Journal of Computational and Applied Mathematics 292 (2016) 562–575

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Efficient wavelets-based valuation of synthetic CDO tranches

Luis Ortiz-Gracia *

Centre de Recerca Matemàtica, Campus de Bellaterra, Edifici C, 08193 Bellaterra (Barcelona), Spain

ARTICLE INFO

Article history:
Received 16 June 2014
Received in revised form 8 June 2015

MSC:
62P05
60E10
65T60

Keywords:
CDO valuation
Factor models
Characteristic function inversion
Haar wavelets
B-splines

ABSTRACT

We present new formulae for the valuation of synthetic collateralized debt obligation (CDO) tranches under a one-factor model. These formulae are based on the wavelet theory and the method used is called $\mathrm{WA}^{[a,b]}$. We approximate the cumulative distribution function (CDF) of the underlying pool by a finite combination of jth order B-spline basis, where the B-spline basis of order zero is typically called a Haar basis. We provide an error analysis and we show that for this type of distributions, the rate of convergence in the approximation is similar regardless of the order of the B-spline basis employed. The resulting formula for the Haar basis case is much easier to implement and performs better than the formula for the B-spline basis of order one in terms of computational time. The numerical experiments confirm the impressive speed and accuracy of the $\mathrm{WA}^{[a,b]}$ method equipped with a Haar basis, independently of the inhomogeneity features of the underlying pool. The method appears to be particularly fast for multiple tranche valuation.

1. Introduction

The last financial crisis has shown that one of the major source of problems for financial institutions was the credit risk management. The credit derivatives market was the most innovative and fastest growing derivative market during the past ten years. The rapid development was due to new possibilities that were offered by credit derivatives. Credit instruments are flexible financial products that enable the efficient repackaging and transfer of credit risk. Credit derivatives are attractive for yield seeking investors and banks that need to hedge their investments and fulfil the capital requirements. The most popular securities traded on open markets are credit default swaps (CDS), default baskets, and CDOs. A traditional CDO is a credit derivative security whose underlying collateral is a portfolio of risky bonds or bank loans. A synthetic CDO is a credit derivative security whose underlying collateral is a portfolio (or *pool*) made up of CDS. In this paper, we focus our attention on synthetic CDOs and will call them simply CDOs.

To offset the pool owner's risk from these default swaps, a portion of the premiums from them is allocated to a collection of securities called *tranches* of the CDO. There is a priority scheme for the tranches to absorb the pool losses up to fixed maximum amounts for each tranche. Losses are based on the recovery adjusted CDS notional values. The *Equity* tranche is the first to absorb the pool losses. After the Equity tranche is exhausted, losses will affect the *Mezzanine* tranches, and finally the *Senior* tranches. Investors take on exposure to a particular tranche, effectively selling credit protection to the CDO issuer, and in turn, collecting the premium.

Valuation of CDOs is an important problem in credit risk management which requires efficient pricing methods. One issue common to CDO pricing and related risk management is how to evaluate the pool's loss distribution efficiently. From a computational point of view, Monte Carlo simulation (MC) is the last resort because of its inefficiency, despite its flexibility.

* Tel.: +34 935868515.
E-mail address: lortiz@crm.cat.

http://dx.doi.org/10.1016/j.cam.2015.07.025
0377-0427/

Widely used methods can be divided into two classes. The first class evaluates a pool's loss distribution exactly, based on the assumption that all obligors' losses-given-default sit on a common lattice. Among these methods are the ABS method by [1], the HW method proposed by [2], the LG method developed by [3] and the JKM method by [4]. The ABS and LG methods are directly applicable to inhomogeneous pools. Although the HW and JKM methods are directly applicable to homogeneous pools only, they can be applied indirectly to inhomogeneous pools by noting that in practice an inhomogeneous pool can usually be partitioned into a small number of homogeneous pools. A closed-form solution is derived in [5] in the case of homogeneous instruments. The second class of methods evaluates a pool's loss distribution approximately. An example of this class is the compound Poisson approximation method (CPA) by [6]. In [4] the authors present an improved compound Poisson approximation method based on [7] to enhance the accuracy of the basic CPA [6]. Both the first class of methods and the compound Poisson approximations, strongly rely on the assumption that the loss-given-default must sit on common lattice. To alleviate this deficiency, [4] introduce the normal power (NP) approximation method, widely used in actuarial science, to approximate the pool's loss distribution. However, this approximation works out well for a large pool as a consequence of the central limit theorem, but it may not capture some important properties such as skewness and fat tails. The numerical experiments carried out in [4] on a wide variety of pools, show that JKM method is always faster than HW and much faster than ABS. For most cases JKM is faster than CPA but slower than NP. Another numerical method belonging to the second class of methods is the saddle point (SP) method by [8], which allows the computation of fully inhomogeneous portfolios. It is well known from the literature that SP method fits the tail of the distributions particularly well, however, as pointed out by [9], the method cannot deal properly with highly concentrated portfolios arising from an unequal distribution of the adjusted notional values. This deficiency may affect the Senior tranches in a CDO pricing problem. Yet another approach called EAP method, presented in [10,11], uses a representation of the hockey stick function to directly approximate tranche prices. The authors approximate the payoff function by a sum of exponentials over the positive real line and consequently they do not need to compute the distribution of losses. As pointed out in [11], EAP method is slower than JKM when dealing with very homogeneous pools. The ratio of computation time between multiple tranches evaluation and single tranche evaluation on a single pool, increases with the size of the portfolio for both JKM and EAP method, although it is significantly higher for EAP (see [10] for details). A method based on Laplace transform inversion is presented in [12] within the multifactor model framework.

In this work, we focus on the recovery of the pool's CDF from its characteristic function by means of a wavelets-based method. This method was originally developed within a credit risk environment to recover a CDF on a bounded domain from its Laplace transform by means of a Haar basis (see [13]). The method was extended in [14] to invert Fourier transforms over the entire real line with B-splines up to order one. Later on, it was applied to an option pricing problem in [15] and it was called $\mathrm{WA}^{[a,b]}$ method. We aim to use this machinery to efficiently price a CDO, assuming that the correlation structure of default events is described by a one-factor model as in the literature. To this end, we approximate the CDF of the underlying pool first by a finite combination of Haar wavelet functions and second by a finite combination of B-spline wavelets of order one, where the Haar basis can be seen as a B-spline basis of order zero. It is well known that under a one-factor model, the resulting CDF is a staircase like function. For this class of functions, we state a proposition showing that we reach a similar rate of convergence regardless of the order of the B-splines employed. Furthermore, the resulting formula for the Haar basis case is much easier to implement and performs better than the formula for the B-spline basis in terms of computational time. To test this new methodology we carry out the numerical experiments under the particular case of the one-factor Gaussian copula model, although any other one-factor Lévy model can be accommodated (as for example in [16,17]) within this approach. These numerical examples show the robustness and confirm the efficiency of the $\mathrm{WA}^{[a,b]}$ method, independently of the inhomogeneity features of the underlying pools. This pricing method is capable to obtain the price of a single tranche of a CDO in about one tenth of a second with a relative error less than one percent. One of the main findings of this work is the ability of this methodology to price multiple tranches of a CDO without adding significant extra computational time, in contrast with JKM and EAP methods.

The rest of the paper is organized as follows. In Section 2 we present the pricing formulae and the default model. In Section 3 we give an overview of some key aspects of the wavelet theory, we briefly explain the $\mathrm{WA}^{[a,b]}$ method and we compute the conditional expected cumulative losses. We also provide an error analysis. Section 4 is devoted to the numerical experiments and finally Section 5 concludes.

2. The pricing framework

We consider a synthetic CDO tranche of size S with an attachment point l, a threshold that determines whether some of the pool losses will be absorbed by this tranche. If the realized losses of the pool are less than l, then the tranche will not suffer any loss, otherwise it will absorb losses up to its size S. The threshold $S + l$ is called the detachment point of the tranche.

We assume there are $\mathcal{K}$ names in the pool underlying the CDO. For name κ, its notional value and the recovery rate of the notional value of the reference asset are N_κ and R_κ, respectively. Then the loss-given-default or the recovery adjusted notional value of name κ is, $L_\kappa = N_\kappa(1 - R_\kappa)$. Let $0 = t_0 < t_1 < \cdots < t_n = T$ be the set of premium dates, where T denotes the maturity of the CDO tranche. Assume that the interest rates are deterministic. Let $\mathcal{L}_i$ be the pool's cumulative losses up to time t_i. Then, the losses absorbed by the specified tranche up to time t_i, denoted by $\mathcal{L}_i$ is, $\mathcal{L}_i = g(\mathcal{L}_i; l, S + l) = \min(S, (\mathcal{L}_i - l)^+)$, where $x^+ = \max(x, 0)$.

