Model-free computation of risk contributions in credit portfolios

Álvaro Leitao[1,2,*] and Luis Ortiz-Gracia[3,†]

[1]CITIC research center, University of A Coruña, Spain
[2]M2NICA research group, Department of Mathematics, University of A Coruña, Spain
[3]Department of Econometrics, Statistics and Applied Economics, University of Barcelona, Spain

April 13, 2020

Abstract

In this work, we propose a non-parametric density estimation technique for measuring the risk in a credit portfolio, aiming at efficiently computing the marginal risk contributions. The novel method is based on wavelets, and we derive closed-form expressions to calculate the Value-at-Risk (VaR), the Expected Shortfall (ES) as well as the individual risk contributions to VaR (VaRC) and ES (ESC). We consider the multi-factor Gaussian and t-copula models for driving the defaults. The results obtained along the numerical experiments show the impressive accuracy and speed of this method when compared with crude Monte Carlo simulation. The presented methodology applies in the same manner regardless of the used model, and the computational performance is invariant under a considerable change in the dimension of the selected model. The speed-up with respect to the classical Monte Carlo approach ranges from twenty-five to one-thousand depending on the used model.

Key words. Credit Risk; Value-at-Risk; Expected Shortfall; Portfolio Risk Contributions; Shannon Wavelets; Non-parametric Density Estimation.

AMS subject classifications. 91G60, 62P05, 65T60

1 Introduction

Portfolio credit risk represents one of the most important sources of risk that any financial institution has to face with. As an important part of the credit risk management process, some common risk measures are usually employed like the well-known VaR and ES. Besides the calculation of these measures, the decomposition of the total risk of a given portfolio into the individual risk contribution of each obligor is a problem of practical importance. Identification of risk concentrations, portfolio optimization or capital allocation are, among others, relevant examples of application. An important amount of papers have addressed the problem of decomposing the total risk into the individual parts (see for instance [21, 11, 9]).

The problem of obtaining the risk contributions represents a great challenge from the computational standpoint. Within the credit risk literature, many authors rely on Monte Carlo simulation, since it is straightforward to implement and can be easily extended to multi-factor

*Corresponding author. E-mail: alvaro.leitao@udc.gal.
†E-mail: luis.ortiz-gracia@ub.edu.

models. This fact makes Monte Carlo methods very attractive for practitioners. By following this approach, the problem of computing the risk contributions is basically reduced to the computation of a certain expectation conditioned on a rare event, which makes the plain application of Monte Carlo method rather inefficient. To overcome this issue, most of the methodologies proposed in the last few years usually include more advanced simulation tools. By far, the so-called *importance sampling* technique (originally applied as a variance reduction technique) appears to be the most promising development, providing a significant performance gain with respect to the plain Monte Carlo. Some representative works were proposed in [7, 8, 22] or, more recently, in [12]. Particularly in [22], the author presents an approach based on *kernel estimators*, but this methodology differs from ours since it is applied within the context of the importance sampling, so the kernel estimation is not employed to recover the loss distribution itself. All in all, the application of importance sampling often requires the knowledge of the portfolio behavioural model, i.e., it is somehow model dependent. Different techniques based on Markov chain Monte Carlo in [10] or Finite Difference in [16] have been recently proposed, but all these approaches have in common that, generally speaking, simulation implies a relatively big computational cost to control the variance.

Following the classical approach the calculation of VaRC and ESC requires the computation either of an expectation given that the total loss equals the VaR value (for VaRC) or an expectation given that the total loss is greater than the VaR value (for ESC). These expectations are therefore calculated by simulation conditioned on a rare event, and this fact makes impractical Monte Carlo methods for big portfolios (which are typically encountered in real situations). From the theoretical point of view, as pointed out by [2], it is not guaranteed that risk contributions to VaR are always smaller than their respective exposures and it is therefore worth computing the ESC as well.

In this work, we propose a non-parametric density estimation based on wavelets. The starting point is the sample of the total loss variable generated by Monte Carlo simulation. Then, the density is estimated either with Haar or Shannon wavelets and the VaR and ES are obtained. It is worth remarking that the main focus of this work is the computation of risk contributions once risk measures are available. Here, we entirely focus on the Euler's capital allocation principle [2, 15] to compute the risk contributions for the VaR and the ES. The Euler's capital allocation principle can be used under the framework of positive homogeneous risk measures of degree one, so our approach is not only valid for these two risk measures, but also for others like, for instance, the unexpected losses, defined as the standard deviation of the total portfolio losses. According to the Euler's capital allocation principle, we take derivatives of the risk measures (VaR or ES) with respect to the exposures and we obtain the VaRC and ESC. While the Haar family has desirable properties like compact support and gives us positive densities by construction, we finally prefer the Shannon family due to its robustness and easy handling. We test our method with one- and multi-factor Gaussian and t-copula models. These models belong to the class of structural models and they are currently used in practice, since they are the models in force given by the regulators. The computation of VaRC by means of Monte Carlo simulation is particularly difficult due to the sensitivity of the results with respect to the length of the interval selected containing the VaR value in the conditional expectation mentioned above. In contrast, the wavelets based method gives impressive results, both in accuracy and speed. For obtaining comparable results in accuracy, Monte Carlo needs between 25 (for the one-factor model) and 1000 times (for multi-factor models) the CPU time required by our method. While the wavelet machinery is not affected by increasing the number of factors in the model, Monte Carlo needs three times more seconds of CPU time when moving from 5 to 25 factors. Moreover, our methodology is model-free in the sense that it stays the same and it applies in the same manner, regardless of the model employed for driving the defaults. To the best of our knowledge, this is the first time that this technique is used for solving the capital allocation problem by means of Euler's capital allocation principle. Multiple tests carried out along this work make us

think that this novel method can be used within the risk management toolkit of financial firms.

The paper is organised as follows. We formulate the credit risk problem and describe the one- and multi-factor models in Section 2. In Section 3, we give an overview of wavelets and multi-resolution analysis, and explain in detail the non-parametric density estimation procedure for obtaining the risk measures and risk contributions. Section 4 is devoted to the numerical experiments and Section 5 concludes.

2 Problem formulation

Let us consider a portfolio consisting in N obligors. Each obligor is characterized by the *exposure at default*, the *probability of default* and the *loss given default*. While the first two parameters will be denoted by E_j and P_j, $j = 1, \ldots, N$, the third parameter is assumed to be 100% for all the obligors. These are the so-called capital parameters and they are typically estimated from default data.

Assume now that we are in the framework of Merton's firm-value model. Let $V_j(t)$ denote the asset value of obligor j at time $t < T$, where T is the time horizon (typically one year). The obligor j defaults when its value at the end of the observation period, $V_j(T)$, falls below a certain threshold, τ_j, i.e, $V_j(T) < \tau_j$. We can therefore define the default indicator as,

$$D_j = \mathbb{1}_{\{V_j(T)<\tau_j\}} \sim Be\left(\mathbb{P}\left(V_j(T) < \tau_j\right)\right),$$

where $Be(p)$ is a Bernoulli distribution with probability of success p. Given D_j, the individual loss of obligor j is defined as,

$$L_j = D_j \cdot E_j,$$

while the total loss in the portfolio reads,

$$L = \sum_{j=1}^{N} L_j. \tag{1}$$

2.1 Factor models for portfolio credit risk

Factor models belong to the class of structural models. Within this class of models, loss only occurs when an obligor defaults in a fixed time horizon T. Based on Merton's firm-value model, to describe the obligor's default and its correlation structure, we assign to each obligor a latent random variable called firm-value. The firm-value (or, more precisely, the asset value log-return) of each obligor is split into two terms: one common component usually called systematic factor, and an idiosyncratic component for each obligor. Depending on the number of factors of the systematic part, the model can be classified into the *one-* or *multi-factor* class. In the following we briefly describe the models used in this work.

2.1.1 One-factor models

In the one-factor model setting, the firm-value of obligor j, V_j, at time T is represented by a common, standard normally distributed single factor Y component and an idiosyncratic Gaussian noise component ε_j. The dependence structure between the latent random variables is given by the use of copulas. Thus, these models are also called one-factor copula models. Two models are usually considered in practice. The so-called Gaussian copula model,

$$V_j = \sqrt{\rho_j} Y + \sqrt{1-\rho_j}\varepsilon_j, \tag{2}$$

where Y and ε_j are i.i.d. standard normal random variables for all $j = 1, \ldots, N$. Alternatively, as an extension of the model in Equation (2), the so-called t-copula model was introduced to

take into account tail dependence [20],

$$V_j = \sqrt{\frac{\nu}{W}} \left(\sqrt{\rho_j} Y + \sqrt{1-\rho_j} \varepsilon_j \right), \tag{3}$$

where $\varepsilon_1, \cdots, \varepsilon_N, Y \sim \mathcal{N}(0,1)$, W follows a chi-square distribution $\chi^2(\nu)$ with ν degrees of freedom and $\varepsilon_1, \cdots, \varepsilon_N$, Y and W are mutually independent. Scaling the model in Equation (2) by the factor $\sqrt{\nu/W}$ transforms standard Gaussian random variables into t-distributed random variables with ν degrees of freedom. For both models, the parameters $\rho_1, \cdots, \rho_N \in (0,1)$ are the correlation coefficients. In case that $\rho_j = \rho$, for all $j = 1, \ldots, N$, the parameter ρ is called the common asset correlation.

According to the Merton's model described above, obligor j defaults when its firm-value falls below the threshold level τ_j. The threshold is therefore defined by $\tau_j := \Phi^{-1}(P_j)$ or $\tau_j := \Phi_\nu^{-1}(P_j)$ for the Gaussian and t-copula models respectively, where Φ^{-1} denotes the inverse of the standard normal cumulative distribution function and Φ_ν^{-1} is the corresponding inverse distribution function of the t-distribution (with ν degrees of freedom).

2.1.2 Multi-factor models

Multi-factor models aim to capture complicated correlation structures. We consider the extension to multiple dimensions of the models presented in Section 2.1.1, i.e., the multi-factor Gaussian copula model and the multi-factor t-copula model.

The d-factor Gaussian copula model assumes that the covariance structure of $[V_1, \ldots, V_N]$ is determined by the multi-factor model,

$$V_j = \boldsymbol{a_j}^T \boldsymbol{Y} + b_j \varepsilon_j, \quad j = 1, \cdots, N. \tag{4}$$

where $\boldsymbol{Y} = [Y_1,\ Y_2, \ldots, Y_d]^T$ denotes the systematic risk factors. Note that we represent vectors by bold symbols throughout the paper. Here, $\boldsymbol{a_j} = [a_{j1}, a_{j2}, \ldots, a_{jd}]^T$ represents the factor loadings satisfying $\boldsymbol{a_j}^T \boldsymbol{a_j} < 1$, and ε_j are standard normally distributed random variables representing the idiosyncratic risks, independent of each other and independent of $\boldsymbol{Y}$. The constant b_j, being the factor loading of the idiosyncratic risk factor, is chosen so that V_j has unit variance, i.e., $b_j = \sqrt{1 - \left(a_{j1}^2 + a_{j2}^2 + \cdots + a_{jd}^2 \right)}$, which ensures that V_j is $\mathcal{N}(0,1)$.

The incentive for considering the multi-factor version of the Gaussian copula model becomes clear when one rewrites it in matrix form,

$$\begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_N \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{N1} \end{bmatrix} Y_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{N2} \end{bmatrix} Y_2 + \cdots + \begin{bmatrix} a_{1d} \\ a_{2d} \\ \vdots \\ a_{Nd} \end{bmatrix} Y_d + \begin{bmatrix} b_1 \varepsilon_1 \\ b_2 \varepsilon_2 \\ \vdots \\ b_N \varepsilon_N \end{bmatrix}.$$

While each ε_j represents the idiosyncratic factor affecting only obligor j, the common factors $Y_1, Y_2 \ldots, Y_d$, may affect all (or a certain group of) obligors. Although the systematic factors are sometimes given economic interpretations (as industry or regional risk factors, for example), their key role being that they allow us to model complicated correlation structures in a non-homogeneous portfolio.

Similarly, the multi-factor t-copula model definition reads,

$$V_j = \sqrt{\frac{\nu}{W}} \left(\boldsymbol{a_j}^T \boldsymbol{Y} + b_j \varepsilon_j \right), \quad j = 1, \cdots, N, \tag{5}$$

where $\boldsymbol{Y}, \varepsilon_j, \boldsymbol{a_j}$ and b_j are defined as before, with $W \sim \chi^2(\nu)$.

2.2 Risk measures and risk contributions

Within this work, we will use two well-known measures of risk, the VaR and the ES. For the sake of completeness, we give an overview of these two measures.

Definition 1. *Given a confidence level* $\alpha \in (0,1)$ *and the vector of exposures* $\boldsymbol{E} = [E_1, E_2, \ldots, E_N]^T$*, we define the portfolio VaR,*

$$VaR_\alpha(\boldsymbol{E}) = \inf\{l \in \mathbb{R} : \mathbb{P}(L \leq l; \boldsymbol{E}) \geq \alpha\} = \inf\{l \in \mathbb{R} : F_L(l; \boldsymbol{E}) \geq \alpha\},$$

where F_L *is the distribution function of the total loss random variable* L *(we emphasize the dependence of VaR with respect to the risk exposures).*

Definition 2. *Given the loss variable* L *with* $\mathbb{E}[|L|] < \infty$ *and distribution function* F_L*, the ES at confidence level* $\alpha \in (0,1)$ *is defined as,*

$$ES_\alpha(\boldsymbol{E}) = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(\boldsymbol{E})\mathrm{d}u.$$

When the loss variable is integrable with continuous distribution function, then the ES satisfies (see Lemma 2.4.2 of [13]) the equation,

$$\mathrm{ES}_\alpha(\boldsymbol{E}) = \mathbb{E}[L|L \geq \mathrm{VaR}_\alpha(\boldsymbol{E})], \tag{6}$$

or, in integral form,

$$\mathrm{ES}_\alpha(\boldsymbol{E}) = \frac{1}{1-\alpha}\int_{\mathrm{VaR}_\alpha(\boldsymbol{E})}^{+\infty} x f_L(x; \boldsymbol{E})\mathrm{d}x, \tag{7}$$

where f_L is the probability density function of the total loss random variable L. Thus, in the continuous case, ES can be interpreted as the expected loss in the event that VaR is exceeded. In the discontinuous case, a more elaborated formula holds (see Section 2.4 of [13] and the references therein).

