MULTIVARIATE STABLE APPROXIMATION IN WASSERSTEIN DISTANCE BY STEIN'S METHOD

PENG CHEN, IVAN NOURDIN, LIHU XU, AND XIAOCHUAN YANG

Abstract. We investigate regularity properties of the solution to Stein's equation associated with multivariate integrable α-stable distribution for a general class of spectral measures and Lipschitz test functions. The obtained estimates induce an upper bound in Wasserstein distance for the multivariate α-stable approximation.

Key words: multivariate α-stable approximation; Stein's method; generalized central limit theorem; rate of convergence; fractional Laplacian.

1. Introduction

This paper is concerned with the multivariate stable approximation by Stein's method. A probability measure π on $\mathbb{R}^d$ with $d \geq 2$ is *strictly stable* if, for any $a > 0$, there is $b > 0$ such that

$$\widehat{\pi}(z)^a = \widehat{\pi}(bz),$$

where $\widehat{\pi}$ is the Fourier transform of π. The distribution π is completely determined by the *stability parameter* $\alpha \in (0, 2)$ and the finite non-zero *spectral measure* ν on the surface $\mathbb{S}^{d-1}$. In this paper, we consider the super-critical regime $\alpha > 1$. The critical $\alpha = 1$ and sub-critical regimes $\alpha < 1$ require different treatment because of the lack of moments, and will be dealt with elsewhere. When $\alpha > 1$, one has the representation [18, Theorem 14.10]:

$$\widehat{\pi}(z) = \exp\left[-\int_{\mathbb{S}^{d-1}} |\langle z, \theta\rangle|^\alpha (1 - \mathrm{i}\tan\frac{\pi\alpha}{2}\mathrm{sgn}\langle z, \theta\rangle)\nu(d\theta)\right]. \tag{1.1}$$

The integral therein is customarily called the *characteristic exponent*, denoted by ψ. We point out that, as a prominent infinitely divisible distribution, π has the Lévy-Khintchine representation [18, Theorem 8.1, Theorem 14.3 and Theorem 14.10]

$$\widehat{\pi}(z) = \exp\left[-\int_0^\infty \int_{\mathbb{S}^{d-1}} (1 - e^{\mathrm{i}r\langle\theta, z\rangle} + \mathrm{i}r\langle\theta, z\rangle)\frac{dr}{r^{1+\alpha}}\nu(d\theta)\right],$$

where $r^{-1-\alpha}dr\nu(d\theta)$ is the Lévy measure on $[0, \infty) \times \mathbb{S}^{d-1} = \mathbb{R}^d$.

The spectral measure ν plays a crucial role in the study of multivariate stable laws. For instance, if ν is the uniform probability measure on $\mathbb{S}^{d-1}$, then $\psi(\lambda) = \sigma|\lambda|^\alpha$ with $\sigma > 0$ so that π is rotationally invariant (sometimes referred to as isotropic). Hereafter $|a|$ denotes the Euclidean norm of a for any $a \in \mathbb{R}^d$. Another example is when ν is supported on $\{\pm e_1, ..., \pm e_d\}$ where $\{e_i, 1 \leq i \leq d\}$ is the canonical basis of $\mathbb{R}^d$, then $\psi(\lambda) = \sum_{i=1}^d \sigma_i|\lambda_i|^\alpha$ so that the marginal distributions of π are independent one-dimensional stable laws. Yet another interesting example is when ν is supported on a fractal subset of $\mathbb{S}^{d-1}$, so that μ is extremely anisotropic. The distributional properties of π change dramatically from one type of ν to another. In this paper, we are going to consider not only each of the aforementioned types of ν, but also mixtures of these types.

Stein's method is a vast range of ideas and tools that allow one to study the proximity between a probability measure and a target distribution. The scope of the method has been considerably extended since Stein [17] proposed his elegant approach for normal approximation. In particular, Barbour [4] devised the generator approach which is applicable to target distributions that can be realized as the stationary distribution of a "nice" Markov process. Barbour's approach is the one adopted in this paper and it takes the following steps. First, one constructs a Markov process $(X_t)_{t\geq 0}$ with infinitesimal generator $\mathcal{A}$ that converges in distribution to π as $t \to \infty$ for any initial condition $X_0 = x \in \mathbb{R}^d$. Second, one considers Stein's equation (or Poisson equation in the PDE literature)

$$\mathcal{A}f(x) = h(x) - \pi(h) \tag{1.2}$$

with $h \in L^1(\pi)$ and $\pi(h) := \int h(x)\pi(dx)$. By exploiting properties of the transition semigroup $(Q_t)_{t\geq 0}$ determined by $\mathcal{A}$, in particular $Q_0 h = h, Q_\infty h = \int h(x)\pi(dx)$ and the relation $\frac{d}{dt}Q_t = \mathcal{A}Q_t$, one argues that

$$f_h(x) := -\int_0^\infty Q_t\big(h(x) - \pi(h)\big)dt \tag{1.3}$$

is in the domain of $\mathcal{A}$ and solves (1.2). Third, one uses the integral form (1.3) and properties of $(Q_t)_{t\geq 0}$ to derive regularity estimates for the solution (1.3). To see why these steps lead to an upper bound for the distance between an arbitrary distribution and π, let Z denote a strictly stable random vector with distribution π, for any $\mathbb{R}^d$-valued random vector F, one has

$$\mathbb{E}[h(F)] - \mathbb{E}[h(Z)] = \mathbb{E}[\mathcal{A}f_h(F)].$$

Ranging h in a class of functions that guarantees convergence in distribution, and using the regularity estimates of (1.3) obtained in the third step, together with the explicit form of $\mathcal{A}$, one would obtain an upper bound for a certain distance between F and Z.

Though conceivable, carrying out rigorously each of the aforementioned steps and claims in the context of stable approximation is certainly a non-trivial task. In dimension one, Xu [21] considered the case of symmetric α-stable law with $\alpha > 1$. The approach of [21] was then generalized in [8] to asymmetric α-stable law with $\alpha > 1$, and in [2] to a class of infinitely divisible distributions with finite first moment. Later, Chen *et al.* [9] considered non-integrable α-stable approximation (necessarily $\alpha \leq 1$). In higher dimension, Arras and Houdré [3] carried out the aforementioned second step (construction of the solution to Stein's equation) for a class of self-decomposable distributions which includes multivariate stable laws. However, regularity estimates of the solution are studied only for smooth test function h (at least twice differentiable with bounded partial derivatives) in [3], therefore, their results cannot be used to derive bounds for multivariate stable approximation in Wasserstein distance that we address in this paper.

The main contribution of this paper is a thorough study of the regularity estimates for the solution to Stein's equation in the context of multivariate stable approximation and Lipschitz test functions, which in turn allows to obtain Wasserstein bounds. As explained already, when the spectral measure ν is general, the solution to Stein's equation is not easy to handle. Our method covers a rich class of spectral measures including the absolutely continuous type, the discrete type, the fractal type, and the mixture of them. Since real life high dimensional data often present anisotropic feature, the rich class of spectral measures that we consider would widen the applicability of our results. In terms of application, we provide the rate of convergence for the classical multivariate stable limit theorem.

The rest of this paper is organized as follows. After introducing the Markov process converging to π, we construct a solution to Stein's equation (Proposition 2.1), present the regularity estimates for the solution (Theorem 2.7) and obtain Stein's bound for multivariate stable approximation (Theorem 2.8). Theorem 2.7 is proved in Section 3 and Theorem 2.8 is proved in Section 4. Examples are given in Section 5.

2. Preliminaries and statement of the main results

2.1. **Ornstein-Uhlenbeck type processes.** The Markov process we construct in the first step of Barbour's program is the so-called Ornstein-Uhlenbeck type process which is a simple stochastic differential equation (SDE) driven by stable Lévy processes. We refer the reader to Applebaum [1] for background on stochastic calculus of Lévy processes, and Sato [18] for general facts about Lévy processes.

Let $(Z_t)_{t\geq 0}$ be a stable Lévy process, a process with independent and stationary increments having marginal distribution $Z_1 \sim \pi$, given by (1.1). Consider the SDE

$$\begin{cases} X_t = X_0 - \frac{1}{\alpha}\int_0^t X_s ds + Z_t \\ X_0 = x \end{cases}, \tag{2.1}$$

Such an equation can be solved explicitly

$$X_t^x = xe^{-\frac{t}{\alpha}} + \int_0^t e^{-\frac{t-s}{\alpha}} dZ_s, \tag{2.2}$$

see [18, p.105], and provides an interpolation between any Dirac mass and π. This follows from the fact that $(X_t^x)_{t\geq 0}$ is a scaled and time-changed Lévy process, i.e.

$$X_t^x \overset{d}{=} xe^{-\frac{t}{\alpha}} + e^{-\frac{t}{\alpha}} Z_{e^t-1} \overset{d}{=} xe^{-\frac{t}{\alpha}} + Z_{1-e^{-t}}, \tag{2.3}$$

see [9, Section 2.3]. For the second equality in distribution we have used the self-similarity of the process $(Z_t)_{t\geq 0}$, namely $Z_{ct} = c^{1/\alpha} Z_t$ in distribution for any $c, t > 0$. One sees that as $t \to \infty$, X_t^x converges in distribution to $Z_1 \sim \pi$. For another proof of the latter fact, one may check the condition of a general result [18, Th. 17.5] for self-decomposable distributions.

An application of Itô's formula for semimartingales with jumps to $(X_t^x)_{t\geq 0}$ shows that (see [1, Chapter 6] for details) its generator is

$$\mathcal{A}^{\alpha,\nu} f(x) := \mathcal{L}^{\alpha,\nu} f(x) - \frac{1}{\alpha}\langle x, \nabla f(x)\rangle, \tag{2.4}$$

where

$$\mathcal{L}^{\alpha,\nu} f(x) = d_\alpha \int_0^\infty \int_{\mathbb{S}^{d-1}} (f(x+r\theta) - f(x) - r\langle \theta, \nabla f(x)\rangle)\nu(d\theta)\frac{dr}{r^{1+\alpha}}.$$

Here $d_\alpha = \left(\int_0^\infty \frac{1-\cos y}{y^{\alpha+1}}\right)^{-1} = \frac{\alpha}{\Gamma(1-\alpha)\cos\frac{\pi\alpha}{2}}$ and ν is normalized on $\mathbb{S}^{d-1}$ so that $\nu(\mathbb{S}^{d-1}) = 1$.

Now one can write out Stein's equation associated with the multivariate stable distribution π as follows

$$\mathcal{A}^{\alpha,\nu} f(x) = h(x) - \pi(h), \tag{2.5}$$

where $h \in L^1(\pi)$. In view of obtaining bounds in Wasserstein distance, consider h belonging to the space Lip_1 of Lipschitz continuous functions with Lipschitz constant at most one. It is standard that $\mathrm{Lip}_1 \subset L^1(\pi)$. We write for simplicity $\mathcal{L}^\alpha = \mathcal{L}^{\alpha,\nu}$ and $\mathcal{A}^\alpha = \mathcal{A}^{\alpha,\nu}$ in the rest of the paper.

