Journal de la Société Française de Statistique
Vol. 159 No. 1 (2018)

On numerical computation for the distribution of the convolution of N independent rectified Gaussian variables

Titre: Éléments de calculs pour la distribution de N variables Gaussiennes rectifiées indépendantes

Maxime Beauchamp[1]

Abstract: For large *N* and when no variables is predominant over the others, the central limit theorem (CLT) shall apply to the sum of random variables with negative values reset to zero. The parameters of the normal distribution are simply obtained by computing the expected value and the variance of each left rectified distributions. But for small *N*, the distribution of the sum is clearly not Gaussian and can present several modes and a strong skewness. In this paper, a way of computing the probability density function of the sum of *N* independent rectified Gaussian variables is presented, so that the calculation issues raised by the convolution product is solved. Some numerical examples are given and the validity of this approach is assessed through a comparison with a Monte-Carlo approach and an application to the PAH's (Polycyclic Aromatic Hydrocarbon) batch filters measurements is provided.

Résumé : Pour *N* grand, le thèorème de la limite centrale s'applique si l'on cherche à sommer des variables aléatoires gaussiennes rectifiées, i.e. dont les valeurs négatives sont remises à zéro. Les paramètres de la gaussienne sont obtenus en sommant les espérances et les variances de chaque gaussienne censurée. Pour *N* petit, la distribution de la somme n'est évidemment pas gaussienne : elle peut présenter plusieurs modes et une forte asymétrie. Dans ce papier, le calcul de la densité de probabilité associée à la somme de variables aléatoires gaussiennes rectifiées indépendantes est présenté. Il est obtenu classiquement à partir d'un produit de convolution mais il présente des complications calculatoires qui sont détaillées et résolues. Une comparaison avec des simulations Monte-Carlo est fournie pour valider les développements. Enfin, une application à des données de mesures par filtres de HAP (Hydrocarbure Aromatique Polycyclique) est également présentée.

Keywords: densité de probabilité, variable gaussienne rectifiée, somme de variables indépendantes
Mots-clés : probability density function, rectified Gaussian variable, sum of independent variables
AMS 2000 subject classifications: 60C99, 62-N99, 62-P12

1. Introduction

The way of dealing with negative values of a random variable, by censoring or rectifying them to zero is a widespread problem in Statistics. While the question of censoring is well documented and already used in a lot of models, whose most famous applications through the truncated Gaussian variable are Tobit models (Tobin, 1958) and Probit models (Bliss, 1934), the second way of handling the negative values through their so-called rectification to zero has been less studied. In particular, the computation of the probability density function (denoted as pdf throughout the

[1] Mines ParisTech, Géosciences, Equipe géostatistique, 35 rue Saint Honoré, 77305 Fontainebleau, France
Institut National de l'Environnement Industriel et des Risques (INERIS), Direction des risques chroniques, Parc Technologique Alata, 60550 Verneuil-en-Halatte, France
E-mail : maxime.beauchamp@mines-paristech.fr

paper) of the sum of N rectified variables is intricate when the normal approximation of the central limit theorem can not be applied. Indeed, when these variables are assumed to be independent, the pdf involves the multiple convolution of densities with singularities at 0.

In this paper, the sum of N independent rectified Gaussian variables is adressed. A rectified Gaussian variables X^+ is basically a Gaussian variable $X \sim \mathcal{N}(\mu, \sigma^2)$ whose negative values are set to zero:

$$X^+ = \max(0, X) \tag{1}$$

This issue has not been explored yet in the literature. In a very recent paper from Krenek et al. (2016), the convolutions of truncated normal random variables (not to be confused with the rectified Gaussian variable) is approximated through the use of the error function but the closed-form of the probability density function is not found. An older work of McConalogue (1981) provides some numerical solutions for the convolution of integrals with densities having discontinuities at the origin, which can be useful to approximate the solution to our problem, but still does not solve it explicitly. However, the rectified Gaussian distribution is involved in many applications, mainly in biological neural network and signal processing (see e.g. Charles and Fyfe, 1998). It has also been applied to factor analysis (Harva and Kabán, 2007) with a lot of applications in computational biology.

In the first section, the rectified Gaussian distribution is defined. Then the analytical expression of the pdf for the sum of N independent rectified Gaussian variables is given: it describes the notion of multiple convolution of Gaussian densities ϕ with singularities at 0 through the use of the Heaviside function H. A theorem is provided with its proof to compute the exact value of these multiple convolution by translating the integral form of the convolution to its probabistic interpretation. The practical implementation of this probabilistic form is given with R-codes, and two additional functions are provided to compute the pdf and the approximation of the cumulative distribution function (cdf) through a Simpson scheme. In the section dedicated to some numerical computations, a comparison is made with a Monte-Carlo approach to confirm that the densities actually follow the simulated histograms. Last, an application is provided for the probability of exceeding a limit value applied to batch filters pollution measurements, namely the BaP (Benzo(a)pyrene). Because the data are assumed to be Gaussian and have a limit of quantification (LQ) below which the values are neglected, i.e. set to this LQ, the rectified Gaussian distribution applies exactly. In addition, the data are built from independent set of filters that cover different time periods so they do not present any obvious temporal correlation. Therefore the pdf of the sum (and the mean) of the batch of measurements is known from the theorem. From this pdf, the probability of exceeding the regulatory limit value can be estimated, using the standard deviation of the batch filters data.

2. Main results

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ be a Gaussian variable with mean μ and variance σ^2. Its pdf is noted $\phi(x)$. Let X^+ denote the rectified Gaussian variable which is constrained to be non-negative, i.e. all its negative values are set to zero. Such a variable can be seen as a zero-inflated continuous process (see e.g Lambert, 1992 for the first use of a zero-inflated discrete Poisson process applied to counting data with frequent zero values) in which the first process is governed by a Dirac measure at $x = 0$ and a Gaussian process for x strictly positive. As a consequence, the resulting distribution of X^+ is a mixture of a discrete distribution at 0 and a continuous distribution on the interval $]0,+\infty[$.

At $x = 0$, the probability density function (pdf) $\phi^+(x)$ equals to:

$$\phi^+(x,\mu,\sigma) = \Phi(-\frac{\mu}{\sigma}).\delta_0(x) \tag{2}$$

where $\Phi(x) = \mathbb{P}(X \leq x)$, $X \sim \mathcal{N}(0,1)$ and $\delta_0(x)$ is the Dirac measure at 0.

On $]0;+\infty[$, the pdf $\phi^+(x)$ is:

$$\phi^+(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{H}(x) = \phi(x)H(x) \triangleq \phi H(x) \tag{3}$$

where $H(x) = \mathbf{1}_{x>0}$ denotes the Heaviside function. From now on, the product of the probability density function $\phi(x)$ by $H(x)$ is noted $\phi H(x)$ to ease the readability of the convolution products.

The mean of X^+ is greater than μ because the probability mass related to the negative values is shifted to 0, doing so the variance of X^+ becomes lower than σ^2, see Appendix A for the details:

$$\begin{aligned}
\mathbb{E}(X^+) &= \mu\left(1-\Phi(-\frac{\mu}{\sigma})\right) + \frac{\sigma}{\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{-\mu}{\sigma})^2} \\
\mathrm{Var}(X^+) &= (\mu^2+\sigma^2)\left(1-\Phi(-\frac{\mu}{\sigma})\right) + \frac{\mu\sigma}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{-\mu}{\sigma}\right)^2} - \left[\mathbb{E}(X^+)\right]^2
\end{aligned}$$

2.1. *Probability density function of* $\sum_{i=1}^N X_i^+$

Let $\phi_i(x)$ denote $\phi(x,\mu_i,\sigma_i)$, the probability density function of the Gaussian variable $X_i \sim \mathcal{N}(\mu_i,\sigma_i)$ with mean μ_i and variance σ_i^2. $\phi_i^+(x)$ stands for the probability density function of the related rectified Gaussian variable.

For $N = 2$ independent rectified Gaussian variables X_1^+ and X_2^+, the density probability function $\phi_{\sum_i X_i^+}(x)$ is:

$$\begin{aligned}
\phi_{X_1^+ + X_2^+}(x) &= (\phi_1^+ * \phi_2^+)(x) \\
&= \left[\Phi(-\frac{\mu_1}{\sigma_1}).\Phi(-\frac{\mu_2}{\sigma_2}).(\delta_0 * \delta_0)(x)\right] + \left[\Phi(-\frac{\mu_1}{\sigma_1}).(\phi_2 H * \delta_0)(x)\right] \\
&+ \left[\Phi(-\frac{\mu_2}{\sigma_2}).(\phi_1 H * \delta_0)(x)\right] + (\phi_1 H * \phi_2 H)(x)
\end{aligned}$$

For N =3, it is:

$$\phi_{\sum_{i=1}^3 X_i^+}(x) = \Phi(-\frac{\mu_1}{\sigma_1})\Phi(-\frac{\mu_2}{\sigma_2})\Phi(-\frac{\mu_3}{\sigma_3}).\delta_0(x)$$

$$
\begin{aligned}
&+ \quad H(x)\left[\Phi(-\frac{\mu_1}{\sigma_1})\Phi(-\frac{\mu_2}{\sigma_2})\phi_3(x)+\Phi(-\frac{\mu_1}{\sigma_1})\Phi(-\frac{\mu_3}{\sigma_3})\phi_2(x)\right.\\
&+ \quad \left.\Phi(-\frac{\mu_2}{\sigma_2})\Phi(-\frac{\mu_3}{\sigma_3})\phi_1(x)\right]+\left[\Phi(-\frac{\mu_1}{\sigma_1}).(\phi_2 H*\phi_3 H)(x)\right]\\
&+ \quad \left[\Phi(-\frac{\mu_2}{\sigma_2})(\phi_1 H*\phi_3 H)(x)\right]+\left[\Phi(-\frac{\mu_3}{\sigma_3})(\phi_1 H*\phi_2 H)(x)\right]\\
&+ \quad \left[\bigotimes_{k=1}^{3}\phi_k H\right](x)
\end{aligned}
$$

For N variables, the pdf of the sum $\sum_i X_i^+$ is:

$$
\phi_{\sum_i X_i^+}(x)=\prod_{i=1,\cdots,N}\Phi(-\frac{\mu_i}{\sigma_i})\delta_0(x)+\sum_{i=1}^{N}\sum_{l=1}^{|S_i|}\prod_{j\in S_{il}}\Phi(-\frac{\mu_j}{\sigma_j})\left[\bigotimes_{k\notin S_{il}}\phi_k H\right](x) \tag{4}
$$

where S_i is the combination set of $(N-i)$ elements among N, S_{il} is the l-th $(N-i)$-tuple of S_i. Last, the operator $\bigotimes$ denotes the multiple convolution product, i.e. $\left[\bigotimes_{k=1\cdots,N} z_k\right](x)=(z_1*z_2*\cdots*z_N)(x)$. Conventionally, $\bigotimes_{\emptyset}=1$ and $\prod_{\emptyset}=1$.

