Distances between probability distributions via characteristic functions and biasing

Benjamin Arras, Guillaume Mijoule, Guillaume Poly and Yvik Swan

Abstract

In a spirit close to classical Stein's method, we introduce a new technique to derive first order ODEs on differences of characteristic functions. Then, using concentration inequalities and Fourier transform tools, we convert this information into sharp bounds for the so-called *smooth Wasserstein metrics* which frequently arise in Stein's method theory. Our methodolgy is particularly efficient when the target density and the object of interest both satisfy biasing equations (including the zero-bias and size-bias mechanisms). In order to illustrate our technique we provide estimates for: (i) the Dickman approximation of the sum of positions of records in random permutations, (ii) quantitative limit theorems for convergence towards stable distributions, (iii) a general class of infinitely divisible distributions and (iv) distributions belonging to the second Wiener chaos, that is to say, in some situations impervious to the standard procedures of Stein's method. Except for the stable distribution, our bounds are sharp up to logarithmic factors. We also describe a general argument and open the way for a wealth of refinements and other applications.

Keywords: Limit theorems, Characteristic function, Bias transformations, Dickman distribution, Stable law, Wiener chaos.

MSC 2010: 60E07, 60E10, 60F05, 60G50.

Contents

1 Introduction

At first thought the most meaningful way to compare two probability distributions μ and ν on $\mathbb{R}^d$ is by means of the Kolmogorov distance

$$\mathrm{Kol}(\mu,\nu) = \sup_{x\in\mathbb{R}^d} |\mu\left((-\infty,x]\right) - \nu\left((-\infty,x]\right)| \tag{1.1}$$

which measures the maximal difference between their cumulative distribution functions. Proximity in Kolmogorov distance is very demanding and it is often also relevant to express probabilistic discrepancies in terms of other probability distances; see for instance [13] for an overview and relevant references. In the sequel we will be interested on expressing proximity in terms of integral probability metrics of the form

$$d_{\mathcal{H}}(\mu,\nu) = \sup_{h\in\mathcal{H}} \left|\int h(d\mu - d\nu)\right| \tag{1.2}$$

with $\mathcal{H}$ a measure separating class of functions. Obviously (1.1) is a particular case of (1.2) where $\mathcal{H}$ is the collection of indicators of semi-open rectangles. Other distances of the form (1.2) include the Wasserstein (or earth-mover's) distance, Fortet-Mourier distance and several others. We refer the reader to [22, Appendix C] for an overview. Our main focus in this paper will be the following particular case of (1.2) (see also (4.1) in the Appendix).

Definition 1. *Let $p \geq 1$ and $\mathcal{H}_p = \{h \in \mathcal{C}^p(\mathbb{R}) \mid ||h^{(k)}||_\infty \leq 1, \forall\, 0 \leq k \leq p\}$. Then*

$$\mathcal{W}_p(\mu,\nu) = \sup_{h\in\mathcal{H}_p} \left|\int h(d\mu - d\nu)\right| \tag{1.3}$$

is the p-smooth Wasserstein distances between μ and ν.

It is a well know fact that for every $p \geq 0$, the p smooth Wasserstein distance is a metric for the convergence in distribution and that $(\mathcal{W})_p$ increases with p. As a matter of fact, the lower is p, the stronger is the bound. Direct computation of (1.2) or (1.3) is generally not tractable and several approaches have been devised in order to produce tight fathomable bounds on these quantities under a wide variety of assumptions on μ and ν. One of the more remarkable such techniques is known as the Stein (or Chen-Stein) method, which can be summarized as follows.

We first assume that there exist two functional spaces, say E_1 and E_2, and some linear operator $T : E_1 \longrightarrow E_2$ such that

$$\nu = \mu \Longleftrightarrow \forall f \in E_1, \int T[f]d\nu = 0, \tag{1.4}$$

i.e. the action of operator T on E_1 *characterizes* the distribution μ. If there exists $\epsilon \ll 1$ such that

$$\sup_{f\in B(E_1)} \left|\int T[f]d\nu\right| \leq \epsilon, \tag{1.5}$$

for $B(E_1)$ some suitably chosen subset of E_1 (e.g. the unit ball associated to the topology on E_1) then it is reasonable to suppose that $\nu \approx \mu$ in some sense. Indeed if $\mathcal{H} = \{T[f], f \in B(E_1)\} \subset E_2$ is measure determining for μ then, using (1.5), we deduce in turn that $d_{\mathcal{H}}(\mu, \nu) < \epsilon$. Simple as it is, this heuristic has proven to be particularly effective in a wide variety of settings (including the all important Gaussian and Poisson cases), particularly in situations wherein more traditional methods do not provide satisfactory bounds. We refer to the monographs [29, 7, 22], to the regularly maintained webpage https://sites.google.com/site/malliavinstein for applications around the so-called fourth moment theorem, as well as to the website https://sites.google.com/site/steinsmethod/ for a broad overview of the different target distributions towards which the method has been setup.

Despite its many successes, Stein's method is by no means a panacea. One of its main defaults comes from the fact that, in order to produce meaningful probabilistic bounds on quantities such as (1.1), (1.2) or (1.3), one first fixes a suitable collection $\mathcal{H} \subset E_2$ (e.g. indicator functions, Lipschitz functions, $\mathcal{H}_p$, etc.), then considers $B(E_1) = T^{-1}(\mathcal{H})$ and the aim becomes that of obtaining a bound such as (1.5) for a fixed collection $B(E_1)$ depending on $\mathcal{H}$. This step in turn requires solving equations of the form $T[f] = h$ and estimating $\|f\|_{E_1}$ with respect to $\|h\|_{E_2}$. This is an extremely difficult problem. It has been solved under specific assumptions on the target distribution μ, see e.g. [11, 10, 17, 28, 27]. There also exist a wide variety of target distributions whose operators T are known but for which the inversion step and consequential bound (1.5) are not tractable, see e.g. [1, 12]. Similar hurdles arise also for stable distributions ([9]). Whenever no good bound on $\|f\|_{E_1}$ is available, Stein's method fails.

In this paper we revisit Stein's method under a new angle and propose a way to set it up so as to provide approximations in smooth metrics (1.3) while bypassing entirely the need to tackle the Stein equations $T[f] = h$. The key to our approach is to start from a characterizing equation (1.4) but then to exploit properties of the characteristic functions. Working with characteristic functions and Stein operators has shown its efficiency in the case of central limit theorems for sums of dependent random variables, see [31]. Here we identify a very general assumption on the target μ under which our method provides near-optimal bounds for a stupefyingly wide class of target distributions including, but not restricted to, stable distributions, distributions belonging to the second chaos as well as compound Poisson distributions.

Let μ be the distribution of a random variable X with characteristic function ϕ, and let us aim at computing $\mathcal{W}_n(X, W) = \mathcal{W}_n(\mu, \nu)$ for ν the distribution of some random variable W with characteristic function ψ. The first step of our approach is the following new representation of smooth Wasserstein distance in terms of a difference of characteristic functions (see Theorem 1 for a complete statement).

Theorem. *Let X, W be random variables with respective characteristic functions ϕ and ψ such that*

(H1) $\exists\, p \in \mathbb{N}$, $\exists C' > 0$ *and* $0 < \epsilon < 1$ *such that* $\forall \xi \in \mathbb{R}, \quad |\psi(\xi) - \phi(\xi)| \le C'\epsilon|\xi|^p$.

(H2) $\exists\, \lambda > 0, \exists C > 0, \exists \alpha > 0 \forall A > 0$, $\mathbb{P}[|X| > A] \le Ce^{-\lambda A^\alpha}$ *and* $\mathbb{P}[|W| > A] \le Ce^{-\lambda A^\alpha}$.

Then, there exists a constant $K = K(p, C, C', \lambda, \alpha)$ *such that*

$$\mathcal{W}_{p+1}(X, W) \le K\epsilon|\ln(\epsilon)|^{\frac{1}{2\alpha}}. \tag{1.6}$$

Our next main idea is to work under the assumption of the existence of a well-defined and tractable *biasing mechanism* $X \mapsto X^*$ which characterizes the law of X. Famous examples are the size bias distribution of positive random variables (see e.g. [4] for a very enlightening overview) or the zero bias transformation of centered distributions introduced in [15]. These two examples are by no means the only such transformations and most classical distributions satisfy some form of biasing assumption, see [16, 8]. The gist of our approach goes as follows: given the transformation $X \mapsto X^\star$, we define by analogy a transformation $W \mapsto W^\star$ and exploit this in order to provide bounds of the form

$$|\psi(\xi) - \phi(\xi)| \leq C|\xi|^p \mathbb{E}\,|X^\star - W^\star| \tag{1.7}$$

for some $C, p > 0$. Plugging (1.7) into (1.6) (with $\epsilon = \mathbb{E}\,|X^\star - W^\star|$) leads to a very general family of near-optimal bounds on smooth Wasserstein distances.

Before proceeding to the more technical results we first illustrate our approach under an additive size-bias assumption.

1.1 Illustration of our approach under an additive size-bias assumption

Let $\mu = P_X$ and $\nu = P_W$ with X and W two positive real random variables with finite means both equal to $\lambda > 0$. Suppose also that X admits an additive size-bias distribution, i.e. there exists $X^\star$ independent of X such that

$$\mathbb{E}\,[Xf(X)] = \lambda \mathbb{E}\,[f(X + X^\star)] \tag{1.8}$$

for all f for which both expectations exist. Conditions under which (1.8) hold are thoroughly addressed e.g. in [3]. Then plugging $f(x) = e^{itx}$ into (1.8) we deduce that the characteristic functions of X and $X^\star$ are related by the ODE

$$(-i)\phi_X'(t) = \lambda \phi_X(t)\phi_{X^\star}(t) \tag{1.9}$$

(here independence of X and $X^\star$ is crucial). Next construct a coupling (W, X) such that $X^\star$ can be chosen independent of W and such that one can also construct a coupling $(W, W^\star)$ which satisfies the relationship

$$\mathbb{E}\,[Wf(W)] = \lambda \mathbb{E}\,[f(W + W^\star)] \tag{1.10}$$

for all bounded test functions f. Then again plugging $f(x) = e^{itx}$ this time into (1.10) we deduce after straightforward manipulations:

$$(-i)\phi_W'(t) = \lambda \phi_W(t)\phi_{X^\star}(t) + \lambda \mathbb{E}\left[e^{itW}\left(e^{itW^\star} - e^{itX^\star}\right)\right]. \tag{1.11}$$

Subtracting (1.9) from (1.11) and letting $\Delta(t) = \phi_X(t) - \phi_W(t)$ we obtain the ODE

$$(-i)\Delta'(t) = \lambda \phi_{X^\star}(t)\Delta(t) + \psi(t) \tag{1.12}$$

where $\psi(t) = \lambda \mathbb{E}\left[e^{iW}\left(e^{itW^\star} - e^{itX^\star}\right)\right]$ which, thanks to (1.9), is easily solved to yield

$$\Delta(t) = \lambda \phi_X(t) \int_0^t i\psi(s)/\phi_X(s) ds. \tag{1.13}$$

Under adequate assumptions on ϕ_X, we can obtain the following estimate :

$$|\Delta(t)| \leq Ct^2 \mathbb{E}\left|W^\star - X^\star\right|. \tag{1.14}$$

for some constant $C > 0$. We are now in a position to apply (1.6) with $p = 2$ to deduce the existence of a universal constant K such that

$$\mathcal{W}_3(W, X) \leq KT\sqrt{1 + \log\left[1/T\right]/\lambda} \tag{1.15}$$

with $T = \mathbb{E}\left|W^\star - X^\star\right|$.

