

Properties and numerical evaluation of the Rosenblatt distribution

MARK S. VEILLETTE* and MURAD S. TAQQU**

*Department of Mathematics and Statistics, Boston University, 111 Cummington Street, Boston, MA 02215, USA. E-mail: *mveillet@bu.edu; **murad@bu.edu*

This paper studies various distributional properties of the Rosenblatt distribution. We begin by describing a technique for computing the cumulants. We then study the expansion of the Rosenblatt distribution in terms of shifted chi-squared distributions. We derive the coefficients of this expansion and use these to obtain the Lévy–Khintchine formula and derive asymptotic properties of the Lévy measure. This allows us to compute the cumulants, moments, coefficients in the chi-square expansion and the density and cumulative distribution functions of the Rosenblatt distribution with a high degree of precision. Tables are provided and software written to implement the methods described here is freely available by request from the authors.

Keywords: Edgeworth expansions; long range dependence; Rosenblatt distribution; self-similarity

1. Introduction

Typical limits of normalized of sums of long-range dependent stationary series are Brownian motion, fractional Brownian motion or the Rosenblatt process. Brownian motion and fractional Brownian motion are Gaussian and can thus be readily tabulated. This is not the case for the Rosenblatt distribution. The goal of this paper is to fill this gap. The tables can be used to compute asymptotic confidence intervals and to implement maximum likelihood methods.

The Rosenblatt distribution is the simplest non-Gaussian distribution which arises in a *non-central limit theorem* involving long-range dependent random variables [11,29,30]. For an overview, see [31]. It also appears in a statistical context as the asymptotic distribution of certain estimators (e.g., [32]).

We shall begin by motivating the Rosenblatt distribution using Rosenblatt’s famous counterexample found in [19]. Consider a stationary Gaussian sequence X_i , $i = 1, 2, \dots$ which has a covariance structure of the form $\mathbb{E}X_0X_k \sim k^{-D}$ as $k \rightarrow \infty$ with $0 < D < 1/2$. Using the transformation

$$Y_i = X_i^2 - 1, \tag{1}$$

one can define a sequence of normalized sums

$$Z_D^N = \frac{\sigma(D)}{N^{1-D}} \sum_{i=1}^N Y_i. \tag{2}$$

Here, $\sigma(D)$ is a normalizing constant and is given by

$$\sigma(D) = \left[\frac{1}{2}(1 - 2D)(1 - D) \right]^{1/2}. \tag{3}$$

The sequence Z_D^N tends to a non-Gaussian limit Z_D as $N \rightarrow \infty$ with mean 0 and variance 1. This limiting distribution was named the Rosenblatt distribution in [29]. The characteristic function of Z_D can be given as the following power series which is only convergent near the origin:

$$\phi(\theta) = \exp\left(\frac{1}{2} \sum_{k=2}^{\infty} (2i\theta\sigma(D))^k \frac{c_k}{k}\right), \tag{4}$$

where

$$c_k = \int_0^1 dx_1 \int_0^1 dx_2 \cdots \int_0^1 dx_k |x_1 - x_2|^{-D} |x_2 - x_3|^{-D} \cdots |x_{k-1} - x_k|^{-D} |x_k - x_1|^{-D}. \tag{5}$$

By Cauchy–Schwarz,

$$c_k \leq \left(\int_0^1 \int_0^1 |x_1 - x_2|^{-2D} dx_1 dx_2\right)^{k/2} = \left(\frac{1}{(1-2D)(1-D)}\right)^{k/2} = \left(\frac{1}{2\sigma^2(D)}\right)^{k/2}, \tag{6}$$

ensuring that the series (4) converges around the origin. [Since (6) is an equality when $k = 2$, Z_D has variance 1 in view of (4) and (3).]

It is interesting to consider the extremes when $D \rightarrow 0^+$ and $D \rightarrow \frac{1}{2}^-$. When $D \rightarrow 0^+$, notice that $c_k \rightarrow 1$ for all k , $\sigma(D) \rightarrow \frac{1}{\sqrt{2}}$ and thus for θ small enough, the characteristic function approaches

$$\begin{aligned} \phi(\theta) &= \exp\left(\frac{1}{2} \sum_{k=2}^{\infty} \frac{(\sqrt{2}i\theta)^k}{k}\right) \\ &= \exp\left(\frac{1}{2}(\log(1 - \sqrt{2}i\theta) - \sqrt{2}i\theta)\right) \\ &= \left(\frac{1}{1 - \sqrt{2}i\theta}\right)^{1/2} e^{-i\theta/\sqrt{2}}, \end{aligned} \tag{7}$$

which is the characteristic function of $\frac{1}{\sqrt{2}}(\varepsilon^2 - 1)$, where ε is $N(0, 1)$. Hence when $D = 0$, the Rosenblatt distribution is simply a chi-squared distribution standardized to have mean 0 and variance 1.

As $D \rightarrow \frac{1}{2}^-$, the limit is $N(0, 1)$. This is not surprising given that the scaling term in (2) approaches \sqrt{N} , hence you would assume the usual central limit theorem to hold and the limiting distribution to be Gaussian. This fact is not obvious however from the characteristic function (4). In Section 4 of this work, we derive an alternative form of the characteristic function from which this Gaussian limit is easier to see.

The distribution Z_D can be given in terms of a weighted sum of chi-squared distributions,

$$Z_D = \sum_{i=1}^{\infty} \lambda_n (\varepsilon_n^2 - 1), \quad \varepsilon_n \text{ i.i.d. } N(0, 1), \tag{8}$$

where the weights $\{\lambda_n\}_{n=1}^\infty$ are such that

$$\sum_{n=1}^\infty \lambda_n^k = \sigma^k(D)c_k, \quad k = 2, 3, \dots \tag{9}$$

The series (8) converges a.s. and in L^2 because

$$\begin{aligned} \sum_{n=1}^\infty \text{Var}[\lambda_n(\varepsilon_n^2 - 1)] &= \mathbb{E}[(\varepsilon_1^2 - 1)^2] \sum_{n=1}^\infty \lambda_n^2 \\ &= 2 \sum_{n=1}^\infty \lambda_n^2 < \infty. \end{aligned}$$

In fact, $\sum_{n=1}^\infty \lambda_n^2 = 1/2$ by (5) and (3). The weights $\{\lambda_n\}$ are given as the eigenvalues of an integral operator which we will discuss in more detail in Section 3. Various integral representations can be found in [32]. Our main focus in this work is on distributional properties of this distribution, namely, cumulants, moments and obtaining a numerical evaluation of the Rosenblatt distribution. A table for the cumulative distribution function (CDF) of the Rosenblatt distribution is useful for obtaining percentiles and confidence intervals.

This paper is organized as follows: In Section 2, we look at the moments and cumulants of the Rosenblatt distribution, as well as detailing a method for computing them. In Section 3, we show that the λ_n 's in the expansion (8) are given by the eigenvalues of an integral operator, and we give asymptotic formulas for this sequence. In Section 4, we state the characteristic function of the Rosenblatt distribution in Lévy–Khintchine form and use it to derive further properties of the Rosenblatt distribution. In Section 6, we compute the moments, cumulants and the λ_n 's and in Section 7, the quantiles of the Rosenblatt distribution are computed for various D values. For more details and software, see the supplemental article [35].

2. Cumulants and moments of the Rosenblatt distribution

It follows from the expansion of the characteristic function (4) around $\theta = 0$ that the cumulants κ_k of the Rosenblatt distribution are given by $\kappa_1 = 0$ and

$$\kappa_k = 2^{k-1}(k-1)!(\sigma(D))^k c_k, \tag{10}$$

where the c_k are given by the multiple integrals (5). Each moment μ_n , $n \geq 1$ can then be expressed as a polynomial using the cumulants κ_k , $k = 1, 2, \dots, n$. These are the *complete Bell Polynomials*

$$\mu_n = B_n(0, \kappa_2, \dots, \kappa_n) \tag{11}$$

as noted, for example, in [17,18]. They can also be computed recursively [24].

Thus, in order to compute any moment or cumulant, it is necessary to compute the multiple integrals c_k . The first two can be computed directly,

$$\begin{aligned}
 c_2 &= \int_0^1 \int_0^1 |x_1 - x_2|^{-2D} dx_1 dx_2 = 2 \int_0^1 \int_0^{x_2} (x_1 - x_2)^{-2D} dx_1 dx_2 = \frac{1}{(1 - 2D)(1 - D)}, \\
 c_3 &= \int_0^1 \int_0^1 \int_0^1 |x_1 - x_2|^{-D} |x_2 - x_3|^{-D} |x_3 - x_1|^{-D} dx_1 dx_2 dx_3 \\
 &= 3 \int_0^1 x_3^{3D} \int_0^{x_3} \int_0^{x_3} \left| \frac{x_1}{x_3} - \frac{x_2}{x_3} \right|^{-D} \left(1 - \frac{x_2}{x_3} \right)^{-D} \left(1 - \frac{x_1}{x_3} \right)^{-D} dx_1 dx_2 dx_3 \\
 &= 3 \left(\int_0^1 x_3^{-3D+2} dx_3 \right) \left(\int_0^1 \int_0^1 |u_1 - u_2|^{-D} (1 - u_2)^{-D} (1 - u_1)^{-D} du_1 du_2 \right) \tag{12} \\
 &= \frac{2}{1 - D} \int_0^1 w_2^{-D} \int_0^{w_2} (w_2 - w_1)^{-D} w_1^{-D} dw_1 dw_2 \\
 &= \frac{2}{1 - D} \left(\int_0^1 w_2^{-3D+1} dw_2 \right) \left(\int_0^1 v^{-D} (1 - v)^{-D} dv \right) \\
 &= \frac{2}{(1 - D)(2 - 3D)} \beta(1 - D, 1 - D),
 \end{aligned}$$

where $\beta(a, b) = \int_0^1 v^{a-1} (1 - v)^{b-1} dv$ is the beta function and we made the following changes of variables above: $u_1 = x_1/x_3, u_2 = x_2/x_3, w_1 = 1 - u_1, w_2 = 1 - u_2$ and $v = w_1/w_2$.

