UNIVERSITAT POLITÈCNICA
DE CATALUNYA

Haar Wavelets-Based Methods for Credit Risk Portfolio Modeling

Departament de Matemàtica Aplicada I (UPC)

Luis Ortiz Gracia

Centre de Recerca Matemàtica

Programa de Doctorat en Matemàtica Aplicada

Memòria per aspirar al grau de Doctor
per la Universitat Politècnica de Catalunya

Certifico que aquesta tesi ha estat realitzada per Luis Ortiz Gracia
i dirigida per mi.

Barcelona, 28 de setembre de 2011.

Josep Joaquim Masdemont Soler

A Eva,
por todo su amor
y comprensión.

Agradecimientos

Quiero dedicar unas palabras de agradecimiento a las personas que directa o indirectamente me han ayudado a llegar hasta el final de mi tesis.

En primer lugar mis agradecimientos van dirigidos al Dr. Josep J. Masdemont por las innumerables horas que ha dedicado a la dirección de este trabajo. Siempre encontró el momento para atender todas las dudas que me surgían y guiarme por el sendero que él consideraba más apropiado, que resultaba ser siempre el mejor de todos los posibles. Entre otras virtudes, hay una que me gustaría destacar de Josep, la facilidad que tiene para hacer que las cosas complejas terminen siendo sencillas. Gracias Josep por el trato tan agradable que me has dispensado siempre y por todo tu apoyo tanto en el terreno científico como en el personal. Gracias también por haber confiado en mi trabajo desde el primer momento.

Un especial agradecimiento al Centre de Recerca Matemàtica por darme la oportunidad de llevar a cabo la parte final de mi tesis. Gracias a mis compañeros por el interés que han mostrado en mi trabajo, Álvaro, Tomás, Graham, Maite, Blanca, Dani, Esther, Francesc, Anna y especialmente a Pilar, de quién he recibido muchos consejos prácticos y con quién más tiempo he pasado hablando de la tesis.

En el terreno personal, quiero tener también palabras de agradecimiento para mis padres, que me han enseñado desde pequeño el valor del sacrificio y de la constancia que tanto he necesitado para terminar esta tesis. Admiro no os imagináis de qué forma, lo que habéis trabajado para que yo tuviese el privilegio de estudiar lo que me gustaba. Gracias a mis hermanas Raquel y Marta por haber admirado mi trabajo desde el principio. Eso aunque quizás vosotras no lo sepáis, me ha dado también alas para llegar hasta aquí.

No menos importante es todo el apoyo que he recibido por parte de mi mujer Eva. Gracias por haber permitido que yo dedicase un número incontable de horas a este trabajo mientras tú te ocupabas de atenderlo todo. Gracias por haberme dado la tranquilidad que necesitaba y sobre todo muchas gracias por el apoyo incondicional que me has dado desde el principio, a pesar de que estabas renunciando al poco tiempo libre en común que nos dejaban nuestros trabajos. Gracias a mi pequeña princesa Lucía por toda la ternura que me inspiras y perdona por no haberte podido dedicar todas las horas que te merecías. Ahora ya no me tendrás que venir a buscar al 'etudio' para ir juntos a jugar al 'paque'. Quiero dedicarte con especial cariño esta tesis a ti Rocío, que aunque todavía no te conozco ya estoy empezando a quererte mucho.

Contents

Abbreviations

ASRF Asymptotic Single Risk Factor

BCBS Basel Committee on Banking Supervision

CDF Cumulative Distribution Function

CDO Collateralized Debt Obligation

CGF Cumulant Generating Function

CLT Central Limit Theorem

EAD Exposure at Default

EC Economic Capital

EL Expected Losses

ES Expected Shortfall

ESC Risk Contribution to ES

GA Granularity Adjustment

HHI Herfindahl-Hirschman Index

IRB Internal Rating Based

LGD Loss Given Default

MC Monte Carlo

MGF Moment Generating Function

MRA Multi-Resolution Analysis

NA Normal Approximation

PD Probability of Default

PDF Probability Density Function

QD Quotient-Difference

QMC Quasi-Monte Carlo

RA Recursive Approximation

SP Saddle Point

UL Unexpected Losses

VaR Value at Risk

VaRC Risk Contribution to VaR

WA Wavelet Approximation

Notations

D_n Default indicator of obligor n

$\mathbb{E}$ Expectation with respect to the probability measure $\mathbb{P}$

E_n Exposure at default of obligor n

$\mathbf{EC}_\alpha$ Economic capital at confidence level α

EL_n Expectation value of L_n

ϵ_n Idiosyncratic shock of obligor n

$\mathbf{ES}_\alpha$ Expected shortfall at confidence level α

$\mathbf{ES}_\alpha^M$ Expected shortfall at confidence level α computed by means of a Monte Carlo method

$\mathbf{ES}_\alpha^{W(m)}$ Expected shortfall at confidence level α computed by means of the WA approximation at scale m

$\mathbf{ESC}_{\alpha,n}$ Risk contribution to ES_α of obligor n at confidence level α

$\mathbf{ESC}_{\alpha,n}^M$ Risk contribution to ES_α of obligor n at confidence level α computed by means of a Monte Carlo method

$\mathbf{ESC}_{\alpha,n}^{W(m)}$ Risk contribution to ES_α of obligor n at confidence level α computed by means of the WA approximation at scale m

$f_{\mathcal{L}}$ Probability density function of the loss variable $\mathcal{L}$

$f_{\mathcal{L}}^A$ Probability density function of $\mathcal{L}$ computed by means of the ASRF method

$F_{\mathcal{L}}$ Cumulative distribution function of the loss variable $\mathcal{L}$

$F_{\mathcal{L}}^A$ Cumulative distribution function of $\mathcal{L}$ computed by means of the ASRF method

$\mathcal{L}$ Portfolio loss random variable

L_n Loss given default of obligor n

m Scale of approximation for the WA approximation

$M_{\mathcal{L}}(s)$ Moment generating function of the loss variable $\mathcal{L}$

N Number of obligors in the portfolio

$\mathbb{P}$ Probability measure

Φ Cumulative standard normal distribution function

P_n Probability of default of obligor n

$P_n(y)$ Probability of default of obligor n conditional to a specification y of the business cycle

ρ_n Default correlation of obligor n

$\mathbb{V}$ Variance with respect to the probability measure $\mathbb{P}$

$\mathbf{VaR}_\alpha$ Value at risk at confidence level α

$\mathbf{VaR}_\alpha^A$ Value at risk at confidence level α computed by means of the ASRF method

$\mathbf{VaR}_\alpha^M$ Value at risk at confidence level α computed by means of a Monte Carlo method

$\mathbf{VaR}_\alpha^{W(m)}$ Value at risk at confidence level α computed by means of the WA approximation at scale m

$\mathbf{VaRC}_{\alpha,n}$ Risk contribution to VaR_α of obligor n at confidence level α

$\mathbf{VaRC}_{\alpha,n}^A$ Risk contribution to VaR_α of obligor n at confidence level α computed by means of the ASRF method

$\mathbf{VaRC}_{\alpha,n}^M$ Risk contribution to VaR_α of obligor n at confidence level α computed by means of a Monte Carlo method

$\mathbf{VaRC}_{\alpha,n}^{W(m)}$ Risk contribution to VaR_α of obligor n at confidence level α computed by means of the WA approximation at scale m

Y Single risk factor (state of the economy or business cycle)

Y_n Firm's composite factor

Chapter 1

Portfolio Credit Risk Modeling

1.1 Introduction

It is important for a bank to manage the risks originated from its business activities. In particular, the credit risk underlying the credit portfolio is often the largest risk in a bank. The measured credit risk is then used to assign risk capital to absorb potential losses arising from its credit portfolio.

The first Basel Accord of 1988, also known as Basel I, laid the basis for international minimum capital standard and banks became subject to regulatory capital requirements, coordinated by the Basel Committee on Banking Supervision (BCBS). This committee has been founded by the Central Bank Governors of the Group of Ten at the end of 1974.

The rules of the Basel Committee do not have any legal force. The supervisory rules are rather intended to provide guidelines for the supervisory authorities of the individual nations such that they can implement them in a suitable way for their banking system. The main focus of the first Basel Accord was on credit risk as the most important risk in the banking industry. Within Basel I banks are supposed to keep at least 8% equity in relation to their assets. The assets are weighted according to their degree of riskiness where the risk weights are determined in four different borrower categories (state, bank, mortgages, companies and retail customers). Hence the portfolio credit risk is measured as a sum of *risk weighted assets*, that is the sum of notional exposures weighted by a coefficient reflecting the credit-worthiness of the counterparty.

Basel I, however, does not account for methods to decrease risk as, for example, by means of portfolio diversification. Moreover, the approach measures risk in an insufficiently differentiated way since minimal capital requirements are computed independent of the borrower's credit-worthiness. These drawbacks lead to the development of the Second Basel Accord from 2001 onwards. In June 2004 the BCBS released a *Revised Framework on International Convergence of Capital Measurement and Capital Standards* (Basel II,[BCBS04]). The rules officially came into force on January 1st, 2008, in the European Union. However, in practice they had been applied already before that date.