Example 1. *For the sake of illustration consider the simple situation where* $W = \sum_{i=1}^n W_i$ *is a sum of independent indicators* $W_i \sim \mathrm{Bern}(p_i)$ *and* $X \sim \mathrm{Poisson}(\sum_i p_i)$*. Then by Chen's identity for the Poisson distribution we know that* $\mathbb{E}\left[Xf(X)\right] = \lambda \mathbb{E}\left[f(X+1)\right]$ *so that* X *satisfies* (1.9) *with* $X^\star = 1$*. Similarly let* I *be randomly chosen* $\{1, \ldots, n\}$ *with probability* $p_i / \sum_i p_i$*. Then* $\mathbb{E}\left[Wf(W)\right] = \lambda \mathbb{E}\left[f(W + 1 - X_I)\right]$ *so that* W *satisfies* (1.11) *with* $W^\star = 1 - X_I$*. Then* $\mathbb{E}\left|W^\star - X^\star\right| = \mathbb{E}\left[X_I\right] = \sum_i p_i^2 / \sum_i p_i$ *and thus* (1.15) *leads to bounds of the correct order up to a supplementary log factor.*

As we will show in Section 3, sufficiently dexterous tweaking of the philosophy outlined above will lead to new informative bounds on stochastic approximations for target distributions which were out of reach of current technology. For instance, we will return to the size-bias assumption (1.8) to show that (1.15) leads to improvements on current available bounds for Dickman approximation of rank statistics, see (3.9) in Section 3.1.

Remark 1. *For the sake of simplicity, we mainly restrict our exposition to the size-bias and zero bias framework. We emphasize however that a much wider class of distributions could be considered, namely those that satisfy an equation of the type*

$$\mathbb{E}[Xf(X)] = \mathbb{E}\left[\sum_i [a_i X + b_i] f^{(m_i)}(X + T_i)\right],$$

where the T_i*'s are independent of* X*. Indeed, in such a case, the characteristic function of* X *satisfies a first order differential equation which can be easily solved.*

1.2 Outline of the paper

The outline of the paper is as follows: in Section 2 we provide a complete statement and a proof of our main Theorem allowing to transfer information on proximity between characteristic functions into information on proximity in smooth Wasserstein distances. In Section 3 we discuss several applications: Section 3.1 contains bounds on approximation by Dickman distribution with applications towards the so-called src statistic; Section 3.3 contains an illustration on Gaussian target under a zero-bias assumption; Section 5 contains the difficult case of approximation towards a symmetric stable distributions and Section 3.5 the case of a target belonging to the second Wiener chaos. Finally an Appendix contains the more technical results. In all cases except Section 3.3 our bounds are the first of their kind and yield, up to the log factor, optimal rates of convergence. We stress that this log factor is irrelevant in all numerical illustrations.

2 Bounds on characteristic functions and smooth Wasserstein

The main result of this Section is the following transfer principle:

Theorem 1. *Let X, Y be random variables with respective characteristic functions ϕ and ψ. We make the following hypothesis :*

(H1) $\exists\, p \in \mathbb{N}$, $C' > 0$ and $0 < \epsilon < 1$ such that

$$\forall \xi \in \mathbb{R}, \quad |\psi(\xi) - \phi(\xi)| \leq C'\epsilon|\xi|^p. \tag{2.1}$$

(H2) $\exists\, \lambda > 0, \exists C > 0, \alpha > 0 \forall A > 0$, $\mathbb{P}[|X| > A] \leq Ce^{-\lambda A^\alpha}$ and $\mathbb{P}[|Y| > A] \leq Ce^{-\lambda A^\alpha}$.

Then,

$$\mathcal{W}_{p+1}(X,Y) \leq 2\epsilon \left[C + \frac{2^{p-1}\sqrt{10}C'C_p}{\pi} \sqrt{1 + (\lambda^{-1}\ln(1/\epsilon))^{\frac{1}{\alpha}}} \right], \tag{2.2}$$

where C_p is a constant depending only on p. In particular, one can take $C_0 = 3$, $C_1 = 12$, $C_2 = 100$.

Proof. Let $f \in \mathcal{C}_c^\infty(\mathbb{R})$, with support in $[-M-1, M+1]$, and $\hat{f}$ its Fourier transform. f is in the Schwarz space $\mathcal{S}(\mathbb{R})$, so that f equals the inverse Fourier transform of $\hat{f}$. We have from Fubini's theorem that

$$\begin{aligned}
\mathbb{E}[f(X) - f(Y)] &= \frac{1}{2\pi}\mathbb{E}\left[\int_{\mathbb{R}} (e^{iX\xi} - e^{iY\xi})\hat{f}(\xi)d\xi\right] \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \mathbb{E}(e^{iX\xi} - e^{iY\xi})\hat{f}(\xi)d\xi \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} (\psi(\xi) - \phi(\xi))\hat{f}(\xi)d\xi \\
&\leq \frac{C'\epsilon}{2\pi}\int_{\mathbb{R}} |\xi|^p|\hat{f}(\xi)|d\xi.
\end{aligned}$$

We use the Cauchy-Schwarz inequality and Plancherel's identity to get

$$\begin{aligned}
\int_{\mathbb{R}} |\xi|^p|\hat{f}(\xi)|d\xi &= \int_{\mathbb{R}} |\xi|^p(1+|\xi|)|\hat{f}(\xi)|\frac{1}{1+|\xi|}d\xi \\
&\leq \sqrt{\int_{\mathbb{R}} |\xi|^{2p}(1+|\xi|)^2|\hat{f}(\xi)|^2 d\xi}\sqrt{\int_{\mathbb{R}} \frac{d\xi}{(1+|\xi|)^2}} \\
&\leq 2\sqrt{\int_{\mathbb{R}} |\xi|^{2p}(1+\xi^2)|\hat{f}(\xi)|^2 d\xi} \\
&= 2\sqrt{\int_{\mathbb{R}} |f^{(p+1)}(x)|^2 + |f^{(p)}(x)|^2 dx}.
\end{aligned}$$

To sum up, if f is a smooth function with support in $[-M-1, M+1]$, then

$$\mathbb{E}[f(X) - f(Y)] \leq \frac{C'\epsilon\sqrt{2(M+1)}}{\pi}\sqrt{||f^{(p+1)}||_\infty^2 + ||f^{(p)}||_\infty^2}. \tag{2.3}$$

Consider now the function $g(x) = a\int_0^x \exp\left(-\frac{1}{t(1-t)}\right) dt$ on $[0,1]$, $g(x) = 0$ for all $x \leq 0$ and $g(x) = 1$ for $x \geq 1$, where $a > 0$ is chosen so that $g \in \mathcal{C}^\infty(\mathbb{R})$. Let $K_n = \sup_{0\leq k\leq n} ||g^{(k)}||_\infty$; one can prove that $K_1 \leq 3$, $K_2 \leq 12$, $K_3 \leq 100$.

Define the even function $\chi_M \in \mathcal{C}_c^\infty(\mathbb{R})$ by $\chi_M(x) = 1$ on $[0, M]$ and $\chi_M(x) = g(M+1-x)$ if $x \geq M$. Let $F \in \mathcal{C}_c^\infty(\mathbb{R}) \cap B_{0,p+1}$. Then

$$\begin{aligned}\mathbb{E}[F(X) - F(Y)] &\leq \mathbb{E}[F\chi_M(X) - F\chi_M(Y)] + \mathbb{E}[F(1-\chi_M)(X) + F(1-\chi_M)(Y)] \\ &\leq \frac{C'\epsilon\sqrt{2(M+1)}}{\pi}\sqrt{||(F\chi_M)^{(p+1)}||_\infty^2 + ||(F\chi_M)^{(p)}||_\infty^2} + 2\mathbb{P}[X \geq M]\end{aligned}$$

However,

$$\begin{aligned}||(F\chi_M)^{(p)}||_\infty &\leq \sum_{k=0}^{p} \binom{p}{k} ||F^{(p-k)}||_\infty ||\chi_M^{(k)}||_\infty \\ &\leq 2^p K_p.\end{aligned}$$

We obtain

$$\begin{aligned}\mathbb{E}[F(X) - F(Y)] &\leq \frac{C'\epsilon\sqrt{2(M+1)}}{\pi}\sqrt{2^{2p+2}K_{p+1}^2 + 2^{2p}K_p^2} + 2\mathbb{P}[X \geq M] \\ &\leq C'\epsilon\frac{2^p\sqrt{10(M+1)}K_{p+1}}{\pi} + 2Ce^{-\lambda M^\alpha}.\end{aligned}$$

Take $M = \left(\lambda^{-1}\ln(1/\epsilon)\right)^{\frac{1}{\alpha}}$ to get

$$\mathbb{E}[F(X) - F(Y)] \leq 2\epsilon\left[C + \frac{2^{p-1}C'\sqrt{10}K_{p+1}}{\pi}\sqrt{1 + (\lambda^{-1}\ln(1/\epsilon))^{\frac{1}{\alpha}}}\right].$$

■

As already mentioned in the Introduction, assumption (H2) is sometimes better replaced by the following.

Theorem 2. *Under the assumptions of Theorem 1 with (H2) replaced by*

(H2') $\exists\gamma > 0, \exists C > 0, \forall A > 0$, $\mathbb{P}[|X| > A] \leq CA^{-\gamma}$ and $\mathbb{P}[|Y| > A] \leq CA^{-\gamma}$,

and if $\epsilon < \frac{\pi\sqrt{5}\gamma C}{5\cdot 2^p C_p}$, then

$$\mathcal{W}_{p+1}(X,Y) \leq \epsilon^{\frac{2\gamma}{2\gamma+1}}\left(\frac{2^{p+1}\sqrt{5M}C_p}{\pi}\right)^{\frac{2\gamma}{2\gamma+1}}(2\gamma C)^{\frac{1}{2\gamma+1}}\left(1 + \frac{1}{2\gamma}\right). \qquad (2.4)$$

Proof. The proof is similar to the one of Proposition 1; only the final optimization step (in M) changes. More precisely, for every test function $F \in \mathcal{C}_c^\infty(\mathbb{R}) \cap B_{0,p+1}$, and if $M > 1$, we have

$$\mathbb{E}[F(X) - F(Y)] \leq \epsilon\frac{2^{p+1}\sqrt{5M}K_{p+1}}{\pi} + 2CM^{-\gamma}.$$

Taking $M = \left(\frac{\epsilon 2^p\sqrt{5}K_{p+1}}{\gamma C\pi}\right)^{-\frac{2}{2\gamma+1}}$ yields the result. ■

3 Applications

3.1 Optimal bounds for Dickman approximation

The so-called Dickman's function $\rho(u)$ is solution to the backward delay differential equation

$$u\rho'(u) + \rho(u-1) = 0 \text{ for } u > 0, \tag{3.1}$$
$$\rho(u) = 0 \text{ for } u < 0 \text{ and } \rho(u) = 1 \text{ for } 0 \leq u \leq 1. \tag{3.2}$$

The Dickman function ρ is positive with finite integral $\int_0^\infty \rho(u)du = e^\gamma$ where γ is the Euler-Mascheroni constant; hence $x \mapsto e^{-\gamma}\rho(x), x \geq 0$ is a probability density (the so-called Dickman density).

