to see that for all $v_1, v_2 \in \mathbb{R}^d$

$$\nabla_{v_2}\nabla_{v_1}\mathbb{E}\left[h_\delta(X_t^x)\right] = \mathbb{E}\left[\nabla h_\delta(X_t^x)\nabla_{v_2}\nabla_{v_1}X_t^x\right] + \mathbb{E}\left[\nabla^2 h_\delta(X_t^x)\nabla_{v_2}X_t^x\nabla_{v_1}X_t^x\right]. \tag{4.1}$$

Now we use (2.3) to get for $t \in [0,1]$,

$$\left|\mathbb{E}\left[\nabla h_\delta(X_t^x)\nabla_{v_2}\nabla_{v_1}X_t^x\right]\right| \le \|\nabla h_\delta\|_\infty \mathbb{E}\left[|\nabla_{v_2}\nabla_{v_1}X_t^x|\right] \le C|v_1||v_2|. \tag{4.2}$$

Recall from Section 2.3, (2.12), that $Z_t^{x;\ell^\epsilon} := X_{\gamma_t^\epsilon}^{x;\ell^\epsilon}$. From (2.17) and (2.6) we get

$$\begin{aligned}
&\mathbb{E}\left[\nabla^2 h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_2}Z_t^{x;\ell^\epsilon}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]\\
&\quad= \mathbb{E}\left[\nabla^2 h_\delta(Z_t^{x;\ell^\epsilon})D_{U_{t,2}}Z_t^{x;\ell^\epsilon}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]\\
&\quad= \mathbb{E}\left[D_{U_{t,2}}\left(\nabla h_\delta(Z_t^{x;\ell^\epsilon})\right)\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]\\
&\quad= \mathbb{E}\left[D_{U_{t,2}}\left(\nabla h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right)\right] - \mathbb{E}\left[\nabla h_\delta(Z_t^{x;\ell^\epsilon})D_{U_{t,2}}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]\\
&\quad= \frac{1}{t-\ell_0^\epsilon}\,\mathbb{E}\left[\nabla h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_1}Z_t^{x;\ell^\epsilon}\int_{\ell_0^\epsilon}^t \nabla_{v_2}Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right] - \mathbb{E}\left[\nabla h_\delta(Z_t^{x;\ell^\epsilon})D_{U_{t,2}}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right].
\end{aligned}$$

Combining this with (2.15) and (2.19), Itô's isometry shows for $t \in (\ell_0^\epsilon, \ell_1^\epsilon]$

$$\begin{aligned}
&\left|\mathbb{E}\left[\nabla^2 h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_2}Z_t^{x;\ell^\epsilon}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]\right|\\
&\le \frac{1}{t-\ell_0^\epsilon}\,\|\nabla h_\delta\|_\infty \mathbb{E}\left[\left|\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right|\left|\int_{\ell_0^\epsilon}^t \nabla_{v_2}Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right|\right] + \|\nabla h_\delta\|_\infty \mathbb{E}\left[\left|D_{U_{t,2}}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right|\right]\\
&\le \frac{C}{t-\ell_0^\epsilon}\,|v_1|\left[\mathbb{E}\left|\int_{\ell_0^\epsilon}^t \nabla_{v_2}Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right|^2\right]^{\frac{1}{2}} + C|v_1||v_2|\\
&\le \frac{C}{t-\ell_0^\epsilon}\,|v_1||v_2|\sqrt{t-\ell_0^\epsilon} + C|v_1||v_2|\\
&= C|v_1||v_2|\left(1+(t-\ell_0^\epsilon)^{-1/2}\right).
\end{aligned}$$

Since $Z_{\ell_t^\epsilon}^{x;\ell^\epsilon} = X_t^{x;\ell^\epsilon}$, one has for $t \in (0,1]$,

$$\left|\mathbb{E}\left[\nabla^2 h_\delta(X_t^{x;\ell^\epsilon})\nabla_{v_2}X_t^{x;\ell^\epsilon}\nabla_{v_1}X_t^{x;\ell^\epsilon}\right]\right| \le C|v_1||v_2|\left(1+(\ell_t^\epsilon-\ell_0^\epsilon)^{-1/2}\right).$$

Letting $\epsilon \downarrow 0$ and using [32, Lemma 2.2] and (2.10), we get

$$\left|\mathbb{E}\left[\nabla^2 h_\delta(X_t^{x;\ell})\nabla_{v_2}X_t^{x;\ell}\nabla_{v_1}X_t^{x;\ell}\right]\right| \le C|v_1||v_2|\left(1+\ell_t^{-1/2}\right).$$