A problem of paramount importance in the practice of quantitative risk management is allocating the risk to the elements of the portfolio, based on their individual contribution to the corresponding risk measure. This leads to the computation of VaRC and ESC. This problem is known as *capital allocation* and a solution to this problem is the *Euler's capital allocation principle* which states that when a risk measure is positive-homogeneous (like VaR and ES) and differentiable with respect to the exposures then,

$$\sum_{j=1}^N E_j \frac{\partial \mathrm{VaR}_\alpha}{\partial E_j}(\boldsymbol{E}) = \mathrm{VaR}_\alpha(\boldsymbol{E}), \quad \text{and,} \quad \sum_{j=1}^N E_j \frac{\partial \mathrm{ES}_\alpha}{\partial E_j}(\boldsymbol{E}) = \mathrm{ES}_\alpha(\boldsymbol{E}). \tag{8}$$

If we define,

$$\mathrm{VaR}C_{\alpha,j} := E_j \frac{\partial \mathrm{VaR}_\alpha}{\partial E_j}(\boldsymbol{E}), \quad \text{and,} \quad \mathrm{ES}C_{\alpha,j} := E_j \frac{\partial \mathrm{ES}_\alpha}{\partial E_j}(\boldsymbol{E}), \tag{9}$$

as the contribution of obligor j to the VaR (respectively ES) at confidence level α, then, subject to technical conditions (see Section 8.5 of [15] and the references therein for details), it can be shown that,

$$\mathrm{VaR}C_{\alpha,j} = \mathbb{E}\left[L_j | L = \mathrm{VaR}_\alpha(\boldsymbol{E})\right], \quad j = 1, \ldots, N, \tag{10}$$

and,

$$\mathrm{ES}C_{\alpha,j} = \mathbb{E}\left[L_j | L \geq \mathrm{VaR}_\alpha(\boldsymbol{E})\right], \quad j = 1, \ldots, N. \tag{11}$$

3 Wavelet-based estimation of the loss distribution

3.1 Multi-resolution analysis

For the sake of clarity we devote a section to the basic theory on wavelets. A basic reference on wavelets is [5].

Consider the space of square-integrable functions, denoted by $L^2(\mathbb{R})$. A general structure for wavelets in $L^2(\mathbb{R})$ is called a *multi-resolution analysis*. We start with a family of closed nested subspaces in $L^2(\mathbb{R})$,

$$\ldots \subset \mathcal{V}_{-2} \subset \mathcal{V}_{-1} \subset \mathcal{V}_0 \subset \mathcal{V}_1 \subset \mathcal{V}_2 \subset \ldots, \quad \bigcap_{m\in\mathbb{Z}} \mathcal{V}_m = \{0\}\ , \quad \overline{\bigcup_{m\in\mathbb{Z}} \mathcal{V}_m} = L^2(\mathbb{R}),$$

where,

$$f(x) \in \mathcal{V}_m \Longleftrightarrow f(2x) \in \mathcal{V}_{m+1}.$$

If these conditions are met, then there exists a function $\phi \in \mathcal{V}_0$ that generates an orthonormal basis for each subspace $\mathcal{V}_m$, denoted by $\{\phi_{m,k}\}_{k\in\mathbb{Z}}$, where $\phi_{m,k}(x) = 2^{m/2}\phi(2^m x - k)$. The function ϕ is usually referred to as the *scaling function* or *father wavelet*.

For any $f \in L^2(\mathbb{R})$, a projection map of $L^2(\mathbb{R})$ onto $\mathcal{V}_m$, denoted by $\mathcal{P}_m : L^2(\mathbb{R}) \to \mathcal{V}_m$, is defined by means of,

$$\mathcal{P}_m f(x) = \sum_{k\in\mathbb{Z}} c_{m,k}\phi_{m,k}(x), \quad \text{with} \quad c_{m,k} = \langle f, \phi_{m,k}\rangle\ ,$$

where $\langle f, g\rangle := \int_{\mathbb{R}} f(x)\bar{g}(x)\mathrm{d}x$ denotes the inner product in $L^2(\mathbb{R})$, with $\bar{g}$ being the complex conjugate of g. As opposed to Fourier series, a key fact regarding the use of wavelets is that wavelets can be moved (by means of the k value), stretched or compressed (by means of the m value) to accurately represent the local properties of a function.

Following wavelets theory, a function $f \in L^2(\mathbb{R})$ can be approximated at the level of resolution m by,

$$f(x) \approx \mathcal{P}_m f(x) = \sum_{k\in\mathbb{Z}} c_{m,k}\phi_{m,k}(x), \tag{12}$$

where $\mathcal{P}_m f$ converges to f in $L^2(\mathbb{R})$, i.e. $\|f - \mathcal{P}_m f\|_2 \to 0$, when $m \to +\infty$. Considering higher values of m (i.e., when more terms are used), the truncated series representation of the function f improves.

3.2 Non-parametric density estimation by wavelets: application to the loss distribution

In this section we briefly describe some concepts of the non-parametric density estimation using wavelets and its application to recover the density function of the loss distribution. For more detailed description of wavelet density estimation, we refer the reader to [24, 25] for example.

Given i.i.d samples from an unknown statistical distribution X with density function denoted by f_X, we apply the wavelet theory to approximate f_X. We consider here the so-called *linear wavelet estimator* (or simply *linear estimator*) which, after the truncation of the infinite sum in Equation (12), is written as,

$$f_X(x) \approx \sum_k c_{m,k}\phi_{m,k}(x),$$

where k varies within a finite range of values that will be determined later on, and the coefficients $c_{m,k}$ are given by,

$$c_{m,k} := \langle f_X, \phi_{m,k}\rangle = \int_{\mathbb{R}} f_X(x)\bar{\phi}_{m,k}(x)\,\mathrm{d}x = \mathbb{E}\left[\bar{\phi}_{m,k}(X)\right], \tag{13}$$

where, as mentioned, $\bar{\phi}$ represents the complex conjugate of ϕ. The last equality comes from the fact that f_X is a density function. As stated by Vannucci in [23], the linear estimator overcomes the classical series estimators (Fourier, Hermite) that appear to be poor in estimating local features of the density since they have poor time/frequency localization properties. On the contrary, wavelets are localized both in time and frequency. Furthermore, the linear estimator can be written in terms of the so-called *reproducing kernels* [1], belonging therefore to the general class of *delta sequence* estimators. This facilitates its asymptotic analysis, showing faster rates of convergence than other classical methodologies, specially for the case of continuous (regular) densities.

Let us estimate the density function of the loss variable L by means of the wavelet estimator presented above. We first generate by Monte Carlo simulation a sample of size n corresponding to the loss variable L and denoted by $L^1, \ldots, L^n$. Given a realization of the systematic factor Y (respectively $\boldsymbol{Y}$) within the class of one-factor models of Section 2.1.1 (respectively multi-factor models of Section 2.1.2), defaults become independent. In what follows, we denote by D_j^i the default indicator of obligor j and realization of the systematic factor i, where $i = 1, \ldots, n, j = 1 \ldots, N$. For convenience, we consider the transformation $Z = \frac{L-a}{b-a}$, and we define $Z^i = \frac{L^i-a}{b-a}, i = 1, \ldots, n$, where,

$$a = \min_{1 \leq i \leq n} \left(L^i\right), \quad b = \max_{1 \leq i \leq n} \left(L^i\right).$$

This normalization transforms the sample interval $[a, b]$ into $[0, 1]$, where the Haar scaling function is naturally defined. Further, it can prevent from potential numerical instabilities and/or computational inaccuracies, produced by the big numbers appearing in practice.

From Equation (13) we obtain the following unbiased estimator for the wavelet series coefficients,

$$\hat{c}_{m,k} := \frac{1}{b-a} \frac{1}{n} \sum_{i=1}^{n} \phi_{m,k}(Z^i). \tag{14}$$

Next, using the wavelet density estimation, the unknown density f_L of L can be therefore approximated as follows,

$$f_L(x; \boldsymbol{E}) \approx \hat{f}_L(x; \boldsymbol{E}) := \sum_{k=0}^{\mathcal{K}} \hat{c}_{m,k} \phi_{m,k} \left(\frac{x-a}{b-a}\right), \tag{15}$$

where, as it will become clear in Section 3.3.1 and Section 3.3.2, by construction, the lower bound for index k is equal to zero, while the upper limit is $\mathcal{K} = 2^m - 1$.

The distribution function of L will be used for computing the risk measures and risk contributions and it can be derived by using expression (15) as follows,

$$F_L(x; \boldsymbol{E}) := \int_{-\infty}^{x} f_L(y; \boldsymbol{E}) \mathrm{d}y \approx \hat{F}_L(x; \boldsymbol{E}) := \sum_{k=0}^{\mathcal{K}} \hat{c}_{m,k} \int_a^x \phi_{m,k} \left(\frac{y-a}{b-a}\right) \mathrm{d}y. \tag{16}$$

We note that the integral in the right hand side of expression (16) will be solved either exactly with Haar wavelets (details in Section 3.3.1) or approximately with Shannon wavelets (details in Section 3.3.2).

3.3 Computation of risk measures and risk contributions

The VaR value is obtained by using a root-finding method to solve the following equation,

$$\hat{F}_L(x; \boldsymbol{E}) = \alpha, \tag{17}$$

where $\hat{F}_L(x;\boldsymbol{E})$ is defined in expression (16) and α is the confidence level given in Definition 1. As pointed out in Section 5.2.4 of [2], although the distribution of the portfolio loss L in (1) is discontinuous, it will appear to be "close to continuous" when the portfolio is large. It is therefore a common practice to compute the ES by means of expression (7). Thus, we replace $\hat{f}_L(x;\boldsymbol{E})$ in expression (7) to obtain an estimation of the ES,

$$\mathrm{ES}_\alpha(\boldsymbol{E}) \approx \frac{1}{1-\alpha}\int_{\mathrm{VaR}_\alpha(\boldsymbol{E})}^{b} x\hat{f}_L(x;\boldsymbol{E})\mathrm{d}x, \tag{18}$$

and then by (15) we get,

$$\mathrm{ES}_\alpha(\boldsymbol{E}) \approx \frac{1}{1-\alpha}\sum_{k=0}^{\mathcal{K}}\hat{c}_{m,k}\int_{\mathrm{VaR}_\alpha(\boldsymbol{E})}^{b} x\phi_{m,k}\left(\frac{x-a}{b-a}\right)\mathrm{d}x. \tag{19}$$

It is worth remarking that the VaR value can be obtained directly from the samples generated by Monte Carlo simulation and the ES can be consequently computed by means of (6). Nevertheless, for the sake of completeness, we use the methodology presented in this work for their estimation.

The risk contributions (VaRC and ESC) will be calculated by following the Euler's capital allocation principle in (8), since the distribution function $\hat{F}_L(x;\boldsymbol{E})$ is differentiable with respect to the exposures (in the case of Haar basis we approximate the scaling function by a differentiable function). Technical conditions under which the equalities (10) and (11) hold, are difficult to verify in practice. Despite of this, a benchmark solution for VaRC (respectively ESC) will be computed by Monte Carlo simulation using the expectation on the right-hand side of Equation (10) (respectively Equation (11)), since those expectations decompose the total risk (see for instance [7]).