Starting from (3.1) it is easy to show that a random variable Z is Dickman distributed if and only if

$$\mathbb{E}\left[Zf(Z)\right] = \mathbb{E}\left[f(Z+U)\right] \tag{3.3}$$

for all f bounded, with U uniform on $[0,1]$ independent of Z (see e.g. [2, 3]). In particular taking $f(x) = e^{i\xi x}$ in (3.3) we deduce that ϕ_Z, the characteristic function of the Dickman distribution, satisfies the simple ordinary differential equation

$$\forall \xi \in \mathbb{R},\ \phi'_Z(\xi) = \frac{e^{i\xi} - 1}{\xi}\phi_Z(\xi), \tag{3.4}$$

Next, we consider a permutation σ chosen uniformly at random in $\mathfrak{S}_n$. A record of σ is a left-to-right maxima, that is, an element $y_k = \sigma(k)$ in $y_1 \dots y_n$ is a record if $y_k \geq y_1, \dots, y_{k-1}$. The position of this record is k. The src statistic is $src = \sum_{k=1}^n k\, Y_k$ with Y_k the random variable which equals 1 if k is a position of a record and 0 otherwise; this statistic and its asymptotic behavior is important e.g. for analysis of algorithms [21, 18]. It is known (see [26]) that the random variables $Y_1, Y_2, \dots, Y_n$ are independent with $P(Y_k = 1) = \frac{1}{k}$. We set

$$W = \sum_{k=1}^n kY_k/n =: \sum_{k=1}^n X_k/n, \tag{3.5}$$

with $X_k = kY_k$. We set $W_k = W - X_k/n$, $1 \leq k \leq n$. Then, for all f bounded:

$$E\left[Wf(W)\right] = \frac{1}{n}\sum_{k=1}^n E\left[X_k f(W)\right] = \frac{1}{n}\sum_{k=1}^n E\left[f\left(W_k + \frac{k}{n}\right)\right] = E\left[f\left(W_I + \frac{I}{n}\right)\right], \tag{3.6}$$

where $I \sim U\{1, 2, \dots, n\}$ is independent of all else. From (3.6) we see that W satisfies

$$\mathbb{E}\left[Wf(W)\right] = \mathbb{E}\left[f(W+T)\right],$$
$$T = \frac{I}{n} - X_I, \quad I \sim U\{1, 2, \dots, n\}.$$

It is known (see e.g. [21, 6]) that $W \overset{(\mathcal{L})}{\Rightarrow} Z$ (convergence in law). We have the following result:

Proposition 1. *There exists a universal constant $c > 0$ such that:*

$$\forall n > 0,\ \forall \xi \in \mathbb{R},\ \mid \phi_W(\xi) - \phi_Z(\xi) \mid \leq \frac{c \mid \xi \mid^2}{n}.$$

Proof. Let $\xi > 0$. By a similar reasoning as for (1.13) we obtain

$$\phi_W(\xi) - \phi_Z(\xi) = i\phi_Z(\xi) \int_0^\xi \frac{1}{\phi_Z(\nu)} \mathbb{E}\left[e^{iW\nu} \left(e^{iT\nu} - e^{iU\nu} \right) \right] d\nu.$$

In order to conclude, we need to bound efficiently the previous quantity. First note, that:

$$\phi_Z(\xi) = \exp\left(\int_0^\xi \frac{e^{i\omega} - 1}{\omega} d\omega \right)$$

Thus,

$$\frac{1}{|\phi_Z(\xi)|} = \exp\left(\int_0^\xi \frac{1 - \cos(\omega)}{\omega} d\omega \right).$$

But the function $\omega \to (1 - \cos(\omega))/\omega$ is positive on $\mathbb{R}_+$. This implies:

$$\forall \nu \in [0, \xi], \ \frac{1}{|\phi_Z(\nu)|} \leq \frac{1}{|\phi_Z(\xi)|}.$$

Then, we have:

$$| \phi_W(\xi) - \phi_Z(\xi) | \leq c \mid \xi \mid^2 \mathbb{E}\left[|T - U| \right].$$

Clearly one can construct a coupling (U, I) in such a way that U is independent of W and

$$\mathbb{E}\left| U - \frac{I}{n} \right| \sim \frac{1}{n}. \tag{3.7}$$

(simply take $U \sim U[0, 1]$ and $I = j$ if $(j-1)/n \leq U \leq j/n$) so that

$$\begin{aligned} |\varphi_W(\xi) - \varphi_Z(\xi)| &\leq c|\xi|^2 \mathbb{E}\left[|T - U| \right], \\ &\leq c|\xi|^2 \left(\mathbb{E}\left| X_I \right| + \mathbb{E}\left| U - \frac{I}{n} \right| \right) \end{aligned}$$

and we conclude

$$|\varphi_W(\xi) - \varphi_Z(\xi)| \leq \frac{c}{n} |\xi|^2 \tag{3.8}$$

with c a universal constant. ■

We conclude this part by transforming (3.8) into a result regarding smooth Wasserstein distance.

Theorem 3. *There exists an universal constant $C > 0$ such that for all $n \geq 2$:*

$$\mathcal{W}_3(W, Z) \leq C \frac{\sqrt{\log n}}{n} \tag{3.9}$$

Proof. In order to apply Theorem 1, we only need to check that W admits an exponential moment (independent of n). We have

$$\mathbb{E}\left[e^{\gamma W} \right] = \prod_{k=1}^{n} \left(1 + \frac{e^{\lambda k/n} - 1}{k} \right),$$

so that $\log \mathbb{E}\left[e^{\gamma W} \right] \leq \frac{1}{n} \sum_{k=1}^{n} \frac{n}{k} \left(e^{\gamma \frac{k}{n} - 1} \right)$, which is uniformly bounded. Markov's inequality thus implies that $(H2)$ of Theorem 1 holds, whereas $(H1)$ has been proven in (3.8). The result follows. ■

Remark 2. *Our bound (3.9) is to be compared with [6, Theorem 3.3] where they apply Stein's method for the compound Poisson to obtain a similar result in Wasserstein distance (with a rate of* $1/n^\gamma$ *for some* $\gamma > 0$*). There is certainly room for improvement of the kind of distance in Theorem 3. In particular Wasserstein and Kolmogorov are perhaps attainable by a more careful analysis.*

3.2 Infinitely divisible distributions and size-biasing

In this section, we revisit the arguments from Section 1.1 in more detail and consider random variables X verifying a Stein identity of the type

$$\mathbb{E}\left[Xf(X)\right] = \lambda \mathbb{E}\left[f(X+U_X)\right], \tag{3.10}$$

for all bounded continuous f and some independent random variable U_X. If we assume moreover that X can be size-biased (that is, $X \geq 0$ and $\mathbb{E}[X] = \lambda \in (0, +\infty)$) , then distributions of X satisfying such an identity are exactly infinitely divisible distributions that can be size-biased, see [4]. In particular, any non-negative integrable compound Poisson distribution

$$X = \sum_{i=1}^{N} Z_i, \tag{3.11}$$

where N is Poisson, Z_i i.i.d. independent of N and positive a.s., satisfies (3.10) with $U_X = Z_1^*$, the size-biased transform of Z_1.

Let Y be another random variable of the previous type, that is, satisfying

$$\mathbb{E}\left[Yf(Y)\right] = \lambda \mathbb{E}\left[f(Y+U_Y)\right],$$

for all bounded continous f and a random variable U_Y independent of Y. We consider the problem of estimating the closeness between the distributions of X and Y in terms of that of U_X and U_Y. For a random variable Z, let ψ_Z denotes its characteristic function.

We assume that:

(i) ψ_X never vanishes and that there exists $p \in \mathbb{N}$ and $C > 0$ (depending only on X) such that for all $\xi \in \mathbb{R}$,

$$|\psi_X(\xi)| \int_0^{|\xi|} \frac{1}{|\psi_X(t)|} t \, dt \leq C|\xi|^p. \tag{3.12}$$

(ii) there exists $\gamma > 0$ such that $\mathbb{E}[e^{\gamma X}]$ and $\mathbb{E}[e^{\gamma Y}]$ are finite.

Assumptions (i) and (ii) are verified, for instance, for any compound Poisson distribution X as defined in (3.11) with $Z_1 \geq 0$ and $\mathbb{E}[e^{\gamma Z_1}] < \infty$. Indeed, in this case, we know from [30], Proposition 2.7 (p. 97) and Theorem 3.2 (p.100), that $|\psi_X| \geq \mathbb{P}[X = 0] > 0$. Thus (3.12) holds with $p = 2$. Assumption (ii) is easily verified by a conditioning argument. Another distribution satisfying assumptions (i) and (ii) is the Dickman distribution, as seen in the previous section.

Applying the Stein identities to $e^{i\xi\cdot}$, we readily obtain (recall identity (1.13))

$$\psi_Y(\xi) - \psi_X(\xi) = i\lambda\psi_X(\xi) \int_0^{\xi} \frac{\psi_Y(t)}{\psi_X(t)} (\psi_{U_Y}(t) - \psi_{U_X}(t)) \, dt.$$

Assume U_X and U_Y live on the same probability space. From (3.12), we get

$$|\psi_Y(\xi) - \psi_X(\xi)| \le C\lambda\mathbb{E}|U_X - U_Y|\,|\xi|^p, \tag{3.13}$$

where C depends only on X. Finally, assumption (ii), Theorem 1 and the fact that we can choose arbitrarily the coupling between U_X and U_Y leads to

$$\mathcal{W}_{p+1}(X,Y) \le C|\log(W(U_X,U_Y))|W(U_X,U_Y),$$

where C does not depend on Y and $W(U_X,U_Y)$ is the (classical) 1-Wasserstein distance between the distributions of U_X and U_Y.

3.3 The zero-bias transformation and Gaussian approximation

Let W have mean 0 and variance σ^2. Following [15] we say that $W^\star$ has the zero-bias distribution for W if

$$\mathbb{E}\left[W\phi(W)\right] = \mathbb{E}\left[\sigma^2\phi'(W^\star)\right] \tag{3.14}$$

for all ϕ bounded and differentiable. Let $T = W^\star - W$. Then taking $\phi(x) = e^{i\xi x}$ in (3.14) and writing φ_W for the characteristic function of W we get (recall identity (1.12))

$$\Delta(t) = e^{-\xi^2\sigma^2/2}\int_0^\xi e^{s^2\sigma^2/2}s\sigma^2\mathbb{E}\left[e^{isW}\left(e^{isT}-1\right)\right]ds. \tag{3.15}$$

with $\Delta(\xi) = \varphi_W(\xi) - \varphi_Z(\xi)$. By integration by parts

$$\begin{aligned}\Delta(t) &= \sigma^2\mathbb{E}\left[e^{i\xi W}\left(e^{i\xi T}-1\right)\right]\\ &\quad + e^{-\xi^2\sigma^2/2}\int_0^\xi e^{s^2\sigma^2/2}\mathbb{E}\left[i(W+T)e^{is(W+T)} - iWe^{isW}\right]ds. \end{aligned} \tag{3.16}$$

Next we cut the second integrand

$$\begin{aligned}&\int_0^\xi e^{s^2\sigma^2/2}\mathbb{E}\left[i(W+T)e^{is(W+T)} - iWe^{isW}\right]ds\\ &= \int_0^1 e^{s^2\sigma^2/2}\mathbb{E}\left[i(W+T)e^{is(W+T)} - iWe^{isW}\right]ds\\ &\quad + \int_1^\xi e^{s^2\sigma^2/2}\mathbb{E}\left[i(W+T)e^{is(W+T)} - iWe^{isW}\right]ds\end{aligned}$$

and treat each term separately : the first one yields

$$\begin{aligned}&\int_0^1 e^{s^2\sigma^2/2}\mathbb{E}\left[(W+T)e^{is(W+T)} - We^{isW}\right]ds\\ &= \int_0^1 e^{s^2\sigma^2/2}\mathbb{E}\left[(W+T)e^{is(W+T)} - We^{is(W+T)} + W\left(e^{is(W+T)} - e^{isW}\right)\right]ds\end{aligned}$$

which shall be good and the second one is after another integration by parts

$$\begin{aligned}&\int_1^{\xi} e^{s^2\sigma^2/2}\mathbb{E}\left[i(W+T)e^{is(W+T)} - iWe^{isW}\right]ds \\ &= \left(e^{s^2\sigma^2/2}E\left[\frac{(W+T)}{s}e^{is(W+T)} - \frac{W}{s}e^{isW}\right]\right) \\ &\quad - \int_1^{\xi} e^{s^2\sigma^2/2}\frac{1}{s^2}\mathbb{E}\left[(W+T)e^{is(W+T)} - We^{isW}\right]ds \\ &\quad + \int_1^{\xi} e^{s^2\sigma^2/2}\frac{1}{s}\mathbb{E}\left[(W+T)^2e^{is(W+T)} - W^2e^{isW}\right]ds\end{aligned}$$

which shall also be good. The first summand yields

$$\sigma^2\mathbb{E}\left[e^{i\xi W}\left(e^{i\xi T}-1\right)\right] = \sigma^2\mathbb{E}\left[e^{i\xi W^\star} - e^{i\xi W}\right] = \sigma^2\left(\varphi_{W^\star}(\xi) - \varphi_W(\xi)\right).$$

The second term

$$\left|e^{-\xi^2\sigma^2/2}\int_0^{\xi} e^{s^2\sigma^2/2}\mathbb{E}\left[i(W+T)e^{is(W+T)} - iWe^{isW}\right]ds\right| \le e^{-\xi^2\sigma^2/2}\int_0^{\xi} se^{s^2\sigma^2/2}ds\mathbb{E}\left[|T|\right]$$

Lemma 1. *Under the above notation*

$$|\Delta(\xi)| \le c\left(|\varphi_{W^\star}(\xi) - \varphi_W(\xi)| + \mathbb{E}\left[|T|\right]\right) \tag{3.17}$$

with c *a universal constant depending only on* σ^2.