We assume that the fair *spread* for the tranche is a constant s per annum. The two important quantities to be determined in synthetic CDO tranche valuation are the present value of the *default leg*, that is, the expected losses of the tranche over the life of the contract, and the present value of the *premium leg*, i.e., the expected premiums that the tranche investor will receive over the life of the contract. Mathematically, the two leg's present values are,

$$\text{PV (premium leg)} = \mathbb{E}\left[\sum_{i=1}^{n} s(S-\mathscr{L}_i)(t_i-t_{i-1})d_i\right],$$

$$\text{PV (default leg)} = \mathbb{E}\left[\sum_{i=1}^{n} (\mathscr{L}_i-\mathscr{L}_{i-1})d_i\right],$$

where PV denotes present value and $d_i,\ i=1,\dots,n$ are the set of discount factors. These factors are deterministic, since we are assuming deterministic interest rates as well. Therefore, the fair spread s is,

$$s = \frac{\mathbb{E}\left[\sum_{i=1}^{n}(\mathscr{L}_i-\mathscr{L}_{i-1})d_i\right]}{\mathbb{E}\left[\sum_{i=1}^{n}(S-\mathscr{L}_i)(t_i-t_{i-1})d_i\right]}, \tag{1}$$

where $\mathbb{E}[\mathscr{L}_0]=0$, due to the assumption that there is no default at time t_0. Once the value of the spread is known, then the value of the tranche to the tranche investor today is,

$$s\cdot\text{PV (premium leg)} - \text{PV (default leg)}. \tag{2}$$

Noting that the tranche size S is constant and the time periods $t_i - t_{i-1}$ are fixed, it follows from (1) and (2) that the valuation problem is now reduced to the computation of the expected cumulative losses $\mathbb{E}[\mathscr{L}_i],\ i=1,\dots,n$.

2.1. The default model

In order to compute the expectations $\mathbb{E}[\mathscr{L}_i],\ i=1,\dots,n$, we have to specify the default processes for each of the names and the correlation structure of the default events. To this end, we consider a one-factor model under the conditional independence framework.

We assume that the risk-neutral default probabilities $\pi_\kappa(t)=\mathbb{P}(\tau_\kappa<t),\ \kappa=1,\dots,\mathcal{K}$, that describe the default-time distributions of all $\mathcal{K}$ names are available, where τ_κ is the default time of name κ. The dependence structure of the default times is determined by the creditworthiness indexes Y_κ through a one-factor model. The indexes Y_κ are defined by,

$$Y_\kappa = \sqrt{\rho_\kappa}X + \sqrt{1-\rho_\kappa}\epsilon_\kappa,\quad \kappa=1,\dots,\mathcal{K}, \tag{3}$$

where X is the systematic risk factor, ϵ_κ are the idiosyncratic factors that are mutually independent and are also independent of X, and ρ_κ captures the correlation between Y_κ and the single risk factor X. Conditional on a given value x of X, all defaults events are independent. The risk-neutral default probabilities and the creditworthiness indexes are related by $\pi_\kappa(t)=\mathbb{P}(\tau_\kappa<t)=\mathbb{P}(Y_\kappa<H_\kappa(t))$, where $H_\kappa(t)$ is the default barrier of the κth name at time t, i.e., $H_\kappa(t)=\Phi^{-1}(\pi_\kappa(t))$, with Φ the cumulative distribution function of the random variable ϵ_κ.

The *conditional* risk-neutral default probabilities are defined by,

$$\pi_\kappa(t;x)=\mathbb{P}(Y_\kappa<H_\kappa(t)|X=x). \tag{4}$$

Therefore, from (3) and (4) we have,

$$\pi_\kappa(t;x)=\Phi\left(\frac{H_\kappa(t)-\sqrt{\rho_\kappa}x}{\sqrt{1-\rho_\kappa}}\right).$$

Within the conditional independence framework, the expected cumulative losses $\mathbb{E}[\mathscr{L}_i]$ can be computed as,

$$\mathbb{E}[\mathscr{L}_i]=\int_{\mathbb{R}}\mathbb{E}[\mathscr{L}_i|X=x]\varphi(x)dx, \tag{5}$$

where $\mathbb{E}[\mathscr{L}_i|X=x]=\mathbb{E}\left[g(\mathcal{L}_i;l,S+l)|X=x\right]=\mathbb{E}\left[\min(S,(\mathcal{L}_i-l)^+)|X=x\right]$ is the expectation of $\mathscr{L}_i$ given $X=x$, $\mathcal{L}_i=\sum_{\kappa=1}^{\mathcal{K}}L_\kappa\mathbf{1}_{\{Y_\kappa<H_\kappa(t_i)\}}$ and φ is the probability density function of the random variable X. Therefore, the fundamental numerical challenge in synthetic CDO tranche pricing, is how to evaluate efficiently the conditional expectation $\mathbb{E}[\mathscr{L}_i|X=x]$ for a fixed abscissa x. Under the conditional independence framework, the default indicators $\mathbf{1}_{\{Y_\kappa<H_\kappa(t_i)\}}$ are independent and therefore the characteristic function of $\mathcal{L}_i$ given $X=x$ can be readily obtained. Finally, the conditional expectation $\mathbb{E}[\mathscr{L}_i|X=x]$ is computed by inverting its characteristic function. We give all the necessary details in Section 3.4.

The integration in (5) need to be approximated numerically using a quadrature rule. As stated in the introduction section, we test the wavelets-based method under the one-factor Gaussian copula model, and therefore X and ϵ_κ are

normally distributed and Φ and φ are, respectively, the distribution and the density function of the standard normal distribution. For this reason, and without loss of generality, we use a Gauss–Hermite quadrature rule with M nodes. If we note $h(x) = \mathbb{E}[\mathscr{L}_i|X = x]$ and make the change of variables $u = x/\sqrt{2}$ in (5), gives us,

$$\mathbb{E}[\mathscr{L}_i] = \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} h(\sqrt{2}u) e^{-u^2} du \simeq \frac{1}{\sqrt{\pi}} \sum_{m=1}^{M} w_m h(\sqrt{2}u_m), \tag{6}$$

where w_m, u_m, $m = 1, \ldots, M$, are the weights and nodes respectively, corresponding to a Gauss–Hermite quadrature of M points.

Guégan and Houdain propose in [17] a factor model based on a Normal Inverse Gaussian (NIG) distributed common factor but with standard normal idiosyncratic risks. In that case Φ remains the same, while φ is NIG-distributed (see [18]) and a more general quadrature rule can be used to numerically integrate (5).

3. Conditional expectation computation: the WA$^{[a,b]}$ method

This section is devoted to the computation of the conditional expected losses $\mathbb{E}[\mathscr{L}_i|X = x]$ in (5). We carry out the approximation by means of the WA$^{[a,b]}$ method based on a family of wavelets called B-spline and we provide an error analysis.

A natural and convenient way to introduce wavelets is following the notion of multiresolution analysis (MRA). Here we provide the basic definitions and properties regarding MRA and B-spline wavelets, for further information see [19–22].

3.1. Wavelets and dual wavelets

We start with the definition of a Riesz basis, as follows,

Definition 1. A countable set $\{f_n\}$ of a Hilbert space is a Riesz basis if every element f of the space can be uniquely written as $f = \sum_n c_n f_n$, and there exist positive constants A and B such that,

$$A\|f\|^2 \leq \sum_n |c_n|^2 \leq B\|f\|^2.$$

Definition 2. A function $\psi \in L^2(\mathbb{R})$ is called an $\mathcal{R}$-function if $\{\psi_{j,k}\}$ defined as $\psi_{j,k}(x) := 2^{j/2}\psi(2^j x - k)$, $j, k \in \mathbb{Z}$, is a Riesz basis of $L^2(\mathbb{R})$.

If we assume that ψ is an $\mathcal{R}$-function, then there exists a unique Riesz basis $\{\psi^{l,m}\}$ of $L^2(\mathbb{R})$, which is dual to $\{\psi_{j,k}\}$ in the sense that $\langle \psi_{j,k}, \psi^{l,m} \rangle = \delta_{j,l}\delta_{k,m}$, for all $j, k, l, m \in \mathbb{Z}$, and $\delta_{p,q}$ is the Kronecker delta defined in the usual way as,

$$\delta_{p,q} = \begin{cases} 1, & p = q, \\ 0, & \text{otherwise.} \end{cases}$$

With the above definitions, we can give the definition of wavelets.

Definition 3. An $\mathcal{R}$-function $\psi \in L^2(\mathbb{R})$ is called an $\mathcal{R}$-wavelet, or simply a wavelet, if there exists a function $\widetilde{\psi} \in L^2(\mathbb{R})$, such that $\{\psi_{j,k}\}$ and $\{\widetilde{\psi}_{j,k}\}$ defined as $\widetilde{\psi}_{j,k}(x) = 2^{j/2}\widetilde{\psi}(2^j x - k)$, $j, k \in \mathbb{Z}$, are dual bases of $L^2(\mathbb{R})$. If ψ is a wavelet, then $\widetilde{\psi}$ is called a dual wavelet corresponding to ψ.

