Electronic Journal of Probability

Electron. J. Probab. **24** (2019), no. 29, 1–33.
ISSN: 1083-6489 https://doi.org/10.1214/19-EJP285

On Stein's method for multivariate self-decomposable laws with finite first moment

Benjamin Arras* Christian Houdré†

Abstract

We develop a multidimensional Stein methodology for non-degenerate self-decomposable random vectors in $\mathbb{R}^d$ having finite first moment. Building on previous univariate findings, we solve an integro-partial differential Stein equation by a mixture of semigroup and Fourier analytic methods. Then, under a second moment assumption, we introduce a notion of Stein kernel and an associated Stein discrepancy specifically designed for infinitely divisible distributions. Combining these new tools, we obtain quantitative bounds on smooth-Wasserstein distances between a probability measure in $\mathbb{R}^d$ and a non-degenerate self-decomposable target law with finite second moment. Finally, under an appropriate Poincaré-type inequality assumption, we investigate, via variational methods, the existence of Stein kernels. In particular, this leads to quantitative versions of classical results on characterizations of probability distributions by variational functionals.

Keywords: infinite divisibility; self-decomposability; Stein's method; Stein's kernel; weak limit theorems; rates of convergence; smooth Wassertein distance.
AMS MSC 2010: 60E07; 60E10; 60F05.
Submitted to EJP on October 4, 2018, final version accepted on March 2, 2019.

1 Introduction

Stein's method is a powerful device to quantify proximity in law between random variables. It has proven to be particularly useful to compute explicit rates of convergence in several limiting results of probability theory (from the standard central limit theorem to more complex paradigms satisfying some specific asymptotic behavior). Moreover, it has been successfully implemented for a large collection of one dimensional target limiting laws (see [45, 46, 13, 43] for standard references on the subject and [27] for a more recent survey). Most of these works essentially focus on the unidimensional setting

*Univ. Lille, CNRS, UMR 8524 - Laboratoire Paul Painlevé, F-59000 Lille, France. E-mail: arrasbenjamin@gmail.com

†Georgia Institute of Technology, School of Mathematics, Atlanta, GA 30332-0160; Research supported in part by the grants # 524678 from the Simons Foundation.
E-mail: houdre@math.gatech.edu

and related multidimensional results are relatively sparse in the literature. Indeed, the multidimensional Stein's method has essentially been developed in the multivariate normal case (see e.g. [2, 22, 20, 41, 39, 11, 40, 31, 34, 42, 35]) and for invariant measures of multidimensional diffusions ([29, 21]). In particular, [21] proposes a general Stein's method framework for target probability measures μ on $\mathbb{R}^d$, $d \geq 1$, which satisfy the following set of assumptions: μ has finite mean, is absolutely continuous with respect to the d-dimensional Lebesgue measure and its density is continuously differentiable with support the whole of $\mathbb{R}^d$.

Below, we introduce and develop a multidimensional Stein's methodology for a different specific class of probability measures on $\mathbb{R}^d$, namely non-degenerate self-decomposable laws with finite first moment (see (2.4) in Section 2 for a definition). This class of probability measures, introduced by Paul Lévy in [26], is rather natural in the context of limit theorems for sum of independent summands and has been thoroughly studied in several classical books (see e.g. [24, 26, 19, 28, 38, 44]). Nevertheless, while being very classical in the context of limit theorems, no systematic Stein's method has been carried out for multivariate non-degenerate self-decomposable distributions. (The whole class of non-degenerate self-decomposable laws with finite first moment is different, but has a non-empty intersection with the class of target probability measures analyzed in [21]. Indeed, non-degenerate self-decomposable laws with finite first moment admit a Lebesgue density, which might not be differentiable on $\mathbb{R}^d$, and whose support might be a half-space of $\mathbb{R}^d$.) Finally, many classical probability measures on $\mathbb{R}^d$ are self-decomposable (see [44, 47] and Section 3 below for some examples).

From our previous univariate work [1], the multidimensional Stein's method we implement is a generalization of the semigroup method "à la Barbour" ([2]). Thanks to the particular structure of self-decomposable characteristic functions, this semigroup approach relies heavily on Fourier analytic tools. Moreover, the generator of the aforementioned semigroup is an integro-differential operator, reflecting the infinite divisibility of the target law, which can be seen as a direct consequence of a characterization identity originating in [23] and further developed and analyzed in [1]. The resulting Stein equation is a non-local partial differential equation and contrasts with the usual second order partial differential equations associated with the multivariate Gaussian distribution or with the invariant measures of Itô diffusions.

Our methodology is then applied to quantify proximity, in smooth Wasserstein distances of order 1 and 2, between an appropriate probability measure on $\mathbb{R}^d$ and a non-degenerate self-decomposable laws with finite second moment. Key quantities used in our analysis are relevant versions of Stein kernels and of Stein discrepancies in this infinitely divisible setting (see Definition 4.1). Stein kernel and Stein discrepancy are concepts which have mostly been well developed in the Gaussian setting and have recently gained a certain momentum in connection with random matrices ([9]), Malliavin calculus ([32, 33]), functional inequalities ([25, 16]), optimal transport ([18]) and rates of convergence in multidimensional central limit theorems ([35]). In particular, [16] investigates the question of existence of a Gaussian Stein kernel for probability measures satisfying a Poincaré inequality or a converse weighted Poincaré inequality (see, e.g., [3] for a definition). Thanks to earlier work on characterizing functionals of infinitely divisible distributions [14], we introduce in the last section of the present manuscript the relevant variational setting which ensures the existence of Stein kernels and implies manageable upper bounds on the Stein discrepancy. In particular, Theorem 4.5 is a quantitative version of the characterizing results contained in [14].

Let us further describe the content of the present notes. In the next section, we introduce notations used throughout this work. In Section 3, we develop the multidimensional Stein methodology for non-degenerate self-decomposable random vector with finite first moment, extending our univariate approach ([1]). In Section 4, we introduce an infinitely divisible version of Stein kernel (and of the Stein discrepancy) and study the existence of the latter under an appropriate version of a Poincaré inequality. We end this section by providing quantitative upper bounds on the smooth Wasserstein distance of order two in terms of Poincaré constants and the second moment of the Lévy measure of the target self-decomposable distribution. A technical appendix finishes our manuscript.

2 Notations

Throughout, let $\|\cdot\|$ and $\langle\cdot;\cdot\rangle$ be respectively the Euclidean norm and the inner product on $\mathbb{R}^d$, $d \geq 1$. Let also $\mathcal{S}(\mathbb{R}^d)$ be the Schwartz space of infinitely differentiable rapidly decreasing real-valued functions defined on $\mathbb{R}^d$, and by $\mathcal{F}$ the Fourier transform operator given, for $f \in \mathcal{S}(\mathbb{R}^d)$, by

$$\mathcal{F}(f)(\xi) = \int_{\mathbb{R}^d} f(x) e^{-i\langle\xi;x\rangle} dx, \quad \xi \in \mathbb{R}^d.$$

On $\mathcal{S}(\mathbb{R}^d)$, the Fourier transform is an isomorphism and the following well known inversion formula holds

$$f(x) = \int_{\mathbb{R}^d} \mathcal{F}(f)(\xi) e^{+i\langle\xi;x\rangle} \frac{d\xi}{(2\pi)^d}, \quad x \in \mathbb{R}^d.$$

Let $\mathcal{C}_b(\mathbb{R}^d)$ be the space of bounded continuous functions on $\mathbb{R}^d$ endowed with the uniform norm $\|f\|_\infty = \sup_{x\in\mathbb{R}^d} |f(x)|$, for $f \in \mathcal{C}_b(\mathbb{R}^d)$. For any bounded linear operator, T, from a Banach space $(\mathcal{X}, \|\cdot\|_\mathcal{X})$ to another Banach space $(\mathcal{Y}, \|\cdot\|_\mathcal{Y})$ the operator norm is, as usual,

$$\|T\|_{op} = \sup_{f\in\mathcal{X},\, \|f\|_\mathcal{X}\neq 0} \frac{\|T(f)\|_\mathcal{Y}}{\|f\|_\mathcal{X}}. \tag{2.1}$$

More generally, for any r-multilinear form F from $(\mathbb{R}^d)^r$, $r \geq 1$, to $\mathbb{R}$, the operator norm of F is

$$\|F\|_{op} := \sup\left(|F(v_1, ..., v_r)| : \, v_j \in \mathbb{R}^d,\, \|v_j\| = 1,\, j = 1, ..., r\right). \tag{2.2}$$

Throughout, a Lévy measure is a positive Borel measure on $\mathbb{R}^d$ such that $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}^d}(1 \wedge \|u\|^2)\nu(du) < +\infty$. An $\mathbb{R}^d$-valued random vector X is infinitely divisible with triplet (b, Σ, ν) (written $X \sim ID(b, \Sigma, \nu)$), if its characteristic function φ writes, for all $\xi \in \mathbb{R}^d$, as

$$\varphi(\xi) = \exp\left(i\langle\xi;b\rangle - \frac{1}{2}\langle\Sigma\xi;\xi\rangle + \int_{\mathbb{R}^d}\left(e^{i\langle\xi;u\rangle} - 1 - i\langle\xi;u\rangle\mathbb{1}_D(u)\right)\nu(du)\right), \tag{2.3}$$

with $b \in \mathbb{R}^d$, Σ a symmetric nonnegative definite $d \times d$ matrix, ν a Lévy measure on $\mathbb{R}^d$ and D the closed Euclidean unit ball of $\mathbb{R}^d$. In the sequel, we are mainly interested in a subclass of infinitely divisible distributions, namely the self-decomposable laws (SD). If φ is the characteristic function of a self-decomposable distribution, then for any $\gamma \in (0, 1)$, there exists, on $\mathbb{R}^d$, a probability measure, say μ_γ, such that, for all $\xi \in \mathbb{R}^d$

$$\int_{\mathbb{R}^d} e^{i\langle\xi;u\rangle}\mu_\gamma(du) = \frac{\varphi(\xi)}{\varphi(\gamma\xi)}, \tag{2.4}$$

(Recall that $\varphi(\xi) \neq 0$, for all $\xi \in \mathbb{R}^d$, e.g., see [44, Lemma 7.5]). Moreover, e.g., see [44, Theorem 15.10], the Lévy measure of a self-decomposable distribution is given, for any Borel set B in $\mathbb{R}^d \setminus \{0\}$, by

$$\nu(B) = \int_{S^{d-1}} \lambda(dx) \int_0^{+\infty} \mathbb{1}_B(rx) k_x(r) \frac{dr}{r}, \tag{2.5}$$

with λ a finite positive measure on the Euclidean unit sphere S^{d-1} and $k_x(r)$ a non-negative function (Lebesgue) measurable in $x \in S^{d-1}$ and decreasing in $r > 0$ (namely, $k_x(s) \leq k_x(r)$, for $0 < r \leq s$). Thanks to [44, Remark 15.12 (iii)], since $\nu \neq 0$, let us assume that $\lambda(S^{d-1}) = 1$, that $\int_0^{+\infty}(r^2 \wedge 1)k_x(r)dr/r$ is finite and independent of x and that $k_x(r)$ is right-continuous in $r > 0$. Finally, since they satisfy the divergence condition (see e.g. [44, Theorem 27.13]), non-degenerate self-decomposable laws on $\mathbb{R}^d$ are absolutely continuous with respect to the d-dimensional Lebesgue measure.

To end this section, let us introduce some natural distances between probability measures on $\mathbb{R}^d$. For $p \geq 1$, the Wasserstein-p distance between two probability measures μ_X and μ_Y with finite p-th moment is

$$W_p(\mu_X, \mu_Y) = \inf_{\Pi \in \Gamma(\mu_X, \mu_Y)} \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} \|x - y\|^p d\Pi(x, y) \right)^{\frac{1}{p}}, \tag{2.6}$$

where $\Gamma(\mu_X, \mu_Y)$ is the collection of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with, respective, first and last d-dimensional marginal given by μ_X and μ_Y. By Hölder inequality, for $1 \leq p \leq q$,

$$W_p(\mu_X, \mu_Y) \leq W_q(\mu_X, \mu_Y), \tag{2.7}$$

while, by duality,

$$W_1(\mu_X, \mu_Y) = \sup_{\|h\|_{Lip} \leq 1} |\mathbb{E}h(X) - \mathbb{E}h(Y)|, \tag{2.8}$$

where $X \sim \mu_X$, $Y \sim \mu_Y$ and where Lip is the space of Lipschitz functions on $\mathbb{R}^d$ endowed with the seminorm

$$\|h\|_{Lip} = \sup_{x \neq y \in \mathbb{R}^d} \frac{|h(x) - h(y)|}{\|x - y\|}. \tag{2.9}$$

Let $\mathbb{N}^d$ be the space of multi-indices of dimension d. For any $\alpha \in \mathbb{N}^d$, let $|\alpha| = \sum_{i=1}^d |\alpha_i|$ and let D^α be the corresponding partial derivatives operator defined on smooth enough functions f, by $D^\alpha(f)(x_1, ..., x_d) = \partial_{x_1}^{\alpha_1} ... \partial_{x_d}^{\alpha_d}(f)(x_1, ..., x_d)$, for all $(x_1, ..., x_d) \in \mathbb{R}^d$. Next, for any r-times continuously differentiable function, h, on $\mathbb{R}^d$, viewing its ℓth-derivative $\mathbf{D}^\ell(h)$ as a ℓ-multilinear form, $1 \leq \ell \leq r$, let

$$M_\ell(h) := \sup_{x \in \mathbb{R}^d} \|\mathbf{D}^\ell(h)(x)\|_{op} = \sup_{x \neq y} \frac{\|\mathbf{D}^{\ell-1}(h)(x) - \mathbf{D}^{\ell-1}(h)(y)\|_{op}}{\|x - y\|}. \tag{2.10}$$

For $r \geq 0$, let $\mathcal{H}_r$ be the space of bounded continuous functions defined on $\mathbb{R}^d$ which are continuously differentiable up to (and including) the order r and such that, for any such function f,

$$\max_{0 \leq \ell \leq r} M_\ell(f) \leq 1, \tag{2.11}$$

with $M_0(f) := \sup_{x \in \mathbb{R}^d} |f(x)|$. In particular, for $f \in \mathcal{H}_r$,

$$\max_{\alpha \in \mathbb{N}^d,\, 0 \leq |\alpha| \leq r} \|D^\alpha(f)\|_\infty \leq 1. \tag{2.12}$$

Therefore, the space $\mathcal{H}_r$ is a subspace of the set of bounded functions which are r-times continuously differentiable on $\mathbb{R}^d$ such that $\|D^\alpha(f)\|_\infty \leq 1$, for all $\alpha \in \mathbb{N}^d$ with $0 \leq |\alpha| \leq r$. Then, the smooth Wasserstein distance of order r between two random vectors X and Y with respective law μ_X and μ_Y is defined by

$$d_{W_r}(\mu_X, \mu_Y) = \sup_{h \in \mathcal{H}_r} |\mathbb{E}h(X) - \mathbb{E}h(Y)|. \tag{2.13}$$

Moreover, the smooth Wasserstein distance of order $r \geq 1$ admits the following representation (see Lemma A.2 of the Appendix)

$$d_{W_r}(\mu_X, \mu_Y) = \sup_{h \in \mathcal{H}_r \cap \mathcal{C}_c^\infty(\mathbb{R}^d)} |\mathbb{E}h(X) - \mathbb{E}h(Y)|, \tag{2.14}$$

where $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is the space of infinitely differentiable compactly supported functions on $\mathbb{R}^d$. In particular, for $p \geq 1$ and $r \geq 1$

$$d_{W_r}(\mu_X, \mu_Y) \leq d_{W_1}(\mu_X, \mu_Y) \leq W_1(\mu_X, \mu_Y) \leq W_p(\mu_X, \mu_Y). \tag{2.15}$$

Finally, and as usual, for two probability measures, μ_1 and μ_2, on $\mathbb{R}^d$, μ_1 is said to be absolutely continuous with respect to μ_2, written $\mu_1 << \mu_2$, if for any Borel set, B, such that $\mu_2(B) = 0$, then $\mu_1(B) = 0$.