2.2. **Solving Stein's equation.** We construct a solution to Stein's equation by using the process $(X_t^x)_{t\geq 0}$, as described in the introduction. Denote by $p(t,x) := p_t(x)$ the density of the driving process $(Z_t)_{t\geq 0}$ in (2.1). Write $p(x) := p_1(x)$. By (2.3), one sees that

$$q(t,x,y) = p_{1-e^{-t}}(y - e^{-t/\alpha}x) = s(t)^{-1/\alpha} p(s(t)^{-1/\alpha}(y - e^{-t/\alpha}x)), \tag{2.6}$$

where $y \mapsto q(t,x,y)$ is the density of X_t^x, $s(t) = 1 - e^{-t}$ and we used the self-similarity of $(Z_t)_{t\geq 0}$ in the second equality.

Proposition 2.1 (Solution to Stein's equation)**.** *Suppose* $h \in \mathrm{Lip}_1$. *Set*

$$\begin{aligned} f(x) &:= -\int_0^\infty \mathbb{E}\big[h\big(X_t^x\big) - \pi(h)\big]dt, \\ &= -\int_0^\infty \int p_{1-e^{-t}}(y - e^{-\frac{t}{\alpha}}x)(h(y) - \pi(h))dydt. \end{aligned} \tag{2.7}$$

Then f *solves Stein's equation* (2.5), *i.e.* f *is in the domain of* $\mathcal{A}^\alpha$ *and*

$$\mathcal{A}^\alpha f(x) = h(x) - \pi(h). \tag{2.8}$$

The proof of this Proposition somewhat standard in view of recent advances [21, 8, 9, 2, 3], we give a proof in the Appendix for completeness.

2.3. **Zoo of spectral measures.** Obtaining density estimates for general multivariate stable law is a genuinely hard problem and is very sensitive to the form of the spectral measure, as pointed out by the seminal work of Watanabe [20]. Ideally, π has a spectral measure that is "close" to a uniform distribution on the sphere, then one may expect that the density is comparable to that of the isotropic stable law, which is indeed the case. When the spectral measure ν becomes less isotropic, the density of π would change accordingly. We distinguish three classes of spectral measures as follows.

1. Absolutely continuous type. We further assume ν is absolutely continuous with respect to the spherical measure $d\theta$ on the unit sphere with density $g(\theta)$ satisfying

$$0 < k_1 \leqslant g(\theta) \leqslant k_2, \tag{2.9}$$

where k_1 and k_2 are positive constants. It follows that

$$\mathcal{L}^\alpha f(x) = d_\alpha \int_{\mathbb{R}^d} \frac{f(x+y) - f(x) - y\cdot\nabla f(x)}{|y|^{d+\alpha}} \cdot g\big(\frac{y}{|y|}\big)\mathrm{d}y. \tag{2.10}$$

In particular, when ν is a uniform distribution on $\mathbb{S}^{d-1}$, one has $g(\theta) = \frac{1}{V(\mathbb{S}^{d-1})}$, where $V(\mathbb{S}^{d-1})$ is the surface area of $\mathbb{S}^{d-1}$ and $V(\mathbb{S}^{d-1}) = \frac{2\pi^{d/2}}{\Gamma(d/2)}$.

When $\nu(d\theta) = g(\theta)d\theta$ satisfies (2.9), we have the following estimates.

Lemma 2.2. *Let $p(1,x)$ be the density of Z_1, we have*

$$|\nabla p(1,x)| \leqslant C_{\alpha,d} \frac{1}{(1+|x|)^{\alpha+d}}, \tag{2.11}$$

$$|\mathcal{L}^\alpha p(1,x)| \leqslant C_{\alpha,d} \frac{1}{(1+|x|)^{\alpha+d}}, \tag{2.12}$$

and

$$|\nabla^2 p(1,x)| \leqslant C_{\alpha,d} \frac{1}{(1+|x|)^{\alpha+d}}, \tag{2.13}$$

where ∇^2 is the Hessian matrix.

2. Purely atomic type. We consider the case that ν is supported on $\{\pm e_1, ..., \pm e_d\}$ where $(e_i)_{1 \leq i \leq d}$ is *any* orthonormal basis of $\mathbb{R}^d$. Changing the coordinate system if necessary, we may and will assume that $(e_i)_{1 \leq i \leq d}$ is the canonical basis. As such, the marginals of the stable vector distributed according to π are independent one-dimensional stable distributions. The following can be easily derived.

Lemma 2.3. *Let $p(1,x)$ be the density of Z_1, for any $x = (x_1, x_2, \cdots, x_d)$, we have*

$$|\nabla p(1,x)| \leqslant C_{\alpha,d} \prod_{i=1}^{d} \frac{1}{(1+|x_i|)^{1+\alpha}}, \tag{2.14}$$

$$|\mathcal{L}^\alpha p(1,x)| \leqslant C_{\alpha,d} \prod_{i=1}^{d} \frac{1}{(1+|x_i|)^{1+\alpha}}, \tag{2.15}$$

and

$$|\nabla^2 p(1,x)| \leqslant C_{\alpha,d} \prod_{i=1}^{d} \frac{1}{(1+|x_i|)^{1+\alpha}}. \tag{2.16}$$

3. Fractal type . This type of Lévy measure is considered in [20, 5, 6]. We first define the so-called γ-measure. For convenience, we denote $\vartheta(drd\theta) = d_\alpha r^{-1-\alpha} dr \nu(d\theta)$.

Definition 2.4. *We say that the measure ϑ defined on $\mathbb{R}^d$ is a γ-measure at $\mathbb{S}^{d-1}$ if $\gamma \geqslant 0$ and*

$$\vartheta(B(x,r)) \leqslant C r^\gamma, \quad |x| = 1, \quad 0 < r < 1/2.$$

Remark 2.5. *The absolutely continuous type mentioned earlier clearly satisfies the condition, and the purely atomic type does not. One prototype non-absolutely continuous γ-measure is the product of $r^{-1-\alpha}dr$ and the uniform probability mass distribution on a Cantor-type subset of $\mathbb{S}^{d-1}$ of Hausdorff dimension $\gamma - 1$, see* [15] *for aspects of fractal measures.*

One always has $1 \leqslant \gamma \leq d$. We further assume that ϑ is a $\gamma-$measure with $1 \leqslant \gamma \leqslant d$ for the case $d = 2$ and $d - \alpha < \gamma \leqslant d$ for the case $d \geqslant 3$. In addition, we assume that ϑ is symmetric, i.e., $\vartheta(A) = \vartheta(-A)$ for any $A \in \mathbb{R}^d$.

When ν satisfies the above condition, we can get the following estimates:.

Lemma 2.6. *Let $p(1,x)$ be the transition probability density of Z_1, we have*

$$|\nabla p(1,x)| \leqslant C_{\alpha,d}\frac{1}{(1+|x|)^{\alpha+\gamma}}, \tag{2.17}$$

$$|\mathcal{L}^{\alpha} p(1,x)| \leqslant C_{\alpha,d}\frac{1}{(1+|x|)^{\alpha+\gamma}}, \tag{2.18}$$

and

$$|\nabla^2 p(1,x)| \leqslant C_{\alpha,d}\frac{1}{(1+|x|)^{\alpha+\gamma}}. \tag{2.19}$$

2.4. **Main results.**

Theorem 2.7 (Regularity estimates for the solution)**.** *Let f be given by (2.7). Let $a,b,c,\sigma_i,\sigma_i' \geq 0$. Suppose that the spectral measure ν of π is given by*

$$\nu(d\theta) = ag(\theta)d\theta + b\sum_{i=1}^{d}(\sigma_i\delta_{e_i} + \sigma_i'\delta_{-e_i}) + c\nu_\gamma(d\theta), \tag{2.20}$$

which are respectively the absolutely continuous, the purely atomic, and the fractal part. Suppose that g satisfies (2.9), $\sigma_i + \sigma_i' = 1$ for each i, and that $r^{-1-\alpha}dr\nu_\gamma(d\theta)$ is a non-absolutely continuous γ-measure, which is symmetric. In addition, we assume that $a+b+c=1$, then we have the following estimates:

$$\|\nabla f\| \leqslant \alpha\|\nabla h\|, \tag{2.21}$$

$$\|\nabla^2 f\| \leqslant C_{\alpha,d}\|\nabla h\|, \tag{2.22}$$

where $\|\cdot\|$ is the L^∞ norm and ∇^2 is the Hessian matrix. Further, for all $x,y \in \mathbb{R}^d$

$$\Big|\mathcal{L}^\alpha f(x) - \mathcal{L}^\alpha f(y)\Big| \leqslant \frac{2d_\alpha\|\nabla^2 f\|}{\alpha(2-\alpha)(\alpha-1)}|x-y|^{2-\alpha}. \tag{2.23}$$

We move to obtaining Wasserstein bounds for CLT with stable limit. Recall that the the Wasserstein distance between a probability measure μ on $\mathbb{R}^d$ and π is defined by

$$d_W(\mu,\pi) \leq \sup_{h\in \mathrm{Lip}_1} |\mu(h) - \pi(h)|.$$

Let $n \in \mathbb{N}$ and let $\zeta_{n,1}, \zeta_{n,2}, \cdots, \zeta_{n,n}$ be a sequence of independent random vectors with $\mathbb{E}|\zeta_{n,i}| < \infty$ for $1 \leqslant i \leqslant n$. Set

$$S_n = \big(\zeta_{n,1} - \mathbb{E}\zeta_{n,1}\big) + \big(\zeta_{n,2} - \mathbb{E}\zeta_{n,2}\big) + \cdots + \big(\zeta_{n,n} - \mathbb{E}\zeta_{n,n}\big);$$

$$S_n(i) = S_n - \zeta_{n,i}, \qquad 1 \leqslant i \leqslant n.$$

Denote $l_n = \frac{\alpha}{d_\alpha}n$ and set $\eta_{n,i} = l_n^{1/\alpha}\zeta_{n,i}$,

Theorem 2.8 (Wassertein bounds)**.** *Let $n \in \mathbb{N}$ and $\zeta_{n,i}$, $\eta_{n,i}$, $i = 1, \cdots, n$ are defined as above. Let μ be an $\alpha-$stable distribution with characteristic function* $\exp\big(-\psi(\lambda)\big)$ *for $\alpha \in (1,2)$. Then, for any $N > 0$, we have*

$$d_W\big(\mathcal{L}(S_n),\mu\big) \leqslant C_{\alpha,d}\Big\{n^{-\frac{2}{\alpha}}\sum_{i=1}^{n}\mathbb{E}|\eta_{n,i}|^{2-\alpha} + n^{-\frac{2}{\alpha}}\sum_{i=1}^{n}\big(\mathbb{E}|\eta_{n,i}|\big)^2 + \mathbb{E}\sum_{i=1}^{n}|\mathcal{R}_{n,i}|\Big\},$$

where

$$\mathcal{R}_{n,i} = n^{-\frac{2}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_0^N r^2 \Big| F_{\eta_{n,i}}(drd\theta) - \frac{\alpha}{r^{\alpha+1}} \mathrm{d}r\nu(d\theta) \Big| + n^{-\frac{1}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_N^\infty r \Big| F_{\eta_{n,i}}(drd\theta) - \frac{\alpha}{r^{\alpha+1}} \mathrm{d}r\nu(d\theta) \Big|.$$

3. Proof of Theorem 2.7

First, we give proofs for the lemmas in Section 2.3.

