

# A fast wavelet expansion technique for Vasicek multi-factor model of portfolio credit risk

Kensuke Ishitani<sup>1</sup>

<sup>1</sup> Mitsubishi UFJ Trust Investment Technology Institute Co., Ltd. (MTEC), 2-6, Akasaka 4-Chome, Minato-ku, Tokyo 107-0052 Japan

E-mail [ishitani@mtec-institute.co.jp](mailto:ishitani@mtec-institute.co.jp)

Received January 19, 2012, Accepted February 22, 2012

## Abstract

This paper presents a new methodology to compute VaR in the portfolio credit loss model. The Wavelet Approximation can be useful to compute non-smooth distributions, often arising in small or concentrated portfolios. We contribute to this technique by extending the Wavelet Approximation for Vasicek one-factor model to multi-factor model. Key features of our new algorithm are: (i) a finite series expansion of the wavelet scaling coefficients, (ii) Wynn's epsilon-algorithm to accelerate convergence of those series, and (iii) an efficient spline interpolation to calculate the Laplace transforms. We illustrate the effectiveness of our algorithm through numerical examples.

**Keywords** Vasicek multi-factor model, Haar wavelets, finite series expansion, Wynn's epsilon-algorithm, spline interpolation method

**Research Activity Group** Mathematical Finance

## 1. Introduction

Credit Risk Models are usually classified as structural or reduced-form models (see, for example, recent research [1]). In the present paper, we consider the structural model called Vasicek multi-factor model.

Let  $(\Omega, \mathcal{F}, P)$  be a probability space. Consider a credit portfolio consisting of  $N$  obligors. Define the exposure weight of obligor  $i$  by  $w_i = E_i / \sum_{j=1}^N E_j$ , where  $E_i$  is the exposure of obligor  $i$ , and the probability of default  $p_i$ . The Vasicek model assumes that the standardized asset log-return  $Z_i$  of obligor  $i$  is standard normally distributed and that obligor  $i$  defaults when  $Z_i$  is less than a pre-specified threshold  $\Phi^{-1}(p_i)$ , where  $\Phi(x)$  is the standard normal cumulative distribution function and  $\Phi^{-1}(p)$  is its inverse. Therefore the default can be modeled as a Bernoulli random variable  $D_i$  such that

$$D_i = \begin{cases} 1, & Z_i \leq \Phi^{-1}(p_i), \\ 0, & Z_i > \Phi^{-1}(p_i). \end{cases}$$

It follows that the portfolio loss is given by

$$L = \sum_{i=1}^N w_i D_i.$$

Let  $\text{VaR}_\alpha(L)$  be  $\alpha$ -quantile of the loss distribution  $L$  defined by

$$\text{VaR}_\alpha(L) = \inf \{x : P(L > x) \leq 1 - \alpha\}. \quad (1)$$

The modeling of the dependence structure among counterparties in the portfolio is simplified by the introduction of systematic risk factors  $Y = (Y_1, \dots, Y_M)$ . For each obligor  $i$ ,  $Z_i$  is represented by standard normally distributed systematic risk factor  $Y$  components

and an idiosyncratic noise component  $\epsilon_i$ :

$$Z_i = \alpha_i \cdot Y + \sqrt{1 - |\alpha_i|^2} \epsilon_i, \quad (2)$$

where  $Y$  and  $\epsilon_1, \dots, \epsilon_N$  are independent and normally distributed and the parameter  $\alpha_i = (\alpha_{i,s})_{s=1}^M \in [0, 1]^M$  is called the loading vector of obligor  $i$ .

Monte Carlo (MC) simulation is a standard method for measuring the risk of a credit portfolio. However this method is very time-consuming when the size of the portfolio increases. For this reason, analytical or fast numerical techniques have been developed during the last years. One of such approach is to give an analytical approximation of the Bromwich integral of the moment generating function [2, 3]. Another approach to numerically invert the Laplace transform is studied by [4, 5] via the wavelet approximation (WA) method [6]. Under the Vasicek one-factor model ( $M = 1$ ), [4, 5] shows accurate and fast results for a wide range of portfolios at very high loss levels.