3 Stein's equation for SD laws via semigroup methods

Let X be a non-degenerate self-decomposable random vector in $\mathbb{R}^d$, without Gaussian component, and with law μ_X. By non-degenerate, we mean that the support of the law of X is not contained in some $d-1$ dimensional subspace of $\mathbb{R}^d$. Denote by X_i, for $i = 1, ..., d$, its coordinates and assume that $\mathbb{E}|X_i| < \infty$, for all $i = 1, ..., d$. Its characteristic function φ is given, for all $\xi \in \mathbb{R}^d$, by

$$\begin{aligned} \varphi(\xi) &= \exp\left(i\langle \xi; \mathbb{E}X\rangle + \int_{\mathbb{R}^d} \left(e^{i\langle \xi; u\rangle} - 1 - i\langle u; \xi\rangle\right)\nu(du)\right) \\ &= \exp\left(i\langle \xi; \mathbb{E}X\rangle + \int_{S^{d-1}\times(0,+\infty)} \left(e^{i\langle \xi; rx\rangle} - 1 - i\langle rx; \xi\rangle\right)\frac{k_x(r)}{r}dr\lambda(dx)\right), \end{aligned} \tag{3.1}$$

where ν is the Lévy measure, while k_x and λ are given in (2.5). Further, assume that, for any $0 < a < b < +\infty$ the functions $k_x(\cdot)$ are such that

$$\sup_{x \in S^{d-1}} \sup_{r \in (a,b)} k_x(r) < +\infty. \tag{3.2}$$

Since the function $k_x(\cdot)$ is a decreasing function in $r > 0$, the previous condition boils down to,

$$\sup_{x \in S^{d-1}} k_x(a^+) < +\infty, \quad a > 0,$$

where $k_x(a^+) = \lim_{r\to a^+} k_x(r)$, $x \in S^{d-1}$. In (3.2), the supremum over x in S^{d-1} has to be understood as the λ-essential supremum of the function $k_x(r)$ in the x variable. In the univariate case, $d = 1$, the polar decomposition of the Lévy measure ν boils down to $\nu(du) = k(u)du/|u|$ where k is non-negative, increasing on $(-\infty, 0)$ and decreasing on $(0, +\infty)$. Thus, the condition (3.2) is automatically satisfied for $d = 1$. For $d \geq 2$, the polar decomposition of the Lévy measure associated with a stable distribution of index $\alpha \in (1, 2)$ is given by

$$\nu(du) = \mathbb{1}_{(0,+\infty)}(r)\mathbb{1}_{S^{d-1}}(x)\frac{dr}{r^{\alpha+1}}\lambda(dx),$$

for some finite positive measure λ on the d-dimensional unit sphere (see [44, Theorem 14.3]). Then, $k_x(r) = 1/r^\alpha$, for all $r > 0$, and condition (3.2) is automatically satisfied (see below for more examples). Next, define a family of operators $(P_t^\nu)_{t\geq 0}$, for all $f \in \mathcal{S}(\mathbb{R}^d)$, all $x \in \mathbb{R}^d$ and all $t \geq 0$, via

$$P_t^\nu(f)(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \mathcal{F}(f)(\xi)e^{ie^{-t}\langle x;\xi\rangle}\frac{\varphi(\xi)}{\varphi(e^{-t}\xi)}d\xi. \tag{3.3}$$

Denoting by $(\mu_t)_{t\geq 0}$ the family of probability measures given by (2.4) with $\gamma = e^{-t}$ and using Fourier inversion on $\mathcal{S}(\mathbb{R}^d)$, it follows that

$$P_t^\nu(f)(x) = \int_{\mathbb{R}^d} f(u + e^{-t}x)\mu_t(du). \tag{3.4}$$

For all $t \geq 0$, the probability measure μ_t is infinitely divisible with finite first moment and its characteristic function φ_t admits, for all $\xi \in \mathbb{R}^d$, the following representation

$$\begin{aligned}\varphi_t(\xi) = \exp\Bigg(i\langle \xi; \mathbb{E}X\rangle(1-e^{-t}) + \int_{S^{d-1}\times(0,+\infty)} &\left(e^{i\langle \xi; rx\rangle} - 1 - i\langle rx; \xi\rangle\right) \\ &\times \frac{k_x(r) - k_x(e^t r)}{r} dr\lambda(dx)\Bigg).\end{aligned} \tag{3.5}$$

The next lemma asserts that the family of operators $(P_t^\nu)_{t\geq 0}$ is a C_0-semigroup on the space $L^1(\mu_X)$. Its proof is very similar to the one dimensional case (see [1, Proposition 5.1]) thanks to the polar decomposition (2.5).

Lemma 3.1. *Let X be a non-degenerate self-decomposable random vector without Gaussian component, with law μ_X, Lévy measure ν and such that $\mathbb{E}\|X\| < \infty$. Moreover, let the functions k_x given by* (2.5) *satisfy* (3.2)*. Let φ be its characteristic function and let $(P_t^\nu)_{t\geq 0}$ be the family of operators defined by* (3.3)*. Then, $(P_t^\nu)_{t\geq 0}$ is a C_0-semigroup on the space $L^1(\mu_X)$ and its generator $\mathcal{A}$ is defined, for all $f \in \mathcal{S}(\mathbb{R}^d)$ and for all $x \in \mathbb{R}^d$, by*

$$\mathcal{A}(f)(x) = \langle \mathbb{E}X - x; \nabla(f)(x)\rangle + \int_{\mathbb{R}^d} \langle \nabla(f)(x+u) - \nabla(f)(x); u\rangle \nu(du). \tag{3.6}$$

Proof of Lemma 3.1. Let $f \in \mathcal{C}_b(\mathbb{R}^d)$. First, by (3.4), for all $s, t \geq 0$ and for all $x \in \mathbb{R}^d$,

$$P_{s+t}^\nu(f)(x) = \int_{\mathbb{R}^d} f(u + e^{-(s+t)}x)\mu_{t+s}(du).$$

Moreover, for all $s, t \geq 0$,

$$\begin{aligned}P_t^\nu \circ P_s^\nu(f)(x) &= \int_{\mathbb{R}^d} P_s^\nu(f)(u + e^{-t}x)\mu_t(du) \\ &= \int_{\mathbb{R}^d\times\mathbb{R}^d} f(v + e^{-s}(u + e^{-t}x))\mu_t(du)\mu_s(dv).\end{aligned} \tag{3.7}$$

Let $\psi_{s,t,x}$ be the measurable function defined by $\psi_{s,t,x}(u,v) = v + e^{-s}(u + e^{-t}x)$, for all $u, v \in \mathbb{R}^d \times \mathbb{R}^d$. Then, from (3.7), for all $s, t \geq 0$ and for all $x \in \mathbb{R}^d$,

$$P_t^\nu \circ P_s^\nu(f)(x) = \int_{\mathbb{R}^d\times\mathbb{R}^d} f(w)(\mu_t \otimes \mu_s) \circ \psi_{s,t,x}^{-1}(dw).$$

Let us now compute the characteristic function of the probability measure $(\mu_t \otimes \mu_s) \circ \psi_{s,t,x}^{-1}$. For all $\xi \in \mathbb{R}^d$,

$$\begin{aligned}\int_{\mathbb{R}^d} e^{i\langle \xi; w\rangle}(\mu_t \otimes \mu_s) \circ \psi_{s,t,x}^{-1}(dw) &= \int_{\mathbb{R}^d\times\mathbb{R}^d} e^{i\langle \xi; \psi_{s,t,x}(u,v)\rangle}\mu_t(du) \otimes \mu_s(dv) \\ &= e^{i\langle \xi; e^{-(s+t)}x\rangle} \int_{\mathbb{R}^d} e^{i\langle \xi; v\rangle} e^{i\langle \xi; e^{-s}u\rangle}\mu_t(du) \otimes \mu_s(dv) \\ &= e^{i\langle \xi; e^{-(s+t)}x\rangle}\varphi_s(\xi)\varphi_t(e^{-s}\xi) \\ &= e^{i\langle \xi; e^{-(s+t)}x\rangle}\frac{\varphi(\xi)}{\varphi(e^{-(s+t)}\xi)}\end{aligned}$$

$$= e^{i\langle \xi; e^{-(s+t)}x\rangle} \int_{\mathbb{R}^d} e^{i\langle \xi; u\rangle} \mu_{s+t}(du)$$
$$= \int_{\mathbb{R}^d} e^{i\langle \xi; \varphi_{s,t,x}(u)\rangle} \mu_{s+t}(du)$$
$$= \int_{\mathbb{R}^d} e^{i\langle \xi; u\rangle} \mu_{s+t} \circ \varphi_{s,t,x}^{-1}(du),$$

where $\varphi_{s,t,x}(u) = e^{-(s+t)}x + u$, for all $x \in \mathbb{R}^d$, $s \geq 0$ and $t \geq 0$. This implies that

$$P_t^\nu \circ P_s^\nu(f)(x) = P_{s+t}^\nu(f)(x),$$

and so the semigroup property is verified on $\mathcal{C}_b(\mathbb{R}^d)$. Moreover,

$$\int_{\mathbb{R}^d} P_t^\nu(f)(x)\mu_X(dx) = \int_{\mathbb{R}^d \times \mathbb{R}^d} f(u + e^{-t}x)\mu_t(du)\mu_X(dx). \tag{3.8}$$

Setting $\omega_t(u,x) = u + e^{-t}x$, for all $u \in \mathbb{R}^d$ and all $x \in \mathbb{R}^d$, (3.8) then becomes

$$\int_{\mathbb{R}^d} P_t^\nu(f)(x)\mu_X(dx) = \int_{\mathbb{R}^d} f(v)(\mu_t \otimes \mu_X) \circ \omega_t^{-1}(dv).$$

The characteristic function of the probability measure $(d\mu_t \otimes d\mu_X) \circ \omega_t^{-1}$ is, for all $\xi \in \mathbb{R}^d$,

$$\begin{aligned}
\int_{\mathbb{R}^d} e^{i\langle \xi; v\rangle} (\mu_t \otimes \mu_X) \circ \omega_t^{-1}(dv) &= \int_{\mathbb{R}^d \times \mathbb{R}^d} e^{i\langle \xi; \omega_t(u,x)\rangle} \mu_t(du)\mu_X(dx) \\
&= \int_{\mathbb{R}^d \times \mathbb{R}^d} e^{i\langle \xi; u + e^{-t}x\rangle} \mu_t(du)\mu_X(dx) \\
&= \varphi_t(\xi)\varphi(e^{-t}\xi) \\
&= \varphi(\xi).
\end{aligned}$$

Therefore, for all $t \in \mathbb{R}^d$,

$$\int_{\mathbb{R}^d} P_t^\nu(f)(x)\mu_X(dx) = \int_{\mathbb{R}^d} f(x)\mu_X(dx),$$

and so the probability measure μ_X is invariant for the family of operators $(P_t^\nu)_{t\geq 0}$. One can further check, by Fourier arguments, that, for all $x \in \mathbb{R}^d$,

$$\lim_{t\to 0^+} P_t^\nu(f)(x) = f(x), \qquad \lim_{t\to +\infty} P_t^\nu(f)(x) = \int_{\mathbb{R}^d} f(x)\mu_X(dx).$$

Next, by Jensen inequality and the invariance property,

$$\begin{aligned}
\int_{\mathbb{R}^d} |P_t^\nu(f)(x)|\mu_X(dx) &\leq \int_{\mathbb{R}^d} P_t^\nu(|f|)(x)\mu_X(dx) \\
&\leq \int_{\mathbb{R}^d} |f(x)|\mu_X(dx).
\end{aligned}$$

Then, by the density of $\mathcal{C}_b(\mathbb{R}^d)$ in $L^1(\mu_X)$ ([4, Corollary 4.2.2]), one can extend the family of operators $(P_t^\nu)_{t\geq 0}$ to functions in $L^1(\mu_X)$, and, this extension, still denoted by $(P_t^\nu)_{t\geq 0}$, is a C_0-semigroup on $L^1(\mu_X)$. To end the proof of the lemma, let us compute the generator of this semigroup on $\mathcal{S}(\mathbb{R}^d)$. Let $f \in \mathcal{S}(\mathbb{R}^d)$. By Fourier inversion, for all $x \in \mathbb{R}^d$ and for all $t > 0$,

$$\frac{1}{t}\left(P_t^\nu(f)(x) - f(x)\right) = \frac{1}{(2\pi)^d t} \int_{\mathbb{R}^d} \mathcal{F}(f)(\xi) e^{i\langle \xi; x\rangle} \left(e^{i\langle \xi; x\rangle (e^{-t}-1)} \frac{\varphi(\xi)}{\varphi(e^{-t}\xi)} - 1 \right) d\xi.$$

First, for all $x \in \mathbb{R}^d$ and for all $\xi \in \mathbb{R}^d$,

$$\lim_{t\to 0^+} \frac{1}{t}\left(e^{i\langle \xi;x\rangle(e^{-t}-1)}\frac{\varphi(\xi)}{\varphi(e^{-t}\xi)}-1\right) = -i\langle \xi; x\rangle + \sum_{i=1}^d \xi_i \frac{\partial_i(\varphi)(\xi)}{\varphi(\xi)},$$

which is a well-defined limit since X has finite first moment. Now, from the representation (3.1), for all $\xi \in \mathbb{R}^d$,

$$\begin{aligned}\sum_{j=1}^d \xi_j \frac{\partial_j(\varphi)(\xi)}{\varphi(\xi)} &= \sum_{j=1}^d \xi_j \left(i\mathbb{E}X_j + i\int_{\mathbb{R}^d} u_j\left(e^{i\langle u;\xi\rangle}-1\right)\nu(du)\right)\\ &= i\left(\langle \xi; \mathbb{E}X\rangle + \int_{\mathbb{R}^d}\langle \xi;u\rangle\left(e^{i\langle u;\xi\rangle}-1\right)\nu(du)\right).\end{aligned}$$

Moreover, by Lemma A.1 (ii) of the Appendix, for all $t \in (0,1)$, $x \in \mathbb{R}^d$ and $\xi \in \mathbb{R}^d$,

$$\frac{1}{t}\left|e^{i\langle \xi;x\rangle(e^{-t}-1)}\frac{\varphi(\xi)}{\varphi(e^{-t}\xi)}-1\right| \le C_1\|\xi\|\|x\| + C_2(\|\xi\| + \|\xi\|\|\mathbb{E}X\| + \|\xi\|^2),$$

for some $C_1 > 0, C_2 > 0$, independent of t, ξ and x. Then, by dominated convergence,

$$\begin{aligned}\lim_{t\to 0^+}\frac{1}{t}\left(P_t^\nu(f)(x) - f(x)\right) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\mathcal{F}(f)(\xi)e^{i\langle\xi;x\rangle}\bigg(&-i\langle\xi;x\rangle + i\langle \xi;\mathbb{E}X\rangle\\ &+ i\int_{\mathbb{R}^d}\langle\xi;u\rangle\left(e^{i\langle u;\xi\rangle}-1\right)\nu(du)\bigg)d\xi,\end{aligned}$$

which is equal, by standard Fourier arguments, to

$$\lim_{t\to 0^+}\frac{1}{t}\left(P_t^\nu(f)(x) - f(x)\right) = \langle \nabla(f)(x); \mathbb{E}X - x\rangle + \int_{\mathbb{R}^d}\langle \nabla(f)(x+u) - \nabla(f)(x); u\rangle \nu(du).$$

This concludes the proof of (3.6) and of the lemma. □

Remark 3.2. The analysis performed in this paper relies heavily on the properties of the semigroup of operators $(P_t^\nu)_{t\ge 0}$. It is, moreover, possible to build a Markov process with values in $\mathbb{R}^d$ such that its induced semigroup corresponds to $(P_t^\nu)_{t\ge0}$. Indeed, for all $t \ge 0$, since P_t^ν is mass conservative and is a contraction on $L^1(\mu_X)$,[3, Proposition 1.2.3] ensures the existence of a kernel $p_t^\nu(\cdot,\cdot)$ such that, for all $f \in \mathcal{C}_b\left(\mathbb{R}^d\right)$,

$$P_t^\nu(f)(x) = \int_{\mathbb{R}^d} f(z)p_t^\nu(x,dz), \quad x \in \mathbb{R}^d.$$

In particular, $p_t^\nu(x,dz) = \mu_t \circ \varphi_{0,t,x}^{-1}(dz)$, for all $t \ge 0$ and $x \in \mathbb{R}^d$, where $\varphi_{0,t,x}(y) = e^{-t}x + y$, for all $y \in \mathbb{R}^d$, and where μ_t is given by (2.4) with $\gamma = e^{-t}$. Besides, the semigroup property of $(P_t^\nu)_{t\ge0}$ implies that the kernels $(p_t^\nu(\cdot,\cdot))_{t\ge0}$ satisfy the Chapman-Kolmogorov equation. In turn, this latter relation ensures the existence of a Markov process $\{Z_t^x : t \ge 0\}$, such that $Z_0^x = x$, with $x \in \mathbb{R}^d$ and such that, for all $f \in \mathcal{C}_b\left(\mathbb{R}^d\right)$,

$$\mathbb{E}f(Z_t^x) = P_t^\nu(f)(x), \quad x \in \mathbb{R}^d, \quad t \ge 0.$$

Finally, for all $t \ge 0$, denoting by X_t the random vector with law μ_t, one has the representation in law

$$Z_t^x =_d e^{-t}x + X_t,$$

where $=_d$ stands for equality in distribution. Based on this representation, the Markov process $\{Z_t^x : t \ge 0\}$ can be thought of as a generalization of the classical Ornstein-Uhlenbeck process whose invariant measure is the multivariate Gaussian distribution on $\mathbb{R}^d$ (see e.g. [3, Part I, Chapter 2, Section 2.7.1]).

Let $h \in \mathcal{H}_r \cap \mathcal{C}_c^{\infty}(\mathbb{R}^d)$, for some $r \geq 1$. The aim of the rest of the section is to solve, for all $x \in \mathbb{R}^d$, the following integro-partial differential equation,

$$\langle \mathbb{E}X - x; \nabla(f)(x)\rangle + \int_{\mathbb{R}^d} \langle \nabla(f)(x+u) - \nabla(f)(x); u\rangle \nu(du) = h(x) - \mathbb{E}h(X), \tag{3.9}$$

which will serve as the fundamental equation in our Stein's methodology for non-degenerate self-decomposable law. As done in the one-dimensional case in [1], we first introduce a potential candidate solution for this equation then study its regularity and finally prove that it actually solves the equation (3.9). The following proposition deals with the existence and the regularity of the candidate solution.