Some definitions relevant to our present work are given below,

Definition 4. Let ψ be the wavelet function in Definition 3.

(i) A wavelet ψ is said to have a vanishing moment of order s if $\int_{\mathbb{R}} x^p \psi(x)dx = 0$, $p = 0, \ldots, s-1$. All wavelets must satisfy this condition for $p = 0$.
(ii) A wavelet ψ is an orthogonal (ON) wavelet if the family $\{\psi_{j,k}\}$ forms an orthonormal basis of $L^2(\mathbb{R})$, that is, $\langle \psi_{s,t}, \psi_{u,v} \rangle = \delta_{s,u}\delta_{t,v}$, for all $s, u, t, v \in \mathbb{Z}$.
(iii) A wavelet ψ is called a semi-orthogonal (SO) wavelet if the family $\{\psi_{j,k}\}$ satisfies, $\langle \psi_{s,t}, \psi_{u,v} \rangle = 0, s \neq u$, for all $s, u, t, v \in \mathbb{Z}$.

A dual wavelet $\widetilde{\psi}$ is unique and is itself a wavelet. The pair $(\psi, \widetilde{\psi})$ is symmetric in the sense that ψ is also the dual wavelet of $\widetilde{\psi}$. If ψ is an orthogonal wavelet, then it is self-dual in the sense of $\widetilde{\psi} \equiv \psi$. Moreover, it is important to emphasize that if $f \in L^2(\mathbb{R})$ then,

$$f = \sum_{j,k=-\infty}^{+\infty} \langle f, \psi_{j,k} \rangle \widetilde{\psi}_{j,k} = \sum_{j,k=-\infty}^{+\infty} \langle f, \widetilde{\psi}_{j,k} \rangle \psi_{j,k}.$$

For each $j \in \mathbb{Z}$, let W_j denote the closure of the linear span of $\{\psi_{j,k} : k \in \mathbb{Z}\}$, namely, $W_j := clos_{L^2(\mathbb{R})} \langle \psi_{j,k} : k \in \mathbb{Z} \rangle$. Then, $L^2(\mathbb{R})$ can be decomposed as a direct sum of the spaces W_j,

$$L^2(\mathbb{R}) = \sum_{j\in\mathbb{Z}}^{\bullet} W_j := \cdots \dotplus W_{-1} \dotplus W_0 \dotplus W_1 \dotplus \cdots, \tag{7}$$

in the sense that every function $f \in L^2(\mathbb{R})$ has a unique decomposition,

$$f(x) = \cdots + g_{-1}(x) + g_0(x) + g_1(x) + \cdots,$$

where $g_j \in W_j,\ j \in \mathbb{Z}$.

If ψ is an orthogonal wavelet, then the subspaces W_j of $L^2(\mathbb{R})$ are mutually orthogonal and consequently, the direct sum in (7) becomes an orthogonal sum,

$$L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j := \cdots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \cdots. \tag{8}$$

Obviously, every SO wavelet generates an orthogonal decomposition (8) of $L^2(\mathbb{R})$, and every ON wavelet is also an SO wavelet.

3.2. Multiresolution analysis, scaling functions and B-splines

Any wavelet, semi-orthogonal or not, generates a direct sum decomposition (7) of $L^2(\mathbb{R})$. For each $j \in \mathbb{Z}$, let us consider the closed subspaces,

$$V_j = \cdots \dotplus W_{j-2} \dotplus W_{j-1}, \quad j \in \mathbb{Z},$$

of $L^2(\mathbb{R})$. These subspaces have the following properties,

(i) $\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots$,
(ii) $clos_{L^2}\left(\bigcup_{j\in\mathbb{Z}} V_j\right) = L^2(\mathbb{R})$,
(iii) $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$,
(iv) $V_{j+1} = V_j \dotplus W_j,\ j \in \mathbb{Z}$,
(v) $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1},\ j \in \mathbb{Z}$.

Observe that every function $f \in L^2(\mathbb{R})$ can be approximated as closely as desirable by its projections $P_j f$ in V_j as described by (ii).

If the reference subspace V_0, say, is generated by a single function $\phi \in L^2(\mathbb{R})$ in the sense that,

$$V_0 := clos_{L^2(\mathbb{R})} \langle \phi_{0,k} : k \in \mathbb{Z} \rangle,$$

where $\phi_{j,k}(x) := 2^{j/2}\phi(2^j x - k)$, then all subspaces V_j are also generated by the same ϕ, namely,

$$V_j := clos_{L^2(\mathbb{R})} \langle \phi_{j,k} : k \in \mathbb{Z} \rangle, \quad j \in \mathbb{Z}. \tag{9}$$

Definition 5. A function $\phi \in L^2(\mathbb{R})$ is said to generate a multiresolution analysis (MRA) if it generates a nested sequence of closed subspaces V_j that satisfy (i), (ii), (iii) and (v) in the sense of (9), such that $\{\phi_{0,k}\}$ forms a Riesz basis of V_0. If ϕ generates a MRA, then ϕ is called a scaling function.

Typical examples of scaling functions ϕ are the jth order cardinal B-splines, $N_j(x)$, defined recursively by a convolution,

$$N_j(x) = \int_{-\infty}^{\infty} N_{j-1}(x-t)N_0(t)dt = \int_0^1 N_{j-1}(x-t)dt, \quad j \geq 1,$$

where,

$$N_0(x) = \chi_{[0,1)}(x) = \begin{cases} 1, & \text{if } x \in [0,1), \\ 0, & \text{otherwise.} \end{cases}$$

We note that cardinal B-spline functions are compactly supported, since the support of the jth order B-spline function N_j is $[0, j+1]$, and they have as the Fourier transform,

$$\widehat{N_j}(w) = \left(\frac{1 - e^{-iw}}{iw}\right)^{j+1}.$$

To describe the space V_0 that N_j generates, we define π_n as the collection of all polynomials of degree at most n, and C^n denotes the collection of all functions f such that $f, f^{(1)}, \ldots, f^{(n)}$ are continuous everywhere.

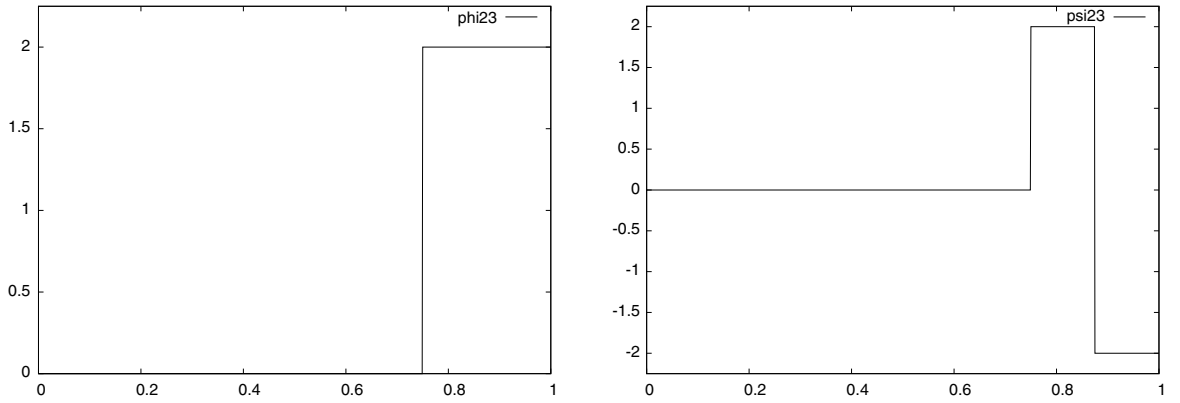

Fig. 1. Scaling ($\phi_{2,3}$) and wavelet ($\psi_{2,3}$) functions.

The subspace V_0 generated by N_j consists of all functions $f \in C^{j-1} \cap L^2(\mathbb{R})$ such that the restriction of each function f to any interval $[k, k+1)$, $k \in \mathbb{Z}$, is in π_j. From property (v) of an MRA, we can now identify all other subspaces V_j, namely,

$$V_j = \{f \in C^{s-1} \cap L^2(\mathbb{R}) : f|_{[\frac{k}{2^j}, \frac{k+1}{2^j})} \in \pi_s, k \in \mathbb{Z}\}.$$

Since splines are only piecewise polynomial functions, they are very easy to implement in a computer code.