By Theorem 1 we get the next result.

Theorem 4. *Let all notations be as above then*

$$d_{\mathrm{FM}}(W,Z) \le c(d_{\mathrm{FM}}(W,W^\star) + \mathbb{E}\left[|T|\right])\sqrt{|\log\mathbb{E}[|T|]|} \tag{3.18}$$

and in particular

$$d_{\mathrm{FM}}(W,Z) \le c\mathbb{E}\left[|T|\right]\sqrt{|\log\mathbb{E}[|T|]|}. \tag{3.19}$$

3.4 Stable distributions

Let $\alpha \in (1,2)$. Let $c = (1-\alpha)/(2\Gamma(2-\alpha)\cos(\pi\alpha/2))$. We denote by Z^α a symmetric α-stable random variable whose characteristic function is given by:

$$\forall\xi\in\mathbb{R},\ \phi_{Z^\alpha}(\xi) = \exp\bigg(-\mid\xi\mid^\alpha\bigg).$$

We define the following differential operator on smooth enough function:

$$\mathcal{D}^{\alpha-1}(\psi)(x) = \frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}(\psi)(\xi)\left(\frac{i\alpha\mid\xi\mid^\alpha}{\xi}\right)e^{i\xi x}d\xi$$

We have the following straightforward Stein-type characterization identity:

$$\mathbb{E}\big[Z^\alpha\psi(Z^\alpha)\big] = \mathbb{E}[\mathcal{D}^{\alpha-1}(\psi)(Z^\alpha)].$$

We denote by $\mathrm{Dom}(Z^\alpha)$ the normal domain of attraction of Z^α. We recall the following result (see Theorem 5 page 81 of [14] and Theorem 2.6.7 page 92 of [19]).

Theorem 5. *A distribution function, F, is in the normal domain of attraction of Z^α if and only if:*

$$\forall x > 0,\ F(x) = 1 - \frac{(c + a_1(x))}{x^\alpha},$$
$$\forall x < 0,\ F(x) = \frac{(c + a_2(x))}{(-x)^\alpha},$$

with $\lim_{x \to +\infty} a_1(x) = \lim_{x \to -\infty} a_2(x) = 0$.

Remark 3. *Let $\lambda = (2c)^{1/\alpha}$ and consider the following Pareto type distribution function:*

$$\forall x > 0,\ F_\lambda(x) = 1 - \frac{1}{2(1 + \frac{x}{\lambda})^\alpha},$$
$$\forall x < 0,\ F_\lambda(x) = \frac{1}{2(1 - \frac{x}{\lambda})^\alpha}.$$

It is easy to check that F_λ is in $\mathrm{Dom}(Z^\alpha)$. In particular, we have:

$$\forall x > 0,\ a_1(x) = a_2(-x) = \frac{x^\alpha}{2\big(1 + \frac{x}{\lambda}\big)^\alpha} - c.$$

Let X be a random variable in $\mathrm{Dom}(Z^\alpha)$ such that $\mathbb{E}[X] = 0$. We make the following assumptions:

- The functions $a_1(.)$ and $a_2(.)$ are continuous and bounded on $\mathbb{R}_+^*$ and on $\mathbb{R}_-^*$ respectively. Moreover, they satisfy:

$$\lim_{x \to +\infty} x a_1(x) < +\infty,\ \lim_{x \to -\infty} x a_2(x) < +\infty.$$

Let (X_i) be independent copies of X. We define the random variable W by:

$$W = \frac{1}{n^{\frac{1}{\alpha}}} \sum_{i=1}^n X_i.$$

We consider the following function on $\mathbb{R}^*$:

$$\phi_X^*(\xi) = \frac{-\xi}{\alpha \mid \xi \mid^\alpha} \mathbb{E}\big[iX \exp\big(i\xi X\big)\big].$$

Note in particular that this function is well-defined, continuous on $\mathbb{R}^*$ and $\lim_{|\xi| \to +\infty} \phi_X^*(\xi) = 0$ since $X \in \mathrm{Dom}(Z^\alpha)$ and $\alpha \in (1, 2)$. We first state the following technical result, whose proof is relegated to the Appendix.

Proposition 2. *The function ϕ_X^* satisfies:*

- $\phi_X^*(0^+) = \phi_X^*(0^-) = 1$.
- *If the law of X is symmetric, then ϕ_X^* is real-valued and even.*
- *There exists a tempered distribution T_X^* such that $\overline{\phi_X^*}$ is the Fourier transform of T_X^*.*

- *For all ψ smooth enough, we have:*

$$\mathbb{E}\big[X\psi(X)\big] = < T_X^*; \mathcal{D}^{\alpha-1}(\psi) >$$

- *We have the following formulae:*

$$\forall \xi > 0,\ \phi_X(\xi) - \phi_{Z^\alpha}(\xi) = \alpha e^{-\xi^\alpha} \int_0^\xi \nu^{\alpha-1} e^{\nu^\alpha} \big(\phi_X(\nu) - \phi_X^*(\nu)\big) d\nu,$$
$$\forall \xi < 0,\ \phi_X(\xi) - \phi_{Z^\alpha}(\xi) = \alpha e^{-(-\xi)^\alpha} \int_\xi^0 (-\nu)^{\alpha-1} e^{(-\nu)^\alpha} \big(\phi_X(\nu) - \phi_X^*(\nu)\big) d\nu.$$

Remark 4. *The functional equality appearing in the fourth bullet point of the previous proposition is very close to the definition of the zero-bias transformation distribution for centred random variable with finite variance. It would be interesting to know if T_X^* is actually a linear functional defined by a probability measure on $\mathbb{R}$.*

Since the X_i are independent and identically distributed, we have the following formula for the characteristic function of W:

$$\forall \xi \in \mathbb{R},\ \phi_W(\xi) = \phi_X^n\big(\frac{\xi}{n^{\frac{1}{\alpha}}}\big)$$

Consequently, this characteristic function is differentiable and we have:

$$\phi_W^*(\xi) = \phi_X^{n-1}\big(\frac{\xi}{n^{\frac{1}{\alpha}}}\big)\phi_X^*\big(\frac{\xi}{n^{\frac{1}{\alpha}}}\big).$$

Thus, we can apply the previous proposition to W and we get the following formula for $\xi > 0$:

$$\begin{aligned}
\phi_W(\xi) - \phi_{Z^\alpha}(\xi) &= \alpha e^{-\xi^\alpha} \int_0^\xi \nu^{\alpha-1} e^{\nu^\alpha} \big(\phi_W(\nu) - \phi_W^*(\nu)\big) d\nu, \\
&= \alpha e^{-\xi^\alpha} \int_0^\xi \nu^{\alpha-1} e^{\nu^\alpha} \phi_X^{n-1}\big(\frac{\nu}{n^{\frac{1}{\alpha}}}\big)\big(\phi_X\big(\frac{\nu}{n^{\frac{1}{\alpha}}}\big) - \phi_X^*\big(\frac{\nu}{n^{\frac{1}{\alpha}}}\big)\big) d\nu.
\end{aligned} \tag{3.20}$$

Now, the objective is to find an efficient way to bound pointwisely the difference between $\phi_X\big(\xi\big)$ and $\phi_X^*\big(\xi\big)$. For this purpose, we use the following formula first obtained in the proof of Proposition 2:

$$\begin{aligned}
\phi_X^*(\xi) = 1 &+ \frac{1}{\alpha}\int_0^{+\infty} e^{ix} \frac{a_1\big(\frac{x}{\xi}\big)}{x^{\alpha-1}} dx \\
&- \frac{i}{\alpha}\int_0^{+\infty} (e^{ix} - 1)\frac{a_1\big(\frac{x}{\xi}\big)}{x^\alpha} dx + \frac{1}{\alpha}\int_{-\infty}^0 e^{ix} \frac{a_2\big(\frac{x}{\xi}\big)}{(-x)^{\alpha-1}} dx \\
&+ \frac{i}{\alpha}\int_{-\infty}^0 (e^{ix} - 1)\frac{a_2\big(\frac{x}{\xi}\big)}{(-x)^\alpha} dx.
\end{aligned} \tag{3.21}$$

Thus, the difference between $\phi_X\big(\xi\big)$ and $\phi_X^*\big(\xi\big)$ brings into play the difference between $\phi_X\big(\xi\big)$ and 1. Since $X \in \text{Dom}(Z^\alpha)$, we expect this difference to be of order $\mid \xi \mid^\alpha$. This is the aim of the next lemma.

Lemma 2. *There exists some strictly positive constant, C_α, depending on α only, such that, for any ξ in $\mathbb{R}$, we have:*

$$| \phi_X(\xi) - 1 | \leq C_\alpha \mid \xi \mid^\alpha \big(1+ \| a_1 \|_\infty + \| a_2 \|_\infty \big)$$

Remark 5. *We note in particular that the previous lemma implies:*

$$\left|\phi_X\big(\frac{\xi}{n^{\frac{1}{\alpha}}}\big) - 1\right| \leq C_\alpha \frac{\mid \xi \mid^\alpha}{n} \big(1+ \| a_1 \|_\infty + \| a_2 \|_\infty \big).$$

Now, we want to bound pointwisely the residual term, namely:

$$R^\alpha(\xi) = \frac{1}{\alpha}\int_0^{+\infty} e^{ix}\frac{a_1\big(\frac{x}{\xi}\big)}{x^{\alpha-1}}dx - \frac{i}{\alpha}\int_0^{+\infty}(e^{ix}-1)\frac{a_1\big(\frac{x}{\xi}\big)}{x^\alpha}dx \\ + \frac{1}{\alpha}\int_{-\infty}^0 e^{ix}\frac{a_2\big(\frac{x}{\xi}\big)}{(-x)^{\alpha-1}}dx + \frac{i}{\alpha}\int_{-\infty}^0(e^{ix}-1)\frac{a_2\big(\frac{x}{\xi}\big)}{(-x)^\alpha}dx.$$

For this purpose, we have the following lemma.

Lemma 3. *For any $i \in \{1, 2\}$, there exist strictly positive constants $C^i_{\alpha,1}$ and $C^i_{\alpha,2}$ depending on the function a_i and α only, such that:*

$$\forall \xi \in \mathbb{R},\ \left|\int_0^{+\infty}(e^{ix}-1)\frac{a_1\big(\frac{x}{\xi}\big)}{x^\alpha}dx\right| \leq C^1_{\alpha,1} \mid \xi \mid^{\frac{2-\alpha}{2}},$$

$$\forall \xi \in \mathbb{R},\ \left|\int_0^{+\infty} e^{ix}\frac{a_1\big(\frac{x}{\xi}\big)}{x^{\alpha-1}}dx\right| \leq C^1_{\alpha,2} \mid \xi \mid^{2-\alpha},$$

and similarly for the integrals with the function $a_2(.)$.

We are now in position to provide a bound for the difference between the characteristic function of W and the one of Z^α.

Theorem 6. *There exists a strictly positive constant C depending on α, a_1 and a_2 only such that:*

$$\forall \xi \in \mathbb{R},\ | \phi_W(\xi) - \phi_{Z^\alpha}(\xi) | \leq C\bigg(\frac{\mid \xi \mid^{2\alpha}}{n} + \frac{\mid \xi \mid^{\frac{\alpha}{2}+1}}{n^{\frac{1}{\alpha}-\frac{1}{2}}} + \frac{\xi^2}{n^{\frac{2}{\alpha}-1}}\bigg).$$

Proof. This a direct application of equations (3.20) and (3.21) together with the bounds obtained in Lemmas 2 and 3. ∎

We conclude the discussion on stable approximation by applying a variant of Theorem 2 to obtain correct-order rates of convergence towards stable distributions.

