ON COMPUTING DISTRIBUTIONS OF PRODUCTS OF RANDOM VARIABLES VIA GAUSSIAN MULTIRESOLUTION ANALYSIS

GREGORY BEYLKIN, LUCAS MONZÓN AND IGNAS SATKAUSKAS

Abstract. We introduce a new approximate multiresolution analysis (MRA) using a single Gaussian as the scaling function, which we call Gaussian MRA (GMRA). As an initial application, we employ this new tool to accurately and efficiently compute the probability density function (PDF) of the product of independent random variables. In contrast with Monte-Carlo (MC) type methods (the only other universal approach known to address this problem), our method not only achieves accuracies beyond the reach of MC but also produces a PDF expressed as a Gaussian mixture, thus allowing for further efficient computations. We also show that an exact MRA corresponding to our GMRA can be constructed for a matching user-selected accuracy.

1. Introduction

While the probability density function (PDF) of the sum of two independent random variables is easily described as the convolution of their PDFs, the expression for the PDF of the product is significantly more complicated. Given two independent random variables X and Y with PDFs f and g, where $\int_{\mathbb{R}} f(x)dx = 1$ and $\int_{\mathbb{R}} g(y)dy = 1$, the PDF p of their product, $Z = XY$, can be succinctly written as

$$p(t) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x)\, g(y)\, \delta(xy - t)\, dxdy, \tag{1.1}$$

where δ is the delta function defined as

$$\int_{-\infty}^{\infty} \delta(t - t')\, u(t')\, dt' = u(t)$$

for a class of test functions u.When written in a more traditional form,

$$p(t) = \int_{\mathbb{R}} f(x)\, g(t/x)\, \frac{1}{|x|} dx,$$

the difficulty of computing p becomes apparent since the integral in (1.1) diverges at $t = 0$ as long as e.g. $g(0) \neq 0$ and $f(0) \neq 0$. Indeed, in the simplest case when independent random variables are selected from two normal distributions with zero means, the PDF p has a logarithmic singularity at $t = 0$ (see (3.1) below).

As far as we know, the only universal method currently available for computing the PDF of the product of two independent random variables relies on a Monte-Carlo approach, where one samples the individual PDFs, computes the products, and collects enough samples to achieve certain accuracy in the computation of p

Key words and phrases. multiresolution analysis, product of independent random variables, probability density function.

This research was partially supported by NSF grant DMS-1320919.

in (1.1). However, due to the slow convergence of such method (typically $1/\sqrt{N}$, where N is the number of samples) achieving high accuracy is not feasible.

In this paper we show how to accurately and efficiently compute the PDF p in (1.1) using an approximate multiresolution analysis (MRA), where the scaling function is a Gaussian. We demonstrate that, in contrast with the sampling method, our new algorithm is both fast and accurate. Moreover, while a sampling method only provides a histogram of p, we directly obtain the result in a functional form that can be used in further computations. Thus, we avoid the need for e.g. kernel density estimation of the result of sampling the product PDF (see e.g. [9, 16] and references therein). In particular, we do not have any obstacles dealing with heavy-tailed distributions (see an example with the Cauchy distribution in Section 4.2.3).

A representation of a function in Gaussian-based MRA (GMRA) can also be viewed as a multiresolution Gaussian mixture. Approximations by Gaussian mixtures have been used early on, see e.g. [17]. We prove that the integral (1.1) can be efficiently evaluated as a multiresolution Gaussian mixture and provide a tight estimate for the accuracy of the result.

Our approach to constructing a GMRA is based on the observation that, for any finite user-selected accuracy ϵ, the function $\phi(x) = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha x^2}$ satisfies the approximate two-scale relation

$$\left| \phi(x) - \sqrt{\frac{4\alpha}{3\pi}} \sum_{k \in \mathbb{Z}} e^{-\frac{\alpha}{3}k^2} \phi(2x-k) \right| \leq \epsilon \phi(x)$$

for an appropriately selected parameter $\alpha = \alpha(\epsilon)$. As a consequence, the scaled and shifted functions

$$\left\{ \phi_{j,k}(x) = 2^{j/2} \phi\left(2^j x - k\right) = 2^{j/2} \sqrt{\frac{\alpha}{\pi}} e^{-\alpha\left(2^j x - k\right)^2} \right\}_{j,k \in \mathbb{Z}}$$

form a Gaussian-based approximate MRA. We use this approximate GMRA to represent PDFs of random variables which we assume to be smooth functions except at a finite number of points where they may have integrable singularities. Under this assumption these PDFs can be represented by a reasonable number of basis functions. The approximate nature of the GMRA does not limit its applicability since any finite accuracy can be selected. For example, in our computations we roughly match the double precision arithmetic of a typical computer by selecting ϵ in the range $10^{-13} - 10^{-16}$.

The integral (1.1) can, in principle, be computed using the Mellin transform. It has been observed that, for non-negative random variables, the Mellin transform of p in (1.1) is equal to the product of the Mellin transforms of f and g (equivalently, p is the so-called Mellin convolution of f and g). In e.g. [18], this approach has been extended to random variables taking both positive and negative values. While analytically appealing, a numerical implementation of Mellin convolution is problematic and has not resulted in a reliable numerical method.

Instead of using the Mellin transform to represent the PDFs of non-negative random variables, we can use linear combinations of decaying exponentials with complex exponents constructed via algorithms developed in [4, 12]. Such representation of PDFs of non-negative random variables leads to a universal method for computing their product. We address such approach in the separate paper [6]. However, for random variables that can take both positive and negative values, it is

more natural to use multiresolution Gaussian mixtures as described in this paper. Moreover, a generalization to higher dimensions becomes problematic if one were to split the treatment of positive and negative values of such random variables.

Computing the PDF of the product of independent random variables using GMRA is just the first of several interesting applications of this approximate basis. Although the idea of using Gaussians to approximate arbitrary functions appeared earlier in [13, 14] –in these papers the coefficients of the shifted Gaussians are associated with function values– it turns out that much more accurate approximations can be achieved if the coefficients are computed differently, as described later in this paper. Since many classical operators of mathematical physics can be accurately approximated via a linear combination of Gaussians (see e.g. [11, 2, 5, 3]), using GMRA to apply these operators can be reduced to the analytic evaluation of integrals, and hence, suggests a new class of fast multiresolution algorithms. However, constructing such algorithms is not straightforward since integer translates of Gaussians are far from being orthogonal (after all, they are positive functions so that no cancellations are possible). Therefore, the use of GMRA differs from the usual MRA with orthogonal or nearly orthogonal scaling functions and we briefly discuss some of these issues in Appendix B.

We present our GMRA construction for functions of one variable and note that the generalization to two and three variables is fairly straightforward and we will describe it in appropriate contexts. Extending our approach to bases for multivariate functions in higher dimensions is more complicated and we plan to address it later.

Finally, we note that the quotient of two independent random variables can be evaluated in the same fashion as the product. Since the PDF of the sums (or differences) of two real valued random variables is obtained by the convolution of their PDFs, the representation provided by GMRA yields a numerical calculus of PDFs. In particular, the expectation

$$\mathrm{E}\left[u\left(Z\right)\right] = \int u\left(t\right) p_Z\left(t\right) dt,$$

where u is a given function and p_Z is the PDF of a random variable Z represented in GMRA, can be easily evaluated (see e.g. Remarks 3.5 and 3.6).

We start by proving a key approximation result and considering its consequences in Section 2. In Section 3 we prove a relative error estimate for the representation of the PDF of the product via GMRA when one of the distributions is Gaussian. We also present two approaches to approximating PDFs via a linear combination of Gaussians as well as demonstrate that the product of two random variables with a joint bivariate normal distribution can also be handled by our approach. We then turn to numerical examples in Section 4. Conclusions and further work is discussed in Section 5. In Appendix A we include a pseudo-code for adaptive integration used in Section 3. Finally, we present a brief discussion of certain features of GMRA in Appendix B.

2. Approximation of shifted Gaussians and approximate multiresolution analysis

We start by proving a key estimate (Theorem 2.1 below) that allows us to replace any shifted Gaussian by an explicit linear combination of Gaussians (with a larger, fixed exponent) shifted at integer values. This estimate yields an approximate

two-scale relation for Gaussians that makes possible to construct an approximate multiresolution basis with Gaussians as the scaling functions. Since the basis of shifted Gaussians is far from being orthogonal, one would expect difficulties in computing projections of functions on such basis. Remarkably, Theorem 2.1 and Theorem 2.5 allow us to compute such projections accurately for any Gaussian with an arbitrary exponent and shift. Since we are interested in applications where all functions and the results of computing integrals can be approximated via linear combinations of Gaussians, the following result is the key to using GMRA as a numerical tool.

Theorem 2.1. *If $0 < \beta < \alpha$ then, for $x, s \in \mathbb{R}$, we have*

(2.1)
$$\left| e^{-\beta(x-s)^2} - \frac{\alpha}{\sqrt{\pi\left(\alpha-\beta\right)}} \sum_{k\in\mathbb{Z}} e^{-\frac{\alpha\beta}{\alpha-\beta}(k-s)^2} e^{-\alpha(x-k)^2} \right| \quad \leq \quad \epsilon\left(\alpha,\beta\right) e^{-\beta(x-s)^2},$$

where

$$\epsilon\left(\alpha,\beta\right) = \vartheta_3\left(0, e^{-\frac{\pi^2}{\alpha}\left(1-\frac{\beta}{\alpha}\right)}\right) - 1$$

and

$$\vartheta_3\left(z,q\right) = \sum_{n\in\mathbb{Z}} q^{n^2} e^{2inz}$$

is the Jacobi theta function.

Proof. We have

$$e^{-\beta(x-s)^2} = \frac{\alpha}{\sqrt{\pi\left(\alpha-\beta\right)}} \int_{\mathbb{R}} e^{-\frac{\alpha\beta}{\alpha-\beta}(t-s)^2} e^{-\alpha(x-t)^2} dt,$$

and need to estimate

$$\int_{\mathbb{R}} F\left(t,s,x\right) dt - \sum_{k\in\mathbb{Z}} F\left(k,s,x\right),$$

where

$$F\left(t,s,x\right) = \frac{\alpha}{\sqrt{\pi\left(\alpha-\beta\right)}} e^{-\frac{\alpha\beta}{\alpha-\beta}(t-s)^2} e^{-\alpha(x-t)^2}.$$

Using Poisson's summation formula, we have

$$\left| \int_{\mathbb{R}} F\left(t,s,x\right) dt - \sum_{k\in\mathbb{Z}} F\left(k,s,x\right) \right| \leq \sum_{n=\pm1,\pm2,\ldots} \left| \widehat{F}\left(n,s,x\right) \right|, \tag{2.2}$$

where

$$\widehat{F}\left(n,s,x\right) = \int_{\mathbb{R}} F\left(t,s,x\right) e^{-2\pi itn} dt = e^{-n^2\frac{\pi^2}{\alpha}\left(1-\frac{\beta}{\alpha}\right)} e^{-\beta(x-s)^2} e^{-2\pi in\frac{\beta}{\alpha}(s-x)-2\pi inx}.$$

We obtain

$$\begin{aligned} \sum_{n=\pm1,\pm2,\ldots} \left| \widehat{F}\left(n,s,x\right) \right| &= \left(\sum_{n=\pm1,\pm2,\ldots} e^{-n^2\frac{\pi^2}{\alpha}\left(1-\frac{\beta}{\alpha}\right)} \right) e^{-\beta(x-s)^2} \\ &\leq \left(\vartheta_3\left(0, e^{-\frac{\pi^2}{\alpha}\left(1-\frac{\beta}{\alpha}\right)}\right) - 1 \right) e^{-\beta(x-s)^2} \end{aligned}$$

yielding the estimate (2.1). For similar type of analysis see [5, Section 2, eq.5], where we use the step size to control the accuracy of approximating integrals by a

sum, whereas here we fix the step size but select the exponent of the Gaussian ϕ to achieve the desired accuracy. □

Corollary 2.2. *Defining* $\phi(x) = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha x^2}$ *and assuming that* $0 < \beta \leq \alpha/4$, *we obtain from* (2.1)

$$\left| e^{-\beta(x-s)^2} - \sqrt{\frac{\alpha}{\alpha-\beta}} \sum_{k\in\mathbb{Z}} e^{-\frac{\alpha\beta}{\alpha-\beta}(k-s)^2} \phi(x-k) \right| \leq \epsilon e^{-\beta(x-s)^2}, \tag{2.3}$$

where

$$\epsilon = \vartheta_3\left(0, e^{-\frac{3\pi^2}{4\alpha}}\right) - 1. \tag{2.4}$$

Also, setting $s = 0$, $\beta = \alpha/4$ *and replacing* x *by* $2x$ *in* (2.1), *we obtain*

$$\left| \phi(x) - \sqrt{\frac{4\alpha}{3\pi}} \sum_{k\in\mathbb{Z}} e^{-\frac{\alpha}{3}k^2} \phi(2x-k) \right| \leq \epsilon \phi(x), \tag{2.5}$$

an approximate two-scale relation for Gaussians.

Consequently, we can use the function ϕ as a scaling function for an approximate multiresolution basis. By selecting an appropriate α we can reduce the value of ϵ in (2.4) and achieve arbitrary finite precision in (2.1). In the examples presented later in the paper, we use $\alpha = 1/4$ so that $\epsilon \approx 2.77 \cdot 10^{-13}$ (see Figures 4.2 and 4.4). We note that selecting $\alpha = 1/5$ yields $\epsilon \approx 2.22 \cdot 10^{-16}$ achieving double precision accuracy.

Remark 2.3. If $\beta = 0$ in (2.1), we recover the approximation described in [14]. However, if $\beta \neq 0$, our approximation (2.1) is substantially different and significantly more accurate than that obtained in [14]. The key difference, of course, is in the approach to the selection of coefficients of the shifted Gaussians.