From the nested sequence of splines subspaces V_j, we have the orthogonal complementary subspaces W_j, such that, $V_{j+1} = V_j \oplus W_j$, $j \in \mathbb{Z}$. Just as the B-spline of order s is the minimally supported generator of $\{V_j\}$ we can find the minimally supported $\psi^s \in W_0$ that generates the mutually orthogonal subspaces W_j. These compactly supported functions will be called B-wavelets of order s. In Chapter 6 in [19], explicit formulas for all ψ^s and their duals are derived.

In this paper we consider $\phi^j = N_j$ as the scaling function which generates a MRA and we restrict ourselves to the orders $j = 0, 1$. Clearly, for $j = 0$ we have the scaling function of the Haar wavelet system. In this case, the wavelet function is (see Fig. 1),

$$\psi^j(x) = \begin{cases} 1, & \text{if } 0 \le x < \dfrac{1}{2}, \\ -1, & \text{if } \dfrac{1}{2} \le x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

In this paper we use the $\mathrm{WA}^{[a,b]}$ method for approximating the CDF associated to the pool's losses, and therefore we carry out the approximation on a bounded interval $[a, b]$. For this reason, it is convenient to give a short review about B-splines on a bounded interval. We refer the reader to [21] for a detailed description of scaling functions on a bounded interval.

For sake of clarity, let us assume that $[0, n]$, $n \in \mathbb{N}$, is the working interval. We must distinguish between *interior* B-splines and *boundary* B-splines. We have,

$$N_j(x-k), \quad k = 0, \ldots, n-j-1.$$

These are the interior B-splines for the bounded interval $[0, n]$. The remaining B-splines,

$$\begin{cases} N_j(x-k), & k = -j, \ldots, -1, \text{ and} \\ N_j(x-k), & k = n-j, \ldots, n-1, \end{cases}$$

are the boundary B-splines for the interval $[0, n]$. Here, the first group is for the boundary $x = 0$, while the second group is for the boundary $x = n$.

3.3. The $\mathrm{WA}^{[a,b]}$ method

In this section, we explain briefly the $\mathrm{WA}^{[a,b]}$ method and refer the reader to [14] for further details.

Let us consider a probability density function $f \in L^2(\mathbb{R})$ associated to a certain continuous random variable X, and its Fourier transform,

$$\widehat{f}(w) = \int_{-\infty}^{+\infty} e^{-iwx} f(x) dx. \tag{10}$$

We can expect that the mass in the tails tends to zero at infinity, so it can be approximated in a finite interval $[a, b]$ by,

$$f^c(x) = \begin{cases} f(x), & \text{if } x \in [a, b], \\ 0, & \text{otherwise.} \end{cases}$$

Following the theory of MRA in a bounded interval, we can approximate $f^c(x) \simeq f^c_{m,j}(x)$ for all $x \in [a, b]$, where,

$$f^c_{m,j}(x) = \sum_{k=0}^{(j+1)\cdot(2^m-1)} c^j_{m,k}\phi^j_{m,k}\left((j+1)\cdot\frac{x-a}{b-a}\right), \quad j \geq 0,$$

with convergence in L^2-norm. Note that we are not considering the left and right boundary scaling functions. For sake of simplicity, we set the coefficients corresponding to the boundary basis functions equal to zero. Observe that in the case that $j = 0$ (Haar wavelets) there are no boundary functions.

The main idea behind the Wavelet Approximation method is to approximate $\widehat{f}$ by $\widehat{f^c_{m,j}}$ and then to compute the coefficients $c^j_{m,k}$ by inverting the Fourier Transform. After some algebraic manipulation and the application of Cauchy's integral formula, we end up with the expressions,

$$c^j_{m,0} \simeq \frac{1}{\pi}\int_0^\pi \Re(Q_{m,j}(re^{iu}))du, \tag{11}$$

and,

$$c^j_{m,k} \simeq \frac{2}{\pi r^k}\int_0^\pi \Re(Q_{m,j}(re^{iu}))\cos(ku)du, \quad k = 1, \ldots, (j+1)\cdot(2^m-1), \tag{12}$$

where $r \neq 1$ is a positive real number and,

$$Q_{m,j}(z) = \frac{2^{\frac{m}{2}}(j+1)z^{-\frac{2^m(j+1)a}{b-a}}\widehat{f}\left(\frac{2^m(j+1)}{b-a}i\cdot\log(z)\right)(\log(z))^{j+1}}{(b-a)(z-1)^{j+1}}.$$

In practice, both integrals in (11) and (12) can be easily computed by means of the Trapezoidal Rule.

3.4. Computation of $\mathbb{E}[\mathscr{L}_i|X = x]$

By definition,

$$\mathbb{E}[\mathscr{L}_i|X = x] = \int_{\mathbb{R}} \min\left(S, (y-l)^+\right) f_{\mathcal{L}_i}(y|x)dy,$$

where $f_{\mathcal{L}_i}(y|x)$ is the conditional probability density function of $\mathcal{L}_i$ given $X = x$. This function is a sum of Dirac delta functions and takes values in the interval $[0, \Sigma]$, where $\Sigma = \sum_{\kappa=1}^{\mathcal{K}} L_\kappa$. Thus,

$$\mathbb{E}[\mathscr{L}_i|X = x] = \int_0^\Sigma \min\left(S, (y-l)^+\right) f_{\mathcal{L}_i}(y|x)dy. \tag{13}$$

Taking into account that,

$$\min\left(S, (y-l)^+\right) = \begin{cases} 0, & \text{if } y < l, \\ y-l, & \text{if } l \leq y < S+l, \\ S, & \text{if } y \geq S+l, \end{cases}$$

and assuming that $0 \leq l < S + l \leq \Sigma$, we can split the integral in expression (13) into two parts,

$$\mathbb{E}[\mathscr{L}_i|X = x] = \int_l^{S+l} (y-l) f_{\mathcal{L}_i}(y|x)dy + \int_{S+l}^\Sigma S f_{\mathcal{L}_i}(y|x)dy. \tag{14}$$

As it is also well known in the context of generalized functions, the derivative of the Heaviside step function is a Dirac delta. In this context (and of course in the context of regular functions) we can integrate by parts the first integral in expression (14) and directly the second integral, and we get,

$$\mathbb{E}[\mathscr{L}_i|X = x] = S - \int_l^{S+l} F_{\mathcal{L}_i}(y|x)dy, \tag{15}$$

where $F_{\mathcal{L}_i}(y|x)$ is the conditional cumulative distribution function of $\mathcal{L}_i$ given $X = x$.

As shown in [13], the Haar wavelets system constitutes a suitable basis to approximate stepped functions like $F_{\mathcal{L}_i}$. Since $F_{\mathcal{L}_i}$ is supported on the interval $[0, \Sigma]$ and $F_{\mathcal{L}_i} \in L^2([0, \Sigma])$, we can apply the $\mathrm{WA}^{[a,b]}$ method with $[a, b] = [0, \Sigma]$ to approximate $F_{\mathcal{L}_i}$.

3.4.1. Approximation with the Haar system ($j = 0$)

We consider the approximation,

$$F_{\mathcal{L}_i}(y|x) \simeq F^j_{\mathcal{L}_i,m}(y|x) := \sum_{k=0}^{2^m-1} c^j_{m,k}(x)\phi^j_{m,k}\left(\frac{y}{\Sigma}\right), \quad y \in [0, \Sigma],$$

where we write $c^j_{m,k}(x)$ to underline the dependence of the coefficients on x. If we replace $F_{\mathcal{L}_i}$ by $F^j_{\mathcal{L}_i,m}$ in expression (15), gives us,

$$\begin{aligned}\mathbb{E}[\mathscr{L}_i|X = x] &\simeq \mathbb{E}^j_m[\mathscr{L}_i|X = x] := S - \sum_{k=0}^{2^m-1}\left(c^j_{m,k}(x)\int_l^{S+l}\phi^j_{m,k}\left(\frac{y}{\Sigma}\right)dy\right)\\ &= S - 2^{m/2}\left[\left(\frac{k_1+1}{2^m}\cdot\Sigma - l\right)c^j_{m,k_1}(x) + \frac{\Sigma}{2^m}\sum_{k=k_1+1}^{k_2-1}c^j_{m,k}(x) + \left(S + l - \frac{k_2}{2^m}\cdot\Sigma\right)c^j_{m,k_2}(x)\right], \qquad (16)\end{aligned}$$

where k_1 and k_2 are integers such that, $k_1 = \left\lfloor \frac{2^m}{\Sigma} l \right\rfloor$ and $k_2 = \left\lfloor \frac{2^m}{\Sigma}(S + l) \right\rfloor$, and $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x, and this completes the proof.