Basel II is structured in a *three pillar framework*. Pillar one sets out details for adopting more risk sensitive minimal capital requirements, so called *regulatory capital*, for banking organizations. Pillar two lays out principles for the *supervisory review process* of capital

adequacy and Pillar three seeks to establish *market discipline* by enhancing transparency in bank's financial reporting.

The main goal of Pillar one is to take care of the specific risk of a bank when measuring minimal capital requirements. Within Basel II banks may opt for the *standard approach* which is quite conservative with respect to capital charge and the more advanced, so called *internal rating based* (IRB) approach when calculating regulatory capital for credit risk. In the standard approach, credit risk is measured by means of external ratings provided by rating agencies such as Standard&Poor's, Moody's or Fitch Ratings. In the IRB approach, the bank evaluates the risk itself. This approach can only been applied when the supervisory authorities accept it. The bank has to prove that certain conditions concerning the method and transparency are fulfilled. Basel II distinguishes between expected loss and unexpected loss. The former directly charges equity whereas for the later banks have to keep the appropriate capital requirements.

The supervisory review process of Pillar two is achieved by the supervisory authorities which evaluate and audit the compliance of regulations with respect to methods and transparency which are necessary for a bank to be allowed to use internal ratings.

The main target of Pillar three is to improve market discipline by means of transparency of information concerning a bank's external accounting.

The Merton model is the basis of the Basel II IRB approach. It is a Gaussian one factor model such that default events are driven by a latent common factor that is assumed to follow the Gaussian distribution. Under this model, loss only occurs when an obligor defaults in a fixed time horizon. If we assume certain homogeneity conditions, this one factor model leads to a simple analytic asymptotic approximation of the loss distribution and *value at risk* (VaR). This approximation, usually called Asymptotic Single Risk Factor (ASRF) model, works well for a large number of small exposures but can underestimate risks in the presence of exposure concentrations (see [Gie06]).

Concentration risks in credit portfolios arise from an unequal distribution of loans to single borrowers (*name concentration*) or different industry or regional sectors (*sector* or *country concentration*). Moreover, certain dependencies as, for example, direct business links between different borrowers, can increase the credit risk in a portfolio since the default of one borrower can cause the default of a dependent second borrower. This effect is called *default contagion* and is linked to both name and sector concentration.

In credit risk management one is particularly interested in the portfolio loss distribution. Since the portfolio loss is usually modeled as a sum of random variables, the main task is to evaluate the probability density function (PDF) of such a sum. The PDF of a sum of random variables is equal to the convolution of the respective PDFs of the individual asset loss distributions. The analytical evaluation of this convolution is a difficult problem, computationally is very intensive and, in full generality, is impractical for any realistically sized portfolio.

Monte Carlo (MC) simulation is a standard method for measuring credit portfolio risk in order to deal with exposure concentration. However this method is very time-consuming when the size of the portfolio increases, making the computation unworkable in many situations, due to the fact that a financial company has to re-balance their credit portfolios frequently.

In summary, credit risk managers in a bank are interested overall in the following issues:

- How can concentration risk be quantified?
- How can risk measures be computed in short times?
- How can the contribution of individual transactions to the total risk be computed?

To answer these questions, several methods have been developed in the last years. The Saddle Point (SP) approximation due to [Mar01a] gives an analytical approximation of the Laplace inversion of the moment generating function (MGF). This method has been improved by [Mar06] based on conditional independence models. Huang and Oosterlee ([Hua07a],[Hua07b]) perform a SP approximation based on the conditional independence framework and compare the results with the Normal Approximation (NA) and the Importance Sampling method. [Gla07] applies the methodology developed by [Aba00] to the multi-factor Merton model. First, the Bromwich integral is approximated by an infinite series using the trapezoidal rule and second, the convergence of the infinite series is accelerated by a method called Euler summation. They have shown that the cumulative distribution function (CDF) is comparatively accurate in the small loss region, whereas the accuracy worsens in the tail region. This is because the infinite series obtained by the Euler summation is an alternating series, each term of which has a very large absolute value. They also compare the results with the Recursive Approximation (RA) which is extremely slow in big portfolios.

Another approach to numerically invert the Laplace transform has been studied by [Hoo82] and [Ahn03] consisting in applying the Poisson algorithm to approximate the Bromwich integral by an infinite series, as in [Aba00] for then to use the quotient-difference (QD) algorithm to accelerate the slow convergence of the infinite series. We will refer to this approach as the *Hoog algorithm*. [Tak08] has applied this methodology to the multi-sector Merton model. The numerical examples presented show that in contrast with the Euler summation technique, Hoog algorithm is quite efficient in measuring tail probabilities.

The Granularity Adjustment (GA) has been developed as an extension of the ASRF model. As an asymptotic approximation, the GA formula might not work well on small portfolios. We give a brief revision of this method in Chapter 3.

In this dissertation we present a novel methodology for computing the risk measures and contributions in a factor model, through numerically inverting the Laplace transform of the CDF of the loss function, once we have approximated it by a finite sum of Haar wavelets basis functions. The idea is up to certain extent similar to the one in [Aba96], which uses Laguerre polynomials instead of wavelets. In the financial context, [Hav09] also performs a Laplace transform inversion for option pricing purposes using a series expansion in terms of the Franklin hat wavelets. The authors numerically compute the coefficients of the approximation by minimizing the average of squared errors between the true option prices and estimated prices. The technique to get the coefficients in our method is quite different in the sense that the analytical treatment that we implement allows us to give an expression of the wavelets coefficients in terms of the Cauchy's integral

theorem for then, to compute them, through an ordinary trapezoidal rule, avoiding this way the infinite series in [Gla07] and [Tak08]. The power of the Wavelet Approximation (WA) method mostly resides in the good balance between the computational time and the accuracy for both small and high loss levels as well as for a wide range of portfolios, independently of the concentration type and size. The Saddle Point approach, as an asymptotic method, tends in general to work better for high VaR confidence levels and when the size of the portfolio increases. Moreover, if the loss distribution is not smooth due to exposure concentration, a straightforward implementation may be insufficient. Finally, it is important to remark that Haar wavelets are naturally capable to reproduce the step-like form distribution derived from the Merton model, even when dealing with extremely small or concentrated portfolios.

It is very important to underline that the WA approach computes the whole loss function, and not only the risk measures, with almost no extra effort. This fact has implications in the pricing of a Collateralized Debt Obligation (CDO), where it is not sufficient to know a quantile of the loss function. Also, once the risk measures have been calculated, and following the Euler allocation principle, the risk contributions can be derived as the partial derivatives of the risk measure with respect to the exposures. Finally, we extend the wavelet approximation method to the multi-factor setting by means of Monte Carlo and quasi-Monte Carlo methods.

1.2 General Model Settings

To represent the uncertainty about future events, we specify a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with sample space Ω, σ-algebra $\mathcal{F}$, probability measure $\mathbb{P}$ and with filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions. We fix a time horizon $T > 0$. Usually T will equal one year.

1.2.1 Risk Parameters

As this dissertation focus on the quantification of name concentration risk, which can only be measured at portfolio level, we consider a credit portfolio consisting of N obligors. Any obligor n can be characterized by three parameters: the *exposure at default* (EAD) E_n, the *loss given default* (LGD) L_n and the *probability of default* (PD) P_n, each of them estimated from empirical default data. The EAD of an obligor denotes the portion of the exposure to the obligor which is lost in case of default. The LGD of a transaction describes the extent of the loss incurred in the event of default, also called *severity*. Finally, the PD measures the uncertainty whether an obligor will default or not, also termed *arrival risk*. Let us explain in detail the three risk parameters briefly described above.

The exposure at default consists in general in two parts, the *outstandings* and the *commitments*. The outstandings refers to the portion of the exposure already drawn by the obligor. In the case of borrower's default, the bank is exposed to the total amount of the outstandings. The commitments can be divided in two portions, *undrawn* and *drawn*, in the time before default. The total amount of commitments is the exposure the bank has promised to lend to the obligor at her or his request. Historical default experience

shows that obligors tend to draw on committed lines of credit in times of financial distress. Therefore, the commitment is also subject to loss in case of the obligor's default, but only the drawn (prior default) amount of the commitments will actually contribute to the loss on loan. The fraction describing the decomposition of commitments in drawn and undrawn portions is a random variable due to the optional character commitments have (the obligor has the right but no the obligation to draw on committed lines of credit). Therefore, it is natural to define the exposure at default by,

$$\text{EAD} = O + \eta \cdot C, \tag{1.1}$$

where O denotes the outstandings and C the commitments of the loan and η is the expected portion of the commitments likely to be drawn prior to default. More precisely, η is the expectation of the random variable capturing the uncertain part of the EAD, namely the utilization of the undrawn part of the commitments. Obviously, η takes place in the unit interval. As we are assuming the EAD to be a deterministic quantity, we directly deal with the expectation η, hereby ignoring the underlying random variable. For the IRB approach, the Basel Committee proposes that banks eligible for this approach will be permitted to use their own internal estimates of EAD for transactions with uncertain exposure. From this perspective, it makes much sense for major banks to carefully think about some rigorous methodology for calibrating EAD to borrower and facility specific characteristics. For example, banks that are able to calibrate the parameter η in (1.1) on a finer scale will have more accurate estimates of the EAD, better reflecting the underlying credit risk. The more the determination of regulatory capital tends towards risk sensitivity, the more will banks with advanced methodology benefit from a more sophisticated calibration of EAD.