For $k \geq 4$, a closed form expression for c_k could not be found, which means they must be computed numerically. Computing the multiple integrals directly is intractable due to the increasing number of singularities in the integrand. It is for this reason that we now develop a more sophisticated method for computing c_k .

Let $L^2(0, 1)$ denote the Hilbert space of all real-valued measurable functions $f(x), 0 < x < 1$, such that $\|f\|_2 \equiv (\int_0^1 f(x)^2 dx)^{1/2} < \infty$, together with the usual inner product $(f, g) \equiv \int_0^1 f(x)g(x) dx$. For $0 < D < 1/2$, define the integral operator $\mathcal{K}_D : L^2(0, 1) \rightarrow L^2(0, 1)$ as

$$(\mathcal{K}_D f)(x) = \int_0^1 |x - u|^{-D} f(u) du. \tag{13}$$

Finally, define the sequence of functions $G_{k,D} \in L^2(0, 1), k \geq 1$, recursively as follows:

$$G_{1,D}(x) = \frac{(1 - x)^{-D}}{\sqrt{1 - D}}; \quad G_{k,D}(x) = (\mathcal{K}_D G_{k-1,D})(x), \quad k \geq 2. \tag{14}$$

Then, we have the following alternative way to express c_k .

Proposition 2.1. *Let μ and ν be any two positive integers such that $\mu + \nu = k$. Then*

$$c_k = (G_{\mu,D}, G_{\nu,D}). \tag{15}$$

Proof. Let μ, ν be as stated. Then, using the circular symmetry of the integrand in (5), if we take x_k as the largest of the $x_i, i = 1, 2, \dots$, and then factor an x_k out of all the terms, we can rewrite c_k as

$$c_k = k \int_0^1 x_k^{-kD} \int_{(0,x_k)^{k-1}} \left(1 - \frac{x_1}{x_k}\right)^{-D} \left|\frac{x_1}{x_k} - \frac{x_2}{x_k}\right|^{-D} \dots \times \left|\frac{x_{k-2}}{x_k} - \frac{x_{k-1}}{x_k}\right|^{-D} \left(1 - \frac{x_{k-1}}{x_k}\right)^{-D} dx_1 \dots dx_k. \tag{16}$$

With the change of variables $u_i = x_i/x_k, i = 1, 2, \dots, k - 1$, one of the k integrals can be separated out, and we obtain

$$\begin{aligned} c_k &= k \left(\int_0^1 x_k^{-kD+(k-1)} dx_k \right) \\ &\quad \times \left(\int_{(0,1)^{k-1}} (1 - u_1)^{-D} |u_1 - u_2|^{-D} \dots |u_{k-1} - u_{k-2}|^{-D} (1 - u_{k-1})^{-D} du_1 \dots du_{k-1} \right) \\ &= \frac{1}{1 - D} \left(\int_{(0,1)^{k-1}} (1 - u_1)^{-D} |u_1 - u_2|^{-D} \dots \right. \\ &\quad \left. \times |u_{k-1} - u_{k-2}|^{-D} (1 - u_{k-1})^{-D} du_1 \dots du_{k-1} \right) \\ &= \int_{(0,1)^{k-1}} G_{1,D}(u_1) \underbrace{|u_1 - u_2|^{-D} \dots |u_{k-1} - u_{k-2}|^{-D}}_{k-2 \text{ terms}} G_{1,D}(u_{k-1}) du_1 \dots du_{k-1} \\ &= \int_{(0,1)^{k-3}} \left[\int_0^1 G_{1,D}(u_1) |u_1 - u_2|^{-D} du_1 \right] \left[|u_3 - u_2|^{-D} \dots |u_{k-2} - u_{k-3}|^{-D} \right] \\ &\quad \times \left[\int_0^1 G_{1,D}(u_{k-1}) |u_{k-1} - u_{k-2}|^{-D} du_{k-1} \right] du_3 \dots du_{k-2} \\ &= \int_{(0,1)^{k-3}} G_{2,D}(u_2) \underbrace{|u_2 - u_3|^{-D} \dots |u_{k-2} - u_{k-3}|^{-D}}_{k-4 \text{ terms}} G_{2,D}(u_{k-2}) du_2 \dots du_{k-2} \\ &\quad \vdots \\ &= \int_0^1 G_{\mu,D}(u_\mu) G_{\nu,D}(u_{k-\nu}) du_\mu = (G_{\mu,D}, G_{\nu,D}). \end{aligned}$$

This finishes the proof. □

Remark. To minimize the number of integrals one needs to compute, it makes sense to choose $\mu = \nu = \frac{k}{2}$ if k is even, and $\mu = \frac{k+1}{2}$ and $\nu = \frac{k-1}{2}$ if k is odd. Proposition 2.1 thus reduces

the problem of computing a k -dimensional integral into computing $\lceil \frac{k}{2} \rceil + 1$ one-dimensional integrals.

$G_{2,D}$ can be given in terms of the beta function and the Gauss hypergeometric function ${}_2F_1(a, b; c; x)$, which has the following integral representation,

$$\begin{aligned}
 {}_2F_1(a, b; c; x) &= \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 v^{b-1}(1-v)^{c-b-1}(1-xv)^{-a} dv, \\
 x < 1, c > b > 0.
 \end{aligned}
 \tag{17}$$

Indeed,

$$\begin{aligned}
 G_{2,D}(x) &= \frac{1}{\sqrt{1-D}} \int_0^1 (1-u)^{-D} |x-u|^{-D} du \\
 &= \frac{1}{\sqrt{1-D}} \left[\int_0^x (1-u)^{-D} (x-u)^{-D} du + \int_x^1 (1-u)^{-D} (u-x)^{-D} du \right] \\
 &= \frac{1}{\sqrt{1-D}} \left[x^{1-D} \int_0^1 (1-xv)^{-D} (1-v)^{-D} dv \right. \\
 &\quad \left. + (1-x)^{1-2D} \int_0^1 w^{-D} (1-w)^{-D} dw \right] \\
 &= \frac{1}{\sqrt{1-D}} \left[\frac{x^{1-D}}{1-D} {}_2F_1(D, 1, 2-D, x) + (1-x)^{1-2D} \beta(1-D, 1-D) \right] \\
 &= \frac{x^{1-D}}{(1-D)^{3/2}} {}_2F_1(D, 1, 2-D, x) + \frac{(1-x)^{1-2D} \beta(1-D, 1-D)}{\sqrt{1-D}},
 \end{aligned}
 \tag{18}$$

where, in the third equality we used the change of variables $v = u/x$ and $w = (1-u)/(1-x)$, and in the fourth, we used (17). The function ${}_2F_1(a, b; c; x)$ is bounded for $x \in (0, 1)$ as long as $c > a + b$ ([16], Section 60:7), which is true in this case. This implies that unlike $G_{1,D}$, $G_{2,D}$ is a bounded function on $(0, 1)$, since $0 < D < \frac{1}{2}$.

In Section C of the supplemental article [35], we outline a technique for computing the c_k numerically based on Proposition 2.1, and tabulate the first 8 cumulants and moments of the Rosenblatt distribution for various values of D .

3. Eigenvalue expansion of the Rosenblatt distribution

In this section, we focus on the expansion of the Rosenblatt distribution in terms of shifted chi-squared distributions (8). The sequence $\{\lambda_i\}_{i=1}^\infty$ can be thought of in two ways.

One way is to start with the integrals $\{c_k, k \geq 2\}$ defined in (5) and to view $\{\lambda_i\}_{i=1}^\infty$ as a non-increasing sequence related to these $\{c_k, k \geq 2\}$ through formula (9), see [29]. While easier to state, this perspective sheds little light on the λ_i 's since the c_k 's are so complicated.

The second way to characterize the sequence $\{\lambda_i\}_{i=1}^\infty$ is more useful in our case, and stems from Proposition 2 in [11]. To recall this proposition, let X be defined through the Wiener–Itô integral

$$X = \int''_{\mathbb{R}^2} H(x, y) Z_G(dx) Z_G(dy), \tag{19}$$

where Z_G is a complex-valued random measure with control measure G such that for all Borel sets $A \in \mathbb{R}$, $Z_G(A) = Z_G(-A)$ and $G(A) = G(-A)$, and the kernel $H(x, y)$ is a complex-valued measurable function such that $H(x, y) = H(y, x) = \overline{H(-x, -y)}$ for all $x, y \in \mathbb{R}$, and $\int_{\mathbb{R}^2} |H(x, y)|^2 G(dx) G(dy) < \infty$. The double prime in the integral means to exclude the diagonal $\{x = \pm y\}$. For background on such integrals, see [17] or [15].