Looking at expression (17), the VaR value satisfies,

$$\hat{F}_L(\mathrm{VaR}_\alpha(\boldsymbol{E});\boldsymbol{E}) = \alpha, \tag{20}$$

and, if we differentiate expression (20) we obtain,

$$\mathrm{VaR}C_{\alpha,j} = E_j\frac{\partial\mathrm{VaR}_\alpha}{\partial E_j}(\boldsymbol{E}) = -E_j\frac{\frac{\partial\hat{F}_L}{\partial E_j}(\mathrm{VaR}_\alpha(\boldsymbol{E});\boldsymbol{E})}{\frac{\partial\hat{F}_L}{\partial x}(x;\boldsymbol{E})|_{x=\mathrm{VaR}_\alpha(\boldsymbol{E})}} = -E_j\frac{\frac{\partial\hat{F}_L}{\partial E_j}(\mathrm{VaR}_\alpha(\boldsymbol{E});\boldsymbol{E})}{\hat{f}_L(\mathrm{VaR}_\alpha(\boldsymbol{E});\boldsymbol{E})}. \tag{21}$$

If we integrate by parts (18) then,

$$\mathrm{ES}_\alpha(\boldsymbol{E}) \approx \frac{1}{1-\alpha}\left(b - \alpha\mathrm{VaR}_\alpha(\boldsymbol{E}) - \int_{\mathrm{VaR}_\alpha(\boldsymbol{E})}^{b}\hat{F}_L(x;\boldsymbol{E})\mathrm{d}x\right), \tag{22}$$

and finally, by taking partial derivatives of (22) with respect to E_j, we end up with the risk contributions to the ES,

$$\begin{aligned}\mathrm{ES}C_{\alpha,j} &= E_j\frac{\partial\mathrm{ES}_\alpha}{\partial E_j}(\boldsymbol{E})\\ &= \frac{1}{1-\alpha}E_j\left(-\alpha\frac{\partial\mathrm{VaR}_\alpha}{\partial E_j}(\boldsymbol{E}) + \frac{\partial\mathrm{VaR}_\alpha}{\partial E_j}(\boldsymbol{E})\hat{F}_L\left(\mathrm{VaR}_\alpha(\boldsymbol{E});\boldsymbol{E}\right) - \int_{\mathrm{VaR}_\alpha(\boldsymbol{E})}^{b}\frac{\partial\hat{F}_L}{\partial E_j}\left(x;\boldsymbol{E}\right)\mathrm{d}x\right)\\ &= -\frac{1}{1-\alpha}E_j\int_{\mathrm{VaR}_\alpha(\boldsymbol{E})}^{b}\frac{\partial\hat{F}_L}{\partial E_j}\left(x;\boldsymbol{E}\right)\mathrm{d}x.\end{aligned} \tag{23}$$

From expressions (21) and (23), we see that VaRC and ESC require the computation of the partial derivative of the distribution function with respect to the exposures, $\frac{\partial\hat{F}_L(x;\boldsymbol{E})}{\partial E_j}$. Given the

wavelet approximation of $\hat{F}_L$ in (16) which depends on E_j only through the coefficients $\hat{c}_{m,k}$, we can compute the required partial derivative as,

$$\frac{\partial \hat{F}_L}{\partial E_j}(x;\boldsymbol{E}) = \frac{\partial}{\partial E_j}\left(\sum_{k=0}^{\mathcal{K}} \hat{c}_{m,k}\int_a^x \phi_{m,k}\left(\frac{y-a}{b-a}\right)\mathrm{d}y\right) = \sum_{k=0}^{\mathcal{K}} \frac{\partial \hat{c}_{m,k}}{\partial E_j}\int_a^x \phi_{m,k}\left(\frac{y-a}{b-a}\right)\mathrm{d}y, \quad (24)$$

where, from expression (14) we have,

$$\frac{\partial \hat{c}_{m,k}}{\partial E_j} = \frac{\partial}{\partial E_j}\left(\frac{1}{b-a}\frac{1}{n}\sum_{i=1}^{n}\phi_{m,k}(Z^i)\right) = \frac{1}{b-a}\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \phi_{m,k}}{\partial E_j}(Z^i) = \frac{2^{3m/2}}{(b-a)^2}\frac{1}{n}\sum_{i=1}^{n} D_j^i \phi'(2^m Z^i - k), \quad (25)$$

and we assume that ϕ is differentiable.

Remark 1. *The computation of the expression on the right hand side of Equation* (25) *can be rewritten as an sparse matrix-vector multiplication, which provides a significant improvement in terms of computational efficiency in our routines.*

The accuracy of the data-driven approach presented so far, depends on the family of wavelets selected to approximate the density function. Within this work, we consider the Haar and Shannon wavelets. While the estimation of the density function by means of Haar wavelets gives always positive values (by construction), their derivatives either do not exist or vanish to zero. This fact hampers the calculation of the partial derivatives in expression (25), which is a crucial step when computing VaRC and ESC. We overcome this problem by using a differentiable approximation of the Haar scaling function. Despite of this approximation, we will show that Shannon wavelets give us a more robust and easy handling solution. We give the details in Section 3.3.1 and Section 3.3.2.

3.3.1 Haar wavelets

The Haar scaling function reads,

$$\phi(x) = \begin{cases} 1, & 0 \le x < 1, \\ 0, & \text{otherwise.} \end{cases} \quad (26)$$

While $\hat{c}_{m,k}$ is easily calculated by replacing the Haar scaling functions $\phi_{m,k}(x) = 2^{m/2}\phi(2^m x - k)$, with ϕ defined in (26), in Equation (14), its partial derivative with respect to E_j, requires the derivative of the Haar scaling function, which does not exist in the context of regular functions. In the context of generalized functions, we can write,

$$\phi'(x) = \delta(x) - \delta(x-1),$$

where δ represents the Dirac delta function. The Haar scaling function can be rewritten in terms of the difference between two Heaviside functions as follows,

$$\phi(x) = H(x) - H(x-1),$$

where function H is defined by,

$$H(x) = \begin{cases} 0, & x < 0, \\ 1, & x \ge 0. \end{cases}$$

Next, we consider a well-studied approximation of the Heaviside function (see for example [3]) which turns out to be differentiable,

$$H(x) \approx H_s(x) := \frac{1}{2} + \frac{1}{\pi}\arctan\left(\frac{x}{s}\right),$$

where s is a parameter close to zero to control the steepness of the approximation. Then, the derivative of the Haar scaling function can be finally approximated by,

$$\phi'(x) \approx H'_s(x) - H'_s(x-1) = \frac{s}{\pi\left(x^2+s^2\right)} - \frac{s}{\pi\left((x-1)^2+s^2\right)}. \tag{27}$$

In what follows, we derive expressions for computing the ES, VaRC and ESC for the particular previous approximation of Haar wavelets.

Lemma 1. *Let $[\mathcal{I}_m^l(k), \mathcal{I}_m^u(k)]$ denote the support of $\phi_{m,k}(x)$, with $\mathcal{I}_m^l(k) := a + \frac{b-a}{2^m}k$ and $\mathcal{I}_m^u(k) := a + \frac{b-a}{2^m}(k+1)$. Then,*

$$\int_{VaR_\alpha(\boldsymbol{E})}^{b} x\phi_{m,k}\left(\frac{x-a}{b-a}\right)\mathrm{d}x = 2^{m/2}\begin{cases} h\left(\mathcal{I}_m^u(k)\right) - h\left(\max\left(VaR_\alpha(\boldsymbol{E}), \mathcal{I}_m^l(k)\right)\right), & \text{if } \mathcal{I}_m^u(k) > VaR_\alpha(\boldsymbol{E}), \\ 0, \quad \text{otherwise}, \end{cases}$$

where $h(x) := \frac{x^2}{2}$.

Proof. The proof follows immediately by observing that $\phi_{m,k}(x)$ has compact support on the bounded interval $[\mathcal{I}_m^l(k), \mathcal{I}_m^u(k)]$. □

Lemma 2. *Let $\mathcal{R}(x)$ be the ramp function with unit high defined as,*

$$\mathcal{R}(x) := \begin{cases} 0, & x < 0, \\ x, & 0 \le x < 1, \\ 1, & x \ge 1. \end{cases}$$

Then,

$$\int_a^x \phi_{m,k}\left(\frac{y-a}{b-a}\right)\mathrm{d}y = 2^{-m/2}(b-a)\mathcal{R}\left(2^m\left(\frac{x-a}{b-a}\right) - k\right).$$

Proof. The proof is immediate by distinguishing whether or not x belongs to the support of $\phi_{m,k}(x)$. □

Lemma 3. *Let us define $h_1(x) := \frac{2^m}{b-a}\frac{(x-a)^2}{2} - kx$ and $h_2(x) := x$. Then,*

$$\int_{VaR_\alpha(\boldsymbol{E})}^{b}\int_a^x \phi_{m,k}\left(\frac{y-a}{b-a}\right)\mathrm{d}y\mathrm{d}x$$
$$= \frac{b-a}{2^{m/2}}\begin{cases} h_1\left(\mathcal{I}_m^u(k)\right) - h_1\left(\max\left(VaR_\alpha(\boldsymbol{E}), \mathcal{I}_m^l(k)\right)\right) + h_2(b) - h_2\left(\mathcal{I}_m^u(k)\right), & \text{if } \mathcal{I}_m^u(k) > VaR_\alpha(\boldsymbol{E}), \\ h_2(b) - h_2\left(VaR_\alpha(\boldsymbol{E})\right), & \text{if } \mathcal{I}_m^u(k) \le VaR_\alpha(\boldsymbol{E}). \end{cases}$$

Proof. The proof directly follows from the use of Lemma 2. □

The ES is calculated by means of expression (19) and Lemma 1. The VaRC is obtained by formula (21) and Lemma 2 and the ESC by formula (23) and Lemma 3. We underline that $\frac{\partial \hat{c}_{m,k}}{\partial E_j}$ is computed from expression (25) by using $H'_s(x) - H'_s(x-1)$ instead of $\phi'(x)$.

3.3.2 Shannon wavelets

The Shannon scaling function reads,

$$\phi(x) = \mathrm{sinc}(x) = \begin{cases} \dfrac{\sin(\pi x)}{\pi x}, & \text{if } x \neq 0, \\ 1, & \text{if } x = 0, \end{cases}$$

where $\mathrm{sinc}(x)$ is usually called *cardinal sine* function. Further details on Shannon wavelets can be found in [4].

One of the advantages of using Shannon wavelets is that the scaling function is differentiable, allowing this way a direct computation of risk measures and risk contributions. More precisely, in Equation (25), we need the derivative of the scaling function which, in the context of Shannon wavelets, reads,

$$\phi'(x) = \begin{cases} \dfrac{\cos(\pi x)}{x} - \dfrac{\sin(\pi x)}{x^2\pi}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases} \tag{28}$$

In what follows, we derive expressions for computing the ES, VaRC and ESC for the particular case of Shannon wavelets.

Lemma 4. *Let x_1, x_2 be real numbers, $K_\gamma = \frac{2\gamma-1}{2^J}\pi$ and $\gamma, J, m, k \in \mathbb{Z}$. Then,*

$$I_1(x_1,x_2) = \int_{x_1}^{x_2} \cos\left(K_\gamma\left(2^m y - k\right)\right) \mathrm{d}y = \frac{1}{2^m K_\gamma}\left(\sin\left(K_\gamma(2^m x_2 - k)\right) - \sin\left(K_\gamma(2^m x_1 - k)\right)\right),$$

$$\begin{aligned} I_2(x_1,x_2) = \int_{x_1}^{x_2} y\cos\left(K_\gamma\left(2^m y - k\right)\right) \mathrm{d}y &= \frac{1}{\left(2^m K_\gamma\right)^2}\left(\cos\left(K_\gamma(2^m x_2 - k)\right) - \cos\left(K_\gamma(2^m x_1 - k)\right)\right) \\ &\quad + \frac{1}{2^m K_\gamma}\left(x_2\sin\left(K_\gamma(2^m x_2 - k)\right) - x_1\sin\left(K_\gamma(2^m x_1 - k)\right)\right), \end{aligned}$$

and,

$$I_3(x_1,x_2) = \int_{x_1}^{x_2} \sin\left(K_\gamma\left(2^m y - k\right)\right) \mathrm{d}y = \frac{1}{2^m K_\gamma}\left(\cos\left(K_\gamma(2^m x_1 - k)\right) - \cos\left(K_\gamma(2^m x_2 - k)\right)\right),$$

Proof. The result directly follows from integration by parts formula. □

Lemma 5. *Let $I_1(x_1,x_2)$ and $I_2(x_1,x_2)$ defined as in Lemma 4. Then,*

$$\int_{VaR_\alpha(\boldsymbol{E})}^{b} x\phi_{m,k}\left(\frac{x-a}{b-a}\right)\mathrm{d}x \approx \frac{2^{m/2}}{2^{J-1}}\sum_{\gamma=1}^{2^{J-1}}(b-a)\left(aI_1\left(\frac{VaR_\alpha(\boldsymbol{E})-a}{b-a},1\right) + (b-a)I_2\left(\frac{VaR_\alpha(\boldsymbol{E})-a}{b-a},1\right)\right).$$

Proof. Using the classical Vieta's formula [6], the cardinal sine can be expressed in terms of the following infinite product:

$$\phi(x) = \mathrm{sinc}(x) = \prod_{\gamma=1}^{+\infty}\cos\left(\frac{\pi x}{2^\gamma}\right). \tag{29}$$

If we truncate the infinite product in expression (29) to a finite product with J terms, then, thanks to the cosine product-to-sum identity in [19], we have,

$$\prod_{\gamma=1}^{J}\cos\left(\frac{\pi x}{2^\gamma}\right) = \frac{1}{2^{J-1}}\sum_{\gamma=1}^{2^{J-1}}\cos\left(\frac{2\gamma-1}{2^J}\pi x\right). \tag{30}$$

By Equations (29) and (30) the Shannon scaling function can thus be approximated as,

$$\phi(x) \approx \frac{1}{2^{J-1}}\sum_{\gamma=1}^{2^{J-1}}\cos\left(\frac{2\gamma-1}{2^J}\pi x\right). \tag{31}$$

Then, we obtain

$$\int_{\mathrm{VaR}_\alpha(\boldsymbol{E})}^{b} x\phi_{m,k}\left(\frac{x-a}{b-a}\right)\mathrm{d}x \approx \frac{2^{m/2}}{2^{J-1}}\sum_{\gamma=1}^{2^{J-1}}\int_{\mathrm{VaR}_\alpha(\boldsymbol{E})}^{b} x\cos\left(K_\gamma\left(2^m\left(\frac{x-a}{b-a}\right)-k\right)\right)\mathrm{d}x.$$

Finally, the expected result holds by applying the change of variables $y = \frac{x-a}{b-a}$ and Lemma 4. □

Remark 2. *The approximation of expression* (31) *is used throughout the paper. It is therefore of utmost importance the determination of integer J. An extensive error analysis is performed in [18]. Based on that analysis, we select the value $J = \lceil \log_2(2^m \pi) \rceil$, where $\lceil x \rceil$ denotes the smallest integer greater or equal than x and m is the scale of approximation. Once the scale is calculated (see Section 3.4 for details) then J remains fixed.*

Lemma 6. *With the same notation as before,*

$$\int_a^x \phi_{m,k}\left(\frac{y-a}{b-a}\right) \mathrm{d}y \approx \frac{2^{m/2}}{2^{J-1}} \sum_{\gamma=1}^{2^{J-1}} (b-a) I_1\left(0, \frac{x-a}{b-a}\right).$$

Proof. We use (31), then we make the change of variables $z = \frac{y-a}{b-a}$ and finally we apply Lemma 4. □

Lemma 7. *Let $I_3(x_1, x_2)$ defined as in Lemma 4. Then,*

$$\begin{aligned}
&\int_{VaR_\alpha}^b \int_a^x \phi_{m,k}\left(\frac{y-a}{b-a}\right) \mathrm{d}y \mathrm{d}x \\
&\approx \frac{b-a}{2^{J-1} 2^{m/2} K_\gamma} \sum_{\gamma=1}^{2^{J-1}} \left((b-a) I_3\left(\frac{VaR_\alpha(\boldsymbol{E}) - a}{b-a}, 1\right) + (b - VaR_\alpha(\boldsymbol{E})) \sin(K_\gamma \cdot k) \right).
\end{aligned}$$

Proof. The proof follows from using (31) and Lemma 4. □

The ES is calculated by means of expression (19) and Lemma 5. The VaRC is obtained by formula (21) and Lemma 6 and the ESC by formula (23) and Lemma 7. In the case of Shannon wavelets, we underline that $\frac{\partial \hat{c}_{m,k}}{\partial E_j}$ from expression (25) is computed using $\phi'(x)$ in (28).