In the present paper, we contribute to this techniques by extending the WA method for Vasicek multi-factor model ( $M \geq 2$ ).

## 2. Moment generating function (MGF) of the portfolio loss

Recall that in the Vasicek model framework, if the systematic risk factors  $Y$  are fixed, default occurs independently because the only remaining uncertainty is the idiosyncratic noise  $(\epsilon_i)_i$ . The MGF conditional on  $Y$  is thus given by the product of each obligor's MGF as

$$M_L(s; Y) \equiv E(e^{-sL} | Y) = \prod_{i=1}^N E(e^{-s w_i D_i} | Y)$$

$$= \prod_{i=1}^N (1 - p_i(Y) + p_i(Y)e^{-sw_i}),$$

where  $p_i(y)$  is the probability of obligor  $i$ 's default conditional on a realization  $Y = y$  given by

$$p_i(y) \equiv P(Z_i \leq T_i | Y = y) = \Phi \left( \frac{T_i - \alpha \cdot y}{\sqrt{1 - |\alpha_i|^2}} \right),$$

$$\text{for } y = (y_1, \dots, y_M), \quad \text{and } T_i \equiv \Phi^{-1}(p_i).$$

Taking the expectation value of this conditional MGF yields the unconditional MGF:

$$\begin{aligned} M_L(s) &\equiv E(e^{-sL}) = E(E(e^{-sL} | Y)) \\ &= E \left[ \prod_{i=1}^N (1 - p_i(Y) + p_i(Y)e^{-sw_i}) \right] \\ &= \int_{\mathbb{R}^M} \frac{1}{(2\pi)^{\frac{M}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left( -\frac{1}{2} y^t \Sigma^{-1} y \right) \\ &\quad \times \prod_{i=1}^N (1 - p_i(y) + p_i(y)e^{-sw_i}) dy, \end{aligned} \quad (3)$$

where  $\Sigma$  is the correlation matrix of a multivariate random vector  $Y = (Y_1, \dots, Y_M)$ .

### 3. Wavelet approximation of cumulative distribution functions (CDF)

Let  $K = L + \eta$  ( $\eta > 0$ ) be a random variable, then the CDF  $F_L(\cdot)$  is equal to  $F_K(\cdot + \eta)$  and  $F_K(x) = 0, x \in [0, \eta]$  holds. Let  $\phi$  be the scaling function of Haar Wavelet

$$\phi(x) = \begin{cases} 1, & \text{if } 0 \leq x < 1, \\ 0, & \text{otherwise,} \end{cases}$$

define  $\{\phi_{j,k}(x)\}_{k \in \mathbb{Z}}$  by  $\phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k)$ , and define a CDF  $F_K^m(x) = \sum_{k=0}^{\infty} c_{m,k} \phi_{m,k}(x)$ , where  $c_{m,k} = 2^{m/2} \int_{k/2^m}^{(k+1)/2^m} F_K(y) dy$ . Recall that  $F_K(\cdot)$  is a monotone nondecreasing function. We can show that for each  $x \in \mathbb{Q}_d \equiv \cup_{m \in \mathbb{N}} \{k/2^m; k \in \mathbb{Z}\}$  there exists some  $m_0 \in \mathbb{N}$  satisfying:

$$F_K^m(x) \geq F_K^{m+1}(x) \geq F_K(x), \quad (m \geq m_0),$$

$$F_K(x) = \lim_{m \rightarrow \infty} F_K^m(x).$$

The sequence  $(F_K^m)_m$  converges in distribution to  $F_K$ , since  $\mathbb{Q}_d$  is a dense subset of  $[0, \infty)$ .