Both, in Theorem 2.1 and Corollary 2.2, the error is described in terms of the quantity $\vartheta_3(0, e^{-\gamma}) - 1$ for some appropriate $\gamma > 0$. In the next lemma we show that, as a function of γ, $\vartheta_3(0, e^{-\gamma}) - 1$ decreases exponentially fast. Thus, slightly increasing the value of γ provides a much improved error. Also, using the lemma, for a given α and target accuracy ϵ, we can accurately estimate the maximum value of β to use in Corollary 2.2.

Lemma 2.4. *For* $\gamma \geq \gamma_0 > 0$,

$$\vartheta_3\left(0, e^{-\gamma}\right) - 1 \leq e^{\gamma_0} \left(\vartheta_3\left(0, e^{-\gamma_0}\right) - 1\right) e^{-\gamma}.$$

In particular, for $\gamma_0 = \pi$, *we have*

$$\vartheta_3\left(0, e^{-\gamma}\right) - 1 \leq c_\pi\, e^{-\gamma}, \text{ for } \gamma \geq \pi, \tag{2.6}$$

and

$$\sqrt{\frac{\pi}{\gamma}} < \vartheta_3\left(0, e^{-\gamma}\right) \leq \sqrt{\frac{\pi}{\gamma}} \left(1 + c_\pi\, e^{-\frac{\pi^2}{\gamma}}\right), \text{ for } \gamma \leq \pi, \tag{2.7}$$

where

$$c_\pi = e^\pi \left(\frac{\sqrt[4]{\pi}}{\Gamma\left(\frac{3}{4}\right)} - 1 \right) < 2.002. \tag{2.8}$$

Proof. Consider

$$f(\gamma) = e^{\gamma}\left(\vartheta_3\left(0, e^{-\gamma}\right) - 1\right) = 2\sum_{n\geq 1} e^{-\gamma\left(n^2-1\right)}.$$

Since $f'(\gamma) < 0$, f is a decreasing function which implies the first claim. For (2.6), use [1, page 325] to write $f(\pi) = e^{\pi}\left(\vartheta_3\left(0, e^{-\pi}\right) - 1\right) = e^{\pi}\left(\frac{\sqrt[4]{\pi}}{\Gamma\left(\frac{3}{4}\right)} - 1\right) < 2.002$. For (2.7), note that

$$\vartheta_3\left(0, e^{-\gamma}\right) = \sqrt{\frac{\pi}{\gamma}}\vartheta_3\left(0, e^{-\frac{\pi^2}{\gamma}}\right) = \sqrt{\frac{\pi}{\gamma}}\sum_{n\in\mathbb{Z}} e^{-\frac{\pi^2}{\gamma}n^2} > \sqrt{\frac{\pi}{\gamma}}. \tag{2.9}$$

Since for $\gamma \leq \pi$ we have $\frac{\pi^2}{\gamma} \geq \pi$, we obtain the result using (2.6) with γ replaced by $\frac{\pi^2}{\gamma}$. □

2.1. **Multiresolution analysis with a Gaussian as the scaling function.** We consider the family of functions

$$\left\{\phi_{j,k}(x) = 2^{j/2}\phi\left(2^j x - k\right) = 2^{j/2}\sqrt{\frac{\alpha}{\pi}}e^{-\alpha\left(2^j x-k\right)^2}\right\}_{j,k\in\mathbb{Z}} \tag{2.10}$$

normalized so that the L^2-norm

$$\|\phi_{j,k}\|_2 = \left(\frac{\alpha}{2\pi}\right)^{1/4}$$

does not depend on the scale j (*n.b.*, setting the norm to 1 would have resulted in bulkier formulas). In the next theorem, we show that a Gaussian with an arbitrary exponent can be projected (approximated) at an appropriate scale of GMRA. It is remarkable that the projection coefficients are obtained explicitly without solving any equations.

Theorem 2.5. *The Gaussian $e^{-\beta(x-s)^2}$ with exponent β and shift s* can be approximated in the GMRA (2.10) as

$$\left|e^{-\beta(x-s)^2} - \sum_{k\in\mathbb{Z}} g_k^j \phi_{jk}(x)\right| \leq \epsilon e^{-\beta(x-s)^2}, \tag{2.11}$$

where $4^{j-2}\alpha < \beta \leq 4^{j-1}\alpha$ and

$$g_k^j = 2^{-j/2}\sqrt{\frac{\alpha}{\pi\left(\alpha - 4^{-j}\beta\right)}}e^{-\frac{\alpha\beta}{\alpha-4^{-j}\beta}\left(2^{-j}k-s\right)^2}. \tag{2.12}$$

Proof. Given β, we select the scale j so that $4^{j-2}\alpha < \beta \leq 4^{j-1}\alpha$. Once j is selected, we rescale the exponent β and the shift s as $\widetilde{\beta} = 4^{-j}\beta$ and $\widetilde{s} = 2^j s$ and compute coefficients using (2.1) for the rescaled Gaussian $e^{-\widetilde{\beta}(x-\widetilde{s})^2}$,

$$\left|e^{-\widetilde{\beta}(x-\widetilde{s})^2} - \frac{\alpha}{\sqrt{\pi\left(\alpha-\widetilde{\beta}\right)}}\sum_{k\in\mathbb{Z}} e^{-\frac{\alpha\widetilde{\beta}}{\alpha-\widetilde{\beta}}(k-\widetilde{s})^2}e^{-\alpha(x-k)^2}\right| \leq \epsilon e^{-\widetilde{\beta}(x-\widetilde{s})^2}.$$

Replacing the variable x by $2^j x$ in the expression above, we arrive at (2.11). □

3. Distribution of the product of independent random variables

To motivate our approach, we first consider two independent Gaussian variables $X \sim N\left(\mu_x, \sigma_x^2\right)$, $Y \sim N\left(\mu_y, \sigma_y^2\right)$, with $\mu_x = \mu_y = 0$. The distribution of their product variable, $Z = XY$, is given by

$$p_{00}(t) = \frac{1}{\pi \sigma_x \sigma_y} K_0\left(\frac{|t|}{\sigma_x \sigma_y}\right), \tag{3.1}$$

where K_0 is modified Bessel function of the second kind. Hence, the PDF p_{00} has an integrable logarithmic singularity at zero even though the normal distributions do not have any singularities. In general, the contribution of such singularity to the cumulative distribution function (CDF) may or may not be significant but it is always present as it follows from the representation of p in (1.1). Specifically, for $p_{00}(t)$ in (3.1), as t approaches 0 we have

$$p_{00}(t) = \frac{1}{\pi \sigma_x \sigma_y}\left(-\gamma + \log\left(2\sigma_x \sigma_y\right) - \log|t|\right) + \mathcal{O}\left(t^2 \log|t|\right),$$

where γ is the Euler constant. For three or more Gaussian PDFs with zero means, their product can be expressed via Meijer G-functions (see [10, Definition 9.301]). However, when the Gaussian PDFs have non-zero means, we are not aware of any explicit expression for the PDF of their product in terms of special functions.

We start by computing the PDF of the product of two independent random variables, X with PDF $f(x)$, $\int_{\mathbb{R}} f(x)\, dx = 1$, and the Gaussian variable $Y \sim N\left(\mu_y, \sigma_y^2\right)$ with PDF

$$g(y) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}}.$$

The PDF of their product, $Z = XY$, is expressed via the integral

$$p(t) = \frac{1}{\sqrt{2\pi}\sigma_y} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)\, e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \delta(t - xy)\, dxdy. \tag{3.2}$$

We assume that the PDF f is a smooth function except at zero and a finite number of points where it may have integrable singularities. Specifically, f may have an integrable logarithmic singularity at $x = 0$, $f(x) \sim (\log|x|)^m$, for some integer $m \geq 1$. We also assume a mild decay of f as $x \to \infty$,

$$|f(x)| \leq \frac{C_0}{|x|^2}, \text{ for } |x| > C_1,$$

where C_0 and C_1 are positive constants. In particular, f can be the Cauchy distribution.

In the next theorem we obtain a multiresolution approximation of $p(t)$, where the coefficients at all scales are given explicitly by integrals of well-behaved functions over a unit interval. These integrals are evaluated numerically via adaptive integration, see Section 3.1 below.

Theorem 3.1. *The PDF of the product of two independent random variables, X and the Gaussian variable $Y \sim N\left(\mu_y, \sigma_y^2\right)$ in (3.2), can be approximated as*

$$\left| p(t) - \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} w_k^j \phi_{jk}(t) \right| \le \epsilon p(t) \tag{3.3}$$

where ϵ is given in (2.4),

$$w_k^j = 2^{-j/2} \log 2 \int_0^1 \frac{w_+(\tau) + w_-(\tau)}{\sqrt{1-4^{\tau-2}}} d\tau, \tag{3.4}$$

$$w_+(\tau) = \frac{1}{\sqrt{2\pi}\sigma_y} f\left(\frac{2^{2-j}}{\sqrt{2\alpha}\sigma_y} 2^{-\tau}\right) e^{-\frac{\alpha 4^{\tau-2}}{1-4^{\tau-2}}\left(k - \frac{\mu_y}{\sqrt{2\alpha}\sigma_y} 2^{-\tau+2}\right)^2},$$

and

$$w_-(\tau) = \frac{1}{\sqrt{2\pi}\sigma_y} f\left(-\frac{2^{2-j}}{\sqrt{2\alpha}\sigma_y} 2^{-\tau}\right) e^{-\frac{\alpha 4^{\tau-2}}{1-4^{\tau-2}}\left(k + \frac{\mu_y}{\sqrt{2\alpha}\sigma_y} 2^{-\tau+2}\right)^2}.$$

Proof. Splitting the integration over x in (3.2), we have

$$\begin{aligned} p(t) &= \frac{1}{\sqrt{2\pi}\sigma_y} \left(\int_0^\infty dx \int_{-\infty}^\infty f(x) e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \delta(xy-t) dy \right. \\ &+ \left. \int_{-\infty}^0 dx \int_{-\infty}^\infty f(x) e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \delta(xy-t) dy \right) \end{aligned}$$

and changing $x \to -x$ in the second integral, obtain

$$\begin{aligned} p(t) &= \frac{1}{\sqrt{2\pi}\sigma_y} \left(\int_0^\infty dx \int_{-\infty}^\infty f(x) e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \delta(xy-t) dy \right. \\ &+ \left. \int_0^\infty dx \int_{-\infty}^\infty f(-x) e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \delta(xy+t) dy \right). \end{aligned}$$

Since $\delta(xy) = \frac{1}{x}\delta(y)$ for $x > 0$, we have

$$p(t) = \frac{1}{\sqrt{2\pi}\sigma_y} \int_0^\infty \left(f(x) e^{-\frac{(\frac{t}{x}-\mu_y)^2}{2\sigma_y^2}} + f(-x) e^{-\frac{(\frac{t}{x}+\mu_y)^2}{2\sigma_y^2}} \right) \frac{1}{x} dx.$$

Changing variables $x = \frac{1}{\sqrt{2\alpha}\sigma_y} 2^{-s}$, we obtain

$$\begin{aligned} p(t) &= c \int_{-\infty}^\infty \left(f\left(\frac{2^{-s}}{\sqrt{2\alpha}\sigma_y}\right) e^{-\alpha 4^s \left(t - \frac{\mu_y}{\sqrt{2\alpha}\sigma_y} 2^{-s}\right)^2} \right. \\ &+ \left. f\left(-\frac{2^{-s}}{\sqrt{2\alpha}\sigma_y}\right) e^{-\alpha 4^s \left(t + \frac{\mu_y}{\sqrt{2\alpha}\sigma_y} 2^{-s}\right)^2} \right) ds, \end{aligned} \tag{3.5}$$

where $c = \log(2) / \left(\sqrt{2\pi}\sigma_y\right)$. Splitting the integration in (3.5) into integrals over unit intervals, we rewrite

$$\begin{aligned} p(t) &= c\sum_{j\in\mathbb{Z}}\int_{j-2}^{j-1}\left(f\left(\frac{2^{-s}}{\sqrt{2\alpha}\sigma_y}\right)e^{-\alpha 4^s\left(t-\frac{\mu_y}{\sqrt{2\alpha}\sigma_y}2^{-s}\right)^2}\right. \\ &+ \left. f\left(-\frac{2^{-s}}{\sqrt{2\alpha}\sigma_y}\right)e^{-\alpha 4^s\left(t+\frac{\mu_y}{\sqrt{2\alpha}\sigma_y}2^{-s}\right)^2}\right)ds \end{aligned}$$

and changing the variable of integration, $s = \tau + j - 2$, we arrive at

$$\begin{aligned} p(t) &= c\sum_{j\in\mathbb{Z}}\int_0^1\left[f\left(\frac{2^{-\tau-j+2}}{\sqrt{2\alpha}\sigma_y}\right)e^{-\alpha 4^{\tau+j-2}\left(t-\frac{\mu_y}{\sqrt{2\alpha}\sigma_y}2^{-\tau-j+2}\right)^2}\right. \\ &+ \left. f\left(-\frac{2^{-\tau-j+2}}{\sqrt{2\alpha}\sigma_y}\right)e^{-\alpha 4^{\tau+j-2}\left(t+\frac{\mu_y}{\sqrt{2\alpha}\sigma_y}2^{-\tau-j+2}\right)^2}\right]d\tau. \end{aligned} \tag{3.6}$$