2.2. *Multiple convolution* $[\bigotimes_k \phi_k H](x)$

The computation of Eq. (4) is complex due to the calculation of the multiple convolution $[\bigotimes_k \phi_k H](x)$. The main result of this paper is a new method to compute this expression:

Theorem 1: Probabilistic closed-form for the multiple convolution $[\bigotimes_k \phi_k H](x)$***]***

Let $(\Omega,\mathscr{F},\mathbb{P})$ *be a probability space, and* $X_k:\Omega\to\mathbb{R},\ k\in\mathbb{N}$, *be independent Gaussian variables defined on that space with expected values* μ_k *and variances* σ_k^2. $\phi_k(x)$ *denote the probability density function of* X_k *and* $H(x)$ *is the Heaviside function.*

The multiple convolution $\left[\bigotimes_{k=1}^N \phi_k H\right](x)$ *is given by:*

$$
\begin{aligned}
\left[\bigotimes_{k=1}^{N}\phi_k H\right](x) \quad = \quad & \int_{b_-^{(N-1)}(x)}^{b_+^{(N-1)}(x)} e^{-\frac{\left[u^{(N-1)}\right]^2}{2}} \int_{\alpha_{-,(N-2)}^{(N-2)}(u^{(N-1)})+\beta_{-,(N-2)}}^{\alpha_{+,(N-2)}^{(N-2)}(u^{(N-1)})+\beta_{+,(N-2)}} e^{-\frac{\left[u^{(N-2)}\right]^2}{2}}\cdots \\
& \cdots \int_{\left(\sum_{k=j+1}^{N-1}\alpha_{-,j}^{(k-1)}(u^{(k)})\right)+\beta_{-,j}(x)}^{\left(\sum_{k=j+1}^{N-1}\alpha_{+,j}^{(k-1)}(u^{(k)})\right)+\beta_{+,j}(x)} e^{-\frac{\left[u^{(j)}\right]^2}{2}}\cdots du^{(1)}du^{(2)}\cdots du^{(N-1)}
\end{aligned} \tag{5}
$$

where:

$$
\begin{aligned}
\mu^{(k)} \quad &= \quad \sum_{i=1}^{k}\mu_i,\quad \sigma^{(k)}=\sum_{i=1}^{k}\sigma_i,\quad l^{(k)}=\frac{\sigma^{(k+1)}}{\sigma^{(k)}\sigma_{k+1}},\quad m^{(k)}=\frac{\sigma^{(k)}}{\sigma^{(k+1)}\sigma_{k+1}}\\
y^{(k)} \quad &= \quad \mu^{(k)}-x,\quad r^{(k)}=s^{(k)}-\mu^{(k)},\quad u^{(k)}=\frac{r^{(k)}\left[\sigma^{(k+1)}\right]^2+y^{(k+1)}\left[\sigma^{(k)}\right]^2}{\sigma^{(k)}\sigma^{(k+1)}\sigma_{k+1}}
\end{aligned}
$$

$$
\begin{aligned}
b_{+}^{(j)}(x) &= \left(\sum_{k=j+1}^{N-1} \alpha_{+,j}^{(k-1)} \cdot u^{(k)}\right) + \beta_{+,j}(x), \quad b_{-}^{(j)}(x) = \left(\sum_{k=j+1}^{N-1} \alpha_{-,j}^{(k-1)} \cdot u^{(k)}\right) + \beta_{-,j}(x) \\
\alpha_{+,j}^{(k)} &= \frac{\prod_{i=j+1}^{k} m^{(i)}(l^{(j)} - m^{(j)})}{\prod_{i=j+1}^{k+1} l^{(i)}}, \quad \alpha_{-,j}^{(k)} = -m^{(j)} \frac{\prod_{i=j+1}^{k} m^{(i)}}{\prod_{i=j+1}^{k+1} l^{(i)}} \\
\beta_{+,j}(x) &= \left(b_{-}^{(1)}(0) - m^{(1)}\mu^{(2)} + m^{(j)} \frac{y^{(N)} \prod_{i=j+1}^{N-1} m^{(i)}}{\prod_{i=j+1}^{N-1} l^{(i)}}\right) \\
\beta_{-,j}(x) &= \left(b_{-}^{(1)}(0) + \mu^{(2)}(l^{(1)} - m^{(1)}) + \frac{y^{(N)} \prod_{i=j+1}^{N-1} m^{(i)}(m^{(j)} - l^{(j)})}{\prod_{i=j+1}^{N-1} l^{(i)}}\right)
\end{aligned}
$$

Let $Z = \left(U^{(N-1)}, T_{+}^{(N-2)}, T_{-}^{(N-2)}, \cdots, T_{+}, T_{-}\right)$ be a Gaussian vector with elements all normally distributed with zero mean and unit standard deviation::

$$
U^{(N-1)} \sim \mathcal{N}(0,1), \;\; T_{+}^{(j)} = \frac{U^{(j)} - \alpha_{+}^{(j)} U^{(j+1)}}{\sqrt{(\alpha_{+}^{(j)})^2 + 1}} \sim \mathcal{N}(0,1), \;\; T_{-}^{(j)} = \frac{\alpha_{-}^{(j)} U^{(j+1)} - U^{(j)}}{\sqrt{(\alpha_{-}^{(j)})^2 + 1}} \sim \mathcal{N}(0,1)
$$

and its covariance matrix defined by:

$$
\begin{aligned}
\mathrm{Cov}(U^{(N-1)}, T_{+}^{(i)}) &= -\frac{\alpha_{+,i}^{(N-2)}}{\sqrt{\sum_{k_i=i+1}^{N-1} \alpha_{+,i}^{(k_i-1)} + 1}}, \quad \mathrm{Cov}(U^{(N-1)}, T_{-}^{(i)}) = \frac{\alpha_{-,i}^{(N-2)}}{\sqrt{\sum_{k_i=i+1}^{N-1} \alpha_{-,i}^{(k_i-1)} + 1}} \\
\mathrm{Cov}(T_{+}^{(i)}, T_{+}^{(j)}) &= \frac{\sum_{k_i=i+1}^{N-1} \sum_{k_j=j+1}^{N-1} \alpha_{+,j}^{(k_j-1)} \alpha_{+,i}^{(k_i-1)} \mathbf{1}_{k_i=k_j} - \sum_{k_i=i+1}^{N-1} \alpha_{+,i}^{(k_i-1)} \mathbf{1}_{k_i=j} - \sum_{k_j=j+1}^{N-1} \alpha_{+,j}^{(k_j-1)} \mathbf{1}_{k_j=i} + \mathbf{1}_{i=j}}{\sqrt{\sum_{k_i=i+1}^{N-1} \alpha_{+,i}^{(k_i-1)} + 1} \sqrt{\sum_{k_j=j+1}^{N-1} \alpha_{+,j}^{(k_j-1)} + 1}} \\
\mathrm{Cov}(T_{-}^{(i)}, T_{-}^{(j)}) &= \frac{\sum_{k_i=i+1}^{N-1} \sum_{k_j=j+1}^{N-1} \alpha_{-,j}^{(k_j-1)} \alpha_{-,i}^{(k_i-1)} \mathbf{1}_{k_i=k_j} - \sum_{k_i=i+1}^{N-1} \alpha_{-,i}^{(k_i-1)} \mathbf{1}_{k_i=j} - \sum_{k_j=j+1}^{N-1} \alpha_{-,j}^{(k_j-1)} \mathbf{1}_{k_j=i} + \mathbf{1}_{i=j}}{\sqrt{\sum_{k_i=i+1}^{N-1} \alpha_{-,i}^{(k_i-1)} + 1} \sqrt{\sum_{k_j=j+1}^{N-1} \alpha_{-,j}^{(k_j-1)} + 1}} \\
\mathrm{Cov}(T_{+}^{(i)}, T_{-}^{(j)}) &= \frac{\sum_{k_i=i+1}^{N-1} \alpha_{+,i}^{(k_i-1)} \mathbf{1}_{k_i=j} + \sum_{k_j=j+1}^{N-1} \alpha_{-,j}^{(k_j-1)} \mathbf{1}_{k_j=i} - \sum_{k_i=i+1}^{N-1} \sum_{k_j=j+1}^{N-1} \alpha_{-,j}^{(k_j-1)} \alpha_{+,i}^{(k_i-1)} \mathbf{1}_{k_i=k_j} - \mathbf{1}_{i=j}}{\sqrt{\sum_{k_i=i+1}^{N-1} \alpha_{-,i}^{(k_i-1)} + 1} \sqrt{\sum_{k_j=j+1}^{N-1} \alpha_{-,j}^{(k_j-1)} + 1}}
\end{aligned}
$$

Then, $\left[\bigotimes_{k=1}^{N} \phi_k H\right](x)$ can be computed as the cumulative distribution function of Z:

$$
\left[\bigotimes_{k=1}^{N} \phi_k H\right](x) = \mathbb{P}\left[b_{-}^{(N-1)}(x) \leq U^{(N-1)} \leq b_{+}^{(N-1)}(x), \bigcap_{j=1}^{N-2} \left(T_{+}^{(j)} \leq \frac{\beta_{+,j}(x)}{\sum_{k=j+1}^{N-1} \alpha_{+,j}^{(k-1)} + 1}\right),\right.
$$

$$\bigcap_{j=1}^{N-2}\left(T_-^{(j)} \le -\frac{\beta_{-,j}(x)}{\sum_{k=j+1}^{N-1}\alpha_{-,j}^{(k-1)}+1}\right)\Bigg] \tag{6}$$

Proof. Let us proceed sequentially by showing how to compute $\left[\bigotimes_{k=1}^{N}\phi_k H\right](x)$ for $N=2$ and then $N=3$. The notations and the results leading to $N=4$ will ease the comprehension of the integral and probabilistic form of the convolution whatever the value of N is.