The main idea of this section is to approximate the mgf  $M_K(\cdot)$  of  $F_K$  by the mgf  $M_K^m(\cdot)$  of  $F_K^m$  [4, 6]. Integration by parts on the integral in  $M_K^m$  gives

$$\begin{aligned} M_K(s) &\approx M_K^m(s) \equiv \int_0^\infty e^{-sx} dF_K^m(x) \\ &= s \int_0^\infty e^{-sx} F_K^m(x) dx \\ &= (1 - z) P_m \left( \exp \left( -\frac{s}{2^m} \right) \right) \end{aligned}$$

for  $s \in \mathcal{D} \equiv \{s \in \mathbb{C}; \operatorname{Re}(s) \geq 0\}$  and  $m \geq -\log_2(\eta)$ , where  $P_m(z) \equiv 2^{m/2} \sum_{k=0}^\infty c_{m,k} z^k$ . The residue theorem

then gives an approximation of  $F_K(x)$ :

$$\begin{aligned} F_K(x) &\approx F_K^m(x) = 2^{\frac{m}{2}} c_{m,k} = \frac{1}{2\pi i} \int_{\mathcal{C}_r} \frac{P_m(z)}{z^{k+1}} dz \\ &\approx \frac{1}{2\pi i} \int_{\mathcal{C}_r} \frac{\tilde{P}_m(z)}{z^{k+1}} dz, \quad \text{for } x \in \left[ \frac{k}{2^m}, \frac{k+1}{2^m} \right), \end{aligned} \quad (4)$$

$$\text{where } \tilde{P}_m(z) \equiv \frac{M_L(-2^m \log z) e^{2^m \eta \log z}}{1 - z}$$

$$\text{and } \mathcal{C}_r \equiv \{z \in \mathbb{C}; |z| = r\} \quad (0 < r < 1).$$

Therefore, by using the change of variables for the integral in (4), we have the following approximation

$$F_K(x) \approx \tilde{F}_K^m(x) \equiv \frac{1}{\pi r^k} \int_0^\pi \operatorname{Re}(\tilde{P}_m(re^{i\theta}) e^{-ik\theta}) d\theta \quad (5)$$

for  $k/2^m \leq x < (k+1)/2^m$ .

### 4. A fast algorithm for wavelet coefficient calculation

In the case of Vasicek one-factor model ( $M = 1$ ), [4, 5] approximate the integral of (5) using the ordinary trapezoidal rule and compute the MGF (3) fast and accurately by using a Gauss-Hermite quadrature formula.

In the case of Vasicek multi-factor model ( $M \geq 2$ ), on the other hand, Monte Carlo Integration is one of the most accurate method to compute the MGF. Thus, in this paper, we use the Monte Carlo Integration to compute the MGF (3). However, as is well known, Monte Carlo Integration is very time consuming. Therefore, we need more efficient method to approximate the integral of (5) than previous method [4, 5] using the ordinary trapezoidal rule.

We introduce fast and efficient methods for calculating the integral of (5) by using the convergence acceleration scheme (Wynn's epsilon algorithm) and a cubic spline interpolation.

### 5. Finite series expansion and Wynn's epsilon-algorithm

Let us consider a finite series expansion of (5) by changing the scale of the variable  $\hat{\theta} = k\theta$ :

$$\tilde{F}_K^m(x) = \sum_{j=0}^{2^m x_k^{(m)} - 1} a_j(x_k^{(m)}), \quad x \in [x_k^{(m)}, x_{k+1}^{(m)}], \quad (6)$$

$$\text{where } a_j(x) \equiv \frac{e^{\gamma x}}{\pi 2^m x} \int_{j\pi}^{(j+1)\pi} \operatorname{Re}(\hat{I}_x^{(m)}(\hat{\theta}) e^{-i\hat{\theta}}) d\hat{\theta},$$

$$\hat{I}_x^{(m)}(\hat{\theta}) \equiv \frac{M_L \left( \gamma - i \frac{\hat{\theta}}{x} \right) \exp \left( -\eta \left( \gamma - i \frac{\hat{\theta}}{x} \right) \right)}{1 - \exp \left( -\frac{\gamma}{2^m} \right) \exp \left( i \frac{\hat{\theta}}{2^m x} \right)},$$

$$x_k^{(m)} \equiv \frac{k}{2^m} \quad \text{and } \gamma \equiv -2^m \log r.$$

The purpose of this finite series expansion is to apply the Wynn's epsilon-algorithm as a method for acceleration of convergence of a complex valued series.