For $\tau \in [0,1]$ we have $1 \le 4^\tau \le 4$, so that $\beta = \alpha 4^{\tau+j-2}$ satisfies $4^{j-2}\alpha \le \beta \le 4^{j-1}\alpha$. Consequently, using (2.11) and (2.12), we obtain

$$\left|e^{-\alpha 4^{\tau+j-2}\left(t\pm\frac{\mu_y}{\sqrt{2\alpha}\sigma_y}2^{-\tau-j+2}\right)^2} - \sum_{k\in\mathbb{Z}} g_k^{j,\pm}\phi_{jk}(t)\right| \le \epsilon e^{-\alpha 4^{\tau+j-2}\left(t\pm\frac{\mu_y}{\sqrt{2\alpha}\sigma_y}2^{-\tau-j+2}\right)^2},$$

where

$$g_k^{j,\pm} = 2^{-j/2}\sqrt{\frac{1}{1-4^{\tau-2}}}e^{-\frac{\alpha 4^{\tau-2}}{1-4^{\tau-2}}\left(k\pm\frac{\mu_y}{\sqrt{2\alpha}\sigma_y}2^{-\tau+2}\right)^2},$$

and use these expressions to replace the corresponding Gaussians in (3.6). Setting w_k^j as in (3.4), we now estimate $d(t) = p(t) - \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} w_k^j\phi_{jk}(t)$. We have

$$\begin{aligned} d(t) &= c\sum_{j\in\mathbb{Z}}\int_0^1\left[f\left(\frac{2^{-\tau-j+2}}{\sqrt{2\alpha}\sigma_y}\right)\left(e^{-\alpha 4^{\tau+j-2}\left(t-\frac{\mu_y}{\sqrt{2\alpha}\sigma_y}2^{-\tau-j+2}\right)^2} - \sum_{k\in\mathbb{Z}} g_k^{j,-}\phi_{jk}(x)\right)\right. \\ &+ \left. f\left(-\frac{2^{-\tau-j+2}}{\sqrt{2\alpha}\sigma_y}\right)\left(e^{-\alpha 4^{\tau+j-2}\left(t+\frac{\mu_y}{\sqrt{2\alpha}\sigma_y}2^{-\tau-j+2}\right)^2} - \sum_{k\in\mathbb{Z}} g_k^{j,+}\phi_{jk}(x)\right)\right]d\tau. \end{aligned} \tag{3.7}$$

Since f is non-negative, we obtain

$$\begin{aligned} |d(t)| &\le \epsilon c\sum_{j\in\mathbb{Z}}\int_0^1\left[f\left(\frac{2^{-\tau-j+2}}{\sqrt{2\alpha}\sigma_y}\right)e^{-\alpha 4^{\tau+j-2}\left(t-\frac{\mu_y}{\sqrt{2\alpha}\sigma_y}2^{-\tau-j+2}\right)^2}\right. \\ &+ \left. f\left(-\frac{2^{-\tau-j+2}}{\sqrt{2\alpha}\sigma_y}\right)e^{-\alpha 4^{\tau+j-2}\left(t+\frac{\mu_y}{\sqrt{2\alpha}\sigma_y}2^{-\tau-j+2}\right)^2}\right]d\tau. \end{aligned} \tag{3.8}$$

Using (3.6) and (3.8), we have

$$|d(t)| \le \epsilon p(t)$$

yielding the estimate in (3.3). □

3.1. **Adaptive method for computing integrals.** We compute the integrals

$$w_{j,k}^{\pm} = \int_0^1 f\left(\mp\frac{2^{2-j}}{\sqrt{2\alpha}\sigma_y}2^{-\tau}\right) e^{-\frac{\alpha 4^{\tau-2}}{1-4^{\tau-2}}\left(k\pm\frac{\mu_y}{\sqrt{2\alpha}\sigma_y}2^{-\tau+2}\right)^2} \frac{1}{\sqrt{1-4^{\tau-2}}}d\tau$$

using adaptive integration. The integrand may have only integrable singularities in the interior of the interval $[0,1]$ due to our assumption on f. Since f may have a logarithmic singularity at $x=0$, i.e. $f(|x|) \sim (\log|x|)^m$ near $x=0$, we have

$$f\left(\frac{2^{2-j}}{\sqrt{2\alpha}\sigma_y}2^{-\tau}\right) \sim \left(\log_2\left(\frac{2^{2-j}}{\sqrt{2\alpha}\sigma_y}\right)-\tau\right)^m,$$

so that, if j is large, f behaves like a polynomial as a function of τ . Such behavior of the integrand explains our choice of a simple and efficient method for adaptive integration. Specifically, we subdivide the integral $[0,1]$ adaptively in a hierarchical fashion while using a fixed number of Legendre nodes to compute the integral on each subinterval. The subdivision terminates when the integral computed on two subintervals matches that computed on the parent interval within an appropriately chosen accuracy. In our examples we used the Gauss-Legendre quadrature with 10 nodes and set the accuracy parameter of the adaptive integrator to 10^{-14}; the local maximum depth of subdivision remained small in all cases. We present a pseudo-code for this algorithm in Appendix A.

3.2. **Representing PDFs in GMRA.** In order to reduce the computation of the PDF of the product in the general case to the assumptions of Theorem 3.1, at least one of the PDFs has to be represented as a Gaussian mixture. We consider two approaches for this purpose. The first approach is applicable to distributions that can be expressed via integrals involving Gaussians. We illustrate this case using the Laplace distribution. The second approach relies on a rapidly decaying Fourier transform of a PDF so that it can be approximated by a bandlimited function. We illustrate this case using the Gumbel distribution. We also refer to [13, 14] for a way to approximate a smooth PDF via a linear combination of Gaussians provided that the accuracy requirements are not too stringent. We note that these approximations can be constructed even for smooth PDFs with compact support since, outside the support, the approximation can be made arbitrarily small.

3.2.1. *Discretization of integrals to approximate PDFs via Gaussian mixtures.* Although the PDF of the Laplace distribution has two parameters, it is sufficient to approximate the function in (4.3) for the particular case $b=1$ and $\mu=0$,

$$f(x) = \frac{1}{2}e^{-|x|}.$$

Following [5, Eq. 37], we have

$$\frac{1}{2}e^{-|x|} = \frac{1}{4\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-e^t/4 - x^2e^{-t} + \frac{1}{2}t}dt.$$

A linear combination of Gaussians is then obtained by discretizing this integral. It is sufficient to use the trapezoidal rule due to the rapid decay of the integrand as $t \to \pm\infty$. We limit integration to the interval $t \in [-40, 10]$ and use $h = 5/12$ as the

step size. Setting $t_n = -40 + 5 \cdot n/12$, and using $N = 120$ terms, we arrive at the Gaussian mixture approximation,

$$\frac{1}{2}e^{-|x|} \approx \frac{h}{4\sqrt{\pi}} \sum_{n=0}^{N-1} e^{-e^{-t_n}x^2} e^{-e^{t_n}/4+\frac{1}{2}t_n}.$$

We use this approximation in Example 4.2.6 to compute the PDF of the product of two independent random variables, one with Laplace distribution and the other with Gumbel distribution. Approximation of the latter via a Gaussian mixture is described next.

3.2.2. *Using an interpolating Gaussian-based scaling function to approximate PDFs via Gaussian mixtures.* An interpolating Gaussian-based scaling function spans the same subspace as the function ϕ, since it is constructed as

$$L(x) = \sum_{k\in\mathbb{Z}} c_k \phi(x-k).$$

Using the interpolating property $L(n) = \delta_{n0}$, $n \in \mathbb{Z}$, we also have

$$\delta_{n0} = \sum_{k\in\mathbb{Z}} c_k \phi(n-k).$$

Multiplying this identity by $e^{-2\pi inp}$, and adding over n, we obtain

$$1 = \left(\sum_{k\in\mathbb{Z}} c_k e^{-2\pi ikp}\right)\left(\sum_{n\in\mathbb{Z}} \phi(n)\, e^{-2\pi inp}\right). \tag{3.9}$$

By evaluating the Fourier transform,

$$\widehat{L}(p) = \left(\sum_{k\in\mathbb{Z}} c_k e^{-2\pi ikp}\right)\widehat{\phi}(p),$$

and using (3.9), we arrive at

$$\widehat{L}(p) = \frac{\widehat{\phi}(p)}{\sum_{n\in\mathbb{Z}} \phi(n)\, e^{-2\pi inp}}.$$

By Poisson summation formula,

$$\sum_{n\in\mathbb{Z}} \phi(n)\, e^{-2\pi inp} = \sqrt{\frac{\alpha}{\pi}} \sum_{n\in\mathbb{Z}} e^{-\alpha n^2} e^{-2\pi inp} = \sqrt{\frac{\alpha}{\pi}}\vartheta_3\left(\pi p, e^{-\alpha}\right),$$

we finally obtain

$$\widehat{L}(p) = \frac{\widehat{\phi}(p)}{\sqrt{\frac{\alpha}{\pi}}\vartheta_3\left(\pi p, e^{-\alpha}\right)}. \tag{3.10}$$

To illustrate the use of the interpolating scaling functions in the construction of Gaussian mixtures, we consider the Gumbel distribution in (4.4) with parameters $\sigma = 1$ and $\mu = 0$,

$$g(x) = e^{-x-e^{-x}}.$$

Gaussian mixtures for general parameters are then obtained from this case by rescaling. First we select an interval to approximate g. For example, outside the interval

$[-6, 50]$ the function g is less than $\approx 2 \cdot 10^{-22}$. We treat g on this interval as approximately periodic and, changing variables, $x = -6 + 56s$, rescale it to the interval $[0, 1]$,

$$\tilde{g}(s) = g(-6 + 56s). \tag{3.11}$$

Sampling $\tilde{g}(s)$ at N points, where N is sufficiently large, we write its approximation using the interpolating scaling function L as

$$\tilde{g}(s) = \sum_{k=0}^{N-1} \tilde{g}\left(\frac{k}{N}\right) N^{1/2} L(Ns - k).$$

Using the scaling function $\phi(s) = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha s^2}$, we seek the coefficients g_k such that

$$\tilde{g}(s) = \sum_{k=0}^{N-1} g_k N^{1/2} \phi(Ns - k). \tag{3.12}$$

Computing the Fourier transform, we obtain

$$\widehat{\tilde{g}}(p) = N^{-1/2} \left(\sum_{k=0}^{N-1} \tilde{g}\left(\frac{k}{N}\right) e^{-2\pi i k p/N} \right) \widehat{L}\left(\frac{p}{N}\right),$$

and

$$\widehat{\tilde{g}}(p) = N^{-1/2} \left(\sum_{k=0}^{N-1} g_k e^{-2\pi i k p/N} \right) \widehat{\phi}\left(\frac{p}{N}\right).$$

Sampling at N points and using (3.10), we obtain

$$\sum_{k=0}^{N-1} g_k e^{-2\pi i k n/N} = \frac{\sum_{k=0}^{N-1} \tilde{g}\left(\frac{k}{N}\right) e^{-2\pi i k n/N}}{\sqrt{\frac{\alpha}{\pi}} \vartheta_3\left(\pi \frac{n}{N}, e^{-\alpha}\right)}, \quad n = 0, \ldots N-1,$$

and compute the coefficients g_k via the Fast Fourier Transform (FFT). In this construction N is chosen to have small prime factors since we use FFT. We illustrate the result in Figure 3.1, where we set $N = 300$. We use this Gaussian mixture as an approximation of the Gumbel distribution in Example 4.2.6.

Remark 3.2. We plan to develop a more efficient algorithm to yield multiresolution representation of PDFs (i.e., with fewer terms) by applying the approach presented in this section in an iterative fashion.

3.3. **Computing the PDFs in GMRA representation.** For computing the PDF of the product of independent random variables where the original PDFs are in GMRA representation, we repeat the derivation in Theorem 3.1 using two representative Gaussians from our approximate GMRA (2.10), namely

$$f(x) = 2^{n/2} e^{-\alpha(2^n x - m)} \text{ and } g(y) = 2^{l/2} e^{-\alpha\left(2^l y - m'\right)}. \tag{3.13}$$

We have

$$p(t) = 2^{(n+l)/2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\alpha(2^n x - m)^2} e^{-\alpha\left(2^l y - m'\right)^2} \delta(xy - t)\, dxdy \tag{3.14}$$

and employing the same changes of variables as in the proof of Theorem 3.1, arrive at

$$p(t) = 2^{(n+l)/2} \int_0^{\infty} \left(e^{-\alpha(2^n x - m)^2} e^{-\alpha\left(2^l \frac{t}{x} - m'\right)^2} + e^{-\alpha(2^n x + m)^2} e^{-\alpha\left(2^l \frac{t}{x} + m'\right)^2} \right) \frac{1}{x} dx.$$

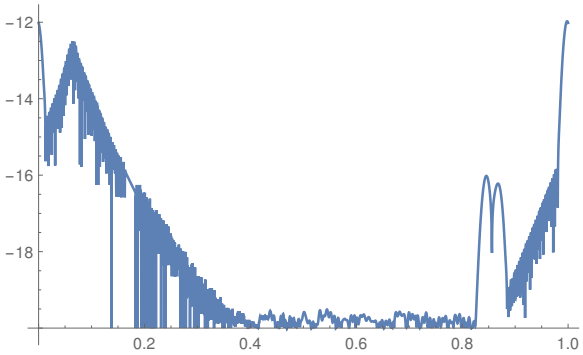

FIGURE 3.1. Logarithm (base 10) of the error of approximating the rescaled Gumbel PDF $\tilde{g}(s)$ in (3.11) as a Gaussian mixture with $N = 300$ terms.