Now we "unfreeze" the fixed path ℓ and use (2.8). Thus, we get for $\alpha \in [\alpha_0, 2)$ and $t \in (0,1]$,

$$\begin{aligned}
\left|\mathbb{E}[\nabla^2 h_\delta(X_t^x)\nabla_{v_2}X_t^x\nabla_{v_1}X_t^x]\right| &\le \int_{\mathbb{S}} \left|\mathbb{E}\left[\nabla^2 h_\delta(X_t^{x;\ell})\nabla_{v_2}X_t^{x;\ell}\nabla_{v_1}X_t^{x;\ell}\right]\right| \mu_{\mathbb{S}}(\mathrm{d}\ell)\\
&\le C|v_1||v_2| \int_{\mathbb{S}} \left(1+\ell_t^{-1/2}\right)\mu_{\mathbb{S}}(\mathrm{d}\ell)\\
&= C|v_1||v_2|\left(1+\mathbb{E}S_t^{-1/2}\right)\\
&\le C_{\alpha_0}|v_1||v_2|t^{-1/\alpha}.
\end{aligned} \tag{4.3}$$

The last estimate uses the following moment estimate from [8, Lemma 4.1]

$$\mathbb{E}S_t^{-1/2} = \sqrt{\frac{2}{\pi}}\,\Gamma\left(1+\frac{1}{\alpha}\right)2^{1/\alpha}t^{-1/\alpha}, \quad t>0. \tag{4.4}$$

It follows from (4.1), (4.2), and (4.3) that for $\alpha \in [\alpha_0, 2)$ and $t \in (0,1]$,

$$\begin{aligned}
|\nabla_{v_2}\nabla_{v_1}\mathbb{E}[h_\delta(X_t^x)]| &\le |\mathbb{E}[\nabla h_\delta(X_t^x)\nabla_{v_2}\nabla_{v_1}X_t^x]| + \left|\mathbb{E}\left[\nabla^2 h_\delta(X_t^x)\nabla_{v_2}X_t^x\nabla_{v_1}X_t^x\right]\right|\\
&\le C_{\alpha_0}|v_1||v_2|t^{-1/\alpha}.
\end{aligned}$$

We can now let $\delta \downarrow 0$ and we get with the help of the dominated convergence theorem

$$\lim_{\delta\downarrow 0}\nabla_{v_2}\nabla_{v_1}\mathbb{E}[h_\delta(X_t^x)] = \nabla_{v_2}\nabla_{v_1}\mathbb{E}[h(X_t^x)] = \nabla_{v_2}\nabla_{v_1}P_t h(x).$$

Therefore, we get for all $v_1, v_2, x \in \mathbb{R}^d$ and $t \in (0,1]$,

$$|\nabla_{v_2}\nabla_{v_1}P_t h(x)| \le C_{\alpha_0}|v_1||v_2|t^{-1/\alpha}$$

and this implies (3.2).

4.3. **Proof of** (3.3)**.** We use the dominated convergence theorem to see that for all $v_1, v_2, v_3 \in \mathbb{R}^d$,

$$\begin{aligned}
&\nabla_{v_3}\nabla_{v_2}\nabla_{v_1}\mathbb{E}\left[h_\delta(X_t^x)\right]\\
&\quad = \mathbb{E}\left[\nabla h_\delta(X_t^x)\nabla_{v_3}\nabla_{v_2}\nabla_{v_1}X_t^x\right] + \mathbb{E}\left[\nabla^2 h_\delta(X_t^x)\nabla_{v_3}X_t^x\nabla_{v_2}\nabla_{v_1}X_t^x\right]\\
&\qquad + \mathbb{E}\left[\nabla^2 h_\delta(X_t^x)\nabla_{v_2}X_t^x\nabla_{v_3}\nabla_{v_1}X_t^x\right] + \mathbb{E}\left[\nabla^2 h_\delta(X_t^x)\nabla_{v_1}X_t^x\nabla_{v_3}\nabla_{v_2}X_t^x\right]\\
&\qquad + \mathbb{E}\left[\nabla^3 h_\delta(X_t^x)\nabla_{v_3}X_t^x\nabla_{v_2}X_t^x\nabla_{v_1}X_t^x\right]\\
&\qquad =: \mathsf{I}_1 + \mathsf{I}_2 + \mathsf{I}_3 + \mathsf{I}_4 + \mathsf{I}_5.
\end{aligned}$$

We will estimate the terms I_k, $k = 1, \dots, 5$ separately. With (2.4) we get for $t \in (0,1]$,

$$|\mathsf{I}_1| \le \|\nabla h_\delta\|_\infty \mathbb{E}\left[|\nabla_{v_3}\nabla_{v_2}\nabla_{v_1}X_t^x|\right] \le C|v_1||v_2||v_3|.$$