3.4 Optimal scale of approximation m

As for any other density estimation algorithm, the free parameters need to be set in order to achieve an optimal convergence in the estimation. In this context the so-called *Mean Integrated Squared Error* (MISE) is the most commonly employed measure of the error. This error can be split into two terms, *bias* and *variance*. It is well established that the bias and variance present an opposite behaviour depending on the number of terms used in the approximation of the density, which in the present work is determined by the scale of approximation m. It is therefore essential to find a balance between these two errors to achieve the minimal overall error. The selection of m has been intensively studied in the literature, being our main reference Section 12.3 and Section 12.4 of [25]. We summarize here some relevant results that we take for our work.

The MISE is defined as,

$$\mathrm{MISE} = \int_{\mathbb{R}} \mathbb{E}\left[\left(\hat{f}_L(x; \boldsymbol{E}) - f_L(x; \boldsymbol{E}) \right)^2 \right] \mathrm{d}x,$$

where $\hat{f}_L$ is the estimated density and f_L is the true density function. In Section 12.4 of [25] the author introduces an expression for the difference in the MISE between two consecutive levels of resolution, i.e., the MISE at scale m (denoted by e_m) and the MISE at scale $m-1$ (denoted by e_{m-1}). We adapt here that expression, noting that we are using a linear wavelet estimator. Thus we have,

$$e_m - e_{m-1} = \frac{1}{n} \int_{-\infty}^{\infty} \phi_{m,k}^2(x) f_L(x; \boldsymbol{E}) \mathrm{d}x - \frac{n+1}{n} \sum_{k=0}^{\mathcal{K}} d_{m-1,k}^2, \tag{32}$$

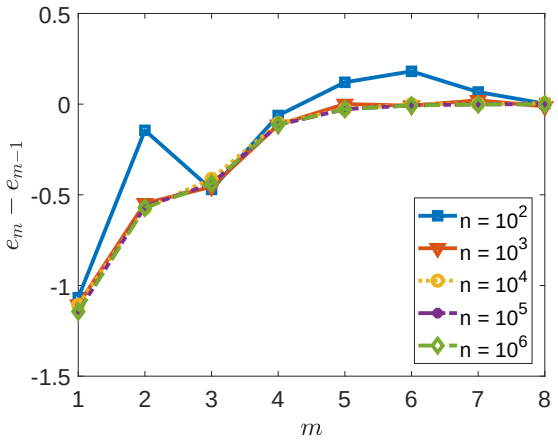


Figure 1: MISE difference. Setting: $\rho = 0.05, N = 20, P_j = 0.5, E_j = C/j$, where C is a constant chosen such that $\sum_{j=1}^{N} E_j = 1$.

where the coefficients $d_{m,k}$ are defined as $d_{m,k} = \mathbb{E}\left[\psi_{m,k}\left(Z\right)\right]$, being $\psi_{m,k}(x) = 2^{m/2}\psi(2^m x - k)$ and ψ the *mother wavelet*. The mother wavelet is the counterpart of the father wavelet ϕ according to a multi-resolution analysis. In 5, the expressions of ψ for Haar and Shannon wavelets are presented. Expression (32) can be therefore approximated, given the samples Z^i, by,

$$e_m - e_{m-1} \approx \frac{1}{n^2}\sum_{i=1}^{n}\sum_{k=0}^{\mathcal{K}}\phi_{m,k}^2\left(Z^i\right) - \frac{n+1}{n}\sum_{k=0}^{\mathcal{K}}\left(\frac{1}{n}\sum_{i=1}^{n}\psi_{m-1,k}\left(Z^i\right)\right)^2. \tag{33}$$

In Figure 1 we depict the approximated difference in the MISE given by (33), for an increasing scale m and several sizes of sample, n. In order to complete the analysis, we present the differences $e_m - e_8$ in Table 1. By observing both tests, we can see that, after reaching a certain scale of approximation, increasing m does not give any further improvement in the MISE. This fact suggests that an iterative algorithm can be developed to determine the optimal value for m. It has been suggested in [25] a recursive procedure to optimally choose the parameter m. This approach starts from a high value of m and decrease it until a big jump in the differences of the MISE is produced. Then, the scale m producing the jump is selected as the optimal one. However, no mechanism of how to choose the initial value of m and the size of the jump is provided. Furthermore, the evaluation of expression (33) for high values of m can be very expensive in terms of computational effort and this is a key aspect in our work.

	$e_m - e_8$							
m	7	6	5	4	3	2	1	0
$n = 10^2$	−0.0029	−0.0698	−0.2508	−0.3712	−0.3088	0.1615	0.3058	1.3730
$n = 10^3$	0.0124	−0.0088	−0.0004	−0.0011	0.1130	0.5686	1.1156	2.2211
$n = 10^4$	−0.0049	−0.0059	−0.0059	0.0209	0.1256	0.5370	1.1122	2.2190
$n = 10^5$	−0.0001	0.0010	0.0075	0.0351	0.1495	0.5886	1.1529	2.2973
$n = 10^6$	0.0005	0.0023	0.0097	0.0372	0.1496	0.5880	1.1593	2.3035

Table 1: MISE difference. Setting: $\rho = 0.05, N = 20, P_j = 0.5, E_j = C/j$, where C is a constant chosen such that $\sum_{j=1}^{N} E_j = 1$.

Following similar arguments as in [25], we present an slightly different approach that turns out to be more efficient and robust. Notice firstly that the computational cost of evaluating the expression in Equation (33) highly depends on the number of terms in the expansion $\mathcal{K}$ which, at the

same time, presents a dependence on m. Then, we propose to start the procedure from $m = 1$. Next, we can proceed by increasing the level of resolution m until some prescribed tolerance[1] in the difference $e_m - e_{m-1}$ is reached. We summarize in Algorithm 1 the optimal scale selection procedure.

Algorithm 1: Optimal selection of m.

Data: n, Z^i and ϵ_m
$m = 1$ `// Initial guess`
while $|\epsilon| > \epsilon_m$ **do**
  $m = m + 1$ `// Increasing` m `by one in each iteration`
  $\epsilon = e_m - e_{m-1}$ `// Compute the difference in the MISE by expression` (33)
return m

Once the optimal value for m has been selected, we summarize in Algorithm 2 the steps to follow for the computation of VaRC and ESC employing the approach presented in this work.

Algorithm 2: Computation of VaRC and ESC based on wavelet density estimation.

`// Portfolio configuration.`
Data: α, N, P_j, E_j.
`// Loss simulation settings.`
Data: n, d, ρ_j (when $d = 1$), $\boldsymbol{a}_j, b_j$ (when $d > 1$).
$[L^i, D_j^i] =$ **Samples_generation(**N**,** P_j**,** E_j**,** n**,** d**,** ρ_j**)**, $\forall i, \forall j$
$a = \min_{1\leq i\leq n}\left(L^i\right)$, $\quad b = \max_{1\leq i\leq n}\left(L^i\right)$, $\quad Z^i = \frac{L^i - a}{b-a}$, $\forall i$ `// Sample transformation.`
$m =$ **Optimal_scale(**Z^i**)** `// By Algorithm 1`
$\hat{c}_{m,k} =$ **Wavelet_coefficients(**Z^i**)**, `// By Equation` (14)
$\hat{f}_L(x; \boldsymbol{E}) =$ **Density_estimation(**$\hat{c}_{m,k}$**,** a**,** b**)** `// By Equation` (15)
$\hat{F}_L(x; \boldsymbol{E}) =$ **Distribution_estimation(**$\hat{c}_{m,k}$**,** a**,** b**)** `// By Equation` (16)
$\mathrm{VaR}_\alpha(\boldsymbol{E}) =$ **VaR_estimation(**α**,** $\hat{F}_L(x; \boldsymbol{E})$**)** `// By Equation` (17)
$\frac{\partial \hat{c}_{m,k}}{\partial E_j} =$ **Partial_wavelet_coefficients(**Z^i**,** D_j^i**)**, $\forall j$ `// By Equation` (25)
$\frac{\partial \hat{F}_L}{\partial E_j}(x; \boldsymbol{E}) =$ **Partial_distribution(**$\frac{\partial \hat{c}_{m,k}}{\partial E_j}$**)** $\forall j$ `// By Equation` (24)
$\mathrm{VaR}C_{\alpha,j} =$ **VaRC_estimation(**$\mathrm{VaR}_\alpha(\boldsymbol{E})$, $\frac{\partial \hat{F}_L}{\partial E_j}(x; \boldsymbol{E})$, $\hat{f}_L(x; \boldsymbol{E})$**)**, $\forall j$ `// By Equation` (21)
$\mathrm{ES}C_{\alpha,j} =$ **ESC_estimation(**$\mathrm{VaR}_\alpha(\boldsymbol{E})$, $\frac{\partial \hat{F}_L}{\partial E_j}(x; \boldsymbol{E})$**)**, $\forall j$ `// By Equation` (23)
return $\mathrm{VaR}C_{\alpha,j}$ and $\mathrm{ES}C_{\alpha,j}$

4 Numerical experiments

This section is devoted to test the technique presented in this work in the context of the credit risk decomposition (computation of the quantities $\mathrm{VaR}C_{\alpha,j}$ and $\mathrm{ES}C_{\alpha,j}$, $\forall j$), and to compare it against the classical approaches, based on crude Monte Carlo simulation, by assessing different important aspects. Particularly, we mainly focus on accuracy, robustness and efficiency of our methodology.

The experiments have been conducted in a computer system with the following characteristics: CPU Intel Core i7-4720HQ 2.6GHz and memory of 16GB RAM. The numerical codes have been implemented in C programming language. We have employed the GNU Scientific Library (GSL), specifically the packages to handle vectors and matrices, random number generation (Mersenne Twister MT19937), statistical distributions (Normal, t-student and χ^2), BLAS[2]

[1] For the experiments in this work, we set $\epsilon_m = 0.1$.
[2] Acronym of *Basic Linear Algebra Subprograms*.

operations (sparse matrix-vector multiplication) and root-finding algorithms (Brent-Dekker).

Two credit portfolio configurations are considered, presented in Table 2. The confidence level, α, is set to 99% for all the experiments. Both portfolios will be coupled with the use of various models given in Section 2.1. In Section 4.1, numerical experiments on a Gaussian one-factor model are conducted, taking advantage that reliable reference values are available. More complex multi-factor models are analysed in Section 4.2, where the gain in robustness provided by our method is highlighted. The computational performance is tested in Section 4.3.

Portfolio	N	P_j	E_j
P1	10000	0.08	$\frac{1}{j}$
P2	25000	0.05	$\frac{1}{j}$

Table 2: Portfolio configurations.

We assume that an initial sample set of size $n = 10^5$ for the loss distribution L is given, as well as the corresponding realizations of the default indicators $D_j, \forall j$. The VaR and ES values are computed from these samples, either directly as a quantile or by means of the recovered density. The computational cost of generating the samples and calculating the risk measures is not included in the reported times.

4.1 Experiments on one-factor model

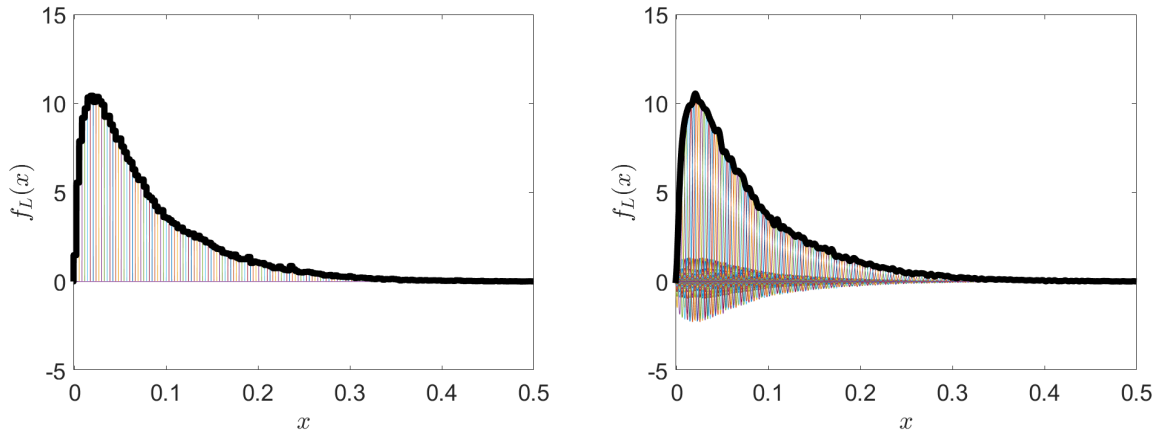
In this section, we consider the one-factor Gaussian copula model (2). The parameter ρ is set to 0.15 for portfolio P1 and 0.12 for P2. The t-copula model (3) will be used in its multi-factor version in Section 4.2.