Let $L_G^2(\mathbb{R})$ denote the space of complex valued functions $h(x)$, $x \in \mathbb{R}$ such that for all $x \in \mathbb{R}$, $h(x) = \overline{h(-x)}$ and $\int |h(x)|^2 G(dx) < \infty$. For random variables X defined as in (19), Dobrushin and Major showed that X has an expansion

$$X = \sum_{n=1}^\infty \eta_n (\varepsilon_n^2 - 1),$$

where the sequence η_n corresponds to the eigenvalues of the integral operator $A : L_G^2(\mathbb{R}) \rightarrow L_G^2(\mathbb{R}^2)$ defined as

$$(Ah)(x) = \int_{-\infty}^\infty H(x, -y) h(y) G(dy). \tag{20}$$

In [30], it is shown that the Rosenblatt distribution has the following representation as a Wiener–Itô integral:

$$Z_D = a(D) \int''_{\mathbb{R}^2} \frac{e^{i(x+y)} - 1}{i(x+y)} Z_{G_D}(dx) Z_{G_D}(dy), \tag{21}$$

where the measure G_D is absolutely continuous and is given by

$$G_D(dx) = |x|^{D-1} dx, \quad x \in \mathbb{R}. \tag{22}$$

The constant

$$a(D) = \frac{\sigma(D)}{2\Gamma(D) \sin((1-D)\pi/2)}, \tag{23}$$

where $\sigma(D)$ is given in (3), ensures a variance of 1.

Thus, Dobrushin and Major’s result implies that the sequence λ_n we seek in (8) is given as the eigenvalues of the operator $A_D : L_{G_D}^2(\mathbb{R}) \rightarrow L_{G_D}^2(\mathbb{R})$ defined as

$$(A_D h)(x) = a(D) \int_{-\infty}^\infty \frac{e^{i(x-y)} - 1}{i(x-y)} h(y) G_D(dy). \tag{24}$$

We shall now reexpress the eigenvalue problem associated to (24) in a much simpler form so that we may both give analytical results about the λ_n 's and develop a method to compute them. This is done in the following proposition.

Proposition 3.1. *The operators $A_D : L^2_{G_D}(\mathbb{R}) \rightarrow L^2_{G_D}(\mathbb{R})$ defined in (24) and $\sigma(D)\mathcal{K}_D : L^2(0, 1) \rightarrow L^2(0, 1)$ defined in (13) have the same eigenvalues.*

Proof. Let $\widehat{g}(z) = (\mathcal{F}g)(z) = \int_{\mathbb{R}} g(y)e^{iyz} dy$ and $\check{g}(z) = (\mathcal{F}^{-1}g)(z) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ixz} g(y) dy$ denote the Fourier transform and inverse Fourier transform. Recall that \mathcal{F} and \mathcal{F}^{-1} are defined on $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$ and can also be extended to generalized functions ([36], Chapter 7).

Let (λ, h) be an eigenpair of the operator A_D . This implies that $h \in L^2_{G_D}(\mathbb{R})$ and thus $\int h(y)^2 |y|^{D-1} dy < \infty$ by (22). Taking inverse Fourier transforms of both sides in $\lambda h = A_D h$, we obtain

$$\lambda \check{h} = \mathcal{F}^{-1}(A_D h) = a(D)\mathcal{F}^{-1}(H_1 * H_2), \tag{25}$$

where

$$H_1(y) = (\exp(iy) - 1)/(iy)$$

and

$$H_2(y) = |y|^{D-1} h(y).$$

We want to apply the convolution theorem to compute the inverse Fourier transform in (25), however some care must be taken, because, while $H_1 \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, H_2 is not necessarily in either $L^1(\mathbb{R})$ or $L^2(\mathbb{R})$. However, this difficulty can be avoided by writing H_2 as a sum:

$$H_2(y) = |y|^{D-1} h(y) \mathbf{1}_{[-1,1]}(y) + |y|^{D-1} h(y) \mathbf{1}_{(1,\infty)}(|y|) := H_2^-(y) + H_2^+(y).$$

Since $\int h(y)^2 |y|^{D-1} dy < \infty$, we have

$$H_2^-(y) = |y|^{D-1} h(y) \mathbf{1}_{[-1,1]}(y) \in L^1(\mathbb{R})$$

and

$$H_2^+(y) = |y|^{D-1} h(y) \mathbf{1}_{(1,\infty)}(|y|) \in L^2(\mathbb{R}).$$

By linearity of the convolution and Fourier transform, we can apply the convolution theorem for L^1 functions ([27], Proposition 6.2.1) and the convolution theorem for L^2 functions ([27], Proposition 6.5.3), we have

$$\begin{aligned} \lambda \check{h} &= a(D)\mathcal{F}^{-1}(H_1 * H_2) = a(D)[\mathcal{F}^{-1}(H_1 * H_2^-) + \mathcal{F}^{-1}(H_1 * H_2^+)] \\ &= 2\pi a(D)\mathbf{1}_{(0,1)}(\check{H}_2^- + \check{H}_2^+) = 2\pi a(D)\mathbf{1}_{(0,1)}\check{H}_2, \end{aligned}$$

where we have used the fact that $\check{H}_1 = \mathbf{1}_{(0,1)}$ and we have picked up an extra factor of 2π since we are applying the inverse Fourier transform \mathcal{F}^{-1} to the convolution. This implies that for any

eigenfunction h of A_D , the support of \check{h} is contained in $(0, 1)$. Viewing H_2 as the product of $|y|^{D-1}$ and $h(y)$, we again apply the convolution theorem. Again, care must be taken as $|y|^{D-1}$ is not integrable or square integrable, but if we view h and $|y|^{D-1}$ as generalized functions, the convolution theorem still applies since \check{h} has compact support ([36], Theorem 7.9-1). Thus,

$$h(y)|y|^{D-1} = \mathcal{F}(\check{h} * \mathcal{F}^{-1}(|y|^{D-1})).$$

Using this, we have

$$\begin{aligned} 2\pi a(D)\mathbf{1}_{(0,1)}\check{H}_2 &= 2\pi a(D)\mathbf{1}_{(0,1)}\mathcal{F}^{-1}[h(y)|y|^{D-1}] \\ &= 2\pi a(D)\mathbf{1}_{(0,1)}\mathcal{F}^{-1}[\mathcal{F}(\check{h} * \mathcal{F}^{-1}(|y|^{D-1}))] \\ &= 2\pi a(D)\mathbf{1}_{(0,1)}(\check{h} * \mathcal{F}^{-1}(|y|^{D-1})). \end{aligned}$$

The inverse Fourier transform of $|y|^{D-1}$ is given by ([13], page 1119),

$$(\mathcal{F}^{-1}|y|^{D-1})(z) = \frac{1}{2\pi} \left[2\Gamma(D) \sin\left(\frac{(1-D)\pi}{2}\right) |z|^{-D} \right] = \frac{\sigma(D)}{2\pi a(D)} |z|^{-D} \tag{26}$$

from (23). Thus, if (h, λ) is an eigenpair for A_D , then

$$\begin{aligned} \lambda\check{h}(z) &= (\mathcal{F}^{-1}A_D h)(z) \\ &= \sigma(D)\mathbf{1}_{[0,1]}(z) \int_{-\infty}^{\infty} |z-y|^{-D}\check{h}(y) dy \\ &= \sigma(D)\mathbf{1}_{[0,1]}(z) \int_0^1 |z-y|^{-D}\check{h}(y) dy, \end{aligned} \tag{27}$$

where we have again used the fact that \check{h} is supported on $(0, 1)$. Thus, if (λ, h) is an eigenpair for A_D , then $(\lambda, \check{h}|_{(0,1)})$ is an eigenpair for $\sigma(D)\mathcal{K}_D$. Reversing this argument shows that there is a one-to-one correspondence between eigenpairs of A_D and $\sigma_D\mathcal{K}$, which preserves the eigenvalues, hence these operators have the same eigenvalues. This completes the proof. \square

The eigenvalues λ_n of \mathcal{K}_D are not known exactly, but their asymptotic behavior is well understood as $n \rightarrow \infty$, see for instance [7,12,14] or [20]. In particular, we have the following result regarding the asymptotic behavior of λ_n in (8).

Theorem 3.2. *Let Z_D denote the Rosenblatt distribution given by*

$$Z_D = \sum_{n=1}^{\infty} \lambda_n(D)(\varepsilon_n^2 - 1), \quad \varepsilon_n \text{ i.i.d. } N(0, 1), \tag{28}$$

where the non-increasing sequence $\{\lambda_n(D)\}$ is given by the eigenvalues of the integral operator $\sigma(D)\mathcal{K}_D: L^2(0, 1) \rightarrow L^2(0, 1)$. The following asymptotic formula holds for any $0 < r < 1$:

$$\lambda_n(D) = C(D)n^{D-1} \left(1 + o\left(\frac{1}{n^r}\right) \right), \tag{29}$$

where,

$$C(D) = \frac{2}{\pi^{1-D}} \sigma(D) \Gamma(1-D) \sin\left(\frac{\pi D}{2}\right). \tag{30}$$

Moreover, the series

$$\sum_{n=1}^{\infty} (\lambda_n(D) - C(D)n^{D-1})$$

converges and equals

$$\sum_{n=1}^{\infty} (\lambda_n(D) - C(D)n^{D-1}) = -2^{1-D} \sigma(D) \zeta(D), \tag{31}$$

where ζ denotes the Riemann zeta function.