Changing variables $y = 2^n x$,

$$p(t) = 2^{(n+l)/2} \int_0^\infty \left(e^{-\alpha(y-m)^2} e^{-\alpha\left(2^{n+l}\frac{t}{y} - m'\right)^2} + e^{-\alpha(y+m)^2} e^{-\alpha\left(2^{n+l}\frac{t}{y} + m'\right)^2} \right) \frac{1}{y} dy,$$

we define

$$\begin{aligned} u_{mm'}(t) &= p\left(2^{-n-l}t\right) 2^{-(n+l)/2} \\ &= \int_0^\infty \left(e^{-\alpha(y-m)^2} e^{-\alpha\left(\frac{t}{y} - m'\right)^2} + e^{-\alpha(y+m)^2} e^{-\alpha\left(\frac{t}{y} + m'\right)^2} \right) \frac{1}{y} dy. \end{aligned}$$

Changing variables $y = 2^{-s}$, $s \in \mathbb{R}$, and splitting the real axis into unit subintervals (see proof of Theorem 3.1), we obtain

$$u_{mm'}(t) = \log 2 \sum_{j \in \mathbb{Z}} \int_{j-2}^{j-1} \left(e^{-\alpha\left(2^{-s} - m\right)^2} e^{-\alpha\left(2^s t - m'\right)^2} + e^{-\alpha\left(2^{-s} + m\right)^2} e^{-\alpha\left(2^s t + m'\right)^2} \right) ds.$$

Changing variables one more time, $s = \tau + j - 2$, we arrive at

$$\begin{aligned} u_{mm'}(t) &= \log 2 \sum_{j \in \mathbb{Z}} \int_0^1 \left(e^{-\alpha\left(2^{2-j-\tau} - m\right)^2} e^{-\alpha\left(2^{\tau+j-2} t - m'\right)^2} \right. \\ &+ \left. e^{-\alpha\left(2^{2-j-\tau} + m\right)^2} e^{-\alpha\left(2^{\tau+j-2} t + m'\right)^2} \right) d\tau. \end{aligned}$$

Finally, we use Theorem 2.5 to obtain

$$u_{mm'}(t) = \log 2 \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \phi_{jk}(t) 2^{-j/2} u^j_{kmm'},$$

where

$$\begin{aligned} u^j_{kmm'} &= \int_0^1 \frac{1}{\sqrt{1 - 4^{\tau-2}}} \left(e^{-\alpha\left(2^{2-j-\tau} - m\right)^2} e^{-\frac{\alpha}{1-4^{\tau-2}}\left(2^{\tau-2}k - m'\right)^2} \right. \\ &+ \left. e^{-\alpha\left(2^{2-j-\tau} + m\right)^2} e^{-\frac{\alpha}{1-4^{\tau-2}}\left(2^{\tau-2}k + m'\right)^2} \right) d\tau. \end{aligned}$$

Since the integrand is comprised of well-behaved functions, we can discretize the integral using the Gauss-Legendre quadrature with M nodes directly (rather than

using an adaptive integrator). It follows that

$$u^j_{kmm'} = \sum_{i=1}^{M} \eta_i \left(U^{-,i}_{j,m} V^{-,i}_{km'} + U^{+,i}_{j,m} V^{+,i}_{km'} \right),$$

where $\eta_i = w_i \left(1 - 4^{\tau_i - 2}\right)^{-1/2}$, w_i are the weights of the Gauss-Legendre quadrature,

$$U^{\pm,i}_{j,m} = e^{-\alpha\left(2^{2-j-\tau_i} \pm m\right)^2},$$

and

$$V^{\pm,i}_{km'} = e^{-\frac{\alpha}{1-4^{\tau_i-2}}\left(2^{\tau_i-2}k+m'\right)^2}.$$

We note that matrices $U^{\pm,i}_{j,m}$ and $V^{\pm,i}_{km'}$ can be precomputed and that they are sparse (banded) since small entries can be removed. Also, these matrices do not depend on the scales n and l of the original Gaussians f and g in (3.13).

3.4. **Computing the PDF of a product of two random variables with a bivariate normal distribution.** We show how to compute the PDF p,

$$p(t) = \int_{\mathbb{R}^2} g(x,y)\, \delta(xy - t)\, dxdy, \tag{3.15}$$

where g is the joint PDF

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2}\begin{pmatrix} x-\mu_x \\ y-\mu_y \end{pmatrix}^T \Sigma^{-1} \begin{pmatrix} x-\mu_x \\ y-\mu_y \end{pmatrix}},$$

with

$$\Sigma^{-1} = \frac{1}{1-\rho^2} \begin{pmatrix} \frac{1}{\sigma_x^2} & -\frac{\rho}{\sigma_x\sigma_y} \\ -\frac{\rho}{\sigma_x\sigma_y} & \frac{1}{\sigma_y^2} \end{pmatrix}.$$

Here ρ is the correlation coefficient and $\rho = 0$ corresponds to the two variables being uncorrelated and independent. Using the next lemma we reduce the computation of p to the results on independent normal variables in Section 3.

Lemma 3.3. *The PDF p in* (3.15) *can be described as*

$$\begin{aligned} p(t) &= \frac{e^{\frac{\rho}{(1-\rho^2)}\left(\frac{t-\mu_x\mu_y}{\sigma_x\sigma_y}\right) + \frac{\rho^2}{2(1-\rho^2)}\left[\left(\frac{\mu_x}{\sigma_x}\right)^2 + \left(\frac{\mu_y}{\sigma_y}\right)^2\right]}}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \\ &\quad \int_{\mathbb{R}^2} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x} + \rho\frac{\mu_y}{\sigma_y}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y} + \rho\frac{\mu_x}{\sigma_x}\right)^2\right]} \delta(xy-t)\, dxdy. \end{aligned}$$

Proof. Denote by Q the quadratic form,

$$\begin{aligned} Q(x,y) &= -\frac{1}{2}\begin{pmatrix} x-\mu_x \\ y-\mu_y \end{pmatrix}^T \Sigma^{-1} \begin{pmatrix} x-\mu_x \\ y-\mu_y \end{pmatrix} \\ &= -\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right] \\ &= -\frac{\tilde{x}^2 + \tilde{y}^2 - 2\rho\tilde{x}\tilde{y}}{2(1-\rho^2)}, \end{aligned}$$

where $\tilde{x} = (x - \mu_x)/\sigma_x$ and $\tilde{y} = (y - \mu_y)/\sigma_y$. Since

$$\begin{aligned}\tilde{x}\tilde{y} &= \frac{xy}{\sigma_x\sigma_y} + \frac{\mu_x\mu_y}{\sigma_x\sigma_y} - \frac{\mu_y}{\sigma_x\sigma_y}(\sigma_x\tilde{x} + \mu_x) - \frac{\mu_x}{\sigma_x\sigma_y}(\sigma_y\tilde{y} + \mu_y) \\ &= \frac{xy - \mu_x\mu_y}{\sigma_x\sigma_y} - \frac{\mu_y}{\sigma_y}\tilde{x} - \frac{\mu_x}{\sigma_x}\tilde{y},\end{aligned}$$

with $q_x = \mu_x/\sigma_x$ and $q_y = \mu_y/\sigma_y$, we rewrite the numerator of $Q(x, y)$ as

$$\begin{aligned}\tilde{x}^2 + \tilde{y}^2 - 2\rho\tilde{x}\tilde{y} &= \tilde{x}^2 + 2\rho q_y\tilde{x} + \tilde{y}^2 + 2\rho q_y\tilde{y} - 2\rho\left(\frac{xy - \mu_x\mu_y}{\sigma_x\sigma_y}\right) \\ &= (\tilde{x} + \rho q_y)^2 + (\tilde{y} + \rho q_x)^2 - 2\rho\left(\frac{xy - \mu_x\mu_y}{\sigma_x\sigma_y}\right) - \rho^2\left(q_x^2 + q_y^2\right).\end{aligned}$$

Therefore, using that xy can be replaced by t in the integrand of (3.15), we have

$$\begin{aligned}\int_{\mathbb{R}^2} e^{Q(x,y)}\delta(xy - t)\,dxdy &= e^{\frac{\rho}{(1-\rho^2)}\frac{t-\mu_x\mu_y}{\sigma_x\sigma_y} + \frac{\rho^2}{2(1-\rho^2)}q_x^2 + q_y^2} \\ &\quad \int_{\mathbb{R}^2} e^{-\frac{1}{2(1-\rho^2)}\left[(\tilde{x}+\rho q_y)^2 + (\tilde{y}+\rho q_x)^2\right]}\delta(xy - t)\,dxdy\end{aligned}$$

which proves the result. □

Remark 3.4. The integral describing the PDF of the product (1.1) does not depend on the order of the factors f and g. However, the computation of (1.1) relies on the evaluation of the coefficients in (3.4) and these coefficients do depend on the order of the factors. While the resulting PDF does not depend on the chosen order, one of the two possible representations may require significantly more coefficients. To explain the reason for such behavior, let us assume that both random variables in the product have normal distributions but different means and deviations, μ_x, σ_x and μ_y, σ_y. According to Theorem 3.1, we need to evaluate

$$\frac{1}{\sqrt{2\pi}\sigma_y} f\left(\mp\frac{2^{2-j}}{\sqrt{2\alpha}\sigma_y}2^{-\tau}\right) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{(\mp 2^{2-j-\tau} - \sqrt{2\alpha}\sigma_y\mu_x)^2}{4\alpha\sigma_y^2\sigma_x^2}}.$$

As j becomes large, we have

$$\lim_{j\to\infty} e^{-\frac{(\mp 2^{2-j-\tau} - \sqrt{2\alpha}\sigma_y\mu_x)^2}{4\alpha\sigma_y^2\sigma_x^2}} = e^{-\frac{\mu_x^2}{2\sigma_x^2}}.$$

Thus, in order to reduce the number of scales, given the ratios,

$$\frac{\mu_x^2}{2\sigma_x^2} \text{ and } \frac{\mu_y^2}{2\sigma_y^2} \tag{3.16}$$

we pick f in (3.2) as the distribution with the largest ratio. The impact of ordering the factors in this way is illustrated in Example 4.2.2.

This remark implies that the ordering of the Gaussians f and g in (3.13)-(3.14) should be such that $m \geq m'$. We also note that in this particular case the ratios in (3.16) do not depend on the scales n and l.

Remark 3.5. Once a PDF f is represented via GMRA,

$$f(x) = \sum_{j,k} c_{j,k}\phi_{j,k}(x),$$

where j and k only take a finite number of integer values, it is immediate to compute its moments since the moments of the scaling functions $\phi_{j,k}$, are readily available. To be precise, for any nonnegative integer n, we compute

$$\mu_{j,k}^{(n)} = \int_{\mathbb{R}} \phi_{j,k}(x)\, x^n dx = 2^{j/2}\sqrt{\frac{\alpha}{\pi}} \int_{\mathbb{R}} e^{-\alpha\left(2^j x-k\right)^2} x^n dx = 2^{-j\left(n+\frac{1}{2}\right)} \mu_k^{(n)},$$

where

$$\mu_k^{(n)} = \sqrt{\frac{\alpha}{\pi}} \int_{\mathbb{R}} e^{-\alpha(x-k)^2} x^n dx. \tag{3.17}$$

Thus we have $\mu_k^{(0)} = 1$, $\mu_k^{(1)} = k$, and higher moments of the scaling functions $\phi_{j,k}$ can be evaluated using the recurrence,

$$\mu_k^{(n)} = k\mu_k^{(n-1)} + \frac{n-1}{2\alpha}\mu_k^{(n-2)},\ n \geq 2,$$

obtained by rewriting (3.17) as

$$\begin{aligned} \mu_k^{(n)} &= \sqrt{\frac{\alpha}{\pi}} \int_{\mathbb{R}} e^{-\alpha(x-k)^2} (k+x-k)\, x^{n-1} dx \\ &= k\mu_k^{(n-1)} + \sqrt{\frac{\alpha}{\pi}} \int_{\mathbb{R}} \frac{d}{dx}\left(\frac{e^{-\alpha(x-k)^2}}{-2\alpha}\right) x^{n-1} dx \end{aligned}$$

and integrating by parts. For the moments of f we arrive at the recurrence

$$\mathcal{M}^{(n)} = \int_{\mathbb{R}} f(x)\, x^n dx = \sum_j 2^{-j\left(n+\frac{1}{2}\right)} \left(\sum_k c_{j,k} k \mu_k^{(n-1)}\right) + \frac{n-1}{\alpha} 2^{-2j} \mathcal{M}^{(n-2)},\ n \geq 2,$$

where $\mathcal{M}^{(0)} = \sum_j 2^{-j/2} \sum_k c_{j,k}$ and $\mathcal{M}^{(1)} = \sum_j 2^{-3j/2} \sum_k c_{j,k} k$. In order to compute the variance, we have

$$E\left[\left(X - \mathcal{M}^{(1)}\right)^2\right] = \mathcal{M}^{(2)} - \left(\mathcal{M}^{(1)}\right)^2,$$

with similar formulas for skewness and kurtosis.

Remark 3.6. More generally, an important use of the PDF p_Z represented in GMRA is to compute, for a function u, the expectation of the variable $u(Z)$,

$$\mathrm{E}\left[u(Z)\right] = \int_{-\infty}^{\infty} u(t)\, p_Z(t)\, dt. \tag{3.18}$$

If the function u is given analytically (i.e., we can evaluate it at any point), the fact that we have a functional representation of p_Z allows us to use an appropriate quadrature to evaluate this integral to a desired accuracy. If only samples of the function u are provided, then we can treat p_Z as a weight and construct an appropriate quadrature with nodes at locations where the values of u are available. On the other hand, if the function u is also represented in GMRA the integral in (3.18) can be evaluated explicitly.

For example, if u is a non-negative function and $\tilde{p}_Z$ is the GMRA approximation of p_Z such that

$$|p_Z(t) - \tilde{p}_Z(t)| \leq \epsilon p_Z(t),$$

(e.g. as in Theorem 3.1), then we have the relative error estimate

$$\left|\mathrm{E}\left[u\left(Z\right)\right]-\tilde{\mathrm{E}}\left[u\left(Z\right)\right]\right|\leq\epsilon\mathrm{E}\left[u\left(Z\right)\right],$$

where

$$\tilde{\mathrm{E}}\left[u\left(Z\right)\right]=\int_{-\infty}^{\infty}u\left(t\right)\tilde{p}_{Z}\left(t\right)dt.$$

4. Examples of computing PDFs of products of independent random variables

Unless noted otherwise, in all our numerical examples we set the parameter $\alpha=1/4$, so that $\epsilon=2.77\cdot10^{-13}$ in (2.4). For representing the PDF of the product of random variables, we use a GMRA with a very large number of scales (one hundred dyadic scales) to demonstrate that we can accurately resolve the singularity at zero. Obviously, there is no need to use so many scales in practical applications; we use them to illustrate that the accuracy of our approach is well beyond the computational feasibility of any Monte-Carlo method. The test code was written in Fortran 90 and run on a laptop with a 2.20 GHz Intel processor. While no speed comparison with a Monte-Carlo method is meaningful, we note the actual time required for our algorithm. Computing the coefficients of the product of two normal distributions using 100 scales takes ≈0.12 seconds and evaluation of the resulting function at $5,000$ points takes ≈0.1 second. While the computation of the coefficients w_k^j in (3.3) is trivially parallelizable, we did not take advantage of it in our code.