Proof of Lemma 2.2. By the same argument as the proof of [11, (2.25) and (2.28)], we can obtain (2.11) and (2.12), respectively. In addition, (2.13) can be obtained by the same argument as the proof of [11, (2.25)]. □

Proof of Lemma 2.3. By independence, if we denote the density of the *ith* component of Z_1 by $p_i(1, x_i)$, we have

$$p(1, x) = \prod_{i=1}^d p_i(1, x_i).$$

Since the $p_i(1, x_i)$ can be consider as the density of 1-dimensional $\alpha-$stable process Z_1, we have by [12, (1.10) and (1.12)],

$$p_i'(1, x_i)| \leqslant \frac{C_\alpha}{(1 + |x_i|)^{1+\alpha}}, \quad |p_i''(1, x_i)| \leqslant \frac{C_\alpha}{(1 + |x_i|)^{1+\alpha}},$$

the case of subordinator can be obtained by the same argument as the proof of [10, lemma 3.1]. These imply

$$\begin{aligned} |\nabla p(1, x)| &= \Big| \big(\frac{\partial p(1, x)}{\partial x_1}, \frac{\partial p(1, x)}{\partial x_2}, \cdots, \frac{\partial p(1, x)}{\partial x_d} \big) \Big| \\ &= \Big| \big(p_1'(1, x_1) \prod_{i=2}^d p_i(1, x_i), \cdots, \prod_{i=1}^{d-1} p_i(1, x_i) p_d'(1, x_d) \big) \Big| \\ &\leqslant C_\alpha \sqrt{d} \prod_{i=1}^d \frac{1}{(1 + |x_i|)^{1+\alpha}}, \end{aligned}$$

(2.14) is proved and (2.15) can be proved by the same argument.

Next, we will prove (2.15). The proof of (2.15) is the same as the proof of [11, (2.28)] and we only need to prove (2.13) in [11], i.e.,

$$\begin{aligned} &\delta_p(1, x; z) := p(1, x + z) + p(1, x - z) - 2p(1, x) \\ &\leqslant C_{\alpha,d}(|z|^2 \wedge 1) \{ \prod_{i=1}^d \frac{1}{(1 + |x_i + z_i|)^{1+\alpha}} + \prod_{i=1}^d \frac{1}{(1 + |x_i - z_i|)^{1+\alpha}} + \prod_{i=1}^d \frac{1}{(1 + |x_i|)^{1+\alpha}} \}. \end{aligned} \tag{3.1}$$

In fact, if $|z| > 1$, then

$$|\delta_p(1, x; z)| \leqslant p(1, x + z) + p(1, x - z) + 2p(1, x).$$

If $|z| \leqslant 1$,

$$\begin{aligned}\delta_p(1,x;z) &= \int_0^1 z\cdot\nabla\big(p(1,x+uz)-p(1,x-uz)\big)\mathrm{d}u \\ &= \int_0^1\int_0^1 \langle z\cdot z^T, \nabla^2 p(1,x+(1-2u')uz)\mathrm{d}u'\mathrm{d}u\end{aligned}$$

then by (2.15), we have

$$\begin{aligned}|\delta_p(1,x;z)| &\leqslant |z|^2\int_0^1\int_0^1 |\nabla^2 p(1,x+(1-2u')uz)|\mathrm{d}u'\mathrm{d}u \\ &\leqslant C_{\alpha,d}|z|^2\int_0^1\int_0^1\prod_{i=1}^d\frac{1}{(1+|x_i+(1-2u')uz_i|)^{\alpha+1}}\mathrm{d}u'\mathrm{d}u \\ &\leqslant C_{\alpha,d}|z|^2\prod_{i=1}^d\frac{1}{(1+|x_i|)^{\alpha+1}},\end{aligned}$$

where the third inequality thanks to [11, (2.9)]. The proof is complete. □

Proof of Lemma 2.6. According to [6, Lemma 2.4], we have

$$|\nabla p(1,x)| \leqslant \frac{C_{\alpha,d}}{(1+|x|)^{\alpha+\gamma}}, \quad |\nabla^2 p(1,x)| \leqslant \frac{C_{\alpha,d}}{(1+|x|)^{\alpha+\gamma}},$$

then, noticing that $\alpha+\gamma>d$, (2.18) can be obtained by the same argument as the proof of [11, (2.28)]. The proof is complete. □

Remark 3.1. *Noticing that*

$$\nu(d\theta) = ag(\theta)d\theta + b\sum_{i=1}^d(\sigma_i\delta_{e_i}+\sigma_i'\delta_{-e_i}) + c\nu_\gamma(d\theta),$$

according to the construction of the Lévy process, one can write

$$p(1,x) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p_{1,a}(1,x-y-z)p_{2,b}(y)p_{3,c}(z)dydz,$$

where $p(1,x)$, $p_{1,a}(1,x)$, $p_{2,b}(1,x)$ and $p_{3,c}(1,x)$ are the transition probability densities of Z_1, corresponding to the $\nu(d\theta)$, $ag(\theta)d\theta$, $b\sum_{i=1}^d(\sigma_i\delta_{e_i}+\sigma_i'\delta_{-e_i})$ and $c\nu_\gamma(d\theta)$, respectively. What's more, since for any $\alpha+\gamma>d$, we always have

$$\int_{\mathbb{R}^d}\frac{1}{(1+|x|)^{\alpha+\gamma}}\mathrm{d}x \leqslant C_{\alpha,d}$$

and

$$\int_{\mathbb{R}^d}\prod_{i=1}^d\frac{1}{(1+|x_i|)^{\alpha+1}}\mathrm{d}x \leqslant C_{\alpha,d}.$$

Hence, we have by dominated convergence theorem and Fubini's theorem that

$$\nabla p(1,x) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\nabla p_{1,a}(1,x-y-z)p_{2,b}(y)p_{3,c}(z)dydz, \tag{3.2}$$

$$\mathcal{L}^\alpha p(1,x) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathcal{L}^\alpha p_{1,a}(1, x-y-z) p_{2,b}(y) p_{3,c}(z) dy dz, \tag{3.3}$$

and

$$\nabla^2 p(1,x) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \nabla^2 p_{1,a}(1, x-y-z) p_{2,b}(y) p_{3,c}(z) dy dz. \tag{3.4}$$

When $a = 0$, *we can consider*

$$p(1,x) = \int_{\mathbb{R}^d} p_{2,b}(x-z) p_{3,c}(z) dz,$$

other steps are similar to the above. Therefore, without loss of generality, in the following proof, we consider the case $a > 0$.

We are ready to prove our first main result.

Proof of Theorem 2.7. Denote $s = (1 - \mathrm{e}^{-t})$ and $z = y - \mathrm{e}^{-\frac{t}{\alpha}} x$, it is easy to check

$$\nabla_x p(s,z) = -\mathrm{e}^{-\frac{t}{\alpha}} \nabla_z p(s,z), \qquad \nabla_y p(s,z) = \nabla_z p(s,z).$$

We have

$$\begin{aligned}
\nabla f(x) &= -\int_0^\infty \int_{\mathbb{R}^d} \nabla_x p(s,z) \big(h(y) - \mu(h)\big) \mathrm{d}y \mathrm{d}t \\
&= \int_0^\infty \int_{\mathbb{R}^d} \mathrm{e}^{-\frac{t}{\alpha}} \nabla_z p(s,z) \big(h(y) - \mu(h)\big) \mathrm{d}y \mathrm{d}t \\
&= \int_0^\infty \int_{\mathbb{R}^d} \mathrm{e}^{-\frac{t}{\alpha}} \nabla_y p(s,z) \big(h(y) - \mu(h)\big) \mathrm{d}y \mathrm{d}t \\
&= -\int_0^\infty \int_{\mathbb{R}^d} \mathrm{e}^{-\frac{t}{\alpha}} p(s,z) \nabla h(y) \mathrm{d}y \mathrm{d}t.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|\nabla f\| &\leqslant \|\nabla h\| \int_0^\infty \mathrm{e}^{-\frac{t}{\alpha}} \int_{\mathbb{R}^d} p(s,z) \mathrm{d}y \mathrm{d}t \\
&= \|\nabla h\| \int_0^\infty \mathrm{e}^{-\frac{t}{\alpha}} \int_{\mathbb{R}^d} p(s,z) \mathrm{d}z \mathrm{d}t = \alpha \|\nabla h\|.
\end{aligned}$$

We further have

$$|\nabla^2 f(x)| \leqslant \int_0^\infty \int_{\mathbb{R}^d} \mathrm{e}^{-\frac{2t}{\alpha}} |\nabla_z p(s,z)| \cdot |\nabla h(y)| \mathrm{d}y \mathrm{d}t.$$

Thanks to the scaling property $p(s,z) = s^{-d/\alpha} p(s^{-1/\alpha} z)$ with $p(x) = p(1,x)$ for $x \in \mathbb{R}^d$, we have

$$\begin{aligned}
\|\nabla^2 f\| &\leqslant \|\nabla h\| \int_0^\infty \mathrm{e}^{-\frac{2t}{\alpha}} \int_{\mathbb{R}^d} s^{-\frac{d+1}{\alpha}} |\nabla p(s^{-\frac{1}{\alpha}} z)| \mathrm{d}y \mathrm{d}t \\
&= \|\nabla h\| \int_0^\infty s^{-1/\alpha} \mathrm{e}^{-\frac{2t}{\alpha}} \int_{\mathbb{R}^d} |\nabla p(u)| \mathrm{d}u \mathrm{d}t,
\end{aligned}$$

where the equality is by taking $u = s^{-1/\alpha} z$. Then, we have by (3.2) and (2.11),

$$\|\nabla^2 f\| \leqslant \|\nabla h\| \int_0^\infty s^{-1/\alpha} \mathrm{e}^{-\frac{2t}{\alpha}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |\nabla p_{1,a}(u-y-z)| p_{2,b}(y) p_{3,c}(z) \mathrm{d}y \mathrm{d}z \mathrm{d}u \mathrm{d}t$$

$$\leqslant C_{\alpha,d}\|\nabla h\| \int_0^\infty s^{-1/\alpha} \mathrm{e}^{-\frac{2t}{\alpha}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} p_{2,b}(y) p_{3,c}(z) \int_{\mathbb{R}^d} \frac{1}{(1+|u-y-z|)^{\alpha+d}} \mathrm{d}u\mathrm{d}y\mathrm{d}z\mathrm{d}t$$

$$\leqslant C_{\alpha,d}\|\nabla h\| \int_0^\infty s^{-1/\alpha} \mathrm{e}^{-\frac{2t}{\alpha}} \mathrm{d}t = C_{\alpha,d} B\big(\frac{\alpha-1}{\alpha}, \frac{2}{\alpha}\big)\|\nabla h\|.$$

The proof is complete. □

Before proving (2.23), we give another representation of the operator $\mathcal{L}^\alpha$.