Theorem 7. *There exists a constant $C > 0$ depending on α, a_1 and a_2 such that*

$$\mathcal{W}_5(W, Z^\alpha) \leq C\, n^{\frac{2\alpha}{2\alpha+1}\left(\frac{1}{2}-\frac{1}{\alpha}\right)}. \tag{3.22}$$

Remark 6. *This result should be compared with the one of [20] (see also [25]). In particular, they obtain bounds of the order $n^{1/2-1/\alpha}$ for the classical 2-Wasserstein distance.*

3.5 Bounds for targets in second Wiener chaos

Recall the following Corollary from [1]:

Corollary 1. *Let $d \geq 1$, $q \geq 1$ and $(m_1, ..., m_d) \in \mathbb{N}^d$ such that $m_1 + ... + m_d = q$. Let $(\lambda_1, ..., \lambda_d) \in \mathbb{R}^*$ pairwise distinct and $(N_i)_{i\geq 1}$ an i.i.d. sequence of standard normal random variables. Consider:*

$$F = \sum_{i=1}^{m_1} \lambda_1(N_i^2 - 1) + \sum_{i=m_1+1}^{m_1+m_2} \lambda_2(N_i^2 - 1) + ... + \sum_{i=m_1+...+m_{d-1}+1}^{q} \lambda_d(N_i^2 - 1),$$

Let Y be a real valued random variable such that $\mathbb{E}[|Y|] < +\infty$. Then $Y \overset{law}{=} F$ if and only if

$$\mathbb{E}\Bigg[\big(Y + \sum_{i=1}^{d} \lambda_i m_i\big)(-1)^d 2^d \Big(\prod_{j=1}^{d} \lambda_j\Big)\phi^{(d)}(Y) + \sum_{l=1}^{d-1} 2^l(-1)^l \Bigg(Y e_l(\lambda_1, ..., \lambda_d)$$
$$+ \sum_{k=1}^{d} \lambda_k m_k \left(e_l(\lambda_1, ..., \lambda_d) - e_l((\underline{\lambda}_k))\right)\Bigg)\phi^{(l)}(Y) + Y\phi(Y)\Bigg] = 0, \tag{3.23}$$

for all $\phi \in S(\mathbb{R})$.

Let us denote by $\mathcal{A}_\infty$ the differential operator appearing in the left-hand side of (3.23). By the proof of Theorem 2.1 of [1], we can rewrite it in the following form:

$$\forall \phi \in S\big(\mathbb{R}\big),\ \mathcal{A}_\infty(\phi) = (x+ < m, \lambda >)\mathcal{A}_{d,\lambda}(\phi)(x) - \mathcal{B}_{d,m,\lambda}(\phi)(x),$$

with,

$$\mathcal{A}_{d,\lambda}(\phi)(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathcal{F}(\phi)(\xi)\Big(\prod_{k=1}^{d}(\frac{1}{2\lambda_k} - i\xi)\Big)\exp(ix\xi)d\xi,$$

$$\mathcal{B}_{d,m,\lambda}(\phi)(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathcal{F}(\phi)(\xi)\Big(\sum_{k=1}^{d}\frac{m_k}{2}\prod_{l=1,l\neq k}^{d}(\frac{1}{2\lambda_l} - i\xi)\Big)\exp(ix\xi)d\xi,$$

$$\mathcal{F}(\phi)(\xi) = \int_{\mathbb{R}} \phi(x)\exp(-ix\xi)dx.$$

Let F_n be a sequence of random variables converging in law towards F as in Corollary 1. In the next proposition, we find a formula linking the difference between the characteristic functions of F_n and of F with the quantity $\mathbb{E}\big[\mathcal{A}_\infty(\exp(is.))(F_n)\big]$ for $s \in \mathbb{R}$.

Proposition 3. *For any ξ in $\mathbb{R}$, we have:*

$$\phi_{F_n}(\xi) - \phi_F(\xi) = i\phi_F(\xi)\int_0^{\xi}\frac{1}{\sigma_{\mathcal{A}}(s)\phi_F(s)}\mathbb{E}\big[\mathcal{A}_\infty(\exp(is.))(F_n)\big]ds. \tag{3.24}$$

Proof. Let ϕ be a smooth enough function which will be specified later on. We are interested in the quantity $\mathbb{E}\big[\mathcal{A}_\infty(\phi)(F_n)\big]$:

$$\begin{aligned}\mathbb{E}\big[\mathcal{A}_\infty(\phi)(F_n)\big] =& \mathbb{E}\big[F_n\mathcal{A}_{d,\lambda}(\phi)(F_n)\big] + < m, \lambda > \mathbb{E}\big[\mathcal{A}_{d,\lambda}(\phi)(F_n)\big] - \mathbb{E}\big[\mathcal{B}_{d,m,\lambda}(\phi)(F_n)\big] \\ =& \frac{1}{2\pi}\Big[< \mathcal{F}\big(x\mathcal{A}_{d,\lambda}(\phi)\big); \phi_{F_n} > + < m, \lambda >< \mathcal{F}\big(\mathcal{A}_{d,\lambda}(\phi)\big); \phi_{F_n} > \\ & - < \mathcal{F}\big(\mathcal{B}_{d,m,\lambda}(\phi)\big); \phi_{F_n} > \Big],\end{aligned}$$

Note that the previous brackets can be understood as duality brackets as soon as F_n as enough moments and $\mathcal{F}\big(\mathcal{A}_{d,\lambda}(\phi)\big)$ and $\mathcal{F}\big(\mathcal{B}_{d,m,\lambda}(\phi)\big)$ are distributions with compact support with orders less or equal to the number of moments of F_n. We denote by $\sigma_{\mathcal{A}}$ and $\sigma_{\mathcal{B}}$ the symbols associated with $\mathcal{A}_{d,\lambda}$ and $\mathcal{B}_{d,m,\lambda}$ respectively. These symbols are infinitely differentiable and non-zero everywhere on $\mathbb{R}$. We have:

$$\begin{aligned}\mathbb{E}\big[\mathcal{A}_\infty(\phi)(F_n)\big] =&\frac{1}{2\pi}\bigg[-i<\mathcal{F}(\phi);\sigma_{\mathcal{A}}(.)(\phi_{F_n})'>+<m,\lambda><\mathcal{F}(\phi);\sigma_{\mathcal{A}}(.)\phi_{F_n}>\\&-<\mathcal{F}(\phi);\sigma_{\mathcal{B}}(.)\phi_{F_n}>\bigg].\end{aligned}$$

Remind that by the proof of Theorem 2.1, we have:

$$\prod_{k=1}^{d}(\frac{1}{2\lambda_k}-i\xi)\frac{d}{d\xi}\bigg(\phi_F(\xi)\bigg)=\bigg[-i<m,\lambda>\prod_{k=1}^{d}(\frac{1}{2\lambda_k}-i\xi)+i\sum_{k=1}^{d}\frac{m_k}{2}\prod_{l=1,l\neq k}^{d}(\frac{1}{2\lambda_l}-i\xi)\bigg]\phi_F(\xi).$$

Namely,

$$\sigma_{\mathcal{A}}(\xi)\frac{d}{d\xi}\bigg(\phi_F(\xi)\bigg)=\bigg[-i<m,\lambda>\sigma_{\mathcal{A}}(\xi)+i\sigma_{\mathcal{B}}(\xi)\bigg]\phi_F(\xi).$$

Thus, we have:

$$\begin{aligned}\mathbb{E}\big[\mathcal{A}_\infty(\phi)(F_n)\big]&=\frac{-i}{2\pi}\bigg[<\mathcal{F}(\phi);\sigma_{\mathcal{A}}(.)(\phi_{F_n})'>-<\mathcal{F}(\phi);\sigma_{\mathcal{A}}(.)\frac{(\phi_F)'}{\phi_F}\phi_{F_N}>\bigg],\\&=\frac{-i}{2\pi}<\mathcal{F}(\phi);\sigma_{\mathcal{A}}(.)\phi(\frac{\phi_{F_n}}{\phi_F})'>.\end{aligned}$$

Now choose ϕ equal to $\exp(i\xi.)$. Thus, its Fourier transform is exactly 2π times the Dirac distribution at ξ. We obtain:

$$\begin{aligned}\mathbb{E}\big[\mathcal{A}_\infty(\exp(i\xi.))(F_n)\big]&=(-i)\sigma_{\mathcal{A}}(\xi)\phi_F(\xi)(\frac{\phi_{F_n}(\xi)}{\phi_F(\xi)})',\\&=(-i)\sigma_{\mathcal{A}}(\xi)\phi_F(\xi)(\frac{\phi_{F_n}(\xi)-\phi_F(\xi)}{\phi_F(\xi)})'.\end{aligned}$$

This ends the proof of the proposition. ∎

Contrarily to the paper [1], we do not restrict ourself to the situation where the sequence (F_n) is inside the second Wiener chaos but instead lies in a finite sum of Wiener chaoses. It is well known ([24]) that such random variables are in $\mathbb{D}^\infty$. We emphasize that the existence of a finite integer p such that the sequence lies in the sum of the p first Wiener chaoses is for simplicity only. It is enough for our purpose and it prevents us from imposing conditions for carrying out interchange of derivative, integration and integration-by-parts. Relying on the papers [1] (page 12 equation 2.9) and [5] (Proposition 3.2 page 13), $\mathbb{E}\big[\mathcal{A}_\infty(\exp(i\xi.))(F_n)\big]$ admits the following suitable representation in terms of iterated gamma operators firstly defined in [23]. Regarding notations and definitions, we refer to [1], (at the beginning of section 2.3 page 10).

$$
\begin{aligned}
\mathbb{E}\left[\mathcal{A}_\infty(\phi)(F_n)\right] &= -\mathbb{E}\Bigg[\phi^{(d)}(F_n) \times \Big(\sum_{r=1}^{d+1} a_r\left[\Gamma_{r-1}(F_n) - \mathbb{E}[\Gamma_{r-1}(F_n)]\right]\Big)\Bigg] \\
&\quad + \sum_{r=2}^{d+1} a_r \sum_{l=0}^{r-2} \Big\{\mathbb{E}[\phi^{(d-l)}(F_n)] \times \Big(\mathbb{E}\left[\Gamma_{r-l-1}(F)\right] - \mathbb{E}\left[\Gamma_{r-l-1}(F_n)\right]\Big)\Big\} \\
&= -\mathbb{E}\Bigg[\phi^{(d)}(F_n) \times \Big(\sum_{r=1}^{d+1} a_r\left[\Gamma_{r-1}(F_n) - \mathbb{E}[\Gamma_{r-1}(F_n)]\right]\Big)\Bigg] \\
&\quad + \sum_{r=2}^{d+1} a_r \sum_{l=0}^{r-2} \frac{\mathbb{E}[\phi^{(d-l)}(F_n)]}{(r-l-1)!} \times \Big(\kappa_{r-l}(F) - \kappa_{r-l}(F_n)\Big).
\end{aligned}
\tag{3.25}
$$

Combining this expression and the formula (3.24) for the difference of the characteristic functions of F_n and F, it holds that F_n converges in distribution towards F if the following conditions are satisfied:

$$
\begin{aligned}
&\forall r = 2, \dots, d+1,\ \kappa_r(F_n) \to \kappa_r(F) \\
&\sum_{r=1}^{d+1} a_r[\Gamma_{r-1}(F_n) - \mathbb{E}[\Gamma_{r-1}(F_n)] \to 0.
\end{aligned}
$$

Our goal is now to give a quantitative bound. Firstly note that:

$$
\begin{aligned}
|\mathbb{E}\left[\mathcal{A}_\infty(e^{i\xi\cdot})(F_n)\right]| \leq& |\xi|^d \mathbb{E}\Bigg[|\sum_{r=1}^{d+1} a_r\left[\Gamma_{r-1}(F_n) - \mathbb{E}[\Gamma_{r-1}(F_n)]\right]|\Bigg] \\
&+ \sum_{r=2}^{d+1} |a_r| \sum_{l=0}^{r-2} \frac{|\xi|^{d-l}}{(r-l-1)!} \times |\kappa_{r-l}(F) - \kappa_{r-l}(F_n)| \\
&\leq (1+|\xi|)^d \Delta_n,
\end{aligned}
$$

with:

$$
\Delta_n = \mathbb{E}\Bigg[|\sum_{r=1}^{d+1} a_r\left[\Gamma_{r-1}(F_n) - \mathbb{E}[\Gamma_{r-1}(F_n)]\right]|\Bigg] + \sum_{r=2}^{d+1} |a_r| \sum_{l=0}^{r-2} \frac{1}{(r-l-1)!} \times |\kappa_{r-l}(F) - \kappa_{r-l}(F_n)|.
$$

Thus, we obtain:

$$
|\phi_{F_N}(\xi) - \phi_F(\xi)| \leq \Delta_n |\phi_F(\xi)| \int_0^\xi \frac{(1+|s|)^d}{|\sigma_{\mathcal{A}}(s)\phi_F(s)|} ds, \tag{3.26}
$$

Before moving on, we need the following simple lemma.