Proof. Proposition 2 in [11] and Proposition 3.1 verify the first claim, namely that (28) holds and that the λ_n 's are the eigenvalues of $\sigma(D)\mathcal{K}_D$. Theorem 1 in [12] describes the eigenvalues of the operator $\int_{-1}^1 |x-u|^{-D} f(u) dy$. Thus, (29)–(31) follow immediately after noting that

$$\int_{-1}^1 |z-u|^{-D} g(u) du = 2^{1-D} \int_0^1 |x-y|^{-D} f(y) dy,$$

where $z = 2x - 1, u = 2y - 1$ and $g(u) = f(\frac{u+1}{2})$. □

Since Theorem 3.2 only gives the asymptotic behavior of the λ_n 's, we need to approximate those for which the asymptotic formula is not applicable. We present a method for approximating these eigenvalues in Section 6 (see also Section D of the supplemental article [35]).

4. Lévy–Khintchine representation of the Rosenblatt distribution

Recall that a distribution X is infinitely divisible if for any integer $n \geq 1$, there exists $X_i^{(n)}, i = 1, 2, \dots, n$, i.i.d. such that $X \stackrel{d}{=} X_1^{(n)} + X_2^{(n)} + \dots + X_n^{(n)}$. The characteristic function of an infinitely divisible distribution X with $\mathbb{E}X^2 < \infty$ can always be written in the following form:

$$\phi(\theta) = \mathbb{E}e^{i\theta X} = \exp\left(ia\theta - \frac{1}{2}b^2\theta^2 + \int_{-\infty}^{\infty} (e^{i\theta u} - 1 - iu\theta)v(du)\right),$$

where $a \in \mathbb{R}, b > 0$ and ν is a positive measure on $\mathbb{R} \setminus \{0\}$ with the property that $\int \min(u^2, 1) \times \nu(du) < \infty$. This is known as the Lévy Khintchine representation of X . Since the chi-square distribution is infinitely divisible, it is not surprising in light of (8) that the Rosenblatt distribution

is also infinitely divisible. In this section, we will make this assertion rigorous, and give the Lévy Khintchine representation of the Rosenblatt distribution.

Before stating the result, we require first a lemma. Given any positive, increasing sequence $\mathbf{c} = \{c_n\}_{n=1}^\infty$ such that $\sum 1/c_n^2 < \infty$, define the function $G_{\mathbf{c}}(x)$ for $0 < x < 1$ as

$$G_{\mathbf{c}}(x) = \sum_{n=1}^\infty x^{c_n}. \tag{32}$$

Since $\mathbf{c}^{-1} = \{c_1^{-1}, c_2^{-1}, \dots\} \in \ell^2$, we have $c_n \rightarrow \infty$ and thus this series converges for all $x \in (0, 1)$ since $\log x < 0$ and hence for n large enough, one has $x^{c_n} = e^{c_n \log(x)} \leq c_n^{-2}$. Notice that $G(0) = 0$, $G_{\mathbf{c}}(x) \rightarrow \infty$ as $x \rightarrow 1$ and is a continuous function for all $x \in (0, 1)$.

We now state a lemma regarding the asymptotic behavior of $G_{\mathbf{c}}$ as $x \rightarrow 0$ and $x \rightarrow 1$. In the following, we will say $a_n \sim b_n$ if $a_n/b_n \rightarrow 1$, and $a_n \lesssim b_n$ if $\limsup a_n/b_n \leq 1$.

Lemma 4.1. *Suppose \mathbf{c} is a positive strictly increasing sequence such that $c_n \sim \beta n^\alpha$ as $n \rightarrow \infty$ for some $1/2 < \alpha < 1$ and constant $\beta > 0$. Then,*

$$G_{\mathbf{c}}(x) \sim x^{c_1}, \quad \text{as } x \rightarrow 0, \tag{33}$$

$$G_{\mathbf{c}}(x) \sim \frac{1}{\alpha \beta^{1/\alpha}} \Gamma\left(\frac{1}{\alpha}\right) (1-x)^{-1/\alpha}, \quad \text{as } x \rightarrow 1. \tag{34}$$

Proof. As $x \rightarrow 0$, we have

$$\frac{G_{\mathbf{c}}(x)}{x^{c_1}} = 1 + \sum_{n=2}^\infty x^{c_n - c_1} \rightarrow 1, \tag{35}$$

since $c_n > c_1$, so that the sum on the right hand side tends to 0. This confirms (33).

For the second assertion, let $0 < \varepsilon < 1$ and let $\beta' = (1 - \varepsilon)\beta$ and $\beta'' = (1 + \varepsilon)\beta$. By assumption, there exists M large enough such that for $n \geq M$,

$$\beta' n^\alpha < c_n < \beta'' n^\alpha. \tag{36}$$

Let $G_{\mathbf{c}}^{(M)}(x)$ be the tail of the series $G_{\mathbf{c}}$:

$$G_{\mathbf{c}}^{(M)}(x) = \sum_{n=M}^\infty x^{c_n}.$$

Now, observe,

$$\int_M^\infty x^{(1+\varepsilon)\beta(y+1)^\alpha} dy \leq \sum_{n=M}^\infty x^{\beta' n^\alpha} \leq G_{\mathbf{c}}^{(M)}(x) \leq \sum_{n=M}^\infty x^{\beta'' n^\alpha} \leq \int_M^\infty x^{(1-\varepsilon)\beta y^\alpha} dy. \tag{37}$$

The integrals on the far ends of this inequality can be computed in terms of the upper incomplete gamma function $\Gamma(a, b) = \int_b^\infty z^{a-1} e^{-z} dz$, since for the right most integral,

$$\int_M^\infty x^{\beta' y^\alpha} dy = \int_M^\infty e^{\beta' y^\alpha \log(x)} dy \tag{38}$$

$$= \frac{1}{\alpha} \frac{1}{(-\beta' \log(x))^{1/\alpha}} \int_{-\beta' \log(x) M^\alpha}^\infty u^{1/\alpha-1} e^{-u} du \tag{39}$$

$$= \frac{(-\log(x))^{-1/\alpha}}{\alpha \beta'^{1/\alpha}} \Gamma\left(\frac{1}{\alpha}, -\beta' \log(x) M^\alpha\right). \tag{40}$$

And similarly, for the left most side of (37),

$$\int_M^\infty x^{\beta''(y+1)^\alpha} dy = \int_{M-1}^\infty x^{\beta'' y^\alpha} dy = \frac{(-\log(x))^{-1/\alpha}}{\alpha \beta''^{1/\alpha}} \Gamma\left(\frac{1}{\alpha}, -\beta'' \log(x)(M-1)^\alpha\right). \tag{41}$$

Using the following asymptotic expansion of $\Gamma(a, b)$ as $b \rightarrow 0$,

$$\Gamma(a, b) \sim \Gamma(a) - \frac{b^a}{a}, \quad a > 0, \text{ as } b \rightarrow 0 \tag{42}$$

(see [16], formula 45:9:7), and $\log(x) \sim x - 1$ as $x \rightarrow 1$, (40) and (41) are asymptotic to

$$\frac{(1-x)^{-1/\alpha}}{\alpha \beta'^{1/\alpha}} \Gamma\left(\frac{1}{\alpha}\right) \quad \text{and} \quad \frac{(1-x)^{-1/\alpha}}{\alpha \beta''^{1/\alpha}} \Gamma\left(\frac{1}{\alpha}\right), \tag{43}$$

respectively. Thus, (37) implies

$$\frac{1}{\alpha \beta''^{1/\alpha}} \Gamma\left(\frac{1}{\alpha}\right) (1-x)^{-1/\alpha} \lesssim G_{\mathbf{e}}^{(M)}(x) \lesssim \frac{1}{\alpha \beta'^{1/\alpha}} \Gamma\left(\frac{1}{\alpha}\right) (1-x)^{-1/\alpha} \tag{44}$$

as $x \rightarrow 1$. Since everything is tending to ∞ , (44) also holds with $G_{\mathbf{e}}^{(M)}(x)$ replaced with $G_{\mathbf{e}}(x)$. And finally, since $\varepsilon > 0$ was arbitrary, we can let $\varepsilon \rightarrow 0$, making $\beta', \beta'' \rightarrow \beta$, which implies (34). □

We are now ready to give the Lévy Khintchine representation of the Rosenblatt distribution.

Theorem 4.2. *Let Z_D have a Rosenblatt distribution with $0 < D < 1/2$ and let*

$$\lambda(D)^{-1} = (\lambda_1(D)^{-1}, \lambda_2(D)^{-1}, \dots)$$

be the sequence of the inverses of the eigenvalues associated to the integral operator $\sigma(D)\mathcal{K}_D$ defined in (13). Then the characteristic function of Z_D can be written as

$$\phi(\theta) = \mathbb{E} e^{i\theta Z_D} = \exp\left(\int_0^\infty (e^{i\theta u} - 1 - i\theta u) \nu_D(u) du\right), \tag{45}$$

where v_D is supported on $(0, \infty)$ and is given by

$$v_D(u) = \frac{1}{2u} G_{\lambda(D)-1}(e^{-u/2}) = \frac{1}{2u} \sum_{n=1}^{\infty} \exp\left(-\frac{u}{2\lambda_n}\right), \quad u > 0. \tag{46}$$

Moreover, v_D has the following asymptotic forms as $u \rightarrow 0^+$ and $u \rightarrow \infty$,

$$v_D(u) \sim \frac{2^{D/(1-D)} C(D)^{1/(1-D)}}{(1-D)} \Gamma\left(\frac{1}{1-D}\right) u^{(D-2)/(1-D)}, \quad u \rightarrow 0, \tag{47}$$

$$v_D(u) \sim \frac{e^{-u/(2\lambda_1)}}{2u}, \quad u \rightarrow \infty, \tag{48}$$

where $C(D)$ is defined in (30).