Lemma 3.2. *Fix $\alpha \in (1,2)$. Let $f \in C^2(\mathbb{R}^d)$ be such that $\|\nabla^2 f\| + \|\nabla f\| < \infty$. We have, for all $x \in \mathbb{R}^d$,*

$$\mathcal{L}^\alpha f(x) = \frac{d_\alpha}{\alpha} \int_{\mathbb{S}^{d-1}} \int_0^\infty \frac{\theta \cdot \nabla f(x+u\theta) - \theta \cdot \nabla f(x)}{u^\alpha} du\nu(d\theta). \tag{3.5}$$

Proof. Recall the definition of operator $\mathcal{L}^\alpha$, one can write

$$\begin{aligned}
\mathcal{L}^\alpha f(x) &= d_\alpha \int_{\mathbb{S}^{d-1}} \int_0^\infty \int_0^r \frac{\theta \cdot \nabla f(x+u\theta) - \theta \cdot \nabla f(x)}{r^{1+\alpha}} dudr\nu(d\theta) \\
&= d_\alpha \int_{\mathbb{S}^{d-1}} \int_0^\infty \int_u^\infty \frac{\theta \cdot \nabla f(x+u\theta) - \theta \cdot \nabla f(x)}{r^{1+\alpha}} drdu\nu(d\theta) \\
&= \frac{d_\alpha}{\alpha} \int_{\mathbb{S}^{d-1}} \int_0^\infty \frac{\theta \cdot \nabla f(x+u\theta) - \theta \cdot \nabla f(x)}{u^\alpha} du\nu(d\theta),
\end{aligned}$$

the desired result follows. □

Now we are in a position to prove (2.23). Using (3.5), we can write

$$\begin{aligned}
&\Big|\mathcal{L}^\alpha f(x) - \mathcal{L}^\alpha f(y)\Big| \\
=&\Big|\frac{d_\alpha}{\alpha} \int_{\mathbb{S}^{d-1}} \int_0^\infty \frac{\theta \cdot \nabla f(x+u\theta) - \theta \cdot \nabla f(x) - \theta \cdot \nabla f(y+u\theta) + \theta \cdot \nabla f(y)}{u^\alpha} du\nu(d\theta)\Big| \\
\leqslant&\frac{d_\alpha}{\alpha} \int_{\mathbb{S}^{d-1}} \int_0^\infty \frac{\Big|f(x+u\theta) - f(x) - f(y+u\theta) + f(y)\Big|}{u^\alpha} du\nu(d\theta) \\
\leqslant&\frac{2d_\alpha}{\alpha}\|\nabla^2 f\||x-y| \int_{\mathbb{S}^{d-1}} \int_{|x-y|}^\infty \frac{1}{u^\alpha} du\nu(d\theta) + \frac{2d_\alpha}{\alpha}\|\nabla^2 f\| \int_{\mathbb{S}^{d-1}} \int_0^{|x-y|} \frac{1}{u^{\alpha-1}} du\nu(d\theta) \\
=&\frac{2d_\alpha\|\nabla^2 f\|_\infty}{\alpha(2-\alpha)(\alpha-1)}|x-y|^{2-\alpha},
\end{aligned}$$

ending the proof. □

4. Proof of Theorem 2.8

We start with two auxiliary lemmas.

Lemma 4.1. *Let X be a d-dimensional random vector with distribution function $F_X(x)$ and $\mathbb{E}|X| < \infty$, then we have*

$$\mathbb{E}[X \cdot \nabla f(X) - X \cdot \nabla f(0)] = \frac{\alpha^2}{d_\alpha}\mathcal{L}^\alpha f(0) + \mathcal{R},$$

where $\mathcal{R} = \int_{\mathbb{S}^{d-1}} \int_0^\infty r\theta \cdot \big(\nabla f(r\theta) - \nabla f(0)\big)\Big[F_X(\mathrm{d}r\mathrm{d}\theta) - \frac{\alpha}{r^{\alpha+1}}\mathrm{d}r\nu(d\theta)\Big]$.

Proof. We have by (3.5)

$$\mathcal{L}^\alpha f(0) = \frac{d_\alpha}{\alpha}\int_{\mathbb{S}^{d-1}}\int_0^\infty \frac{u\theta\cdot\nabla f(u\theta) - u\theta\cdot\nabla f(0)}{u^{\alpha+1}} du\nu(d\theta),$$

which implies

$$\begin{aligned}&\mathbb{E}[X\cdot\nabla f(X) - X\cdot\nabla f(0)]\\ =&\frac{\alpha^2}{d_\alpha}\mathcal{L}^\alpha f(0) + \int_{\mathbb{S}^{d-1}}\int_0^\infty \big(r\theta\cdot\nabla f(r\theta) - r\theta\cdot\nabla f(0)\big)\Big[F_X(\mathrm{d}r\mathrm{d}\theta) - \frac{\alpha}{r^{\alpha+1}}\mathrm{d}r\nu(d\theta)\Big].\end{aligned}$$

The proof is complete. □

Lemma 4.2. *Let* $\zeta_{n,i}$ *and* $\eta_{n,i}$ $i = 1, \cdots, n$ *are defined as above. Denote the distribution function of* $\eta_{n,i}$ *by* $F_{\eta_{n,i}}$, *then we have*

$$\begin{aligned}\mathbb{E}[S_n\cdot\nabla f(S_n)] =&\frac{\alpha}{n}\sum_{i=1}^n \mathbb{E}[\mathcal{L}^\alpha f\big(S_n(i)\big)] + \mathbb{E}\sum_{i=1}^n \mathcal{R}_i\\ &+\sum_{i=1}^n l_n^{-1/\alpha}\mathbb{E}\eta_{n,i}\cdot\mathbb{E}\big[\nabla f\big(S_n(i)\big) - \nabla f\big(S_n(i) + l_n^{-1/\alpha}\eta_{n,i}\big)\big],\end{aligned}$$

where

$$\mathcal{R}_i = l_n^{-1/\alpha}\int_{\mathbb{S}^{d-1}}\int_0^\infty r\theta\cdot\Big(\nabla f\big(S_n(i) + l_n^{-1/\alpha}r\theta\big) - \nabla f\big(S_n(i)\big)\Big)\Big[F_{\eta_{n,i}}(drd\theta) - \frac{\alpha}{r^{\alpha+1}}dr\nu(d\theta)\Big].$$

Proof.

$$\begin{aligned}\mathbb{E}[S_n\cdot\nabla f(S_n)] &= \sum_{i=1}^n \mathbb{E}\big[\zeta_{n,i}\cdot\nabla f\big(S_n(i) + \zeta_{n,i}\big)\big]\\ &= \sum_{i=1}^n \mathbb{E}\big[\big(l_n^{-1/\alpha}\eta_{n,i} - l_n^{-1/\alpha}\mathbb{E}\eta_{n,i}\big)\cdot\nabla f\big(S_n(i) + l_n^{-1/\alpha}\eta_{n,i}\big)\big]\\ &= \sum_{i=1}^n \mathbb{E}\Big[l_n^{-1/\alpha}\eta_{n,i}\cdot\nabla f\big(S_n(i) + l_n^{-1/\alpha}\eta_{n,i}\big) - l_n^{-1/\alpha}\eta_{n,i}\cdot\nabla f\big(S_n(i)\big)\Big]\\ &\quad+\sum_{i=1}^n l_n^{-1/\alpha}\mathbb{E}\eta_{n,i}\cdot\mathbb{E}\big[\nabla f\big(S_n(i)\big) - \nabla f\big(S_n(i) + l_n^{-1/\alpha}\eta_{n,i}\big)\big],\end{aligned}$$

and we have by independence and Lemma 4.1

$$\begin{aligned}&\sum_{i=1}^n \mathbb{E}\Big[l_n^{-1/\alpha}\eta_{n,i}\cdot\nabla f\big(S_n(i) + l_n^{-1/\alpha}\eta_{n,i}\big) - l_n^{-1/\alpha}\eta_{n,i}\cdot\nabla f\big(S_n(i)\big)\Big]\\ =&\sum_{i=1}^n \mathbb{E}\big[\eta_{n,i}\cdot\nabla\eta_{n,i} f\big(S_n(i) + l_n^{-1/\alpha}\eta_{n,i}\big) - \eta_{n,i}\cdot\nabla z f\big(S_n(i) + l_n^{-1/\alpha}z\big)\mathbf{1}_{\{z=0\}}\big]\\ =&\sum_{i=1}^n \Big\{\frac{\alpha^2}{d_\alpha}\mathbb{E}\big[l_n^{-1}\mathcal{L}^\alpha f\big(S_n(i)\big)\big] + \mathbb{E}\mathcal{R}_i\Big\}\end{aligned}$$

$$= \frac{\alpha}{n} \sum_{i=1}^{n} \mathbb{E}[\mathcal{L}^{\alpha} f\big(S_n(i)\big)] + \mathbb{E} \sum_{i=1}^{n} \mathcal{R}_i,$$

which get the desired results. □

Proof of Theorem 2.8. By Stein's equation (2.5) and Lemma 4.2, we have

$$\begin{aligned}
&\Big|\mathbb{E}\big[h(S_n)\big] - \mu(h)\Big| \\
=&\Big|\mathbb{E}[\mathcal{L}^{\alpha} f(S_n) - \frac{1}{\alpha} S_n \cdot \nabla f(S_n)]\Big| \\
\leqslant&\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\Big|\mathcal{L}^{\alpha} f(S_n) - \mathcal{L}^{\alpha} f\big(S_n(i)\big)\Big| + \frac{1}{\alpha}\mathbb{E}\sum_{i=1}^{n} |\mathcal{R}_i| \\
&+ l_n^{-1/\alpha} \sum_{i=1}^{n} \big|\mathbb{E}\eta_{n,i}\big| \cdot \mathbb{E}\big|\nabla f\big(S_n(i)\big) - \nabla f\big(S_n(i) + l_n^{-1/\alpha}\eta_{n,i}\big)\big| \\
\leqslant&\frac{C_{\alpha,d}}{n}\|\nabla h\| \sum_{i=1}^{n} \mathbb{E}|\zeta_{n,i}|^{2-\alpha} + \frac{1}{\alpha}\mathbb{E}\sum_{i=1}^{n} |\mathcal{R}_i| + C_{\alpha,d}\|\nabla h\| l_n^{-2/\alpha} \sum_{i=1}^{n} \big(\mathbb{E}|\eta_{n,i}|\big)^2 \\
\leqslant&\frac{C_{\alpha,d}}{n} l_n^{\frac{2-\alpha}{\alpha}} \sum_{i=1}^{n} \mathbb{E}|\eta_{n,i}|^{2-\alpha} + \frac{1}{\alpha}\mathbb{E}\sum_{i=1}^{n} |\mathcal{R}_i| + C_{\alpha,d} l_n^{-2/\alpha} \sum_{i=1}^{n} \big(\mathbb{E}|\eta_{n,i}|\big)^2
\end{aligned}$$