4.1.1 Comparison: Haar versus Shannon

In this section, we compare the performance of the novel methodology presented here when it is used either with Haar or with Shannon wavelets. Our preliminary experiment consists of presenting the estimated densities of the loss random variable, L, for portfolio P1. We can see in Figure 2 the resulting densities for both Haar- and Shannon-based estimations. The individual scaling functions, $\phi_{m,k}(x), k = 1, \ldots, \mathcal{K}$, are also depicted, to give an insight of how the density is locally approximated by each of the wavelet bases. Note that, by construction, Haar basis always provides a positive density function, a desirable statistical property. That is not the case when employing Shannon basis since, due its trigonometric nature, negative values are theoretically achievable. In practical situations, given a sufficiently high number of samples, we can assume that Shannon-based estimation also keeps the positiveness in the density (i.e., the negative values are negligible).

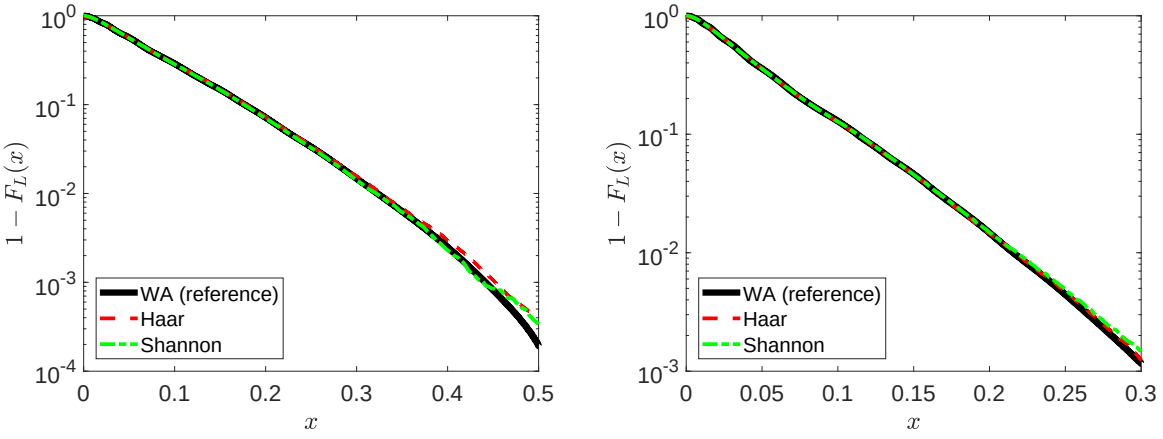
Next, we perform a numerical test on credit portfolio management, focusing especially on risk contributions calculation. As a reference, we will employ the wavelet-based closed-form solution introduced in [14], extended to the computation of credit risk contributions in [17]. This numerical method provides a highly accurate approximation for the case of the one-factor Gaussian model, and it is based on the inversion of the Laplace transform associated to the loss variable L. We will denote it here by WA method, which originally stands for *Wavelet Approximation*. In Figure 3 the tail probabilities obtained by our method are depicted. We observe that our approach employing both Haar and Shannon wavelets results in good approximations to the tail probability, for both Portfolios P1 and P2.

However, the main goal of this work is the credit risk contributions computation which is much more involved and some differences between Haar- and Shannon-based estimations arise and become clearer. The risk decomposition obtained by the Haar wavelet estimator is affected by the steepness parameter choice required in expression (27). We recall at this point that smaller s gives better approximations of the Heaviside function, i.e., this parameter must tend



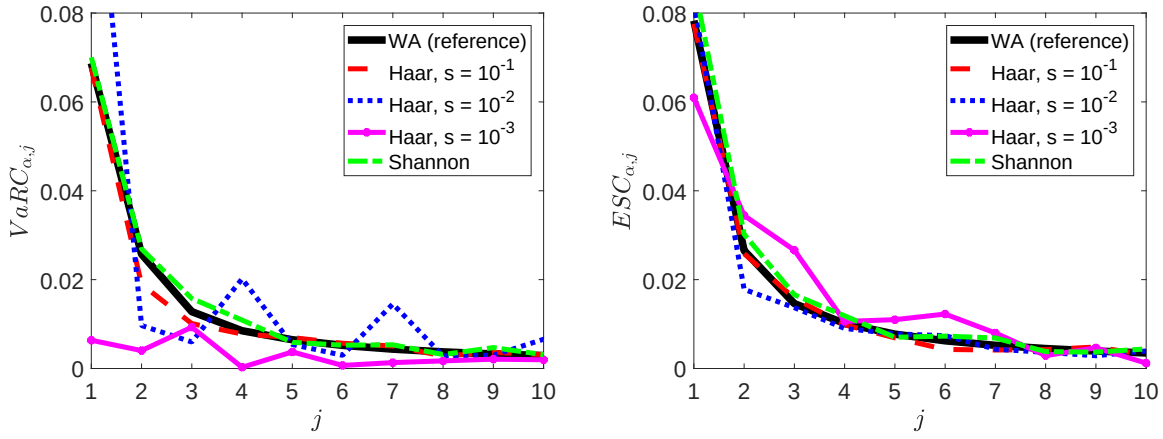
(a) Haar. (b) Shannon.

Figure 2: Estimation of the densities for portfolio P1 with Haar (left plot) and Shannon (right plot).



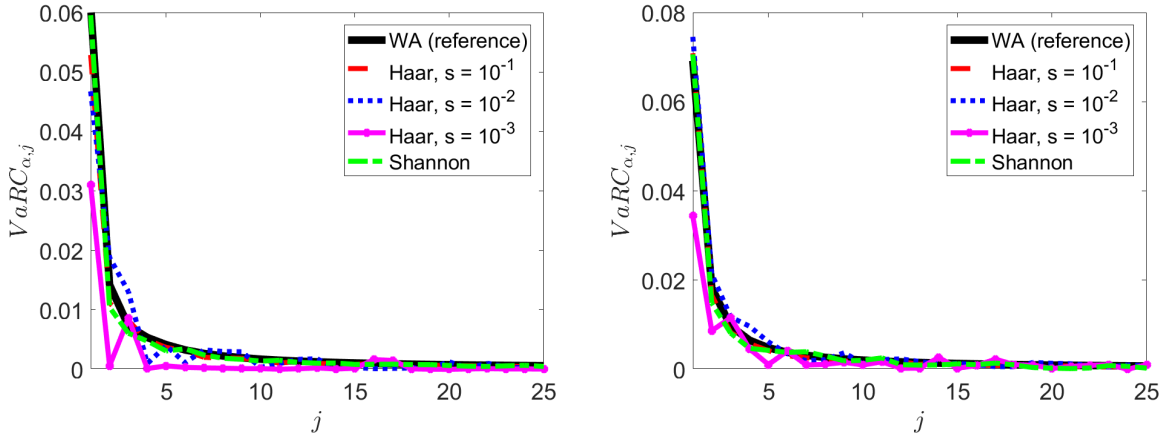
(a) Portfolio P1. (b) Portfolio P2.

Figure 3: Tail probabilities for portfolios P1 and P2.



(a) VaR contributions. (b) ES contributions.

Figure 4: Haar vs. Shannon - Portfolio P1: risk contributions ($j = 1, \ldots, 10$).



(a) VaR contributions. (b) ES contributions.

Figure 5: Haar vs. Shannon - Portfolio P2: risk contributions ($j = 1, \ldots, 25$).

to zero ($s \to 0$) in order to accurately approximate the Haar scaling function. However, the smaller s the worse is the approximation obtained to the risk contributions. Moreover, there is not a prescription on how to choose s to get better approximations. In contrast, the estimation by means of Shannon wavelets is highly accurate. All these facts are shown in Figure 4 where the risk contributions to VaR and ES for portfolio P1 are presented. For the sake of clarity, we only show the outcomes for the first ten obligors which, according to the portfolio configuration, have the bigger exposure, being of major interest. We want to point out the decreasing shape of the Shannon plot in accordance with the reference as well as with the decreasing exposures of obligors. The use of Shannon wavelets estimator turns to be more precise and stable in the approximations.

We carry out a similar experiment as above, but now on portfolio P2. We can see the first 25 VaRC and ESC obtained values in Figure 5. Again, the results employing Haar basis are rather unstable, specially in the case of low values of s. As before, Shannon basis gives more accurate and robust estimations.

We also provide the total sum of the individual risk contributions in Table 3 which, as established in (8), has to reproduce the risk measure itself, either the VaR or the ES. As expected, the use of Haar basis (for several choices of parameter s) results in poor and uncontrolled approxi-

mations while Shannon wavelet estimation preserves the allocation properties. For the reasons given above, from now on we discard Haar wavelets and use the Shannon family in the following experiments.

	Portfolio P1		Portfolio P2	
	$\sum \text{VaR}C_{\alpha,j}$	$\sum \text{ES}C_{\alpha,j}$	$\sum \text{VaR}C_{\alpha,j}$	$\sum \text{ES}C_{\alpha,j}$
WA ($m = 10$)	0.3227	0.3658	0.2153	0.2429
Haar ($s = 10^{-1}$)	0.2667	0.3440	0.1847	0.2315
Haar ($s = 10^{-2}$)	0.4016	0.4684	0.1762	0.2799
Haar ($s = 10^{-3}$)	0.6411	0.4236	0.0753	0.1599
Shannon	0.3236	0.3681	0.2091	0.2457

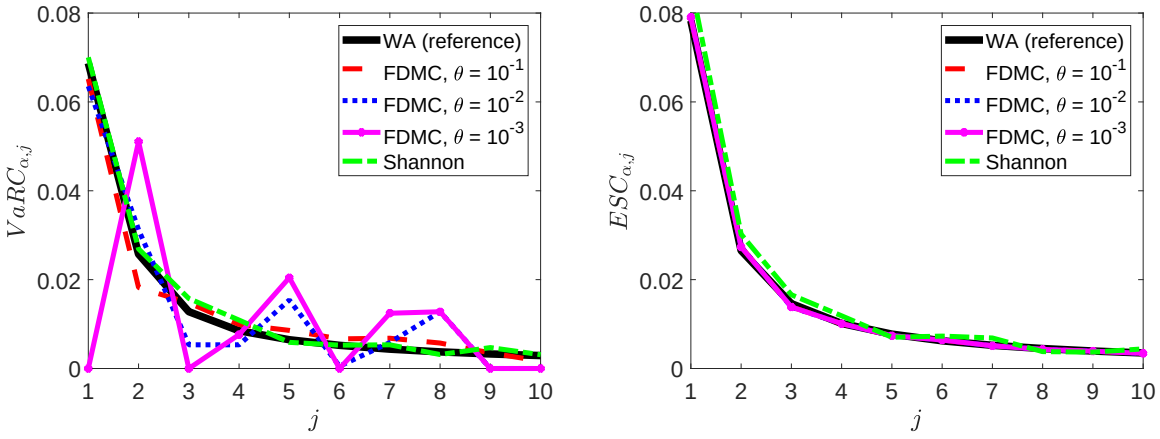
Table 3: Influence of the steepness parameter s, in the Haar-based data-driven approximation.

4.1.2 Comparison: Shannon versus Monte Carlo methods

Our methodology intends to be an alternative to pure Monte Carlo simulation-based techniques for the computation of the risk allocation problem, particularly in terms of efficiency and robustness. As mentioned, Monte Carlo simulation is rather inefficient due to its poor theoretical rate of convergence. According to the reasons given in Section 4.1.1, we choose Shannon basis for our next experiments. We therefore carry out an experiment comparing the outcomes provided by the Shannon wavelets density estimation approach presented in this work and a couple of plain Monte Carlo approaches in computing the risk contributions. As shown in Section 2.2, under the Euler's capital allocation principle, there are two equivalent ways of expressing the risk contributions, one in terms of partial derivatives, see (9), and other one in terms of expectations, see (10) and (11). Monte Carlo simulation can be employed on both equations, by introducing some numerical mechanisms to approximately solve them. As a first Monte Carlo based alternative, we consider the well-known *Finite Difference* numerical method[3], applied to the partial derivative appearing in expression (9). Given the samples of default indicators, $D_j^i, j = 1, \ldots, N, i = 1, \ldots, n$, and a proportional shift parameter θ[4], the risk decomposition can be addressed as presented in Algorithm 3, corresponding to the *central* finite difference version. We denote this approach as *FDMC*. As usual, the main issue of FDMC is the proper selection of θ, whose optimal value is not known a priori, i.e. it is problem dependent.

[3] In [16] for example, finite difference method was employed in the risk contributions computation context.

[4] The shift parameter θ is given in percentage, since the individual exposures can differ much (in size) one from each other, producing inaccurate estimations.



(a) VaR contributions.

(b) ES contributions.

Figure 6: FDMC vs. Shannon - Portfolio P1: risk contributions ($j = 1, \ldots, 10$).

Algorithm 3: Finite Difference Monte Carlo (FDMC) method.