<u>N=2</u>:

The convolution $(\phi_1 H * \phi_2 H)(x)$ is easily obtained by the convolution between two Gaussian variables:

$$\left[\bigotimes_{k=1}^{2}\phi_k H\right](x) = \int_{-\infty}^{+\infty}\phi_1(x-s^{(1)})\phi_2(s^{(1)})H(x-s^{(1)})H(s^{(1)})ds^{(1)}$$

where $H(x-s^{(1)})H(s^{(1)}) = \begin{cases} 1 \text{ if } 0 < s^{(1)} < x \\ 0 \text{ else}\end{cases}$

$$\left[\bigotimes_{k=1}^{2}\phi_k H\right](x) \quad = \quad \int_0^x \phi_1(x-s^{(1)})\phi_2(s^{(1)})ds^{(1)} = \frac{1}{2\pi\sigma_1\sigma_2}\int_0^x e^{-\frac{(s^{(1)}-\mu_1)^2}{2\sigma_1^2}} e^{-\frac{(x-s^{(1)}-\mu_2)^2}{2\sigma_2^2}} ds^{(1)}$$

In what follows, $\mu^{(k)} = \sum_{i=1}^{k}\mu_i$ is the sum of the means, $\sigma^{(k)} = \sum_{i=1}^{k}\sigma_i$ is the sum of the standard deviations and $y^{(k)} = \mu^{(k)} - x$. Let $\phi^{(k)}(x) = \mathbb{P}[X^{(k)} \le x]$ denote the pdf of $X^{(k)} \sim \mathcal{N}(\mu^{(k)}, \sigma^{(k)})$. Last, two substitutions of the variables $r^{(k)}$ and $s^{(k)}$ are needed in the computation of the multiple convolution:

$$\begin{aligned} r^{(k)} &= s^{(k)} - \mu^{(k)} \\ u^{(k)} &= \frac{r^{(k)}\left[\sigma^{(k+1)}\right]^2 + y^{(k+1)}\left[\sigma^{(k)}\right]^2}{\sigma^{(k)}\sigma^{(k+1)}\sigma_{k+1}} \Rightarrow \frac{du^{(k)}}{dr^{(k)}} = \frac{\sigma^{(k+1)}}{\sigma^{(k)}\sigma_{k+1}} \Rightarrow dr^{(k)} = \sigma^{(k)}\sigma_{k+1}\frac{du^{(k)}}{\sigma^{(k+1)}} \end{aligned}$$

According to the previous notations, the convolution becomes (see the details in Appendix B):

$$\left[\bigotimes_{k=1}^{2}\phi_k H\right](x) = \frac{1}{\sqrt{2\pi}}\phi^2(x)\int_{b_-(x)}^{b_+(x)} e^{-\frac{\left[u^{(1)}\right]^2}{2}} du^{(1)} = \phi^2(x)\left[\Phi(b_+(x)) - \Phi(b_-(x))\right] \tag{7}$$

where the upper and lower limits of the integration, $b_-(x)$ and $b_+(x)$ are modified according to the two substitutions. They are now functions of $y^{(2)}$ (and thus x):

$$b_+ : x \mapsto \frac{x-\mu_1\left[\sigma^{(2)}\right]^2 + y^{(2)}\left[\sigma^{(1)}\right]^2}{\sigma^{(1)}\sigma^{(2)}\sigma_2}, \quad b_- : x \mapsto \frac{-\mu_1\left[\sigma^{(2)}\right]^2 + y^{(2)}\left[\sigma^{(1)}\right]^2}{\sigma^{(1)}\sigma^{(2)}\sigma_2}$$

This particular density mixture model can also be seen as the difference between two skew normal distributions $X_S(\alpha_+)$ and $X_S(\alpha_-)$ (O'Hagan and Leonard, 1976; Ashour and Abdel-Hamid, 2010; Mudholkar and Hutson, 2000) with α_+ and α_- their shape parameters:

$$\begin{aligned} \frac{1}{\sqrt{2\pi}}\phi^2(x)\int_{b_-(x)}^{b_+(x)} e^{-\frac{[u^{(1)}]^2}{2}}du^{(1)} &= \frac{1}{\sqrt{2\pi}}\phi^2(x)\left(\int_{-\infty}^{b_+(x)} e^{-\frac{[u^{(1)}]^2}{2}}du - \int_{-\infty}^{b_-(x)} e^{-\frac{[u^{(1)}]^2}{2}}du\right) \\ &= \frac{1}{2}(\phi_S(x,\alpha_+) - \phi_S(x,\alpha_-)) \end{aligned}$$

where $\phi_S(x,\alpha) = 2\phi(x)\Phi(\alpha x)$, ($\phi$ and Φ are respectively the density and the cumulative distribution function of a standard normal variable) denotes the probability density function of the skew normal distribution $X_S(\alpha)$. After rewriting $b_+(x)$ and $b_-(x)$, the shape parameters are defined by:

$$\alpha_+ = -\frac{(x-\mu_1)(\sigma_1^2+\sigma_2^2)}{y^{(2)}\sigma_1\sigma_2} - \frac{\sigma_1^2}{\sigma_1\sigma_2}, \quad \alpha_- = -\frac{\mu_1(\sigma_1^2+\sigma_2^2)}{y^{(2)}\sigma_1\sigma_2} - \frac{\sigma_1^2}{\sigma_1\sigma_2}$$

Note that α_+ and α_- are functions of x which implies that the two skew normal variables involved in the calculation differs for every value of x.

<u>N =3</u>:
The same calculations can be done for $N = 3$ (see the details in Appendix B). Starting from:

$$\left[\bigotimes_{k=1}^{2}\phi_k H\right](x) = \frac{1}{2\pi\sigma^{(2)}} e^{\frac{[y^{(2)}]^2}{2[\sigma^{(2)}]^2}} \int_{b_-(x)}^{b_+(x)} e^{-\frac{[u^{(1)}]^2}{2}} du^{(1)}$$

The convolution with the additional term $\phi_3 H(x)$ leads to:

$$\begin{aligned} \left[\bigotimes_{k=1}^{3}\phi_k H\right](x) &= \left[\left(\bigotimes_{k=1}^{2}\phi_k H\right) * \phi_3 H\right](x) \\ &= \frac{1}{2\pi}\phi^{(3)}(x)\int_{b_-^{(2)}(x)}^{b_+^{(2)}(x)} e^{-\frac{[u^{(2)}]^2}{2}} \left[\int_{b_-^{(1)}(s^{(2)})}^{b_+^{(1)}(s^{(2)})} e^{-\frac{[u^{(1)}]^2}{2}} du^{(1)}\right] du^{(2)} \end{aligned} \tag{8}$$

The lower and upper limits of the first integral, with $u^{(2)}$ used as variable of integration, are:

$$b_+^{(k)} : x \mapsto l^{(k)}(s^{(k)} - \mu^{(k)}) + y^{(k+1)}m^{(k)}, \quad b_-^{(k)} : x \mapsto -\mu^{(k)}l^{(k)} + y^{(k+1)}m^{(k)}$$

where $l^{(k)} = \frac{\sigma^{(k+1)}}{\sigma^{(k)}\sigma_{k+1}}$ and $m^{(k)} = \frac{\sigma^{(k)}}{\sigma^{(k+1)}\sigma_{k+1}}$.

The lower and upper limits $b_+^{(1)}(s^{(2)})$ and $b_-^{(1)}(s^{(2)})$ of the second integral with $u^{(1)}$ as variable of integration can be written by introducing successively the two substitutions $r^{(2)}$ and $u^{(2)}$.

$$\begin{aligned} b_+^{(1)}(s^{(2)}) &= b_+^{(1)}(r^{(2)}) + \mu^{(2)}(l^{(1)} - m^{(1)}) \\ &= \frac{b_+^{(1)}(u^{(2)}) - b_+^{(1)}(0)(1-l^{(2)}) - y^{(3)}m^{(2)}(l^{(1)} - m^{(1)})}{l^{(2)}} + \mu^{(2)}(l^{(1)} - m^{(1)}) \end{aligned}$$

$$\begin{aligned} b_-^{(1)}(s^{(2)}) &= b_-^{(1)}(r^{(2)}) + \mu^{(2)} m^{(1)} \\ &= \frac{b_-^{(1)}(u^{(2)}) - b_-^{(1)}(0)(1-l^{(2)}) + y^{(3)} m^{(2)} m^{(1)}}{l^{(2)}} + \mu^{(2)} m^{(1)} \end{aligned}$$

They can be rearranged as two affine functions of $u^{(2)}$:

$$b_+^{(1)}(s^{(2)}) = \alpha_+^{(1)} \cdot u^{(2)} + \beta_+^{(1)}$$

with $\alpha_+^{(1)} = \frac{l^{(1)} - m^{(1)}}{l^{(2)}}$ and $\beta_+^{(1)}(x) = b^{(1)}(0) + \frac{y^{(3)} m^{(2)} (m^{(1)} - l^{(1)})}{l^{(2)}} + \mu^{(2)}(l^{(1)} - m^{(1)})$

$$b_-^{(1)}(s^{(2)}) = \alpha_-^{(1)} \cdot u^{(2)} + \beta_-^{(1)}$$

with $\alpha_-^{(1)} = \frac{-m^{(1)}}{l^{(2)}}$ and $\beta_-^{(1)}(x) = b^{(1)}(0) + \frac{y^{(3)} m^{(1)} m^{(2)}}{l^{(2)}} - m^{(1)} \mu^{(2)}$.

Doing so, this integral can be translated in probabilistic terms. Let f denote the function that maps x to $e^{-\frac{x^2}{2}}$, then:

$$\begin{aligned} & \frac{1}{2\pi} \int_{b_-^{(2)}(x)}^{b_+^{(2)}(x)} f(u^{(2)}) \int_{b_-^{(1)}(s)}^{b_+^{(1)}(s)} f(u) du du^{(2)} \\ = & \int_{b_-^{(2)}(x)}^{b_+^{(2)}(x)} \frac{1}{\sqrt{2\pi}} f(u^{(2)}) \int_{\alpha_-^{(1)}(u^{(2)}) + \beta_-^{(1)}}^{\alpha_+^{(1)}(u^{(2)}) + \beta_+^{(1)}} \frac{1}{\sqrt{2\pi}} f(u) du du^{(2)} \\ = & \int_{-\infty}^{+\infty} \int_{-\infty}^{-\infty} 1_{b_-^{(2)}(x) \le U^{(2)} \le b_+^{(2)}(x)} 1_{\alpha_-^{(1)} U^{(2)} - \beta_-^{(1)} \le U \le \alpha_+^{(1)} U^{(2)} - \beta_+^{(1)}} \frac{1}{\sqrt{2\pi}} f(u^{(2)}) \frac{1}{\sqrt{2\pi}} f(u) du du^{(2)} \\ = & \mathbb{P}\left[b_-^{(2)}(x) \le U^{(2)} \le b_+^{(2)}(x), U - \alpha_+^{(1)} U^{(2)} \le \beta_+^{(1)}, U - \alpha_-^{(1)} U^{(2)} > \beta_-^{(1)} \right] \\ = & \mathbb{P}\left[b_-^{(2)}(x) \le U^{(2)} \le b_+^{(2)}(x), T_+^{(1)} < \frac{\beta_+^{(1)}}{\sqrt{(\alpha_+^{(1)})^2 + 1}}, T_-^{(1)} < -\frac{\beta_-^{(1)}}{\sqrt{(\alpha_-^{(1)})^2 + 1}} \right] \end{aligned}$$

where $(U^{(2)}, T_+^{(1)}, T_-^{(1)})$ is a Gaussian vector with:

$$U^{(2)} \sim \mathcal{N}(0,1), \quad T_+^{(1)} = \frac{U - \alpha_+^{(1)} U^{(2)}}{\sqrt{(\alpha_+^{(1)})^2 + 1}} \sim \mathcal{N}(0,1), \quad T_-^{(1)} = \frac{\alpha_-^{(1)} U^{(2)} - U}{\sqrt{(\alpha_-^{(1)})^2 + 1}} \sim \mathcal{N}(0,1)$$