In all our examples, we display significant coefficients of the GMRA representation, w_k^j in (3.3), as a matrix where the horizontal direction corresponds to the shift k and the vertical direction to the scale j. In interpreting these images (see Figures 4.1, 4.4, 4.6, 4.9, 4.11, 4.13, 4.15 and 4.17 below), it is important to remember that the shift k of the scaling functions $2^{j/2}\phi\left(2^{j}t-k\right)$ indicates the distance to the singularity at $t=0$, which is measured in scale dependent units, i.e. 2^{-j} at the scale j.

Just as an extra check, in the examples where explicit answers are not available, we verify the computed distributions by evaluating zero, first and second moments. In all cases these moments were within the accuracy of computation. The moments of PDFs in GMRA representation were computed as described in Remark 3.5.

4.1. **Accuracy tests.**

The PDFs of the product of two and three normal random variables with zero means are available analytically and we use them to ascertain accuracy of our algorithm.

4.1.1. *Two normal PDFs with zero means.* The exact PDF of the product of two random variables $X\sim N\left(\mu_x,\sigma_x^2\right)$, $Y\sim N\left(\mu_y,\sigma_y^2\right)$, with $\mu_x=\mu_y=0$, is given in (3.1). Setting $\sigma_x=\sigma_y=1$, we compute the PDF of the product $Z=XY$ via our algorithm and compare it with the exact PDF. The computed PDF p_Z and its GMRA coefficients are displayed in Figure 4.1 and the relative errors

$$\epsilon\left(x\right)=\log_{10}\left(\frac{\left|p_Z\left(10^x\right)-\frac{1}{\pi}K_0\left(10^x\right)\right|}{\frac{1}{\pi}K_0\left(10^x\right)}\right),\quad-30\leq x\leq0, \tag{4.1}$$

for different values of the parameter α of the GMRA, are displayed in Figure 4.2. The exact PDF in (3.1) has a logarithmic singularity at the origin and we use

100 scales to recover the PDF up to the interval $\left[-10^{-27}, 10^{-27}\right]$ with accuracy $\epsilon \approx 10^{-13}$. The limit of resolution of this singularity is illustrated in Figure 4.2 (again, well beyond what may be needed in applications).

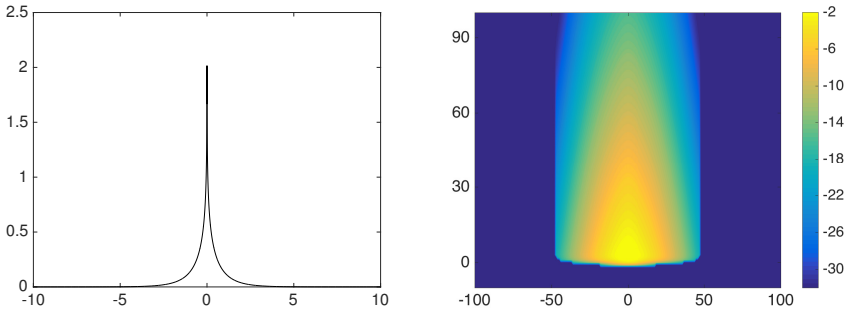

FIGURE 4.1. PDF p_Z of the product in Example 4.1.1 (left) and logarithm (base 10) of GMRA coefficients of p_Z (right).

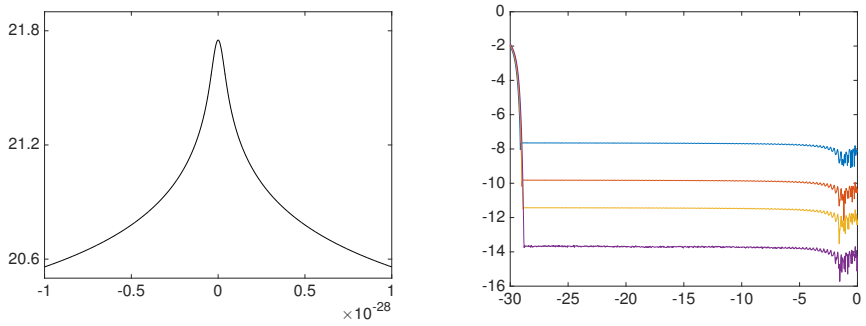

FIGURE 4.2. The limit of resolution of singularity of the PDF p_Z in Example 4.1.1 using GMRA with 100 scales and $\alpha = 0.25$ (left) and the relative error curves (4.1) shown, on logarithmic scale, over the interval $\left[10^{-30}, 0\right]$ for different parameters α in the GMRA (2.10). The bottom curve corresponds to $\alpha = 0.25$, followed by $\alpha = 0.3$, $\alpha = 0.35$ and $\alpha = 0.4$, as we go up in the display (right).

4.1.2. *Three normal PDFs with zero means.* We compute the PDF of the product of three zero-mean Gaussian random variables, starting with the GMRA representation of the product of two of them computed in the previous example. Given three zero-mean Gaussian random variables, $X_1 \sim N\left(0, \sigma_1^2\right)$, $X_2 \sim N\left(0, \sigma_2^2\right)$, and $X_3 \sim N\left(0, \sigma_3^2\right)$, the PDF of their product, $W = X_1 X_2 X_3$, is available analytically as

$$p(t) = \frac{G_{0,3}^{3,0}\left(0, 0, 0; \frac{t^2}{8\sigma_1^2\sigma_2^2\sigma_3^2}\right)}{(2\pi)^{3/2}\,\sigma_1\sigma_2\sigma_3},$$

where $G_{0,3}^{3,0}$ is a special case of a Meijer G-function, see [10, Definition 9.301]. The computed product PDF p_W is obtained by multiplying $X_3 \sim N(0,1)$ by the GMRA

representation of the PDF of $Z = X_1X_2$, as computed in Example 4.1.1. The PDFs of the random variables Z, X_3, and W are illustrated in Figure 4.3. As in Example 4.1.1, we use 100 scales and $\alpha = 0.25$ to recover the PDF p with accuracy $\epsilon \approx 10^{-13}$. In Figure 4.4 we display the GMRA coefficients of p_W and its logarithmic singularity resolved to the interval $\left[-10^{-28}, 10^{-28}\right]$. We compute the relative error,

$$\epsilon(x) = \log_{10}\left(\frac{|p_W(10^x) - p(10^x)|}{p(10^x)}\right), \quad -30 \le x \le 0, \tag{4.2}$$

for different values of the parameter α in the GMRA and display them in Figure 4.4.

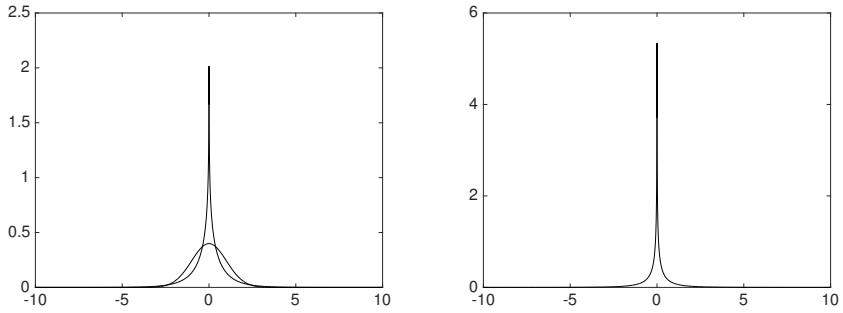

FIGURE 4.3. PDFs of the random variables $Z = X_1X_2$ and X_3 in Example 4.1.2 (left) and PDF of the product random variable $W = ZX_3$ in Example 4.1.2 (right).

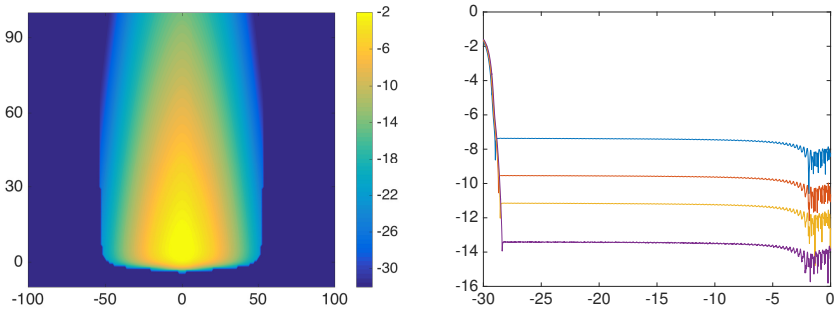

FIGURE 4.4. Logarithm (base 10) of GMRA coefficients of the PDF p_W in Example 4.1.2 (left) and the relative error curves (4.2) shown, on logarithmic scale, over the interval $\left[10^{-30}, 0\right]$ for different parameters α in the GMRA (2.10). The bottom curve corresponds to $\alpha = 0.25$, followed by $\alpha = 0.3$, $\alpha = 0.35$ and $\alpha = 0.4$, as we go up in the display, demonstrating resolution up to the interval $\left[-10^{-28}, 10^{-28}\right]$.

4.2. **Examples of computing the PDFs of products of random variables with various distributions.** In the next examples we use our algorithm to compute the PDFs of the product of two independent random variables with different distributions. In Section 4.2.1 we compute the product of two independent normal random variables with non-zero means. We comment on the impact of the order of integrands in Section 4.2.2. In Sections 4.2.3-4.2.5 we compute the product of a random variable with either Cauchy, Laplace or Gumbel distributions with a second normally distributed random variable. Finally, in Section 4.2.6 we compute the product of two independent random variables with Laplace and Gumbel distributions where both distributions are approximated via Gaussian mixtures as described in Section 3.2.

4.2.1. *Normal PDFs with non-zero means.* For the random variables $X \sim N(\mu_x, \sigma_x^2)$, $Y \sim N(\mu_y, \sigma_y^2)$ with $\mu_x = 2$, $\sigma_x = 1$ and $\mu_y = 1$, $\sigma_y = 1$, we compute the PDF p_Z of their product, $Z = XY$. Figure 4.5 shows the PDFs of the random variables X, Y and Z. As far as we know, there is no analytic expression for the result which we compute with accuracy $\approx 10^{-13}$. Using 100 scales, we resolve the logarithmic singularity at zero within an interval as small as $\left[-10^{-26}, 10^{-26}\right]$ (see Figure 4.6). We observe that to compute the PDF of Z to such accuracy and resolution via a Monte-Carlo method is not feasible. Moreover, we obtain the PDF of Z in a functional form which can be used for further computations. In contrast, a Monte-Carlo method using 10^8 samples and 10^3 bins of approximate size 0.04 yields just the histogram shown in Figure 4.7.

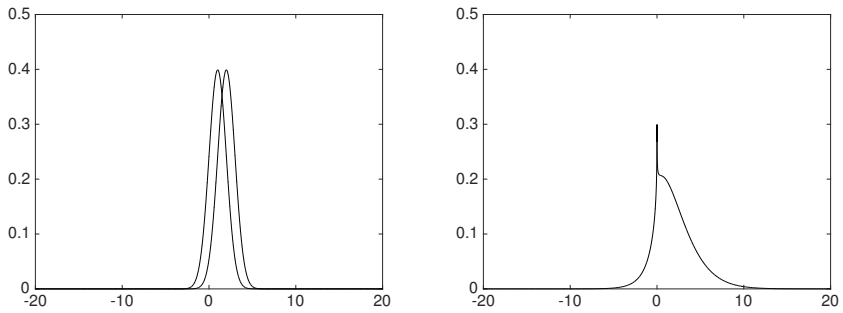

FIGURE 4.5. PDFs of the random variables X and Y of Example 4.2.1 (left) and the computed PDF p_Z of Example 4.2.1 (right).

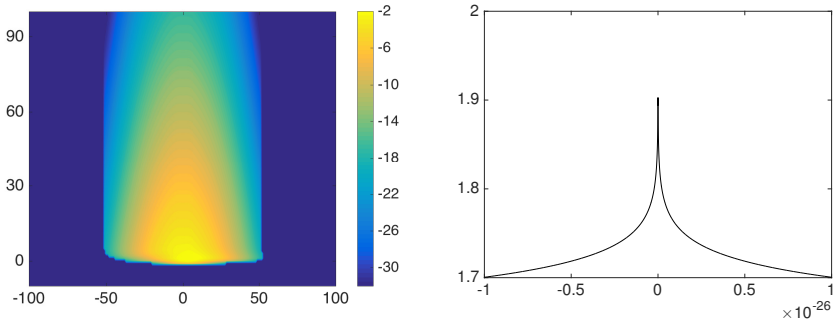

FIGURE 4.6. Logarithm (base 10) of GMRA coefficients of the PDF p_Z in Example 4.2.1 (left) and logarithmic singularity of p_Z resolved to the interval $\left[-2 \cdot 10^{-26}, 2 \cdot 10^{-26}\right]$.

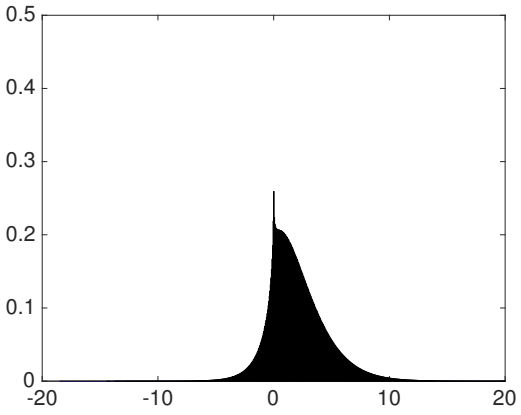

FIGURE 4.7. A histogram of the distribution of Example 4.2.1 computed with 10^8 samples using 10^3 bins of approximate size 0.04. The logarithmic singularity at the origin is poorly resolved (cf. the resolution of our method illustrated in Figure 4.6).

4.2.2. *Order of the integrands when computing with normal PDFs with non-zero means.* We demonstrate the effect of the order of factors on the coefficients of the representation of the product as described in Remark 3.4. We compute the PDFs of $Z_1 = XY$ and $Z_2 = YX$, where $X \sim N(6,1)$ and $Y \sim N(2,1)$, and display the results in Figures 4.8 and 4.9. While the values of the PDFs p_{Z_1} and p_{Z_2} are the same within the accuracy of computation, the coefficients of p_{Z_1} and p_{Z_2} do depend on the order of the factors as illustrated in Figure 4.9.