The loss given default is far from being straightforward, because it depends on many driving factors, for example on the quality of *collateral* (securities, mortgages, guarantees, etc.) and on the seniority of the bank's claim on the borrower's assets. It is usually modeled as a random variable describing the severity of losses in the default event. In case the applied credit risk model admits only a constant value for LGD, one usually chooses the expectation of this severity. A bank external source for recovery data comes from the rating agencies. For example, Moody's provide recovery values of defaulted bonds, hereby distinguishing between different seniorities. Unfortunately many banks do not have good internal data for estimating recovery rates. In fact, although LGD is a key driver for risk measurement, there is little progress in comparison with other risk drivers like PD made in moving towards a sophisticated calibration.

Assigning a default probability to every borrower in the bank's credit portfolio is quite complicated. One approach is to calibrate default probabilities from market data. This is done, for example, in the concept of *expected default frequencies* from KMV Corporation. Another method is to compute the default probabilities from the credit risk inherent in certain credit spreads of traded products, e.g. credit derivatives such as *credit default swaps*. Default probabilities can also be calibrated from *ratings* that are assigned to borrowers either by external rating agencies such Moody's, Standard and Poor's or Fitch, or by bank-internal rating methods. Basically, ratings describe the *creditworthiness* of customers. Hereby quantitative as well as qualitative information is used to evaluate a

client. In practice, the rating procedure is often more based on the experience of the rating analyst rather than on pure mathematical procedures with strictly defined outcomes. In other words, statistical tools provide a first indication regarding the rating of a customer, but due to the various soft factors underlying a rating, the responsibility to assign a final rating remains the duty of the rating analyst. It is important to know that the rating agencies have established and ordered scale of ratings in terms of a letter system describing the creditworthiness of rated companies. The rating categories of Moody's and Standard and Poor's are slightly different, but it is not difficult to find a mapping between the two. Table 1.1 shows the Standard and Poor's rating categories. The process of assigning a default probability to a rating is called a *calibration*.

AAA	Best credit quality. Extremely reliable with regard to financial obligations.
AA	Very good credit quality. Very reliable.
A	More susceptible to economic conditions. Still good credit quality.
BBB	Lowest rating in investment grade.
BB	Caution is necessary. Best sub-investment credit quality.
B	Vulnerable to changes in economic conditions. Currently showing the ability to meet its financial obligations.
CCC	Currently vulnerable to non payment. Dependent on favourable economic conditions.
CC	Highly vulnerable to a payment default.
C	Close to or already bankrupt. Payments on the obligation currently continued.
D	Payment default on some financial obligation has actually occurred.

Table 1.1: Standard and Poor's rating categories.

1.2.2 Risk Measures and Contributions

Let us consider an obligor n subject to default in the fixed time horizon T. We introduce D_n, the default indicator of obligor n taking the following values:

$$D_n = \begin{cases} 1, & \text{if obligor } n \text{ is in default,} \\ 0, & \text{if obligor } n \text{ is not in default,} \end{cases}$$

where $\mathbb{P}(D_n = 1) = P_n$ and $\mathbb{P}(D_n = 0) = 1 - P_n$.

Let $\mathcal{L}$ be the portfolio loss given by,

$$\mathcal{L} = \sum_{n=1}^{N} \mathcal{L}_n,$$

where $\mathcal{L}_n = E_n \cdot L_n \cdot D_n$.

In general, credit risk can be split in *expected losses* (EL), which can be forecasted and thus can easily be managed, and *unexpected losses* (UL) which are more complicated

to quantify. The latter can be distinguished in *systematic risk* and *idiosyncratic risk*. The former arises from dependencies across individual obligors in the portfolio and from common shocks, while the latter arises from obligor specific shocks. Idiosyncratic risk can be diversified away, whereas systematic risk can not be eliminated but can be reduced by shifting exposures in a way to reduce correlation. *Economic capital* (EC) is held for unexpected losses that arise from systematic and idiosyncratic risk. From now on, we will always suppose that the following assumption holds.

Assumption 1.2.1. *The exposure at default E_n, the loss given default L_n and the default indicator D_n of an obligor n are independent.*

Denote by EL_n the expectation value of L_n, therefore,

$$\mathrm{EL} = \mathbb{E}(\mathcal{L}) = \sum_{n=1}^{N} E_n \cdot EL_n \cdot P_n.$$

The deviation of losses from the EL is usually measured by means of the standard deviation of the loss variable, however, since there is a significant likelihood that losses will exceed the portfolio's EL by more than one standard deviation of the portfolio loss, holding the UL of a portfolio as a risk capital for cases of financial distress might not be appropriate. The concept of EC is a widely used approach for bank internal credit risk models and will be defined latter.

Let $\alpha \in (0,1)$ be a given confidence level, the α-quantile of the loss distribution of $\mathcal{L}$ in this context is called *Value at Risk* (VaR). Thus,

$$\mathrm{VaR}_\alpha = \inf\{l \in \mathbb{R} : \mathbb{P}(\mathcal{L} \leq l) \geq \alpha\} = \inf\{l \in \mathbb{R} : F_{\mathcal{L}}(l) \geq \alpha\},$$

where $F_{\mathcal{L}}$ is the cumulative distribution function of the loss variable. Usually the α of interest is very close to 1. This is the measure chosen in the Basel II Accord (at a confidence level of $\alpha = 0.999$) for the computation of capital requirement, which means a bank that manages its risks with Basel II must to reserve capital of an amount of VaR_α to cover extreme loss. By its definition, VaR gives no information about the severity of losses which occur with a probability less than $1-\alpha$. If the loss distribution is heavy tailed, this can be quite problematic. This is a major drawback of the concept as a risk measure and also the main intention behind the innovation of the alternative risk measure *expected shortfall* (ES). Moreover, VaR is not a coherent risk measure since it is not sub-additive (see [Art97], [Art99], [Blu03]). This means that, if we have two loss distributions $F_{\mathcal{L}_1}$ and $F_{\mathcal{L}_2}$ for two portfolios and we denote the overall loss distribution of the merged portfolio $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$ by $F_{\mathcal{L}}$, then we do not necessarily have that $\mathrm{VaR}_\alpha(F_{\mathcal{L}}) \leq \mathrm{VaR}_\alpha(F_{\mathcal{L}_1}) + \mathrm{VaR}_\alpha(F_{\mathcal{L}_2})$. Hence, the VaR of the merged portfolio is not necessarily bounded above by the sum of the VaR of the individual portfolios which contradicts the intuition of diversification benefits associated with merging portfolios.

Example 1.2.1. *Consider two independent loans with default indicators following a Bernoulli distribution $B(1,p)$ with $0.006 \leq p < 0.01$ and exposures equal to 1. Define*

two portfolios A and B, each of them consisting of one unit of the above introduced loans. Then if we denote the corresponding portfolio losses by $\mathcal{L}_A$ *and* $\mathcal{L}_B$*,*

$$VaR_{0.99}(\mathcal{L}_A) = VaR_{0.99}(\mathcal{L}_B) = 0.$$

Now if we consider a portfolio C defined as the union of portfolios A and B and denote by $\mathcal{L}_C = \mathcal{L}_A + \mathcal{L}_B$*. Then,*

$$\mathbb{P}(\mathcal{L}_C = 0) = (1-p)^2 < 0.99,$$

and therefore,

$$VaR_{0.99}(\mathcal{L}_C) > 0,$$

so that,

$$VaR_{0.99}(\mathcal{L}_C) > VaR_{0.99}(\mathcal{L}_A) + VaR_{0.99}(\mathcal{L}_B).$$