Its covariance matrix is:

$$\Sigma_3 = \begin{bmatrix} 1 & \mathrm{Cov}(U^{(2)}, T_+^{(1)}) & \mathrm{Cov}(U^{(2)}, T_-^{(1)}) \\ \mathrm{Cov}(U^{(2)}, T_+^{(1)}) & 1 & \mathrm{Cov}(T_+^{(1)}, T_-^{(1)}) \\ \mathrm{Cov}(U^{(2)}, T_-^{(1)}) & \mathrm{Cov}(T_+^{(1)}, T_-^{(1)}) & 1 \end{bmatrix}$$

with

$$\mathrm{Cov}(U^{(2)}, T_+^{(1)}) = \mathrm{Cov}(U^{(2)}, \frac{U - \alpha_+^{(1)} U^{(2)}}{\sqrt{(\alpha_+^{(1)})^2 + 1}}) = -\frac{\alpha_+^{(1)}}{\sqrt{(\alpha_+^{(1)})^2 + 1}} \sigma_{U^{(2)}}^2 = -\frac{\alpha_+^{(1)}}{\sqrt{(\alpha_+^{(1)})^2 + 1}}$$

$$\begin{aligned}
\mathrm{Cov}(U^{(2)},T_-^{(1)}) &= \mathrm{Cov}(U^{(2)},\frac{\alpha_-^{(1)}U^{(2)}-U}{\sqrt{(\alpha_-^{(1)})^2+1}}) = \frac{\alpha_-^{(1)}}{\sqrt{(\alpha_-^{(1)})^2+1}}\sigma^2_{U^{(2)}} = \frac{\alpha_-^{(1)}}{\sqrt{(\alpha_-^{(1)})^2+1}} \\
\mathrm{Cov}(T_+^{(1)},T_-^{(1)}) &= \mathrm{Cov}(\frac{U-\alpha_+^{(1)}U^{(2)}}{\sqrt{(\alpha_+^{(1)})^2+1}},\frac{\alpha_-^{(1)}U^{(2)}-U}{\sqrt{(\alpha_-^{(1)})^2+1}}) = -\frac{\alpha_+^{(1)}\alpha_-^{(1)}+1}{\sqrt{(\alpha_+^{(1)})^2+1}\sqrt{(\alpha_-^{(1)})^2+1}}
\end{aligned}$$

$N=4$ and so on:

For $N=4$, the same process is used but the affine terms $\beta_+^{(1)}$ and $\beta_-^{(1)}$ are functions of x. By convolving, x becomes $s^{(k)}$ and goes through two substitutions $r^{(k)}$ and $u^{(k)}$. For $\beta_+^{(1)}(x)$, :

$$\begin{aligned}
& \beta_+^{(1)}(x) \\
= & \beta_+^{(1)}(s^{(3)}) \\
= & \beta_+^{(1)}\left(\frac{u^{(3)}-y^{(4)}m^{(3)}}{l^{(3)}}-\mu^{(3)}\right) \\
= & b_-^{(1)}(0)+\frac{\left[\mu^{(3)}-\left(\frac{u^{(3)}-y^{(4)}m^{(3)}}{l^{(3)}}+\mu^{(3)}\right)\right]m^{(2)}(m^{(1)}-l^{(1)})}{l^{(2)}}+\mu^{(2)}(l^{(1)}-m^{(1)}) \\
= & \left(\frac{u^{(3)}}{l^{(3)}l^{(2)}}m^{(2)}(l^{(1)}-m^{(1)})\right)+\left(b^{(1)}(0)+\frac{y^{(4)}m^{(3)}m^{(2)}(m^{(1)}-l^{(1)})}{l^{(3)}l^{(2)}}+\mu^{(2)}(l^{(1)}-m^{(1)})\right)
\end{aligned}$$

Finally,

$$b_+^{(1)}(x)=\alpha_+^{(1)}\cdot u^{(2)}+\beta_+^{(1)}(x)=\alpha_+^{(1)}(u^{(2)})+\alpha_+^{(2)}\cdot u^{(3)}+\beta_+^{(2)}(x)$$

For any value of N, the recursive process that sets the value for upper and lower limits is:

$$b_+^{(j)}(x)=\left(\sum_{k=j+1}^{N-1}\alpha_{+,j}^{(k-1)}\cdot u^{(k)}\right)+\beta_{+,j}(x),\quad b_-^{(j)}(x)=\left(\sum_{k=j+1}^{N-1}\alpha_{-,j}^{(k-1)}\cdot u^{(k)}\right)+\beta_{-,j}(x)$$

where:

$$\begin{aligned}
\alpha_{+,j}^{(k)} &= \frac{\prod_{i=j+1}^{k}m^{(i)}(l^{(j)}-m^{(j)})}{\prod_{i=j+1}^{k+1}l^{(i)}},\quad \alpha_{-,j}^{(k)}=-m^{(j)}\frac{\prod_{i=j+1}^{k}m^{(i)}}{\prod_{i=j+1}^{k+1}l^{(i)}} \\
\beta_{+,j}(x) &= \left(b_-^{(1)}(0)-m^{(1)}\mu^{(2)}+m^{(j)}\frac{y^{(N)}\prod_{i=j+1}^{N-1}m^{(i)}}{\prod_{i=j+1}^{N-1}l^{(i)}}\right) \\
\beta_{-,j}(x) &= \left(b_-^{(1)}(0)+\mu^{(2)}(l^{(1)}-m^{(1)})+\frac{y^{(N)}\prod_{i=j+1}^{N-1}m^{(i)}(m^{(j)}-l^{(j)})}{\prod_{i=j+1}^{N-1}l^{(i)}}\right)
\end{aligned}$$

The multiple convolution formula can thus be generalized for any generic value of N with Eq. (5) and Eq. (6) of Theorem 1 that respectively provides the integral and the probabilistic form of $\left[\bigotimes_{k=1}^{N}\phi_k H\right](x)$. □

2.3. Cumulative density function of $\sum_i X_i^+$

The computation of the cumulative distribution function (cdf) is even more complex than the pdf because of the additional integration along x. However, since the formula of the exact pdf is known from Theorem 1, an efficient way of approximating the cdf is to use a Simpson approximation of the integral:

$$\int_0^x \phi_{\sum_i X_i^+}(x)dx \approx \int_0^x P(x)dx = \frac{x}{6n}\sum_{i=0}^{N-1}\left[\phi_{\sum_i X_i^+}(x_i) + 4\phi_{\sum_i X_i^+}\left(\frac{x_i+x_{i+1}}{2}\right) + \phi_{\sum_i X_i^+(x_{i+1})}\right] \quad (9)$$

where $P(x)$ is the quadratic polynomial which takes the same values as $\phi_{\sum_i X_i^+}(x)$ at the end points x_i and x_{i+1}, and the midpoint $\frac{x_{i+1}-x_i}{2}$.

2.4. Asymptotic normal approximation

When N increases, the computational time for Eq. (4) with multiple convolution computations given by Eq. (6) increases and becomes too large (Figure 1).

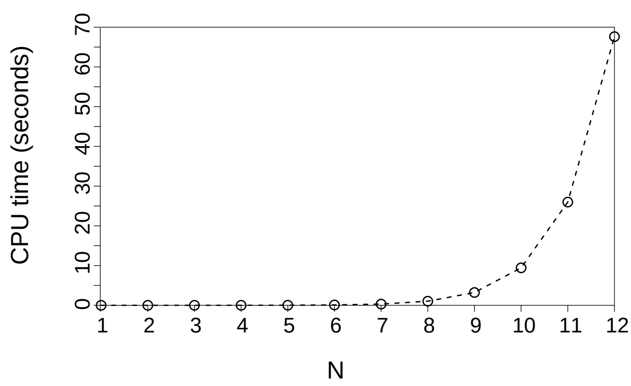


FIGURE 1: *CPU time for computing the probability density function of N rectified Gaussian variables* $\{X_i^+\}_{i=1,\cdots,N}$, $(N=12)$ *using the R-codes provided in Section 3*

However in this case, the Lyapunov or Lindeberg version (Billingsley, 1995) of the Central Limit Theorem (CLT) can be applied. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and $X_i^+ : \Omega \to \mathbb{R},\ i \in \mathbb{N}$, be independent rectified Gaussian variables defined on that space with expected values and variances that exists and are, by construction from their related original Gaussian variable, finite. Also let s_N^2 denote the sum of their variances:

$$s_N^2 = \sum_{i=1}^{N} \mathrm{Var}[X_i^+] \quad (10)$$

Then, the Lyapunov condition:

$$\lim_{N\to\infty} \frac{1}{s_N^{2+\delta}} \sum_{i=1}^{N} \mathbb{E}\left[|X_i^+ - \mathbb{E}[X_i^+]|^{2+\delta}\right] = 0 \quad (11)$$

is satisfied under specific but very common conditions. Indeed, it can always be written that:

$$|X_i^+ - \mathbb{E}[X_i^+]| < X_i^+ + \mathbb{E}[X_i^+]$$

Let denote $\mu_i^+ = \mathbb{E}[X_i^+]$, then the third moment for the variable $X_i^+ + \mu_i^+$ is defined by:

$$\mathbb{E}\left[(X_i^+ + \mu_i^+)^3\right] = \int_{\mu_i^+}^{\infty} x^3 e^{-\frac{1}{2}(\frac{x-\mu_i}{\sigma_i})^2} dx$$

For simplifications, $\mu_i^+ - \mu_i > 0$ is denoted as $\Delta_{\mu i}$ and the substitution $u = \frac{x-\mu_i}{\sigma_i}$ is now introduced. Let denote:

$$I_k(a) = \int_a^{\infty} u^k e^{-\frac{u^2}{2}} du \tag{12}$$

Then, developing the expression and using the notation I_k leads to:

$$\begin{aligned}
& \int_{\mu_i^+}^{\infty} x^3 e^{-\frac{1}{2}(\frac{x-\mu_i}{\sigma_i})^2} dx \\
= & \int_{\Delta_{\mu i}}^{\infty} (u\sigma_i + \mu i)^3 e^{-\frac{u^2}{2}} dx \\
= & \int_{\Delta_{\mu i}}^{\infty} (u^3\sigma_i^3 + 3u^2\sigma_i^2\mu_i + 3u\sigma_i\mu_i^2 + \mu_i^3) e^{-\frac{u^2}{2}} du \\
= & \sigma_i^3 \int_{\Delta_{\mu i}}^{\infty} u^3 e^{-\frac{u^2}{2}} du + 3\sigma_i^2\mu_i \int_{\Delta_{\mu i}}^{\infty} u^2 e^{-\frac{u^2}{2}} du + 3\sigma_i\mu_i^2 \int_{\Delta_{\mu i}}^{\infty} u e^{-\frac{u^2}{2}} du + \mu_i^3 \int_{\Delta_{\mu i}}^{\infty} e^{-\frac{u^2}{2}} du \\
= & \sigma_i^3 I_3(\Delta_{\mu i}) + 3\sigma_i^2\mu_i I_2(\Delta_{\mu i}) + 3\sigma_i\mu_i^2 I_1(\Delta_{\mu i}) + \mu_i^3 I_0(\Delta_{\mu i})
\end{aligned} \tag{13}$$

According to the values of $I_k(\Delta_{\mu i})$ given in Appendix A for $k = 0, \cdots, 3$, see Eqs. (21a), (21b), (21c) and (21d), Eq. (13) can be upper-bounded:

$$\begin{aligned}
\int_{\mu_i^+}^{\infty} x^3 e^{-\frac{1}{2}(\frac{x-\mu_i}{\sigma_i})^2} dx & \leq \sigma_i^3(\Delta_{\mu i}^2 + 2) + 3\sigma_i^2\mu_i(\Delta_{\mu i} + 1) + 3\sigma_i\mu_i^2 + \mu_i^3 \\
& \leq (\sigma_i + \mu_i)^3 + \sigma_i^3(\Delta_{\mu i}^2 + 1) + 3\sigma_i^2\mu_i\Delta_{\mu i}
\end{aligned} \tag{14}$$