Data: N, n, D_j^i and θ

for $j = 1, \ldots, N$ **do**

$E_j^{up} = (1+\theta)E_j$ `// Shifting up.`

$E_j^{dw} = (1-\theta)E_j$ `// Shifting down.`

for $i = 1, \ldots, n$ **do**

$L_{up}^i = D_j^i E_j^{up} + \sum_{k=1, k\neq j}^N D_j^i E_j^{up}$ `// Shifted up loss sample.`

$L_{dw}^i = D_j^i E_j^{dw} + \sum_{k=1, k\neq j}^N D_j^i E_j^{dw}$ `// Shifted down loss sample.`

$\boldsymbol{E}^{up} = [E_1, \ldots, E_j^{up}, \ldots, E_N]^T$ `// Shifted up vector of exposures.`

$\boldsymbol{E}^{dw} = [E_1, \ldots, E_j^{dw}, \ldots, E_N]^T$ `// Shifted down vector of exposures.`

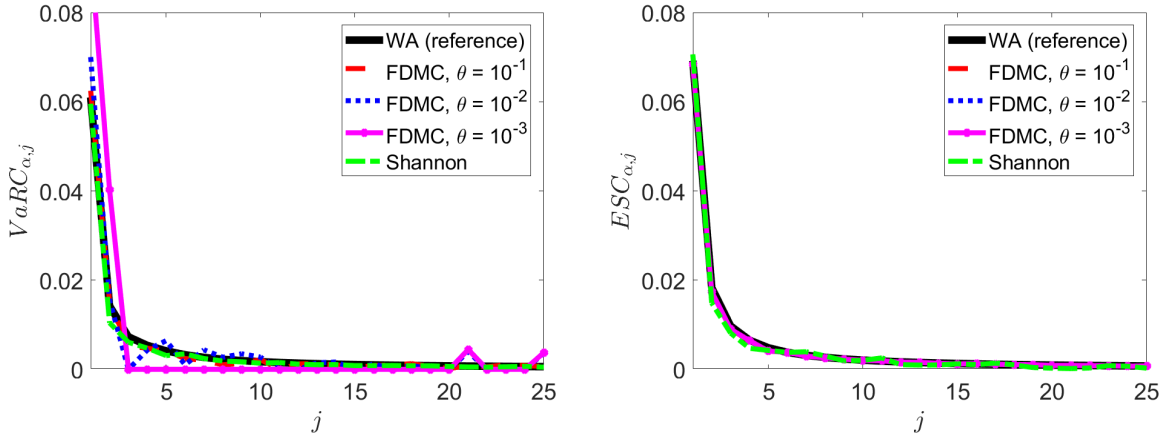
$\mathrm{VaR}C_{\alpha,j} \approx E_j \left(\frac{\mathrm{VaR}_\alpha(\boldsymbol{E}^{up}) - \mathrm{VaR}_\alpha(\boldsymbol{E}^{dw})}{2\theta E_j} \right) = \frac{\mathrm{VaR}_\alpha(\boldsymbol{E}^{up}) - \mathrm{VaR}_\alpha(\boldsymbol{E}^{dw})}{2\theta}$ `// FDMC` VaRC.

$\mathrm{ES}C_{\alpha,j} \approx E_j \left(\frac{\mathrm{ES}_\alpha(\boldsymbol{E}^{up}) - \mathrm{VaR}_\alpha(\boldsymbol{E}^{dw})}{2\theta E_j} \right) = \frac{\mathrm{ES}_\alpha(\boldsymbol{E}^{up}) - \mathrm{ES}_\alpha(\boldsymbol{E}^{dw})}{2\theta}$ `// FDMC` ESC.

return $\mathrm{VaR}C_{\alpha,j}$ and $\mathrm{ES}C_{\alpha,j}$, $\forall j$

In Figures 6 and 7, the FDMC contributions for several values of θ are presented. We observe that the precision provided by FDMC highly depends on θ, as expected. In general terms, the VaRC estimations by FDMC are unsatisfactory and worse than the ones provided by our method with Shannon wavelets. Note however that FDMC turns to be an acceptable candidate to compute the ESC, although it slightly violates the allocation principle, see Table 4. Furthermore, FDMC does not require any extra simulation, since it receives the initial sample set ($n = 10^5$) as an input, just like our wavelet density estimation based method. Nevertheless, as the FDMC procedure implies to reconstruct the loss distribution for each obligor, the technique becomes significantly inefficient in terms of computational effort, see Section 4.3.

Let us now propose a second Monte Carlo based alternative to compute the risk decomposition. We focus on the VaR contributions, since their calculation is rather involved for pure simulation methods, as seen above. According to (10), each obligor's contribution to the VaR is defined as an expectation, for whose estimation the plain Monte Carlo method is a natural candidate. Since the VaR is nothing else than a quantile (typically obtained at high confidence levels), the calculation of VaRC by crude Monte Carlo requires an artificial artifact to avoid an impractically huge number of simulations. As it is commonly done in the industry, we consider an interval around the VaR and we select the scenarios falling into that interval. To be more concrete, instead of using exactly the right hand side of expression (10), we instead compute



(a) VaR contributions.

(b) ES contributions.

Figure 7: FDMC vs. Shannon - Portfolio P2: risk contributions ($j = 1, \ldots, 25$).

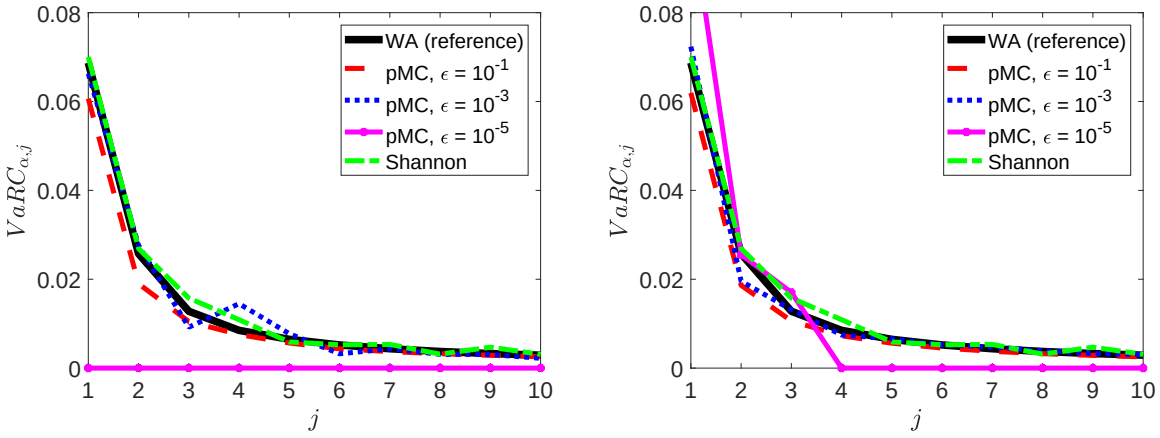
	Portfolio P1		Portfolio P2	
	$\sum \text{VaR}C_{\alpha,j}$	$\sum \text{ES}C_{\alpha,j}$	$\sum \text{VaR}C_{\alpha,j}$	$\sum \text{ES}C_{\alpha,j}$
WA ($m = 10$)	0.3227	0.3658	0.2153	0.2429
FDMC ($\theta = 10^{-1}$)	0.3706	0.3775	0.1979	0.2559
FDMC ($\theta = 10^{-2}$)	0.3667	0.3777	0.1902	0.2562
FDMC ($\theta = 10^{-3}$)	0.3266	0.3777	0.2087	0.2562
Shannon	0.3236	0.3681	0.2091	0.2457

Table 4: Influence of the shift parameter θ, in the FD-MC approximation.

$\mathbb{E}\left[L_j | L \in (\text{VaR}_\alpha(\boldsymbol{E}) - \epsilon, \text{VaR}_\alpha(\boldsymbol{E}) + \epsilon)\right]$, for several choices of ϵ. In accordance with [10], we will call this alternative as the *pseudo* Monte Carlo (pMC) method. Note that, since the VaR value is obtained from the initial sample set, pMC requires to perform an extra Monte Carlo simulation to preserve the statistical properties. As it will be shown in the numerical examples, the pMC estimator on $\mathbb{E}\left[L_j | L \in (\text{VaR}_\alpha(\boldsymbol{E}) - \epsilon, \text{VaR}_\alpha(\boldsymbol{E}) + \epsilon)\right]$ is quite sensitive to ϵ, which needs to be arbitrarily selected. Further, recalling that the VaR is a very extreme quantile, the number of generated scenarios falling around it might be insufficient to provide accurate estimations. We therefore include the approximations given by the pMC approach with an increasing number n_{MC} of Monte Carlo scenarios, at the cost of more computational effort, specially when a high number of factor are employed. They are compared with our methodology based on density estimation with Shannon wavelets, where it is worth recalling that the calculations have been carried out with the initial sample set of size $n = 10^5$.

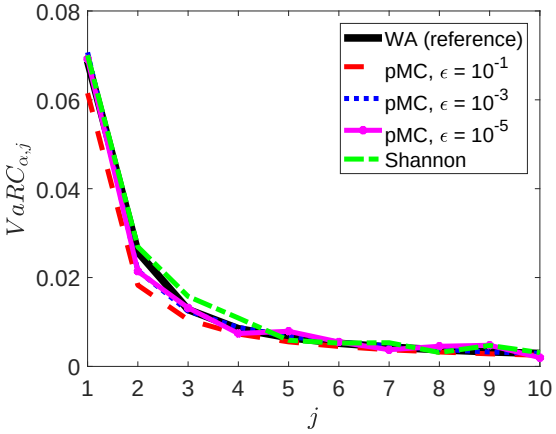
Next, we present the results of the pseudo Monte Carlo approach in computing the VaR contributions. First, in Figures 8 and 9, the VaRC curves are plotted. An straight line on the horizontal axis represents that the pMC method with the chosen configuration (n_{MC} and ϵ) was unable to provide results, i.e., no single scenario has fallen into the selected interval around the VaR. We can easily observe that the accuracy of the pMC estimations for the VaR contributions highly depends on a particular choice of n_{MC} and ϵ, resulting in rather imprecise approximations, even when we employ very high number of Monte Carlo simulations. When the length of the interval is too wide, the approximation is very smooth but tends to underestimate the true value. In contrast, if the interval is very narrow some peaks can appear, requiring higher number of scenarios to remove them. Note that the presence of peaks in the curves, also observed in the approximations provided by the FDMC approach, is not desirable, since it would mean that smaller exposures involve higher risks (and vice versa), and this fact is nonsensical

when the probabilities are equal across the whole portfolio. Thus, in order to produce accurate estimations employing Monte Carlo methods, an appropriate balance between n_{MC} and ϵ for pMC, or an optimal choice of θ for FDMC, are crucial aspects, for the selection of which no prior knowledge is available. In contrast, our wavelets-based technique overcomes most of these problems, producing stable and precise estimations obtained with an impressive reduction of the computational effort. This is due to the fact that the derivatives are computed analytically instead of numerically.



(a) $n_{\mathrm{MC}} = 10^5$.

(b) $n_{\mathrm{MC}} = 10^6$.



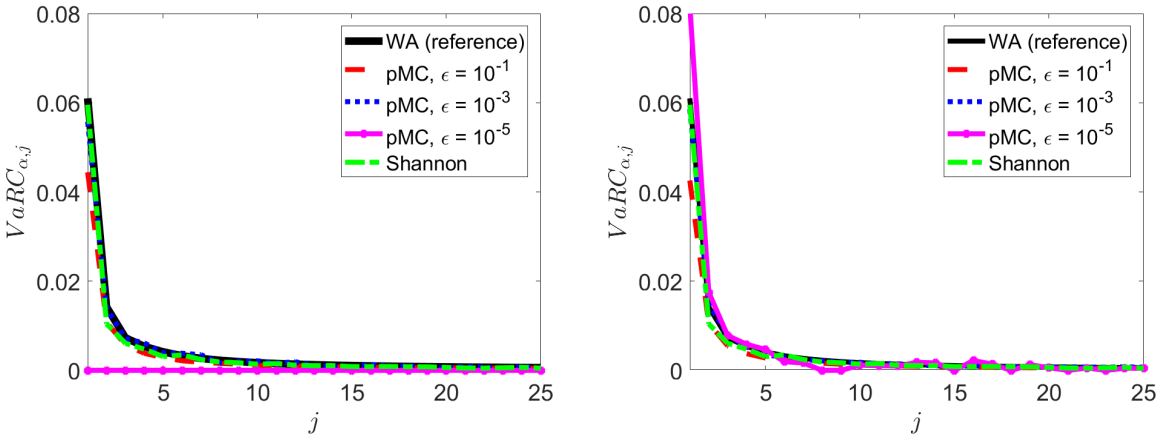
(c) $n_{\mathrm{MC}} = 10^7$.

Figure 8: VaR contributions (VaRC) with Monte Carlo varying ϵ. Portfolio P1.

Finally, we give in Table 5 the three smallest (MinRE) and the three biggest (MaxRE) relative errors when computing the contributions to VaR for portfolio P1. Again, the benchmark is the VaRC calculated with the WA method. As expected, our Shannon wavelets-based method provides accurate estimations. This experiment brings out that FDMC produces bigger errors, specially important when $\theta \to 0$. Generally speaking, pMC approach provides comparable precision as Shannon with $n_{\mathrm{MC}} = 10^7$ and $\epsilon \in \{10^{-3}, 10^{-5}\}$, but, as we will show in Section 4.3, at much higher computational cost.