Lemma 4. *Let F be as above and ϕ_F its characteristic function. We have:*

$$\forall \xi \in \mathbb{R},\ \phi_F(\xi) = \exp(-i\xi < m, \lambda >) \prod_{k=1}^{d} \frac{1}{(1-2i\xi\lambda_k)^{\frac{m_k}{2}}},$$
$$\forall \xi \in \mathbb{R},\ \frac{1}{(1+4\lambda_{max}^2\xi^2)^{\frac{q}{4}}} \leq |\phi_F(\xi)| \leq \frac{1}{(1+4\lambda_{min}^2\xi^2)^{\frac{q}{4}}},$$
$$\forall \xi \in \mathbb{R},\ \frac{1}{|\sigma_{\mathcal{A}}(s)|} \leq \frac{2^d \prod_{j=1}^d |\lambda_j|}{(1+4\lambda_{min}^2\xi^2)^{\frac{d}{2}}}.$$

Using these inequalities into (3.26), we obtain:

$$|\phi_{F_N}(\xi) - \phi_F(\xi)| \leq \Delta_n \frac{2^d \prod_{j=1}^d |\lambda_j|}{(1+4\lambda_{min}^2\xi^2)^{\frac{q}{4}}} \int_0^{\xi} \frac{(1+|s|)^d (1+4\lambda_{max}^2 s^2)^{\frac{q}{4}}}{(1+4\lambda_{min}^2 s^2)^{\frac{d}{2}}} ds,$$
$$|\phi_{F_N}(\xi) - \phi_F(\xi)| \leq C_{d,\lambda}\Delta_n \int_0^{\xi} \frac{(1+|s|)^d}{(1+s^2)^{\frac{d}{2}}} ds,$$
$$|\phi_{F_N}(\xi) - \phi_F(\xi)| \leq C_{d,\lambda}\Delta_n |\xi|. \tag{3.27}$$

Now, we perform a similar reasoning as in the previous section in order to use the bound (3.27) to obtain a quantitative bound on smooth Wasserstein metric between F_n and F. Finally, we get the following Theorem.

Theorem 8. *Let F_n be a sequence of random variables lying in the finite sum of the p first Wiener chaoses and let*

$$F = \sum_{i=1}^{m_1} \lambda_1 (N_i^2 - 1) + \sum_{i=m_1+1}^{m_1+m_2} \lambda_2 (N_i^2 - 1) + ... + \sum_{i=m_1+...+m_{d-1}+1}^{q} \lambda_d (N_i^2 - 1),$$

for parameters $(\lambda_i)_{1\leq i \leq d}$ and $m_1, \cdots, m_d \in \mathbb{N}^d$ as in Corollary 1. Then, there exist some constants $C > 0$ and $\theta > 0$ depending only on the target and p such that

$$\mathcal{W}_2(F_n, F) \leq C\Delta_n \log(\Delta_n)^{\theta},$$

with

$$\Delta_n = \mathbb{E}\left[|\sum_{r=1}^{d+1} a_r \left[\Gamma_{r-1}(F_n) - \mathbb{E}[\Gamma_{r-1}(F_n)]\right] | \right] + \sum_{r=2}^{d+1} |a_r| \sum_{l=0}^{r-2} \frac{1}{(r-l-1)!}$$
$$\times |\kappa_{r-l}(F) - \kappa_{r-l}(F_n)|.$$

4 Appendix

4.1 Remarks on smooth Wasserstein distance

Let $n \geq 1$, $0 \leq p \leq n$, and

$$B_{p,n} = \{\phi \in \mathcal{C}^n(\mathbb{R}) \mid ||\phi^{(k)}||_\infty \leq 1, \forall\, p \leq k \leq n\}.$$

For two $\mathbb{R}$-valued random variables X and Y, we define the following distances between their distributions :

$$\begin{aligned} d_W^{p,n}(X,Y) &= \sup_{\phi\in B_{p,n}} \mathbb{E}[\phi(X)-\phi(Y)], \\ d_\infty^{p,n}(X,Y) &= \sup_{\phi\in B_{p,n}\cap\mathcal{C}^\infty(\mathbb{R})} \mathbb{E}[\phi(X)-\phi(Y)], \\ d_c^{p,n}(X,Y) &= \sup_{\phi\in B_{p,n}\cap\mathcal{C}_c^\infty(\mathbb{R})} \mathbb{E}[\phi(X)-\phi(Y)], \end{aligned}$$

where $\mathcal{C}_c^\infty(\mathbb{R})$ the space of smooth functions with compact support. We also note

$$\mathcal{W}_n(X,Y) = d_W^{0,n}(X,Y). \tag{4.1}$$

Lemma 5. *Assume X and Y admit a moment of order p. Then*

$$d_W^{p,n}(X,Y) = d_\infty^{p,n}(X,Y).$$

Proof. It is clear that $d_\infty^{p,n}(X,Y) \leq d_W^{p,n}(X,Y)$.

Assume first that $\mathbb{E}[X^k] > \mathbb{E}[Y^k]$ for some $k \leq p-1$. Let $\phi_A(x) = Ax^k$ with $A > 0$. Then $\phi \in B_{p,n} \cap \mathcal{C}^\infty(\mathbb{R})$, so $d_\infty^{p,n}(X,Y) \geq A(\mathbb{E}[X^k]-\mathbb{E}[Y^k])$, leading to $d_\infty^{p,n}(X,Y) = d_W^{p,n}(X,Y) = +\infty$. From now on, we assume that $\mathbb{E}[X^k] = \mathbb{E}[Y^k]$ for every $0 \leq p \leq k-1$.

Let ω be the pdf of the standard normal distribution, $0 < \epsilon < 1$ and $\omega_\epsilon(x) = \epsilon^{-1}\omega(\epsilon^{-1}x)$. Let $\phi \in B_{p,n}$ and define

$$\phi_\epsilon(x) = \int \phi(y)\omega_\epsilon(x-y)dy.$$

Since for every $k \leq p-1$, the moments of order k of X and Y match, on can assume without loss of generality that $\phi(0) = \ldots = \phi^{(p-1)}(0) = 0$. Since $||\phi^{(p)}||_\infty \leq 1$, then

$$\forall\, 0 \leq k \leq p, |\phi^{(k)}(x)| \leq C|x|^{p-k}, \tag{4.2}$$

where C is a constant not depending on ϕ. From (4.2) applied to $k = 0$, et by Lebesgue's dominated convergence theorem, we have that $\phi_\epsilon \in \mathcal{C}^\infty(\mathbb{R})$, and for all $k \leq n$, $\phi_\epsilon^{(k)}(x) = \int \phi^{(k)}(x-y)\omega_\epsilon(y)dy$. Thus $\phi_\epsilon^{(k)} \in B_{p,n} \cap \mathcal{C}^\infty(\mathbb{R})$. We deduce

$$\begin{aligned} \mathbb{E}[\phi(X)-\phi(Y)] &= \mathbb{E}[\phi_\epsilon(X)-\phi_\epsilon(Y)] + \mathbb{E}[\phi(X)-\phi_\epsilon(X)] + \mathbb{E}[\phi_\epsilon(Y)-\phi(Y)] \\ &\leq d_\infty^{p,n}(X,Y) + \mathbb{E}[\phi_\epsilon(X)-\phi(X)] + \mathbb{E}[\phi_\epsilon(Y)-\phi(Y)]. \end{aligned} \tag{4.3}$$

Now, note that $\mathbb{E}[\phi_\epsilon(X)] = \mathbb{E}[\phi(X+\epsilon Z)]$, where Z is a standard normal r.v. independent of X. Then we have

$$\begin{aligned} |\mathbb{E}[\phi(X+\epsilon Z)-\phi(X)]| &\leq \epsilon \sum_{k=1}^{p-1} \epsilon^{k-1}\mathbb{E}[\frac{|\phi^{(k)}(X)|}{k!}|Z|^k] + \epsilon^p\mathbb{E}[\frac{|\phi^{(p)}(\tilde{X})|}{p!}|Z|^p] \\ &\leq C\epsilon \sum_{k=1}^{p-1} \mathbb{E}[|X|^{p-k}]\mathbb{E}[|Z|^k] + C\epsilon^p\mathbb{E}[|Z|^p] \\ &\leq C\epsilon, \end{aligned}$$

where $\tilde{X} \in [X, X+\epsilon Z]$ and C does not depend on ϕ nor on ϵ. Apply this inequality to (4.3), take the supremum over ϕ and let ϵ go to 0 to achieve the proof. ■

We now state a technical Lemma.

Lemma 6. *For every $n \geq 1$, there exists $C > 0$ and $\epsilon_0 > 0$ such that, for all $\epsilon \in (0, \epsilon_0)$ and $M \geq 1$, there exists $f \in \mathcal{C}_c^\infty(\mathbb{R})$ with values in $[0,1]$ verifying the following :*

$$f(x) = 1, \; \forall x \in [-M, M],$$

and

$$|f^{(k)}(x)| \leq \frac{C\epsilon}{|x|}, \quad \forall x \in \mathbb{R}, \forall k \in [|1, n|].$$

Proof. Let $g \in \mathcal{C}^\infty(\mathbb{R})$ with values in $[0,1]$ such that $g(x) = 0$ for all $x \leq 0$ and $g(x) = 1$ for all $x \geq 1$. We will first define the derivative of the function f, then integrate. Let $\epsilon > 0$, $M \geq 1$ and $A \geq M + 1$. We define $f_{A,\epsilon}$ as follows :

$$f_{A,\epsilon}(x) = \begin{cases} \frac{\epsilon g(x-M)}{x} & \text{if } x \leq M+1 \\ \frac{\epsilon g(A+1-x)}{x} & \text{if } x \geq A \\ \frac{\epsilon}{x} & \text{if } x \in [M+1, A]. \end{cases}$$

Let $h(A) = \int_{\mathbb{R}} f_{A,\epsilon}(x)dx$. Then $h(M+1) \leq 2\epsilon ||g||_\infty$, so that if $\epsilon_0 < \frac{1}{2||g||_\infty}$, we have that for all $\epsilon < \epsilon_0$, $h(M+1) < 1$. On the other hand, $h(A) \to +\infty$ when $A \to +\infty$, and since h is continuous one can choose A such that $h(A) = 1$. Then we define for every $x \in \mathbb{R}$,

$$f(x) = 1 - \int_0^{|x|} f_{A,\epsilon}(t)dt.$$

Let us prove that f has the required properties. It is clear that $f \in \mathcal{C}_c^\infty(\mathbb{R})$ and that $f(x) = 1$ if $x \in [-M, M]$. If $1 \leq k \leq n$ and $A \leq x \leq A + 1$,

$$\begin{aligned} |f^{(k)}(x)| &\leq \epsilon \sum_{l=0}^{k-1} C_{k-1}^l l! \frac{||g^{(k-1-l)}||_\infty}{|x|^{1+l}} \\ &\leq \frac{\epsilon C}{|x|}, \end{aligned}$$

where $C = n! 2^n \max_{0 \leq l \leq n} ||g^{(l)}||_\infty$. The same bound holds on $[M, M+1]$. On $[M+1, A]$, we have $|f^{(k)}(x)| \leq \epsilon n!/|x|$. Since f is even, the same bounds hold for $x \leq 0$. ∎

Lemma 7. *Assume $p = 0$ or $p = 1$, and $n \geq 1$. Then, for every random variables X, Y with finite moment of order p,*

$$d_W^{p,n}(X, Y) = d_c^{p,n}(X, Y).$$

Proof. From Lemma 5, it suffices to show that $d_\infty^{p,n} = d_c^{p,n}$.