In the most common case where N is sufficiently large and there is no variable X_i^+ predominant over the others, meaning that $\forall i,\ (\sigma_i^+)^2 << s_N^2$, it is clear that Eq. (14) implies that:

$$\frac{\mathbb{E}\left[(X_i^+ + \mu_i^+)^3\right]}{s_N^2} \longrightarrow 0 \Longrightarrow \frac{1}{s_N} \sum_{i=1}^{N} \frac{\mathbb{E}\left[(X_i^+ + \mu_i^+)^3\right]}{s_N^2} \underset{N \to +\infty}{\longrightarrow} 0 \tag{15}$$

and because the Lyapunov condition is satisfied for $\delta = 1$ in this very common assumption that the Gaussian variables X_i, even if not identically distributed, have expectation and variance ranging in similar order of magnitudes, the Central Limit Theorem holds, meaning that:

$$\mathscr{L}\left(\sum_{i=1}^{N} X_i^+\right) \xrightarrow{\mathscr{L}} \mathscr{N}\left(\sum_{i=1}^{n} \mathbb{E}[X_i^+], s_N^2\right) \tag{16}$$

where $\mathscr{L}$ stands for the distribution of the variable.
Last, the Lyapunov condition being satisfied, the so-called Lindeberg condition is also ensured

$$\lim_{N\to\infty} \frac{1}{s_N^2} \sum_{i=1}^{N} \mathbb{E}\left[(X_i^+ - \mathbb{E}[X_i^+])^2 \cdot \mathbf{1}_{\{|X_i^+ - \mathbb{E}[X_i^+]| > \varepsilon s_N\}} \right] = 0 \tag{17}$$

for all $\varepsilon > 0$, where $\mathbf{1}$ is the indicator function. It implies that:

$$\max_{i=1,\cdots,N} \frac{\sigma_i^2}{s_N^2} \underset{N\to\infty}{\longrightarrow} 0 \tag{18}$$

which is in fact the intuitive condition required to satisfy the Lyapunov condition (which is here a specific case because the inverse implication, meaning that Eq. (18) would imply the Lindeberg condition, does not hold in general).

A simple counterexample is given to show that in the case of a predominant variable X_i^+, the sum is no longer Gaussian. Let assume that:

$$Y \sim \mathscr{N}\left(\sum_{i=1}^{N} \mathbb{E}\left[X_i^+\right], s_N^2 \right) \tag{19}$$

is the normal approximation of $\sum_{i=1}^{N} X_i^+$ with N large enough and no predominant variables so the the Lyapunov-CLT applies. Now, let $X_{N+1}^+ \sim \mathscr{N}^+(\mu_{N+1}, \sigma_{N+1}^2)$ be a new variable added to the previous sum Y . If $s_N^2 << \sigma_{N+1}^2$, then the resulting distribution of $Y + X_{N+1} = \sum_{i=1}^{N+1} X_i^+$ is clearly non-Gaussian, for instance if the parameters $\sum_{i=1}^{N} \mathbb{E}\left[X_i^+\right]$ and s_N^2 are respectively set to 0.1 and 0.01 and parameters μ_{N+1} and σ_{N+1} are set to 0.1 and 4, $Y + X_{N+1}$ will basically looks like a rectified Gaussian with a Dirac at 0 due to the high variance σ_{N+1}^2.

3. Practical implementation with R

The R-codes (R Development Core Team, 2008) referenced in the Listings 1, 2 and 3 respectively provide the R-version of the algorithmic codes that are necessary to compute the implementation of the probabilistic form of the multiple convolution $\left[\bigotimes_k \phi_k H\right](x)$, see Eq. (6), the pdf of $\sum_i X_i^+$, see Eq. (4), and the approximation of its cdf, see Eq. (9).
This code can be easily optimized by implementing the same functions in a compiled language (Fortran or C for instance). But because the computation of Eq. (6) is needed, the R-library **mvtnorm** is involved (Genz et al., 2017; Genz and Bretz, 2009) and more specifically the function **pmvnorm** that enables the computation of the distribution function of the multivariate normal distribution for arbitrary limits, given its correlation matrix. Thus, it would be even easier to improve the R-code by using the already existing interface between R and C++ towards the R library Rcpp (Eddelbuettel and François, 2011; Eddelbuettel, 2013).

LISTING 1: *Implementation in R of the convolution* $[\bigotimes_k \phi_k H](x)$

```
library(mvtnorm)
convol_NR<-function(x,M,S){
  N=length(M)
  if(N>1){
    A=numeric(N-1)
    B=numeric(N-1)
    for(k in 1:(N-1)) {
      A[k]=sqrt(sum((S[1:(k+1)])^2))/(sqrt(sum((S[1:k])^2))*S[k+1])
      B[k]=sqrt(sum((S[1:k])^2))/(sqrt(sum((S[1:(k+1)])^2))*S[k+1])
    }
    b_s<-function(x,k,A,B,M){return(A[k]*(x-sum(M[1:k]))+(sum(M[1:(k+1)])-x)*B[k])
       }
    b_i<-function(x,k,A,B,M){return(A[k]*(-sum(M[1:k]))+(sum(M[1:(k+1)])-x)*B[k])}
    if(N>2){
      # computation of alpha's and beta's
      alpha_s=matrix(NA,nrow=N-2,ncol=N-2)
      alpha_i=matrix(NA,nrow=N-2,ncol=N-2)
      beta_s=numeric(N-2)
      beta_i=numeric(N-2)
      for(j in 1:(N-2)) {
        for(k in j:(N-2)){
        if((j+1)>k){
            alpha_s[j,k]=(A[j]-B[j])/prod(A[(j+1):(k+1)])
          alpha_i[j,k]=-B[j]/prod(A[(j+1):(k+1)])
        }
        if((j+1)<=k){
            alpha_s[j,k]=(A[j]-B[j])*(prod(B[(j+1):(k)])/prod(A[(j+1):(k+1)]))
          alpha_i[j,k]=(-B[j])*(prod(B[(j+1):(k)])/prod(A[(j+1):(k+1)]))
        }
      }
      beta_s[j]=b_s(0,j,A,B,M) + (sum(M[1:(j+1)])*(A[j]-B[j])) + (B[j]-A[j])*(sum(
          M[1:N])-x)*(prod(B[(j+1):(N-1)])/prod(A[(j+1):(N-1)]))
      beta_i[j]=b_i(0,j,A,B,M) + (sum(M[1:(j+1)])*(-B[j])) + (B[j])*(sum(M[1:N])-x
          )*(prod(B[(j+1):(N-1)])/prod(A[(j+1):(N-1)]))
      }
      # lower and upper limits ; covariances terms
      upper=numeric(2*(N-2)+1)
      upper[1]=b_s(x,N-1,A,B,M)
      lower=rep(-Inf,2*(N-2)+1)
      lower[1]=b_i(x,N-1,A,B,M)
      sigma=diag(1,nrow=2*(N-2)+1,ncol=2*(N-2)+1)
      # filling the sigma matrix for the first and last column
      for(j in 1:(2*(N-2))){
        if(j%%2 !=0) {
        sigma[1,j+1]=-alpha_s[floor(j/2)+1,N-2]/sqrt(sum(alpha_s[floor(j/2)+1,(
            floor(j/2)+1):(N-2)]^2)+1)
        sigma[j+1,1]=sigma[1,j+1]
      }
      if(j%%2 ==0) {
        sigma[1,j+1]=alpha_i[(j/2),N-2]/sqrt(sum(alpha_i[(j/2),((j/2):(N-2))]^2)
            +1)
        sigma[j+1,1]=sigma[1,j+1]
      }
      }
      # filling the sigma matrix
      for(j in 1:(2*(N-2))){
        # upper limits (lower limits=-inf)
      if(j%%2 !=0) {
        upper[j+1]=beta_s[floor(j/2)+1]/sqrt(sum(alpha_s[floor(j/2)+1,(floor(j/2)
            +1):(N-2)]^2)+1)
        }
```

```
     if(j%%2 ==0) {
      upper[j+1]=-beta_i[(j/2)]/sqrt(sum(alpha_i[(j/2),(j/2):(N-2)]^2)+1)
    }
    # covariances
     for(i in 1:(2*(N-2))){
      # T+ et T+
     if(j%%2 !=0 & i%%2!=0) {
      res=0
      for(k_i in (floor(i/2)+1):(N-2)){
       for(k_j in (floor(j/2)+1):(N-2)){
        res=res+(ifelse(k_i==k_j,1,0)*alpha_s[floor(i/2)+1,k_i]*alpha_s[(
          floor(j/2)+1),k_j])
       }
      }
      for(k_i in (floor(i/2)+1):(N-2)){
       res=res-(ifelse(k_i+1==floor(j/2)+1,1,0)*alpha_s[floor(i/2)+1,k_i])
      }
      for(k_j in (floor(j/2)+1):(N-2)){
       res=res-(ifelse(k_j+1==floor(i/2)+1,1,0)*alpha_s[(floor(j/2)+1),k_j])
      }
      if(i==j){res=res+1}
      sigma[j+1,i+1]=res/(sqrt(sum(alpha_s[floor(i/2)+1,(floor(i/2)+1):(N-2)
        ]^2)+1)*sqrt(sum(alpha_s[(floor(j/2)+1),(floor(j/2)+1):(N-2)]^2)+1))
     }
      # T+ et T-
     if(j%%2 !=0 & i%%2==0) {
      res=0
      for(k_i in (i/2):(N-2)){
       for(k_j in (floor(j/2)+1):(N-2)){
       res=res-(ifelse(k_i==k_j,1,0)*alpha_i[(i/2),k_i]*alpha_s[(floor(j/2)
         +1),k_j])
       }
      }
      for(k_i in (i/2):(N-2)){
       res=res+(ifelse(k_i+1==floor(j/2)+1,1,0)*alpha_i[(i/2),k_i])
      }
      for(k_j in (floor(j/2)+1):(N-2)){
       res=res+(ifelse(k_j+1==(i/2),1,0)*alpha_s[(floor(j/2)+1),k_j])
      }
      if(i==(j+1)){res=res-1}
      sigma[j+1,i+1]=res/(sqrt(sum(alpha_i[(i/2),(i/2):(N-2)]^2)+1)*sqrt(sum(
        alpha_s[(floor(j/2)+1),(floor(j/2)+1):(N-2)]^2)+1))
     }
     # T- et T+
     if(j%%2 ==0 & i%%2!=0) {
      res=0
      for(k_i in (floor(i/2)+1):(N-2)){
       for(k_j in (j/2):(N-2)){
       res=res-(ifelse(k_i==k_j,1,0)*alpha_s[floor(i/2)+1,k_i]*alpha_i[(j/2),
         k_j])
       }
      }
      for(k_i in (floor(i/2)+1):(N-2)){
       res=res+(ifelse(k_i+1==(j/2),1,0)*alpha_s[floor(i/2)+1,k_i])
      }
       for(k_j in (j/2):(N-2)){
       res=res+(ifelse(k_j+1==floor(i/2)+1,1,0)*alpha_i[(j/2),k_j])
      }
      if(i==(j-1)){res=res-1}
      sigma[j+1,i+1]=res/(sqrt(sum(alpha_s[floor(i/2)+1,(floor(i/2)+1):(N-2)
        ]^2)+1)*sqrt(sum(alpha_i[(j/2),(j/2):(N-2)]^2)+1))
     }
     # T- et T-
```

```
        if(j%%2 ==0 & i%%2==0) {
          res=0
          for(k_i in (i/2):(N-2)){
            for(k_j in (j/2):(N-2)){
            res=res+(ifelse(k_i==k_j,1,0)*alpha_i[(i/2),k_i]*alpha_i[(j/2),k_j])
              }
          }
            for(k_i in (i/2):(N-2)){
              res=res-(ifelse(k_i+1==(j/2),1,0)*alpha_i[(i/2),k_i])
          }
          for(k_j in (j/2):(N-2)){
              res=res-(ifelse(k_j+1==(i/2),1,0)*alpha_i[(j/2),k_j])
            }
          if(i==j){res=res+1}
          sigma[j+1,i+1]=res/(sqrt(sum(alpha_i[(i/2),(i/2):(N-2)]^2)+1)*sqrt(sum(
              alpha_i[(j/2),(j/2):(N-2)]^2)+1))
        }
        }
      }
      val=dnorm(x,sum(M),sqrt(sum(S^2)))*pmvnorm(lower=lower,upper=upper,rep(0,2*(
          N-2)+1),sigma=sigma)
    }
    else{
      val=dnorm(x,sum(M),sqrt(sum(S^2)))*(pnorm(b_s(x,N-1,A,B,M),0,1)-pnorm(b_i(x,
          N-1,A,B,M),0,1))
    }
  }
  else{
    val=dnorm(x,M,S)
  }
  return(val)
}
```