Proof. Let

$$Z_D^{(M)} = \sum_{n=1}^M \lambda_n (\varepsilon_i^2 - 1).$$

We have $Z_D^{(M)} \xrightarrow{d} Z_D$, and since $Z_D^{(M)}$ is a sum of shifted i.i.d. chi-squared distributions, we can use the Lévy Khintchine representation of a chi-square ([4], Example 1.3.22), which is a gamma distribution with shape parameter $1/2$ and scale parameter 2 :

$$\mathbb{E}e^{i\theta Z_D^{(M)}} = \prod_{n=1}^M e^{i\theta \lambda_n (\varepsilon_i^2 - 1)} = \prod_{n=1}^M \exp\left(-i\theta \lambda_n + \int_0^{\infty} (e^{i\theta u} - 1) \left[\frac{e^{-u/(2\lambda_i)}}{2u}\right] du\right). \tag{49}$$

Using $(1/2) \int_0^{\infty} e^{-u/(2\lambda)} du = \lambda$, (49) can be rewritten as

$$\begin{aligned} & \prod_{n=1}^M \exp\left(\int_0^{\infty} (e^{i\theta u} - 1 - i\theta u) \left[\frac{e^{-u/(2\lambda_i)}}{2u}\right] du\right) \\ &= \exp\left(\int_0^{\infty} (e^{i\theta u} - 1 - i\theta u) \left[\frac{1}{2u} G_{\lambda(D)-1}^{(M)}(e^{-u/2})\right] du\right), \end{aligned}$$

where

$$G_{\lambda(D)-1}^{(M)}(x) = \sum_{n=1}^M x^{\lambda_i^{-1}}.$$

Now, we let $M \rightarrow \infty$. In order to justify passing the limit through the integral, notice that

$$\left| (e^{i\theta u} - 1 - i\theta u) \left[\frac{1}{2u} G_{\lambda(D)-1}^{(M)}(e^{-u/2})\right] \right| \leq \frac{\theta^2}{4} u G_{\lambda(D)-1}^{(M)}(e^{-u/2}) \leq \frac{\theta^2}{4} u G_{\lambda(D)-1}(e^{-u/2}), \tag{50}$$

where we have used the identity $|e^{iz} - 1 - z| \leq \frac{z^2}{2}$ for $z \in \mathbb{R}$. Notice that (50) is continuous for $0 < u < \infty$, and by (29) together with Lemma 4.1 using $\alpha = 1 - D$ and $\beta = C(D)^{-1}$, we have

$$uG_{\lambda(D)^{-1}}(e^{-u/2}) \sim ue^{-u/(2\lambda_1)} \quad \text{as } u \rightarrow \infty \tag{51}$$

and,

$$uG_{\lambda(D)^{-1}}(e^{-u/2}) \sim C'u(1 - e^{-u/2})^{-1/(1-D)} \sim C''u^{D/(1-D)} \quad \text{as } u \rightarrow 0 \tag{52}$$

for some constants C' and C'' . Since $0 < \frac{D}{1-D} < 1$, (51) and (52) imply that (50) is integrable on $(0, \infty)$, and hence the dominated convergence theorem applies and

$$\mathbb{E}e^{i\theta Z_D^{(M)}} \rightarrow \mathbb{E}e^{i\theta Z_D} = \exp\left(\int_0^\infty (e^{i\theta u} - 1 - i\theta u) \left[\frac{1}{2u}G_{\lambda(D)^{-1}}(e^{-u/2})\right] du\right)$$

which verifies (45).

The final assertions (47) and (48) also follow from (29) and Lemma 4.1 with $\alpha = 1 - D$ and $\beta = C(D)^{-1}$, since these imply

$$\begin{aligned} \frac{1}{2u}G_{\lambda(D)^{-1}}(e^{-u/2}) &\sim \frac{1}{2u} \frac{C(D)^{-1/(1-D)}}{(1-D)} \Gamma\left(\frac{1}{1-D}\right) \left(\frac{u}{2}\right)^{-1/(1-D)} \\ &= \frac{2^{D/(1-D)}C(D)^{1/(1-D)}}{(1-D)} \Gamma\left(\frac{1}{1-D}\right) u^{(D-2)/(1-D)}, \quad u \rightarrow 0 \end{aligned} \tag{53}$$

and

$$\frac{1}{2u}G_{\lambda(D)^{-1}}(e^{-u/2}) \sim \frac{1}{2u}e^{-u/(2\lambda_1)}, \quad u \rightarrow \infty. \tag{54}$$

This concludes the proof. □

Remark. Notice that for any $0 < D < 1/2$, the Lévy measure is normalized in the sense that

$$\int_0^\infty u^2 v_D(u) du = \mathbb{E}Z_D^2 = 1. \tag{55}$$

As $D \rightarrow 1/2$, $\sigma(D) \rightarrow 0$ by (3) $\Rightarrow C(D) \rightarrow 0$ by (30) $\Rightarrow \lambda_n(D) \rightarrow 0$ by (29) $\Rightarrow v_D(u) \rightarrow 0$ by (46) and hence $u^2 v_D(u) \rightarrow 0$ for all u positive, but in light of (55), the function $u^2 v_D(u)$ approaches a dirac mass at $u = 0$. Thus, as $D \rightarrow 1/2$, since $(e^{iu\theta} - 1 - iu\theta) \rightarrow -1/2\theta^2$, one gets

$$\phi(\theta) = \exp\left(\int_0^\infty \left(\frac{e^{iu\theta} - 1 - i\theta u}{u^2}\right) u^2 v_D(u) du\right) \rightarrow \exp\left(-\frac{1}{2}\theta^2\right),$$

which verifies that $Z_D \xrightarrow{d} N(0, 1)$.

Understanding the Lévy measure of a distribution has some immediate implications pertaining to its probability density function and distribution function. We will state three such results as

corollaries. The first is not surprising given that Z_D is an infinite sum of chi-squared distributions, however it is now easy to prove given what we know about the Lévy measure:

Corollary 4.3. *For $0 < D < 1/2$, the probability density function of Z_D is infinitely differentiable with all derivatives tending to 0 as $|x| \rightarrow \infty$.*

Proof. Using Proposition 23.8 in [21], this follows as long as there exists an $\alpha \in (0, 2)$ such that

$$\liminf_{r \rightarrow 0} \frac{\int_{[-r,r]} u^2 v_D(u) du}{r^{2-\alpha}} > 0. \tag{56}$$

And indeed, (47) implies that for some constant $\tilde{C}(D)$,

$$\int_{[-r,r]} u^2 v_D(u) du \sim \tilde{C}(D) \int_0^r u^{-D/(1-D)} du = \tilde{C}(D) \left(\frac{1-D}{1-2D} \right) r^{2-1/(1-D)}. \tag{57}$$

Thus, choosing $\alpha = 1/(1 - D)$ verifies the result. □

The second corollary gives a simple bound on the left-hand tail of the CDF of Z_D . The proof is similar to Proposition 9.5(ii) in [28].

Corollary 4.4 (Left tail). *Let Z_D denote the Rosenblatt distribution. Then,*

$$P[Z_D < -x] \leq \exp(-\frac{1}{2}x^2), \quad x > 0. \tag{58}$$

Proof. For any $M \geq 1$, recall the random variable $Z_D^{(M)}$ defined in the proof of Theorem 4.2. Applying Markov’s inequality we see for any $x > 0$ and $s > 0$,

$$\begin{aligned} P(Z_D^{(M)} \leq -x) &\leq e^{-sx} \mathbb{E} e^{-sZ_D^{(M)}} \\ &= e^{-sx} \exp\left(\int_0^\infty (e^{-su} - 1 + su) dv^{(M)}(u) du\right), \end{aligned} \tag{59}$$

where the last equality follows since $Z_D^{(M)}$ is a finite sum of weighted chi-square distributions, and hence has a moment generating function defined for $s < s_0$ for some $s_0 > 0$. Using the inequality $e^{-su} - 1 + su \leq \frac{1}{2}s^2u^2$, and since $v^{(M)}(u)$ increases in M , (59) becomes

$$\begin{aligned} e^{-sx} \exp\left(\int_0^\infty (e^{-su} - 1 + su) dv^{(M)}(u) du\right) &\leq e^{-sx} \exp\left(\frac{s^2}{2} \int_0^\infty u^2 v^{(M)}(u) du\right) \\ &\leq e^{-sx} \exp\left(\frac{s^2}{2} \int_0^\infty u^2 v(u) du\right) \\ &\leq \exp\left(-sx + \frac{1}{2}s^2\right) \end{aligned} \tag{60}$$

since $\text{Var } Z_D = \int_0^\infty u^2 v(u) \, du = 1$. The minimum of (60) over s is attained at $s = x$. Thus,

$$P(Z_D^{(M)} \leq -x) \leq \exp\left(-\frac{1}{2}x^2\right) \tag{61}$$

which finishes the proof. \square

We obtain a result similar to Theorem 2 in [3] involving the rate of decay of the distribution function.