Proposition 3.3. *Let X be a non-degenerate self-decomposable random vector in $\mathbb{R}^d$, without Gaussian component, with law μ_X, characteristic function φ and such that $\mathbb{E}\|X\| < \infty$. Moreover, let the functions k_x given by* (2.5) *satisfy* (3.2)*. Let $(P_t^{\nu})_{t\geq 0}$ be the semigroup of operators obtained in Lemma 3.1. Then, for any $h \in \mathcal{H}_2$, the function f_h given, for all $x \in \mathbb{R}^d$, by*

$$f_h(x) = -\int_0^{+\infty} (P_t^{\nu}(h)(x) - \mathbb{E}h(X))dt, \tag{3.10}$$

is well defined and twice continuously differentiable on $\mathbb{R}^d$. Moreover, for any $\alpha \in \mathbb{N}^d$ such that $|\alpha| = 1$,

$$\|D^{\alpha}(f_h)\|_{\infty} \leq 1, \tag{3.11}$$

and, for any $\alpha \in \mathbb{N}^d$ such that $|\alpha| = 2$,

$$\|D^{\alpha}(f_h)\|_{\infty} \leq \frac{1}{2}. \tag{3.12}$$

Proof of Proposition 3.3. Let $h \in \mathcal{H}_2$. By (3.4) and Theorem A.4 of the Appendix, for all $x \in \mathbb{R}^d$,

$$\begin{aligned} |P_t^{\nu}(h)(x) - \mathbb{E}h(X)| &\leq \|x\|e^{-t}M_1(h) + d_{W_2}(X, X_t) \\ &\leq \|x\|e^{-t} + C_d e^{-\frac{t}{2^{d+1}(d+1)}}, \end{aligned}$$

and so the function

$$f_h(x) = -\int_0^{+\infty} (P_t^{\nu}(h)(x) - \mathbb{E}h(X))dt,$$

is well defined for all $x \in \mathbb{R}^d$. Moreover, by (3.4) and the regularity of $h \in \mathcal{H}_2$, for all $1 \leq j \leq d$ and for all $x \in \mathbb{R}^d$,

$$\partial_j(f_h)(x) = -\int_0^{+\infty} e^{-t}(P_t^{\nu}(\partial_j(h))(x))dt,$$

which implies that, for all $1 \leq j \leq d$ and for all $x \in \mathbb{R}^d$,

$$|\partial_j(f_h)(x)| \leq 1.$$

Similarly for all $i, j \in \{1, .., d\}$ and for all $x \in \mathbb{R}^d$,

$$\partial_{ij}^2(f_h)(x) = -\int_0^{+\infty} e^{-2t}(P_t^{\nu}(\partial_{ij}^2(h))(x))dt,$$

which implies that

$$\left|\partial_{ij}^2(f_h)(x)\right| \leq \frac{1}{2},$$

and concludes the proof of the proposition. □

Proposition 3.4. *Let X be a non-degenerate self-decomposable random vector in $\mathbb{R}^d$ without Gaussian component, with law μ_X, characteristic function φ, and such that $\mathbb{E}\|X\| < \infty$. Moreover, let the function k_x given by* (2.5) *satisfy* (3.2)*, and let $(X_t)_{t\geq 0}$ be the collection of random vectors such that, for all $t \geq 0$, X_t has law μ_t given by* (2.4) *with $\gamma = e^{-t}$. For each $t > 0$, let μ_t be absolutely continuous with respect to the d-dimensional Lebesgue measure and let its Radon-Nikodym derivative, denoted by q_t, be continuously differentiable on $\mathbb{R}^d$ and such that, for all $1 \leq i \leq d$,*

$$\lim_{y_i \to \pm\infty} q_t(y) = 0, \quad \int_{\mathbb{R}^d} |\partial_i(q_t)(y)|\, dy < \infty, \quad \int_0^{+\infty} e^{-2t} \left(\int_{\mathbb{R}^d} |\partial_i(q_t)(y)|\, dy \right) dt < \infty. \tag{3.13}$$

Let $h \in \mathcal{H}_1$ and $(P_t^\nu)_{t\geq 0}$ be the semigroup of operators obtained in Lemma 3.1. Then, the function f_h given, for all $x \in \mathbb{R}^d$, by

$$f_h(x) = -\int_0^{+\infty} (P_t^\nu(h)(x) - \mathbb{E}h(X))dt, \tag{3.14}$$

is well defined and twice continuously differentiable on $\mathbb{R}^d$. Moreover, for any $\alpha \in \mathbb{N}^d$ such that $|\alpha| = 1$

$$\|D^\alpha(f_h)\|_\infty \leq 1, \tag{3.15}$$

and, for any $\alpha \in \mathbb{N}^d$ such that $|\alpha| = 2$

$$\|D^\alpha(f_h)\|_\infty \leq C_d. \tag{3.16}$$

for some $C_d > 0$ only depending on d.

Proof of Proposition 3.4. Let $h \in \mathcal{H}_1$. By (3.4) and Theorem A.4, for all $x \in \mathbb{R}^d$

$$\begin{aligned}|P_t^\nu(h)(x) - \mathbb{E}h(X)| &\leq \|x\|e^{-t}M_1(h) + d_{W_1}(X, X_t)\\ &\leq \|x\|e^{-t} + C_d e^{-\frac{t}{2^{d+1}(d+1)}},\end{aligned}$$

and so the function

$$f_h(x) = -\int_0^{+\infty} (P_t^\nu(h)(x) - \mathbb{E}h(X))dt,$$

is well defined for all $x \in \mathbb{R}^d$. Moreover, by (3.4) and the regularity of $h \in \mathcal{H}_1$, for all $1 \leq j \leq d$ and for all $x \in \mathbb{R}^d$,

$$\partial_j(f_h)(x) = -\int_0^{+\infty} e^{-t}(P_t^\nu(\partial_j(h))(x))dt,$$

which implies that, for all $1 \leq j \leq d$ and for all $x \in \mathbb{R}^d$,

$$|\partial_j(f_h)(x)| \leq 1.$$

Let us fix $1 \leq i \leq d$ and $x \in \mathbb{R}^d$. By definition and an integration by parts (thanks to (3.13)),

$$\begin{aligned}P_t^\nu(\partial_i(h))(x) &= \int_{\mathbb{R}^d} \partial_i(h)(xe^{-t} + y)q_t(y)dy\\ &= -\int_{\mathbb{R}^d} h(xe^{-t} + y)\partial_i(q_t)(y)dy.\end{aligned}$$

Thus, for all $1 \leq i, j \leq d$ and $x \in \mathbb{R}^d$,

$$\partial_j \left(P_t^\nu (\partial_i(h))(x) \right) = -e^{-t} \int_{\mathbb{R}^d} \partial_j(h)(xe^{-t} + y) \partial_i(q_t)(y) dy.$$

This representation, together with the third condition in (3.13), ensures that f_h is twice continuously differentiable on $\mathbb{R}^d$ and that, for all $x \in \mathbb{R}^d$ and for all $1 \leq i, j \leq d$,

$$\partial_{i,j}(f_h)(x) = \int_0^{+\infty} e^{-2t} \left(\int_{\mathbb{R}^d} \partial_j(h)(xe^{-t} + y) \partial_i(q_t)(y) dy \right) dt.$$

Finally, for all $x \in \mathbb{R}^d$ and all $1 \leq i, j \leq d$,

$$|\partial_{i,j}(f_h)(x)| \leq \int_0^{+\infty} e^{-2t} \left(\int_{\mathbb{R}^d} |\partial_i(q_t)(y)| \, dy \right) dt < +\infty.$$

This concludes the proof of the proposition. □

Before showing that f_h, as given in the two previous propositions, is a solution to the integro-partial differential equation

$$\langle \mathbb{E}X - x; \nabla(f)(x) \rangle + \int_{\mathbb{R}^d} \langle \nabla(f)(x+u) - \nabla(f)(x); u \rangle \nu(du) = h(x) - \mathbb{E}h(X), \qquad x \in \mathbb{R}^d,$$

for h respectively in $\mathcal{H}_2 \cap \mathcal{C}_c^\infty(\mathbb{R}^d)$ and $\mathcal{H}_1 \cap \mathcal{C}_c^\infty(\mathbb{R}^d)$, we provide some examples of non-degenerate self-decomposable random vectors satisfying the assumptions of the aforementioned propositions.

Some examples

Rotationally invariant α-stable random vector in $\mathbb{R}^d$. Let $\alpha \in (1,2)$ and let X be an α-stable random vector whose law is rotationally invariant. Then, its characteristic function φ is, for all $\xi \in \mathbb{R}^d$,

$$\varphi(\xi) = \exp\left(-C_{\alpha,d} \|\xi\|^\alpha\right),$$

for some constant $C_{\alpha,d} > 0$ depending on α and d. Hence, the characteristic function of μ_t is given, for all $\xi \in \mathbb{R}^d$, and $t \geq 0$ by

$$\varphi_t(\xi) = \exp\left(-C_{\alpha,d}(1 - e^{-\alpha t}) \|\xi\|^\alpha\right).$$

Thus, for all $t > 0$, μ_t is absolutely continuous with respect to the Lebesgue measure and its density q_t is given for all $x \in \mathbb{R}^d$, by the Fourier inversion formula,

$$q_t(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\langle \xi; x \rangle} \overline{\varphi_t(\xi)} d\xi = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\langle \xi; x \rangle} \exp\left(-C_{\alpha,d}(1 - e^{-\alpha t}) \|\xi\|^\alpha\right) d\xi. \quad (3.17)$$

From (3.17), it is clear that q_t is continuously differentiable on $\mathbb{R}^d$ and that, for all $x \in \mathbb{R}^d$ and $1 \leq j \leq d$,

$$\partial_j(q_t)(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\langle \xi; x \rangle} (i\xi_j) \exp\left(-C_{\alpha,d}(1 - e^{-\alpha t}) \|\xi\|^\alpha\right) d\xi.$$

Moreover, the characteristic function φ_t is linked to the Fourier transform of the transition density of a rotationally invariant d-dimensional α-stable process after the time change $\tau = (1 - e^{-\alpha t})$. Indeed, if $(Z_t^\alpha)_{t \geq 0}$ is a rotationally invariant d-dimensional α-stable

Lévy process, then its characteristic function at time t is given, for all $t \geq 0$ and all $\xi \in \mathbb{R}^d$, by

$$\mathbb{E}\, e^{i\langle \xi; Z_t^\alpha \rangle} = \exp\left(-\frac{t\|\xi\|^\alpha}{2^{\frac{\alpha}{2}}}\right).$$

Thus, for all $\xi \in \mathbb{R}^d$ and all $t \geq 0$,

$$\varphi_t(\xi) = \mathbb{E}\, e^{i\langle \xi; \sqrt{2} C_{\alpha,d}^{\frac{1}{\alpha}} Z^\alpha_{1-e^{-\alpha t}} \rangle}.$$

Finally, Lemma 2.2 of [15] implies that the density q_t and its gradient satisfy the following inequalities, for all $x \in \mathbb{R}^d$,

$$|q_t(x)| \leq C'_{\alpha,d} \frac{(1-e^{-\alpha t})}{\left((1-e^{-\alpha t})^{\frac{1}{\alpha}} + \|x\|\right)^{\alpha+d}}, \quad \|\nabla(q_t)(x)\| \leq C''_{\alpha,d} \frac{(1-e^{-\alpha t})}{\left((1-e^{-\alpha t})^{\frac{1}{\alpha}} + \|x\|\right)^{\alpha+d+1}},$$

for some $C'_{\alpha,d} > 0, C''_{\alpha,d} > 0$, only depending on α and d. It follows that q_t satisfies the conditions (3.13).

Symmetric α-stable random vector in $\mathbb{R}^d$. Let $\alpha \in (1,2)$ and let X be a symmetric α-stable random vector on $\mathbb{R}^d$. By [44, Theorem 14.13], the characteristic function of X is given, for all $\xi \in \mathbb{R}^d$ with $\|\xi\| \neq 0$, by

$$\varphi(\xi) = \exp\left(-\int_{S^{d-1}} |\langle x;\xi\rangle|^\alpha \lambda_1(dx)\right) = \exp\left(-\|\xi\|^\alpha \int_{S^{d-1}} \left|\left\langle x; \frac{\xi}{\|\xi\|}\right\rangle\right|^\alpha \lambda_1(dx)\right),$$

where λ_1 is a symmetric positive finite measure on S^{d-1}. Then, for all $\xi \in \mathbb{R}^d$ and all $t \geq 0$,

$$\varphi_t(\xi) = \exp\left(-(1-e^{-\alpha t})\|\xi\|^\alpha \int_{S^{d-1}} \left|\left\langle x; \frac{\xi}{\|\xi\|}\right\rangle\right|^\alpha \lambda_1(dx)\right).$$

Moreover, let us assume that there exists $c_0 > 0$ such that $\int_{S^{d-1}} |\langle x;u\rangle|^\alpha \lambda_1(dx) \geq c_0$, for any $u \in S^{d-1}$. Then, μ_t is absolutely continuous with respect to the d-dimensional Lebesgue measure and its density q_t is given, for all $t > 0$ and all $x \in \mathbb{R}^d$, by

$$q_t(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\langle \xi; x\rangle} \overline{\varphi_t(\xi)} d\xi = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\langle \xi; x\rangle} \exp\left(-(1-e^{-\alpha t})\|\xi\|^\alpha \eta_\alpha(\xi)\right) d\xi,$$

with $\eta_\alpha(\xi) = \int_{S^{d-1}} |\langle x; \xi/\|\xi\|\rangle|^\alpha \lambda_1(dx)$. For all $t > 0$, q_t is continuously differentiable on $\mathbb{R}^d$ and its partial derivative in the jth direction, $j \in \{1, ..., d\}$, is given, for all $x \in \mathbb{R}^d$, by

$$\partial_j(q_t)(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\langle \xi; x\rangle} (i\xi_j) \exp\left(-(1-e^{-\alpha t})\|\xi\|^\alpha \eta_\alpha(\xi)\right) d\xi.$$

Takano distribution I [49]. Let $\alpha \in (0, +\infty)$ and let $\mu_{\alpha,d}$ be the probability measure on $\mathbb{R}^d$ given by

$$\mu_{\alpha,d}(dx) = c_{\alpha,d}\left(1 + \|x\|^2\right)^{-\alpha-d/2} dx,$$

where $c_{\alpha,d} > 0$ is a normalizing constant. As shown in [49] such a probability measure is self-decomposable. Moreover, from ([49, Theorem II]), its characteristic function is given, for all $\xi \in \mathbb{R}^d$, by

$$\varphi(\xi) = \exp\left(\int_0^{+\infty} 2(2\pi)^{-d/2} v^{d/2} \left(\int_0^{+\infty} (\sqrt{2w})^{(d+2)/2} K_{(d-2)/2}(\sqrt{2w}v) g_\alpha(2w) dw\right) dv\right.$$

$$\times \int_{S^{d-1}} dx \int_0^v \left(e^{iux\xi} - 1 - \frac{iux\xi}{1+u^2} \right) \frac{du}{u} \bigg),$$

where $K_{(d-2)/2}$ denotes the modified Bessel function of order $(d-2)/2$ while $g_\alpha(w) = 2/(\pi^2 w) \times 1/\left(J_\alpha^2(\sqrt{w}) + Y_\alpha^2(\sqrt{w})\right)$, $w > 0$, with J_α and Y_α the Bessel functions of the first kind and of the second kind, respectively (see [36, Chapter 10] for definitions of Bessel functions). Finally, the Lévy measure of $\mu_{\alpha,d}$ is given (see [49, Representation (2) page 23]) by:

$$\nu(du) = \frac{2}{\|u\|^d} \left(\int_0^{+\infty} g_\alpha(2w) L_{d/2}\left(\sqrt{2w}\|u\|\right) dw \right) du,$$

where $L_{d/2}(v) = (2\pi)^{-d/2} v^{d/2} K_{d/2}(v)$, for all $v > 0$, implying the following polar decomposition

$$\nu(du) = \mathbb{1}_{(0,+\infty)}(r) \mathbb{1}_{S^{d-1}}(x) \frac{2}{r} \left(\int_0^{+\infty} g_\alpha(2w) L_{d/2}\left(\sqrt{2w}r\right) dw \right) dr\sigma(dx),$$

where σ is the uniform measure on S^{d-1}. Hence, the condition (3.2) is automatically satisfied. Therefore, our methodology applies as soon as $\alpha > 1/2$ (which ensures that $\int_{\mathbb{R}^d} \|x\| \mu_{\alpha,d}(dx) < +\infty$).