Now we tern to I_2. Recall from Section 2.3, (2.12), that $Z_t^{x;\ell^\epsilon} := X_{\gamma_t^\epsilon}^{x;\ell^\epsilon}$. By (2.17) and (2.6),

$$\begin{aligned}
&\mathbb{E}\left[\nabla^2 h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_3}Z_t^{x;\ell^\epsilon}\nabla_{v_2}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]\\
&= \mathbb{E}\left[\nabla^2 h_\delta(Z_t^{x;\ell^\epsilon})D_{U_{t,3}}Z_t^{x;\ell^\epsilon}\nabla_{v_2}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]\\
&= \mathbb{E}\left[D_{U_{t,3}}\left(\nabla h_\delta(Z_t^{x;\ell^\epsilon})\right)\nabla_{v_2}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]\\
&= \mathbb{E}\left[D_{U_{t,3}}\left(\nabla h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_2}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right)\right] - \mathbb{E}\left[\nabla h_\delta(Z_t^{x;\ell^\epsilon})D_{U_{t,3}}\nabla_{v_2}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]
\end{aligned}$$

$$= \frac{1}{t-\ell_0^\epsilon}\,\mathbb{E}\left[\nabla h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_2}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\int_{\ell_0^\epsilon}^t \nabla_{v_3}Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right] - \mathbb{E}\left[\nabla h_\delta(Z_t^{x;\ell^\epsilon})D_{U_{t,3}}\nabla_{v_2}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right].$$

Combining this with (2.15), (2.16), (2.20) and Itô's isometry, we get for $t\in(\ell_0^\epsilon,\ell_1^\epsilon]$,

$$\begin{aligned}
&\left|\mathbb{E}\left[\nabla^2 h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_3}Z_t^{x;\ell^\epsilon}\nabla_{v_2}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]\right|\\
&\le \frac{1}{t-\ell_0^\epsilon}\,\|\nabla h_\delta\|_\infty|v_1||v_2|\left[\mathbb{E}\left|\int_{\ell_0^\epsilon}^t \nabla_{v_3}Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right|^2\right]^{\frac12} + \|\nabla h_\delta\|_\infty\mathbb{E}\left[\left|D_{U_{t,3}}\nabla_{v_2}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right|\right]\\
&\le \frac{C}{t-\ell_0^\epsilon}\,|v_1||v_2||v_3|\sqrt{t-\ell_0^\epsilon}+C|v_1||v_2||v_3|\\
&= C|v_1||v_2||v_3|\left(1+(t-\ell_0^\epsilon)^{-1/2}\right).
\end{aligned}$$

Since $Z_{\ell_t^\epsilon}^{x;\ell^\epsilon}=X_t^{x;\ell^\epsilon}$, one has for $t\in(0,1]$,

$$\left|\mathbb{E}\left[\nabla^2 h_\delta(X_t^{x;\ell^\epsilon})\nabla_{v_3}X_t^{x;\ell^\epsilon}\nabla_{v_2}\nabla_{v_1}X_t^{x;\ell^\epsilon}\right]\right| \le C|v_1||v_2||v_3|\left(1+(\ell_t^\epsilon-\ell_0^\epsilon)^{-1/2}\right).$$

Letting $\epsilon\downarrow 0$ and using [32, Lemma 2.2] and (2.10), we get

$$\left|\mathbb{E}\left[\nabla^2 h_\delta(X_t^{x;\ell})\nabla_{v_3}X_t^{x;\ell}\nabla_{v_2}\nabla_{v_1}X_t^{x;\ell}\right]\right| \le C|v_1||v_2||v_3|\left(1+\ell_t^{-1/2}\right).$$

Now we "unfreeze" the fixed path ℓ and use (2.8). The moment estimate (4.4) gives for $t\in(0,1]$,

$$\begin{aligned}
|\mathsf{I}_2| &\le \int_{\mathbb{S}}\left|\mathbb{E}\left[\nabla^2 h_\delta(X_t^{x;\ell})\nabla_{v_3}X_t^{x;\ell}\nabla_{v_2}\nabla_{v_1}X_t^{x;\ell}\right]\right|\mu_{\mathbb{S}}(\mathrm{d}\ell)\\
&\le C|v_1||v_2||v_3|\int_{\mathbb{S}}\left(1+\ell_t^{-1/2}\right)\mu_{\mathbb{S}}(\mathrm{d}\ell)\\
&= C|v_1||v_2||v_3|\left(1+\mathbb{E}S_t^{-1/2}\right)\\
&\le C_{\alpha_0}|v_1||v_2||v_3|t^{-1/\alpha}.
\end{aligned}$$

In a similar way we can show that

$$|\mathsf{I}_3|+|\mathsf{I}_4|\le C_{\alpha_0}|v_1||v_2||v_3|t^{-1/\alpha},\quad t\in(0,1].$$