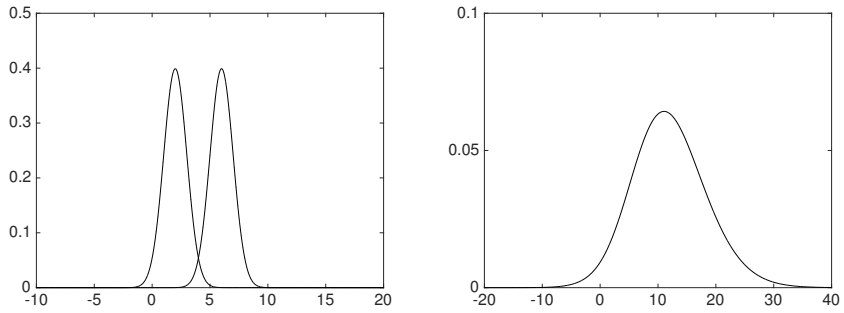

FIGURE 4.8. PDFs of random variables X and Y in Example 4.2.2 (left) and computed product PDF p_{Z_1} (right).

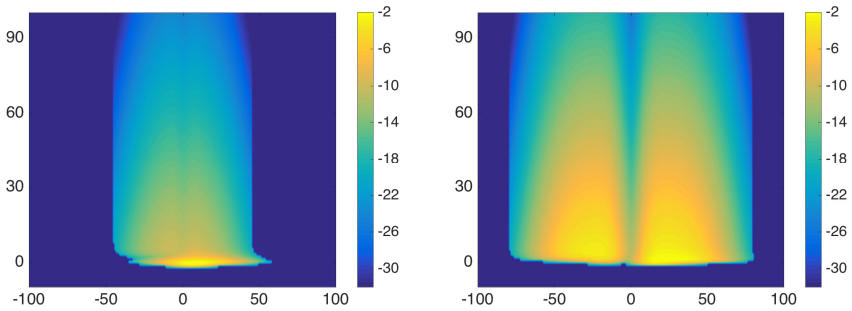

FIGURE 4.9. Logarithm (base 10) of GMRA coefficients of the PDF p_{Z_1} (left) and that of the PDF p_{Z_2} (right).

4.2.3. *Normal distribution and Cauchy distribution.* We now turn to computing the PDF of the product of random variables where one of them has a non-Gaussian distribution. First, we consider the PDF of the Cauchy distribution

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma}\left(\frac{\gamma^2}{(x-x_0)^2 + \gamma^2}\right),$$

where x_0 is the location parameter and γ is the scale parameter. The Cauchy distribution is defined on the real line and is an example of a heavy-tailed distribution for which the moments are infinite. In this example, we compute the PDF p_Z of $Z = XY$, where $X \sim f(x; -2, 1)$ and $Y \sim N(1.5, 1)$. We used 40 coarse ("negative") scales, i.e. $j = -40, \ldots, 100$, in order to capture the heavy-tail behavior of the product distribution (which also has no finite moments). By introducing these coarse scales, the integral of the resulting PDF over the real line differs from 1 by approximately $\approx 10^{-14}$. PDFs of the factors X and Y and the product Z are illustrated in Figure 4.10. In Figure 4.11, we display the GMRA coefficients of p_Z and resolve the singularity at zero of p_Z within the interval $\left[-10^{-28}, 10^{-28}\right]$.

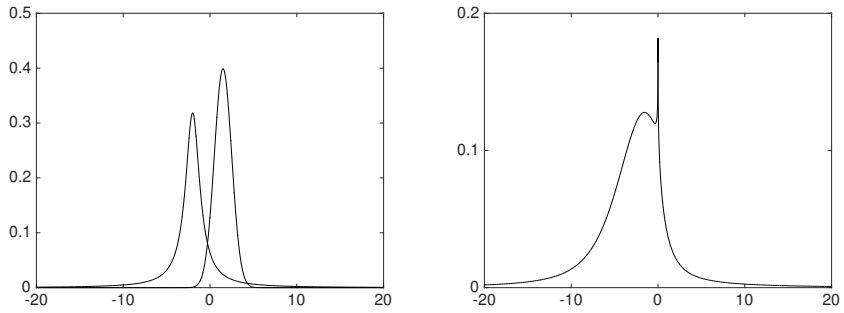

FIGURE 4.10. The PDFs of the random variables X and Y in Example 4.2.3 (left) and computed product PDF p_Z (right).

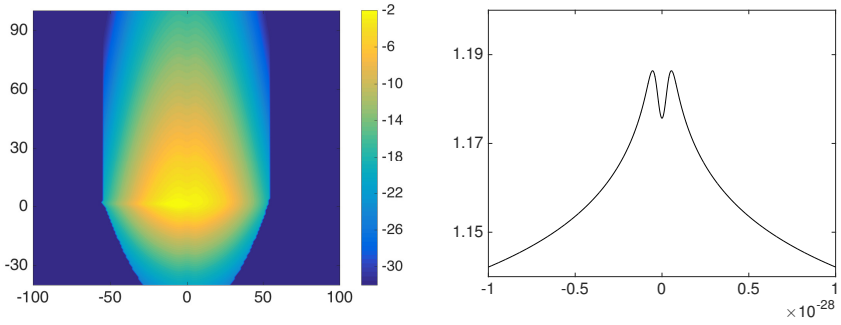

FIGURE 4.11. Logarithm (base 10) of GMRA coefficients of the PDF p_Z in Example 4.2.3 (left) and logarithmic singularity of p_Z resolved to the interval $\left[-10^{-28}, 10^{-28}\right]$ (right).

4.2.4. *Normal distribution and Laplace distribution.* We consider the PDF of the Laplace distribution

$$f(x;\mu,b) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}, \tag{4.3}$$

where μ is the location parameter and $b > 0$ is the scale parameter. The density function, $f(x;\mu,b)$, has a cusp at $x = \mu$. In this example, we compute the PDF p_Z of the product $Z = XY$, where $X \sim f(x;2,1)$ and $Y \sim N(1,1)$ and display these PDFs in Figure 4.12. In Figure 4.13 we display the GMRA coefficients of p_Z and resolve, within the interval $\left[-10^{-28}, 10^{-28}\right]$, the singularity of p_Z at zero. Note that the resulting product distribution does not have a singularity at $x = \mu$ but has a singularity at zero as in the other examples.

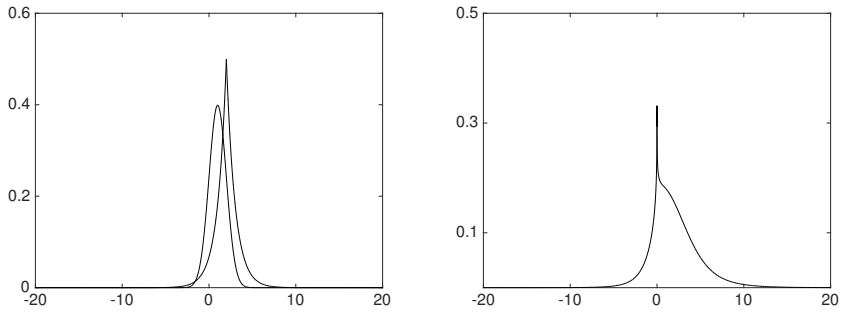

FIGURE 4.12. The PDFs of the random variables X and Y in Example 4.2.4 (left) and computed product PDF p_Z (right).

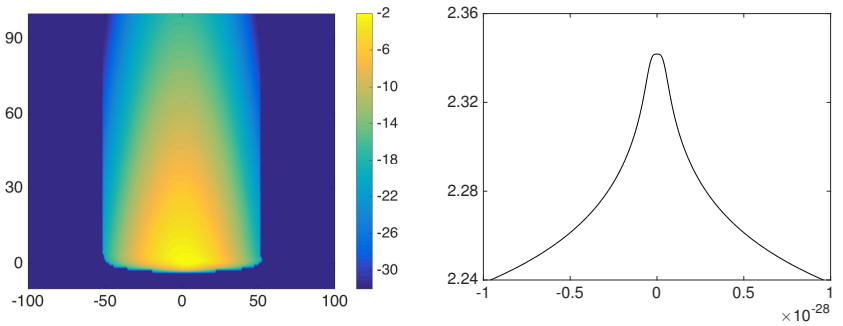

FIGURE 4.13. Logarithm (base 10) of GMRA coefficients of the PDF p_Z in Example 4.2.4 and (b) logarithmic singularity of p_Z resolved to the interval $\left[-10^{-28}, 10^{-28}\right]$.

4.2.5. *Normal distribution and Gumbel distribution.* We also consider the Gumbel distribution which is a special case of the generalized extreme value distribution. Its PDF is not symmetric and is given by

$$g\left(x;\mu,\sigma\right) = \frac{1}{\sigma} e^{\frac{-(x-\mu)}{\sigma} - e^{\frac{-(x-\mu)}{\sigma}}}, \tag{4.4}$$

where μ is the location parameter and σ is the scale parameter. Here, we compute the PDF p_Z of the product random variable $Z = XY$, where $X \sim g(x; 2, 1.5)$ and $Y \sim N(1, 3)$ and display these PDFs in Figure 4.14. In Figure 4.15 we display the GMRA coefficients of p_Z and resolve, within the interval $\left[-10^{-28}, 10^{-28}\right]$, the singularity of p_Z at zero .

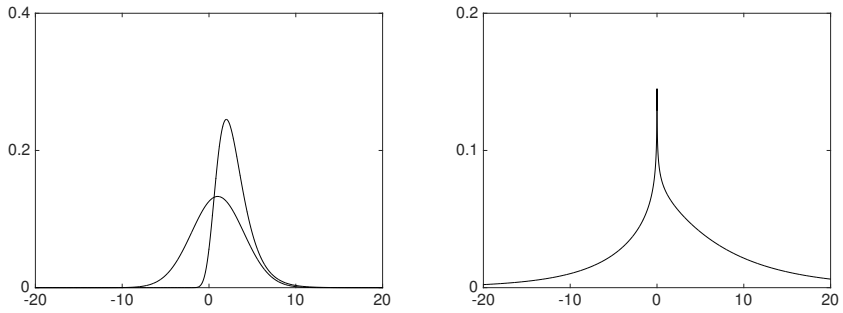

FIGURE 4.14. PDFs of random variables X and Y in Example 4.2.5 (left) and computed product PDF p_Z (right).

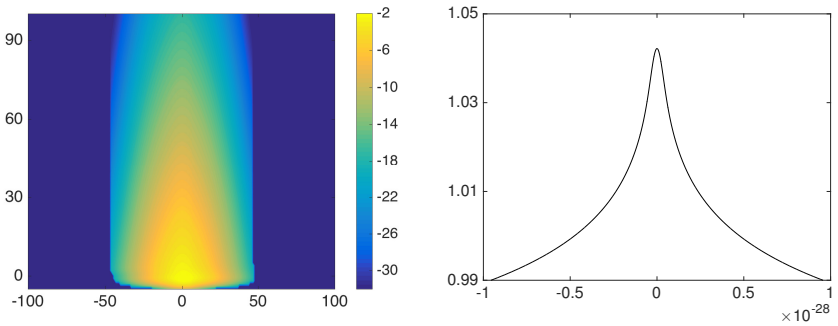

FIGURE 4.15. Logarithm (base 10) of GMRA coefficients of p_Z in Example 4.2.5 (left) and logarithmic singularity of p_Z resolved to the interval $\left[-10^{-28}, 10^{-28}\right]$ (right).

4.2.6. *Laplace distribution and Gumbel distribution.* In this example we compute the PDF p_Z of the product random variable $Z = XY$, where $X \sim f(x; 3, 1)$ is the Laplace distribution (4.3) and $Y \sim g(x; 2, 3)$ is the Gumbel distribution (4.4). Both distributions are approximated as Gaussian mixtures as described in Section 3.2. In computing p_Z we ordered the parameters of Gaussians to reduce the total number of coefficients as explained in Remark 3.4. The Laplace distribution was approximated by a Gaussian mixture with 120 terms and the Gumbel distribution with 300 terms, see Section 3.2 for details. We display the results in Figure 4.16 and Figure 4.17. The accuracy of the result is $\approx 10^{-6}$ as indicated by computing moments of p_Z.

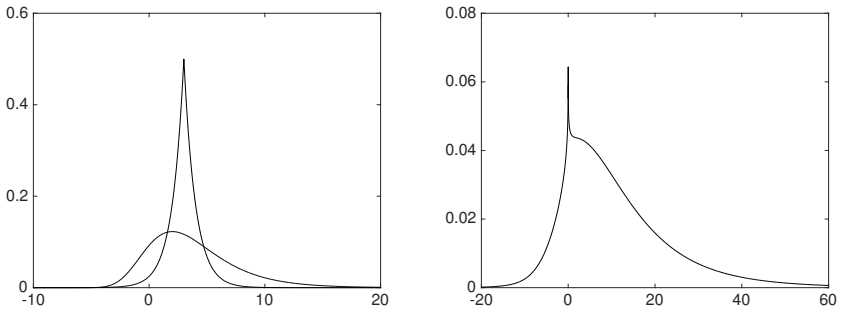

FIGURE 4.16. The PDFs of the random variables X and Y in Example 4.2.6 (left) and computed product PDF p_Z (right).

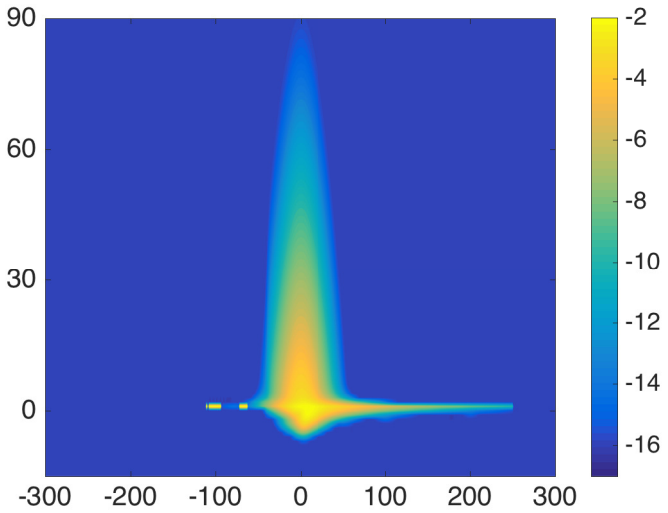

FIGURE 4.17. Logarithm (base 10) of GMRA coefficients of the PDF p_Z in Example 4.2.6.