Takano distribution II [48]. Let μ be the probability measure on $\mathbb{R}^d$ given by

$$\mu(dx) = C \exp\left(-\|x\|\right) dx,$$

where $C > 0$ is a normalizing constant. Thanks to [48, Result 1], its characteristic function φ is given by

$$\varphi(\xi) = \exp\left(\int_{\mathbb{R}^d} \left(e^{i\langle \xi; u \rangle} - 1 \right) \frac{M(\|u\|)}{\|u\|^d} du \right), \quad \xi \in \mathbb{R}^d,$$

where $M(w) = (2\pi)^{-d/2}(d+1) w^{d/2} K_{d/2}(w)$, for $w > 0$. Hence, μ is an infinitely divisible probability measure on $\mathbb{R}^d$. Moreover, the function M admits the following representation (see the last formula page 64 of [48])

$$M(w) = C_d \int_w^{+\infty} v^{d/2} K_{(d-2)/2}(v) dv, \quad w > 0,$$

which is non-negative and decreasing on $(0, +\infty)$ (and $C_d > 0$). Thus, μ is self-decomposable and its Lévy measure admits the following polar decomposition

$$\nu(du) = \mathbb{1}_{(0,+\infty)}(r) \mathbb{1}_{S^{d-1}}(x) \frac{M(r)}{r} dr\sigma(dx),$$

where σ is the uniform measure on the Euclidean unit sphere. Finally, the associated functions k_x satisfy (3.2) and the probability measure μ admits finite moments of any orders.

Multivariate gamma distribution. Let $(\alpha_1, \dots, \alpha_d) \in (0, +\infty)^d$ and let $X = (X_1, \dots, X_d)$ be a random vector whose independent coordinates are distributed according to gamma laws with parameters $(\alpha_i, 1)$, $1 \leq i \leq d$. The characteristic function of X is given by

$$\varphi(\xi) = \prod_{j=1}^d \left(1 - i\xi_j\right)^{-\alpha_j}, \quad \xi \in \mathbb{R}^d.$$

For any $b > 1$, there exists ρ_b, a probability measure on $\mathbb{R}^d$, such that, $\varphi(\xi) = \varphi(\xi/b)\hat{\rho}_b(\xi)$, for all $\xi \in \mathbb{R}^d$. Indeed, take $\rho_b = \rho_{1,b} \otimes ... \otimes \rho_{d,b}$, where, for any $1 \leq j \leq d$, $\rho_{j,b}$ is defined, for all $\xi_j \in \mathbb{R}$, by

$$\int_{\mathbb{R}} e^{i\xi_j x}\rho_{j,b}(dx) = \frac{(1-i\xi_j)^{-\alpha_j}}{\left(1-i\frac{\xi_j}{b}\right)^{-\alpha_j}}.$$

Therefore, our methodology applies to these multivariate gamma distributions. Moreover, for all $\xi \in \mathbb{R}^d$ and all $t > 0$,

$$e^{-t\sum_{j=1}^d \alpha_j} \leq \left|\frac{\varphi(\xi)}{\varphi(e^{-t}\xi)}\right| \leq 1.$$

Other types of self-decomposable multivariate gamma distributions have been considered in the literature. In particular, [37] considers a class of infinitely divisible multivariate gamma distributions with Lévy measure having the following polar decomposition,

$$\nu(du) = \mathbb{1}_{(0,+\infty)}(r)\mathbb{1}_{S^{d-1}}(x)\alpha\frac{\exp(-\beta r)}{r}dr\lambda(dx),$$

where α, β are positive reals and λ is a finite positive measure on the d-dimensional Euclidean unit sphere. Therefore, $k_x(r) = \alpha e^{-\beta r}$, $r > 0$ and so (3.2) is again satisfied. Moreover, such infinitely divisible multivariate gamma distributions are self-decomposable and admit absolute moment of any orders.

Another way to build probability measures on $\mathbb{R}^d$ which are self-decomposable is through mixtures. For example, thanks to [50, Corollary p. 40], the mixture of a d-multivariate normal distribution, $\mathcal{N}(m, \Gamma I_d)$, $m \in \mathbb{R}^d$, and of a generalized gamma convolution Γ (see, [5] for a definition) is self-decomposable.

Finally, let us explain how one can build an example of a self-decomposable distribution for which the k_x function defined in (2.5) does not satisfied the condition (3.2). Let $d = 2$ and let σ be the uniform measure on the circle. Let $\alpha \in (1, 2)$ and let β be a positive function defined on S^1, measurable on S^1 and such that

$$\int_{S^1} \beta(x)\sigma(dx) < +\infty, \quad \sup_{x\in S^1}\beta(x) = +\infty.$$

Then, the self-decomposable distribution whose Lévy measure is given through the polar decomposition

$$\nu(du) = \mathbb{1}_{(0,+\infty)}(r)\mathbb{1}_{S^1}(x)\frac{\beta(x)}{r^{\alpha+1}}dr\sigma(dx),$$

has finite first moment but does not satisfied the condition (3.2) since then $k_x(r) = \beta(x)/r^\alpha$, for all $r > 0$ and a.e. $x \in S^1$.

Let us next return to the integro-partial differential equation (3.9) and solve it for any $h \in \mathcal{H}_2 \cap \mathcal{C}_c^\infty(\mathbb{R}^d)$. In fact, under the assumption of Proposition 3.4, it is possible to solve, *mutatis mutandis*, the Stein equation (3.9) for $h \in \mathcal{H}_1 \cap \mathcal{C}_c^\infty(\mathbb{R}^d)$.

Proposition 3.5. *Let X be a non-degenerate self-decomposable random vector in $\mathbb{R}^d$ without Gaussian component, with law μ_X, with Lévy measure ν and such that $\mathbb{E}\|X\| < \infty$. Moreover, let the functions k_x given by* (2.5) *satisfy* (3.2)*. Let $h \in \mathcal{H}_2 \cap \mathcal{C}_c^\infty(\mathbb{R}^d)$ and let f_h be the function given by Proposition 3.3. Then, for all $x \in \mathbb{R}^d$,*

$$\langle \mathbb{E}X - x; \nabla(f_h)(x)\rangle + \int_{\mathbb{R}^d}\langle \nabla(f_h)(x+u) - \nabla(f_h)(x); u\rangle\nu(du) = h(x) - \mathbb{E}h(X).$$

Proof of Proposition 3.5. Let $h \in \mathcal{H}_2 \cap \mathcal{C}_c^\infty(\mathbb{R}^d)$, let f_h be given by (3.10) and let $\hat{h} = h - \mathbb{E}h(X)$.
Step 1: Let us prove that for all $t > 0$ and all $x \in \mathbb{R}^d$,

$$\frac{d}{dt}\left(P_t^\nu(h)(x)\right) = \mathcal{A}(P_t^\nu(h))(x). \tag{3.18}$$

Since $h \in \mathcal{C}_c^\infty(\mathbb{R}^d)$, by Fourier inversion, for all $t > 0$ and for all $x \in \mathbb{R}^d$,

$$P_t^\nu(h)(x) = \int_{\mathbb{R}^d} \mathcal{F}(h)(\xi) e^{ie^{-t}\langle x;\xi\rangle} \frac{\varphi(\xi)}{\varphi(e^{-t}\xi)} \frac{d\xi}{(2\pi)^d},$$

which will be used to compute $d(P_t^\nu(h)(x))/dt$. First, note that, for all $x \in \mathbb{R}^d$, $\xi \in \mathbb{R}^d$ and $t > 0$,

$$\begin{aligned}
&\frac{d}{dt}\left(e^{ie^{-t}\langle x;\xi\rangle} \frac{\varphi(\xi)}{\varphi(e^{-t}\xi)}\right) \\
&\quad = -ie^{-t}\langle x;\xi\rangle e^{ie^{-t}\langle x;\xi\rangle} \frac{\varphi(\xi)}{\varphi(e^{-t}\xi)} + e^{ie^{-t}\langle x;\xi\rangle} \varphi(\xi) e^{-t} \left(\sum_{j=1}^d \xi_j \frac{\partial_j(\varphi)(e^{-t}\xi)}{\varphi(\xi e^{-t})^2}\right) \\
&\quad = e^{ie^{-t}\langle x;\xi\rangle} \frac{\varphi(\xi)}{\varphi(e^{-t}\xi)} \left(-ie^{-t}\langle x;\xi\rangle + e^{-t}\sum_{j=1}^d \xi_j \frac{\partial_j(\varphi)(e^{-t}\xi)}{\varphi(\xi e^{-t})}\right) \\
&\quad = e^{ie^{-t}\langle x;\xi\rangle} \frac{\varphi(\xi)}{\varphi(e^{-t}\xi)} ie^{-t} \left(\langle \mathbb{E}X - x;\xi\rangle + \int_{\mathbb{R}^d} \langle u;\xi\rangle (e^{i\langle u;e^{-t}\xi\rangle} - 1)\nu(du)\right).
\end{aligned}$$

Moreover, for all $x \in \mathbb{R}^d$, for all $\xi \in \mathbb{R}^d$ and all $t > 0$,

$$\begin{aligned}
\left|\frac{d}{dt}\left(e^{ie^{-t}\langle x;\xi\rangle} \frac{\varphi(\xi)}{\varphi(e^{-t}\xi)}\right)\right| &\le e^{-t}\Bigg(\|\mathbb{E}X - x\|\|\xi\| + \|\xi\|^2 e^{-t} \int_{\|u\|\le 1} \|u\|^2 \nu(du) \\
&\qquad + 2\|\xi\| \int_{\|u\|\ge 1} \|u\|\nu(du)\Bigg) \\
&\le \Bigg(\|\mathbb{E}X - x\|\|\xi\| + \|\xi\|^2 \int_{\|u\|\le 1} \|u\|^2 \nu(du) \\
&\qquad + 2\|\xi\| \int_{\|u\|\ge 1} \|u\|\nu(du)\Bigg).
\end{aligned}$$

Hence,

$$\begin{aligned}
\frac{d}{dt}(P_t^\nu(h)(x)) = \int_{\mathbb{R}^d} \mathcal{F}(h)(\xi) e^{ie^{-t}\langle x;\xi\rangle} \frac{\varphi(\xi)}{\varphi(e^{-t}\xi)} ie^{-t}\Bigg(&\langle \mathbb{E}X - x;\xi\rangle \\
&+ \int_{\mathbb{R}^d} \langle u;\xi\rangle (e^{i\langle u;e^{-t}\xi\rangle} - 1)\nu(du)\Bigg) \frac{d\xi}{(2\pi)^d}.
\end{aligned}$$

To conclude the first step, let us precisely compute the right-hand side of the previous equality. First,

$$\begin{aligned}
\int_{\mathbb{R}^d} \mathcal{F}(h)(\xi) e^{ie^{-t}\langle x;\xi\rangle} \frac{\varphi(\xi)}{\varphi(e^{-t}\xi)} ie^{-t} \langle \mathbb{E}X - x;\xi\rangle \frac{d\xi}{(2\pi)^d} &= \langle \mathbb{E}X - x; e^{-t}P_t^\nu(\nabla(h))(x)\rangle \\
&= \langle \mathbb{E}X - x; \nabla(P_t^\nu(h))(x)\rangle.
\end{aligned}$$

using $e^{-t}P_t^\nu(\partial_j(h))(x) = \partial_j(P_t^\nu(h))(x)$, for all $t \ge 0$, all $x \in \mathbb{R}^d$ and all $1 \le j \le d$. Moreover, by Fubini Theorem and Fourier arguments,

$$\begin{aligned}(I) &:= \int_{\mathbb{R}^d} \mathcal{F}(h)(\xi) e^{ie^{-t}\langle x;\xi\rangle} \frac{\varphi(\xi)}{\varphi(e^{-t}\xi)} ie^{-t} \left(\int_{\mathbb{R}^d} \langle u;\xi\rangle (e^{i\langle u;e^{-t}\xi\rangle}-1)\nu(du)\right) \frac{d\xi}{(2\pi)^d} \\ &= \sum_{j=1}^d \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} u_j \Big(\mathcal{F}(\partial_j(h)(.+e^{-t}u))(\xi) - \mathcal{F}(\partial_j(h))(\xi)\Big)\nu(du)\right) \\ &\quad \times e^{ie^{-t}\langle x;\xi\rangle} \frac{\varphi(\xi)}{\varphi(e^{-t}\xi)} e^{-t} \frac{d\xi}{(2\pi)^d} \\ &= \sum_{j=1}^d \int_{\mathbb{R}^d} u_j e^{-t} \left(P_t^\nu(\partial_j(h))(x+u) - P_t^\nu(\partial_j(h))(x)\right)\nu(du) \\ &= \int_{\mathbb{R}^d} \langle u; \nabla(P_t^\nu(h))(x+u) - \nabla(P_t^\nu(h))(x)\rangle \nu(du).\end{aligned}$$

Thus, for all $x \in \mathbb{R}^d$ and all $t > 0$,

$$\frac{d}{dt}\left(P_t^\nu(h)\right)(x) = \mathcal{A}(P_t^\nu(h))(x),$$

which gives (3.18) and finishes the proof of the first step.
Step 2: Let $0 < b < +\infty$. Integrating out the equality (3.18) gives,

$$P_b^\nu(h)(x) - h(x) = \int_0^b \mathcal{A}(P_t^\nu(h))(x)dt,$$

then, letting $b \to +\infty$ and using Lemma 3.1 lead to:

$$\lim_{b\to+\infty}\left(P_b^\nu(h)(x) - h(x)\right) = -\hat{h}(x), \qquad x \in \mathbb{R}^d.$$

Next, let us show that $\int_0^{+\infty} |\mathcal{A}(P_t^\nu(h))(x)|\, dt < +\infty$, for all $x \in \mathbb{R}^d$. To do so, we need to estimate the quantities $\|\nabla(P_t^\nu(h))(x)\|$ and $\|\nabla(P_t^\nu(h))(x+u) - \nabla(P_t^\nu(h))(x)\|$, for all $x \in \mathbb{R}^d$, all $u \in \mathbb{R}^d$, and all $t \geq 0$. Since $\partial_j\left(P_t^\nu(h)\right)(x) = e^{-t}P_t^\nu(\partial_j(h))(x)$ and since $h \in \mathcal{H}_2$,

$$\begin{aligned}\|\nabla(P_t^\nu(h))(x)\| &\leq \sqrt{d}e^{-t}, \\ \|\nabla(P_t^\nu(h))(x+u) - \nabla(P_t^\nu(h))(x)\| &\leq \sqrt{d}e^{-t}\mathbb{1}_{\|u\|\geq 1} + \sqrt{d}e^{-t}\|u\|\mathbb{1}_{\|u\|\leq 1}.\end{aligned}$$

Then, by the very definitions of $\mathcal{A}$ and of P_t^ν and standard inequalities, for all $x \in \mathbb{R}^d$, and all $t \geq 0$,

$$|\mathcal{A}(P_t^\nu(h))(x)| \leq \sqrt{d}e^{-t}\left(\|\mathbb{E}X - x\| + \left(\int_{\|u\|\leq 1} \|u\|^2\nu(du) + \int_{\|u\|\geq 1} \|u\|\nu(du)\right)\right),$$

which implies that $\int_0^{+\infty} |\mathcal{A}(P_t^\nu(h))(x)|\, dt < +\infty$, for all $x \in \mathbb{R}^d$. Moreover, $\mathcal{A}(P_t^\nu(h))(x) = \mathcal{A}(P_t^\nu(\hat{h}))(x)$, thus, for all $x \in \mathbb{R}^d$,

$$-\hat{h}(x) = \int_0^{+\infty} \mathcal{A}(P_t^\nu(\hat{h}))(x)dt.$$

To conclude, one needs to prove that, for all $x \in \mathbb{R}^d$,

$$\int_0^{+\infty} \mathcal{A}(P_t^\nu(\hat{h}))(x)dt = \mathcal{A}\left(\int_0^{+\infty} P_t^\nu(\hat{h})(x)dt\right) = -\mathcal{A}(f_h)(x).$$

But, this follows from standard arguments as well as from Proposition 3.3. □

We end this section by proving regularity estimates for the solution f_h of the Stein equation under the assumptions of Proposition 3.3. Similar estimates hold true under the assumptions of Proposition 3.4. In particular, under these latter assumptions, it is sufficient to have $h \in \mathcal{H}_1$ to obtain a bound on $M_2(f_h)$, and this is in line with the Gaussian case (see [39, 11, 31]).