We still have to estimate $|\mathsf{I}_5|$. By (2.17), (2.6) and (2.5),

$$\begin{aligned}
&\mathbb{E}\left[\nabla^3 h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_3}Z_t^{x;\ell^\epsilon}\nabla_{v_2}Z_t^{x;\ell^\epsilon}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]\\
&\quad= \mathbb{E}\left[\nabla^3 h_\delta(Z_t^{x;\ell^\epsilon})D_{U_{t,3}}Z_t^{x;\ell^\epsilon}\nabla_{v_2}Z_t^{x;\ell^\epsilon}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]\\
&\quad= \mathbb{E}\left[D_{U_{t,3}}\left(\nabla^2 h_\delta(Z_t^{x;\ell^\epsilon})\right)\nabla_{v_2}Z_t^{x;\ell^\epsilon}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]\\
&\quad= \mathbb{E}\left[D_{U_{t,3}}\left(\nabla^2 h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_2}Z_t^{x;\ell^\epsilon}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right)\right]\\
&\qquad- \mathbb{E}\left[\nabla^2 h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_1}Z_t^{x;\ell^\epsilon}D_{U_{t,3}}\nabla_{v_2}Z_t^{x;\ell^\epsilon}\right]\\
&\qquad- \mathbb{E}\left[\nabla^2 h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_2}Z_t^{x;\ell^\epsilon}D_{U_{t,3}}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{t-\ell_0^\epsilon}\,\mathbb{E}\left[\nabla^2 h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_2} Z_t^{x;\ell^\epsilon}\nabla_{v_1} Z_t^{x;\ell^\epsilon}\int_{\ell_0^\epsilon}^t \nabla_{v_3} Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right]\\
&\quad - \mathbb{E}\left[\nabla^2 h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_1} Z_t^{x;\ell^\epsilon} D_{U_{t,3}}\nabla_{v_2} Z_t^{x;\ell^\epsilon}\right]\\
&\quad - \mathbb{E}\left[\nabla^2 h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_2} Z_t^{x;\ell^\epsilon} D_{U_{t,3}}\nabla_{v_1} Z_t^{x;\ell^\epsilon}\right]\\
&=: \mathsf{I}_{51} + \mathsf{I}_{52} + \mathsf{I}_{53}.
\end{aligned}
$$

For I_{51}, a similar argument as above gives

$$
\begin{aligned}
\mathsf{I}_{51} &= \frac{1}{t-\ell_0^\epsilon}\,\mathbb{E}\left[D_{U_{t,2}}\left(\nabla h_\delta(Z_t^{x;\ell^\epsilon})\right)\nabla_{v_1} Z_t^{x;\ell^\epsilon}\int_{\ell_0^\epsilon}^t \nabla_{v_3} Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right]\\
&= \frac{1}{t-\ell_0^\epsilon}\,\mathbb{E}\left[D_{U_{t,2}}\left(\nabla h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_1} Z_t^{x;\ell^\epsilon}\int_{\ell_0^\epsilon}^t \nabla_{v_3} Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right)\right]\\
&\quad - \frac{1}{t-\ell_0^\epsilon}\,\mathbb{E}\left[\nabla h_\delta(Z_t^{x;\ell^\epsilon}) D_{U_{t,2}}\nabla_{v_1} Z_t^{x;\ell^\epsilon}\int_{\ell_0^\epsilon}^t \nabla_{v_3} Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right]\\
&\quad - \frac{1}{t-\ell_0^\epsilon}\,\mathbb{E}\left[\nabla h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_1} Z_t^{x;\ell^\epsilon}\int_{\ell_0^\epsilon}^t D_{U_{t,2}}\nabla_{v_3} Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right]\\
&\quad - \frac{1}{(t-\ell_0^\epsilon)^2}\,\mathbb{E}\left[\nabla h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_1} Z_t^{x;\ell^\epsilon}\int_{\ell_0^\epsilon}^t \nabla_{v_3} Z_s^{x;\ell^\epsilon}\nabla_{v_2} Z_s^{x;\ell^\epsilon}\,\mathrm{d}s\right]\\
&=: \mathsf{I}_{511} + \mathsf{I}_{512} + \mathsf{I}_{513} + \mathsf{I}_{514}.
\end{aligned}
$$

By (2.15) and Itô's isometry, we get for $t \in (\ell_0^\epsilon, \ell_1^\epsilon]$,

$$
\begin{aligned}
|\mathsf{I}_{511}| &= \left|\frac{1}{(t-\ell_0^\epsilon)^2}\,\mathbb{E}\left[\nabla h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_1} Z_t^{x;\ell^\epsilon}\int_{\ell_0^\epsilon}^t \nabla_{v_3} Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\int_{\ell_0^\epsilon}^t \nabla_{v_2} Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right]\right|\\
&\le \frac{C}{(t-\ell_0^\epsilon)^2}\,\|\nabla h_\delta\|_\infty |v_1| \left(\mathbb{E}\left|\int_{\ell_0^\epsilon}^t \nabla_{v_3} Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right|^2\right)^{1/2}\left(\mathbb{E}\left|\int_{\ell_0^\epsilon}^t \nabla_{v_2} Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right|^2\right)^{1/2}\\
&\le \frac{C}{t-\ell_0^\epsilon}\,|v_1||v_2||v_3|
\end{aligned}
$$