LISTING 2: *Implementation in R of the probability density function* $\phi_{\sum_i X_i^+}(x)$

```
PDF<-function(x,M,S){
  N=length(M)
  val=rep(0,(2^N)-1)
  k=1
  for(i in 1:N){
    tab=combn(N,i)
    if(ncol(tab)>1){
      for(j in 1:ncol(tab)){
        val[k]=prod(sapply(-1*(M[-tab[,j]]/S[-tab[,j]]),FUN=pnorm,0,1))*convol_NR(
            x,M[tab[,j]],S[tab[,j]])
        k=k+1
      }
    }
    else{
      val[k]=convol_NR(x,M[tab[,1]],S[tab[,1]])
    }
  }
  return(sum(val))
}
```

LISTING 3: *Implementation in R of the cumulative density function* $\Phi_{\sum_i X_i^+}(x)$

```
CDF<-function(x,M,S){
  I=0
  Xn=seq(1e-7,x,length.out=101)
  for(i in 1:100){
    I=I+(PDF(Xn[i],M,S)+4*PDF((Xn[i]+Xn[i+1])/2,M,S)+PDF(Xn[i+1],M,S))
  }
  I=(x/(6*100))*I
}
```

4. Numerical computations

4.1. *Large N*

A comparison is made between the normal approximation of the sum:

$$\sum_{i=1}^{N} X_i^+ \sim \mathcal{N}\left(\sum_{i=1}^{N} \mathbb{E}\left[X_i^+\right], s_N^2\right) \tag{20}$$

and a Monte-Carlo experiment in which $1e6$ realizations of $N = 10,\ 50,\ 100$ and 1000 independent Gaussian variables with mean and standard deviations uniformly sampled in $[-3,3]$ and $[0,10]$. The simulation of a specific rectified Gaussian distribution $X^+ \sim \mathcal{N}^+(\mu,\sigma^2)$ from the original Gaussian distribution $X \sim \mathcal{N}(\mu,\sigma^2)$ is easily obtained by Eq. (1).
Figure 2 shows the Monte-Carlo resulting histograms (gray bars) and the pdf of the normal approximation (dashed red lines).

It is clear that when N increases, the skewness of the sum tends to 0, making the normal approximation valid. Even when the rectified Gaussian variables are badly shaped, for instance heavily tailed, meaning the probability mass of the variables in 0 took most part of the pdf, which is the case if μ_i is weakly negative and σ_i^2 is small enough so that $\mathbb{P}\left[X_i^+ > 0\right] < \varepsilon$, the normal approximation behaves well when $N \to \infty$.

Let denote $\widetilde{\phi}_{\sum_i X_i^+}(x)$ the probability density function of the asymptotic normal approximation defined in Eq. (16). Figure 3 shows the boxplots of $\max_x |\phi(x) - \widetilde{\phi}(x)|$, the maximum absolute error between the asymptotic normal approximation and the true pdf according to the number N of rectified Gaussian variables, built according to the same sampling scheme already used for Figure 2.

As expected, the error decreases when N becomes larger. But depending on the means and variances of the N variables $\{X_i^+\}_{i=1,\cdots,N}$, the value of N for which the error becomes negligible can not be precisely known. According to the trend followed by the mean of $\max_x |\phi(x) - \widetilde{\phi}(x)|$ and its very small variance from $N = 8$, it seems reasonable to consider the approximation made by the CLT valid for a dozen of rectified Gaussian variables. Basically, Figure 2b shows that for $N = 50$, except of a very small value of the skewness, the approximation is fully valid. It is in fact already the case for $N = 10$, see Figure 2a, when rectifying to 0 the negative values of the normal approximation (because the sum of N rectified Gaussian variables is always positive). If not completely satisfied by the normal approximation for the sum of a few dozen of variables, the computational work of McConalogue (1981), on the numerical treatment of convolution integrals involving distributions with densities having singularities at the origin, could be useful because it treats exactly the case of a rectified Gaussian distribution.

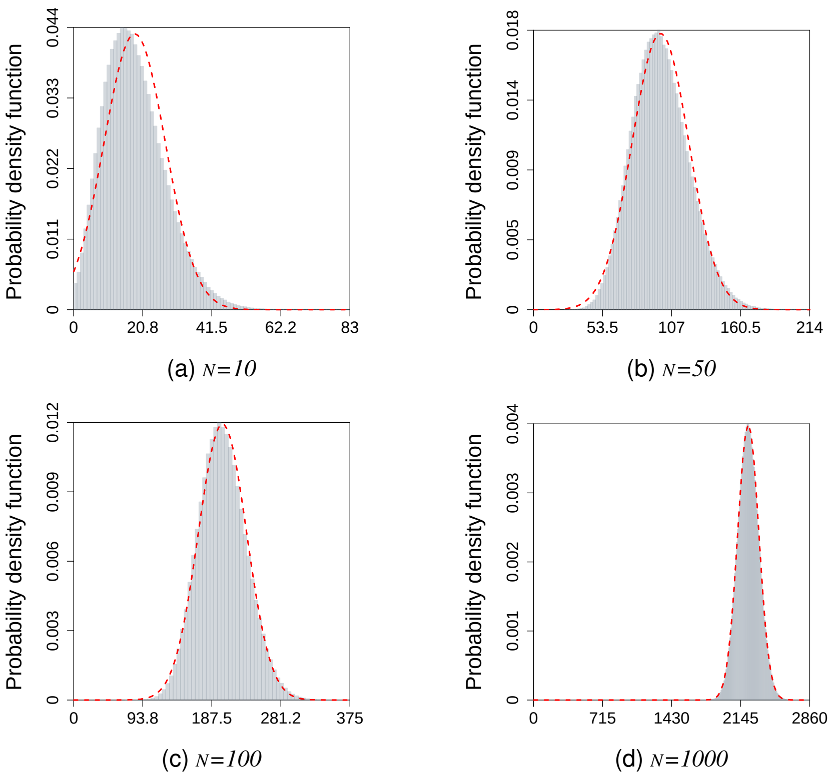


FIGURE 2: *Monte-Carlo histograms and probability density functions of the normal approximations. Examples for the sum of 10, 50, 100 and 1000 independent rectified Gaussian variables with mean and standard deviations of their related Gaussian variables uniformly sampled in* $[-3,3]$ *and* $[0,10]$

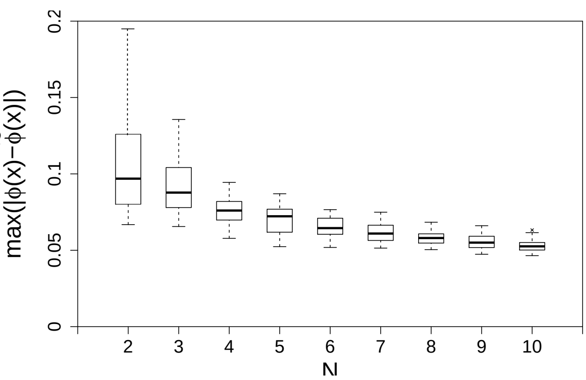
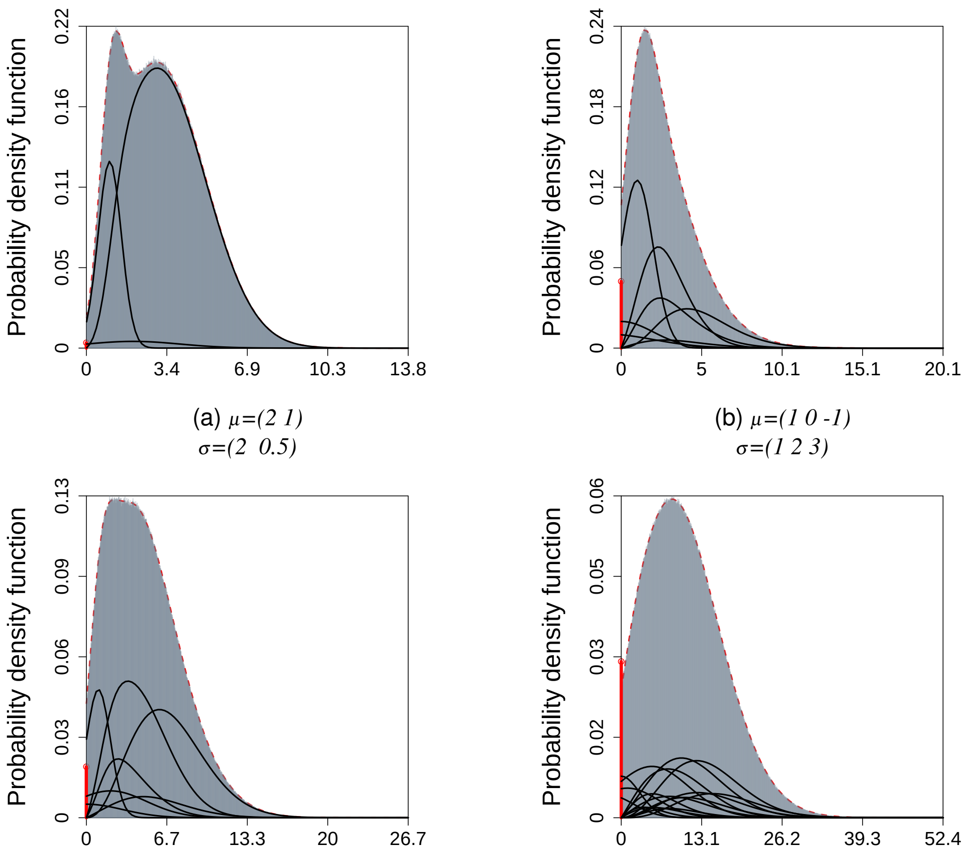


FIGURE 3: *Boxplots of* $\max_x |\phi(x) - \widetilde{\phi}(x)|$*, the maximum absolute error between the asymptotic normal approximation* $\widetilde{\phi}$ *and the true pdf* ϕ

4.2. Small N

Once again, a comparison is made with a Monte-Carlo approach: $1e6$ realizations of N independent Gaussian variables with different means and variances are simulated and N is small enough so that the CLT does not apply. Figure 4 shows both the histograms empirically obtained by Monte-Carlo simulations and the theoretical probability density functions (red dashed line) computed from Eq. (4) and Eq. (6) of the Main results section. The black solid lines are the terms $\Phi(-\frac{\mu_j}{\sigma_j})\left[\bigotimes_{k\notin S_{il}} \phi_k H\right](x)$, $i = 1,\cdots,N$, $j \in S_{il}$, from the analytical form of the pdf, see Eq. (4).