The remaining part of this section is devoted to the computation of the coefficients $c_{m,k}(x)$. Let $\widehat{f}_{\mathcal{L}_i}(\omega|x) := \mathbb{E}\left(e^{-i\omega\mathcal{L}_i}|X = x\right)$ be the *characteristic function* of the random variable $\mathcal{L}_i$ given $X = x$, i.e., the Fourier transform of the corresponding density $f_{\mathcal{L}_i}$. Under the conditional independence framework, the default indicators $\mathbf{1}_{\{Y_\kappa < H_\kappa(t_i)\}}$ are independent, and therefore,

$$\mathbb{E}\left(e^{-i\omega\mathcal{L}_i}|X = x\right) = \prod_{\kappa=1}^{\mathcal{K}}\mathbb{E}\left(e^{-i\omega L_\kappa \mathbf{1}_{\{Y_\kappa < H_\kappa(t_i)\}}}|X = x\right) = \prod_{\kappa=1}^{\mathcal{K}}\left(\pi_\kappa(t_i; x)e^{-i\omega L_\kappa} + 1 - \pi_\kappa(t_i; x)\right).$$

It is worth noting that if $\widehat{F}_{\mathcal{L}_i}$ represents the Fourier transform of $F_{\mathcal{L}_i}$, it is straightforward to see that,

$$\widehat{F}_{\mathcal{L}_i}(\omega|x) = \frac{\widehat{f}_{\mathcal{L}_i}(\omega|x) - e^{-i\Sigma\omega}}{i\omega}.$$

Finally, if we consider,

$$Q_m(z) = \frac{2^{\frac{m}{2}}\widehat{F}_{\mathcal{L}_i}\left(\frac{2^m}{\Sigma}i\cdot\log(z)|x\right)\log(z)}{\Sigma(z-1)},$$

then we can recover the coefficients $c_{m,k}(x)$ following (11) and (12).

3.4.2. Approximation with B-splines scaling functions ($j = 1$)

We consider the approximation,

$$F_{\mathcal{L}_i}(y|x) \simeq F^j_{\mathcal{L}_i,m}(y|x) := \sum_{k=0}^{2(2^m-1)} c^j_{m,k}(x)\phi^j_{m,k}\left(\frac{2y}{\Sigma}\right), \quad y \in [0, \Sigma],$$

where we write $c^j_{m,k}(x)$ to underline the dependence of the coefficients on x. If we replace $F_{\mathcal{L}_i}$ by $F^j_{\mathcal{L}_i,m}$ in expression (15), gives us,

$$\mathbb{E}[\mathscr{L}_i|X = x] \simeq \mathbb{E}^j_m[\mathscr{L}_i|X = x] := S - \sum_{k=0}^{2(2^m-1)}\left(c^j_{m,k}(x)\int_l^{S+l}\phi^j_{m,k}\left(\frac{2y}{\Sigma}\right)dy\right),$$

where,

$$\phi^j_{m,k}\left(\frac{2y}{\Sigma}\right) = \begin{cases} 2^{m/2}\left(1 - \left|2^{m+1}\cdot\frac{y}{\Sigma} - k - 1\right|\right), & \frac{k}{2^{m+1}}\cdot\Sigma \le y < \frac{k+2}{2^{m+1}}\cdot\Sigma,\\ 0, & \text{otherwise.}\end{cases}$$

If we define $\alpha_k = \frac{k}{2^{m+1}}\Sigma$, $\gamma_k = \frac{k+2}{2^{m+1}}\Sigma$ and $\beta_k = (\alpha_k + \beta_k)/2$, then,

$$\begin{aligned}\mathbb{E}_m^j[\mathscr{L}_i|X=x] = S - 2^{m/2}\Bigg[&\sum_{k\in\mathcal{I}_1} c_{m,k}^j(x)\Bigg(\mathbf{1}_{\{\beta_k\le l\}}\cdot\int_l^{\gamma_k}\left(-\frac{2^{m+1}}{\Sigma}y+k+2\right)dy\\ &+\mathbf{1}_{\{\beta_k>l\}}\cdot\left(\int_l^{\beta_k}\left(\frac{2^{m+1}}{\Sigma}y-k\right)dy+\int_{\beta_k}^{\gamma_k}\left(-\frac{2^{m+1}}{\Sigma}y+k+2\right)dy\right)\Bigg)\\ &+\sum_{k\in\mathcal{I}_2} c_{m,k}^j(x)\left(\int_{\alpha_k}^{\beta_k}\left(\frac{2^{m+1}}{\Sigma}y-k\right)dy+\int_{\beta_k}^{\gamma_k}\left(-\frac{2^{m+1}}{\Sigma}y+k+2\right)dy\right)\\ &+\sum_{k\in\mathcal{I}_3} c_{m,k}^j(x)\Bigg(\mathbf{1}_{\{\beta_k<S+l\}}\cdot\left(\int_{\alpha_k}^{\beta_k}\left(\frac{2^{m+1}}{\Sigma}y-k\right)dy+\int_{\beta_k}^{S+l}\left(-\frac{2^{m+1}}{\Sigma}y+k+2\right)dy\right)\\ &+\mathbf{1}_{\{\beta_k\ge S+l\}}\cdot\int_{\alpha_k}^{S+l}\left(\frac{2^{m+1}}{\Sigma}y-k\right)dy\Bigg)\Bigg],\end{aligned} \tag{17}$$

where $\mathcal{I}_1 = \{k : \alpha_k \le l < \gamma_k\}$, $\mathcal{I}_2 = \{k : \alpha_k > l, \gamma_k \le S + l\}$ and $\mathcal{I}_3 = \{k : \alpha_k \le S + l < \gamma_k\}$.

Result 1. *If we define* $I_1(a,b) := \int_a^b \left(\frac{2^{m+1}}{\Sigma}y - k\right)dy$ *and* $I_2(c,d) := \int_c^d \left(-\frac{2^{m+1}}{\Sigma}y + k + 2\right)dy$, *then,*

$$I_1(a,b) = (b-a)\left(\frac{2^m}{\Sigma}(b+a) - k\right),$$
$$I_2(c,d) = (d-c)\left(-\frac{2^m}{\Sigma}(c+d) + k + 2\right).$$

Finally, if we use Result 1, then the expression (17) can be written in compact form as,

$$\begin{aligned}\mathbb{E}_m^j[\mathscr{L}_i|X=x] = S - 2^{m/2}\Bigg[&\sum_{k\in\mathcal{I}_1} c_{m,k}^j(x)\Big(\mathbf{1}_{\{\beta_k\le l\}}\cdot I_2(l,\gamma_k) + \mathbf{1}_{\{\beta_k>l\}}\cdot\Big(I_1(l,\beta_k) + I_2(\beta_k,\gamma_k)\Big)\Big)\\ &+\sum_{k\in\mathcal{I}_2} c_{m,k}^j(x)\Big(I_1(\alpha_k,\beta_k) + I_2(\beta_k,\gamma_k)\Big) + \sum_{k\in\mathcal{I}_3} c_{m,k}^j(x)\Big(\mathbf{1}_{\{\beta_k<S+l\}}\cdot\Big(I_1(\alpha_k,\beta_k) + I_2(\beta_k,S+l)\Big)\\ &+\mathbf{1}_{\{\beta_k\ge S+l\}}\cdot I_1(\alpha_k,S+l)\Big)\Bigg].\end{aligned}$$

The coefficients $c_{m,k}^j(x)$ are to be computed following a similar procedure like in Section 3.4.1. In this case, we consider,

$$Q_m(z) = \frac{2^{\frac{m}{2}+1}\widehat{F}_{\mathcal{L}_i}\left(\frac{2^{m+1}}{\Sigma}i\cdot\log(z)|x\right)(\log(z))^2}{\Sigma(z-1)^2}.$$

3.5. Error analysis

We distinguish three sources of error in the computation of the conditional expected losses $\mathbb{E}[\mathscr{L}_i|X=x]$ in (5) by means of the $\mathrm{WA}^{[a,b]}$ method. These are:

(A) The approximation error at scale m,

$$\mathcal{E}(x) := \mathbb{E}[\mathscr{L}_i|X=x] - \mathbb{E}_m^j[\mathscr{L}_i|X=x] = \int_l^{S+l}\left[F_{\mathcal{L}_i}(y|x) - F_{\mathcal{L}_i,m}^j(y|x)\right]dy.$$

(B) The discretization error, which results when computing the integrals (11) and (12) by means of the Trapezoidal Rule.
(C) The roundoff error. This type of error arises after multiplying by the factor $\frac{1}{r^k}$ in (12).

Here, we focus on the study of the error of type (A), and we refer the reader to [14] for a detailed analysis of the error of types (B) and (C).