Proposition 3.6. *Let X be a non-degenerate self-decomposable random vector in $\mathbb{R}^d$ without Gaussian component, with law μ_X, characteristic function φ and such that $\mathbb{E}\|X\| < \infty$. Moreover, let the functions k_x given by* (2.5) *satisfy* (3.2)*. Let $h \in \mathcal{H}_2$ and $(P_t^\nu)_{t\geq 0}$ be the semigroup of operators obtained in Lemma 3.1. Then, f_h given, for all $x \in \mathbb{R}^d$, by*

$$f_h(x) = -\int_0^{+\infty} (P_t^\nu(h)(x) - \mathbb{E}h(X))dt,$$

is such that

$$M_1(f_h) \leq 1, \qquad M_2(f_h) \leq \frac{1}{2}.$$

Proof of Proposition 3.6. By definition,

$$M_1(f_h) = \sup_{x\in\mathbb{R}^d} \|\nabla(f_h)(x)\|_{op}.$$

Let $u \in \mathbb{R}^d$ with $\|u\| = 1$. Then, by the commutation relation $\nabla(P_t^\nu(h))(x) = e^{-t} P^\nu(\nabla(h))(x)$, for all $x \in \mathbb{R}^d$ and for all $t \geq 0$, one has

$$\nabla(f_h)(x) = -\int_0^{+\infty} e^{-t} P_t^\nu(\nabla(h))(x)dt, \quad x \in \mathbb{R}^d.$$

Hence, for all $x \in \mathbb{R}^d$,

$$\langle \nabla(f_h)(x); u\rangle = -\int_0^{+\infty} e^{-t} P_t^\nu \left(\langle \nabla(h); u\rangle\right)(x)dt,$$

which readily implies that $M_1(f_h) \leq M_1(h) \leq 1$ since $h \in \mathcal{H}_2$. The bound on $M_2(f_h)$ follows similarly using the commutation relation twice and the fact that $h \in \mathcal{H}_2$. ☐

4 Stein kernels for SD laws with finite second moment

In the Gaussian setting, a major finding in the context of Stein's method is the introduction of the notion of Stein kernel (see e.g. [46, 7, 8, 9, 32, 10, 33, 35, 25, 16]). Recall that γ_d, a centered Gaussian measure on $\mathbb{R}^d$, satisfies the following integration by parts formula,

$$\int_{\mathbb{R}^d} \langle x; f(x)\rangle \gamma_d(dx) = \int_{\mathbb{R}^d} \operatorname{div}(f(x))\gamma_d(dx),$$

for all smooth enough, $\mathbb{R}^d$-valued function $f = (f_1, ..., f_d)$ and where $\operatorname{div}(f(x)) = \sum_{j=1}^d \partial_j(f_j)(x)$. For a centered probability measure ρ on $\mathbb{R}^d$, the Gaussian Stein kernel of ρ is the measurable function τ_ρ, from $\mathbb{R}^d$ to $\mathcal{M}_{d\times d}(\mathbb{R})$, the space of $d \times d$ real matrices, such that, for all smooth enough $\mathbb{R}^d$-valued function f,

$$\int_{\mathbb{R}^d} \langle x; f(x)\rangle \rho(dx) = \int_{\mathbb{R}^d} \langle \tau_\rho(x); \nabla(f)(x)\rangle_{HS} \rho(dx),$$

where $\langle A; B\rangle_{HS} = \operatorname{Tr}(A^t B)$, for $A, B \in \mathcal{M}_{d\times d}(\mathbb{R})$. Recall, also from the previous section, that a non-degenerate self-decomposable random vector X without Gaussian component, with law μ_X and with finite first moment satisfies, for f smooth enough, the following characterizing equation:

$$\int_{\mathbb{R}^d} \langle x - \mathbb{E}X; \nabla(f)(x)\rangle \mu_X(dx) = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \langle \nabla(f)(x+u) - \nabla(f)(x); u\rangle \nu(du) \right) \mu_X(dx).$$

Then, quite naturally, in the infinitely divisible framework, let us introduce the following definitions of Stein kernels and of the Stein discrepancy.

Definition 4.1. *Let X be a centered non-degenerate infinitely divisible random vector without Gaussian component, with law μ_X, Lévy measure ν and such that $\mathbb{E}\|X\|^2 < \infty$. Let Y be a centered random vector with law μ_Y and such that $\mathbb{E}\|Y\|^2 < \infty$. A Stein kernel of Y with respect to X is a measurable function τ_Y from $\mathbb{R}^d$ to $\mathbb{R}^d$ such that,*

$$\int_{\mathbb{R}^d} \langle y; f(y)\rangle \mu_Y(dy) = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \langle f(y+u) - f(y); \tau_Y(y+u) - \tau_Y(y)\rangle \nu(du) \right) \mu_Y(dy),$$

for all $\mathbb{R}^d$-valued test function f for which both sides of the previous equality are well defined. The Stein's discrepancy of μ_Y with respect to μ_X is given by

$$S(\mu_Y || \mu_X) = \inf \left(\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \|\tau_Y(y+u) - \tau_Y(y) - u\|^2 \nu(du) \mu_Y(dy) \right)^{1/2},$$

where the infimum is taken over all Stein kernels of Y with respect to X, and is equal to $+\infty$ if no such Stein kernel exists.

The next result ensures that the Stein's discrepancy provides a good control of some classical metrics between probability measures on $\mathbb{R}^d$.

Theorem 4.2. *Let X be a centered non-degenerate self-decomposable random vector without Gaussian component, with law μ_X, Lévy measure ν, such that $\mathbb{E}\|X\|^2 < +\infty$ and let also the functions k_x given by* (2.5) *satisfy* (3.2)*. Let Y be a centered non-degenerate random vector with law μ_Y, such that $\mathbb{E}\|Y\|^2 < +\infty$, and for which a Stein kernel with respect to X exists. Then,*

$$d_{W_2}(\mu_X, \mu_Y) \leq \frac{1}{2} \left(\int_{\mathbb{R}^d} \|u\|^2 \nu(du) \right)^{1/2} S(\mu_Y || \mu_X).$$

Proof of Theorem 4.2. Let $h \in \mathcal{H}_2 \cap \mathcal{C}_c^\infty(\mathbb{R}^d)$. By Proposition 3.5, f_h is a solution to,

$$-\langle x; \nabla(f_h)(x)\rangle + \int_{\mathbb{R}^d} \langle \nabla(f_h)(x+u) - \nabla(f_h)(x); u\rangle \nu(du) = h(x) - \mathbb{E}h(X), \qquad x \in \mathbb{R}^d,$$

and thus,

$$\mathbb{E}\left(-\langle Y; \nabla(f_h)(Y)\rangle + \int_{\mathbb{R}^d} \langle \nabla(f_h)(Y+u) - \nabla(f_h)(Y); u\rangle \nu(du) \right) = \mathbb{E}h(Y) - \mathbb{E}h(X).$$

Now, since Y admits a Stein kernel with respect to X,

$$\mathbb{E}h(Y) - \mathbb{E}h(X) = \mathbb{E}\left(\int_{\mathbb{R}^d} \langle \nabla(f_h)(Y+u) - \nabla(f_h)(Y); u - \tau_Y(Y+u) + \tau_Y(Y)\rangle \nu(du) \right).$$

Taking the absolute values and applying the Cauchy-Schwarz inequality,

$$|\mathbb{E}h(Y) - \mathbb{E}h(X)| \leq \mathbb{E}\int_{\mathbb{R}^d} \|\nabla(f_h)(Y+u) - \nabla(f_h)(Y)\| \|\tau_Y(Y+u) - \tau_Y(Y) - u\| \nu(du).$$

Now, by the very definition of $M_2(f_h)$ and by the Cauchy-Schwarz inequality (applied twice), the following bound holds true

$$|\mathbb{E}h(Y) - \mathbb{E}h(X)| \leq M_2(f_h)\sqrt{\int_{\mathbb{R}^d} \|u\|^2\nu(du)}\sqrt{\mathbb{E}\int_{\mathbb{R}^d} \|\tau_Y(Y+u) - \tau_Y(Y) - u\|^2\nu(du)}.$$

To conclude use the definition of the Stein discrepancy and Proposition 3.6. □

In the sequel, we wish to discuss sufficient conditions for the existence of Stein kernels as defined above. For this purpose, let us recall, in connection with Poincaré type inequalities in an infinitely divisible setting, some definitions and results from [12, 14]. First, if X is a non-degenerate infinitely divisible random vector in $\mathbb{R}^d$ without Gaussian component, with law μ_X and with Lévy measure ν and if $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ is such that $\mathbb{E}f(X)^2 + \mathbb{E}\int_{\mathbb{R}^d} |f(X+u) - f(X)|^2\nu(du) < +\infty$, then [12, Theorem 4.1] gives

$$\operatorname{Var} f(X) \leq \mathbb{E}\int_{\mathbb{R}^d} |f(X+u) - f(X)|^2\nu(du). \tag{4.1}$$

Further, if Y is a centered non-degenerate random vector in $\mathbb{R}^d$ such that $\mathbb{E}\|Y\|^2 < +\infty$, if ν is a Lévy measure in $\mathbb{R}^d$ such that $\int_{\mathbb{R}^d} \|u\|^2\nu(du) < +\infty$ and if $\mathcal{H}_Y$ is the space of real valued functions f on $\mathbb{R}^d$ such that $\mathbb{E}f(Y)^2 < +\infty$ and $0 < \mathbb{E}\int_{\mathbb{R}^d} |f(Y+u) - f(Y)|^2\nu(du) < +\infty$, then the Poincaré constant $U(Y,\nu)$ defined as

$$U(Y,\nu) = \sup_{f\in\mathcal{H}_Y} \frac{\operatorname{Var}(f(Y))}{\mathbb{E}\int_{\mathbb{R}^d} |f(Y+u) - f(Y)|^2\nu(du)}, \tag{4.2}$$

characterizes the proximity in law of Y to a centered infinitely divisible random vector with finite second moment and Lévy measure ν. Indeed, [14, Theorem 2.1] proves the following: if $\mathbb{E}|Y_i|^2 = \int_{\mathbb{R}^d} |u_i|^2\nu(du)$, for all $1 \leq i \leq d$, then $U(Y,\nu) \geq 1$ and $U(Y,\nu) = 1$ if and only if the characteristic function of Y is given by

$$\varphi_Y(\xi) = \exp\left(\int_{\mathbb{R}^d} \left(e^{i\langle\xi;u\rangle} - 1 - i\langle\xi;u\rangle\right)\nu(du)\right), \quad \xi \in \mathbb{R}^d,$$

i.e., if and only if $Y \sim ID(b,0,\nu)$, with $b = -\int_{\|u\|\geq 1} u\nu(du)$.

In the Gaussian case, the existence of a Stein kernel for multivariate distributions has been investigated with the help of variational methods. Indeed, in [16], under a spectral gap assumption, the existence of a Gaussian Stein kernel has been ensured thanks to the classical Lax-Milgram Theorem. Thus, in view of (4.2) and the associated characterization, it is natural to introduce the following variational setting: let Y be a centered non-degenerate random vector with finite second moment and with law μ_Y and let ν be a Lévy measure on $\mathbb{R}^d$ such that $\int_{\|u\|\geq 1} \|u\|^2\nu(du) < +\infty$. Moreover, assume that $\nu * \mu_Y << \mu_Y$, with $\nu * \mu_Y$ denoting the convolution of the two positive measures ν and μ_Y. Now, let $H_\nu(\mu_Y)$ be the vector space of Borel measurable $\mathbb{R}^d$-valued functions on $\mathbb{R}^d$ such that $\int_{\mathbb{R}^d} \|f(y)\|^2\mu_Y(dy) < +\infty$ and $\int_{\mathbb{R}^d\times\mathbb{R}^d} \|f(y+u) - f(y)\|^2\nu(du)\mu_Y(dy) < +\infty$ and let $H_{\nu,0}(\mu_Y)$ be the subspace of $H_\nu(\mu_Y)$ such that $\mathbb{E}f(Y) = 0$. (Two functions f and g of $H_\nu(\mu_Y)$ are identified as soon as $f = g$, μ_Y-almost everywhere.) Then, let us assume that Y satisfies a Poincaré inequality of the following type: there exists a positive and finite constant U_Y such that, for all $f \in H_\nu(\mu_Y)$

$$\mathbb{E}\|f(Y) - \mathbb{E}f(Y)\|^2 \leq U_Y\, \mathbb{E}\int_{\mathbb{R}^d} \|f(Y+u) - f(Y)\|^2\nu(du). \tag{4.3}$$

In particular, note that if Y is such that $U(Y,\nu) < +\infty$ in (4.2), then, for all $f \in H_\nu(\mu_Y)$ such that $f_j \in \mathcal{H}_Y$, $1 \leq j \leq d$,

$$\mathbb{E}|f_j(Y) - \mathbb{E}f_j(Y)|^2 \leq U(Y,\nu)\mathbb{E}\int_{\mathbb{R}^d} |f_j(Y+u) - f_j(Y)|^2\nu(du),$$

so that Y satisfies (4.3) with $U_Y = U(Y,\nu)$.

Moreover, let A be the bilinear functional defined, for all test functions f and g, by

$$A(f,g) = \mathbb{E}\int_{\mathbb{R}^d} \langle f(Y+u)-f(Y); g(Y+u)-g(Y)\rangle \nu(du), \tag{4.4}$$

and let L be the linear functional defined, for all test functions f, by

$$L(f) = \mathbb{E}\langle Y; f(Y)\rangle. \tag{4.5}$$

Before solving the variational problem associated with A, L and $H_\nu(\mu_Y)$, we need the following technical lemma.

Lemma 4.3. *The vector space $H_\nu(\mu_Y)$ endowed with the bilinear functional*

$$\langle f; g\rangle_{H_\nu(\mu_Y)} = \mathbb{E}\langle f(Y); g(Y)\rangle + A(f,g), \tag{4.6}$$

is a Hilbert space. Moreover, A, defined by (4.4), is continuous on $H_\nu(\mu_Y)\times H_\nu(\mu_Y)$, coercive on $H_{\nu,0}(\mu_Y)$ while, L, defined by (4.5), is continuous on $H_\nu(\mu_Y)$.

Proof of Lemma 4.3. First, it is clear that the bilinear symmetric functional $\langle \cdot;\cdot\rangle_{H_\nu(\mu_Y)}$ is an inner product on $H_\nu(\mu_Y)$. Then, let $\|\cdot\|_{H_\nu(\mu_Y)}$ be the induced norm defined via $\|f\|^2_{H_\nu(\mu_Y)} = \mathbb{E}\|f(Y)\|^2 + A(f,f)$, for all $f\in H_\nu(\mu_Y)$. Let us prove that $H_\nu(\mu_Y)$ endowed with this norm is complete. Let $(f_n)_{n\geq 1}$ be a Cauchy sequence in $H_\nu(\mu_Y)$. Therefore $(f_n)_{n\geq 1}$ is a Cauchy sequence in $L^2(\mu_Y)$, and there exists $f\in L^2(\mu_Y)$ such that $f_n\to f$, as $n\to+\infty$ in $L^2(\mu_Y)$. Now, pick a subsequence $(f_{n_k})_{k\geq 1}$ such that $f_{n_k}\to f$, μ_Y-almost everywhere, as $k\to+\infty$. Fatou's lemma together with the assumption that $\nu*\mu_Y << \mu_Y$ and the fact that $(f_n)_{n\geq 1}$ is a Cauchy sequence in $H_\nu(\mu_Y)$ (thus is bounded), imply that

$$A(f,f)\leq \liminf_{k\to+\infty} A(f_{n_k},f_{n_k}) \leq \sup_{n\geq 1}\|f_n\|^2_{H_\nu(\mu_Y)} < +\infty. \tag{4.7}$$

Therefore, $f\in H_\nu(\mu_Y)$. Another application of Fatou's lemma, and since $(f_n)_{n\geq 1}$ is Cauchy in $H_\nu(\mu_Y)$, shows that $f_n\to f$ in $H_\nu(\mu_Y)$. Now, by the Cauchy-Schwarz inequality, for all $f,g\in H_\nu(\mu_Y)$,

$$\begin{aligned}|A(f,g)| &\leq \left(\mathbb{E}\int_{\mathbb{R}^d}\|f(Y+u)-f(Y)\|^2\nu(du)\right)^{1/2}\left(\mathbb{E}\int_{\mathbb{R}^d}\|g(Y+u)-g(Y)\|^2\nu(du)\right)^{1/2}\\ &\leq \|f\|_{H_\nu(\mu_Y)}\|g\|_{H_\nu(\mu_Y)},\end{aligned}$$

which proves that A is a continuous bilinear functional on $H_\nu(\mu_Y)$. Moreover, since Y satisfies (4.3), for all $f\in H_{\nu,0}(\mu_Y)$

$$\begin{aligned}A(f,f) &= \mathbb{E}\int_{\mathbb{R}^d}\|f(Y+u)-f(Y)\|^2\nu(du),\\ &\geq \frac{1}{2}\mathbb{E}\int_{\mathbb{R}^d}\|f(Y+u)-f(Y)\|^2\nu(du) + \frac{1}{2U_Y}\mathbb{E}\|f(Y)\|^2,\\ &\geq C_Y\|f\|^2_{H_\nu(\mu_Y)},\end{aligned}$$

for $2C_Y = \min(1, 1/(U_Y)) > 0$ and so A is coercive on $H_{\nu,0}(\mu_Y)$. Finally, the continuity property of the linear functional L on $H_\nu(\mu_Y)$ follows from the Cauchy-Schwarz inequality, from $\mathbb{E}\|Y\|^2 < +\infty$, and from the continuous embedding $H_\nu(\mu_Y)\hookrightarrow L^2(\mu_Y)$. □

Note that since $H_{\nu,0}(\mu_Y)$ is a closed subspace of $H_\nu(\mu_Y)$, it is as well a Hilbert space with the inner product $\langle .;.\rangle_{H_\nu(\mu_Y)}$.

Based on Lemma 4.3, a direct application of the Lax-Milgram Theorem (see, e.g., [17, Theorem 1 page 297] for a precise statement as well as a proof) ensures the existence of a Stein kernel in the sense of Definition 4.1 for probability measures μ_Y satisfying the Poincaré-type inequality (4.3). This is the content of the next theorem.