5. Conclusions and further work

GMRA allows us to develop a numerical calculus of PDFs since we can now compute sums and products of independent random variables (as well as quotients, since the integral for the quotient of two independent random variables can be evaluated in a manner similar to that of the product). Indeed, if the initial PDFs are expressed in the GMRA, the resulting PDF is also expressed in the GMRA, i.e. as a multiresolution Gaussian mixture. Similar results hold for CDFs of random variables. We expect that generalizations of the GMRA to two and three dimensions can be accomplished in a straightforward manner, perhaps supplemented by employing the singular value decomposition (SVD) to simplify the representation of

coefficients on each scale. We plan to address such generalizations within relevant applications. A practical generalization to dimensions higher than three is more challenging and we plan to work towards it.

6. Appendix A

The Algorithm 1 describes an adaptive integrator for computing (with absolute accuracy ϵ) the integral

$$I = \int_a^b f(x)\,dx,$$

where f can have integrable singularities within $[a, b]$. Let $\{x_m\}_{m=1}^M$ and$\{w_m\}_{m=1}^M$ be nodes and weights of the Gauss-Legendre quadrature defined on the interval $[-1, 1]$. We use this quadrature to evaluate the sum

$$s = \sum_{m=1}^{M} \tilde{w}_m f(\tilde{x}_m),$$

where nodes and weights are adjusted to an interval $[a_0, b_0]$ as follows,

$$\tilde{x}_m = \frac{b_0 - a_0}{2} x_m + \frac{b_0 + a_0}{2}, \quad \tilde{w}_m = w_m \frac{b_0 - a_0}{2}.$$

We denote this sum as $s = \textbf{quadrature\_sum}\,(a_0, b_0, f)$. The computed value of the integral, I, is returned by Algorithm 1.

7. Appendix B

Computing the PDF of the product of independent random variables does not require all of the features of MRA since we only compute projections on the scaling functions at appropriate scales. In this appendix we briefly present some of the usual MRA constructs. We assume that the reader is familiar with the basic concepts of orthogonal and bi-orthogonal wavelet bases, see e.g. [8]. We note that the approximate GMRA is in many respects similar to MRA based on high order splines, see e.g. [7]. However, there are several special properties of GMRA that are not shared by spline based MRA. For example, as shown in Theorem 2.5, a Gaussian with an arbitrary exponent and an arbitrary shift can be represented at an appropriate subspace which does not depend on the shift. Another advantage of using GMRA is that it becomes possible to apply operators to functions in a semi-analytic fashion which is important in a number of applications we plan to address separately. Finally, we would like to emphasize that the fact that the GMRA is approximate is not an obstacle in practice since we can adjust the accuracy of the GMRA as needed.

7.1. Quadrature Mirror filters for GMRA.

We start by considering the ratio $\widehat{\phi}(2p)/\widehat{\phi}(p)$, where

$$\widehat{\phi}(p) = \int_{\mathbb{R}} \phi(x) e^{-2\pi i x p} dx = e^{-\frac{\pi^2}{\alpha} p^2}, \tag{7.1}$$

is the Fourier transform of the approximate scaling function in (2.10). We have

$$\frac{\widehat{\phi}(2p)}{\widehat{\phi}(p)} = e^{-\frac{3\pi^2}{\alpha} p^2} \tag{7.2}$$

Algorithm 1 Adaptive integrator

Input: interval $[a,b]$, function f, the target accuracy ϵ, and a fixed number M of quadrature nodes and weights.

Initialization:

$\mathbf{left\_p}\,(1) = a$

$\mathbf{right\_p}\,(1) = b$

$s\,(1) = \mathbf{quadrature\_sum}\,(a,b,f)$

$j = 1$

$I = 0$

Main loop: ($N_{iter}^{\max}$ is large enough to accommodate any reasonable recursion depth).

do $i = 1, N_{iter}^{\max}$

$c = \frac{1}{2}\left(\mathbf{left\_p}\,(j) + \mathbf{right\_p}\,(j)\right)$

$s_1 = \mathbf{quadrature\_sum}\left(\mathbf{left\_p}\,(j), c, f\right)$

$s_2 = \mathbf{quadrature\_sum}\left(\mathrm{c}, \mathbf{right\_p}\,(j), f\right)$

if $|s_1 + s_2 - s\,(j)| > \epsilon$ **then**

$\mathbf{left\_p}\,(j+1) = \mathbf{left\_p}\,(j)$

$\mathbf{right\_p}\,(j+1) = \frac{1}{2}\left(\mathbf{left\_p}\,(j) + \mathbf{right\_p}\,(j)\right)$

$s\,(j+1) = s_1$

$\mathbf{left\_p}\,(j) = \frac{1}{2}\left(\mathbf{left\_p}\,(j) + \mathbf{right\_p}\,(j)\right)$

$s\,(j) = s_2$

$j = j + 1$

else

$I = I + s_1 + s_2$

$j = j - 1$

If $j = 0$ **return** (the integral has been evaluated)

endif

enddo

and note that for an exact MRA this ratio should be a periodic function. This leads us to define the filter m_0 for GMRA as a periodization of (7.2),

$$m_0\,(p) = \sum_{n\in\mathbb{Z}} e^{-\frac{3\pi^2}{\alpha}(p-n)^2} = \sqrt{\frac{\alpha}{3\pi}}\vartheta_3\left(\pi p, e^{-\frac{\alpha}{3}}\right). \tag{7.3}$$

Using the Poisson summation formula applied to $\vartheta_3\left(\pi p, e^{-\frac{\alpha}{3}}\right)$, we have

$$\begin{aligned}\left|e^{-\frac{3\pi^2}{\alpha}p^2}-\sqrt{\frac{\alpha}{3\pi}}\vartheta_3\left(\pi p,e^{-\frac{\alpha}{3}}\right)\right| &= \left|e^{-\frac{3\pi^2}{\alpha}p^2}-\sum_{n\in\mathbb{Z}}e^{-\frac{3\pi^2}{\alpha}(p-n)^2}\right| \\ &= \sum_{n\neq 0}e^{-\frac{3\pi^2}{\alpha}(p-n)^2}.\end{aligned}$$

For $|p|\leq 1/2$, we obtain the error estimate

$$\left|e^{-\frac{3\pi^2}{\alpha}p^2}-\sqrt{\frac{\alpha}{3\pi}}\vartheta_3\left(\pi p,e^{-\frac{\alpha}{3}}\right)\right|\leq 2\sum_{n\geq 1}e^{-\frac{3\pi^2}{\alpha}\left(n-\frac{1}{2}\right)^2}=\vartheta_2\left(0,e^{-\frac{3\pi^2}{\alpha}}\right), \tag{7.4}$$

where the last equality can be found in [10, eq. 8.180.3]. Since

$$\begin{aligned}\sum_{n\geq 1}e^{-\frac{3\pi^2}{\alpha}\left(n-\frac{1}{2}\right)^2}=\sum_{n\geq 0}e^{-\frac{3\pi^2}{\alpha}\left(n+\frac{1}{2}\right)^2} &\leq e^{-\frac{3\pi^2}{4\alpha}}\sum_{n\geq 0}e^{-\frac{3\pi^2}{\alpha}n^2} \\ &= \frac{1}{2}e^{-\frac{3\pi^2}{4\alpha}}\left(1+\vartheta_3\left(0,e^{-\frac{3\pi^2}{\alpha}}\right)\right),\end{aligned}$$

Lemma 2.4 shows that the error (7.4) decays exponentially fast as a function of the parameter α. For example, if $\alpha=1/5$, the estimate (7.4) provides an error bound which is approximately 1.69×10^{-16}. and observe that (for any $\alpha\leq 3\pi$) (2.6) implies

$$\begin{aligned}|m_0(0)-1| &= m_0(0)-1=\sqrt{\frac{\alpha}{3\pi}}\vartheta_3\left(0,e^{-\frac{\alpha}{3}}\right)-1 \\ &= \vartheta_3\left(0,e^{-\frac{3\pi^2}{\alpha}}\right)-1\leq c_\pi\, e^{-\frac{3\pi^2}{\alpha}},\end{aligned} \tag{7.5}$$

where $c_\pi\approx 2.002$, see (2.8). Therefore, for the values of α of interest, $m_0(0)$ is extremely close to one. For example, for $\alpha=1/5$, (7.5) yields

$$|m_0(0)-1|\leq 1.015\times 10^{-64}.$$

7.1.1. *Quadrature Mirror filters for orthogonal approximate GMRA.* Next, we compute

$$\sum_{n\in\mathbb{Z}}\left|\widehat{\phi}(p+n)\right|^2=\sum_{n\in\mathbb{Z}}e^{-\frac{2\pi^2}{\alpha}(p+n)^2}=\sqrt{\frac{\alpha}{2\pi}}\sum_{n\in\mathbb{Z}}e^{-\frac{\alpha}{2}n^2}e^{2\pi inp}=\sqrt{\frac{\alpha}{2\pi}}\vartheta_3\left(\pi p,e^{-\frac{\alpha}{2}}\right),$$

to define the Fourier transform of the approximate, orthogonal Gaussian-based scaling function as

$$\widehat{\varphi}(p)=\frac{\widehat{\phi}(p)}{\sqrt{\sqrt{\frac{\alpha}{2\pi}}\vartheta_3\left(\pi p,e^{-\frac{\alpha}{2}}\right)}}. \tag{7.6}$$

The corresponding ~~approximate~~ quadrature mirror filter (QMF), determined by the identity $\widehat{\varphi}(2p)=M_a(p)\,\widehat{\varphi}(p)$, is then approximated~~given~~ by

$$M_a(p)=m_0(p)\left(\frac{\vartheta_3\left(\pi p,e^{-\frac{\alpha}{2}}\right)}{\vartheta_3\left(2\pi p,e^{-\frac{\alpha}{2}}\right)}\right)^{1/2}, \tag{7.7}$$

where we used (7.3) instead of (7.2). We note that $M_a(0) = m_0(0)$ so that, for $\alpha = 1/5$, $|M_a(0) - 1| \leq 1.015 \times 10^{-64}$ as well. The real valued filter M_a satisfies an approximate QMF equation,

$$M_a^2(p) + M_a^2\left(p + \frac{1}{2}\right) \approx 1. \tag{7.8}$$

For $\alpha = 1/5$, we numerically verify that

$$\left| M_a^2(p) + M_a^2\left(p + \frac{1}{2}\right) - 1 \right| < 0.11 \times 10^{-20}.$$

7.1.2. *Quadrature Mirror filters for biorthogonal approximate GMRA.* In order to obtain the filter that is dual to m_0, we rewrite (7.8) as

$$M_a^2(p) + M_a^2\left(p + \frac{1}{2}\right) = m_0(p) M_{00}(p) + m_0\left(p + \frac{1}{2}\right) M_{00}\left(p + \frac{1}{2}\right) \approx 1,$$

so that the dual filter M_{00} is given by

$$M_{00}(p) = \frac{m_0(p)\,\vartheta_3\left(\pi p, e^{-\frac{\alpha}{2}}\right)}{\vartheta_3\left(2\pi p, e^{-\frac{\alpha}{2}}\right)} = \sqrt{\frac{\alpha}{3\pi}} \frac{\vartheta_3\left(\pi p, e^{-\frac{\alpha}{3}}\right)\vartheta_3\left(\pi p, e^{-\frac{\alpha}{2}}\right)}{\vartheta_3\left(2\pi p, e^{-\frac{\alpha}{2}}\right)}. \tag{7.9}$$

We use M_{00} to obtain a formula for the projection of coefficients onto a coarser scale. It is sufficient to consider a projection from zero scale into the next coarser scale. Therefore, starting with f of the form

$$f(x) = \sum_{k \in \mathbb{Z}} f_k \phi(x - k),$$

so that

$$\widehat{f}(p) = \left(\sum_k f_k e^{-2\pi i k p} \right) \widehat{\phi}(p),$$

the coefficients of f in the next coarser scale are computed via (see e.g. [8, Section 8.3.1]),

$$\begin{aligned} g_n &= \frac{1}{\sqrt{2}} \int_{-1/2}^{1/2} \left[\left(\sum_k f_k e^{-2\pi i k p/2} \right) M_{00}\left(\frac{p}{2}\right) + \right. \\ &\quad \left. \left(\sum_k f_k e^{-2\pi i k (\frac{p}{2} + \frac{1}{2})} \right) M_{00}\left(\frac{p}{2} + \frac{1}{2}\right) \right] e^{2\pi i n p} dp. \end{aligned} \tag{7.10}$$

We would like to point out a numerical problem associated with using the filter M_{00}. The plot of M_{00} for $\alpha = 0.2$ in Figure 7.1 shows that one can expect a loss of significant digits due to its large dynamic range. To avoid the possible loss of significant digits we have several options. First, as is done in this paper, we can avoid using M_{00} altogether. The second option is to accept a lower (e.g. single precision) accuracy in computations with GMRA. Finally, one can look for an alternative (nonlinear) algorithm for computing coefficients g_n in (7.10). To illustrate the second option, we choose $\alpha = 0.4$ and observe the relatively small dynamic range of M_{00} for this value of α in Figure 7.1. Using $\alpha = 0.4$ in computing the PDFs results in roughly single precision accuracy as illustrated in Figures 4.2 and4.4.