By (2.15), Itô's isometry and (2.19), we get for $t \in (\ell_0^\epsilon, \ell_1^\epsilon]$,

$$
\begin{aligned}
|\mathsf{I}_{512}| &\le \frac{1}{t-\ell_0^\epsilon}\,\|\nabla h_\delta\|_\infty |v_1||v_2| \left(\mathbb{E}\left|\int_{\ell_0^\epsilon}^t \nabla_{v_3} Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right|^2\right)^{1/2}\\
&\le \frac{C}{\sqrt{t-\ell_0^\epsilon}}\,|v_1||v_2||v_3|
\end{aligned}
$$

and, with a similar calculation,

$$
|\mathsf{I}_{513}| \le \frac{C}{\sqrt{t-\ell_0^\epsilon}}\,|v_1||v_2||v_3|.
$$

Using (2.15) we see for $t \in (\ell_0^\epsilon, \ell_1^\epsilon]$ that

$$\begin{aligned}|\mathsf{I}_{514}| &\le \frac{\|\nabla h_\delta\|_\infty}{t-\ell_0^\epsilon}|v_1||v_2||v_3| \\ &\le \frac{C}{t-\ell_0^\epsilon}|v_1||v_2||v_3|.\end{aligned}$$

If we combine the estimates for $\mathsf{I}_{511}, \dots, \mathsf{I}_{514}$, we obtain

$$|\mathsf{I}_{51}| \le C\left(\frac{1}{\sqrt{t-\ell_0^\epsilon}} + \frac{1}{t-\ell_0^\epsilon}\right)|v_1||v_2||v_3|, \quad t \in (\ell_0^\epsilon, \ell_1^\epsilon].$$

The structure of I_{52} is similar to I_2, and so we can adapt the arguments of the estimate of I_2 and use (2.15), (2.19) and (2.21), to get

$$\begin{aligned}|\mathsf{I}_{52}| &\le \frac{1}{t-\ell_0^\epsilon}\|\nabla h_\delta\|_\infty \mathbb{E}\left[\left|D_{U_{t,3}}\nabla_{v_2}Z_t^{x;\ell^\epsilon}\right|\left|\int_{\ell_0^\epsilon}^t \nabla_{v_1}Z_s^{x;\ell^\epsilon}\,\mathrm{d}W_s\right|\right] \\ &\quad + \|\nabla h_\delta\|_\infty \mathbb{E}\left[\left|D_{U_{t,1}}D_{U_{t,3}}\nabla_{v_2}Z_t^{x;\ell^\epsilon}\right|\right] \\ &\le C\left(\frac{1}{\sqrt{t-\ell_0^\epsilon}}+1\right)|v_1||v_2||v_3|, \quad t \in (\ell_0^\epsilon, \ell_1^\epsilon].\end{aligned}$$

The same argument gives

$$|\mathsf{I}_{53}| \le C\left(\frac{1}{\sqrt{t-\ell_0^\epsilon}}+1\right)|v_1||v_2||v_3|, \quad t \in (\ell_0^\epsilon, \ell_1^\epsilon].$$

If we combine the estimates for $\mathsf{I}_{51}, \mathsf{I}_{52}, \mathsf{I}_{53}$, we obtain for all $t \in (\ell_0^\epsilon, \ell_1^\epsilon]$

$$\left|\mathbb{E}\left[\nabla^3 h_\delta(Z_t^{x;\ell^\epsilon})\nabla_{v_3}Z_t^{x;\ell^\epsilon}\nabla_{v_2}Z_t^{x;\ell^\epsilon}\nabla_{v_1}Z_t^{x;\ell^\epsilon}\right]\right| \le C\left(\frac{1}{t-\ell_0^\epsilon}+1\right)|v_1||v_2||v_3|.$$

Since $Z_{\ell_t^\epsilon}^{x;\ell^\epsilon} = X_t^{x;\ell^\epsilon}$, we get for all $t \in (0,1]$

$$\left|\mathbb{E}\left[\nabla^3 h_\delta(X_t^{x;\ell^\epsilon})\nabla_{v_3}X_t^{x;\ell^\epsilon}\nabla_{v_2}X_t^{x;\ell^\epsilon}\nabla_{v_1}X_t^{x;\ell^\epsilon}\right]\right| \le C|v_1||v_2||v_3|\left(1+(\ell_t^\epsilon-\ell_0^\epsilon)^{-1}\right).$$

Letting $\epsilon \downarrow 0$ and using [32, Lemma 2.2] and (2.10), we arrive at

$$\left|\mathbb{E}\left[\nabla^3 h_\delta(X_t^{x;\ell})\nabla_{v_3}X_t^{x;\ell}\nabla_{v_2}X_t^{x;\ell}\nabla_{v_1}X_t^{x;\ell}\right]\right| \le C|v_1||v_2||v_3|\left(1+\ell_t^{-1}\right).$$