where the last inequality follows from Theorem 2.7. Furthermore, for any $N > 0$, by Theorem 2.7, one has

$$\begin{aligned}
|\mathcal{R}_i| \leqslant C_{\alpha,d}\Big[&n^{-\frac{2}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_0^N r^2 \Big|F_{\eta_{n,i}}(drd\theta) - \frac{\alpha}{r^{\alpha+1}} \mathrm{d}r\nu(d\theta)\Big| \\
&+ n^{-\frac{1}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_N^{\infty} r \Big|F_{\eta_{n,i}}(drd\theta) - \frac{\alpha}{r^{\alpha+1}} \mathrm{d}r\nu(d\theta)\Big|\Big],
\end{aligned}$$

finishing the proof. □

Remark 4.3. *By the scaling property of stable distribution, if X has a stable distribution μ with characteristic function* $\exp\big(-\sigma\psi(\lambda)\big)$, *then $\sigma^{-1/\alpha}X$ has a distribution $\tilde{\mu}$ with characteristic function* $\exp\big(-\psi(\lambda)\big)$. *Recall the definition of the Wasserstein distance, we have*

$$d_W\big(\mathcal{L}(\sigma^{-1/\alpha} S_n), \tilde{\mu}\big) = \sigma^{-1/\alpha} d_W\big(\mathcal{L}(S_n), \mu\big).$$

5. Examples

5.1. **Example 1: Approximation of Multidimensional Stable Laws**[14]**.** In [14], Davydov and Nagaev consider a random variable ξ have the Pareto distribution with the density

$$p(u) = \begin{cases} \alpha u^{-1-\alpha} & if \quad u \geqslant 1, \\ 0 & if \quad u < 1. \end{cases} \tag{5.1}$$

It is convenient to adhere the following definition.

Definition 5.1. *We call a distribution $\nu-$Paretian if it corresponds to a random vector τ admitting the representation $\xi\varepsilon$, where ξ and ε are independent, ξ has the density (5.1) while ε is a random unit vector satisfying*

$$P(\varepsilon \in E) = \nu(E), E \in \mathfrak{B}_{\mathbb{S}^{d-1}}, \tag{5.2}$$

In [14], the authors assumed that ν is symmetric and

$$m_\nu = \min_{e\in\mathcal{S}^{d-1}} \Sigma_\alpha(e,\nu) > 0,$$

where $\Sigma_\alpha(e,\nu) = \int_{\mathcal{S}^{d-1}} |\langle e,\theta\rangle|^\alpha \nu(\mathrm{d}\theta)$. That means the $v-$Paretian distribution is strictly $d-$dimensional. Consider a sequence of i.i.d. random vectors such that

$$\tau_i =^d \tau, \qquad i = 1, 2, \cdots .$$

Set

$$T_n = n^{-1/\alpha} \sum_{i=1}^n \tau_i.$$

Let a random vector T have the stable distribution determined by the characteristic function

$$f(\lambda) = E\mathrm{e}^{i\langle\lambda,\zeta\rangle} = \exp(-\frac{\alpha}{d_\alpha}|\lambda|^\alpha \Sigma_\alpha(\mathrm{e}_\lambda,\nu)), \qquad \lambda \in \mathbb{R}^d, d \geqslant 1.$$

Set

$$S_\alpha(A) = P(T \in A), \qquad P_n(A) = P(T_n \in A). \tag{5.3}$$

Let $\mathbf{d}(P,Q)$ denote the uniform distance between two measures P and Q; that is

$$\mathbf{d}(P,Q) = \sup_{A\in\mathfrak{B}_{R^d}} |P(A) - Q(A)|.$$

Based on above, we recall the approximation of multidimensional stable law:

Theorem 5.2. [14, Theorem 3.2] *Let S_α, P_n be defined as in (5.3). If the underlying distribution is $\nu-$Paretian then as $n \to \infty$*

$$\mathbf{d}(P_n, S_\alpha) = \mathbf{O}(n^{-\beta}),$$

where $\beta = \frac{\min(\alpha,2-\alpha)}{d+\alpha}$.

By the above Theorem, we immediately get

Lemma 5.3. *As $n \to \infty$, $T_n \Rightarrow S_\alpha$, where S_α is a symmetric stable distribution with characteristic function $\exp(-\frac{\alpha}{d_\alpha}|\lambda|^\alpha\Sigma_\alpha(\mathrm{e}_\lambda,\nu))$. In particular, it follows from the scaling property of stable distribution that as $n \to \infty$,*

$$(\frac{\alpha}{d_\alpha})^{-\frac{1}{\alpha}} T_n \Rightarrow \hat{\mu}, \tag{5.4}$$

where $\hat{\mu}$ is a symmetric stable distribution with characteristic function $\exp(-|\lambda|^\alpha\Sigma_\alpha(\mathrm{e}_\lambda,\nu))$.

By Lemma 5.3, denote

$$\zeta_{n,i} = (\frac{\alpha}{d_\alpha})^{-\frac{1}{\alpha}} \frac{\tau_i}{n^{\frac{1}{\alpha}}}$$

for $i = 1, \cdots, n$, S_n weakly converges to a stable distribution μ with characteristic function $\exp(-|\lambda|^\alpha\Sigma_\alpha(\mathrm{e}_\lambda,\nu))$.

However, according to Theorem 2.8, we can consider the more general ν defined by (2.20) and get a convergence rate $n^{-\frac{2-\alpha}{\alpha}}$. That is, set

$$\zeta_{n,i} = \big(\frac{\alpha}{d_\alpha}\big)^{-\frac{1}{\alpha}} \frac{\tau_i}{n^{\frac{1}{\alpha}}}, \quad S_n = \big(\zeta_{n,1} - \mathbb{E}\zeta_{n,1}\big) + \big(\zeta_{n,2} - \mathbb{E}\zeta_{n,2}\big) + \cdots + \big(\zeta_{n,n} - \mathbb{E}\zeta_{n,n}\big).$$

Then,

$$d_W\big(\mathcal{L}(S_n), \mu\big) \leqslant C_{\alpha,d} n^{-\frac{2-\alpha}{\alpha}}.$$

Proof. By definition 5.1, we obtain

$$F_{\tau_i}(drd\theta) = \begin{cases} \frac{\alpha}{r^{\alpha+1}} dr\nu(d\theta), & r \geqslant 1, \\ 0, & r < 1. \end{cases}$$

Let $\zeta_{n,i} = l_n^{-1/\alpha}\tau_i$ and $\eta_{n,i} = l_n^{1/\alpha}\zeta_{n,i} = \tau_i$, it follows that

$$F_{\eta_{n,i}}(drd\theta) = \begin{cases} \frac{\alpha}{r^{\alpha+1}} dr\nu(d\theta), & r \geqslant 1, \\ 0, & r < 1. \end{cases}$$

According to Theorem 2.8, since

$$n^{-\frac{2}{\alpha}} \sum_{i=1}^{n} \mathbb{E}|\eta_{n,i}|^{2-\alpha} \leqslant C_{\alpha,d} n^{\frac{\alpha-2}{\alpha}},$$

$$n^{-\frac{2}{\alpha}} \sum_{i=1}^{n} \big(\mathbb{E}|\eta_{n,i}|\big)^2 \leqslant C_{\alpha,d} n^{\frac{\alpha-2}{\alpha}}$$

and choose $N > 1$

$$|\mathcal{R}_{n,i}| = n^{-\frac{2}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_0^1 \frac{\alpha}{r^{\alpha-1}} dr\nu(d\theta) = \frac{\alpha}{2-\alpha} n^{-\frac{2}{\alpha}},$$

we have

$$d_W\big(\mathcal{L}(S_n), \mu\big) \leqslant C_{\alpha,d} n^{\frac{\alpha-2}{\alpha}}.$$

The proof is complete. □

Let us compare our result with the known results in literatures. When $\alpha \in (1, 2)$, the authors of [14] obtained a rate $n^{-\frac{2-\alpha}{d+\alpha}}$ for d dimensional stable law in total variation distance and conjectured that the rate can be improved to $n^{-\frac{2-\alpha}{\alpha}}$ in L^1 or total variation distance. Our results gives a positive answer to their conjecture for the L^1 distance case.

5.2. **Example 2: Convergence rate of Pareto densities with modified tails.** Assume that $\xi_1, \cdots, \xi_n, \cdots$ be i.i.d. random vectors with a distribution function $F_{\xi_i}(r\theta)$ satisfying

$$F_{\xi_i}(drd\theta) = \begin{cases} \frac{A}{r^{\alpha+1}} dr\nu(d\theta) + \frac{B(r\theta)}{r^{\beta+1}} drd\theta, & r \geqslant 1; \\ 0, & r < 1, \end{cases}$$

with $\beta \in (\alpha, \infty)$, $|B(r\theta)| \leqslant B$ for some constant $B > 0$. We denote $L_n = \frac{A}{d_\alpha} n$ and let $\zeta_{n,i} = L_n^{-1/\alpha}\xi_i$, then we know $\eta_{n,i} = (\frac{A}{\alpha})^{-1/\alpha}\xi_i$ and the distribution function of $\eta_{n,i}$ satisfying

$$F_{\eta_{n,i}}(drd\theta) = \begin{cases} \frac{\alpha}{r^{\alpha+1}} dr\nu(d\theta) + \frac{B\left((\frac{A}{\alpha})^{\frac{1}{\alpha}} r\theta\right)}{(\frac{A}{\alpha})^{\frac{\beta}{\alpha}} r^{\beta+1}} drd\theta, & r \geqslant (\frac{A}{\alpha})^{-1/\alpha}; \\ 0, & r < (\frac{A}{\alpha})^{-1/\alpha}. \end{cases}$$

According to Theorem 2.8, it is straightforward to check that

$$n^{-\frac{2}{\alpha}} \sum_{i=1}^{n} \mathbb{E}|\eta_{n,i}|^{2-\alpha} \leqslant C_{\alpha,d,A} n^{\frac{\alpha-2}{\alpha}},$$

and

$$n^{-\frac{2}{\alpha}} \sum_{i=1}^{n} \left(\mathbb{E}|\eta_{n,i}|\right)^2 \leqslant C_{\alpha,d,A} n^{\frac{\alpha-2}{\alpha}}.$$