Wynn's epsilon-algorithm is the following nonlinear recursive scheme:

$$\epsilon_{k+1}^n = \epsilon_{k-1}^{n+1} + \frac{1}{\epsilon_k^{n+1} - \epsilon_k^n}, \quad (7)$$

$$\epsilon_{-1}^n = 0, \quad \epsilon_0^n = S_n,$$

where  $S_n$  is  $(n+1)$ th partial sum of the series (6). For  $q \in \mathbb{N}$ , it is known that the converted sequence  $(\epsilon_{2q}^n)_n$  converges drastically faster than the original series. Assume that  $S_0, S_1, \dots, S_{n(\epsilon)-1}$  ( $n(\epsilon) < k$ ) is given, then we can calculate

$$\epsilon_{2q}^0, \epsilon_{2q}^1, \dots, \epsilon_{2q}^{n(\epsilon)-2q-1}$$

by using the recursive equation (7) and approximate  $\tilde{F}_K^m(x)$  by  $\tilde{F}_K^{m,(\epsilon)}(x) \equiv \epsilon_{2q}^{n(\epsilon)-2q-1}$  for  $k/2^m \leq x < (k+1)/2^m$ . Therefore, the original CDF  $F_K(x)$  can be approximated by  $\tilde{F}_K^{m,(\epsilon)}(x)$ :

$$F_K(x) \approx \tilde{F}_K^{m,(\epsilon)}(x), \quad x \in \left[ \frac{k}{2^m}, \frac{k+1}{2^m} \right).$$

In order to calculate  $\tilde{F}_K^m(x)$ , we need to calculate  $M_L(\gamma - i\theta)$  for all  $\theta \in [0, 2^m\pi)$ . But on the other hand, in order to calculate  $\tilde{F}_K^{m,(\epsilon)}(x)$ , we only need to calculate  $M_L(\gamma - i\theta)$  for  $\theta \in [0, 2^m\pi n(\epsilon)/k)$ . We can thus expect that Wynn's epsilon-algorithm is quite efficient in calculating the integral of (5).

## 6. Cubic spline interpolation

It is required to calculate the moment generating function  $\{M_L(\gamma - i\theta); \theta \in [0, 2^m\pi n(\epsilon)/k_0)\}$  in order to obtain the CDF function  $\tilde{F}_K^{m,(\epsilon)}(x)$  for  $x \in [k_0/2^m, 1)$ . If a cubic spline interpolation method introduced below is used, the computation time of the moment generating function can be shortened significantly.

First, we calculate the MGF at grid points  $\Delta_\xi$

$$\Delta_\xi \equiv \{0 = \xi_0 < \xi_1 < \dots < \xi_{N_X}\}, \quad (8)$$

where  $\frac{2^m\pi n(\epsilon)}{k_0} < \xi_{N_X-1}$ ,

using the Monte Carlo methods:

$$M_L(\gamma - i\xi_i) \approx \frac{1}{N_I} \sum_{k=1}^{N_I} M_L(\gamma - i\xi_i; Y^{(k)}), \quad i \leq N_X, \quad (9)$$

where  $N_I$  is the sample size of monte carlo integration and  $Y^{(1)}, \dots, Y^{(N_I)}$  are samples from the probability distribution of random variable  $Y = (Y_1, \dots, Y_M)$ .

For  $i = 1, \dots, N_X - 1$ , we then find a quadratic function  $f_i(\theta)$  whose graph  $\{(\theta, f_i(\theta)); \theta \geq 0\}$  contains all three data points  $\{(\xi_j, M_L(\gamma - i\xi_j))\}_{j=i-1}^{i+1}$ .

Thus, we obtain the following cubic spline interpolation

$$M_L(\gamma - i\theta) \approx \frac{\xi_{i+1} - \theta}{\xi_{i+1} - \xi_i} f_i(\theta) + \frac{\theta - \xi_i}{\xi_{i+1} - \xi_i} f_{i+1}(\theta)$$

for  $\theta \in [\xi_i, \xi_{i+1}]$  ( $i \geq 1$ ) and quadratic spline interpolation  $M_L(\gamma - i\theta) \approx f_1(\theta)$  for  $\theta \in [\xi_0, \xi_1]$ .