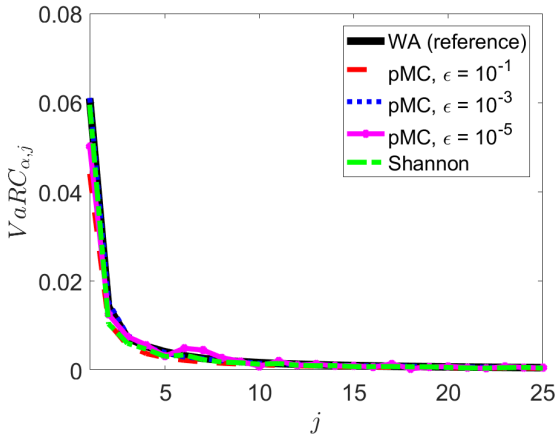
4.2 Experiments on multi-factor models

We consider the multi-factor models presented in Section 2.1.2. This type of models entails computational challenges from the numerical standpoint due to the curse of dimensionality. In accordance to the results presented in Section 4.1.2, the FDMC method is discarded.



(a) $n_{\mathrm{MC}} = 10^5$.

(b) $n_{\mathrm{MC}} = 10^6$.



(c) $n_{\mathrm{MC}} = 10^7$.

Figure 9: VaR contributions (VaRC) with Monte Carlo varying ϵ. Portfolio P2.

		Portfolio P1	
Method	Samples	MinRE	MaxRE
FDMC ($\theta = 10^{-1}$)	$n = 10^5$	$[4.95 \times 10^{-3}, 5.36 \times 10^{-3}, 8.34 \times 10^{-3}]$	$[2.90, 2.90, 2.90]$
FDMC ($\theta = 10^{-3}$)	$n = 10^5$	$[6.81 \times 10^{-3}, 6.83 \times 10^{-2}, 7.54 \times 10^{-2}]$	$[2.90, 2.90, 2.90]$
FDMC ($\theta = 10^{-3}$)	$n = 10^5$	$[1.02 \times 10^{-1}, 9.85 \times 10^{-1}, 1.00]$	$[2.90, 2.90, 2.90]$
pMC ($\epsilon = 10^{-1}$)	$n_{\text{MC}} = 10^5$	$[3.27 \times 10^{-2}, 3.31 \times 10^{-2}, 3.92 \times 10^{-2}]$	$[2.07 \times 10^{-1}, 2.19 \times 10^{-1}, 2.62 \times 10^{-1}]$
	$n_{\text{MC}} = 10^6$	$[9.02 \times 10^{-2}, 9.04 \times 10^{-2}, 9.09 \times 10^{-2}]$	$[1.43 \times 10^{-1}, 1.70 \times 10^{-1}, 2.75 \times 10^{-1}]$
	$n_{\text{MC}} = 10^7$	$[1.05 \times 10^{-1}, 1.15 \times 10^{-1}, 1.15 \times 10^{-1}]$	$[1.45 \times 10^{-1}, 1.79 \times 10^{-1}, 2.89 \times 10^{-1}]$
pMC ($\epsilon = 10^{-3}$)	$n_{\text{MC}} = 10^5$	$[3.83 \times 10^{-3}, 1.50 \times 10^{-2}, 1.90 \times 10^{-2}]$	$[1.00, 1.11, 1.21]$
	$n_{\text{MC}} = 10^6$	$[3.76 \times 10^{-4}, 1.13 \times 10^{-3}, 2.00 \times 10^{-3}]$	$[3.63 \times 10^{-1}, 3.78 \times 10^{-1}, 3.89 \times 10^{-1}]$
	$n_{\text{MC}} = 10^7$	$[4.03 \times 10^{-6}, 1.95 \times 10^{-5}, 3.24 \times 10^{-5}]$	$[1.49 \times 10^{-1}, 1.52 \times 10^{-1}, 1.62 \times 10^{-1}]$
pMC ($\epsilon = 10^{-5}$)	$n_{\text{MC}} = 10^5$	–	–
	$n_{\text{MC}} = 10^6$	$[7.50 \times 10^{-3}, 3.31 \times 10^{-1}, 4.86 \times 10^{-1}]$	$[2.90, 2.90, 2.90]$
	$n_{\text{MC}} = 10^7$	$[1.54 \times 10^{-4}, 4.63 \times 10^{-4}, 5.41 \times 10^{-4}]$	$[1.14, 1.26, 1.26]$
Shannon	$n = 10^5$	$[8.07 \times 10^{-5}, 1.03 \times 10^{-4}, 1.35 \times 10^{-4}]$	$[8.07 \times 10^{-1}, 8.09 \times 10^{-1}, 8.90 \times 10^{-1}]$

Table 5: VaRC: minimum and maximum relative errors: one-factor Gaussian copula model, portfolio P1.

The classically employed pMC simulation approach, besides the inconvenients already mentioned above, is further impacted by the model dimensionality which, in combination with an increasing number of simulations, makes its application very expensive. In contrast, the methodology presented here is model-free in the sense that, given the initial samples of the loss distribution (independently on the model used for generating them), we compute the VaRC (and ESC) values in an efficient, robust and accurate manner.

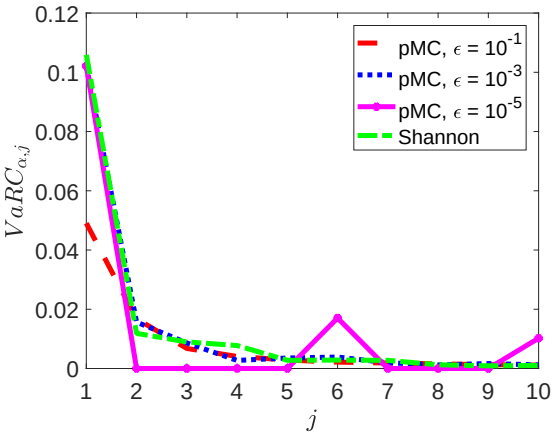
We consider first the multi-factor Gaussian copula model given by Equation (4). In terms of model dimensionality, two cases are considered, $d = 5$ and $d = 25$. The factor loadings $\boldsymbol{a_j}$ are randomly generated in both cases following a uniform distribution in $(0, 1)$. The obtained VaRC values (for portfolio P1) are depicted in Figure 10 (for $d = 5$) and Figure 11 (for $d = 25$), varying again the width of the interval by means of ϵ and the number of systematic scenarios n_{MC} for the pseudo Monte Carlo approach. Since no reference is available, the most reliable reference comes with the combination of a narrower interval ($\epsilon \in \{10^{-3}, 10^{-5}\}$) and very high number of scenarios ($n_{\text{MC}} = 10^7$). Taking this into account, the results confirm the excellent estimation provided by our methodology which matches almost perfectly that reference value.

We repeat the previous experiment but considering this time the multi-factor t-copula model described in Equation (5). The model parameters are selected in the same way as before, including in this case the degrees of freedom parameter, set to $\nu = 7$. In Figure 12 and Figure 13 the results are presented for dimensions $d = 5$ and $d = 25$, respectively. As expected, the wavelets-based approach performs very well compared to Monte Carlo.

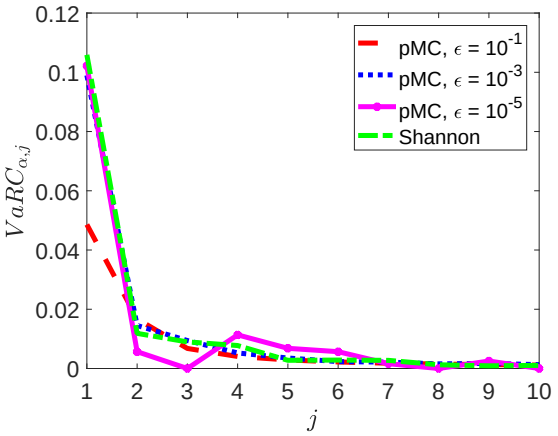
4.3 Computational performance

So far, it has been shown that the technique proposed in this work provides accurate and robust estimations. A third important component is the computational performance of the methodology with respect to the approaches relying on Monte Carlo. In simulation-based approaches, this is a key aspect since a desirable balance between precision and efficiency is achieved by reducing either the execution time given a prescribed accuracy or the approximation error given a computational budget. We therefore conduct an experiment to compare the computational performance of our method based on wavelets density estimation against the Monte Carlo methods. By analysing the previous experiments, we observe that, in order to get satisfactory approximations using pMC, the number of systematic scenarios needs to be set to at least $n_{\text{MC}} = 10^6$, requiring in many cases a higher amount for stability purposes.

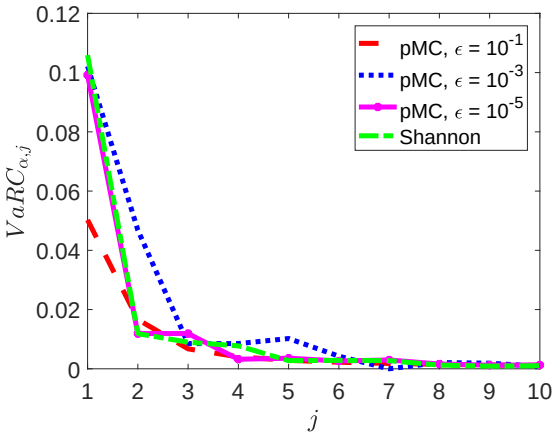
The execution times (given in seconds) and the scale of approximation m needed for different models are presented in Table 6 (one-factor Gaussian copula), Table 7 (multi-factor Gaussian



(a) $n_{\mathrm{MC}} = 10^5$.



(b) $n_{\mathrm{MC}} = 10^6$.



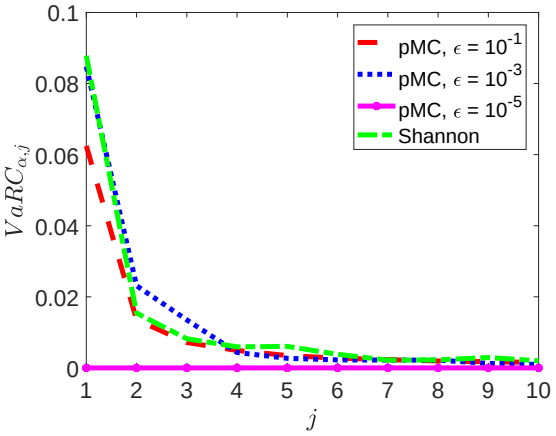
(c) $n_{\mathrm{MC}} = 10^7$.

Figure 10: Multi-factor Gaussian copula: VaR contributions portfolio P1 and $d = 5$.

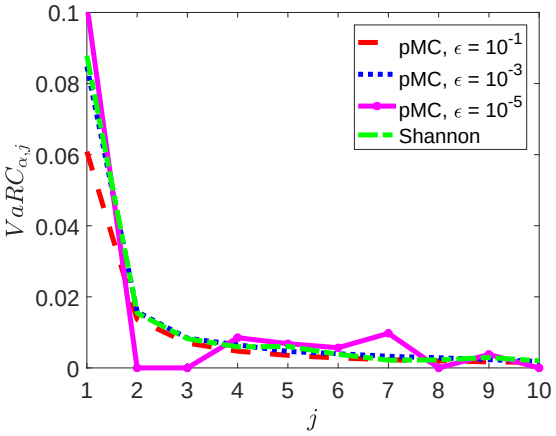
copula) and Table 8 (multi-factor t-copula). We also include the speed-up taken with respect to our technique. In all the cases, the presented method outperforms Monte Carlo-based methods, achieving an impressive speed-up in some particular situations, especially when multi-factor copula models are considered. It is worth pointing out that, as observed in Table 7 and Table 8, the CPU time employed with Shannon wavelets does not depend on the dimension d of the model considered (while pMC is highly dependent on the number of factors of the model). As mentioned, the performance of our technique is basically impacted by the scale parameter m, chosen from the MISE. As a general insight, we get higher m for sharper distributions or distributions including non-common features, like a significant mass at zero or several modes, barely appearing in practical situations. This is partially observed, for example, in the cases of the one-factor Gaussian model (Table 6) or the multi-factor t-copula (Table 8), which achieve relatively high m's. Still, the presented machinery based on wavelet density estimation provides a remarkable reduction of the computational cost with respect to Monte Carlo methods.

5 Conclusions

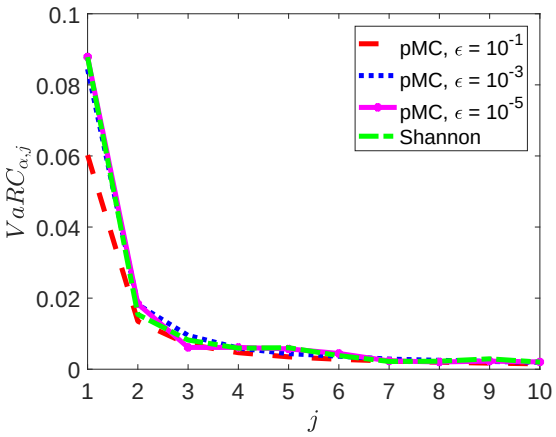
In this work, we have investigated the computation of risk contributions to VaR and ES in a credit portfolio by means of non-parametric density estimation based on wavelets. This problem is usually referred as *capital allocation* and it is of paramount importance for financial firms. It is a



(a) $n_{MC} = 10^5$.