It remains to compute the remainder in the bound of Theorem 2.8. We choose $N = n^{\frac{1}{\alpha}} > (\frac{A}{\alpha})^{-1/\alpha}$, then

$$\begin{aligned} |\mathcal{R}_{n,i}| \leqslant{}& n^{-\frac{2}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_0^{(\frac{A}{\alpha})^{-1/\alpha}} \frac{\alpha}{r^{\alpha-1}} dr\nu(d\theta) + n^{-\frac{2}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_{(\frac{A}{\alpha})^{-1/\alpha}}^{n^{\frac{1}{\alpha}}} \frac{\left|B\left((\frac{A}{\alpha})^{\frac{1}{\alpha}} r\theta\right)\right|}{(\frac{A}{\alpha})^{\frac{\beta}{\alpha}} r^{\beta-1}} \mathrm{d}rd\theta \\ &+ n^{-\frac{1}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_{n^{\frac{1}{\alpha}}}^{\infty} \frac{\left|B\left((\frac{A}{\alpha})^{\frac{1}{\alpha}} r\theta\right)\right|}{(\frac{A}{\alpha})^{\frac{\beta}{\alpha}} r^{\beta}} drd\theta \\ :={}& \mathrm{I} + \mathrm{II} + \mathrm{III}, \end{aligned}$$

and it is easy to compute that

$$\mathrm{I} \leqslant C_{\alpha,d,A} n^{-\frac{2}{\alpha}},$$

$$\mathrm{III} \leqslant n^{-\frac{1}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_{n^{\frac{1}{\alpha}}}^{\infty} \frac{B}{(\frac{A}{\alpha})^{\frac{\beta}{\alpha}} r^{\beta}} drd\theta \leqslant C_{\alpha,d,A,B} n^{-\frac{\beta}{\alpha}},$$

and

$$\mathrm{II} \leqslant n^{-\frac{2}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_{(\frac{A}{\alpha})^{-1/\alpha}}^{n^{\frac{1}{\alpha}}} \frac{B}{(\frac{A}{\alpha})^{\frac{\beta}{\alpha}} r^{\beta-1}} \mathrm{d}rd\theta,$$

when $\beta \neq 2$, we have

$$\mathrm{II} \leqslant C_{\alpha,d,A,B} n^{-\frac{2}{\alpha}} \left[n^{\frac{2-\beta}{\alpha}} - (\frac{A}{\alpha})^{-\frac{2-\beta}{\alpha}}\right];$$

when $\beta = 2$, we have

$$\mathrm{II} \leqslant C_{\alpha,d,A,B} n^{-\frac{2}{\alpha}} \left[\log n^{\frac{1}{\alpha}} + \frac{1}{\alpha} \log(\frac{A}{\alpha})\right].$$

Therefore, we have
(1). When $\beta \neq 2$,

$$d_W\left(\mathcal{L}(S_n), \mu\right) \leqslant C_{\alpha,d,A,B}\left(n^{\frac{\alpha-2}{\alpha}} + n^{\frac{\alpha-\beta}{\alpha}}\right).$$

(2). When $\beta = 2$,

$$d_W\big(\mathcal{L}(S_n), \mu\big) \leqslant C_{\alpha,d,A,B} n^{\frac{\alpha-2}{\alpha}} \log n.$$

Now, we can consider a more general case:

$$F_{\xi_i}(drd\theta) = \frac{A}{r^{\alpha+1}} dr\nu(\mathrm{d}\theta) + \frac{\epsilon(r\theta)}{r^{\alpha+1}} drd\theta,$$

where $\lim_{r\to\infty} |\epsilon(r\theta)| = 0$.

We denote $L_n = \frac{A}{d_\alpha} n$ and let $\zeta_{n,i} = L_n^{-\frac{1}{\alpha}} \xi_i$, then we know $\eta_{n,i} = (\frac{A}{\alpha})^{-1/\alpha} \xi_i$ and the distribution of $\eta_{n,i}$ satisfies

$$F_{\xi_i}(r\theta) = \frac{\alpha}{r^{\alpha+1}} dr\nu(\mathrm{d}\theta) + \frac{\alpha\epsilon\big((\frac{A}{\alpha})^{\frac{1}{\alpha}} r\theta\big)}{Ar^{\alpha+1}} drd\theta.$$

According to Theorem 2.8, it is straightforward to check that

$$n^{-\frac{2}{\alpha}} \sum_{i=1}^n \mathbb{E}|\eta_{n,i}|^{2-\alpha} \leqslant C_{\alpha,d,A} n^{\frac{\alpha-2}{\alpha}},$$

and

$$n^{-\frac{2}{\alpha}} \sum_{i=1}^n \big(\mathbb{E}|\eta_{n,i}|\big)^2 \leqslant C_{\alpha,d,A} n^{\frac{\alpha-2}{\alpha}}.$$

It remains to compute the remainder in the bound of Theorem 2.8. We choose $N = n^{\frac{1}{\alpha}}$, then

$$\begin{aligned} |\mathcal{R}_{n,i}| &\leqslant n^{-\frac{2}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_0^{n^{\frac{1}{\alpha}}} \frac{\alpha\big|\epsilon\big((\frac{A}{\alpha})^{\frac{1}{\alpha}} r\theta\big)\big|}{Ar^{\alpha-1}} drd\theta + n^{-\frac{1}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_{n^{\frac{1}{\alpha}}}^\infty \frac{\alpha\big|\epsilon\big((\frac{A}{\alpha})^{\frac{1}{\alpha}} r\theta\big)\big|}{Ar^{\alpha}} drd\theta \\ &\leqslant n^{-\frac{2}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_0^{n^{\frac{1}{\alpha}}} \frac{\alpha\big|\epsilon\big((\frac{A}{\alpha})^{\frac{1}{\alpha}} r\theta\big)\big|}{Ar^{\alpha-1}} drd\theta + n^{-\frac{1}{\alpha}} \sup_{r \geqslant (\frac{An}{\alpha})^{\frac{1}{\alpha}}} |\epsilon(r\theta)| \int_{\mathbb{S}^{d-1}} \int_{n^{\frac{1}{\alpha}}}^\infty \frac{\alpha}{Ar^\alpha} drd\theta \\ &\leqslant C_{\alpha,d,A} \Big[n^{-\frac{2}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_0^{n^{\frac{1}{\alpha}}} \frac{\big|\epsilon\big((\frac{A}{\alpha})^{\frac{1}{\alpha}} r\theta\big)\big|}{r^{\alpha-1}} drd\theta + n^{-1} \sup_{r \geqslant (\frac{An}{\alpha})^{\frac{1}{\alpha}}} |\epsilon(r\theta)| \Big]. \end{aligned}$$

Therefore, we have

$$d_W\big(\mathcal{L}(S_n), \mu\big) \leqslant C_{\alpha,d,A,B} \Big[n^{\frac{\alpha-2}{\alpha}} + \sup_{r \geqslant (\frac{An}{\alpha})^{\frac{1}{\alpha}}} |\epsilon(r\theta)| + n^{\frac{\alpha-2}{\alpha}} \int_{\mathbb{S}^{d-1}} \int_0^{n^{\frac{1}{\alpha}}} \frac{\big|\epsilon\big((\frac{A}{\alpha})^{\frac{1}{\alpha}} r\theta\big)\big|}{r^{\alpha-1}} drd\theta \Big].$$

5.3. **Example 3 in** [19]**.** Let us assume that $\xi_1, \cdots, \xi_n, \cdots$ be s sequence of i.i.d. random vectors with a density

$$p(x) = \begin{cases} \frac{K_0\big[\alpha(\log|x|)^\beta - \beta(\log|x|)^{\beta-1}\big]}{|x|^{\alpha+d}} g(\frac{x}{|x|}), & |x| \geqslant \mathrm{e}; \\ 0, & |x| < \mathrm{e}. \end{cases}$$

By [13, Theorem 3.7.2], $A_n = \inf\{x > 0 : \mathbb{P}(|\xi_1| > x) \leqslant \frac{1}{n}\}$ can be determined by $\frac{K_0(\log A_n)^\beta}{A_n^\alpha} = \frac{1}{n}$, which gives

$$\frac{n}{A_n^\alpha} = \frac{1}{K_0(\log A_n)^\beta},$$

and it is easy to see $C_{\alpha,\beta} n^{\frac{1}{\alpha}} \leqslant A_n \leqslant C_{\alpha,\beta} n^{\frac{1}{\alpha}} (\log n)^{\frac{\beta}{\alpha}}$.

Now, we consider

$$\zeta_{n,i} = \frac{1}{\tilde{A}_n}\xi_i \quad with \quad \tilde{A}_n = (\frac{\alpha}{d_\alpha})^{1/\alpha} A_n.$$

Then we know $\eta_{n,i} = \frac{n^{1/\alpha}\xi_i}{A_n}$ and the density of $\eta_{n,i}$ is

$$p_{\eta_{n,i}}(x) = \begin{cases} \frac{\alpha(\log|\frac{A_n}{n^{1/\alpha}}x|)^\beta - \beta(\log|\frac{A_n}{n^{1/\alpha}}x|)^{\beta-1}}{(\log A_n)^\beta |x|^{\alpha+d}} g(\frac{x}{|x|}), & |x| \geqslant \frac{n^{1/\alpha}}{A_n}\mathrm{e}; \\ 0, & |x| < \frac{n^{1/\alpha}}{A_n}\mathrm{e}. \end{cases}$$

By Theorem 2.8, it is straightforward to check that

$$n^{-\frac{2}{\alpha}} \sum_{i=1}^n \mathbb{E}|\eta_{n,i}|^{2-\alpha} \leqslant C_{\alpha,d,K_0,\beta} A_n^{\alpha-2},$$

and

$$n^{-\frac{2}{\alpha}} \sum_{i=1}^n \left(\mathbb{E}|\eta_{n,i}|\right)^2 \leqslant C_{\alpha,d,K_0,\beta} n A_n^{-2} \leqslant C_{\alpha,d,K_0,\beta} n^{\frac{\alpha-2}{\alpha}}.$$

It remains to compute the remainder in the bound of Theorem 2.8. When $\nu(d\theta) = g(\theta)d\theta$, let $x = r\theta$, we have

$$\frac{\alpha}{r^{\alpha+1}} dr \nu(d\theta) = \frac{\alpha}{r^{\alpha+1}} g(\theta) dr d\theta = \frac{\alpha}{|x|^{\alpha+d}} g(\frac{x}{|x|}) \mathrm{d}x,$$