Therefore, we can calculate the approximate MGF  $\{M_L(\gamma - i\theta); \theta \in [0, 2^m\pi n(\epsilon)/k_0)\}$  and compute the

integral of  $\int_{j\pi}^{(j+1)\pi} \text{Re}(\tilde{I}_x^{(m)}(\hat{\theta})e^{-i\hat{\theta}})d\hat{\theta}$ ,  $x \in [k_0/2^m, 1)$ ,  $j < n(\epsilon)$ , via the trapezoidal rule with a partition of the interval  $[j\pi, (j+1)\pi]$  into  $N_T^{(j)}$  equal parts.

## 7. Computation of VaR

We use a numerical root-finding algorithm of bisection method to invert the CDF  $\tilde{F}_K^{m,(\epsilon)}(\cdot)$ . Suppose  $\tilde{F}_K^{m,(\epsilon)}(k_0/2^m) < \alpha$ , then there exists  $j^*$  such that

$$\tilde{F}_K^{m,(\epsilon)}\left(\frac{j^*}{2^m}\right) \leq \alpha < \tilde{F}_K^{m,(\epsilon)}\left(\frac{j^*+1}{2^m}\right)$$

and approximate the Value-at-Risk

$$\text{VaR}_\alpha(L) = \text{VaR}_\alpha(K) - \eta \approx \frac{2j^*+1}{2^{m+1}} - \eta.$$

## 8. Numerical examples

In this section we illustrate the performance of our method through examples. All experiments are performed on a personal computer, Intel(R) Core(TM) 2 Duo CPU 3.00GHz, 2.00GB RAM.

We assume that the number of systematic risk factors is set as  $M = 33$  which is the number of TOPIX sector indices, each obligor  $i$  belong to an industrial sector  $sec(i) \in \{1, \dots, 33\}$  and  $Z_i$  is represented by a standard normally distributed sectoral factor component  $Y_{sec(i)}$  and an idiosyncratic noise component  $\epsilon_i$ :

$$Z_i = \sqrt{\rho_i} Y_{sec(i)} + \sqrt{1 - \rho_i} \epsilon_i. \quad (10)$$

In this multi-sector model (10), we can use an efficient monte carlo sampling algorithm [3] for the calculation of (9), whose computational time moderately depends on portfolio size. Note that efficient monte carlo sampling algorithms for the calculation of general multi-factor model (2) are not established. The reason is that discretization method [3] for the fast MGF computation does not work effectively as for the multi-sector model.

The factor correlations  $\Sigma = (\Sigma_{s,t})_{s,t=1}^{33}$  are estimated by the historical correlation of TOPIX sector index returns. We use monthly data of TOPIX sector indices from January 2007 to December 2011.

We consider eight sample portfolios, as shown in the following table,

Portfolio	$N$	$w_i$	$\rho_i$
P1	1,000	$C/i$	0.2
P2	1,000	$C/i$	0.25
P3	1,000	$C/i$	0.3
P4	1,000	$C/i$	0.35
P5	1,000	$C/i$	0.4
P6	3,000	$1/N$	0.3
P7	10,000	$C/i$	0.3
P8	30,000	$C/i$	0.3

where  $C$  is a positive constant such that  $\sum_{i=1}^N w_i = 1$  holds. The exposure distribution indicates that portfolios P1, P2, P3, P4, P5 and P7 are concentrated according to a power law distribution, whereas portfolio P6 is completely diversified.

We consider a rating system with 20 ratings, and assume their PD to be given by the following table.

$r_i$	1	2	3	4	5
$PD_{r_i}$	0.19%	0.21%	0.23%	0.25%	0.27%
$r_i$	6	7	8	9	10
$PD_{r_i}$	0.29%	0.30%	0.33%	0.35%	0.39%
$r_i$	11	12	13	14	15
$PD_{r_i}$	0.42%	0.5%	0.6%	0.9%	1.1%
$r_i$	16	17	18	19	20
$PD_{r_i}$	1.3%	1.73%	1.9%	3.0%	10.0%

PD for 20 ratings

In all portfolios, the rating  $r(i) \in \{1, \dots, 20\}$  of obligor  $i$  is chosen with

$$r(i) \equiv i \pmod{20}$$

holds for  $i = 1, \dots, N$ . Similarly to the rating allocation explained above, the sector  $sec(i) \in \{1, \dots, 33\}$  and the correlation parameter  $\rho_i$  of obligor  $i$  is chosen such that

$$sec(i) \equiv i \pmod{33}$$

holds for  $i = 1, \dots, N$ .