(a) μ=(2 1) σ=(2 0.5)

(b) μ=(1 0 -1) σ=(1 2 3)

(c) μ=(2 0 1) σ=(3 3 1)

(d) μ=(0 -2 1 5) σ=(3 4 5 6)

FIGURE 4: *Monte-Carlo histograms and exact probability density functions. Examples for the sum of 2, 3 and 4 rectified and independent Gaussian variables*

The results provided by Eq. (6) fit perfectly to the Monte-Carlo simulations. The first class of the histogram can deviate a bit from the true pdf $\phi_{\sum_i X_i^+}(0)$ (red vertical line symbolizing the Dirac measure) because it integrates values of x close to 0 while the true pdf is only representative of the discrete component of the variable in 0.

5. Application: the case of batch filters PAH's measurements

PAHs (polycyclic aromatic hydrocarbons) are the chemical organic compounds with a ring structure including at least two aromatic rings. From the hundred PAHs listed, sixteen are commonly analyzed in the different components of the environment, including the air. Many PAHs in the atmosphere exist simultaneously as gaseous and particulate.
Benzo[a]pyrene (B[a]P), one of the most known PAHs, is defined as an agent that causes cancer to human kind (Group 1) by IARC (International Agency for Research on Cancer). Other PAHs are also defined by IARC as carcinogenic probable agents (Group 2A) or carcinogenic potential (group 2B). B[a]P is regarded as a marker of risk for the family of the PAHs causing cancer. PAHs are regulated by Directive 2004/107/EC of 15 December 2004. In particular, the target value assigned for the B[a]P annual mean concentration is fixed at 1 $ng.m^3$.

PAHs measurements are collected by filters. The number of data available by monitoring sites thus depends on the time coverage and how filters are analyzed at the laboratory. The minimum time coverage set by the Directive 2008/50/EC (2008) is 33% for fixed measurement - which is mandatory in areas where the concentration of B[a]P is greater than the minimum assessment threshold - and 14% for the indicative measurement.

In order to meet these quality objectives, the LCSQA (the French organism in charge of monitoring the air quality) recommends the implementation of a daily sampling every 3 days (fixed measurements) or every 6 days (indicative measurement) (Albinet, 2011). This sampling strategy was adopted rather widely by the AASQA's (French local agencies in charge of monitoring the air quality in French regions). The main part of the AASQA's (21 of 26) analyses each collected sample and compute the annual mean accordingly. For financial reasons or when PAH's concentration levels are particularly low (summer period or sampling low-speed-based approach), it is allowed to group the filters over a maximum period of one month and make what is called a "grouped analysis" (Albinet, 2011). The result of the analysis is thus an averaged concentration by batch of filters. Five AASQA's are working that way, either for part or all of their PAHs monitoring sites.

When grouped analysis are involved, the number of available data on the year can be very small (between 4 and 12 according to the time sampling strategy). If the values are low (and sometimes lower than the limit of quantification LQ of the laboratory), the value is set at LQ/2. However, the measure has some uncertainty that is defined in NF EN ISO/CEI 17025 (2005) and has to be compared to the requirements of the Directive 2004/107/CE (2004) for B[a]P and the technical prescriptions of the standard XP CEN/TS 166453 (2014).

In Section 2, is is shown how to compute the probability density function of a few number N of rectified normal variables $\phi_{\sum_{i=1}^{N} X_i^+}$. Here, X_i will be the average concentration of a batch of filters. It can be seen as a normal random variable, following the definition of its standard deviation in standard NF EN ISO/CEI 17025. Its mean is the value analyzed by the laboratory. Its standard deviation could be estimated using standard NF EN ISO/CEI 17025 or, if not possible, taken as the pessimistic case for the limit value of uncertainty defined by both the Directive 2004/107/CE (2004) and XP CEN/TS 166453 (2014).

Since the value below the LQ are not considered, the variable X_i has a rectified normal related variable X_i^+ for which the values of X_i lower than LQ are 0. The pdf of the mean $\frac{1}{N}\sum_{i=1}^{N} X_i^+$ directly comes from Theorem 1. It can be strongly assymetric depending on the number of batch filters available to compute the annual mean. Thus, knowing this pdf (and the cdf by extension) gives more information than the only value provided by the laboratory, see e.g. Figure 5 for the pdf of annual means computed from B[a]P batch filters.

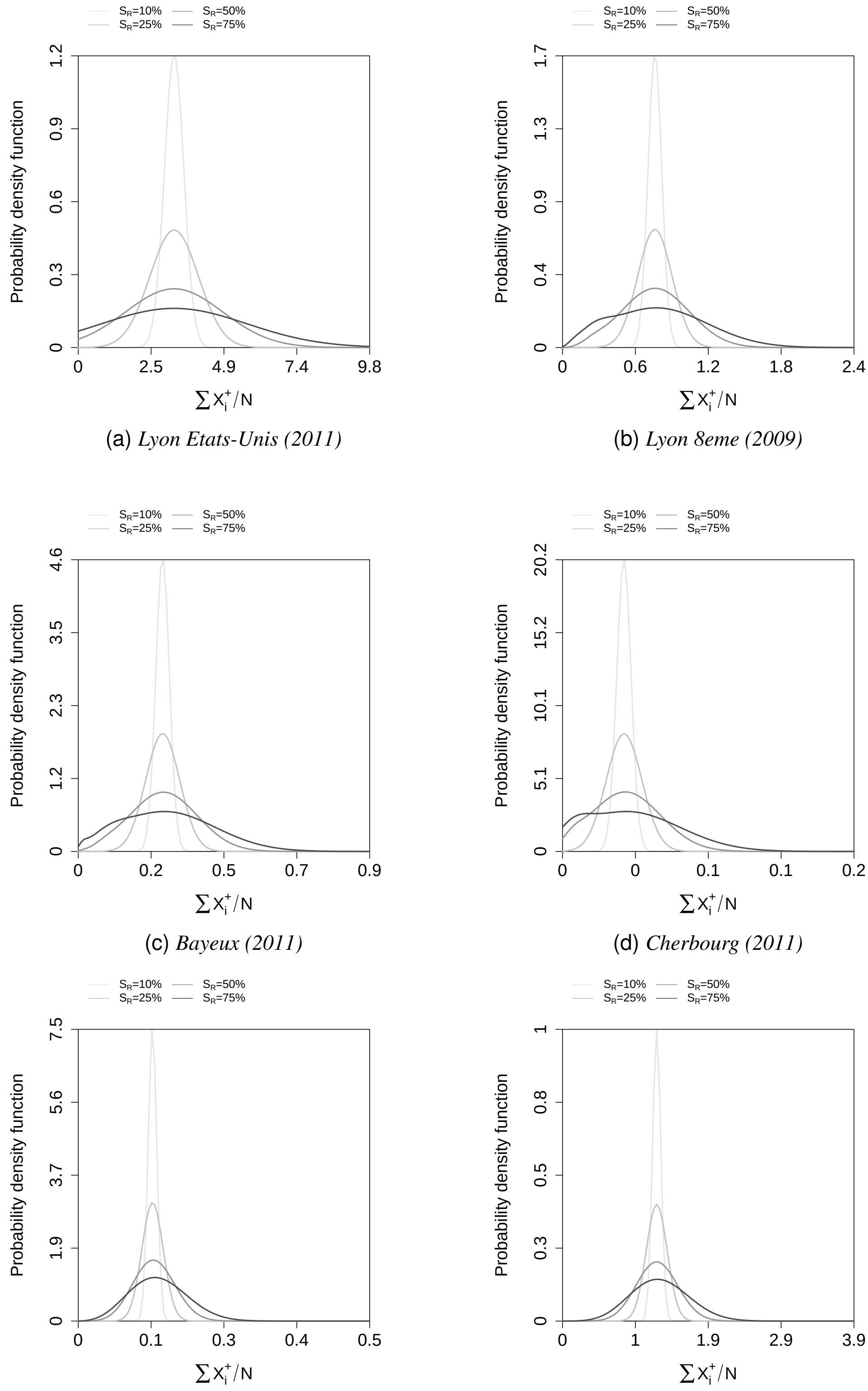
In addition, it is now possible to give the probability that the annual mean exceeds the limit value. Table 1 shows the probability $\mathbb{P}\left[\frac{1}{N}\sum_{i=1}^{N} X_i^+ > 1\right]$ for the same years and monitoring sites as in Figure 5. This quantity may significantly change with the value set for the uncertainty u.

TABLE 1. *Probability of exceeding the regulatory BaP limit value (1* ng.m^{-3}*) depending of the uncertainty level u*

Stations	Year	Mean	u=10%	u=25%	u=50%	u=75%
Lyon Etats-Unis	2001	3.24	1	1	0.94	0.91
L'Hôpital-Bois-Richard	2009	1.26	1	0.97	0.83	0.78
Lyon 8e	2009	0.77	0	0.05	0.21	0.3
Bayeux	2011	0.26	0	0	0	0
Cherbourg	2011	0.03	0	0	0.01	0.05
St-Louis	2011	0.13	0	0	0	0

6. Conclusion

In this paper, a new method to compute the probability density function for the sum of N rectified Gaussian variables is provided. The exact calculation of the pdf is made possible by a new probabilistic interpretation for the integral form of the multiple convolution $\left[\bigotimes_k \phi_k H\right](x)$ involved in the pdf. Its computational time is reasonable when the number N of variables is too small for applying the normal approximation of the central limit theorem. A comparison with Monte-Carlo simulations is made to ensure the calculation is correct. An application for the mean of batch filters BaP measurements shows how the method can be applied to environmental data without temporal correlation confronted to limit values. A probability of exceedance is now provided instead of a simple mean evaluation calculated by the laboratory. Monte-Carlo simulations could be used to reach almost the same results but because of the low CPU time necessary to compute the cumulative distribution function (similar and even lower for very small values of N), it is clearly more efficient to use the exact calculation.