From [8, Lemma 4.1] we know the following moment identity

$$\mathbb{E}S_t^{-1} = \frac{1}{2}\Gamma\left(1+\frac{2}{\alpha}\right)2^{2/\alpha}t^{-2/\alpha}, \quad t > 0.$$

Thus, we obtain for $t \in (0,1]$ and $\alpha \in [\alpha_0, 2)$,

$$\begin{aligned}|\mathsf{I}_5| &\le \int_{\mathbb{S}}\left|\mathbb{E}\left[\nabla^3 h_\delta(X_t^{x;\ell})\nabla_{v_3}X_t^{x;\ell}\nabla_{v_2}X_t^{x;\ell}\nabla_{v_1}X_t^{x;\ell}\right]\right|\mu_{\mathbb{S}}(\mathrm{d}\ell) \\ &\le C|v_1||v_2||v_3|\int_{\mathbb{S}}\left(1+\ell_t^{-1}\right)\mu_{\mathbb{S}}(\mathrm{d}\ell) \\ &= C|v_1||v_2||v_3|\left(1+\mathbb{E}S_t^{-1}\right) \\ &\le C_{\alpha_0}|v_1||v_2||v_3|t^{-2/\alpha}.\end{aligned}$$

Finally we can combine the bounds for $|\mathsf{I}_i|$, $i = 1, 2, 3, 4, 5$, and we get for $t \in (0, 1]$,

$$\begin{aligned}|\nabla_{v_3}\nabla_{v_2}\nabla_{v_1}\mathbb{E}\,[h_\delta(X_t^x)]| &\le |\mathsf{I}_1| + \cdots + |\mathsf{I}_5| \\ &\le C_{\alpha_0}|v_1||v_2||v_3|\left(1 + t^{-1/\alpha} + t^{-2/\alpha}\right) \\ &\le C_{\alpha_0}|v_1||v_2||v_3|t^{-2/\alpha}.\end{aligned}$$

We can now let $\delta \downarrow 0$ and use dominated convergence to see for all $v_1, v_2, v_3, x \in \mathbb{R}^d$ and $t \in (0, 1]$,

$$|\nabla_{v_3}\nabla_{v_2}\nabla_{v_1}P_t h(x)| = \lim_{\delta\downarrow 0}|\nabla_{v_3}\nabla_{v_2}\nabla_{v_1}\mathbb{E}\,[h_\delta(X_t^x)]| \le C_{\alpha_0}|v_1||v_2||v_3|t^{-2/\alpha},$$

which yields (3.3). This completes the proof.

5. A lower bound for the Ornstein–Uhlenbeck case

In this section we establish a lower bound for the Ornstein–Uhlenbeck case, i.e. for $b(x) = -x$. In this case, **(H1)** holds with $\theta_0 = 1$ and $K = 0$, and **(H2)** holds with $\theta_1 = 1$ and $\theta_2 = \theta_3 = 0$. Recall that μ_α and μ are the ergodic measures of the solutions to the SDEs (1.1) and (1.2).

Proposition 5.1. *Let $\alpha_0 \in (1, 2)$ be an arbitrary fixed number. For any $\alpha \in (\alpha_0, 2)$,*

$$W_1(\mu_\alpha, \mu) \ge C_{\alpha_0,d}(2 - \alpha).$$

Proof. Since $X_t = \mathrm{e}^{-t}x + \mathrm{e}^{-t}\int_0^t \mathrm{e}^s\,\mathrm{d}L_s$, we get

$$\begin{aligned}\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\,\xi X_t}\right] &= \mathrm{e}^{\mathrm{i}\,\xi \mathrm{e}^{-t}x}\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\int_0^t \xi\mathrm{e}^{-t}\mathrm{e}^s\,\mathrm{d}L_s}\right] = \mathrm{e}^{\mathrm{i}\,\xi \mathrm{e}^{-t}x}\mathbb{E}\left[\mathrm{e}^{2^{-1}\int_0^t|\xi\mathrm{e}^{-t}\mathrm{e}^s|^\alpha\,\mathrm{d}s}\right] \\ &= \mathrm{e}^{\mathrm{i}\,\xi \mathrm{e}^{-t}x}\mathrm{e}^{-(2\alpha)^{-1}|\xi|^\alpha(1-\mathrm{e}^{-\alpha t})} \xrightarrow{t\to\infty} \mathrm{e}^{-(2\alpha)^{-1}|\xi|^\alpha} = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}\,\xi\alpha^{-1/\alpha}L_1}\right].\end{aligned}$$