Theorem 4.4. *Let Y be a centered non-degenerate random vector with finite second moment and with law μ_Y, let ν be a Lévy measure on $\mathbb{R}^d$ such that $\int_{\|u\|\geq 1}\|u\|^2\nu(du) < +\infty$ and let $\nu * \mu_Y << \mu_Y$. Let Y satisfy the Poincaré-type inequality* (4.3) *for some $0 < U_Y < +\infty$. Then, there exists a unique $\tau_Y \in H_{\nu,0}(\mu_Y)$, such that, for all $f \in H_{\nu,0}(\mu_Y)$*

$$A(f, \tau_Y) = L(f). \tag{4.8}$$

Moreover,

$$\mathbb{E}\int_{\mathbb{R}^d}\|\tau_Y(Y+u) - \tau_Y(Y)\|^2\nu(du) \leq U_Y\mathbb{E}\|Y\|^2. \tag{4.9}$$

Proof of Theorem 4.4. The first part of the theorem is a direct application of the Lax-Milgram Theorem with A, L and $H_{\nu,0}(\mu_Y)$. To obtain the inequality (4.9), note that thanks to (4.8) with $f = \tau_Y$,

$$\mathbb{E}\int_{\mathbb{R}^d}\|\tau_Y(Y+u) - \tau_Y(Y)\|^2\nu(du) = A(\tau_Y, \tau_Y) = L(\tau_Y) \leq \sqrt{\mathbb{E}\|Y\|^2}\sqrt{\mathbb{E}\|\tau_Y(Y)\|^2}. \tag{4.10}$$

Then, (4.3) applied to $f(\cdot) = \tau_Y(\cdot)$ (note that $\mathbb{E}\tau_Y(Y) = 0$) ensures that

$$\mathbb{E}\|\tau_Y(Y)\|^2 \leq U_Y\,\mathbb{E}\int_{\mathbb{R}^d}\|\tau_Y(Y+u) - \tau_Y(Y)\|^2\nu(du). \tag{4.11}$$

Finally, (4.10) together with (4.11) implies

$$\mathbb{E}\int_{\mathbb{R}^d}\|\tau_Y(Y+u) - \tau_Y(Y)\|^2\nu(du) \leq \sqrt{U_Y}\sqrt{\mathbb{E}\|Y\|^2}\sqrt{\mathbb{E}\int_{\mathbb{R}^d}\|\tau_Y(Y+u) - \tau_Y(Y)\|^2\nu(du)},$$

which concludes the proof. □

The next theorem is the main result of this section.

Theorem 4.5. *Let X be a centered non-degenerate self-decomposable random vector without Gaussian component, with law μ_X, with Lévy measure ν, such that $\mathbb{E}\|X\|^2 < +\infty$ and let also the functions k_x given by* (2.5) *satisfy* (3.2)*. Let Y be a centered non-degenerate random vector with law μ_Y, with $\mathbb{E}\|Y\|^2 < +\infty$ and such that $\nu * \mu_Y << \mu_Y$. Let Y satisfy a Poincaré-type inequality* (4.3) *with $1 \leq U_Y < +\infty$. Then,*

$$d_{W_2}(\mu_X, \mu_Y) \leq \frac{1}{2}\left(\int_{\mathbb{R}^d}\|u\|^2\nu(du)\right)^{1/2}\left(U_Y\mathbb{E}\|Y\|^2 + \int_{\mathbb{R}^d}\|u\|^2\nu(du) - 2\mathbb{E}\|Y\|^2\right)^{1/2}. \tag{4.12}$$

Moreover, if $\mathbb{E}\|Y\|^2 = \int_{\mathbb{R}^d}\|u\|^2\nu(du)$, then

$$d_{W_2}(\mu_X, \mu_Y) \leq \frac{1}{2}\left(\int_{\mathbb{R}^d}\|u\|^2\nu(du)\right)\sqrt{U_Y - 1}. \tag{4.13}$$

Proof of Theorem 4.5. Let us start with the proof of (4.13). First, note that since $M_1(f_h) < +\infty$ and $M_2(f_h) < +\infty$, for $h \in \mathcal{H}_2 \cap \mathcal{C}_c^\infty(\mathbb{R}^d)$, $\nabla(f_h)$ belongs to $H_\nu(\mu_Y)$ with f_h given by Proposition 3.5. Thus, since $\mathbb{E}Y = 0$, by Theorem 4.2,

$$d_{W_2}(\mu_X, \mu_Y) \leq \frac{1}{2}\left(\int_{\mathbb{R}^d}\|u\|^2\nu(du)\right)^{1/2} S\left(\mu_Y||\mu_X\right). \tag{4.14}$$

We continue by estimating $\mathbb{E}\int_{\mathbb{R}^d}\|\tau_Y(Y+u) - \tau_Y(Y) - u\|^2\nu(du)$. By the Pythagorean Theorem, Definition 4.1 and the fact that $\mathbb{E}\|Y\|^2 = \int_{\mathbb{R}^d}\|u\|^2\nu(du)$

$$\begin{aligned}\mathbb{E}\int_{\mathbb{R}^d}\|\tau_Y(Y+u)-\tau_Y(Y)-u\|^2\nu(du) &= \mathbb{E}\int_{\mathbb{R}^d}\|\tau_Y(Y+u)-\tau_Y(Y)\|^2\nu(du)+\int_{\mathbb{R}^d}\|u\|^2\nu(du)\\ &\quad -2\,\mathbb{E}\int_{\mathbb{R}^d}\langle u;\tau_Y(Y+u)-\tau_Y(Y)\rangle\nu(du),\\ &= \mathbb{E}\int_{\mathbb{R}^d}\|\tau_Y(Y+u)-\tau_Y(Y)\|^2\nu(du)+\int_{\mathbb{R}^d}\|u\|^2\nu(du)\\ &\quad -2\,\mathbb{E}\|Y\|^2,\\ &= \mathbb{E}\int_{\mathbb{R}^d}\|\tau_Y(Y+u)-\tau_Y(Y)\|^2\nu(du)-\int_{\mathbb{R}^d}\|u\|^2\nu(du),\end{aligned}$$

where the definition of the Stein kernel has been used in the second equality with $f(y)=y$, for all $y\in\mathbb{R}^d$. Moreover, (4.9) implies that

$$\mathbb{E}\int_{\mathbb{R}^d}\|\tau_Y(Y+u)-\tau_Y(Y)-u\|^2\nu(du)\le (U_Y-1)\int_{\mathbb{R}^d}\|u\|^2\nu(du),$$

so that

$$S\left(\mu_Y||\mu_X\right)\le\sqrt{U_Y-1}\left(\int_{\mathbb{R}^d}\|u\|^2\nu(du)\right)^{1/2}. \tag{4.15}$$

Combining (4.14) and (4.15) concludes the proof of the theorem. The proof of (4.12) follows in a completely similar manner. □

Remark 4.6. (i) When $\mathbb{E}\|Y\|^2=\int_{\mathbb{R}^d}\|u\|^2\nu(du)$ and $\mathbb{E}Y=0$, note that $U_Y\ge 1$, since, in (4.3), one can take $f(y)=y$, for all $y\in\mathbb{R}^d$.
(ii) If Y is as in Theorem 4.5 with $\mathbb{E}\|Y\|^2=\int_{\mathbb{R}^d}\|u\|^2\nu(du)$, and if $U_Y=1$, then, clearly from Theorem 4.5, $Y=_d X$ since $d_{W_2}(\mu_X,\mu_Y)=0$. Conversely, if $Y=_d X$, with X as in Theorem 4.5, then, for all $f=(f_1,...,f_d)$, (4.1) asserts that

$$\mathbb{E}|f_j(Y)-\mathbb{E}f_j(Y)|^2\le\mathbb{E}\int_{\mathbb{R}^d}|f_j(Y+u)-f_j(Y)|^2\nu(du), \tag{4.16}$$

for all $1\le j\le d$. Therefore, $U_Y=1$.
(iii) The following inequality on the Stein discrepancy is a direct byproduct of the proof of the previous theorem

$$S\left(\mu_Y||\mu_X\right)\le\left(U_Y\mathbb{E}\|Y\|^2+\int_{\mathbb{R}^d}\|u\|^2\nu(du)-2\mathbb{E}\|Y\|^2\right)^{1/2}.$$

(iv) All the results presented above should be compared with the analogous Gaussian ones obtained in [16] (see [16, Theorem 2.4 and Corollary 2.5]).
(v) Finally, let us mention the work [51] where similar one-dimensional quantitative results have been obtained in the Gaussian and Poisson cases. More specifically, [51, Theorem 1] (respectively [51, Theorem 2]) provides upper bound on the total variation distance between a probability measure, for which an appropriate version of $U(Y,\nu)$ defined by (4.2) is finite, and the Gaussian distribution (respectively the Poisson distribution).

The following convergence result is a straightforward consequence of Theorem 4.5.

Corollary 4.7. *Let X be a centered non-degenerate self-decomposable random vector without Gaussian component, with law μ_X, Lévy measure ν, such that $\mathbb{E}\|X\|^2<+\infty$ and let also the functions k_x given by* (2.5) *satisfy* (3.2)*. Let $(Y_n)_{n\ge1}$ be a sequence of centered square-integrable non-degenerate random vectors with laws $(\mu_n)_{n\ge1}$, such that $\nu*\mu_n<<\mu_n$, for all $n\ge1$, and such that Y_n satisfies the Poincaré type inequality* (4.3) *with $1\le U_n<+\infty$, for all $n\ge1$. If $\mathbb{E}\|Y_n\|^2\to\int_{\mathbb{R}^d}\|u\|^2\nu(du)$ and $U_n\to1$, as n tends to $+\infty$, then, $(Y_n)_{n\ge1}$ converges in distribution towards X.*

To end this section, we briefly discuss the condition $\nu * \mu_Y << \mu_Y$ appearing in Theorems 4.4 and 4.5. For this purpose, let ν be the Lévy measure of a non-degenerate infinitely divisible random vector, X, in $\mathbb{R}^d$ with law μ_X. Now, let $\mathcal{P}(\nu)$ be the set of probability measures, μ, on $\mathbb{R}^d$, such that $\nu * \mu << \mu$. First of all, thanks to [12, Lemma 4.1], the set $\mathcal{P}(\nu)$ is non-empty and contains the probability measure μ_X. Moreover, it is clearly a convex set. Next, let us describe some further non-trivial examples of probability measures belonging to $\mathcal{P}(\nu)$. For this purpose, we say that two probability measures μ_1 and μ_2 on $\mathbb{R}^d$ are equivalent (denoted by $\mu_1 \sim \mu_2$) if for any Borel set B of $\mathbb{R}^d$, $\mu_1(B) = 0$ if and only if $\mu_2(B) = 0$.

Proposition 4.8. *Let X be a non-degenerate infinitely divisible random vector in $\mathbb{R}^d$ with law μ_X and Lévy measure ν and $\mathcal{P}(\nu)$ be the set of probability measures, μ, in $\mathbb{R}^d$ such that $\nu * \mu << \mu$. Let Y be a non-degenerate random vector in $\mathbb{R}^d$ with law μ_Y such that $\mu_Y \sim \mu_X$. Then, $\mu_Y \in \mathcal{P}(\nu)$.*

Proof of Proposition 4.8. Let B be a Borel set of $\mathbb{R}^d$ such that $\mu_Y(B) = 0$. Then, $\mu_X(B) = 0$ since $\mu_Y \sim \mu_X$. But, $\mu_X \in \mathcal{P}(\nu)$, thus $\nu * \mu_X(B) = 0$. Finally, $\nu * \mu_Y << \nu * \mu_X$, since $\mu_Y \sim \mu_X$, and therefore, $\nu * \mu_Y(B) = 0$, which concludes the proof. □

As a further straightforward corollary, the following result holds true.

Corollary 4.9. *Let X be a non-degenerate infinitely divisible random vector in $\mathbb{R}^d$ without Gaussian component, with law μ_X, Lévy measure ν_X and parameter $b_X \in \mathbb{R}^d$ and let $\mathcal{P}(\nu_X)$ be the set of probability measures, μ, on $\mathbb{R}^d$ such that $\nu_X * \mu << \mu$. Let Y be a non-degenerate infinitely divisible random vector in $\mathbb{R}^d$ without Gaussian component, with law μ_Y, Lévy measure ν_Y and parameter $b_Y \in \mathbb{R}^d$. Assume that $\nu_X \sim \nu_Y$ and that*

$$\int_{\mathbb{R}^d} \left(e^{\Phi(u)/2} - 1\right)^2 \nu_X(du) < +\infty, \quad b_Y - b_X - \int_{\|u\| \leq 1} u(\nu_Y - \nu_X)(du) = 0,$$

where $\exp(\Phi(u)) = d\nu_Y/d\nu_X$, for all $u \in \mathbb{R}^d$. Then, $\mu_Y \in \mathcal{P}(\nu_X)$.

Proof of Corollary 4.9. This is a direct application of Proposition 4.8 together with [44, Theorem 33.1]. □

A Appendix

The aim of this section is to provide technical results (which are often multivariate versions of univariate ones proved in [1]) used throughout the previous sections.

Lemma A.1. *Let X be a non-degenerate self-decomposable random vector in $\mathbb{R}^d$, without Gaussian component, with law μ_X, characteristic function φ and such that $\mathbb{E}\|X\| < \infty$. Assume further that, for any $0 < a < b < +\infty$ the functions k_x given by (2.5) satisfy the following condition*

$$\sup_{x \in S^{d-1}} \sup_{r \in (a,b)} k_x(r) < +\infty. \tag{A.1}$$

Let X_t, $t \geq 0$, be the random vectors each with characteristic functions, φ_t, given, for all $\xi \in \mathbb{R}^d$ by

$$\varphi_t(\xi) = \frac{\varphi(\xi)}{\varphi(e^{-t}\xi)}. \tag{A.2}$$

Then,
(i)

$$\sup_{t>0} \mathbb{E}\|X_t\| < +\infty, \tag{A.3}$$

and,
(ii) for all $\xi \in \mathbb{R}^d$ *and all* $t \in (0,1)$*,*

$$\frac{1}{t}\,|\varphi_t(\xi) - 1| \leq C(\|\xi\|\|\mathbb{E}X\| + \|\xi\| + \|\xi\|^2), \tag{A.4}$$

for some $C > 0$ *independent of* ξ *and* t*.*

Proof of Lemma A.1. Let us start with the proof of (i). First note that, for all $t > 0$,

$$X_t =_d (1 - e^{-t})\mathbb{E}X + Y_t + Z_t,$$

where Y_t and Z_t are independent, with, for all $\xi \in \mathbb{R}^d$,

$$\mathbb{E}e^{i\langle \xi; Y_t\rangle} = \exp\left(\int_{u\in D}\left(e^{i\langle \xi;u\rangle} - 1 - i\langle \xi; u\rangle\right)\nu_t(du)\right),$$
$$\mathbb{E}e^{i\langle \xi; Z_t\rangle} = \exp\left(\int_{u\in D^c}\left(e^{i\langle \xi;u\rangle} - 1 - i\langle \xi; u\rangle\right)\nu_t(du)\right),$$

with ν_t the Lévy measure of X_t. Then, for all $t > 0$,

$$\mathbb{E}\|X_t\| \leq (1 - e^{-t})\mathbb{E}\|X\| + \mathbb{E}\|Y_t\| + \mathbb{E}\|Z_t\|,$$

and from [30, Lemma 1.1],

$$\mathbb{E}\|X_t\| \leq (1 - e^{-t})\mathbb{E}\|X\| + \left(\int_{\|u\|\leq 1}\|u\|^2\nu_t(du)\right)^{1/2} + 2\int_{\|u\|\geq 1}\|u\|\nu_t(du).$$

Now, thanks to the representation (3.5),

$$\int_{\|u\|\leq 1}\|u\|^2\nu_t(du) \leq \int_{\|u\|\leq 1}\|u\|^2\nu(du), \qquad \int_{\|u\|\geq 1}\|u\|\nu_t(du) \leq \int_{\|u\|\geq 1}\|u\|\nu(du).$$

Thus,

$$\sup_{t\geq 0}\mathbb{E}\|X_t\| \leq \mathbb{E}\|X\| + \left(\int_{\|u\|\leq 1}\|u\|^2\nu(du)\right)^{\frac{1}{2}} + 2\int_{\|u\|\geq 1}\|u\|\nu(du) < +\infty.$$

To prove (ii), first note that, for all $\xi \in \mathbb{R}^d$ with $\|\xi\| \neq 0$ and all $t > 0$,

$$\mathbb{E}e^{i\langle \xi; X_t\rangle} - 1 = \int_0^{\|\xi\|}\left\langle \nabla(\varphi_t)\left(s\frac{\xi}{\|\xi\|}\right); \frac{\xi}{\|\xi\|}\right\rangle ds,$$

and thus

$$\left|\mathbb{E}e^{i\langle \xi; X_t\rangle} - 1\right| \leq \|\xi\| \max_{s\in[0,\|\xi\|]}\left\|\nabla(\varphi_t)\left(s\frac{\xi}{\|\xi\|}\right)\right\|.$$