It is clear that $d_c \leq d_\infty$.

Let $\epsilon > 0$. Let $M > 1$ such that if $K = [-M, M]$,

$$\mathbb{E}[|X|^p \, \mathbf{1}_{X \in K^c}] \leq \epsilon,$$

$$\mathbb{E}[|Y|^p \, \mathbf{1}_{Y \in K^c}] \leq \epsilon.$$

Let $\phi \in B_{p,n} \cap \mathcal{C}^\infty(\mathbb{R})$. We may assume that $\phi(0) = 0$, in which case, since $||\phi'||_\infty \leq 1$, $|\phi(x)| \leq |x|$.

Let f be the function defined in Lemma 6 and $\tilde{\phi} = \phi f$. In the following, C is a constant independent of ϵ and ϕ which may vary from line to line. Then for all $k \in [|1, n|]$:

$$\begin{aligned}
|\tilde{\phi}^{(k)}(x)| &= \left| \sum_{l=0}^{k} C_k^l f^{(l)}(x) \phi^{(k-l)}(x) \right| \\
&= \left| f(x)\phi^{(k)}(x) + \sum_{l=1}^{k-p} C_k^l f^{(l)}(x)\phi^{(k-l)}(x) + \sum_{l=k-p+1}^{k} C_k^l f^{(l)}(x)\phi^{(k-l)}(x) \right| \\
&\leq 1 + C\epsilon + |f^{(k)}(x)||\phi(x)| \\
&\leq 1 + C\epsilon + \frac{C\epsilon}{|x|}|x| \\
&\leq 1 + C\epsilon
\end{aligned}$$

Thus, $\tilde{\phi}/(1 + C\epsilon) \in B_n \cap \mathcal{C}_c^\infty(\mathbb{R})$. This leads to

$$\begin{aligned}
\mathbb{E}[\phi(X) - \phi(Y)] &= \mathbb{E}[\tilde{\phi}(X) - \tilde{\phi}(Y)] + \mathbb{E}[\phi(X)(1 - f(X))] - \mathbb{E}[\phi(Y)(1 - f(Y))] \\
&\leq (1 + C\epsilon) d_c(X, Y) + 2\epsilon.
\end{aligned}$$

By taking the supremum over ϕ and letting ϵ go to zero, we obtain $d_\infty \leq d_c$, which achieves the proof. ■

4.2 Technical proofs

Proof of Proposition 2. Let us prove the first bullet point. We denote by F_X the distribution function of X. Let $\xi > 0$. By definition and using the fact that $\mathbb{E}[X] = 0$, we have:

$$\begin{aligned}
\mathbb{E}\big[iX \exp\big(i\xi X\big)\big] &= i \int_{\mathbb{R}} x e^{i\xi x} dF_X(x), \\
&= i \int_{\mathbb{R}} x(e^{i\xi x} - 1) dF_X(x), \\
&= i \int_0^{+\infty} x(e^{i\xi x} - 1) dF_X(x) + i \int_{-\infty}^0 x(e^{i\xi x} - 1) dF_X(x)
\end{aligned}$$

Let us deal with the first integral. For this purpose, we fix some interval $[a, b]$ strictly contained in $(0, +\infty)$. We note that F_X is of bounded variation on $[a, b]$ as well as the function $x \to c/x^\alpha$. Thus, the function $x \to a_1(x)/x^\alpha$ is of bounded variation on $[a, b]$. We can write:

$$\begin{aligned}
\int_a^b x(e^{i\xi x} - 1) dF_X(x) &= \alpha c \int_a^b (e^{i\xi x} - 1) \frac{dx}{x^\alpha} - \int_a^b x(e^{i\xi x} - 1) d\big(\frac{a_1(x)}{x^\alpha}\big), \\
&= \alpha c \int_a^b (e^{i\xi x} - 1) \frac{dx}{x^\alpha} + i\xi \int_a^b x e^{i\xi x} \frac{a_1(x)}{x^\alpha} dx + \int_a^b (e^{i\xi x} - 1) \frac{a_1(x)}{x^\alpha} dx \\
&- \left[x(e^{i\xi x} - 1) \frac{a_1(x)}{x^\alpha} \right]_a^b,
\end{aligned}$$

where we have performed an integration by parts on the last line. Using the assumptions upon the function $a_1(.)$, we can let a and b tend to 0^+ and $+\infty$ respectively in order to obtain:

$$\int_0^{+\infty} x(e^{i\xi x}-1)dF_X(x) = \alpha c \int_0^{+\infty} (e^{i\xi x}-1)\frac{dx}{x^\alpha} + i\xi \int_0^{+\infty} x e^{i\xi x}\frac{a_1(x)}{x^\alpha}dx + \int_0^{+\infty} (e^{i\xi x}-1)\frac{a_1(x)}{x^\alpha}dx$$

Note in particular that the boundary terms disappear since $\alpha \in (1,2)$. Thus,

$$\int_0^{+\infty} x(e^{i\xi x}-1)dF_X(x) = c\alpha\xi^{\alpha-1} \int_0^{+\infty} (e^{ix}-1)\frac{dx}{x^\alpha} + i\xi \int_0^{+\infty} x e^{i\xi x}\frac{a_1(x)}{x^\alpha}dx + \int_0^{+\infty} (e^{i\xi x}-1)\frac{a_1(x)}{x^\alpha}dx$$

To compute the first integral, we perform the following contour integration: we consider the function of the complex variable $f(z) = (e^{iz}-1)/z^\alpha$ and we integrate it along the contour formed by the upper quarter circle of radius R minus the upper quarter circle of radius $0 < r < R$. Letting $r \to 0$ and $R \to +\infty$, we obtain:

$$\int_0^{+\infty} (e^{ix}-1)\frac{dx}{x^\alpha} = ie^{-\frac{i\pi\alpha}{2}}\frac{\Gamma(2-\alpha)}{1-\alpha}.$$

Thus, we have:

$$\int_0^{+\infty} x(e^{i\xi x}-1)dF_X(x) = ic\alpha\xi^{\alpha-1}e^{-\frac{i\pi\alpha}{2}}\frac{\Gamma(2-\alpha)}{1-\alpha} + i\xi^{\alpha-1}\int_0^{+\infty} e^{ix}\frac{a_1\left(\frac{x}{\xi}\right)}{x^{\alpha-1}}dx + \xi^{\alpha-1}\int_0^{+\infty} (e^{ix}-1)\frac{a_1\left(\frac{x}{\xi}\right)}{x^\alpha}dx.$$

Similarly, for the negative values of X, we have the following:

$$\int_{-\infty}^{0} x(e^{i\xi x}-1)dF_X(x) = i\alpha c\xi^{\alpha-1}e^{+\frac{i\pi\alpha}{2}}\frac{\Gamma(2-\alpha)}{1-\alpha} + i\xi^{\alpha-1}\int_{-\infty}^{0} e^{ix}\frac{a_2\left(\frac{x}{\xi}\right)}{(-x)^{\alpha-1}}dx - \xi^{\alpha-1}\int_{-\infty}^{0} (e^{ix}-1)\frac{a_2\left(\frac{x}{\xi}\right)}{(-x)^\alpha}dx.$$

Thus, using the explicit expression for c, we obtain:

$$\mathbb{E}\big[iX\exp\big(i\xi X\big)\big] = -\alpha\xi^{\alpha-1} - \xi^{\alpha-1}\int_0^{+\infty} e^{ix}\frac{a_1\left(\frac{x}{\xi}\right)}{x^{\alpha-1}}dx + i\xi^{\alpha-1}\int_0^{+\infty} (e^{ix}-1)\frac{a_1\left(\frac{x}{\xi}\right)}{x^\alpha}dx - \xi^{\alpha-1}\int_{-\infty}^{0} e^{ix}\frac{a_2\left(\frac{x}{\xi}\right)}{(-x)^{\alpha-1}}dx - i\xi^{\alpha-1}\int_{-\infty}^{0} (e^{ix}-1)\frac{a_2\left(\frac{x}{\xi}\right)}{(-x)^\alpha}dx.$$

Now, using the assumptions upon the functions $a_1(.)$ and $a_2(.)$ and Lebesgue dominated convergence theorem, we have:

$$\phi_X^*(0^+) = 1.$$

We proceed similarly for $\phi_X^*(0^-)$. The second bullet point is trivial. Let us prove the third and fourth ones. We consider the following linear functional on $S\big(\mathbb{R}\big)$:

$$F:\ \psi \longrightarrow \frac{1}{2\pi}\int_{\mathbb{R}} \mathcal{F}\big(\psi\big)(\xi)\phi_X^*(\xi)d\xi.$$

By the previous computations, for any $\xi > 0$ (and similarly for $\xi < 0$), we have:

$$\begin{aligned}\phi_X^*(\xi) = 1 &+ \frac{1}{\alpha}\int_0^{+\infty} e^{ix}\frac{a_1\big(\frac{x}{\xi}\big)}{x^{\alpha-1}}dx\\ &-\frac{i}{\alpha}\int_0^{+\infty}(e^{ix}-1)\frac{a_1\big(\frac{x}{\xi}\big)}{x^{\alpha}}dx + \frac{1}{\alpha}\int_{-\infty}^{0} e^{ix}\frac{a_2\big(\frac{x}{\xi}\big)}{(-x)^{\alpha-1}}dx\\ &+\frac{i}{\alpha}\int_{-\infty}^{0}(e^{ix}-1)\frac{a_2\big(\frac{x}{\xi}\big)}{(-x)^{\alpha}}dx.\end{aligned}$$

This implies the following simple bound on $\phi_X^*(\xi)$:

$$\mid \phi_X^*(\xi) \mid \leq 1 + C_\alpha\big(\mid\mid a_1 \mid\mid_\infty + \mid\mid a_2 \mid\mid_\infty + \mid\mid xa_1(.) \mid\mid_\infty \mid \xi \mid + \mid\mid xa_2(.) \mid\mid_\infty \mid \xi \mid\big),$$

for some constant, C_α, strictly positive depending on α only. Thus, we have:

$$\mid F(\psi) \mid \leq C_{\alpha,X,p,q} \mid\mid \psi \mid\mid_{p,q,\infty},$$

for some $p, q \in \mathbb{N}$ such that,

$$\mid\mid \psi \mid\mid_{p,q,\infty} = \sup_{x\in\mathbb{R}}\big((1+\mid x \mid)^p \mid \psi^{(q)}(x) \mid\big).$$

This proves the third bullet point. For the fourth one, we introduce the following subspace of $S\big(\mathbb{R}\big)$:

$$S_0\big(\mathbb{R}\big) = \{\psi \in S\big(\mathbb{R}\big):\ \forall n \in \mathbb{N},\ \int_{\mathbb{R}} \psi(x)x^n dx = 0\}.$$

Note that this space is invariant under the action of the pseudo-differential operator, $\mathcal{D}^{\alpha-1}$. Let $\psi \in S_0\big(\mathbb{R}\big)$. We have:

$$\begin{aligned}< T_X^*; \mathcal{D}^{\alpha-1}(\psi) > &= F\big(\mathcal{D}^{\alpha-1}(\psi)\big),\\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \mathcal{F}\big(\mathcal{D}^{\alpha-1}(\psi)\big)(\xi)\phi_X^*(\xi)d\xi,\\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \mathcal{F}\big(\psi\big)(\xi)\big(\frac{i\alpha \mid \xi \mid^\alpha}{\xi}\big)\frac{-\xi}{\alpha \mid \xi \mid^\alpha}(\phi_X(\xi))' d\xi,\\ &= \frac{i}{2\pi}\int_{\mathbb{R}} (\mathcal{F}\big(\psi\big)(\xi))'\phi_X(\xi)d\xi,\\ &= \mathbb{E}\big[X\psi(X)\big]\end{aligned}$$