Thus, the ergodic measure μ_α is given by the law of $\alpha^{-1/\alpha}L_1$. Similarly, the ergodic measure μ is given by the law of $2^{-1/2}B_1$. We know from [8, Lemma 4.2] that

$$\mathbb{E}\,|L_1| = \frac{2\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{d}{2}\right)}2^{-1/\alpha}\Gamma\left(1 - \tfrac{1}{\alpha}\right) \quad \text{and} \quad \mathbb{E}\,|B_1| = \frac{2\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{d}{2}\right)}2^{-1/2}\Gamma\left(\tfrac{1}{2}\right),$$

Note that

$$W_1(\mu_\alpha, \mu) = \inf_{\Pi\in\mathscr{C}(\mu_\alpha,\mu)}\iint |x - y|\,\Pi(\mathrm{d}x, \mathrm{d}y),$$

where $\mathscr{C}(\mu_\alpha, \mu)$ denotes the set of all couplings of μ_α and μ. Then we obtain

$$\begin{aligned}W_1(\mu_\alpha, \mu) &\ge \inf_{\Pi\in\mathscr{C}(\mu_\alpha,\mu)}\left|\iint |x|\,\Pi(\mathrm{d}x, \mathrm{d}y) - \iint |y|\,\Pi(\mathrm{d}x, \mathrm{d}y)\right| \\ &= \left|\int |x|\,\mu_\alpha(\mathrm{d}x) - \int |y|\,\mu(\mathrm{d}y)\right| \\ &= \left|\mathbb{E}\left|\alpha^{-1/\alpha}L_1\right| - \mathbb{E}\left|2^{-1/2}B_1\right|\right| \\ &= \frac{2\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{d}{2}\right)}\left|(2\alpha)^{-1/\alpha}\Gamma\left(1 - \tfrac{1}{\alpha}\right) - 2^{-1}\Gamma\left(\tfrac{1}{2}\right)\right| \\ &\ge C_{\alpha_0,d}(2 - \alpha),\end{aligned}$$

where the last inequality follows from Lemma 5.2 below. □

Lemma 5.2. *Let $\alpha_0 \in (1,2)$ be an arbitrary number. There are constants $c_1, c_2 \in (0,\infty)$ such that for any $\alpha \in [\alpha_0, 2)$,*

$$c_1(2-\alpha) \le \left|(2\alpha)^{-1/\alpha}\Gamma\left(1-\tfrac{1}{\alpha}\right) - 2^{-1}\Gamma\left(\tfrac{1}{2}\right)\right| \le c_2(2-\alpha).$$

Proof. Let

$$\phi(x) := (2x)^{-1/x}\Gamma\left(1-\frac{1}{x}\right), \quad \alpha_0 \le x \le 2.$$

It is not hard to verify that there exist constants $c_i = c_i(\alpha_0) > 0$ (independent of α), $i = 1,2$, such that $\phi'|_{[\alpha_0,2]} \in [-c_2, -c_1]$. Since

$$\begin{aligned} |\phi(\alpha) - \phi(2)| &= (2-\alpha)\left|\int_0^1 \phi'\left(\alpha + r(2-\alpha)\right)\mathrm{d}r\right| \\ &= (2-\alpha)\int_0^1 \left[-\phi'\left(\alpha + r(2-\alpha)\right)\right]\mathrm{d}r, \end{aligned}$$

we get

$$c_1(2-\alpha) \le |\phi(\alpha) - \phi(2)| \le c_2(2-\alpha).$$

This is what we have claimed. □

References

[1] M. Abramowitz and I.A. Stegun (1972): *Handbook of Mathematical Formulas.* Dover, New York (reprint of the 10th edition).

[2] J. Bao, F.-Y. Wang and C. Yuan (2013): Bismut formulae and applications for functional SPDEs. *Bulletin des Sciences Mathématiques* **137**, 509–522.

[3] P. Chen, C.-S. Deng, R.L. Schilling and L. Xu (2022): Approximation of the invariant measure of stable SDEs by an Euler–Maruyama scheme. Preprint *arXiv*[math.PR] `https://arxiv.org/abs/2205.01342`.

[4] P. Chen, Q.-M. Shao and L. Xu (2022+): A probability measure approximation framework: Markov processes approach. To appear in: *Annals of Applied Probability.*

[5] Z.-Q. Chen, Z. Hao and X. Zhang (2020): Hölder regularity and gradient estimates for SDEs driven by cylindrical α-stable processes. *Electronic Journal of Probability* **25**, no. 137.

[6] Z.-Q. Chen, X. Zhang and G. Zhao (2021): Supercritical SDEs driven by multiplicative stable-like Lévy processes. *Transactions of the American Mathematical Society* **374**, 7621–7655.

[7] S. Cho, P. Kim, R. Song and Z. Vondracek (2022): Heat kernel estimates for subordinate Markov processes and their applications. *Journal of Differential Equations* **316**, 28–93.