Noting that, for all $\xi \in \mathbb{R}^d$ and all $1 \leq j \leq d$,

$$\partial_j(\varphi_t)(\xi) = \left(i\mathbb{E}X_j(1 - e^{-t}) + i\int_{\mathbb{R}^d} u_j\left(e^{i\langle u;\xi\rangle} - 1\right)\nu_t(du)\right)\varphi_t(\xi),$$

it follows that

$$\left|\mathbb{E}e^{i\langle \xi; X_t\rangle} - 1\right| \leq \|\xi\|(1 - e^{-t})\sum_{j=1}^d \mathbb{E}|X_j| + \sqrt{d}\|\xi\|^2\int_{\|u\|\leq 1}\|u\|^2\nu_t(du)$$

$$+ 2\|\xi\|\sqrt{d}\int_{\|u\|\geq 1}\|u\|\nu_t(du).$$

Next, the polar decomposition of ν_t allows to bound the terms $\int_{\|u\|\leq 1}\|u\|^2\nu_t(du)$ and $\int_{\|u\|\geq 1}\|u\|\nu_t(du)$. Let us start with $\int_{\|u\|\geq 1}\|u\|\nu_t(du)$. By (3.5),

$$\begin{aligned}\int_{\|u\|\geq 1}\|u\|\nu_t(du) &= \int_{S^{d-1}\times(1,+\infty)} r\frac{k_x(r)-k_x(e^t r)}{r}dr\lambda(dx)\\ &= \int_{S^{d-1}}\left(\int_1^{e^t}k_x(r)dr + (1-e^{-t})\int_{e^t}^{+\infty}k_x(r)dr\right)\lambda(dx)\\ &\leq (e^t-1)\sup_{x\in S^{d-1}}|k_x(1^+)| + (1-e^{-t})\int_{\|u\|\geq 1}\|u\|\nu(du),\end{aligned}$$

which is finite in view of (A.1). For $\int_{\|u\|\leq 1}\|u\|^2\nu_t(du)$,

$$\begin{aligned}\int_{\|u\|\leq 1}\|u\|^2\nu_t(du) &= \int_{S^{d-1}\times(0,1)} r(k_x(r)-k_x(e^t r))dr\lambda(dx)\\ &= \int_{S^{d-1}}\left(\int_0^1 r(k_x(r)-k_x(e^t r))dr\right)\lambda(dx)\\ &= \int_{S^{d-1}}\left(-e^{-2t}\int_1^{e^t} rk_x(r)dr + (1-e^{-2t})\int_0^1 rk_x(r)dr\right)\lambda(dx)\\ &\leq (1-e^{-2t})\int_{S^{d-1}\times(0,1)} rk_x(r)dr\lambda(dx).\end{aligned}$$

This concludes the proof of the lemma. □

Lemma A.2. *Let X and Y be two random vectors in $\mathbb{R}^d$ with respective law μ_X and μ_Y. Let $r\geq 1$. Then,*

$$d_{W_r}(\mu_X,\mu_Y) = \sup_{h\in\mathcal{H}_r\cap\mathcal{C}_c^\infty(\mathbb{R}^d)}|\mathbb{E}h(X)-\mathbb{E}h(Y)|. \tag{A.5}$$

Proof of Lemma A.2. Let $r\geq 1$. First, it is clear that

$$d_{W_r}(\mu_X,\mu_Y) \geq \sup_{h\in\mathcal{H}_r\cap\mathcal{C}_c^\infty(\mathbb{R}^d)}|\mathbb{E}h(X)-\mathbb{E}h(Y)|.$$

Now, let $h\in\mathcal{H}_r$ and let $(h_\varepsilon)_{\varepsilon>0}$ be the regularization of h with the Gaussian kernel, namely, for all $x\in\mathbb{R}^d$ and all $\varepsilon>0$,

$$h_\varepsilon(x) := \int_{\mathbb{R}^d} h(x-y)\exp\left(-\frac{\|y\|^2}{2\varepsilon^2}\right)\frac{dy}{(2\pi)^{\frac{d}{2}}\varepsilon^d}.$$

Note that $h_\varepsilon\in\mathcal{C}^\infty(\mathbb{R}^d)$, the space of infinitely differentiable functions on $\mathbb{R}^d$, for all $\varepsilon>0$. Moreover,

$$\|h-h_\varepsilon\|_\infty \leq d\varepsilon, \qquad M_\ell(h_\varepsilon)\leq 1, \quad 0\leq\ell\leq r.$$

Next, let Ψ be a compactly supported infinitely differentiable function with values in $[0,1]$ such that $\mathrm{supp}(\Psi)\subseteq D(0,2)$, the closed Euclidean ball centered at the origin and of radius 2, and such that $\Psi(x)=1$, for all $x\in D$. Then, for any $R\geq 1$ and any $\varepsilon>0$, set, for all $x\in\mathbb{R}^d$,

$$h_{\varepsilon,R}(x) := \Psi\left(\frac{x}{R}\right)h_\varepsilon(x).$$

Then, for X and Y two random vectors on $\mathbb{R}^d$ with respective law μ_X and μ_Y,

$$\begin{aligned}|\mathbb{E}h(X)-\mathbb{E}h(Y)| &\leq |\mathbb{E}h_{\varepsilon,R}(X)-\mathbb{E}h_{\varepsilon,R}(Y)|+2d\varepsilon+\int_{\mathbb{R}^d}\left(1-\Psi\left(\frac{x}{R}\right)\right)d\mu_X(x)\\ &\quad+\int_{\mathbb{R}^d}\left(1-\Psi\left(\frac{x}{R}\right)\right)d\mu_Y(x),\\ &\leq |\mathbb{E}h_{\varepsilon,R}(X)-\mathbb{E}h_{\varepsilon,R}(Y)|+2d\varepsilon+\mathbb{P}\left(\|X\|\geq R\right)+\mathbb{P}\left(\|Y\|\geq R\right).\end{aligned}$$

Now, for $R\geq 1$ such that $\max\left\{\mathbb{P}\left(\|X\|\geq R\right),\mathbb{P}\left(\|Y\|\geq R\right)\right\}\leq\varepsilon$,

$$|\mathbb{E}h(X)-\mathbb{E}h(Y)|\leq|\mathbb{E}h_{\varepsilon,R}(X)-\mathbb{E}h_{\varepsilon,R}(Y)|+(2d+2)\varepsilon.$$

To continue, one needs to estimate the quantities $M_\ell(h_{\varepsilon,R})$, for all $0\leq\ell\leq r$. First, since $h\in\mathcal{H}_r$,

$$M_0(h_{\varepsilon,R}):=\sup_{x\in\mathbb{R}^d}|h_{\varepsilon,R}(x)|\leq 1.$$

But, for $v\in\mathbb{R}^d$ such that $\|v\|=1$ and $x\in\mathbb{R}^d$,

$$\begin{aligned}\mathbf{D}(h_{\varepsilon,R})(v)(x)&=\sum_{i=1}^d v_i\partial_i\left(h_{\varepsilon,R}\right)(x)\\ &=\frac{1}{R}\sum_{i=1}^d v_ih_\varepsilon(x)\partial_i(\Psi)\left(\frac{x}{R}\right)+\Psi\left(\frac{x}{R}\right)\sum_{i=1}^d v_i\partial_i\left(h_\varepsilon\right)(x)\\ &=\frac{h_\varepsilon(x)}{R}\langle\nabla\left(\Psi\right)\left(\frac{x}{R}\right);v\rangle+\Psi\left(\frac{x}{R}\right)\langle\nabla(h_\varepsilon)(x);v\rangle.\end{aligned}$$

Thus, for all $R\geq 1$ and all $\varepsilon>0$

$$M_1(h_{\varepsilon,R})\leq\frac{1}{R}\sup_{x\in\mathbb{R}^d}\|\nabla\left(\Psi\right)(x)\|+1.$$

With a similar reasoning, it follows that $M_\ell(h_{\varepsilon,R})\leq C_{\ell,\Psi}\left(\sum_{k=1}^\ell 1/R^k\right)+1$, for all $1\leq\ell\leq r$, and for some $C_{\ell,\Psi}>0$ only depending on ℓ and Ψ. Hence, the function $\widetilde{h}_{\varepsilon,R}$ defined, for all $x\in\mathbb{R}^d$, by

$$\widetilde{h}_{\varepsilon,R}(x):=\frac{h_{\varepsilon,R}(x)}{\max_{1\leq\ell\leq r}(C_{\ell,\Psi})\left(\sum_{k=1}^r 1/R^k\right)+1},$$

belongs to $\mathcal{H}_r\cap\mathcal{C}_c^\infty(\mathbb{R}^d)$. Finally,

$$\begin{aligned}|\mathbb{E}h(X)-\mathbb{E}h(Y)|&\leq\left(\max_{1\leq\ell\leq r}(C_{\ell,\Psi})\left(\sum_{k=1}^r\frac{1}{R^k}\right)+1\right)\left|\mathbb{E}\widetilde{h}_{\varepsilon,R}(X)-\mathbb{E}\widetilde{h}_{\varepsilon,R}(Y)\right|+(2d+2)\varepsilon\\ &\leq\left(\max_{1\leq\ell\leq r}(C_{\ell,\Psi})\left(\sum_{k=1}^r\frac{1}{R^k}\right)+1\right)\sup_{h\in\mathcal{H}_r\cap\mathcal{C}_c^\infty(\mathbb{R}^d)}|\mathbb{E}h(X)-\mathbb{E}h(Y)|+(2d+2)\varepsilon.\end{aligned}$$

Letting first R tend to $+\infty$ and then ε tend to 0^+ concludes the proof of the lemma. □

The objective of Theorem A.4 below is to prove that the d_{W_1} distance, between the law of X and the law of X_t, decreases exponentially fast as t tends to $+\infty$. For this purpose, for any $r\geq 1$ and any random vectors X and Y, let

$$d_{\widetilde{W}_r}(X,Y)=\sup_{h\in\widetilde{\mathcal{H}}_r}|\mathbb{E}h(X)-\mathbb{E}h(Y)|,\qquad\text{(A.6)}$$

where $\widetilde{\mathcal{H}}_r$ is the set of functions which are r-times continuously differentiable on $\mathbb{R}^d$ such that $\|D^\alpha(f)\|_\infty \leq 1$, for all $\alpha \in \mathbb{N}^d$ with $0 \leq |\alpha| \leq r$. Since, for any $r \geq 1$, $\mathcal{H}_r \subset \widetilde{\mathcal{H}}_r$,

$$d_{W_r}(X,Y) \leq d_{\widetilde{W}_r}(X,Y). \tag{A.7}$$

The next lemma shows that in (A.6), it is enough to take the supremum over smooth compactly supported functions in $\widetilde{H}_r$, $r \geq 1$.

Lemma A.3. *Let X and Y be two random vectors in $\mathbb{R}^d$ with respective law μ_X and μ_Y. Let $r \geq 1$. Then,*

$$d_{\tilde{W}_r}(\mu_X, \mu_Y) = \sup_{h \in \widetilde{\mathcal{H}}_r \cap \mathcal{C}_c^\infty(\mathbb{R}^d)} |\mathbb{E}h(X) - \mathbb{E}h(Y)|. \tag{A.8}$$

Proof of Lemma A.3. By definition,

$$d_{\tilde{W}_r}(\mu_X, \mu_Y) \geq \sup_{h \in \widetilde{\mathcal{H}}_r \cap \mathcal{C}_c^\infty(\mathbb{R}^d)} |\mathbb{E}h(X) - \mathbb{E}h(Y)|.$$

Let $h \in \widetilde{\mathcal{H}}_r$ and let $(h_\varepsilon)_{\varepsilon>0}$ be a regularization by convolution of h, such that $h_\varepsilon \in \mathcal{C}^\infty(\mathbb{R}^d)$ and

$$\|h - h_\varepsilon\|_\infty \leq d\varepsilon, \qquad \|D^\alpha(h_\varepsilon)\|_\infty \leq 1, \quad \alpha \in \mathbb{N}^d, \quad 0 \leq |\alpha| \leq r, \quad r \geq 1.$$

Let ψ be a compactly supported, even, infinitely differentiable function on $\mathbb{R}$ with values in $[0,1]$ such that $\psi(x) = 1$, for $x \in [-1,1]$. Then, for all $M \geq 1$, $\varepsilon > 0$ and $x \in \mathbb{R}^d$ set $\Psi_M(x) = \prod_{i=1}^d \psi(x_i/M)$ and set also,

$$h_{M,\varepsilon}(x) = \Psi_M(x) h_\varepsilon(x).$$

Clearly, by construction, $h_{M,\varepsilon} \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. Then, for all $M \geq 1$ and $\varepsilon > 0$,

$$|\mathbb{E}h(X) - \mathbb{E}h(Y)| \leq |\mathbb{E}h_{M,\varepsilon}(X) - \mathbb{E}h_{M,\varepsilon}(Y)| + 2d\varepsilon + \int_{\mathbb{R}^d} |1 - \Psi_M(x)|\, d\mu_X(x) + \int_{\mathbb{R}^d} |1 - \Psi_M(y)|\, d\mu_Y(y).$$

Choosing $M \geq 1$ large enough,

$$|\mathbb{E}h(X) - \mathbb{E}h(Y)| \leq |\mathbb{E}h_{M,\varepsilon}(X) - \mathbb{E}h_{M,\varepsilon}(Y)| + (2d+2)\varepsilon.$$

Next, by the very definition of $h_{M,\varepsilon}$

$$\|h_{M,\varepsilon}\|_\infty \leq 1,$$

and, moreover, by Leibniz formula, for all $\alpha \in \mathbb{N}^d$ with $1 \leq |\alpha| \leq r$ and $x \in \mathbb{R}^d$,

$$\begin{aligned} |D^\alpha(h_{M,\varepsilon})(x)| &\leq \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} |D^\beta(\Psi_M)(x)||D^{\alpha-\beta}(h_\epsilon)(x)| \\ &\leq |D^\alpha(h_\epsilon)(x)| + \sum_{\beta \leq \alpha,\, \beta \neq 0} \binom{\alpha}{\beta} |D^\beta(\Psi_M)(x)||D^{\alpha-\beta}(h_\epsilon)(x)| \\ &\leq 1 + \sum_{\beta \leq \alpha,\, \beta \neq 0} \binom{\alpha}{\beta} |D^\beta(\Psi_M)(x)|. \end{aligned}$$

Now, for all $\beta \leq \alpha$, $\beta \neq 0$ and $x \in \mathbb{R}^d$,

$$|D^\beta(\Psi_M)(x)| \leq \frac{1}{M^{|\beta|}} \prod_{1 \leq j \leq d} \sup_{x \in \mathbb{R}} |\psi^{(\beta_j)}(x)|.$$

Thus,

$$|D^\alpha(h_{M,\varepsilon})(x)| \leq 1 + C_{d,\alpha} \sum_{\beta \leq \alpha,\, \beta \neq 0} \frac{1}{M^{|\beta|}},$$

for some $C_{d,\alpha} > 0$ only depending on d, α and ψ. This implies that

$$|\mathbb{E}h(X) - \mathbb{E}h(Y)| \leq \left(1 + C_{d,r} \sum_{1 \leq |\alpha| \leq r} \sum_{\beta \leq \alpha,\, \beta \neq 0} \frac{1}{M^{|\beta|}}\right) \sup_{h \in \widetilde{\mathcal{H}}_r \cap C_c^\infty(\mathbb{R}^d)} |\mathbb{E}h(X) - \mathbb{E}h(Y)| + (2d+2)\varepsilon,$$

for some $C_{d,r} > 0$ only depending on d, r and ψ. The conclusion follows by, first taking $M \to +\infty$, and then $\varepsilon \to 0^+$. □

Theorem A.4. *Let X be a non-degenerate self-decomposable random vector in $\mathbb{R}^d$, without Gaussian component, with law μ_X, characteristic function φ and such that $\mathbb{E}\|X\| < \infty$. Assume further that, for any $0 < a < b < +\infty$ the functions k_x given by* (2.5) *satisfy the following condition*

$$\sup_{x \in S^{d-1}} \sup_{r \in (a,b)} k_x(r) < +\infty. \tag{A.9}$$

Let X_t, $t > 0$ be random vectors each with law μ_{X_t}, with characteristic function φ_t, given, for all $\xi \in \mathbb{R}^d$ by

$$\varphi_t(\xi) = \frac{\varphi(\xi)}{\varphi(e^{-t}\xi)}. \tag{A.10}$$

Then, for $t > 0$,

$$d_{W_1}(\mu_{X_t}, \mu_X) \leq C_d e^{-\frac{t}{2^{d+1}(d+1)}}, \tag{A.11}$$

for some $C_d > 0$ independent of t.