Corollary 4.5 (Right tail). *Let Z_D denote the Rosenblatt distribution. Then for $\alpha > 0$,*

$$\lim_{u \rightarrow \infty} \frac{P[Z_D > u + \alpha]}{P[Z_D > u]} = e^{-\alpha/(2\lambda_1)},$$

where λ_1 is the largest eigenvalue of $\sigma(D)\mathcal{K}_D$ defined in (13).

Proof. For $u \geq 1$, define \bar{A}_{v_D} as

$$\bar{A}_{v_D}(u) = \frac{\int_u^\infty v_D(v) \, dv}{\int_1^\infty v_D(v) \, dv}.$$

By Theorem 1 in [8], it suffices to show

$$\lim_{u \rightarrow \infty} \frac{\bar{A}_{v_D}(u + \alpha)}{\bar{A}_{v_D}(u)} = e^{-\alpha/(2\lambda_1)}, \tag{62}$$

and

$$\int_0^\infty e^{u/(2\lambda_1)} \bar{A}_{v_D}(du) = \infty. \tag{63}$$

Both follow from (48) in Theorem 4.2. Indeed, as $u \rightarrow \infty$,

$$\int_u^\infty v_D(v) \, dv \sim \int_u^\infty \frac{1}{2v} e^{-v/(2\lambda_1)} \, dv \sim \frac{\lambda_1}{u} e^{-u/(2\lambda_1)} \tag{64}$$

which follows because

$$-\frac{d}{dx} \frac{\lambda_1}{u} e^{-u/(2\lambda_1)} = \frac{1}{2u} e^{-u/(2\lambda_1)} + \frac{\lambda_1}{u^2} e^{-u/(2\lambda_1)} \sim \frac{1}{2u} e^{-u/(2\lambda_1)},$$

and thus integrating both sides implies (64). From this, (62) follows. Also,

$$\begin{aligned} \int_0^\infty e^{u/(2\lambda_1)} \bar{A}_{v_D}(du) &= \frac{1}{\int_1^\infty v_D(v) \, dv} \int_0^\infty e^{u/(2\lambda_1)} \left(\frac{1}{2u} \sum_{n=1}^\infty e^{-u/(2\lambda_n)} \right) du \\ &= \frac{1}{\int_1^\infty v_D(v) \, dv} \int_0^\infty \frac{1}{2u} \left(1 + \sum_{n=1}^\infty e^{-u/(2\lambda_n)+u/(2\lambda_1)} \right) du = \infty, \end{aligned}$$

since the integrand on the right is asymptotic to $1/(2u)$. This verifies (63). □

5. Approximating the distribution of the tail of the series

The representation of the Rosenblatt distribution Z_D as an infinite sum of shifted chi-squared distributions (8) has proven to be quite useful for obtaining theoretical results about the distribution. In this section, we aim to take advantage of this representation to compute the CDF and PDF of the Rosenblatt distribution.

For $M \geq 1$, define X_M and Y_M as

$$Z_D = \sum_{n=1}^{M-1} \lambda_n (\varepsilon_n^2 - 1) + \sum_{n=M}^{\infty} \lambda_n (\varepsilon_n^2 - 1) := X_M + Y_M. \tag{65}$$

Notice that Y_M has mean 0 and variance

$$\sigma_M^2 := \mathbb{E}Y_M^2 = 2 \sum_{n=M}^{\infty} \lambda_n^2. \tag{66}$$

Notice that from Theorem 3.2, as $M \rightarrow \infty$,

$$\sigma_M^2 \sim 2C(D)^2 \sum_{n=M}^{\infty} n^{2D-2} \sim 2C(D)^2 (1 - 2D)^{-1} M^{2D-1}, \tag{67}$$

where we approximated the sum $\sum_{n=M}^{\infty} \lambda_n^2$ with an integral. This suggests that X_M alone is not a very good approximation of Z_D since this variance tends to 0 slowly with M , especially when D is close to $1/2$. As an alternative, we will see below that Y_M is approximated by a normal distribution as $M \rightarrow \infty$. By taking advantage of this property, we can obtain accurate approximations of the distribution of Z_D .

In [33], random variables expressed as infinite sums of gamma and hence chi-squared distributions were studied. It was shown in particular that the asymptotic form of the λ_n 's in the expansion implies that the distribution of the tail Y_M is approximately normal as $M \rightarrow \infty$.

Let $\kappa_{k,M}$ be the normalized cumulants of Y_M . These are given by

$$\kappa_{k,M} = 2^{k-1} (k - 1)! \sigma_M^{-k} \sum_{n=M}^{\infty} \lambda_n^k. \tag{68}$$

Notice that from the asymptotics (29) of λ_n ,

$$\begin{aligned} \kappa_{k,M} &\sim 2^{k-1} (k - 1)! \cdot 2^{-k/2} (1 - 2D)^{k/2} M^{k/2-kD} \cdot \int_M^{\infty} n^{k(D-1)} \, dn \\ &= \frac{(k - 1)! (2 - 4D)^{k/2}}{2 \cdot k - kD - 1} M^{1-k/2}, \end{aligned}$$

which is indeed equal to 1 when $k = 2$.

The convergence to normality of Y_M is implied by the following Berry–Essen bound proved in [33].

Theorem 5.1. *Let Y_M and σ_M be defined as in (65) and (66). Then^{1,2},*

$$\sup_{x \in \mathbb{R}} |\mathbb{P}[\sigma_M^{-1} Y_M \leq x] - \Phi(x)| \leq 0.7056 \kappa_{3,M} = O(M^{-1/2}). \tag{69}$$

While this result gives an idea of the distribution of Y_M as $M \rightarrow \infty$, the bound on the right-hand side of (69) may not be satisfactory for small M . In order to better approximate the CDF of Y_M , we will use an Edgeworth expansion which is also considered in [33]. These expansions use the higher cumulants to better approximate the CDF of Y_M and give a faster convergence than what we see in (69).

Recall the *Hermite polynomials*, which are defined as

$$H_k(x) = (-1)^k e^{x^2/2} \frac{d^k}{dx^k} e^{-x^2/2}, \quad k \geq 0.$$

The first few are $H_2(x) = x^2 - 1$, $H_3(x) = x^3 - 3x$, $H_4(x) = x^4 - 6x^2 + 3$ and $H_5(x) = x^5 - 10x^3 + 15x$. A simple induction gives

$$\frac{d}{dx} H_k(x) \phi(x) = -H_{k+1}(x) \phi(x), \tag{70}$$

where ϕ is the standard normal PDF. The following theorem is also proved in [33].

Theorem 5.2. *The CDF of the tail Y_M satisfies*

$$\begin{aligned} \mathbb{P}[\sigma_M^{-1} Y_M \leq x] &= \Phi(x) - \phi(x) \left\{ \sum_{\eta(N)} \left[\prod_{m=1}^N \frac{1}{k_m!} \left(\frac{\kappa_{m,M}}{m!} \right)^{k_m} \right] H_{\zeta(k_3, \dots, k_N)}(x) \right\} \\ &+ O(M^{-(N-1)/2}), \end{aligned} \tag{71}$$

where $\kappa_{k,M}$ is defined in (68), $\eta(N)$ denotes all k_3, k_4, \dots, k_N such that

$$1 \leq k_3 + 2k_4 + \dots + (N - 2)k_N \leq N - 2 \tag{72}$$

and

$$\zeta(k_3, \dots, k_N) = 3k_3 + 4k_4 + \dots + Nk_N - 1. \tag{73}$$

¹Multiplicative factors: they are denoted $\sigma(D)$ in (2) and σ_M^{-1} in (69). Using (9) and (66), compare also (10) with (68).

²The constant 0.7056 appearing in this inequality is the smallest known to date, see [22].

Notice by taking a derivative of (71) and using (70), the PDF of Y_M is approximated by

$$f_{\sigma_M^{-1}Y_M}(x) = \phi(x) \left[1 + \sum_{\eta(N)} \left[\prod_{m=1}^N \frac{1}{k_m!} \left(\frac{\kappa_{m,M}}{m!} \right)^{k_m} \right] H_{\zeta(k_3, \dots, k_N)+1}(x) \right] + O(M^{-(N-1)/2}). \tag{74}$$

These expansions allows us to compute the CDF and PDF of Y_M (and hence of Z_D) to high accuracy. We provide the details and results of this method in Section 7.

6. Obtaining the eigenvalues, cumulants and moments numerically

There exists an extensive literature regarding the problem of approximating eigenvalues of integral operators like \mathcal{K}_D , see for instance [2,6,23,25], or [10]. Many, if not all, of these methods boil down to approximating \mathcal{K}_D with a finite-dimensional linear operator, a technique often referred to as the ‘‘Nyström method’’ (see [26] and [5]). We shall approximate \mathcal{K}_D by a $(J + 1) \times (J + 1)$ matrix \mathbf{K}_D , defined as follows.