Hence,

$$\begin{aligned} |\mathcal{R}_{n,i}| &= n^{-\frac{2}{\alpha}} \int_{|x|\leqslant N} |x|^2 \Big| p_{\eta_{n,i}}(x) - \frac{\alpha}{|x|^{\alpha+d}} g(\frac{x}{|x|}) \Big| \mathrm{d}x + n^{-\frac{1}{\alpha}} \int_{|x|>N} |x| \Big| p_{\eta_{n,i}}(x) - \frac{\alpha}{|x|^{\alpha+d}} g(\frac{x}{|x|}) \Big| \mathrm{d}x \\ &= n^{\frac{d}{\alpha}} \int_{|y|\leqslant n^{-\frac{1}{\alpha}}N} |y|^2 \Big| p_{\eta_{n,i}}(n^{\frac{1}{\alpha}}y) - \frac{\alpha}{n^{\frac{\alpha+d}{\alpha}}|y|^{\alpha+d}} g(\frac{y}{|y|}) \Big| \mathrm{d}y \\ &\quad + n^{\frac{d}{\alpha}} \int_{|y|>n^{-\frac{1}{\alpha}}N} |y| \Big| p_{\eta_{n,i}}(n^{\frac{1}{\alpha}}y) - \frac{\alpha}{n^{\frac{\alpha+d}{\alpha}}|y|} g(\frac{y}{|y|}) \Big| \mathrm{d}y \\ &:= \mathcal{J}_1 + \mathcal{J}_2. \end{aligned}$$

Furthermore, we choose $N = (n \log A_n)^{\frac{1}{\alpha}}$, then

$$\begin{aligned} \mathcal{J}_1 \leqslant & \frac{\alpha}{n} \int_{|y|\leqslant A_n^{-1}\mathrm{e}} \frac{1}{|y|^{\alpha+d-2}} g(\frac{y}{|y|}) \mathrm{d}y + \frac{\beta}{n} \int_{A_n^{-1}\mathrm{e}\leqslant|y|\leqslant(\log A_n)^{\frac{1}{\alpha}}} \frac{(\log|A_n y|)^{\beta-1}}{(\log A_n)^\beta |y|^{\alpha+d-2}} g(\frac{y}{|y|}) \mathrm{d}y \\ & + \frac{\alpha}{n} \int_{A_n^{-1}\mathrm{e}\leqslant|y|\leqslant(\log A_n)^{\frac{1}{\alpha}}} \Big| \frac{(\log|A_n y|)^\beta}{(\log A_n)^\beta |y|^{\alpha+d-2}} - \frac{1}{|y|^{\alpha+d}} \Big| g(\frac{y}{|y|}) \mathrm{d}y \\ & := \mathcal{J}_{11} + \mathcal{J}_{12} + \mathcal{J}_{13}. \end{aligned}$$

Throughout some computations, we have

$$\mathcal{J}_{11} \leqslant C_{\alpha,d} n^{-1} A_n^{\alpha-2}$$

and

$$\begin{aligned}\mathcal{J}_{12} &= \frac{\beta}{n}\int_{A_n^{-1}\mathrm{e}\leqslant|y|\leqslant(\log A_n)^{\frac{1}{\alpha}}} \frac{(\log A_n+\log|y|)^{\beta-1}}{(\log A_n)^{\beta}|y|^{\alpha+d-2}} g(\frac{y}{|y|})\mathrm{d}y \\ &= \frac{\beta}{n}\int_{A_n^{-1}\mathrm{e}\leqslant|y|\leqslant(\log A_n)^{\frac{1}{\alpha}}} \frac{(\log A_n+\log|y|)^{\beta-1}-(\log A_n)^{\beta-1}+(\log A_n)^{\beta-1}}{(\log A_n)^{\beta}|y|^{\alpha+d-2}} g(\frac{y}{|y|})\mathrm{d}y.\end{aligned}$$

By the fact $A_n^{-1} \leqslant |y| \leqslant A_n$ and $|1-(1+x)^{\beta-1}| \leqslant C_\beta|x|$ for any $|x|<1$, we have

$$\mathcal{J}_{12} \leqslant \frac{C_\beta}{n}\frac{1}{\log A_n}\int_{A_n^{-1}\mathrm{e}\leqslant|y|\leqslant(\log A_n)^{\frac{1}{\alpha}}} \frac{\frac{\log|y|}{\log A_n}+1}{|y|^{\alpha+d-2}} g(\frac{y}{|y|})\mathrm{d}y \leqslant \frac{C_{\alpha,d,\beta}}{n}(\log n)^{-1+\frac{1}{\alpha}}.$$

By the same argument as the proof of $\mathcal{J}_{12}$, we can obtain

$$\mathcal{J}_{13} \leqslant \frac{C_{\alpha,d,\beta}}{n}(\log n)^{-1+\frac{1}{\alpha}}.$$

For $\mathcal{J}_2$, we have

$$\begin{aligned}\mathcal{J}_2 \leqslant & \frac{\beta}{n}\int_{|y|>(\log A_n)^{\frac{1}{\alpha}}} \frac{(\log|A_n y|)^{\beta-1}}{(\log A_n)^{\beta}|y|^{\alpha+d-2}} g(\frac{y}{|y|})\mathrm{d}y \\ &+ \frac{\alpha}{n}\int_{|y|>(\log A_n)^{\frac{1}{\alpha}}} \Big|\frac{(\log|A_n y|)^{\beta}}{(\log A_n)^{\beta}|y|^{\alpha+d-2}} - \frac{1}{|y|^{\alpha+d}}\Big| g(\frac{y}{|y|})\mathrm{d}y := \mathcal{J}_{21}+\mathcal{J}_{22}.\end{aligned}$$

According to $A_n \geqslant C_{\alpha,\beta} n^{\frac{1}{\alpha}}$, we have

$$\begin{aligned}\mathcal{J}_{21} &= \frac{\beta}{n}(\log A_n)^{\frac{1-\alpha}{\alpha}}\int_{|x|>1} \frac{(\log|A_n(\log A_n)^{\frac{1}{\alpha}}x|)^{\beta-1}}{(\log A_n)^{\beta}|x|^{\alpha+d-1}} g(\frac{x}{|x|})\mathrm{d}x \\ &\leqslant \frac{C_{\alpha,d,\beta}}{n}(\log A_n)^{\frac{1-\alpha}{\alpha}} \leqslant \frac{C_{\alpha,d,\beta}}{n}(\log n)^{-1+\frac{1}{\alpha}}.\end{aligned}$$

By the same argument as above, we also can obtain

$$\mathcal{J}_{22} \leqslant \frac{C_{\alpha,d,\beta}}{n}(\log n)^{-1+\frac{1}{\alpha}}.$$

Combining all of above inequalities, we have

$$d_W\big(\mathcal{L}(S_n),\mu\big) \leqslant C_{\alpha,d,K_0,\beta}(\log n)^{-1+\frac{1}{\alpha}}.$$

Remark 5.4. *For the above example, we can also consider the mixture ν defined by (2.20), and get the same conclusion.*

Appendix A. Proof of Proposition 2.1

We first give the following lemma:

Lemma A.1. *Let $(Q_t)_{t\geqslant 0}$ be a Markovian semigroup with transition density $q(t,x,y) = p_{1-e^{-t}}(y-e^{-\frac{t}{\alpha}}x)$. Then for any $h \in Lip(1)$, we have*

$$\partial_t Q_t h(x) = \mathcal{A} Q_t h(x). \tag{1.1}$$

Proof. Here, we give the proof for the case $a = 1$, that is, the absolutely continuous type, and other cases can be obtained by the same argument according to Remark 3.1.

Recall that $q(t,x,y) = p_{1-e^{-t}}(y - e^{-\frac{t}{\alpha}}x)$ and $s(t) = 1 - e^{-t}$. Then

$$\begin{aligned}
\left|\frac{\partial}{\partial t}q(t,x,y)\right| &= \left|e^{-t}\frac{\partial}{\partial s(t)}p_{1-e^{-t}}(y - e^{-\frac{t}{\alpha}}x) + \alpha^{-1}e^{-\frac{t}{\alpha}}x\frac{\partial}{\partial y}p_{1-e^{-t}}(y - e^{-\frac{t}{\alpha}}x)\right| \\
&\le \frac{C_{\alpha,d}}{((1-e^{-t})^{1/\alpha} + |y - e^{-\frac{t}{\alpha}}x|)^{\alpha+d}} + \frac{|x|}{\alpha}e^{-\frac{t}{\alpha}}\frac{C_{\alpha,d}(1-e^{-t})^{(\alpha-1)/\alpha}}{((1-e^{-t})^{1/\alpha} + |y - e^{-\frac{t}{\alpha}}x|)^{\alpha+d}} \\
&\le \frac{C_{\alpha,d}(1 + |x|(e^{t}-1)^{-1/\alpha})}{((1-e^{-t})^{1/\alpha} + |y - e^{-\frac{t}{\alpha}}x|)^{\alpha+d}},
\end{aligned}$$

where the second inequality above follows from $\frac{\partial}{\partial t}p_{t^{\frac{1}{\alpha}}}(x) = \mathcal{L}^{\alpha,\beta}p_{t^{\frac{1}{\alpha}}}(x)$ and (2.12). Thus, for $t > 0$, $s > 0$ small enough such that $(1 - e^{-s/\alpha})|x| \le \frac{1}{2}(e^{t}-1)^{1/\alpha}$,

$$|q(t+s,x,y) - q(t,x,y)| \le s\frac{C_{\alpha,d}2^{\alpha+d}(1 + |x|(e^{t}-1)^{-1/\alpha})}{((1-e^{-t})^{1/\alpha} + |y - e^{-\frac{t}{\alpha}}x|)^{\alpha+d}}.$$

In addition, according to (2.4) and (2.1), we have

$$\partial_t q(t,x,y) = \mathcal{A}q(t,x,y). \tag{1.2}$$

Hence, using dominated convergence theorem, (1.2) and Fubini's theorem, we have

$$\begin{aligned}
\partial_t Q_t h(x) &= \partial_t \int_{\mathbb{R}^d} q(t,x,y)h(y)dy = \int_{\mathbb{R}^d} \partial_t q(t,x,y)h(y)dy \\
&= \int_{\mathbb{R}^d} \mathcal{A}q(t,x,y)h(y)dy = \mathcal{A}\int_{\mathbb{R}^d} q(t,x,y)h(y)dy = \mathcal{A}Q_t h(x),
\end{aligned}$$

the desired conclusion follows. □

Proof of Proposition 2.1. Here, we also give the proof for the case $a = 1$, that is, the absolutely continuous type, and other cases can be obtained by the same argument according to Remark 3.1.