We refer to all other parameters as algorithm parameters, which determine the performance of the numerical approximation but do not affect the risk profile of portfolios. These algorithm parameters are listed in the following table.

Image resolution parameter	$m = 13$
Parallel shift parameter	$\eta = 0.0003$
Real part of integration path	$\gamma = -0.25 \times \log 10^{-14}$
Acceleration index	$q = 4$
Truncation parameter	$n(\epsilon) = 14$
Number of integration points	$N_I = 1,000,000$
Partition size of $[j\pi, (j+1)\pi]$	$N_T^{(j)} = 190,000$

The grid points (8) are set as

$$\Delta_\xi = \{i_1\}_{i_1=0}^{99} \cup \{100 + 4 \times i_2\}_{i_2=0}^{99}.$$

We examine the performance of our method (WA) by computing VaR for sample portfolios. Monte Carlo Simulation used as benchmarks are performed with 1 million loss scenarios (MC).

We present  $VaR_{0.999}(L)$  measurement results obtained by our method and Monte Carlo Simulation in the following table,

Portfolio	VaR(WA)	VaR(MC)	RE
P1	0.1454	0.1453	0.09%
P2	0.1471	0.1473	0.16%
P3	0.1495	0.1499	0.27%
P4	0.1531	0.1541	0.68%
P5	0.1586	0.1594	0.53%
P6	0.1001	0.0997	0.44%
P7	0.1228	0.1227	0.05%
P8	0.1148	0.1154	0.55%

where the relative error (RE) is defined by

$$RE = \frac{|VaR(WA) - VaR(MC)|}{VaR(MC)}.$$

All relative errors are smaller than 0.68%.

We provide as well the computational time in minutes in the following table both for WA method and MC method.

$N$	Portfolio	Time(WA)	Time(MC)
1,000	P1~P5	3.2 min	5.3 min
3,000	P6	3.3 min	17.9 min
10,000	P7	3.5 min	51.7 min
30,000	P8	3.9 min	158.3 min

WA method takes a few minutes even for a portfolio with 30,000 obligors. On the other hand, MC method requires computation time roughly proportional to portfolio size.

We have shown the suitability of the fast wavelet expansion method using epsilon-algorithm and cubic spline interpolation to measure the credit portfolio risk, based on the multi-factor Vasicek model ( $M \geq 2$ ).

## Acknowledgments

We thank Masaaki Otaka, Tetsu Kukita, Hideyuki Tanaka, Hiroshi Minaguchi, Takeshi Hirose, Takayoshi Yamamoto, Yuki Tai and Mari Ozeki for helpful discussions and comments.

## References

- [1] S. Yamanaka, M. Sugihara and H. Nakagawa, Analysis of credit event impact with self-exciting intensity model, JSIAM Letters, **3** (2011), 49–52.
- [2] P. Glasserman and J. Ruiz-Mata, Computing the credit loss distribution in the Gaussian copula model: a comparison of methods, J. Credit Risk, **2** (2007), 33–66.
- [3] Y. Takano and J. Hashiba, A Novel Methodology for Credit Portfolio Analysis: Numerical approximation approach, Available at [www.defaultrisk.com](http://www.defaultrisk.com).
- [4] J. J. Masdemont and L. O. Gracia, Haar wavelets-based approach for quantifying credit portfolio losses, Quantitative Finance, DOI:10.1080/14697688.2011.595731.
- [5] L. O. Gracia and J. J. Masdemont, Credit Risk Contributions under the Vasicek One-Factor Model : A fast wavelet expansion approximation, Available at [www.defaultrisk.com](http://www.defaultrisk.com).
- [6] G. G. Walter, Wavelets and Other Orthogonal Systems With Applications, CRC Press, Inc., Boca Raton, Florida, USA, 1994.