Example 1.2.2. *Consider a portfolio of* 100 *defaultable corporate bonds. We assume that default of corporate bonds are independent, moreover the default probability is identical for all bonds and is equal to* 2%*. The current price of the bonds is* 100*. If there is no default, a bond pays in T an amount of* 105*, otherwise there is no repayment. Hence,* $\mathcal{L}_n$*, the loss of bond n, is equal to* 100 *when the bond defaults and to* -5 *otherwise. Denote by* D_n *the default indicator of firm n, i.e.* D_n *is equal to one if bond n defaults and equal to zero otherwise. We get* $\mathcal{L}_n = 100D_n - 5(1-D_n) = 105D_n - 5$*. Hence* $\mathcal{L}_n$ *is a sequence of i.i.d. random variables with* $\mathbb{P}(\mathcal{L}_n = -5) = 0.98$ *and* $\mathbb{P}(\mathcal{L}_n = 100) = 0.02$*. Let us compare now two portfolios, both with current value equal to* 10000*. Portfolio A is fully concentrated and consists of* 100 *units of bond one. Portfolio B is completely diversified, it consists of one unit of each of the bonds. Economic intuition suggests that portfolio B is less risky than portfolio A and hence it should have a lower VaR. Let us compute VaR at a confidence level of* 95% *for both portfolios. For portfolio A the loss is given by* $\mathcal{L}_A = 100\mathcal{L}_1$*, so* $VaR_{0.95}(\mathcal{L}_A) = 100\, VaR_{0.95}(\mathcal{L}_1)$*. Now* $\mathbb{P}(\mathcal{L}_1 \leq -5) = 0.98 \geq 0.95$ *and* $\mathbb{P}(\mathcal{L}_1 \leq l) = 0 < 0.95$ *for* $l < -5$*. Hence,* $VaR_{0.95}(\mathcal{L}_1) = -5$ *and therefore* $VaR_{0.95}(\mathcal{L}_A) = -500$*. For portfolio B we have,*

$$\mathcal{L}_B = \sum_{n=1}^{100} \mathcal{L}_n = 105 \sum_{n=1}^{100} D_n - 500.$$

Noting that the sum $\sum_{n=1}^{100} D_n$ *has a binomial distribution* $B(100, 0.02)$*, we get by inspection that* $\mathbb{P}(\sum_{n=1}^{100} D_n \leq 5) \simeq 0.984 \geq 0.95$ *and* $\mathbb{P}(\sum_{n=1}^{100} D_n \leq 4) \simeq 0.949 < 0.95$*, so* $VaR_{0.95}(\mathcal{L}_B) = 105 \cdot 5 - 500 = 25$*. In this case, a bank would need an additional risk capital of 25 to satisfy a regulatory working with VaR at the 95% level. Clearly, the risk capital required for portfolio B is higher than for portfolio A.*

The two examples presented above illustrate that measuring risk with VaR can lead to nonsensical results.

Definition 1.2.1. *Denote by* L^∞ *the space of bounded real random variables defined on the probability space stated before. A mapping* $\gamma : L^\infty \to \mathbb{R}$ *is called a coherent risk measure if the following properties hold,*

1. *Subadditivity: for all* $X, Y \in L^{\infty}, \gamma(X+Y) \leq \gamma(X) + \gamma(Y)$.
2. *Monotonicity: for all* $X, Y \in L^{\infty}$ *with* $X \leq Y a.s., \gamma(X) \leq \gamma(Y)$.
3. *Positive homogeneity: for all* $\lambda > 0, X \in L^{\infty}, \gamma(\lambda X) = \lambda\gamma(X)$.
4. *Translation invariance: for all* $x \in \mathbb{R}, X \in L^{\infty}, \gamma(X+x) = \gamma(X) + x$.

The first axiom, as mentioned before, reflects the fact that due to diversification effects the risk inherent in the union of two portfolios should be less than the sum of the two portfolios considered separately. The second axiom reflects the fact that if almost surely the losses of portfolio X are lower than the losses of portfolio Y then the required risk capital for portfolio X should be less than the required risk capital of portfolio Y. The homogeneity axiom states that if we have a portfolio with loss X and we scale all exposures by a factor of λ then, the loss X changes to a scaled loss λX. Accordingly, the originally required risk capital $\gamma(X)$ will also change to $\lambda\gamma(X)$. Finally, translation invariance axiom tell us that if x is some capital which will be lost (gained) on a portfolio with certainty at the considered horizon, then the risk capital required for covering losses in this portfolio can be increased (reduced) accordingly. Translation invariance implies the natural property $\gamma(X - \gamma(X)) = 0$ for every loss $X \in L^{\infty}$.

VaR as a risk measure defined on L^{∞} has translation invariance, is positively homogeneous and monotone, but not subadditive as showed in the examples above. [Tas00b] showed that expected shortfall to a great extent enjoys the coherence properties.

Expected shortfall can be interpreted as the expected loss that is incurred in the event that VaR is exceeded. Thus, the tail behavior of the loss distribution is taken into account. Formally, we define ES as follows,

$$\mathrm{ES}_{\alpha} = \mathbb{E}\left(\mathcal{L} | \mathcal{L} \geq \mathrm{VaR}_{\alpha}\right),$$

or alternatively,

$$\mathrm{ES}_{\alpha} = \frac{1}{1-\alpha} \int_{VaR_{\alpha}}^{+\infty} x f_{\mathcal{L}}(x) dx,$$

where $f_{\mathcal{L}}$ is the density function of the portfolio loss variable $\mathcal{L}$.

Another important risk measure, as mentioned before, is the so called economic capital EC_{α} for a given confidence level α. It is defined as the VaR at level α of the portfolio loss $\mathcal{L}$ minus the expected loss of the portfolio,

$$\mathrm{EC}_{\alpha} = \mathrm{VaR}_{\alpha} - \mathrm{EL}.$$

Hence, it represents the capital that a bank should reserve to limit the probability of default to a given confidence level. The VaR is reduced by the EL due to the common decomposition of total risk capital, that is VaR, into a part covering expected losses and a part reserved for unexpected losses.

In the worst case, a bank could lose its entire credit portfolio in a year. Holding capital against such an unlikely event is economically inefficient. Since banks want to spend most of their capital for profitable investments, there is a strong incentive to minimize the

capital a bank holds. Hence the problem of risk management in a financial institution is to find the balance between holding enough capital to be able to meet all debt obligations (also accounting for times of financial distress) on the one hand, but minimizing economic capital to make profits, on the other hand.

We can consider also how to decompose the risk measured by the VaR or the ES into individual transactions. In this purpose and following [Tas00a], we consider the allocation principle simply given by the partial derivative of the risk measure with respect the exposure of an obligor. Thus, let us define the *risk contribution to VaR* (VaRC) of obligor n at confidence level α by,

$$\mathrm{VaRC}_{\alpha,n} \equiv E_n \cdot \frac{\partial \mathrm{VaR}_\alpha}{\partial E_n}, \tag{1.2}$$

and the *risk contribution to ES* (ESC) of obligor n at confidence level α by,

$$\mathrm{ESC}_{\alpha,n} \equiv E_n \cdot \frac{\partial \mathrm{ES}_\alpha}{\partial E_n}. \tag{1.3}$$

These definitions satisfy the additivity condition,

$$\sum_{n=1}^{N} \mathrm{VaRC}_{\alpha,n} = \mathrm{VaR}_\alpha, \quad \sum_{n=1}^{N} \mathrm{ESC}_{\alpha,n} = \mathrm{ES}_\alpha.$$

1.3 The Merton Model

Credit risk models can be divided into two fundamental classes of models, *structural* or *asset-value models* and *reduced-form* or *default-rate models*.

Asset-value models have their origin on the famous Merton model, where the default of a firm is modeled in terms of the relationship between its assets and the liabilities that it faces at the end of a given time period. In models of this type default risk depends mainly on the stochastic evolution of the asset value and default occurs when the random variable describing this asset value falls below a certain threshold which represents the liabilities. Two famous industry models descending from the Merton approach are the KMV model and the CreditMetrics model, the first one developed by Moody's KMV and the second one by JPMorgan and the RiskMetrics Group.

In reduced-form models one directly models the process of credit defaults instead of constructing a stochastic process of the firm's asset value which indirectly leads to a model of the firm's default. This class of models is also called *mixture models*. The most representative industry model for this kind of models is CreditRisk$^+$, developed by Credit Suisse Financial Products.

Further information about these industry models can be found in [Lut09].

1.3.1 The General Framework

The model proposed by [Mer74] is the precursor of all asset-value models and, although a lot of extensions haven been developed since that time, the original Merton model remains

influential and is still popular in the practice of credit risk analysis. The Merton model assumes that the asset value of a firm follows an stochastic process $(V_t)_{t\geq 0}$. There are only two classes of securities; equity and debt. It is assumed that equity receives no dividends and that the firm cannot issue new debt. The model also assumes that the company's debt is given by a zero-coupon bond with face value B that will become due at a future time T. The firm defaults if the value of its assets is less than the promised debt repayment at time T. In the Merton model, default can occur only at the maturity T of the bond. Let us denote the value at time t of equity and debt by S_t and B_t, respectively. In the market, assuming that there are no taxes or transaction costs, the value of the firm's assets is given by the sum of debt and equity, i.e., $V_t = S_t + B_t,\ 0 \leq t \leq T$. At maturity there are only two possible scenarios:

(i) $V_T > B$: the value of the firm's assets exceeds the debt. In this case the debtholders receive $B_T = B$, the shareholders receive the residual value $S_T = V_T - B$ and there is no default.