Fix $J > 0$ and choose nodes $0 = x_0 < x_1 < \dots < x_J = 1$. For such f , let $\mathbf{f} = (f(x_0), \dots, f(x_J))$ and define

$$f_J(x) = f(x_{j-1}) \frac{x - x_j}{x_{j-1} - x_j} + f(x_j) \frac{x - x_{j-1}}{x_j - x_{j-1}} \quad \text{for } x \in [x_{j-1}, x_j]. \tag{75}$$

Then, for any $i = 0, 1, \dots, J$,

$$(\mathcal{K}_D f_J)(x_i) = \sum_{j=1}^J \int_{x_{j-1}}^{x_j} \left(f(x_{j-1}) \frac{u - x_j}{x_{j-1} - x_j} + f(x_j) \frac{u - x_{j-1}}{x_j - x_{j-1}} \right) |x_i - u|^{-D} du = (\mathbf{K}_D \mathbf{f})_i,$$

where J indicates the level of approximation. For more details on the matrix \mathbf{K}_D , see Sections A and B of the supplemental article [35]. By taking the eigenvalues of the matrix \mathbf{K}_D , we can approximate the M largest eigenvalues of \mathcal{K}_D where $M \ll J$.

In order to test that the approximate eigenvalues we are obtaining are accurate, we have three methods:

1. Check for numerical convergence, that is, for J large, does increasing J lead to a negligible change in the approximations of λ_n . Thus, to approximate λ_n , we increased J until sufficient convergence was met. By ‘‘sufficient convergence,’’ we mean that J is increased by multiples of 200 until the values of the λ_n ’s no longer changed in the 4 significant decimal digit. Table 1 gives the results of this method applied to the first 10 values of λ_n ’s. For more values, see the supplemental article [35].
2. Compare λ_n with the asymptotic formula given by Theorem 3.2 for n large. The first 30 λ_n are approximated and plotted on a log scale in the supplemental article [35] and compared to the asymptotic formula in Theorem 3.2. For large n , our approximations appear to be in agreement with the asymptotic formula as $n \rightarrow \infty$. In fact, the asymptotic formula for

Table 1. First 10 eigenvalues of \mathcal{K}_D for various D . Note that for $D = 0$, we have $\lambda_1 = 1$ and $\lambda_n = 0$ for $n \geq 2$. As $D \rightarrow 1/2$, $\lambda_n \rightarrow 0$

n	$D = 0.1$	$D = 0.2$	$D = 0.3$	$D = 0.4$
1	0.70200	0.68130	0.63050	0.51250
2	0.05648	0.11260	0.16040	0.17840
3	0.03477	0.07341	0.11070	0.13010
4	0.02453	0.05411	0.08512	0.10420
5	0.01947	0.04409	0.07119	0.08948
6	0.01600	0.03710	0.06129	0.07879
7	0.01374	0.03240	0.05446	0.07121
8	0.01198	0.02872	0.04903	0.06512
9	0.01069	0.02596	0.04489	0.06037
10	0.009629	0.02366	0.04140	0.05635

the λ_n is a good approximation even for moderate values of n (about $n \geq 20$ for $D = 0.1, 0.2, 0.3, 0.4$).

3. The cumulants κ_k of Z_D are given in terms of the eigenvalues as

$$\kappa_k = 2^{k-1}(k-1)! \sum_{n=1}^{\infty} \lambda_n^k = 2^{k-1}(k-1)! \left[\sum_{n=1}^{M-1} \lambda_n^k + \sum_{n=M}^{\infty} \lambda_n^k \right]. \tag{76}$$

Since $\kappa_2 = 1$, and κ_3, κ_4 can be computed numerically exactly, we can compute the absolute error made by approximating the cumulants $\kappa_2, \kappa_3, \kappa_4$ by using (76). In order to compute the sums on the right-hand side of (76), we used the first $M = 50$ values of λ_n using the matrix \mathbf{K}_D , and for $n > M$, the asymptotic formula given in Theorem 3.2. The results of this test are given in a table in the supplemental article [35]. Because we are using the asymptotic formula for n large, we expect some error, but as it turns out, the absolute errors are very small compared to the size of $\kappa_k, k = 2, 3, 4$.

One can obtain also higher cumulants and moments of Z_D . The cumulants of Z_D are expressed in terms of the functions $G_{k,D}$ by 10 and Proposition 2.1. These functions are defined recursively in (14). By using \mathbf{K}_D to approximate the functions $G_{k,D}$, as explained in Section C of the supplemental article [35], one obtains approximations for the cumulants. One derives the corresponding moments by applying (11). We have tabulated the first 8 moments and cumulants of the Rosenblatt distribution for various D in the supplemental article [35].

7. Obtaining the CDF numerically

We now present a technique for computing the CDF and PDF of the Rosenblatt distribution. Computing the distribution of $Z_D = X_M + Y_M$ (see (65)) requires three steps:

1. Approximate the eigenvalues λ_n for $n = 1, 2, \dots, M$ for some $M > 0$.

2. Compute separately the CDF of X_M and the PDF of Y_M in the decomposition (65).

For X_M , methods exist already for accurately computing the CDF of a finite sum of chi-squared distributions, see for instance methods based on Laplace transform inversion, [9,34], or Fourier transform inversion, [1].

For Y_M , use an Edgeworth expansion of order $N \geq 2$ which is found in Theorem 5.2. To compute σ_M in (66), take $\sigma_M = (1 - 2 \sum_{n=1}^{M-1} \lambda_n^2)^{1/2}$. The Edgeworth expansion in Theorem 5.2 also involves the cumulants $\kappa_{k,M}$, $k \geq 1$ of Y_M . If M is sufficiently large, $\kappa_{k,M}$ can be approximated using the asymptotic formula given in Theorem 3.2:

$$\begin{aligned} \kappa_{k,M} &= \sigma_M^{-k} \left(2^{k-1} (k-1)! \sum_{n=M}^{\infty} \lambda_n^k \right) \approx \sigma_M^{-k} \left(2^{k-1} (k-1)! \sum_{n=M}^{\infty} C(D)^k n^{k(D-1)} \right) \\ &= 2^{k-1} (k-1)! \sigma_M^{-k} C(D)^k \zeta(k(1-D), M), \end{aligned} \tag{77}$$

where $\zeta(s, M) = \sum_{n=M}^{\infty} n^{-s}$ denotes the *Hurwitz Zeta function* ([16], Chapter 64). Using (77) introduces a small amount of error since we are approximating λ_n for n large, however we will see below this is negligible for M large.

3. Finally, the CDF of Z_D is given by a convolution

$$F_{Z_D}(x) = \int_{-\infty}^{\infty} F_{X_M}(x-y) f_{Y_M}(y) dy,$$

where F_{X_M} is the CDF of X_M and f_{Y_M} is the PDF of Y_M . We compute this convolution in MATLAB using standard numerical integration techniques.

The choice of M (number of chi-squared distributions) and N (order of Edgeworth expansion) was determined by increasing both until the value of the CDF changed by less than 10^{-5} . In Table 2, we show the effects of increasing M and N for the case of $D = 0.3$ and $x = 0$. Observe that for fixed M , the values of the CDF converge rapidly as N increases. Nevertheless, if M is small ($M \leq 10$) the approximation will have a slight error since we are approximating the $\kappa_{k,M}$ using (77). Moreover, by choosing $N = 2$ and $M \geq 10$, the only noticeable improvement by increasing N happens in the fifth decimal place and beyond.

We have tabulated in Table 3 quantiles of Z_D for various D . These are useful for obtaining confidence intervals. To obtain these values, we solved the equation $F_{Z_D}(x) = q$ for various

Table 2. Approximation of $P[Z_{0.3} \leq 0]$ for various M and N values. This shows that one gets an excellent approximation with relatively small values of M and N

N	$M = 10$	$M = 20$	$M = 30$	$M = 50$
2	0.616883	0.616909	0.616909	0.616907
3	0.616817	0.616878	0.616890	0.616896
4	0.616885	0.616898	0.616899	0.616899
5	0.616895	0.616900	0.616900	0.616900
6	0.616895	0.616900	0.616900	0.616900

Table 3. Various quantiles of the Rosenblatt distribution for selected values of D

Quantile	$D = 0.1$	$D = 0.2$	$D = 0.3$	$D = 0.4$	$D = 0.45$
0.01	-0.8472	-1.0567	-1.3546	-1.7838	-2.0603
0.025	-0.8142	-0.9827	-1.7669	-1.5536	-1.7639
0.05	-0.7808	-0.9122	-1.0958	-1.3493	-1.5051
0.10	-0.7340	-0.8201	-0.9419	-1.1053	-1.6462
0.25	-0.6200	-0.6277	-0.6479	-0.6713	-0.6789
0.50	-0.3622	-0.3055	-0.2329	-0.1332	-0.0673
0.75	0.2370	0.2781	0.3666	0.5059	0.5955
0.90	1.2047	1.2031	1.2110	1.2417	1.2673
0.95	2.0015	1.9726	1.9114	1.8022	1.7262
0.975	2.8312	2.7759	2.6483	2.3858	2.1774
0.99	3.9618	3.8718	3.6579	3.1909	2.7892

quantiles q in MATLAB. To compute the CDF $F_{Z_D}(x)$, we fixed $N = 6$, and increased M in increments of 10 until the approximation of the CDF changed by less than 10^{-5} . For the CDF values, see Table 9 in the supplemental article [35]. We have also plotted the PDF and CDF in Figure 1.