[8] C.-S. Deng and R.L. Schilling (2019): Exact Asymptotic formulas for the heat kernels of space and time-fractional equations. *Fractional Calculus & Applied Analysis* **22**, 968–989.

[9] C.-S. Deng, R.L. Schilling and Y.-H. Song (2017): Subgeometric rates of convergence for Markov processes under subordination. *Advances in Applied Probability* **49**, 162–181.

[10] Z. Dong, L. Xu and X. Zhang (2014): Exponential ergodicity of stochastic Burgers equations driven by α-stable processes. *Journal of Statistical Physics* **154**, 929–949.

[11] R. Durrett (2019): *Probability: Theory and Examples* (5th edn). Cambridge University Press, Cambridge.

[12] K. Grobys (2021): What do we know about the second moment of financial markets? *International Review of Financial Analysis* **78**, 101891.

[13] M. Hairer (2021): The convergence of Markov processes. `https://www.hairer.org/notes/Convergence.pdf` (accessed January 13, 2023).

[14] A. Janicki, Z. Michna and A. Weron (1996): Approximation of stochastic differential equations driven by α-stable Lévy motion. *Applicationes Mathematicae* **24**, 149–168.

[15] F. Kühn and R.L. Schilling (2019): Strong convergence of the Euler-Maruyama approximation for a class of Lévy-driven SDEs. *Stochastic Processes and Their Applications* **129**, 2654–2680.

[16] M. Liang and J. Wang (2020): Gradient estimates and ergodicity for SDEs driven by multiplicative Lévy noises via coupling. *Stochastic Processes and Their Applications* **130**, 3053–3094.

[17] W. Liu, R. Song and L. Xie (2020): Gradient estimates for the fundamental solution of Lévy type operator. *Advances in Nonlinear Analysis* **9**, 1453–1462.

[18] X. Liu (2022): On the α-dependence of stochastic differential equations with Hölder drifts and driven by α-stable Lévy processes. *Journal of Mathematical Analysis and Applications* **506**, 125642.

[19] X. Liu (2022): Limits of invariant measures of stochastic Burgers equations driven by two kinds of α-stable processes. *Stochastic Processes and Their Applications* **146**, 1–21.

[20] X. Liu (2022): The α-dependence of the invariant measure of stochastic real Ginzburg-Landau equation driven by α-stable Lévy processes. *Journal of Differential Equations* **314**, 418–445.

[21] S.P. Meyn and R.L. Tweedie (1993): Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes. *Advances in Applied Probability* **25**, 518–548.

[22] S.P. Meyn and R.L. Tweedie (2009): *Markov Chains and Stochastic Stability* (2nd edn). Cambridge University Press, Cambridge.

[23] M. Peligrad, H. Sang, Y. Xiao and G. Yang (2022): Limit theorems for linear random fields with innovations in the domain of attraction of a stable law. *Stochastic Processes and Their Applications* **150**, 596-621.

[24] E. Priola, A. Shirikyan, L. Xu and J. Zabczyk (2012): Exponential ergodicity and regularity for equations with Lévy noise. *Stochastic Processes and their Applications* **122**, 106–133.

[25] E. Priola and J. Zabczyk (2011): Structural properties of semilinear SPDEs driven by cylindrical stable processes. *Probability Theory and Related Fields* **149**, 97–137.

[26] J. Wang (2016): L^p-Wasserstein distance for stochastic differential equations driven by Lévy processes. *Bernoulli*, **22**, 1598–1616.

[27] R. Wang, J. Xiong and L. Xu (2017): Irreducibility of stochastic real Ginzburg-Landau equation driven by α-stable noises and applications. *Bernoulli* **23**, 1179–1201.

[28] K. Sato (1999): *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Studies in Advanced Mathematics **68**, Cambridge.

[29] R.L. Schilling, R. Song and Z. Vondracek (2012): *Bernstein Functions: Theory and Applications*. De Gruyter, Studies in Mathematics **37**, Berlin.

[30] R. Song and Z. Vondracek (2003): Potential theory of subordinate killed Brownian motion in a domain. *Probability Theory and Related Fields* **125**, 578–592.

[31] L. Xu (2013): Ergodicity of the stochastic real Ginzburg-Landau equation driven by α-stable noises. *Stochastic Processes and their Applications* **123**, 3710–3736.

[32] X. Zhang (2013): Derivative formulas and gradient estimates for SDEs driven by α-stable processes. *Stochastic Processes and their Applications* **123**, 1213–1228.

(C.-S. Deng) School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China

Email address: dengcs@whu.edu.cn

(R.L. Schilling) TU Dresden, Fakultät Mathematik, Institut für Mathematische Stochastik, 01062 Dresden, Germany

Email address: rene.schilling@tu-dresden.de

(L. Xu) Department of Mathematics, Faculty of Science and Technology, University of Macau, Macau S.A.R., China

Email address: lihuxu@um.edu.mo