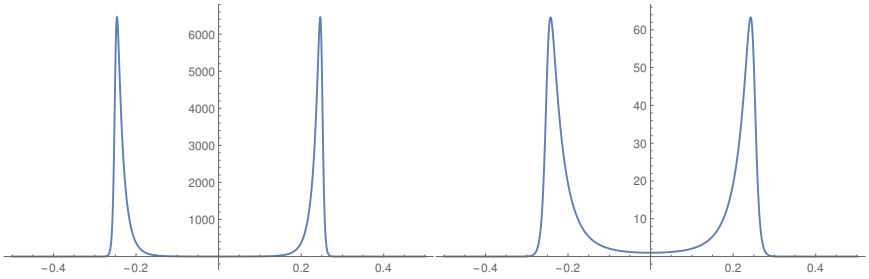

FIGURE 7.1. Filter $M_{00}(p)$ in (7.9) for $\alpha = 0.2$ (left) and $\alpha = 0.4$ (right).

7.2. **Constructing an exact MRA.** An alternative approach motivated by the construction of GMRA is to develop an exact MRA such that the resulting scaling function is closely approximated by a single Gaussian. Although we do not use this approach in this paper, we briefly indicate how to attain such construction. We can demonstrate that the nonnegative filter

$$M_0(p) = \left(\frac{\vartheta_3\left(\pi p, e^{-\frac{\alpha}{8}}\right)}{2\vartheta_3\left(2\pi p, e^{-\frac{\alpha}{2}}\right)} + \left(1 - \frac{\vartheta_3\left(0, e^{-\frac{\alpha}{8}}\right)}{2\vartheta_3\left(0, e^{-\frac{\alpha}{2}}\right)}\right)\cos 2\pi p \right)^{1/2}, \tag{7.11}$$

is an excellent approximation of the filter M_a in (7.7). In fact, for $\alpha = 0.2$, we verify numerically that

$$|M_a(p) - M_0(p)| < 1.05 \times 10^{-21}, \ \ p \in \left[-\frac{1}{2}, \frac{1}{2}\right]. \tag{7.12}$$

In the next theorem we show that the filter M_0 has the necessary properties to generate an exact orthogonal MRA.

Theorem 7.1. *The filter M_0 in (7.11) satisfies the QMF condition,*

$$M_0^2(p) + M_0^2\left(p + \frac{1}{2}\right) = 1,$$

and it is an even, positive function in $\left(-\frac{1}{2}, \frac{1}{2}\right)$, monotonically decreasing in $\left[0, \frac{1}{2}\right]$ with $M_0(0) = 1$.

Proof. Since ϑ_3 satisfies the well-known property

(7.13)

$$\begin{aligned}\vartheta_3\left(\pi p, e^{-\gamma}\right) + \vartheta_3\left(\pi p + \frac{\pi}{2}, e^{-\gamma}\right) &= \sum_{n\in\mathbb{Z}} e^{-\gamma n^2} e^{2\pi i p n} + \sum_{n\in\mathbb{Z}} (-1)^n e^{-\gamma n^2} e^{2\pi i p n}, \\ &= 2\sum_{n\in\mathbb{Z}} e^{-\gamma(2n)^2} e^{2\pi i p(2n)} \\ &= 2\vartheta_3\left(2\pi p, e^{-4\gamma}\right),\end{aligned}$$

and $\cos 2\pi\left(p + \frac{1}{2}\right) = -\cos 2\pi p$, we obtain

$$\begin{aligned}M_0^2(p) + M_0^2\left(p + \frac{1}{2}\right) &= \frac{\vartheta_3\left(\pi p, e^{-\frac{\alpha}{8}}\right) + \vartheta_3\left(\pi p + \frac{\pi}{2}, e^{-\frac{\alpha}{8}}\right)}{2\vartheta_3\left(2\pi p, e^{-\frac{\alpha}{2}}\right)} \\ &= 1.\end{aligned} \tag{7.14}$$

By construction, we have $M_0(0) = 1$ and, since $\vartheta_3(\pi p, e^{-\gamma})$ is an even function, M_0 is even. Using (2.9) and observing that $\vartheta_3(0, e^{-\gamma})$ is strictly decreasing for $\gamma > 0$, we have

$$\frac{\vartheta_3\left(0, e^{-\frac{\alpha}{8}}\right)}{2\vartheta_3\left(0, e^{-\frac{\alpha}{2}}\right)} = \frac{\vartheta_3\left(0, e^{-\frac{8\pi^2}{\alpha}}\right)}{\vartheta_3\left(0, e^{-\frac{2\pi^2}{\alpha}}\right)} < 1.$$

Hence, denoting by

$$\eta = 1 - \frac{\vartheta_3\left(0, e^{-\frac{\alpha}{8}}\right)}{2\vartheta_3\left(0, e^{-\frac{\alpha}{2}}\right)},$$

the constant in (7.11), we know that $\eta \in (0,1)$. For $p \in \left[0, \frac{1}{2}\right)$, we denote by f the function

$$f(p) = M_0^2(p) = \frac{\vartheta_3\left(\pi p, e^{-\frac{\alpha}{8}}\right)}{2\vartheta_3\left(2\pi p, e^{-\frac{\alpha}{2}}\right)} + \eta\cos(2\pi p) = g(p) + \eta\cos(2\pi p),$$

and notice that, due to (7.14), we have $f(0) = 1$ and $f\left(\frac{1}{2}\right) = 0$. Since $\eta > 0$ and $\cos(2\pi p)$ is strictly decreasing on $(0, 1/2)$, to show that f is strictly decreasing in this interval, it is enough to show that g has the same property. By (7.13), we have

$$g(p) = \frac{\vartheta_3\left(\pi p, e^{-\frac{\alpha}{8}}\right)}{\vartheta_3\left(\pi p, e^{-\frac{\alpha}{8}}\right) + \vartheta_3\left(\pi\left(p + \frac{1}{2}\right), e^{-\frac{\alpha}{8}}\right)}$$

and, taking derivatives, $g'(p) < 0$ is equivalent to the inequality

$$\frac{d}{dp}\ln\vartheta_3\left(\pi p, e^{-\frac{\alpha}{8}}\right) < \frac{d}{dp}\ln\vartheta_3\left(\pi\left(p + \frac{1}{2}\right), e^{-\frac{\alpha}{8}}\right). \tag{7.15}$$

From [15, 20.5.12], we have, for any $\gamma > 0$,

$$\frac{d}{dp}\ln\vartheta_3\left(\pi p, e^{-\gamma}\right) = -\frac{4}{\pi}\sin 2\pi p \prod_{n\geq 1}\frac{1}{\left|1 + e^{-\gamma(2n-1)}e^{2\pi p i}\right|}$$

and, hence,

$$\frac{d}{dp}\ln\vartheta_3\left(\pi\left(p + \frac{1}{2}\right), e^{-\frac{\alpha}{8}}\right) = \frac{4}{\pi}\sin 2\pi p \prod_{n\geq 1}\frac{1}{\left|1 - e^{-\frac{\alpha}{8}(2n-1)}e^{2\pi p i}\right|},$$

which implies (7.15) and, thus, that f is strictly decreasing on $(0, 1/2)$. Therefore, f is also positive in that interval. It follows that $\sqrt{f} = M_0$ is strictly decreasing on $(0, 1/2)$. □

Since M_0 has no zeros in $[-1/3, 1/3]$, using Corollary 6.3.2 in [8, p. 186], we know that the QMF M_0 generates an orthogonal MRA. The Fourier transform of the resulting scaling function Φ can now be obtained as

$$\widehat{\Phi}(p) = \prod_{n\geq 1} M_0\left(\frac{p}{2^n}\right).$$

Due to (7.12), if we use a finite number of scales, the exact MRA and approximate GMRA generated by the function $\widehat{\varphi}(p)$ in (7.6) are exchangeable for an appropriately selected parameter α tuned to the user-selected accuracy.

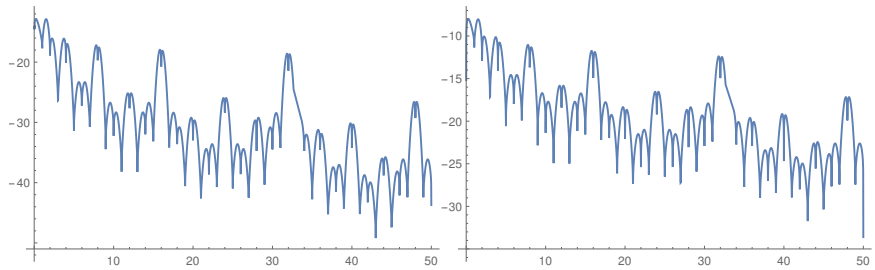

FIGURE 7.2. The logarithm (base 10) of the difference between the scaling function defined by the product in (7.16) and the Gaussian (7.1) for $\alpha = 0.25$ (left) and $\alpha = 0.4$ (right) on the interval $[0, 50]$.

7.2.1. *An alternative approach to constructing exact MRA.* We have strong numerical evidence but no proofs that the following construction also leads to an exact MRA. We modify m_0 in (7.3) to define

$$M_{exact}(p) = \frac{m_0(p) - m_0(1/2)}{m_0(0) - m_0(1/2)} = \frac{\vartheta_3\left(\pi p, e^{-\frac{\alpha}{3}}\right) - \vartheta_3\left(\frac{\pi}{2}, e^{-\frac{\alpha}{3}}\right)}{\vartheta_3\left(0, e^{-\frac{\alpha}{3}}\right) - \vartheta_3\left(\frac{\pi}{2}, e^{-\frac{\alpha}{3}}\right)},$$

so that $M_{exact}(0) = 1$ and $M_{exact}(1/2) = 0$. Then the scaling function for the exact (non-orthogonal) MRA would be given by the infinite product,

$$\widehat{\phi}_{exact}(p) = \prod_{j=1}^{\infty} M_{exact}\left(\frac{p}{2^j}\right), \tag{7.16}$$

in which case it holds that $\widehat{\phi}_{exact}(0) = 1$ and $\widehat{\phi}_{exact}(n) = 0$, $n \in \mathbb{Z}/\{0\}$. We verify numerically that

$$\left|\widehat{\phi}_{exact}(p) - \widehat{\phi}(p)\right| \leq \epsilon,$$

where ϵ is of the same order as in (2.4) and $\widehat{\phi}(p)$ is given in (7.1). In Figure 7.2 we plot the error $\log_{10}\left|\widehat{\phi}_{exact}(p) - \widehat{\phi}(p)\right|$ for two choices of the parameter $\alpha = 0.25$ and $\alpha = 0.4$. Effectively we force contributions outside the main band to be below the desired accuracy ϵ, which allows us to use a single Gaussian as an approximate scaling function.

Remark 7.2. Scaling functions of all MRAs have "bumps" in the Fourier domain, i.e. contributions outside the main band, due to their construction as an infinite product of periodic functions. As an example, see the discussion in [8, p. 245]. It is particularly easy to see this phenomenon for the B-splines. The Fourier transform of the non-orthogonal scaling function forB-splines (of odd degree m) is given by

$$\widehat{\varphi}_{spline}(p) = \left(\frac{\sin \pi p}{\pi p}\right)^{m+1},$$

so that the periodic contributions of the numerator outside the main band are suppressed (but not eliminated) by the denominator. We observe the same phenomenon in the behavior of $\widehat{\phi}_{exact}$.

References

[1] B. C. Berndt. *Ramanujan's notebooks. Part V*. Springer-Verlag, New York, 1998.

[2] G. Beylkin, V. Cheruvu, and F. Pérez. Fast adaptive algorithms in the non-standard form for multidimensional problems. *Appl. Comput. Harmon. Anal.*, 24(3):354–377, 2008.

[3] G. Beylkin, G. Fann, R. J. Harrison, C. Kurcz, and L. Monzón. Multiresolution representation of operators with boundary conditions on simple domains. *Appl. Comput. Harmon. Anal.*, 33:109–139, 2012. http://dx.doi.org/10.1016/j.acha.2011.10.001.

[4] G. Beylkin and L. Monzón. On approximation of functions by exponential sums. *Appl. Comput. Harmon. Anal.*, 19(1):17–48, 2005.

[5] G. Beylkin and L. Monzón. Approximation of functions by exponential sums revisited. *Appl. Comput. Harmon. Anal.*, 28(2):131–149, 2010.

[6] G. Beylkin, L. Monzón, and I. Satkauskas. On computing distributions of products of non-negative random variables. *Appl. Comput. Harmon. Anal.*, 2017. http://dx.doi.org/10.1016/j.acha.2017.08.008, also see arXiv:1707.07762.

[7] C. Chui. *An Introduction to Wavelets*. Academic Press, Boston, MA, 1992.

[8] I. Daubechies. *Ten Lectures on Wavelets*. CBMS-NSF Series in Applied Mathematics. SIAM, 1992.

[9] S. Efromovich. *Nonparametric curve estimation: methods, theory, and applications*. Springer Science & Business Media, 2008.

[10] I. S. Gradshteyn, I. M. Ryzhik, A. Jeffrey, and D. Zwillinger. *Table of integrals, series, and products*. Academic Press, 7 edition, 2007.

[11] R.J. Harrison, G.I. Fann, T. Yanai, Z. Gan, and G. Beylkin. Multiresolution quantum chemistry: basic theory and initial applications. *J. Chem. Phys.*, 121(23):11587–11598, 2004.

[12] T. S. Haut and G. Beylkin. Fast and accurate con-eigenvalue algorithm for optimal rational approximations. *SIAM J. Matrix Anal. Appl.*, 33(4):1101–1125, 2012. doi: 10.1137/110821901.

[13] F. Lanzara, V. Maz'ya, and G. Schmidt. Numerical solution of the Lippmann-Schwinger equation by approximate approximations. *J. Fourier Anal. Appl.*, 10(6):645–660, 2004.

[14] V. Maz'ya and G. Schmidt. *Approximate approximations*, volume 141 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2007.

[15] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, editors. *NIST Handbook of Mathematical Functions*. Cambridge University Press, New York, NY, 2010.

[16] J. S. Simonoff. *Smoothing methods in statistics*. Springer Science & Business Media, 2012.

[17] H. W. Sorenson and D. L. Alspach. Recursive Bayesian estimation using Gaussian sums. *Automatica—J. IFAC*, 7:465–479, 1971.

[18] M.D. Springer. *Algebra of Random Variables*. John Wiley & Sons, 1979.

Department of Applied Mathematics, University of Colorado at Boulder, UCB 526, Boulder, CO 80309-0526