Looking at,

$$|\mathcal{E}(x)| = \left| \int_l^{S+l} \left[F_{\mathcal{L}_i}(y|x) - F_{\mathcal{L}_i,m}(y|x) \right] dy \right|,$$

then, by the Cauchy–Schwarz inequality, we have,

$$|\mathcal{E}(x)| \leq \left(\int_l^{S+l} 1\, dy \right)^{1/2} \cdot \|F_{\mathcal{L}_i}(\cdot|x) - F_{\mathcal{L}_i,m}(\cdot|x)\|_2 = \sqrt{S} \cdot \|F_{\mathcal{L}_i}(\cdot|x) - F_{\mathcal{L}_i,m}(\cdot|x)\|_2, \tag{18}$$

where $\|\cdot\|_2$ indicates the norm restricted to the space $L^2([l, S+l])$.

Following the theory of MRA in Section 3.2, we can write,

$$F_{\mathcal{L}_i}(y|x) = \sum_{k=0}^{(j+1)\cdot(2^m-1)} c^j_{m,k}(x)\phi^j_{m,k}\left((j+1)\cdot\frac{y}{\Sigma}\right) + \sum_{p\geq m}\sum_{k=0}^{(j+1)\cdot(2^p-1)} d^j_{p,k}(x)\psi^j_{p,k}\left((j+1)\cdot\frac{y}{\Sigma}\right), \tag{19}$$

where $d^j_{p,k}$ are the wavelet coefficients defined as,

$$d^j_{p,k} := \int_0^{\Sigma} F_{\mathcal{L}_i}(y|x)\cdot\tilde{\psi}^j_{p,k}\left((j+1)\cdot\frac{y}{\Sigma}\right) dy, \tag{20}$$

with $\tilde{\psi}^j$ the jth order dual wavelet (note that in the case of the Haar system, the jth order wavelet is the same as the jth order dual wavelet, that is, $\psi^j = \tilde{\psi}^j$).

If we consider expression (19) and taking into account that $\{\psi^j_{p,k}\}_{p,k\in\mathbb{Z}}$ is a Riesz basis, then,

$$\|F_{\mathcal{L}_i}(\cdot|x) - F^j_{\mathcal{L}_i,m}(\cdot|x)\|_2^2 = \left\| \sum_{p\geq m}\sum_{k=0}^{(j+1)\cdot(2^p-1)} d^j_{p,k}(x)\psi^j_{p,k}\left((j+1)\cdot\frac{y}{\Sigma}\right) \right\|_2^2 \leq \mathcal{C}\sum_{p\geq m}\sum_{k=0}^{(j+1)\cdot(2^p-1)} |d^j_{p,k}(x)|^2,$$

where $\mathcal{C}$ is a certain constant. It is worth remarking that part of the coefficients on the right hand side of the inequality are zero, since the norm has been taken over the interval $[l, S+l]$. As we can observe, the approximation error depends on the size of the coefficients $d^j_{p,k}$. We aim to demonstrate that for stepped functions like $F_{\mathcal{L}_i}$, the size of the coefficients $d^j_{p,k}$ is similar in the presence of a jump discontinuity, regardless of the order of the B-spline basis employed. To this end, we state the following proposition (for sake of simplicity we drop the explicit dependence of $d^j_{p,k}(x)$ on x and thus, we write $d^j_{p,k}$).

Proposition 1. *Let us assume that $F_{\mathcal{L}_i}$ has a unique jump discontinuity at $y_0 \in [0, \Sigma]$. Then,*

$$|d^j_{p,k}| = \begin{cases} \mathcal{O}\left(2^{-\frac{p}{2}}\right), & \text{if } y_0 \in [0,\Sigma] \cap \operatorname{supp}\tilde{\psi}^j_{p,k}, \\ 0, & \text{otherwise.} \end{cases}$$

Proof. If we assume first that $j = 0$ and we take into account the fact that $\operatorname{supp}\psi^j_{p,k}\left(\frac{y}{\Sigma}\right) = \left[\frac{k}{2^p}\Sigma, \frac{k+1}{2^p}\Sigma\right]$, then,

$$\begin{aligned} d^j_{p,k} &= \int_0^{\Sigma} F_{\mathcal{L}_i}(y|x)\cdot\tilde{\psi}^j_{p,k}\left(\frac{y}{\Sigma}\right) dy = \int_{\frac{k}{2^p}\Sigma}^{\frac{k+1}{2^p}\Sigma} F_{\mathcal{L}_i}(y|x)\cdot\psi^j_{p,k}\left(\frac{y}{\Sigma}\right) dy \\ &= 2^{\frac{p}{2}}\int_{\frac{k}{2^p}\Sigma}^{\frac{k+1}{2^p}\Sigma} F_{\mathcal{L}_i}(y|x)\cdot\psi^j\left(2^p\frac{y}{\Sigma} - k\right) dy. \end{aligned}$$

If we make the change of variables $u = 2^p\frac{y}{\Sigma} - k$, gives us,

$$d^j_{p,k} = 2^{-\frac{p}{2}}\Sigma\int_0^1 F_{\mathcal{L}_i}\left(\frac{u+k}{2^p}\Sigma|x\right)\cdot\psi^j(u)du. \tag{21}$$

Let us define $u_0 = 2^p\frac{y_0}{\Sigma} - k$. If $y_0 \notin \operatorname{supp}\psi^j_{p,k}$ then $u_0 \notin [0,1]$ and therefore $d^j_{p,k} = 0$, since $F_{\mathcal{L}_i}$ is constant and ψ^j has one vanishing moment. If $y_0 \in \operatorname{supp}\psi^j_{p,k}$ then $u_0 \in [0,1]$. Without loss of generality we can assume that $u_0 < \frac{1}{2}$ and we can split the integral in expression (21) into three parts,

$$d^j_{p,k} = 2^{-\frac{p}{2}}\Sigma\left(\int_0^{u_0} F^-_{\mathcal{L}_i}\cdot\psi^j(u)du + \int_{u_0}^{\frac{1}{2}} F^+_{\mathcal{L}_i}\cdot\psi^j(u)du + \int_{\frac{1}{2}}^1 F^+_{\mathcal{L}_i}\cdot\psi^j(u)du\right),$$

where $F^-_{\mathcal{L}_i}$ and $F^+_{\mathcal{L}_i}$ are the constant values of $F_{\mathcal{L}_i}$ evaluated on left and the right side of the jump discontinuity, respectively.

Finally, following the definition of ψ^j, we have,

$$|d^j_{p,k}| = 2^{-\frac{p}{2}}\Sigma|F^-_{\mathcal{L}_i} - F^+_{\mathcal{L}_i}|u_0.$$

In case that $j = 1$, we note that supp $\psi^j(t) = [0, 3]$ and $\tilde{\psi}^j_{p,k}(t) = \sum_{s=0}^{\bar{s}} h^k_{p,s}\psi^j_{p,s}(t)$, for certain coefficients $h^k_{p,s}$ and $\bar{s} > 0$, $\bar{s} \in \mathbb{N}$. Therefore, $\tilde{\psi}^j_{p,k}\left((j+1)\cdot\frac{y}{\Sigma}\right)$ is supported on the whole interval $[0, \Sigma]$ (see [19,23] for details). Thus,

$$\begin{aligned} d^j_{p,k} &= \int_0^{\Sigma} F_{\mathcal{L}_i}(y|x)\cdot\tilde{\psi}^j_{p,k}\left(\frac{2y}{\Sigma}\right)dy = \int_0^{\Sigma} F_{\mathcal{L}_i}(y|x)\cdot\sum_{s=0}^{\bar{s}} h^k_{p,s}\psi^j_{p,s}\left(\frac{2y}{\Sigma}\right)dy \\ &= 2^{\frac{p}{2}}\sum_{s=0}^{\bar{s}} h^k_{p,s}\int_{\frac{s}{2^p}\frac{\Sigma}{2}}^{\frac{s+3}{2^p}\frac{\Sigma}{2}} F_{\mathcal{L}_i}(y|x)\cdot\psi^j\left(2^p\frac{2y}{\Sigma} - s\right)dy. \end{aligned}$$

If we make the change of variables $u = 2^p\frac{2y}{\Sigma} - s$, gives us,

$$d^j_{p,k} = 2^{-\frac{p}{2}}\frac{\Sigma}{2}\sum_{s=0}^{\bar{s}} h^k_{p,s}\int_0^3 F_{\mathcal{L}_i}\left(\frac{u+s}{2^p}\cdot\frac{\Sigma}{2}|x\right)\cdot\psi^j(u)du. \tag{22}$$

Let us define $u_0 = 2^p\frac{2y_0}{\Sigma} - s$. If $y_0 \notin$ supp $\psi^j_{p,s}$ then $u_0 \notin [0, 3]$ and therefore the integral in (22) is zero, since $F_{\mathcal{L}_i}$ is constant and ψ^j has two vanishing moments. If $y_0 \in$ supp $\psi^j_{p,s}$ then $u_0 \in [0, 3]$ and then, only some of the terms on the right hand side of (22) remain. □

Remark 1. The result in Proposition 1 can be easily extended to the general case of having more than one jump discontinuity and therefore, the approximation error is similar regardless of the order of the B-spline basis employed when dealing with stepped shape functions. Further, we note that an analogous argument holds when the order of the B-spline basis is greater than one. We can conclude that enlarging the order of the wavelet basis does not improve the accuracy of the proposed method, while the difficulty of the implementation increases as well as the CPU time needed. We confirm the theoretical results along the numerical examples section.