To conclude, we have to prove the fifth bullet point. Let $\xi > 0$. Note that the characteristic function of Z^α can not be equal to zero. Using straightforward computations, we have:

$$\begin{aligned}\frac{d}{d\xi}\bigg(\frac{\phi_X(\xi) - \phi_{Z^\alpha}(\xi)}{\phi_{Z^\alpha}(\xi)}\bigg)\phi_{Z^\alpha}(\xi) &= \phi_X'(\xi) - \phi_X(\xi)\frac{\phi_{Z^\alpha}'(\xi)}{\phi_{Z^\alpha}(\xi)},\\ &= -\alpha\xi^{\alpha-1}\phi_X^*(\xi) + \phi_X(\xi)\alpha\xi^{\alpha-1},\\ &= \alpha\xi^{\alpha-1}\big(\phi_X(\xi) - \phi_X^*(\xi)\big).\end{aligned}$$

The result follows by integration. ■

Proof of Lemma 2. Let $\xi > 0$. By definition and using the fact that X is centered, we have:

$$\begin{aligned}
\phi_X(\xi) - 1 &= \int_{\mathbb{R}} \big(e^{ix\xi} - 1\big) dF_X(x), \\
&= \int_{\mathbb{R}} \big(e^{ix\xi} - 1 - ix\xi\big) dF_X(x), \\
&= \int_0^{+\infty} \big(e^{ix\xi} - 1 - ix\xi\big) dF_X(x) + \int_{-\infty}^0 \big(e^{ix\xi} - 1 - ix\xi\big) dF_X(x)
\end{aligned}$$

Let us deal with the first term. By integration by parts, we have:

$$\int_0^{+\infty} \big(e^{ix\xi} - 1 - ix\xi\big) dF_X(x) = i\xi \int_0^{+\infty} \big(e^{ix\xi} - 1\big) \frac{c + a_1(x)}{x^\alpha} dx$$

Note that the boundary terms disappear since $\alpha \in (1, 2)$. We have:

$$\begin{aligned}
\int_0^{+\infty} \big(e^{ix\xi} - 1 - ix\xi\big) dF_X(x) &= i\xi \Big[c \int_0^{+\infty} \big(e^{ix\xi} - 1\big) \frac{dx}{x^\alpha} + \int_0^{+\infty} \big(e^{ix\xi} - 1\big) \frac{a_1(x)}{x^\alpha} dx\Big] \\
&= -c\xi^\alpha e^{-i\frac{\pi}{2}\alpha} \frac{\Gamma(2-\alpha)}{1-\alpha} + i\xi^\alpha \int_0^{+\infty} \big(e^{ix} - 1\big) \frac{a_1\big(\frac{x}{\xi}\big)}{x^\alpha} dx.
\end{aligned}$$

Similarly, for the term concerning negative values of X, we obtain the following expression:

$$\int_{-\infty}^0 \big(e^{ix\xi} - 1 - ix\xi\big) dF_X(x) = -c\xi^\alpha e^{i\frac{\pi}{2}\alpha} \frac{\Gamma(2-\alpha)}{1-\alpha} - i\xi^\alpha \int_{-\infty}^0 \big(e^{ix} - 1\big) \frac{a_2\big(\frac{x}{\xi}\big)}{(-x)^\alpha} dx$$

Thus, we have:

$$\begin{aligned}
\phi_X(\xi) - 1 = -\xi^\alpha &+ i\xi^\alpha \int_0^{+\infty} \big(e^{ix} - 1\big) \frac{a_1\big(\frac{x}{\xi}\big)}{x^\alpha} dx \\
&- i\xi^\alpha \int_{-\infty}^0 \big(e^{ix} - 1\big) \frac{a_2\big(\frac{x}{\xi}\big)}{(-x)^\alpha} dx.
\end{aligned}$$

This expression leads to the following easy bound:

$$\mid \phi_X(\xi) - 1 \mid \leq C_\alpha \mid \xi \mid^\alpha \big(1 + \| a_1 \|_\infty + \| a_2 \|_\infty \big).$$

A similar reasoning for $\xi < 0$ ends the proof of the lemma. ■

Proof of Lemma 3. Let $R > 0$ and $\xi > 0$. We have:

$$\Big| \int_0^{+\infty} (e^{ix} - 1) \frac{a_1\big(\frac{x}{\xi}\big)}{x^\alpha} dx \Big| \leq \Big| \int_0^R (e^{ix} - 1) \frac{a_1\big(\frac{x}{\xi}\big)}{x^\alpha} dx \Big| + \Big| \int_R^{+\infty} (e^{ix} - 1) \frac{a_1\big(\frac{x}{\xi}\big)}{x^\alpha} dx \Big|.$$

Let us deal with the first term of the right-hand side of the previous inequality:

$$\begin{aligned}
\Big| \int_0^R (e^{ix} - 1) \frac{a_1\big(\frac{x}{\xi}\big)}{x^\alpha} dx \Big| &\leq \| a_1 \|_\infty \int_0^R \frac{dx}{x^{\alpha-1}}, \\
&\leq \| a_1 \|_\infty \frac{R^{2-\alpha}}{2-\alpha}.
\end{aligned}$$

For the second term, we have:

$$\Big| \int_R^{+\infty} (e^{ix}-1)\frac{a_1\big(\frac{x}{\xi}\big)}{x^\alpha}dx \Big| \leq 2\int_R^{+\infty}\frac{\mid a_1\big(\frac{x}{\xi}\big)\mid}{x^\alpha}dx,$$
$$\leq 2\mid\mid xa_1(.)\mid\mid_\infty \mid \xi\mid \int_R^{+\infty}\frac{dx}{x^{\alpha+1}},$$
$$\leq 2\mid\mid xa_1(.)\mid\mid_\infty \mid \xi\mid \frac{R^{-\alpha}}{\alpha}.$$

Optimising in $R>0$ leads to:

$$\Big| \int_0^{+\infty} (e^{ix}-1)\frac{a_1\big(\frac{x}{\xi}\big)}{x^\alpha}dx \Big| \leq C^1_{\alpha,1}\mid \xi\mid^{\frac{2-\alpha}{2}},$$

for some $C^1_{\alpha,1}>0$ depending on $\mid\mid xa_1(.)\mid\mid_\infty$, $\mid\mid a_1\mid\mid_\infty$ and α only. We proceed exactly in the same way to obtain the appropriate estimates for the other integrals. ■

Acknowledgements

YS gratefully acknowledges support from the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy). YS also thanks Guy Louchard for presenting the problem of approximating the distribution of the *src* statistic, as well as Giovanni Peccati for enlightening discussions which led to some of the main arguments in Section 3.1. The research of BA and GM is funded by a Welcome Grant from Université de Liège.

References

[1] B. Arras, E. Azmoodeh, G. Poly et Y. Swan. Stein's method on the second Wiener chaos: 2-Wasserstein distance. *arXiv preprint arXiv:1601.03301*, 2016.

[2] R. Arratia and P. Baxendale. Bounded size bias coupling: a gamma function bound, and universal dickman-function behavior. *arXiv preprint arXiv:1306.0157*, 2013.

[3] R. Arratia and L. Goldstein. Size bias, sampling, the waiting time paradox, and infinite divisibility: when is the increment independent? *arXiv preprint arXiv:1007.3910*, 2010.

[4] R. Arratia, L. Goldstein and F. Kochman. Size bias for one and all. *arXiv preprint arXiv:1308.2729*, 2013.

[5] E. Azmoodeh, G. Peccati and G. Poly Convergence towards linear combinations of chi-squared random variables: a Malliavin-based approach. *Séminaire de Probabilités* XLVII (Special volume in memory of Marc Yor), 339–367, 2014.

[6] A. D. Barbour and B. Nietlispach. Approximation by the Dickman distribution and quasi-logarithmic combinatorial structures. *Electron. J. Probab.* **16**, 880–902, 2011.

[7] L. H. Y. Chen, L. Goldstein, and Q.-M. Shao. *Normal Approximation by Stein's Method.* Probability and its Applications (New York). Springer, Heidelberg, 2011.

[8] L. H. Y. Chen and A. Röllin. Stein couplings for normal approximation. *arXiv preprint arXiv:1003.6039*, 2010.

[9] P. Dey Stein's method for alpha-stable distributions Oral presentation at Workshop on new directions on Stein's method, Singapore, May 2015.

[10] C. Döbler, *Stein's method of exchangeable pairs for the beta distribution and generalizations*, Electronic Journal of Probability **20** (2015), 1–34.

[11] C. Döbler, R. E. Gaunt, and S. J. Vollmer, *An iterative technique for bounding derivatives of solutions of Stein equations*, arXiv preprint arXiv:1510.02623 (2015).

[12] R. Gaunt, G. Mijoule and Y. Swan. Stein operators for product distributions, with applications. *arXiv preprint arXiv:1604.06819* (2016).

[13] A. L. Gibbs and F. E. Su. On choosing and bounding probability metrics. *Int. stat. rev.* **70**(3), 419–435, 2002.

[14] 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.

[15] L. Goldstein and G. Reinert. Stein's method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.*, **7**(4), 935–952, 1997.

[16] L. Goldstein and G. Reinert, Distributional transformations, orthogonal polynomials, and Stein characterizations, *J. Theor. Probab.* **18**, 237–260, 2005.

[17] J. Gorham and L. Mackey. Multivariate Stein Factors for Strongly Log-concave Distributions. *arXiv preprint arXiv:1512.07392*, 2015

[18] H.-K. Hwang and T.-H. Tsai. Quickselect and the dickman function. *Combinatorics Probability and Computing*, 11(4):353–371, 2002.

[19] I. A. Ibragimov and Ju. V. Linnik. Ju. V. *Nezavisimye stalionarno svyazannye velichiny.* Izdat. "Nauka", Moscow, 1965.

[20] O. Johnson and R. Samworth. Central limit theorem and convergence to stable laws in Mallows distance. *Bernoulli*, 11(5): 829–845, 2005.

[21] G. Louchard. Sum of positions of records in random permutations: asymptotic analysis. *Online Journal of Analytic Combinatorics* **9** (2014).

[22] I. Nourdin and G. Peccati. *Normal Approximations Using Malliavin Calculus: from Stein's Method to Universality.* Cambridge Tracts in Mathematics. Cambridge University, 2012.

[23] I. Nourdin and G. Peccati. Cumulants on the Wiener space. *J. Funct. Anal.*, 258(11): 3775–3791, 2010.

[24] D. Nualart. The Malliavin calculus and related topics. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[25] S. T. Rachev and L. Rüschendorf. On the rate of convergence in the CLT with respect to the Kantorovich metric. *Probability in Banach Spaces, 9*, 193–207, 1994.

[26] A. Rényi. Théorie des éléments saillants d'une suite d'observations. *Annales de la Faculté des sciences de l'Université de Clermont. Série Mathématiques*, 8(2):7–13, 1962.

[27] G. Reinert and A. Röllin, Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition *Ann. Prob.* **37**, 2150–2173, 2009.

[28] A. Röllin, *On the optimality of Stein factors*, *Propb. Approx. and Beyond*, 61–72, 2012.

[29] C. Stein. *Approximate Computation of Expectations.* Institute of Mathematical Statistics Lecture Notes—Monograph Series, 7. Institute of Mathematical Statistics, Hayward, CA, 1986.

[30] F.W. Steutel and K. van Harn Infinite divisibility of probability distributions on the real line *CRC Press*, 2003.

[31] A.N. Tikhomirov. On the convergence rate in the central limit theorem for weakly dependent random variables. *Theory of Probability & Its Applications*, **25**(4), pp.790–809, 1981