First of all, we show that f is well defined. Noticing that μ has a density $p_1(x)$ and $h \in Lip(1)$, we have

$$\begin{aligned}
&\Big|\int_{\mathbb{R}^d} p_{1-e^{-t}}(y - e^{-\frac{t}{\alpha}}x)\big(h(y) - \mu(h)\big)dy\Big| \\
=&\Big|\int_{\mathbb{R}^d} p_1(y)\big[h\big((1-e^{-t})^{\frac{1}{\alpha}}y + e^{-\frac{t}{\alpha}}x\big) - h(y)\big]dy\Big| \\
\leqslant& C_\alpha\|\nabla h\|_\infty e^{-\frac{t}{\alpha}}\big(|x| + \int_{\mathbb{R}^d}|y|p_1(y)dy\big) \leqslant C_{\alpha,d}\|\nabla h\|_\infty e^{-\frac{t}{\alpha}}(|x|+1).
\end{aligned}$$

Hence,

$$\Big|\int_0^\infty \int_{\mathbb{R}} p_{(1-e^{-t})^{\frac{1}{\alpha}},\beta}(y - e^{-\frac{t}{\alpha}}x)\big(h(y) - \mu(h)\big)dydt\Big| \leqslant C_{\alpha,\beta}\|h'\|_\infty(1+|x|),$$

that is, f is well defined. Then, we continue the proof. Observing

$$Q_t h(x) = \int_{\mathbb{R}^d} p_{1-e^{-t}}(y - e^{-\frac{t}{\alpha}}x)h(y)dy = \int_{\mathbb{R}^d} p_1(y)h\big((1-e^{-t})^{\frac{1}{\alpha}}y + e^{-\frac{t}{\alpha}}x\big)dy, \tag{1.3}$$

and

$$\Big|h\big((1-e^{-t})^{\frac{1}{\alpha}}y+e^{-\frac{t}{\alpha}}(x+z)\big)-h\big((1-e^{-t})^{\frac{1}{\alpha}}y+e^{-\frac{t}{\alpha}}x\big)\Big|\leqslant e^{-\frac{t}{\alpha}}|z|.$$

By (1.3), we immediately have

$$\Big|Q_th(x+z)-Q_th(x)\Big|\leqslant\int_{\mathbb{R}^d}p_1(y)e^{-\frac{t}{\alpha}}|z|dy=e^{-\frac{t}{\alpha}}|z|. \tag{1.4}$$

Recall $\mathcal{A}f(x)=\mathcal{L}^\alpha f(x)-\frac{1}{\alpha}x\cdot\nabla f(x)$. By (1.4), using the dominated convergence theorem, we get that

$$\nabla f(x)=-\int_0^\infty\nabla Q_th(x)\,dt.$$

Furthermore, we have by (2.11)

$$\big|\nabla_xp_{1-e^{-t}}(y-e^{-\frac{t}{\alpha}}x)\big|=\big|e^{-\frac{t}{\alpha}}\nabla_yp_{1-e^{-t},\beta}(y-e^{-\frac{t}{\alpha}}x)\big|\leqslant\frac{C_{\alpha,d}(e^t-1)^{-1/\alpha}}{((1-e^{-t})^{1/\alpha}+|y-e^{-\frac{t}{\alpha}}x|)^{\alpha+d}},$$

then for $x\in\mathbb{R}$, $z\in\mathbb{R}$ such that $|z|\leqslant\frac{1}{2}(e^t-1)^{\frac{1}{\alpha}}$,

$$\Big|p_{1-e^{-t}}\big(y-e^{-\frac{t}{\alpha}}(x+z)\big)-p_{1-e^{-t}}(y-e^{-\frac{t}{\alpha}}x)\Big|\leqslant|z|\frac{C_{\alpha,d}2^{d+\alpha}(e^t-1)^{-1/\alpha}}{((1-e^{-t})^{1/\alpha}+|y-e^{-\frac{t}{\alpha}}x|)^{\alpha+d}}.$$

Hence, by dominated convergence theorem and integration by parts, we have

$$\begin{aligned}\partial_xQ_t\big(h(x)-\mu(h)\big)&=\int_{\mathbb{R}^d}\nabla_xp_{1-e^{-t}}(y-e^{-\frac{t}{\alpha}}x)\big(h(y)-\mu(h)\big)dy\\&=-e^{-\frac{t}{\alpha}}\int_{\mathbb{R}^d}\nabla_yp_{1-e^{-t}}(y-e^{-\frac{t}{\alpha}}x)\big(h(y)-\mu(h)\big)dy\\&=e^{-\frac{t}{\alpha}}\int_{\mathbb{R}^d}p_{1-e^{-t}}(y-e^{-\frac{t}{\alpha}}x)\nabla h(y)dy,\end{aligned}$$

and similarly,

$$\nabla^2Q_t\big(h(x)-\mu(h)\big)=-e^{-\frac{2t}{\alpha}}\int_{\mathbb{R}^d}\nabla_yp_{(1-e^{-t})^{\frac{1}{\alpha}},\beta}(y-e^{-\frac{t}{\alpha}}x)\cdot\nabla h(y)^T)dy,$$

these imply

$$\begin{aligned}\big|\mathcal{L}^\alpha f(x)\big|\leqslant&d_\alpha\int_{\mathbb{S}^{d-1}}\int_0^\infty\int_0^\infty\frac{\big|Q_th(x+r\theta)-Q_th(x)-r\theta\cdot\nabla Q_th(x)\big|}{r^{\alpha+1}}dtdr\nu(d\theta)\\\leqslant&d_\alpha\int_{\mathbb{S}^{d-1}}\int_0^1\int_0^\infty\int_0^1\int_0^1\frac{sr^2\big|\nabla^2Q_th(x+sur\theta)\big|}{r^{\alpha+1}}dudsdtdr\nu(d\theta)\\&+d_\alpha\int_{\mathbb{S}^{d-1}}\int_1^\infty\int_0^\infty\int_0^1\frac{r\big|\nabla Q_th(x+sr\theta)-\nabla Q_th(x)\big|}{r^{\alpha+1}}dsdtdr\nu(d\theta)\leqslant C_{\alpha,d}.\end{aligned}$$

Thus, by Fubini's theorem, we have

$$\mathcal{L}^\alpha f(x)=-\int_0^\infty\mathcal{L}^\alpha Q_th(x)dt.$$

Hence, according to Lemma A.1, we can obtain

$$\mathcal{A}f=-\int_0^\infty\mathcal{A}Q_thdt=-\int_0^\infty\partial_tQ_thdt=Q_0h-Q_\infty h,$$

here $Q_\infty = \mu$, the unique invariant distribution of the semigroup $(Q_t)_{t\geq 0}$ associated with $\mathcal{A}$ by [18, Cor. 17.9]. The proof is complete.

Acknowledgements: I. Nourdin is supported by the FNR OPEN grant APOGee at Luxembourg University. L. Xu is supported by Macao S.A.R. FDCT 038/2017/A1, NNSFC 11571390, University of Macau MYRG (2016-00025-FST, 2018-00133-FST). X. Yang is supported by the FNR Grant MISSILe (R-AGR-3410-12-Z) at Luxembourg and Singapore Universities.

References

[1] D. Applebaum (2004): Lévy processes and stochastic calculus. Cambridge Studies inAdvanced Mathematics, 93. Cambridge University Press, Cambridge, xxiv+384 pp.

[2] B. Arras, C. Houdré (2019): On Stein's Method for Infinitely Divisible Laws with Finite First Moment. Springer International Publishing.

[3] B. Arras, C. Houdré (2019): On Stein's Method for Multivariate Self-Decomposable Laws. Preprint, arXiv:1907.10050.

[4] A. D. Barbour (1988): Stein's method and Poisson process convergence. A celebration of applied probability. J. Appl. Probab. Special Vol. 25A, 175–184.

[5] K. Bogdan, P. Sztonyk (2007): *Estimates of the potential kernel and Harnack's inequality for the anisotropic fractional Laplacian. Studia Math.* **181**(2), pp. 101–123.

[6] K. Bogdan, P. Sztonyk and V. Knopova (2017): Heat kernel of anisotropic nonlocal operators. Preprint, arXiv:1704.03705.

[7] A. H. Cook (1973): *Foundations of Modern Potential Theory Physics Bulletin.* **24**(11), pp. 681.

[8] P. Chen, I. Nourdin and L. Xu (2018): Stein's method for asymmetric α-stable distributions, with application to the stable CLT. Preprint, arXiv:1808.02405.

[9] P. Chen, I. Nourdin, L. Xu, X. Yang and R. Zhang (2019): Non-integrable stable approximation by Stein's method. Preprint, arXiv:1903.12315.

[10] P. Chen, L. Xu (2019): *Lindeberg's method for $\alpha-$stable central limit theorems.* Accepted by *Journal of Mathematical Analysis and Applications*, doi.org/10.1016/j.jmaa.2019.07.028.

[11] Z. Q. Chen, X. Zhang (2016): *Heat kernels and analyticity of non-symmetric jump diffusion semigroups. Probab. Theory Related Fields.* , **165**, pp. 267-312.

[12] Z. Q. Chen, X.C. Zhang (2018): Heat kernels for time-dependent non-symmetric stable-like operators. *Journal of Mathematical Analysis and Applications.* , **465**, pp. 1-21.

[13] R. Durrett (2010): Probability: theory and examples. Fourth edition. CambridgeSeries in Statistical and Probabilistic Mathematics, 31. Cambridge University Press, Cambridge, x+428 pp

[14] Y. Davydov, A. V. Nagaev (2002): *On Two Aproaches to Approximation of Multidimensional Stable Laws. Journal of Multivariate Analysis.* [J]. **82**(1), pp. 210-239.

[15] Falconer, Kenneth. Fractal geometry. Mathematical foundations and applications. Third edition. John Wiley & Sons, Ltd., Chichester, 2014. xxx+368 pp.

[16] J. R. Partington (2004): Linear operators and linear systems. London Mathematical Society Student Texts 60, Cambridge University Press.

[17] C. Stein (1986): Approximate computation of expectations. Institute of Mathematical Statistics Lecture Notes - Monograph Series, 7. Institute of Mathematical Statistics, Hayward, CA, iv+164 pp.

[18] K.I. Sato (1999): *Lévy processes and infinitely divisible distributions.* [M]. Cambridge University Press.

[19] A. Yuozulinas, V. Paulauskas (1998): *Some remarks on the rate of convergence to stable laws. Lithuanian Mathematical Journal.* , **38**(4), pp. 335-347.

[20] Watanabe, Toshiro. Asymptotic estimates of multi-dimensional stable densities and their applications. Trans. Amer. Math. Soc. 359 (2007), no. 6, 2851–2879.

[21] L. Xu (2019): *Approximation of stable law in Wasserstein-1 distance by Stein's method. Ann. Appl. Probab.* , **29**(1), pp. 458–504.

1. Department of Mathematics, Faculty of Science and Technology, University of Macau, Av. Padre Tomás Pereira, Taipa Macau, China; 2. UM Zhuhai Research Institute, Zhuhai, China.
E-mail address: yb77430@connect.umac.mo

Université du Luxembourg, Unité de Recherche en Mathématiques, Maison du Nombre, 6, avenue de la Fonte, L-4364 Esch-sur-Alzette, Grand Duchy of Luxembourg
E-mail address: ivan.nourdin@uni.lu

1. Department of Mathematics, Faculty of Science and Technology University of Macau Av. Padre Tomás Pereira, Taipa Macau, China; 2. UM Zhuhai Research Institute, Zhuhai, China.
E-mail address: lihuxu@um.edu.mo

Université du Luxembourg, Unité de Recherche en Mathématiques, Maison du Nombre, 6, avenue de la Fonte, L-4364 Esch-sur-Alzette, Grand Duchy of Luxembourg
E-mail address: xiaochuan.yang@uni.lu