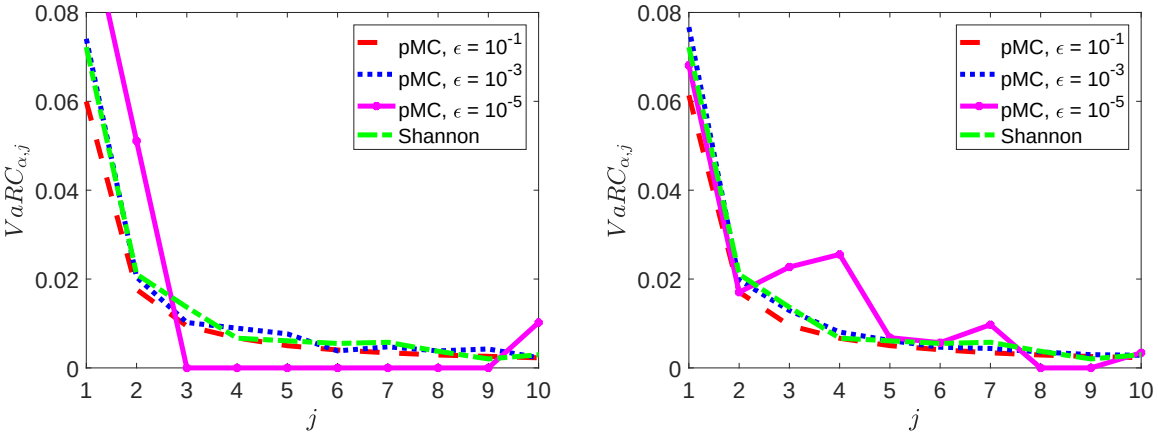


(b) $n_{MC} = 10^6$.



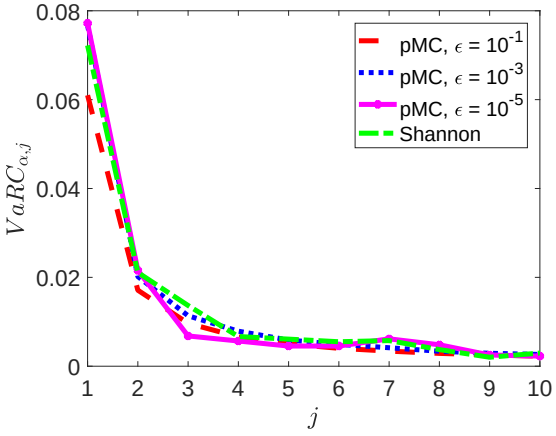
(c) $n_{MC} = 10^7$.

Figure 11: Multi-factor Gaussian copula: VaR contributions for portfolio P1 and $d = 25$.



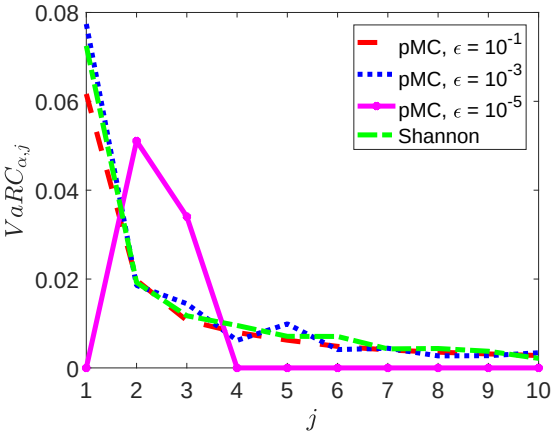
(a) $n_{MC} = 10^5$.

(b) $n_{MC} = 10^6$.

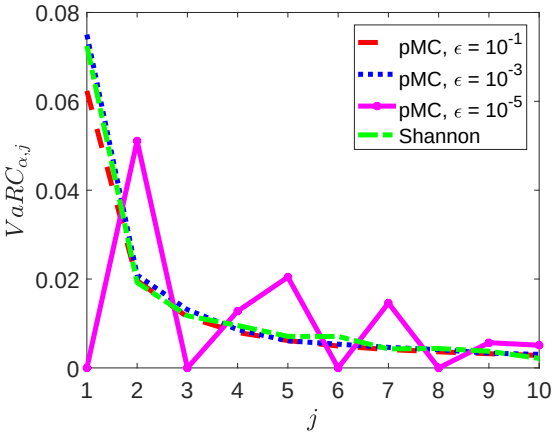


(c) $n_{MC} = 10^7$.

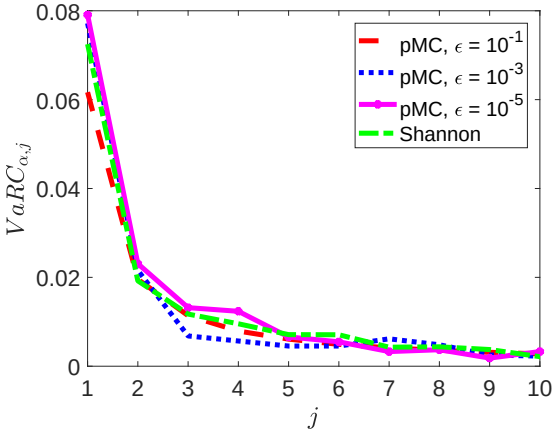
Figure 12: Multi-factor t-copula: VaR contributions for portfolio P1 and $d = 5$.



(a) $n_{MC} = 10^5$.



(b) $n_{MC} = 10^6$.



(c) $n_{MC} = 10^7$.

Figure 13: Multi-factor t-copula: VaR contributions for portfolio P1 and $d = 25$.

		Portfolio P1 ($m = 8$)		Portfolio P2 ($m = 9$)	
Method	Samples	Time	Speed-up	Time	Speed-up
Shannon	$n = 10^5$	212	$\times 1$	1257	$\times 1$
FDMC	$n = 10^5$	8646	$\times 41$	35122	$\times 28$
pMC	$n_{\mathrm{MC}} = 10^6$	1246	$\times 6$	3140	$\times 2$
pMC	$n_{\mathrm{MC}} = 10^7$	12397	$\times 58$	30636	$\times 24$

Table 6: Time and speed-up: one-factor Gaussian copula model.

		$d = 5$ ($m = 7$)		$d = 25$ ($m = 7$)	
Method	Samples	Time	Speed-up	Time	Speed-up
Shannon	$n = 10^5$	90	$\times 1$	91	$\times 1$
pMC	$n_{\mathrm{MC}} = 10^6$	3330	$\times 37$	9759	$\times 107$
pMC	$n_{\mathrm{MC}} = 10^7$	33260	$\times 370$	99252	$\times 1091$

Table 7: Time and speed-up: multi-factor Gaussian copula model. Portfolio P1.

common practice the use of Monte Carlo simulation to perform the calculation. However, Monte Carlo-based techniques turns out to be a very demanding method from the computational point of view, since once the loss random variable L has been simulated, the VaRC and ESC require the computation of either partial derivatives with respect to the corresponding risk measures, or expectations given that the total loss equals the VaR value (for VaRC) or is greater than the VaR value (for ESC). When combined with finite difference numerical derivation, Monte Carlo simulation entails the reconstruction of the entire loss distribution over the number of obligors. Moreover the estimations are inaccurate and unstable, specially in the case of the contributions to the VaR. If a direct application of Monte Carlo is considered, the expectations conditioned on a rare event have to be calculated by additional simulations, and this fact makes Monte Carlo methods prohibitive for large portfolios (which are typically encountered in real situations).

We propose a novel methodology based on non-parametric density estimation with wavelets. The starting point is the sample of the loss variable L generated by simulation. Then, the density is estimated either with Haar or Shannon wavelets. Finally, we take derivatives of the risk measures (VaR or ES) with respect to the exposures and we obtain the VaRC and ESC. While the Haar family has desirable properties like compact support and gives us positive densities, we finally prefer the Shannon family due to its robustness and easy handling.

We test our method with one- and multi-factor Gaussian and t-copula models. These models belong to the class of structural models and they are currently used in practice, since they are the models in force given by the regulators. The computation of VaRC by means of Monte Carlo simulation is particularly difficult due to the sensitivity of the results with respect to free parameters, i.e., the numerical derivative spacing or the length of the interval selected containing the VaR value. In contrast, the Shannon-wavelets based method gives impressive results, both in accuracy and speed. For obtaining comparable results in a reasonable accuracy, Monte Carlo needs between 25 (for the one-factor model) and 1000 times (for multi-factor models) the CPU time required by our method. Furthermore, while the Shannon-wavelets machinery is not affected by increasing the number of factors in the model, Monte Carlo needs three times more execution time when moving from 5 to 25 factors. Moreover, our proposed methodology is model-free in the sense that it stays the same and it applies in the same manner, regardless of the selected model.

To the best of our knowledge, this is the first time that this technique is used for solving the capital allocation problem by means of Euler's capital allocation principle. Multiple tests carried out along this work make us think that this novel method can be used within the risk management toolkit of financial firms.

		$d = 5$ $(m = 8)$		$d = 25$ $(m = 8)$	
Method	Samples	Time	Speed-up	Time	Speed-up
Shannon	$n = 10^5$	210	$\times 1$	213	$\times 1$
pseudo MC	$n_{\mathrm{MC}} = 10^6$	3181	$\times 15$	10376	$\times 49$
pseudo MC	$n_{\mathrm{MC}} = 10^7$	33135	$\times 158$	99932	$\times 469$

Table 8: Time and speed-up: multi-factor t-copula model. Portfolio P1.

Appendix A. Mother wavelet functions

Here we present the definition of the mother wavelet functions for both Haar and Shannon families. Thus, in the case of Haar basis, the mother wavelet reads,

$$\psi(x) := \begin{cases} 1, & 0 \leq x < \dfrac{1}{2}, \\ -1, & \dfrac{1}{2} \leq x < 1, \\ 0, & \text{otherwise}, \end{cases}$$

while the Shannon mother wavelet is defined as,

$$\psi(x) := \frac{\sin\left(\pi\left(x - \frac{1}{2}\right)\right) - \sin\left(2\pi\left(x - \frac{1}{2}\right)\right)}{\pi\left(x - \frac{1}{2}\right)} = 2\mathrm{sinc}(2x - 1) - \mathrm{sinc}(x).$$

Acknowledgments

Álvaro Leitao acknowledges financial support from the Spanish Ministry of Science and Innovation, through the María de Maeztu Programme for Units of Excellence in R&D (MDM-2014-0445).
Luis Ortiz-Gracia acknowledges the Spanish Ministry of Economy and Competitiveness for funding under grants ECO2016-76203-C2-2 and MTM2016-76420-P (MINECO/FEDER, UE).

References

[1] N. Aronszajn. Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68(3):337–404, 1950.

[2] Christian Bluhm, Ludger Overbeck, and Christoph Wagner. *Introduction to credit risk modeling*. Chapman & Hall/CRC, 2nd edition, 2010.

[3] Ronald N. Bracewell. *The Fourier transform and its applications*. Electrical Engineering. McGraw-Hill, 3rd edition, 2000.

[4] Carlo Cattani. Shannon wavelets theory. *Mathematical problems in Engineering*, 2008:164808, 2008.

[5] Ingrid Daubechies. *Ten lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992.

[6] William B. Gearhart and Harris S. Shultz. The function sin(x)/x. *The College Mathematics Journal*, 21(2):90–99, 1990.

[7] Paul Glasserman. Measuring marginal risk contributions in credit portfolios. *Journal of Computational Finance*, 9(2):1–41, 2005.

[8] Paul Glasserman and Jingyi Li. Importance sampling for portfolio credit risk. *Management Science*, 51(11):1643–1656, 2005.

[9] Michael Kalkbrener. An axiomatic approach to capital allocation. *Mathematical Finance*, 15(3):425–437, 2005.

[10] Takaaki Koike and Mihoko Minami. Estimation of risk contributions with MCMC, 2017.

[11] Alexandre Kurth and Dirk Tasche. Contributions to credit risk. *Risk Magazine*, pages 84–88, March 2003.

[12] Guangwu Liu. Simulating risk contributions of credit portfolios. *Operations Research*, 63(1):104–121, 2015.

[13] Eva Lütkebohmert. *Concentration risk in credit portfolios*. European Actuarial Academy (EAA) Lecture Notes. Springer-Verlag, 2009.

[14] Josep J. Masdemont and Luis Ortiz-Gracia. Haar wavelets-based approach for quantifying credit portfolio losses. *Quantitative Finance*, 14(9):1587–1595, 2014.

[15] Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative risk management: concepts, techniques and tools*. Princeton University Press, revised edition, 2015.

[16] Michael Olschewsky, Stefan Lüdemann, and Thorsten Poddig. Finite difference methods for estimating marginal risk contributions in asset management. *Journal of Risk*, 18(5):63–99, 2016.

[17] Luis Ortiz-Gracia and Josep J. Masdemont. Credit risk contributions under the Vasicek one-factor model: a fast wavelet expansion approximation. *Journal of Computational Finance*, 17(4):59–97, 2014.

[18] Luis Ortiz-Gracia and Cornelis W. Oosterlee. A highly efficient Shannon wavelet inverse Fourier technique for pricing European options. *SIAM Journal on Scientific Computing*, 38(1):118–143, 2016.

[19] Brendan M. Quine and Sanjar M. Abrarov. Application of the spectrally integrated Voigt function to line-by-line radiative transfer modelling. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 127:37–48, 2013.

[20] Marek Rutkowski and Silvio Tarca. Regulatory capital modeling for credit risk. *International Journal of Theoretical and Applied Finance*, 18(05):1550034, 2015.

[21] Dirk Tasche. Risk contributions and performance measurement. Technical report, Zentrum Mathematik (SCA), 2000.

[22] Dirk Tasche. Capital allocation for credit portfolios with kernel estimators. *Quantitative Finance*, 9(5):581–595, 2009.

[23] Marina Vannucci. Nonparametric density estimation using wavelets, 1995. Discussion Paper 95-26, ISDS, Duke University.

[24] Brani Vidakovic. *Statistical modeling by wavelets*. John Wiley & Sons Ltd, 1999.

[25] Gilbert G. Walter. *Wavelets and other orthogonal systems with applications*. CRC Press, Boca Raton, FL, 1994.