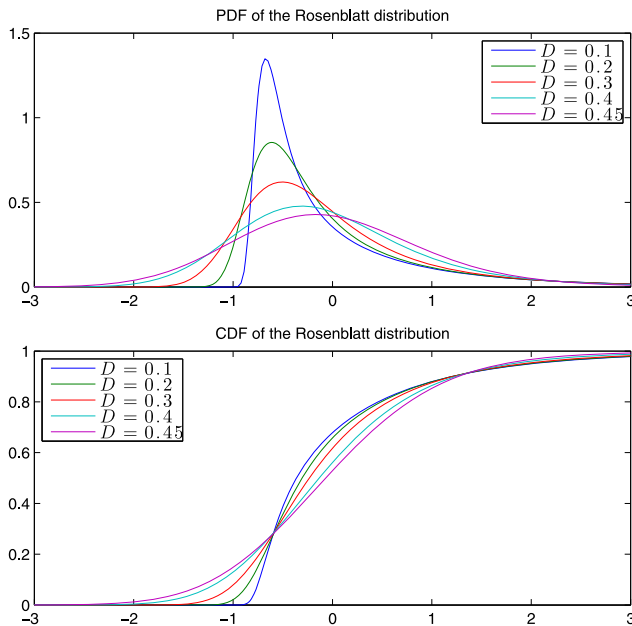


Figure 1. Plots of the PDF and CDF of Z_D for various D . The CDF with the steepest slope and the PDF with the highest mode correspond to $D = 0.1$.

Acknowledgement

This work was partially supported by the NSF Grants DMS-07-06786 and DMS-10-07616 at Boston University.

Supplementary Material

Supplemental article: Supplement to *Properties and numerical evaluation of the Rosenblatt distribution* (DOI: [10.3150/12-BEJ421SUPP](https://doi.org/10.3150/12-BEJ421SUPP); .pdf). The supplement [35] to this article details the approximation of the integral operator \mathcal{K}_D and the computation of the cumulants, moments and CDF of Z_D . It also contains an extensive table of the CDF of Z_D and a guide to the software.

References

- [1] Abate, J. and Whitt, W. (1992). The Fourier-series method for inverting transforms of probability distributions. *Queueing Systems Theory Appl.* **10** 5–87. [MR1149995](#)
- [2] Ahues, M., Largillier, A. and Limaye, B.V. (2001). *Spectral Computations for Bounded Operators. Applied Mathematics (Boca Raton)* **18**. Boca Raton, FL: Chapman & Hall/CRC. [MR1886113](#)
- [3] Albin, J.M.P. (1998). A note on Rosenblatt distributions. *Statist. Probab. Lett.* **40** 83–91. [MR1650532](#)
- [4] Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus. Cambridge Studies in Advanced Mathematics* **93**. Cambridge: Cambridge Univ. Press. [MR2072890](#)
- [5] Atkinson, K. and Han, W. (2009). *Theoretical Numerical Analysis: A Functional Analysis Framework*, 3rd ed. *Texts in Applied Mathematics* **39**. Dordrecht: Springer. [MR2511061](#)
- [6] Atkinson, K.E. (1967). The numerical solutions of the eigenvalue problem for compact integral operators. *Trans. Amer. Math. Soc.* **129** 458–465. [MR0220105](#)
- [7] Birman, M.Š. and Solomjak, M.Z. (1977). Estimates for the singular numbers of integral operators. *Uspehi Mat. Nauk* **32** 17–84, 271. [MR0438186](#)
- [8] Braverman, M. (2005). On a class of Lévy processes. *Statist. Probab. Lett.* **75** 179–189. [MR2210548](#)
- [9] Castaño-Martínez, A. and López-Blázquez, F. (2005). Distribution of a sum of weighted noncentral chi-square variables. *Test* **14** 397–415. [MR2211387](#)
- [10] Chen, Z., Nelakanti, G., Xu, Y. and Zhang, Y. (2009). A fast collocation method for eigen-problems of weakly singular integral operators. *J. Sci. Comput.* **41** 256–272. [MR2550370](#)
- [11] Dobrushin, R.L. and Major, P. (1979). Non-central limit theorems for nonlinear functionals of Gaussian fields. *Z. Wahrsch. Verw. Gebiete* **50** 27–52. [MR0550122](#)
- [12] Dostanić, M.R. (1998). Spectral properties of the operator of Riesz potential type. *Proc. Amer. Math. Soc.* **126** 2291–2297. [MR1451794](#)
- [13] Gradshteyn, I.S. and Ryzhik, I.M. (2007). *Table of Integrals, Series, and Products*, 7th ed. Amsterdam: Elsevier/Academic Press. Translated from the Russian, Translation edited and with a preface by Alan Jeffrey and Daniel Zwillinger, With one CD-ROM (Windows, Macintosh and UNIX). [MR2360010](#)
- [14] Kac, M. (1955–1956). Distribution of eigenvalues of certain integral operators. *Michig. Math. J.* **3** 141–148. [MR0085650](#)
- [15] Major, P. (1981). *Multiple Wiener–Itô Integrals: With Applications to Limit Theorems. Lecture Notes in Math.* **849**. Berlin: Springer. [MR0611334](#)
- [16] Oldham, K., Myland, J. and Spanier, J. (2009). *An Atlas of Functions*, 2nd ed. New York: Springer. With Equator, the atlas function calculator, With 1 CD-ROM (Windows). [MR2466333](#)

- [17] Peccati, G. and Taqqu, M.S. (2011). *Wiener Chaos: Moments, Cumulants and Diagrams: A Survey with Computer Implementation. Bocconi & Springer Series 1*. Milan: Springer. Supplementary material available online. [MR2791919](#)
- [18] Pitman, J. (2006). *Combinatorial Stochastic Processes. Lecture Notes in Math. 1875*. Berlin: Springer. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard. [MR2245368](#)
- [19] Rosenblatt, M. (1961). Independence and dependence. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. II* 431–443. Berkeley, CA: Univ. California Press. [MR0133863](#)
- [20] Rosenblatt, M. (1963). Some results on the asymptotic behavior of eigenvalues for a class of integral equations with translation kernels. *J. Math. Mech.* **12** 619–628. [MR0150551](#)
- [21] Sato, K.i. (1999). *Lévy Processes and Infinitely Divisible Distributions. Cambridge Studies in Advanced Mathematics 68*. Cambridge: Cambridge Univ. Press. Translated from the 1990 Japanese original, Revised by the author. [MR1739520](#)
- [22] Shevtsova, I.G. (2006). Sharpening the upper bound for the absolute constant in the Berry–Esseen inequality. *Teor. Veroyatn. Primen.* **51** 622–626. [MR2325552](#)
- [23] Sloan, I.H. (1976). Iterated Galerkin method for eigenvalue problems. *SIAM J. Numer. Anal.* **13** 753–760. [MR0428707](#)
- [24] Smith, P.J. (1995). A recursive formulation of the old problem of obtaining moments from cumulants and vice versa. *Amer. Statist.* **49** 217–218. [MR1347727](#)
- [25] Spence, A. (1975/76). On the convergence of the Nyström method for the integral equation eigenvalue problem. *Numer. Math.* **25** 57–66. [MR0395278](#)
- [26] Spence, A. (1977/78). Error bounds and estimates for eigenvalues of integral equations. *Numer. Math.* **29** 133–147. [MR0659480](#)
- [27] Stade, E. (2005). *Fourier Analysis. Pure and Applied Mathematics (New York)*. Hoboken, NJ: Wiley-Interscience. [MR2125315](#)
- [28] Steutel, F.W. and van Harn, K. (2004). *Infinite Divisibility of Probability Distributions on the Real Line. Monographs and Textbooks in Pure and Applied Mathematics 259*. New York: Dekker. [MR2011862](#)
- [29] Taqqu, M.S. (1974/75). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrsch. Verw. Gebiete* **31** 287–302. [MR0400329](#)
- [30] Taqqu, M.S. (1979). Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrsch. Verw. Gebiete* **50** 53–83. [MR0550123](#)
- [31] Taqqu, M.S. (2011). The Rosenblatt process. In *Selected Works of Murray Rosenblatt* (R. Davis, K.S. Lii and D. Politis, eds.). New York: Springer.
- [32] Tudor, C.A. (2008). Analysis of the Rosenblatt process. *ESAIM Probab. Stat.* **12** 230–257. [MR2374640](#)
- [33] Veillette, M.S. and Taqqu, M.S. (2010). Berry–Esseen and Edgeworth approximations for the tail of an infinite sum of weighted gamma random variables. *Stochastic Process. Appl.* To appear. Preprint available online at <http://arxiv.org/abs/1010.3948>.
- [34] Veillette, M.S. and Taqqu, M.S. (2011). A technique for computing the PDFs and CDFs of nonnegative infinitely divisible random variables. *J. Appl. Probab.* **48** 217–237. [MR2809897](#)
- [35] Veillette, M.S. and Taqqu, M.S. (2012). Supplement to “Properties and numerical evaluation of the Rosenblatt distribution.” DOI:10.3150/12-BEJ421SUPP.
- [36] Zemanian, A.H. (1987). *Distribution Theory and Transform Analysis: An Introduction to Generalized Functions, with Applications*, 2nd ed. New York: Dover. [MR0918977](#)