4. Numerical experiments

In this section, we carry out several numerical experiments[1] to evaluate the efficiency and robustness of the WA$^{[a,b]}$ method by means of Haar and B-spline basis. We use a Monte Carlo method based on 100 samples with one million scenarios each as our benchmark. A point estimate for the price of the tranche is given by the mean and we provide a 95% confidence interval based on the calculation of the 2.5% and 97.5% quantiles.

To compute the coefficients for the WA$^{[a,b]}$ method at the scale of approximation m we consider $r = 0.9995, 2^m$ subintervals when applying the Trapezoidal Rule in the Haar case and 2^{m+1} in the B-spline case. These parameters are the optimal values to control the discretization and roundoff errors of the method, as shown in [14]. The Gauss–Hermite quadrature is carried out with $M = 20$ nodes. With this number of nodes we keep a good balance between the accuracy and the speed of the method, as demonstrated in [13]. Proceeding this way, the number of coefficients used for the Haar basis is 2^m and for the B-spline basis is $2^{m+1} - 1$.

The parameters corresponding to the numerical examples have been taken from the literature (mainly from [11]) and we list them below. The results presented are based on a sample of five pools of a variety of sizes and different types in terms of the notional values of the underlying reference entities. The notional values and sizes for each pool are summarized as follows,

Pool 1. $N_\kappa = 100,\ \kappa = 1, \ldots, 100.$

Pool 2. $N_\kappa = 50,\ \kappa = 1, \ldots, 50,\ N_\kappa = 100,\ \kappa = 51, \ldots, 100.$

Pool 3. $N_\kappa = 50,\ \kappa = 1, \ldots, 50, N_\kappa = 100,\ \kappa = 51, \ldots, 100, N_\kappa = 150, \kappa = 101, \ldots, 150,\ N_\kappa = 200,\ \kappa = 151, \ldots, 200.$

Pool 4. $N_\kappa = 20,\ \kappa = 1, \ldots, 80,\ N_\kappa = 50,\ \kappa = 81, \ldots, 160,\ N_\kappa = 100, \kappa = 161, \ldots, 240,\ N_\kappa = 150,\ \kappa = 241, \ldots, 320,\ N_\kappa = 200,\ \kappa = 321, \ldots, 400.$

Pool 5. $N_\kappa = \kappa,\ \kappa = 1, \ldots, 125.$

Thus, the higher the pool number, the more heterogeneous is the pool, with Pool 1 indicating a completely homogeneous pool in terms of notional values.

[1] The programs were coded in C language and run on a Dell Vostro 320 with Intel Core 2 Duo E7500 2.93 GHz processor and 4 GB RAM.

Table 1
Risk-neutral cumulative default probabilities.

Time (years)	Probability
1	0.72%
2	1.85%
3	3.28%
4	4.95%
5	6.80%

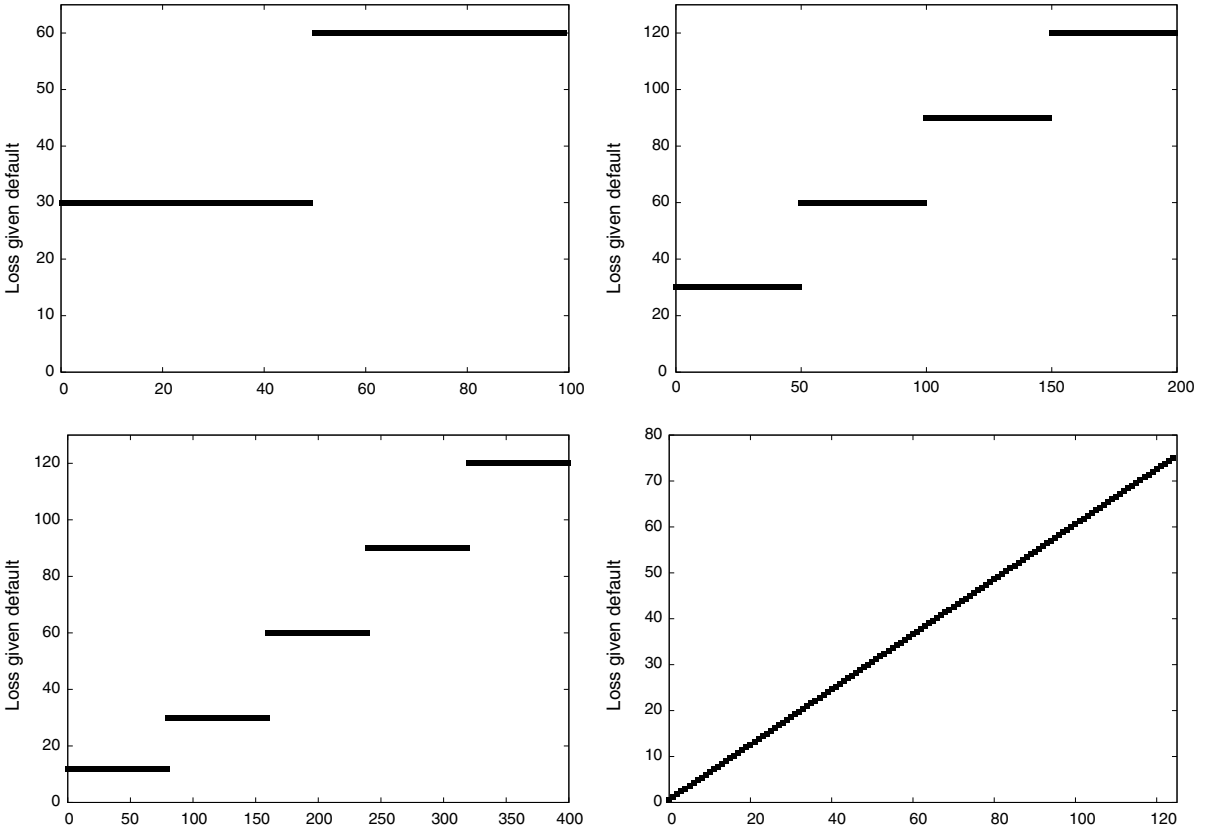

Fig. 2. Loss-given-default associated to Pool 2 (top left), Pool 3 (top right), Pool 4 (bottom left) and Pool 5 (bottom right).

For each name, the common risk-neutral cumulative default probabilities are given in Table 1. The recovery rate is assumed to be 40% for all names. Thus, the loss-given-default of name κ is $0.6N_\kappa$. Fig. 2 shows the loss-given-default corresponding to the pools ranging from 2 to 5. The maturity of a CDO deal is $T = 5$ years and the premium dates are $t_i = 1,\ i = 1, \ldots, 5$ years from today ($t_0 = 0$). The continuously compounded interest rates are $r_1 = 4.6\%$, $r_2 = 5\%$, $r_3 = 5.6\%$, $r_4 = 5.8\%$ and $r_5 = 6\%$. Thus, the corresponding discount factors are $d_i = e^{-r_i t_i}$, $i = 1, \ldots, 5$. All CDO pools have four tranches that are determined by the attachment points 0, 3%, 7% and 10%, with the last detachment point being 15%. The constants ρ_κ are 0.25, $\kappa = 1, \ldots, \mathcal{K}$.

We evaluate the accuracy and CPU time of the $\mathrm{WA}^{[a,b]}$ method at different scales of approximation, ranging from 6 to 9, and we present the results in Tables 2 and 3, where in the first case we deal with a completely homogeneous pool (Pool 1), while in the second case the pool is the most heterogeneous (Pool 5) in our set of examples. In both examples, we price the tranche with attachment point 3% and detachment point 7%. We observe that, in most of the scales, we reach the same accuracy with Haar and B-spline wavelets, as expected. However, the computational effort is greater with B-splines. It is worth remarking that Haar wavelets at scale six, are capable to compute the CDO spread with an error less than 1% in about one tenth of a second.

Thus, for the next examples, we consider the $\mathrm{WA}^{[a,b]}$ provided with a Haar basis. In Table 4 we present the minimum and maximum relative errors calculated with $M = 20$ and $M = 40$ nodes in the expression (6), arising in the computation of the first four tranches with reference pools Pool 2, Pool 3 and Pool 4 respectively. For $M = 20$, the minimum errors are reached in the first tranches, while the maximum errors correspond to the last tranches. The maximum relative error can

Table 2
Relative errors arising in the computation of the CDO spread for Pool 1 corresponding to the tranche with attachment point 3% and detachment point 7%. The reference value is 642.47 and the 95% confidence interval is [640.98, 644.18].