(ii) $V_T \leq B$: the value of the firm's assets is less than its debt. Hence the firm cannot meet its financial obligations and defaults. In this case, the debtholders take ownership of the firm, and the shareholders are left with nothing, so that we have $B_T = V_T, S_T = 0$.

Combining the above two results, the payment to the shareholders at time T is given by,

$$S_T = \max(V_T - B, 0) = (V_T - B)^+, \tag{1.4}$$

and debtholders receive,

$$B_T = \min(V_T, B) = B - (B - V_T)^+. \tag{1.5}$$

This shows that the value of the firm's equity is the payoff of an European call option on the assets of the firm with strike price equal to the promised debt payment. By put-call parity, the firm's debt comprises a risk-free bond that guarantees payment of B plus a short European put option on the firm's assets with exercise price equal to the promised debt payment B. The Merton model thus treats the asset value V_t as any underlying. It assumes that under the real-world probability measure $\mathbb{P}$ the asset value process $(V_t)_{t\geq 0}$ follows a geometric Brownian motion of the form,

$$dV_t = \mu_V V_t dt + \sigma_V V_t dW_t, \quad 0 \leq t \leq T, \tag{1.6}$$

for constants $\mu_V \in \mathbb{R}$, $\sigma_V > 0$, and a standard Brownian motion $(W_t)_{t\geq 0}$. Furthermore, it makes all the other simplifying assumptions of the Black-Scholes option pricing formula (see [Bla73]). The solution at time T of the stochastic differential equation (1.6) with initial value V_0 can be computed and is given by,

$$V_T = V_0 e^{\left(\mu_V - \frac{1}{2}\sigma_V^2\right)T + \sigma_V W_T}.$$

This implies in particular that,

$$\ln V_T \sim N\left(\ln V_0 + \left(\mu_V - \frac{1}{2}\sigma_V^2\right)T, \sigma_V^2 T\right).$$

Hence the market value of the firm's equity at maturity T can be determined as the price of a European call option on the asset value V_t with exercise price B and maturity T. The risk neutral pricing theory then yields that the market value of equity at time $t < T$ can be computed as the discounted expectation (under the risk neutral equivalent measure $\mathbb{Q}$) of the payoff function (1.4), i.e.

$$S_t = \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}(V_T - B)^+ \middle| \mathcal{F}_t\right],$$

and is given by,

$$S_t = V_t \cdot \Phi(d_{t,1}) - B \cdot e^{-r(T-t)} \cdot \Phi(d_{t,2}),$$

where $d_{t,1} = \frac{\ln(V_t/B)+(r+\frac{1}{2}\sigma_V^2)(T-t)}{\sigma_V\sqrt{T-t}}$ and $d_{t,2} = d_{t,1} - \sigma_V\sqrt{T-t}$. Here r denotes the risk-free rate interest rate which is assumed to be constant. Also, according to equation (1.5) we can value the firm's debt at time $t \leq T$ as

$$\begin{aligned} B_t &= \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}(B - (B - V_T)^+) \middle| \mathcal{F}_t\right] \\ &= Be^{-r(T-t)} - \left(Be^{-r(T-t)}\Phi(-d_{t,2}) - V_t\Phi(-d_{t,1})\right). \end{aligned}$$

We discount the payment B at the risk-free rate because that payment is risk-free since we stripped out the credit risk as a put option.

The default probability of the firm by time T is the probability that shareholders will not exercise their call option to buy the assets of the company for B at time T, i.e. it is precisely the probability of the call option expiring *out-of-the-money*[1]. It can be computed as,

$$\mathbb{P}(V_T \leq B) = \mathbb{P}(\ln V_T \leq \ln B) = \Phi\left(\frac{\ln\frac{B}{V_0} - (\mu_V - \frac{1}{2}\sigma_V^2)T}{\sigma_V\sqrt{T}}\right). \tag{1.7}$$

Equation (1.7) shows that the default probability is increasing in B, decreasing in V_0 and μ_V and for $V_0 > T$, increasing in σ_V, which is all perfectly in line with economic intuition. Under the risk neutral measure $\mathbb{Q}$ we have,

$$\mathbb{Q}(V_T \leq B) = \mathbb{Q}\left(\frac{\ln\frac{B}{V_0} - (r - \frac{1}{2}\sigma_V^2)T}{\sigma_V\sqrt{T}} \leq -d_{0,2}\right) = 1 - \Phi(d_{0,2}).$$

Hence, the risk-neutral default probability, given information up to time t, is given by $1 - \Phi(d_{t,2})$.

[1] An option is said to be *in-the-money* if it has positive intrinsic value, respectively *out-of-the-money* if it has zero intrinsic value. A call is in-the-money if the value of the underlying is above the strike price. A put is in-the-money if the value of the underlying is below the strike price.

Although the Merton model provides a useful context for modeling credit risk and practical implementations of the model are being used by many financial institutions, it also has some drawbacks like to assume that the firm's debt financing consists of a one-year zero coupon bond, which is an oversimplification. Moreover, the simplifying assumptions of the Black-Scholes model are questionable in the context of corporate debt. In particular, the assumption of normally distributed losses can lead to an underestimation of the potential risk in a loan portfolio. Alternatively, [Alb07] describes the portfolio loss within the context of a one-factor Lévy model. Compared to a model with normally distributed asset returns, using a distribution with fatter tails as, for example, the Variance Gamma distribution, leads to an increase in the economic capital of the portfolio. Finally, and this might be the most important shortcoming of the Merton model, the firm's value is not observable which makes assigning values to it and its volatility problematic.

1.3.2 The Multi-Factor Merton Model

In credit risk analysis it is desirable to explain the firm's economic success by means of some global underlying influences. This leads to the introduction of the so-called *factor models*. Factor models provide a possibility to interpret the correlation between single loss variables in terms of some underlying economic variables such that large portfolio losses can be explained by these economic factors. The Merton model can be understood as a multi-factor model as it will be explained in this section. Here we focus only on the default mode version, where the Merton model is of Bernoulli type and the decision about default or survival of a firm at the end of a time period is made by comparing the firm's asset value with respect to a certain threshold value. If the firm value is below this threshold, the firm defaults and otherwise it survives.

Let us consider the portfolio of N obligors described in Section 1.2. Each of the obligors has exactly one loan with principal E_n. We fix a time horizon $T > 0$. The Merton model is a so-called asset-value model, meaning that the loss distribution is derived by focusing on a description of the firm's asset value. Therefore, we define $V_t^{(n)}$ to be the asset value of the counterparty n at time $t \leq T$. For every counterparty there exists a threshold T_n such that counterparty n defaults in the time period $[0, T]$ if $V_T^{(n)} < T_n$, i.e., when the asset value at maturity T is less than the threshold value. If we understand for example T_n as counterparty n's liabilities then, $V_T^{(n)}$ can be viewed as a latent variable driving the default event. Therefore, for $n = 1, \ldots, N$, we define,

$$D_n = \chi_{\{V_T^{(n)} < T_n\}} \sim B\left(1, \mathbb{P}(V_T^{(n)} < T_n)\right), \tag{1.8}$$

where $B(1, p)$ denotes the Bernoulli distribution such that the event 1 occurs with probability p.

Let us now consider the borrower n's asset-value log return r_n: $\log\left(V_T^{(n)}/V_0^{(n)}\right)$. In the factor model approach, the main assumption is that the asset value process depends on some underlying factors which represent the industrial and regional influences as well as on some idiosyncratic terms.

Assumption 1.3.1. *Asset returns r_n depend linearly on K standard normally distributed risk factors $X = (X_1, \ldots, X_K)$ affecting the borrowers' defaults in a systematic way as well*

as on a standard normally distributed idiosyncratic term ϵ_n. Moreover, ϵ_n are independent of the systematic factors X_k for every $k \in \{1, \dots, K\}$ and the ϵ_n are uncorrelated.