(a) *Lyon Etats-Unis (2011)*

(b) *Lyon 8eme (2009)*

(c) *Bayeux (2011)*

(d) *Cherbourg (2011)*

(e) *St-Louis (2011)*

(f) *L'Hôpital-Bois-Richard (2009)*

FIGURE 5: *Probability density functions* $\phi_{\frac{1}{N}\sum_{i=1}^{N} X_i^+}$ *for the average concentration of N batch filters* ($N < 10$) *collected on 6 independent monitoring sites over France*

References

Albinet, A. (2011). Rapport de recommandations pour le prélèvement et l'analyse des hydrocarbures aromatiques polycycliques (hap) dans l'air ambiant. Technical report, LCSQA.

Ashour, S. and Abdel-Hamid, M. (2010). Approximate skew normal distribution. journal of advanced research. *Journal of Advanced Research,*, 1:341–350.

Billingsley, P. (1995). *Probability and Measure. 3rd Edition*. Probability and Mathematical Statistics. Wiley.

Bliss, C. I. (1934). The method of probits. *Science*, 79(2037):38–39.

Charles, D. and Fyfe, C. (1998). Modelling multiple-cause structure using rectification constraints. *Network: Computation in Neural Systems*, 9(2):167–182. PMID: 9861984.

Directive 2004/107/CE (2004). *Directive 2004/107/EC of the European Parliament and of the Council of 15 December 2004 relating to arsenic, cadmium, mercury, nickel and polycyclic aromatic hydrocarbons in ambient air*.

Directive 2008/50/EC (2008). *Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe*.

Eddelbuettel, D. (2013). *Seamless R and C++ Integration with Rcpp*. Springer, New York. ISBN 978-1-4614-6867-7.

Eddelbuettel, D. and François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, 40(8):1–18.

Genz, A. and Bretz, F. (2009). *Computation of Multivariate Normal and t Probabilities*. Lecture Notes in Statistics. Springer-Verlag, Heidelberg.

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., and Hothorn, T. (2017). *mvtnorm: Multivariate Normal and t Distributions*. R package version 1.0-6.

Harva, M. and Kabán, A. (2007). Variational learning for rectified factor analysis. *Signal Processing*, 87(3):509 – 527.

Krenek, R., Cha, J., and Cho, B. R. (2016). Development of the convolutions of truncated normal random variables with three different quality characteristics in engineering applications. *Comput. Ind. Eng.*, 94(C):125–137.

Lambert, D. (1992). Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1):1–14.

McConalogue, D. (1981). Numerical treatment of convolution integrals involving distributions with densities having singularities at the origin. *Commun. Statist. B Simulation Comput.*, 10:265–280.

Mudholkar, G. S. and Hutson, A. D. (2000). The epsilon-skew-normal distribution for analyzing near-normal data. *Journal of Statistical Planning and Inference*, 83:291–309.

NF EN ISO/CEI 17025 (September 2005). *Exigences générales concernant la compétence des laboratoires d'étalonnages et d'essais*.

O'Hagan, A. and Leonard, T. (1976). Bayes estimation subject to uncertainty about parameter constraints. *Biometrika*, 63:201–202.

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, 26(1):24–36.

XP CEN/TS 166453 (May 2014). *Air ambiant -Mesurage pour la mesure de benz(a)anthracène, benzo(b)fluoranthène, benzo(j)fluoranthène, benzo(k)fluoranthène, dibenz(a,h)anthracène, indéno(1,2,3-cd)pyrène et benzo(ghi)perylène*.

Appendices

A. Expectation and Variance of a rectified Gaussian variable

Let $I_k(a)$ denote:

$$I_k(a) = \int_a^{\infty} u^k e^{-\frac{u^2}{2}} du$$

Then, for $0 \leq k \leq 3$, integration by parts leads to:

$$\begin{aligned}
I_3(a) &= \int_a^{\infty} u^3 e^{-\frac{u^2}{2}} du = \int_a^{\infty} u^2 \left[u e^{-\frac{u^2}{2}} \right] du \\
&= -\left[u^2 e^{-\frac{u^2}{2}} \right]_a^{\infty} + 2\int_a^{\infty} u \left[e^{-\frac{u^2}{2}} \right] du = a^2 e^{-\frac{a^2}{2}} + 2e^{-\frac{a^2}{2}} \quad (21a) \\
I_2(a) &= \int_a^{\infty} u^2 e^{-\frac{u^2}{2}} du \\
&= -\left[u e^{-\frac{u^2}{2}} \right]_a^{\infty} + \int_a^{\infty} \left[e^{-\frac{u^2}{2}} \right] du = a e^{-\frac{a^2}{2}} + (1 - \Phi(a)) \quad (21b) \\
I_1(a) &= \int_a^{\infty} u e^{-\frac{u^2}{2}} du \\
&= -\left[e^{-\frac{u^2}{2}} \right]_a^{\infty} = e^{-\frac{a^2}{2}} \quad (21c) \\
I_0(a) &= \int_a^{\infty} e^{-\frac{u^2}{2}} du = 1 - \Phi(a) \quad (21d)
\end{aligned}$$

Based on this essential results and the substitution $u = \frac{x-\mu}{\sigma}$, the expectation and variance of a rectified Gaussian variable $X^+ \sim \mathcal{N}^+(\mu, \sigma^2)$, with μ and σ^2 the expectation and variance of the original Gaussian variable X, are:

$$\begin{aligned}
\mathbb{E}(X^+) &= \frac{1}{\sigma\sqrt{2\pi}} \int_0^{+\infty} x e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx \\
&= \frac{\mu}{\sqrt{2\pi}} \int_{-\frac{\mu}{\sigma}}^{+\infty} e^{-\frac{u^2}{2}} du + \frac{\sigma}{\sqrt{2\pi}} \int_{-\frac{\mu}{\sigma}}^{+\infty} u e^{-\frac{u^2}{2}} du = \mu(1 - \Phi(-\frac{\mu}{\sigma})) + \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{-\mu}{\sigma})^2} \\
\mathrm{Var}(X^+) &= \mathbb{E}[(X^+ - \mathbb{E}(X^+))^2] = \frac{1}{\sigma\sqrt{2\pi}} \left(\int_0^{+\infty} x^2 e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx - \left(\int_0^{+\infty} x e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx \right)^2 \right) \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\frac{\mu}{\sigma}}^{+\infty} u^2 e^{-\frac{u^2}{2}} du + \frac{2\mu\sigma}{\sqrt{2\pi}} \int_{-\frac{\mu}{\sigma}}^{+\infty} u e^{-\frac{u^2}{2}} du + \frac{\mu^2}{\sqrt{2\pi}} \int_{-\frac{\mu}{\sigma}}^{+\infty} e^{-\frac{u^2}{2}} du - \left(\mathbb{E}(X^+)\right)^2 \\
&= (\mu^2 + \sigma^2)(1 - \Phi(-\frac{\mu}{\sigma})) + \frac{\mu\sigma}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{-\mu}{\sigma})^2} - \left(\mathbb{E}(X^+)\right)^2
\end{aligned}$$

B. Calculation details for the proof of Theorem 1

B.1. Equation 7

$$
\begin{aligned}
\left[\bigotimes_{k=1}^{2}\phi_k H\right](x) &= \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\mu_1}^{x-\mu_1} e^{-\frac{\left[r^{(1)}\right]^2}{2\sigma_1^2}} e^{-\frac{(r^{(1)}+y^{(2)})^2}{2\sigma_2^2}} dr^{(1)} \\
&= \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\mu_1}^{x-\mu_1} e^{-\frac{(r^{(1)}\sigma^{(2)}+\frac{y^{(2)}\sigma_1^2}{\sigma^{(2)}})+\frac{\left[y^{(2)}\right]^2\sigma_1^2\sigma_2^2}{\sigma^2}}{2\sigma_1^2\sigma_2^2}} dr^{(1)} \\
&= \frac{1}{2\pi\sigma_1\sigma_2} e^{\frac{\left[y^{(2)}\right]^2}{2\left[\sigma^{(2)}\right]^2}} \int_{-\mu_1}^{x-\mu_1} e^{-\frac{r^{(1)}\left[\sigma^{(2)}\right]^2+y^{(2)}\left[\sigma^{(1)}\right]^2}{\sigma^{(1)}\sigma^{(2)}\sigma_2}} dr^{(1)} \\
&= \frac{1}{2\pi\sigma^{(2)}} e^{\frac{\left[y^{(2)}\right]^2}{2\left[\sigma^{(2)}\right]^2}} \int_{b_-(x)}^{b_+(x)} e^{-\frac{\left[u^{(1)}\right]^2}{2}} du^{(1)} \\
&= \frac{1}{\sqrt{2\pi}}\phi^2(x)\int_{b_-(x)}^{b_+(x)} e^{-\frac{\left[u^{(1)}\right]^2}{2}} du^{(1)}
\end{aligned}
$$

B.2. Equation 8

$$
\begin{aligned}
\left[\bigotimes_{k=1}^{3}\phi_k H\right](x) &= \left[\left(\bigotimes_{k=1}^{2}\phi_k H\right) * \phi_3 H\right](x) \\
&= \int_0^x \frac{1}{2\pi\sigma^{(2)}} e^{\frac{\left[y^{(2)}\right]^2}{2\left[\sigma^{(2)}\right]^2}} \int_{b_-(s^{(2)})}^{b_+(s^{(2)})} e^{-\frac{\left[u^{(1)}\right]^2}{2}} du^{(1)} \frac{1}{\sqrt{2\pi}\sigma_3} e^{-}\frac{(x-s^{(2)}-\mu_3)^2}{2\sigma_3^2} ds^{(2)} \\
&= \frac{1}{\sqrt{2\pi}^{3/2}\sigma^{(2)}\sigma_3}\int_0^x e^{\frac{(\mu^{(2)}-s^{(2)})^2}{2\left[\sigma^{(2)}\right]^2}} e^{\frac{(x-s^{(2)}-\mu_3)^2}{2\sigma_3^2}} \left[\int_{b_-(s^{(2)})}^{b_+(s^{(2)})} e^{-\frac{\left[u^{(1)}\right]^2}{2}} du^{(1)}\right] ds^{(2)} \\
&= \frac{1}{\sqrt{2\pi}^{3/2}\sigma^{(2)}\sigma_3}\int_{-\mu^{(2)}}^{x-\mu^{(2)}} e^{-\frac{\left[r^{(2)}\right]^2}{2\left[\sigma^{(2)}\right]^2}} e^{-\frac{\left[r^{(2)}+y^{(3)}\right]^2}{2\sigma_3^2}} \left[\int_{b_-^{(1)}(s^{(2)})}^{b_+^{(1)}(s^{(2)})} e^{-\frac{\left[u^{(1)}\right]^2}{2}} du^{(1)}\right] dr^{(2)} \\
&= \frac{1}{2\pi}\phi^{(3)}(x)\int_{b_-^{(2)}(x)}^{b_+^{(2)}(x)} e^{-\frac{\left[u^{(2)}\right]^2}{2}} \left[\int_{b_-^{(1)}(s^{(2)})}^{b_+^{(1)}(s^{(2)})} e^{-\frac{\left[u^{(1)}\right]^2}{2}} du^{(1)}\right] du^{(2)}
\end{aligned}
$$