Scale m	Haar	Error	CPU time (s)	B-splines	Error	CPU time (s)
6	643.89	$2.22 \cdot 10^{-3}$	0.09	645.34	$4.47 \cdot 10^{-3}$	0.19
7	645.88	$5.31 \cdot 10^{-3}$	0.18	643.53	$1.66 \cdot 10^{-3}$	0.45
8	643.49	$1.59 \cdot 10^{-3}$	0.39	643.03	$8.78 \cdot 10^{-4}$	0.84
9	643.33	$1.35 \cdot 10^{-3}$	0.89	642.85	$5.98 \cdot 10^{-4}$	1.87

Table 3
Relative errors arising in the computation of the CDO spread for Pool 5 corresponding to the tranche with attachment point 3% and detachment point 7%. The reference value is 645.96 and the 95% confidence interval is [644.62, 647.51].

Scale m	Haar	Error	CPU time (s)	B-splines	Error	CPU time (s)
6	645.01	$1.47 \cdot 10^{-3}$	0.12	646.33	$5.74 \cdot 10^{-4}$	0.24
7	646.88	$1.43 \cdot 10^{-3}$	0.23	646.11	$2.34 \cdot 10^{-4}$	0.53
8	646.11	$2.34 \cdot 10^{-4}$	0.52	646.08	$1.87 \cdot 10^{-4}$	1.12
9	646.11	$2.34 \cdot 10^{-4}$	1.00	646.08	$1.87 \cdot 10^{-4}$	2.19

Table 4
Minimum (mE) and maximum (ME) relative errors when computing the first four tranches with Haar wavelets at scale of approximation 6.

Pool	MC	mE ($M = 20$)	ME ($M = 20$)	mE ($M = 40$)	ME ($M = 40$)
2	(2142.75, 646.65, 278.34, 124.24)	$7.51 \cdot 10^{-4}$	$1.29 \cdot 10^{-2}$	$6.21 \cdot 10^{-4}$	$5.15 \cdot 10^{-3}$
3	(2231.47, 638.95, 269.88, 119.34)	$7.30 \cdot 10^{-4}$	$3.03 \cdot 10^{-2}$	$8.74 \cdot 10^{-4}$	$3.44 \cdot 10^{-3}$
4	(2272.25, 631.43, 265.10, 116.81)	$1.72 \cdot 10^{-4}$	$5.55 \cdot 10^{-2}$	$1.58 \cdot 10^{-5}$	$1.30 \cdot 10^{-2}$

Table 5
Ratio of CPU times used by the method to compute simultaneously the first four tranches over those used to compute only the first tranche.

Pool	$m = 6$	$m = 7$	$m = 8$	$m = 9$
2	1.09	1.12	1.21	1.39
3	1.03	1.07	1.11	1.21
4	1.02	1.03	1.06	1.11

be reduced when considering $M = 40$ instead of $M = 20$ at the cost of doubling the CPU time. For sake of simplicity, we have not provided confidence intervals in this table.

Finally, we present in Table 5 the ratio of CPU times employed to compute simultaneously the spread corresponding to the first four tranches, over those used to compute only the first tranche. The ratio is computed at different scales of approximation and considering Pools 2–4 as the underlying pools. We can observe that at scale six, almost no extra time of computation is required, in contrast to EAP method which has a considerable higher ratio even when using only a few terms, as shown in [10]. Further, with the $\mathrm{WA}^{[a,b]}$ method, the ratio decreases when increasing the number of entities in the reference pool. However, for JKM method, this ratio tends to grow rapidly with the size of the pool (see [10] for details). The efficient behaviour of $\mathrm{WA}^{[a,b]}$ method is partly due to the compact support of the Haar basis, since each time that we compute the conditional expected losses in (16), only a few coefficients per tranche have to be considered.

5. Conclusions

In this work we have investigated and developed new formulae to value synthetic CDO tranches within a one-factor model framework, which leads to a stepped shape CDF. The methodology is based on the $\mathrm{WA}^{[a,b]}$ method and uses Haar and B-spline wavelets as the bases. Haar wavelets can be seen as order zero B-spline wavelets. We approximate the CDF associated to the pool's losses and we use the resulting expansion to price either one or multiple CDO tranches. From a theoretical point of view, we show that for staircase like functions the accuracy is the same regardless of the order of the B-spline basis employed. This result is confirmed along the numerical experiments, which are carried out with the Gaussian copula default model. Further, the method based on Haar wavelets is very easy to implement and the computational cost is lower when compared to the B-splines based method. Therefore, the Haar basis happens to be the most suitable one to perform the valuation, keeping a good balance between efficiency and easiness of implementation. We have shown that $\mathrm{WA}^{[a,b]}$ method works out well independently of the inhomogeneous features of the underlying pool and it is capable to price simultaneously several tranches of a CDO at the same computational effort than for only one tranche.

References

[1] L. Andersen, S. Basu, J. Sidenius, All your hedges in one basket, Risk (November) (2003) 67–72.
[2] J. Hull, A. White, Valuation of a CDO and an *n*th to default CDS without Monte Carlo simulation, J. Deriv. 12 (2) (2004) 8–23.
[3] J.-P. Laurent, J. Gregory, Basket default swaps, CDOs, and factor copulas, J. Risk 7 (4) (2005) 103–122.
[4] K. Jackson, A. Kreinin, X. Ma, Loss Distribution Evaluation for Synthetic CDOs, Working Paper, 2007.
[5] I. Iscoe, A. Kreinin, Valuation of synthetic CDOs, J. Banking Finance 31 (11) (2007) 3357–3376.
[6] B. De Prisco, I. Iscoe, A. Kreinin, Loss in translation, Risk (June) (2005) 77–82.
[7] C. Hipp, Improved approximations for the aggregate claims distribution in the individual model, ASTIN Bull. 16 (2) (1986) 89–100.
[8] J. Yang, T.R. Hurd, X. Zhang, Saddlepoint approximation method for pricing CDOs, J. Comput. Finance 10 (1) (2006) 1–20.
[9] L. Ortiz-Gracia, J.J. Masdemont, Credit risk contributions under the Vasicek one-factor model: a fast wavelet expansion approximation, J. Comput. Finance 17 (4) (2014) 59–97.
[10] I. Iscoe, K. Jackson, A. Kreinin, X. Ma, Pricing Correlation-dependent Derivatives Based on Exponential Approximations to the Hockey Stick Function, Working Paper, 2007.
[11] I. Iscoe, K. Jackson, A. Kreinin, X. Ma, Pricing synthetic collateralized debt obligations based on exponential approximations to the payoff function, J. Comput. Finance 16 (3) (2013) 127–150.
[12] P. Glasserman, S. Suchintabandid, Quadratic transform approximation for CDO pricing in multifactor models, SIAM J. Financ. Math. 3 (1) (2012) 137–162.
[13] J.J. Masdemont, L. Ortiz-Gracia, Haar wavelets-based approach for quantifying credit portfolio losses, Quant. Finance 14 (9) (2014) 1587–1595.
[14] L. Ortiz-Gracia, J.J. Masdemont, Peaks and jumps reconstruction with B-splines scaling functions, J. Comput. Appl. Math. 272 (2014) 258–272.
[15] L. Ortiz-Gracia, C.W. Oosterlee, Robust pricing of European options with wavelets and the characteristic function, SIAM J. Sci. Comput. 35 (5) (2013) B1055–B1084.
[16] H. Albrecher, S.A. Ladoucette, W. Schoutens, A generic one-factor Lévy model for pricing synthetic CDOs, in: R.J. Elliott, et al. (Eds.), Advances in Mathematical Finance, 2007, pp. 259–277.
[17] D. Guégan, J. Houdain, Collateralized debt obligations pricing and factor models: a new methodology using normal inverse Gaussian distributions. Research Report IDHE-MORA, n. 07–2005, ENS Cachan, 2005.
[18] O.E. Barndorff-Nielsen, Normal inverse Gaussian distributions and stochastic volatility modelling, Scand. J. Statist. 24 (1) (1997) 1–13.
[19] C.K. Chui, An Introduction to Wavelets, Academic Press, 1992.
[20] C.K. Chui, J.-Z. Wang, On compactly supported spline wavelets and a duality principle, Trans. Amer. Math. Soc. 330 (2) (1992) 903–915.
[21] C.K. Chui, Wavelets: A Mathematical Tool for Signal Analysis, in: SIAM Monographs on Mathematical Modeling and Computation, 1997.
[22] I. Daubechies, Ten Lectures on Wavelets, in: CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, 1992.
[23] E. Quak, N. Weyrich, Decomposition and reconstruction algorithms for spline wavelets on a bounded interval, Appl. Comput. Harmon. Anal. 1 (1994) 217–231.