Under this assumption and after standardization, the borrower n's asset value *log*-return admits a representation of the form,

$$r_n = \beta_n Y_n + \sqrt{1 - \beta_n^2}\epsilon_n, \tag{1.9}$$

where Y_n denotes the firm's *composite factor* and ϵ_n represents the *idiosyncratic shock*. The factor loading β_n illustrates borrower n's sensitivity to the systematic risk. Hence, it captures the linear correlation between r_n and Y_n. Y_n can be decomposed into K independent factors $X = (X_1, \dots, X_K)$ by,

$$Y_n = \sum_{k=1}^{K} \alpha_{n,k} X_k.$$

The weights $\alpha_{n,k}$ describe the dependence of obligor n on an industrial or regional sector k represented by factor X_k. Since the idiosyncratic shocks and the risk factors are assumed to be independent, the correlation of the counterparties' asset returns depend only on the correlation of the composite factors Y_n. Computing the variances of both sides of equation (1.9) yields,

$$\mathbb{V}(r_n) = \beta_n^2 \cdot \mathbb{V}(Y_n) + (1 - \beta_n^2) \cdot \mathbb{V}(\epsilon_n).$$

This expression can be understood as splitting the total risk of counterparty n into a *systematic* and an *idiosyncratic* risk component. Then, the quantity $\beta_n^2 \cdot \mathbb{V}(Y_n)$ expresses how much of the volatility of r_n can be explained by the volatility of Y_n and, thus, quantifies the systematic risk of the counterparty n. The term $(1 - \beta_n^2) \cdot \mathbb{V}(\epsilon_n)$ captures the idiosyncratic risk which cannot be explained by the common factors X_k. Since we assumed that the asset returns r_n, the systematic risk factors X_k and the idyosincratic terms ϵ_n are all standard normally distributed, we have to make sure that Y_n has unit variance. Therefore, the coefficients $\alpha_{n,k}$ must satisfy $\sum_{k=1}^{K} \alpha_{n,k}^2 = 1$.

Now we can rewrite equation (1.8) as,

$$D_n = \chi_{\{r_n < t_n\}} \sim B(1, \mathbb{P}(r_n < t_n)),$$

where t_n is the threshold corresponding to T_n after exchanging $V_T^{(n)}$ by r_n. Assuming the time horizon to equal 1 year, i.e. $T = 1$. We denote the one-year default probability of obligor n, by P_n. We have $P_n = \mathbb{P}(r_n < t_n)$ and, since $r_n \sim \mathcal{N}(0, 1)$, we obtain $t_n = \Phi^{-1}(P_n)$, where Φ denotes the cumulative standard normal distribution function and Φ^{-1} its inverse. Hence, the default probability of obligor n conditional on a specification $y_n = \sum_{k=1}^{K} \alpha_{n,k} x_k$ can be written as,

$$P_n(y_n) \equiv \Phi\left(\frac{t_n - \beta_n y_n}{\sqrt{1 - \beta_n^2}}\right). \tag{1.10}$$

The left plot of Figure 1.1 represents the dependence of the conditional default probability on the state of the economy in a single factor setting. Here we used an unconditional

default probability of 0.01 and an asset correlation of $\beta^2 = 0.15$. The right plot shows the dependence of the conditional default probability $P_n(y)$ with respect to the unconditional default probability P_n for a fixed correlation of $\beta^2 = 0.15$ and three different stages of the economy. The solid graph corresponds to a bad state of the economy where the systematic risk factor takes the values $y = -4$. The dotted graph corresponds to a risk factor $y = 0$ and the dotted-dashed graph to a good state of the economy with $y = 5$.

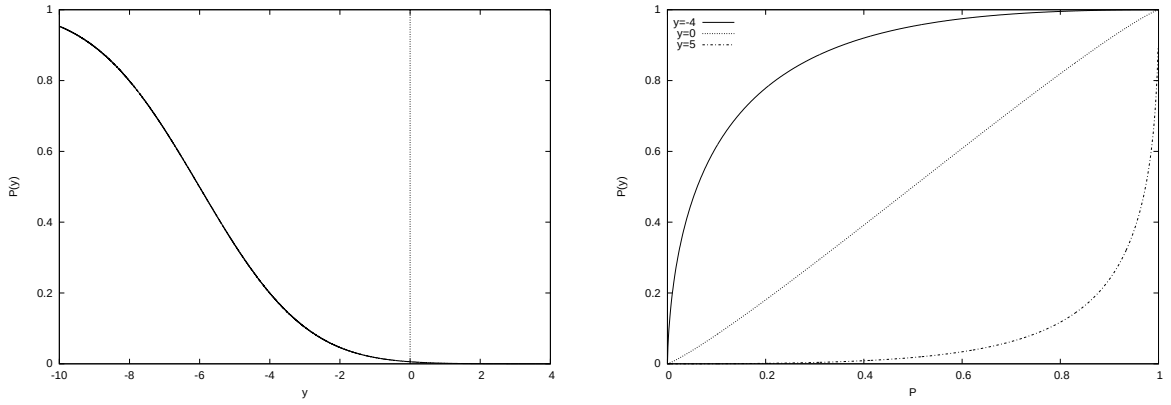

Figure 1.1: Conditional default probabilities.

The next purpose is to find an expression for the portfolio loss variable $\mathcal{L}$. Since L_n is assumed to be independent from the default indicator D_n and, since the conditional expectation of the default indicator equals the probability that r_n lies below the threshold t_n conditional on the risk factors, we obtain,

$$\mathbb{E}\left(\mathcal{L}|Y_n = y_n\right) = \sum_{n=1}^{N} E_n \cdot EL_n \cdot \Phi\left(\frac{t_n - \beta_n y_n}{\sqrt{1-\beta_n^2}}\right). \tag{1.11}$$

The determination of the portfolio loss distribution requires a Monte Carlo simulation of the systematic risk factors $X_1, \ldots, X_K$. The default indicator variables D_n are Bernoulli distributed with parameter $P_n = \mathbb{P}(D_n = 1)$ such that for any $d_n \in \{0, 1\}$ we have $\mathbb{P}(D_n = d_n) = P_n^{d_n} \cdot (1 - P_n)^{1-d_n}$. Now, instead of the default probabilities P_n, consider the conditional default probabilities $P_n(y_n)$ given by the composite factor Y_n. Then we can compute the joint distribution of the default indicator variables by integrating the conditioning on the factors Y_n as,

$$\mathbb{P}(D_1 = d_1, \ldots, D_N = d_N) = \int_{\mathbb{R}^N} \prod_{n=1}^{N} P_n(y)^{d_n} \cdot (1 - P_n(y))^{1-d_n} dF_Y(y),$$

for $d_n \in \{0, 1\}$ and $n \in \{1, \ldots, N\}$, where F_Y denotes the distribution function of the composite factors $Y_1, \ldots, Y_N$. Now we make the substitution $q_n = P_n(y)$, i.e. $y = P_n^{-1}(q_n)$. Note that the joint distribution function of the q_n's is given by the joint distribution

function of the conditional P_n's which are normally distributed. So we obtain,

$$\mathbb{P}(D_1 = d_1, \ldots, D_N = d_N) = \int_{[0,1]^N} \prod_{n=1}^{N} q_n^{d_n} \cdot (1 - q_n)^{1-d_n} dF(q_1, \ldots, q_N),$$

with the distribution function F explicitly given by,

$$F(q_1, \ldots, q_N) = \Phi_N(P_1^{-1}(q_1), \ldots, P_N^{-1}(q_N); \Gamma),$$

where $\Phi_N(\cdot; \Gamma)$ denotes the cumulative multivariate centered Gaussian distribution with correlation matrix $\Gamma = (\gamma_{nm})_{1 \leq n,m \leq N}$ containing the correlation of the asset returns r_n.

Assuming a constant loss given default equal to L_n for obligor n, the portfolio loss distribution can then be derived from the joint distribution of the default indicators as,

$$\mathbb{P}(\mathcal{L} \leq l) = \sum_{\substack{(d_1,\ldots,d_N) \in \{0,1\}^N \\ \sum_{n=1}^{N} E_n \cdot L_n \cdot d_n \leq l}} \left(\sum_{n=1}^{N} E_n \cdot L_n \cdot d_n \right) \cdot \mathbb{P}(D_1 = d_1, \ldots, D_N = d_N).$$

Remark 1.3.1. *In the next section we present an analytical approximation for the q^{th} percentile of the loss distribution in the one-factor framework, under the assumption that portfolios are infinitely fine-grained such that the idiosyncratic risk is completely diversified.*

1.4 The Asymptotic Single Risk Factor Model

The already mentioned Revised Framework incorporates new developments in credit risk management as it is more flexible and risk sensitive than the former Basel I accord. Moreover, within Basel II banks may opt for the standard approach, which is quite conservative with respect to capital charge, and the more advanced IRB approach when calculating regulatory capital for credit risk. Financial companies that opt for the IRB approach are allowed to use their own internal credit risk measures as inputs to the capital calculation whenever these are approved by the supervisory authorities. Therefore banks have to prove that certain conditions concerning the method and transparency are fulfilled. In the IRB approach, banks are allowed to determine the borrower's default probabilities using their own methods while those using the advanced IRB approach are further permitted to provide own estimates of LGD and EAD parameters. The Basel II risk weight formulas then translate these risk measures into risk weights and regulatory capital requirements which are intended to ensure that unexpected losses can be covered up to a certain confidence level prescribed by the supervisors.