Proof of Theorem A.4. Step 1: Let $r \geq 2$ and let $h \in \widetilde{\mathcal{H}}_{r-1}$. Let $(h_\varepsilon)_{\varepsilon > 0}$ be a regularization by convolution of h such that

$$\|h - h_\varepsilon\|_\infty \leq d\varepsilon, \quad \|D^\alpha(h_\varepsilon)\|_\infty \leq 1, \quad 0 \leq |\alpha| \leq r-1.$$

For $\alpha \in \mathbb{N}^d$ such that $|\alpha| = r$, let us estimate $\|D^\alpha(h_\varepsilon)\|_\infty$. By definition, for all $x \in \mathbb{R}^d$,

$$h_\varepsilon(x) = \int_{\mathbb{R}^d} h(y) \exp\left(-\frac{\|x-y\|^2}{2\varepsilon^2}\right) \frac{dy}{(2\pi)^{\frac{d}{2}} \varepsilon^d}.$$

Now, by Rodrigues formula, for all $j \in 1, ..., d$,

$$\partial_{x_j}^{\alpha_j}\left(\exp\left(-\frac{x_j^2}{2}\right)\right) = (-1)^{\alpha_j} H_{\alpha_j}(x_j) \exp\left(-\frac{x_j^2}{2}\right).$$

where H_{α_j} is the Hermite polynomial of degree α_j. Thus, for all $\alpha \in \mathbb{N}^d$ and $x \in \mathbb{R}^d$,

$$D^\alpha\left(\exp\left(-\frac{\|x\|^2}{2}\right)\right) = (-1)^\alpha H_\alpha(x) \exp\left(-\frac{\|x\|^2}{2}\right),$$

where $H_\alpha(x)=\prod_{j=1}^d H_{\alpha_j}(x_j)$. Hence, for all $\alpha\in\mathbb{N}^d$ such that $|\alpha|=r$, for all $x\in\mathbb{R}^d$ and for some $\beta\in\mathbb{N}^d$ such that $|\beta|=r-1$ and $\alpha-\beta\geq 0$

$$D^\alpha(h_\varepsilon)(x)=\int_{\mathbb{R}^d} D^\beta(h)(y)D^{\alpha-\beta}\left(\exp\left(-\frac{\|x-y\|^2}{2\varepsilon^2}\right)\right)\frac{dy}{(2\pi)^{\frac{d}{2}}\varepsilon^d},$$
$$D^\alpha(h_\varepsilon)(x)=\frac{(-1)}{\varepsilon}\int_{\mathbb{R}^d} D^\beta(h)(y)H_{\alpha-\beta}\left(\frac{x-y}{\varepsilon}\right)\exp\left(-\frac{\|x-y\|^2}{2\varepsilon^2}\right)\frac{dy}{(2\pi)^{\frac{d}{2}}\varepsilon^d}.$$

Then,

$$\begin{aligned}\|D^\alpha(h_\varepsilon)\|_\infty&\leq\frac{1}{\varepsilon}\int_{\mathbb{R}^d}|H_{\alpha-\beta}(y)|\exp\left(-\frac{\|y\|^2}{2}\right)\frac{dy}{(2\pi)^{\frac{d}{2}}}\\&\leq C_\alpha\varepsilon^{-1}\leq C_r\varepsilon^{-1},\end{aligned}$$

for some $C_\alpha>0, C_r>0$ depending only on α, on d and on r. Let Z and Y be two random vectors with respective law μ_Z and μ_Y such that $d_{\widetilde{W}_r}(Z,Y)<1$. Then,

$$|\mathbb{E}h(Z)-\mathbb{E}h(Y)|\leq 2d\varepsilon+|\mathbb{E}h_\varepsilon(Z)-\mathbb{E}h_\varepsilon(Y)|.$$

Choosing $\varepsilon\in(0,C_r)$,

$$\begin{aligned}|\mathbb{E}h(Z)-\mathbb{E}h(Y)|&\leq 2d\varepsilon+\frac{C_r}{\varepsilon}d_{\widetilde{W}_r}(Z,Y)\\&\leq\max(2d,C_r)\left(\varepsilon+\varepsilon^{-1}d_{\widetilde{W}_r}(Z,Y)\right).\end{aligned}$$

Taking $\varepsilon\leq C_r/(1+C_r)\sqrt{d_{\widetilde{W}_r}(Z,Y)}$ yields,

$$d_{\widetilde{W}_{r-1}}(Z,Y)\leq\tilde{C}_r\sqrt{d_{\widetilde{W}_r}(Z,Y)},$$

for some $\tilde{C}_r>0$ only depending on r and on d. Now, let Z and Y be two random vectors such that $d_{\widetilde{W}_2}(Z,Y)<1$. Then, thanks to (2.15), $d_{\widetilde{W}_m}(Z,Y)<1$, for all $2\leq m\leq r$. By induction, we get

$$d_{\widetilde{W}_1}(Z,Y)\leq\overline{C}_r\left(d_{\widetilde{W}_r}(Z,Y)\right)^{\frac{1}{2^{r-1}}},\tag{A.12}$$

for some $\overline{C}_r>0$ only depending on r and on d.

Step 2: Let g be an infinitely differentiable function with compact support contained in the closed Euclidean ball centered at the origin of radius $R+1$, for some $R>0$. Then by Fourier inversion and Fubini theorem, for all $t>0$,

$$\begin{aligned}|\mathbb{E}g(X)-\mathbb{E}g(X_t)|&\leq e^{-t}\mathbb{E}\|X\|\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}|\mathcal{F}(g)(\xi)|\|\xi\|d\xi\\&\leq e^{-t}\mathbb{E}\|X\|\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}|\mathcal{F}(g)(\xi)|\frac{(1+\|\xi\|)^{d+2}}{(1+\|\xi\|)^{d+2}}\|\xi\|d\xi\\&\leq e^{-t}\mathbb{E}\|X\|\sup_{\xi\in\mathbb{R}^d}\left(|\mathcal{F}(g)(\xi)|(1+\|\xi\|^{d+2})\right)\left(\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\frac{\|\xi\|d\xi}{(1+\|\xi\|)^{d+2}}\right).\end{aligned}$$

Moreover, for all $p\geq 2$

$$\sup_{\xi\in\mathbb{R}^d}\left(|\mathcal{F}(g)(\xi)|(1+\|\xi\|^p)\right)\leq C_d(R+1)^d\left(\|g\|_\infty+\max_{1\leq j\leq d}\|\partial_j^p(g)\|_\infty\right),$$

for some $C_d > 0$ depending on the dimension d only. Thus, for all $t > 0$

$$|\mathbb{E}g(X) - \mathbb{E}g(X_t)| \leq \tilde{C}_d e^{-t}\mathbb{E}\|X\|(R+1)^d \left(\|g\|_\infty + \max_{1\leq j\leq d}\|\partial_j^{d+2}(g)\|_\infty\right). \tag{A.13}$$

Step 3: Let $h \in \mathcal{C}_c^\infty(\mathbb{R}^d)\cap\widetilde{\mathcal{H}}_{d+2}$. Let Ψ_R be a compactly supported infinitely differentiable function on $\mathbb{R}^d$ whose support is contained in the closed Euclidean ball centered at the origin of radius $R+1$, with values in [0,1] and such that $\Psi_R(x) = 1$, for all x such that $\|x\| \leq R$. Then, for all $t > 0$

$$\begin{aligned}|\mathbb{E}h(X) - \mathbb{E}h(X_t)| \leq & |\mathbb{E}h(X)\psi_R(X) - \mathbb{E}h(X_t)\Psi_R(X_t)| + |\mathbb{E}h(X)(1-\Psi_R(X))| \\ & + |\mathbb{E}h(X_t)(1-\Psi_R(X_t))|.\end{aligned}$$

Next, note that

$$\begin{aligned}|\mathbb{E}h(X_t)(1-\Psi_R(X_t))| &\leq \int_{\mathbb{R}^d}(1-\Psi_R(x))d\mu_t(x) \\ &\leq \mathbb{P}\left(\|X_t\| \geq R\right) \\ &\leq \frac{\mathbb{E}\|X_t\|}{R} \\ &\leq \frac{1}{R}\sup_{t>0}\mathbb{E}\|X_t\|,\end{aligned}$$

using Lemma A.1. A similar bound holds true for $|\mathbb{E}h(X)(1-\Psi_R(X))|$. Moreover, from (A.13),

$$|\mathbb{E}h(X) - \mathbb{E}h(X_t)| \leq \frac{C_d}{R} + \tilde{C}_d e^{-t}\mathbb{E}\|X\|(R+1)^d\left(\|h\Psi_R\|_\infty + \max_{1\leq j\leq d}\|\partial_j^{d+2}(h\Psi_R)\|_\infty\right),$$

for some constant C_d depending on d. Now,

$$\|h\Psi_R\|_\infty \leq 1,$$

and, by taking for Ψ_R an appropriate tensorization of one dimensional bump functions ψ_R,

$$\max_{1\leq j\leq d}\|\partial_j^{d+2}(h\Psi_R)\|_\infty \leq \tilde{D},$$

for some $\tilde{D} > 0$ independent of R and h. Hence,

$$|\mathbb{E}h(X) - \mathbb{E}h(X_t)| \leq C_d\left(\frac{1}{R} + (R+1)^d e^{-t}\mathbb{E}\|X\|\right).$$

Choosing $R = e^{t/(d+1)}$, for all $t > 0$, it follows that

$$d_{\widetilde{W}_{d+2}}(X, X_t) \leq \tilde{C}_d e^{-\frac{t}{d+1}},$$

for some $\tilde{C}_d > 0$, and from (A.12) with $r = d+2$,

$$d_{\widetilde{W}_1}(X, X_t) \leq \overline{C}_d\left(d_{\widetilde{W}_{d+2}}(X, X_t)\right)^{\frac{1}{2^{d+1}}} \leq C_d e^{-\frac{t}{2^{d+1}(d+1)}}.$$

The inequality (A.7) concludes the proof of the theorem. □

References

[1] B. Arras and C. Houdré. On Stein's method for infinitely divisible laws with finite first moment. *To appear: SpringerBriefs in Probability and Mathematical Statistics*, arXiv:1712.10051, 2019.

[2] A. D. Barbour. Stein's method for diffusion approximations. *Probab. Theory Relat. Fields.* 84(3), 297-322, 1990. MR-1035659

[3] D. Bakry, I. Gentil and M. Ledoux. *Analysis and Geometry of Markov diffusion operators*. Springer, 2014. MR-3155209

[4] V. I. Bogachev. *Measure Theory* Springer, Vol. 1, 2007. MR-2267655

[5] L. Bondesson. *Generalized Gamma Convolutions and Related Classes of Distributions and Densities* Lectures Notes in Statistics, Springer, 1992. MR-1224674

[6] A.A. Borovkov and S.A. Utev. On an inequality and a related characterisation of the normal distribution. *Theory Probab. Appl.*, 28:219–228, 1984. MR-0700206

[7] T. Cacoullos, V. Papathanasiou and S. Utev. Variational inequalities with examples and an application to the central limit theorem. *Ann. Probab.*, 22:1607–1618, 1994. MR-1303658

[8] S. Chatterjee. A new method of normal approximation. *Ann. Probab.* 36(4):1584–1610, 2008. MR-2435859

[9] S. Chatterjee. Fluctuations of eigenvalues and second order Poincaré inequalities. *Probab. Theory Related Fields* 143:1–40, 2007. MR-2449121

[10] S. Chatterjee. A new approach to strong embeddings. *Probab. Theory Related Fields* 152:231–264, 2012. MR-2875758

[11] S. Chatterjee and E. Meckes. Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.* 4, 257–283, 2008. MR-2453473

[12] L.H.Y. Chen. Poincaré-type inequalities via stochastic integrals. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 69:251–277, 1985. MR-0779459

[13] L.H.Y. Chen, L. Goldstein, and Q.M. Shao. *Normal Approximation by Stein's Method* Probability and its Application, Springer, Heidelberg, 2011. MR-2732624

[14] L. H. Y. Chen and J. H. Lou. Characterization of probability distributions by Poincaré-type inequalities. *Ann. Inst. Henri Poincaré Probab. Stat.* 23(1): 91–110, 1987. MR-0877386

[15] Z.-Q. Chen and X. Zhang. Heat kernels and analyticity of non-symmetric jump diffusion semigroups. *Probab. Theory Relat. Fields* 165:267–312, 2016. MR-3500272

[16] T. A. Courtade, M. Fathi and A. Pananjady. Existence of Stein kernels under a spectral gap, and discrepancy bounds. *To appear: Ann. Inst. Henri Poincaré Probab. Stat.*, arXiv:1703.07707, 2018.

[17] L. C. Evans. *Partial Differential Equations* Graduate Studies in Mathematics, Vol. 19, 2002. MR-2597943

[18] M. Fathi. Stein kernel and moment maps. *To appear: Ann. Probab.*, arXiv:1804.04699, 2018.

[19] B. V. Gnedenko and A. N. Kolmogorov. *Limit Distributions for Sums of Independent Random Variables*. Translated from the Russian, annotated and revised by K.-L. Chung, with appendices by J. L. Doob and P. L. Hsu. Addison-Wesley, revised edition, 1968. MR-0233400

[20] L. Goldstein and Y. Rinott. Multivariate normal approximations by Stein's method and size bias couplings. *J. Appl. Probab.* 33(1):1–17, 1996. MR-1371949

[21] J. Gorham, A. B. Duncan, S.J. Vollmer and L. Mackey. Measuring sample quality with diffusions. *To appear: Ann. Appl. Probab.* arXiv:1611.06972, 2016.

[22] F. Götze. On the rate of convergence in the multivariate CLT. *Ann. Probab.* 19(2):724–739, 1991. MR-1106283

[23] C. Houdré, V. Pérez-Abreu and D. Surgailis. Interpolation, correlation identities and inequalities for infinitely divisible variables. *J. Fourier Anal. Appl.* 4(6):651–668, 1998. MR-1665993

[24] A. Ya. Khintchine. *Limit Laws for Sums of Independent Random Variables.* ONTI, Moscow–Leningrad (in Russian), 1938.

[25] M. Ledoux, I. Nourdin and G. Peccati. Stein's method, logarithmic Sobolev and transport inequalities. *Geom. Funct. Anal.*, 25:256–306, 2015. MR-3320893

[26] P. Lévy. *Théorie de l'Addition des Variables Aléatoires* 2nd, Gauthier-Villars, Paris, (1st ed. 1937), 1954.

[27] C. Ley, G. Reinert and Y. Swan. Stein's method for comparison of univariate distributions. *Probab. Surveys*, 14:1–52, 2017. MR-3595350

[28] M. Loève. *Probability Theory, I and II* Springer, New-York. (1st ed., Van Nostrand, Princeton, NJ, 1955), 1977,1978. MR-651017 MR-651018

[29] L. Mackey and J. Gorham. Multivariate Stein factors for a class of strongly log-concave distributions. *Electron. Commun. Probab.*, 21(56):1–14, 2016. MR-3548768

[30] M. B. Marcus and J. Rosinski. L^1-norm of infinitely divisible random vectors and certain stochastic integrals. *Electron. Commun. Probab.* 6:15–29, 2001. MR-1817886

[31] E. Meckes. On Stein's method for multivariate normal approximation. *High Dimensional Probability V: The Luminy Volume*, Institute of Mathematical Statistics, 153–178, 2009. MR-2797946

[32] I. Nourdin and G. Peccati. Stein's method on Wiener chaos. *Probab. Theory Related Fields*, 145(1):75–118, 2008. MR-2520122

[33] I. Nourdin and G. Peccati. *Normal Approximations with Malliavin calculus*, volume 192 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2012. From Stein's method to universality. MR-2962301

[34] I. Nourdin, G. Peccati and A. Réveillac. Multivariate normal approximation using Stein's method and Malliavin calculus. *Ann. Inst. H. Poincaré Probab. Statist.*, 46(1):45–58, 2010. MR-2641769

[35] I. Nourdin, G. Peccati and Y. Swan. Entropy and the fourth moment phenomenon. *J. Funct. Anal.*, 266:3170–3207, 2014. MR-3158721

[36] F. W. J. Olver, D. W. Lozier, R. F. Boisvert and C. W. Clark. *NIST Handbook of Mathematical Functions* Cambridge University Press, 2010. MR-2723248

[37] V. Pérez-Abreu and R. Stelzer. Infinitely divisible multivariate and matrix gamma distributions. *J. Multivariate Anal.*, 130:155–175, 2014. MR-3229530

[38] V.V. Petrov *Limit Theorems of Probability Theory.* Oxford University Press, Oxford, 1995. MR-1353441

[39] M. Raic. A multivariate CLT for decomposable random vectors with finite second moments. *J. Theoret. Probab.* 17(3), 573–603, 2004. MR-2091552

[40] G. Reinert and A. Röllin. Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Ann. Probab.*, 37(6):2150–2173, 2009. MR-2573554

[41] Y. Rinott and V. Rotar. A multivariate CLT for local dependence with $n^{-1/2}\log(n)$ rate and applications to multivariate graph related statistics. *J. Multivariate Anal.*, 56:333–350, 1996. MR-1379533

[42] A. Röllin. Stein's method in high dimensions with applications. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(2):529–549, 2013. MR-3088380

[43] N.F. Ross Fundamentals of Stein's method. *Probab. Surv.* 8:210–293, 2011. MR-2861132

[44] K-I. Sato, *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Corrected Printing with Supplements, 2015. MR-3185174

[45] C. Stein, A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, 583–602, 1972. MR-0402873

[46] C. Stein. Approximate computation of expectations. *Institute of Mathematical Statistics Lecture Notes* Monograph Series, 7. Institute of Mathematical Statistics, 1986. MR-0882007

[47] F.W. Steutel and K. Van Harn. *Infinite Divisibility of Probability Distributions on the Real Line* CRC Press, 2003. MR-2011862

[48] K. Takano. The Lévy representation of the characteristic function of the probability density $Ce^{-\|x\|}dx$. *Bull. Fac. Sci., Ibaraki Univ.* 20:61–65,1988. MR-0951729

[49] K. Takano. The Lévy representation of the characteristic function of the probability density $\Gamma(m+d/2)\left(\pi^2\Gamma(m)\right)^{-1}(1+|x|^2)^{-m-d/2}$. *Bull. Fac. Sci., Ibaraki Univ.* 21:21–27,1989. MR-1001807

[50] K. Takano. On mixtures of the normal distribution by the generalized gamma convolutions. *Bull. Fac. Sci., Ibaraki Univ.* 21:29–41,1989. MR-1001808

[51] S.A. Utev. Probability problems connected with a certain integrodifferential inequality. *Sib. Math. J.* 30(3):490–493, 1989. MR-1010851

Acknowledgments. We are grateful to an anonymous referee and to an associate editor for insightful remarks which led to improvements in the presentation of this manuscript.