The risk weight formulas for the computation of regulatory capital for unexpected losses are based on the already mentioned Asymptotic Single Risk Factor model developed by the Basel Committee and in particular by [Gor03]. This model was constructed in a way to ensure that the capital required for any risky loan should not depend on the particular portfolio decomposition it is added to. This so-called *portfolio invariance* was

necessary for reasons of applicability relying on straightforward and fast computations of capital requirements. However, portfolio invariance comes along with some drawbacks as it makes recognition of diversifications effects very difficult. Judging whether a loan fits well into an existing portfolio requires the knowledge of the portfolio decomposition and therefore contradicts portfolio invariance. The ASRF model is based on the assumption of a well diversified portfolio. Actually, no real bank can exactly fulfill this assumption and therefore banks are expected to account for this existence of concentration risk in Pillar two. Models for this task will be presented in the following chapters.

The ASRF model developed by [Gor03] is based on the law of large numbers. An ordinary portfolio consists of a large number of exposures of different sizes. When exposure sizes are equally distributed, the idiosyncratic risk associated with every single exposure is almost diversified away, meaning that the idiosyncratic risk cancel out one-another. Note that this requires a very huge portfolio. When idiosyncratic risk is diversified away, only systematic risk affecting many exposures remains. Such a portfolio is also called *infinitely fine-grained*. It is needless to say that such perfectly fine-grained portfolios does not exist in practice. Real bank portfolios have a finite number of obligors and lumpy distributions of exposure sizes. The asymptotic assumption might be approximately valid for some of the largest bank portfolios, but clearly would be much less satisfactory for portfolios of smaller or more specialized institutions. Thus, any capital charges computed under the assumption of an asymptotically fine-grained portfolio can underestimate the required capital for a real finite portfolio. Therefore, banks have to account for this non-diversified idiosyncratic risk under Pillar two. In the ASRF model all systematic risks in a portfolio are modeled by a single systematic risk factor representing for example the economic cycle. The two following assumptions are the two main assumptions of the ASRF.

Assumption 1.4.1. *1. Portfolios are infinitely fine-grained, i.e. no exposure accounts for more than an arbitrarily small share of total portfolio exposure.*

2. Dependence across exposures is driven by a single systematic risk factor Y.

Consider the portfolio of N risky loans described in Section 1.2. We express loss not in absolute value but in percentage of total exposure. Therefore, if we denote the exposure share of obligor n by $s_n = \frac{E_n}{\sum_{n=1}^{N} E_n}$, then the *portfolio loss ratio* $\mathcal{L}$ is given by $\mathcal{L} = \sum_{n=1}^{N} D_n \cdot L_n \cdot s_n$. The first condition in Assumption 1.4.1 is satisfied when the sequence of positive constant exposures E_n satisfies the following conditions.

Assumption 1.4.2. *1. $\sum_{n=1}^{N} E_n \uparrow \infty$ and*

2. there exist a positive ζ such that the largest exposure share is of order $\mathcal{O}(N^{-(\frac{1}{2}+\zeta)})$.

These assumptions are sufficient to guarantee that the share of the largest single exposure of the total portfolio exposure vanishes to zero as the number of exposures in the portfolio increases. As a practical matter, the restrictions are quite weak and would be satisfied by any conceivable real-word large bank portfolio. Under these conditions the following theorem follows directly from the strong law of large numbers (see [Gor03], Proposition 1 for a formal proof).

Theorem 1.4.1. *Under assumptions 1.4.1 and 1.4.2 the portfolio loss ratio conditional on any realization y of the systematic risk factor Y satisfies,*

$$\mathcal{L} - \mathbb{E}(\mathcal{L}|Y = y) \to 0 \text{ almost surely as } N \to \infty.$$

In intuitive terms, this theorem says that as the exposure share of each asset in the portfolio goes to zero, idiosyncratic risk in portfolio loss is diversified away perfectly. In the limit, the loss ratio converges to a fixed function of the systematic factor Y given by $\mathbb{E}(\mathcal{L}|Y)$. We refer to this limiting portfolio as *infinitely fine-grained* or as an *asymptotic portfolio*.

If we now want a practical model for the calculation of regulatory capital, we first have to find a way to derive conditional default probabilities. This is done by an adaptation of Merton's model. We recall that in the Merton model an obligor defaults if its asset value falls below a threshold given by its obligations and within the Merton model the asset value is described by a normally distributed random variable. The Basel Committee adopted this assumption of normally distributed risk factors.

Hence, the ASRF model can be described as a factor model such that the return on the firm's assets is of the form (1.9), i.e.,

$$r_n = \sqrt{\rho_n} Y + \sqrt{1 - \rho_n}\epsilon_n, \quad n = 1, \ldots, N,$$

with normally distributed systematic risk factor Y an idiosyncratic shocks ϵ_n. Here r_n denotes the log-asset return of obligor n and ρ_n captures the correlation between r_n and the single risk factor Y:

$$\begin{aligned} corr(r_n, Y) &= \frac{cov(r_n, Y)}{dev(r_n) \cdot dev(Y)} = \mathbb{E}(r_n \cdot Y) - \mathbb{E}(r_n) \cdot \mathbb{E}(Y) \\ &= \mathbb{E}\left(\left(\sqrt{\rho_n} Y + \sqrt{1 - \rho_n}\epsilon_n\right) \cdot Y\right) = \sqrt{\rho_n} \cdot \mathbb{E}\left(Y^2\right) = \sqrt{\rho_n}, \end{aligned}$$

due to the fact that $1 = \mathbb{V}(Y) = \mathbb{E}\left(Y^2\right) - (\mathbb{E}(Y))^2$ and $\mathbb{E}(Y) = 0$.

Moreover, the return on the firm's assets of two obligors are correlated as follows,

$$\begin{aligned} corr(r_n, r_m) &= \frac{cov(r_n, r_m)}{dev(r_n) \cdot dev(r_m)} = \mathbb{E}(r_n \cdot r_m) - \mathbb{E}(r_n) \cdot \mathbb{E}(r_m) \\ &= \mathbb{E}\left(\left(\sqrt{\rho_n} Y + \sqrt{1 - \rho_n}\epsilon_n\right) \cdot \left(\sqrt{\rho_m} Y + \sqrt{1 - \rho_m}\epsilon_m\right)\right) \\ &= \mathbb{E}\Big(\sqrt{\rho_n \rho_m} Y^2 + \sqrt{\rho_n(1 - \rho_m)} Y \cdot \epsilon_m + \sqrt{(1 - \rho_n)\rho_m}\epsilon_n \cdot Y + \\ &+ \sqrt{(1 - \rho_n)(1 - \rho_m)}\epsilon_n \cdot \epsilon_m\Big) \\ &= \sqrt{\rho_n \rho_m}. \end{aligned}$$

If the loss given default L_n is assumed to be deterministic, the conditional expectation can be written as in (1.11),

$$\mathbb{E}\left(\mathcal{L}|Y = y\right) = \sum_{n=1}^{N} s_n \cdot L_n \cdot \Phi\left(\frac{t_n - \sqrt{\rho_n} y}{\sqrt{1 - \rho_n}}\right). \tag{1.12}$$

For quantification of the credit risk, the Value at Risk on confidence level α can be used, that is the α-quantile VaR_α of the loss variable, in which $\alpha \in (0,1)$ is the target solvency probability. In the context of the ASRF framework, taking into account the Theorem 1.4.1 and the fact that the α-quantile of $\mathcal{L}$ is simply the $(1-\alpha)$-quantile of the systematic risk factor Y, we have,

$$\mathrm{VaR}_\alpha(\mathcal{L}) - \mathbb{E}(\mathcal{L}|Y = l_{1-\alpha}(Y)) \to 0 \text{ almost surely as } N \to \infty, \tag{1.13}$$

where $l_{1-\alpha}(Y)$ stands for the $(1-\alpha)$-quantile of the systematic risk factor Y. By (1.12), (1.13) and recalling that Y is normally distributed, the VaR of the portfolio $\mathcal{L}$ equals,

$$\mathrm{VaR}_\alpha^A = \sum_{n=1}^{N} s_n \cdot L_n \cdot \Phi\left(\frac{t_n + \sqrt{\rho_n}\Phi^{-1}(\alpha)}{\sqrt{1-\rho_n}}\right), \tag{1.14}$$

where the index A denotes that the risk measure has been calculated by means of the ASRF method. The VaR contributions for obligor n can now be calculated easily making use of expression (1.2),

$$\mathrm{VaRC}_{\alpha,n}^A = s_n \cdot \frac{\partial \mathrm{VaR}_\alpha^A}{\partial s_n} = s_n \cdot L_n \cdot \Phi\left(\frac{t_n + \sqrt{\rho_n}\Phi^{-1}(\alpha)}{\sqrt{1-\rho_n}}\right).$$

We have already noted that in general, VaR is not a coherent risk measure because it is not necessarily sub-additive. As long as we stay in the ASRF framework, this characteristic is not problematic because in this context the VaR is exactly sub-additive[2]. But if we leave the ASRF framework, this behavior is not guaranteed anymore. For this reason in the following chapters we will also consider the ES as an alternatively risk measure to deal with risk concentrated portfolios.

The variable ρ_n describes the degree of the obligor's exposure to the systematic risk factor by means of the asset correlation and is determined by the borrower's asset class. In the Revised Framework the asset value correlation for a given asset class, represented by the according unconditional default probability P_n, is given by,

$$\rho_n = 0.12 \cdot \frac{1-e^{-50\cdot P_n}}{1-e^{-50}} + 0.24 \cdot \left(1 - \frac{1-e^{-50\cdot P_n}}{1-e^{-50}}\right). \tag{1.15}$$

We see that asset correlation ρ_n decrease with increasing P_n. This is based both on empirical evidence and intuition. Intuitively, for instance, the effect can be explained as follows: the higher the P_n, the higher the idiosyncratic (individual) risk components of a borrower, i.e., the default risk depends less on the overall state of the economy and more on individual risk drivers. The behavior of the correlation coefficient ρ_n with respect P_n is shown in Figure 1.2. The dashed lines denote the upper and lower bound of the correlation coefficient which are given by 24% and 12% respectively. Correlations between these limits are given by the exponential weighting function (1.15) describing the dependence on the probability of default.

[2]This can be seen in formula 1.13 considering that the expectation operator is additive.

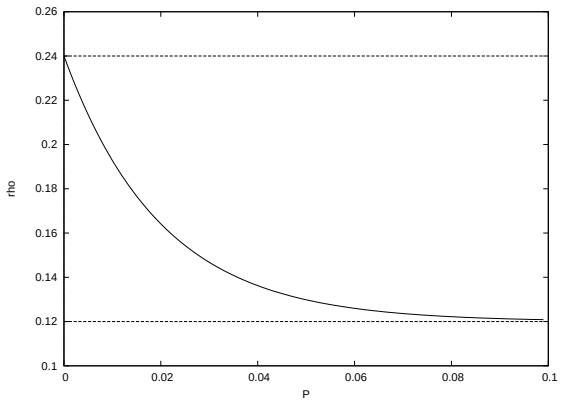

Figure 1.2: Correlation in dependence of the default probability.

Now, we want to compute the loss distribution function in an infinitely granular portfolio. For simplicity we start assuming a homogeneous portfolio in the sense that all obligors have the same default probability $P_n = P$, for $n = 1, \ldots, N$, and we assume that $L_n = 100\%$ for $n = 1, \ldots, N$. The correlation parameter ρ is constant since default probabilities are also constant. Then,

$$\mathbb{E}\left(\mathcal{L}|Y = y\right) = \sum_{n=1}^{N} s_n \cdot \Phi\left(\frac{\Phi^{-1}(P) - \sqrt{\rho}y}{\sqrt{1-\rho}}\right) = \Phi\left(\frac{\Phi^{-1}(P) - \sqrt{\rho}y}{\sqrt{1-\rho}}\right),$$

since the sum over the exposure shares equals 1. Hence, the percentage portfolio loss $\mathcal{L}$ tends to the conditional default probability $\mathcal{L} \to P(y)$ almost surely as $N \to \infty$. Thus, in an infinitely fine-grained portfolio, the conditional default probability $P(y)$ describes the fraction of defaulted obligors. Denote the percentage number of defaults in such a portfolio by $\mathcal{L}$. Then we can compute the probability density function $f_{\mathcal{L}}^{A}$ of $\mathcal{L}$ as follows. For every $0 \leq x \leq 1$ we have,

$$\begin{aligned} F_{\mathcal{L}}^{A}(x) \equiv \mathbb{P}(\mathcal{L} \leq x) = \mathbb{P}(P(y) \leq x) &= \mathbb{P}\left(-y \leq \frac{1}{\sqrt{\rho}}\left(\sqrt{1-\rho}\cdot\Phi^{-1}(x) - \Phi^{-1}(P)\right)\right) \\ &= \Phi\left(\frac{1}{\sqrt{\rho}}\left(\sqrt{1-\rho}\cdot\Phi^{-1}(x) - \Phi^{-1}(P)\right)\right), \end{aligned}$$

and differentiating respect to x we obtain the probability density function,

$$\begin{aligned} f_{\mathcal{L}}^{A}(x) = \frac{\partial F^{A}(x)}{\partial x} &= \sqrt{\frac{1-\rho}{\rho}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\rho}\left(\sqrt{1-\rho}\cdot\Phi^{-1}(x) - \Phi^{-1}(P)\right)^2} \cdot \frac{\partial}{\partial x}\left(\Phi^{-1}(x)\right) \\ &= \sqrt{\frac{1-\rho}{\rho}} \cdot e^{-\frac{1}{2\rho}\left(\sqrt{1-\rho}\cdot\Phi^{-1}(x) - \Phi^{-1}(P)\right)^2} \cdot e^{\frac{1}{2}\left(\Phi^{-1}(x)\right)^2}. \end{aligned}$$

Let us now consider an infinitely granular portfolio A with a default probability of 5% and a correlation of $\rho = 13\%$ and an infinitely granular portfolio B with default probability

of 10% and a correlation of 12%. Here, the correlation values have been computed via the correlation function (1.15). The left plot of Figure 1.3 shows the graph of the density function $f_{\mathcal{L}}^{A}$, represented by a continuous line for portfolio A and a dashed line for portfolio B. The VaR value computed at 95% confidence level for portfolio A is 0.13 and the VaR for portfolio B at the same confidence level is 0.22. The right plot of Figure 1.3 shows the cumulative distribution function of both portfolios.

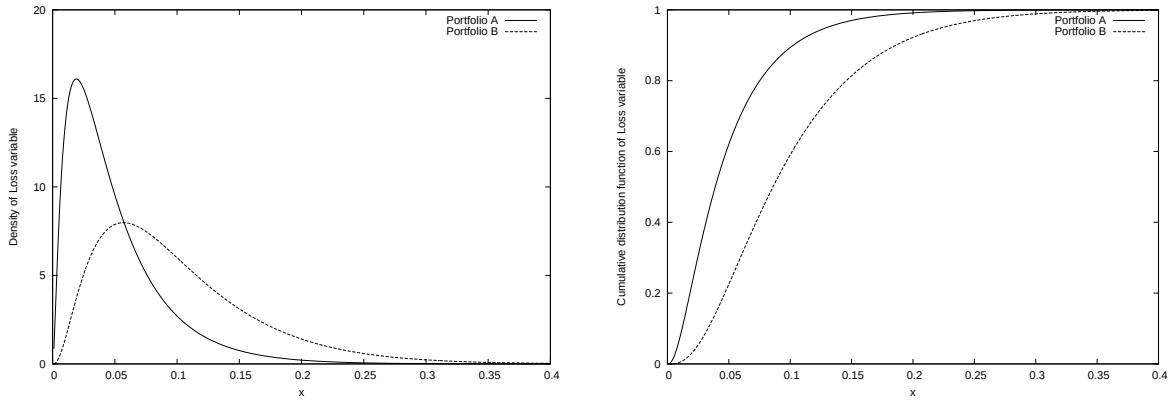

Figure 1.3: Densities and distributions for portfolios A and B.

1.5 Concentration Risk

Historical experience shows that concentration of risk in asset portfolios has been one of the major causes of bank distress. Failures of large borrowers like Enron or Parmalat where the source of sizable losses in a number of banks. Furthermore, the relevance of sector concentration risk is demonstrated by the recent developments in conjunction with the sub-prime mortgage crisis.

Let us explain the issue of concentration risk in more detail based on the example of the Asymptotic Single Risk Factor model. Under the ASRF framework it is assumed that bank portfolios are perfectly fine-grained, that is, idiosyncratic risk has been fully diversified away, so that economic capital depends only on systematic risk. Real world portfolios are not, of course, perfectly fine-grained. The asymptotic assumption might be approximately valid for some of the largest bank portfolios, but clearly would be much less satisfactory for portfolios of smaller or more specialized institutions. When there are material name concentrations of exposure, there will be a residual of undiversified idiosyncratic risk in the portfolio, and the IRB formula will underestimate the required economic capital. Moreover, the single risk factor assumption of the ASRF model does not allow for the explicit measurement of sector concentration risk.

There are in the literature some simple ad-hoc measures for concentration risk. All of them have advantages and drawbacks based on a set of desirable properties which ensure a consistent measurement of concentration risk. We give a brief overview of these simple measures.

A concentration index for a portfolio of N loans should satisfy the following properties:

1. The reduction of a loan exposure and an equal increase of a bigger loan must not decrease the concentration measure (*transfer principle*).

2. The measure of concentration attains its minimum value, when all loans are of equal size (*uniform distribution principle*).

3. If two portfolios, which are composed of the same number of loans, satisfy that the aggregate size of the k biggest loans of the first portfolio is greater or equal to the size of the k biggest loans in the second credit portfolio for $1 \leq k \leq N$, then the same inequality must hold between the measures of concentration in the two portfolios (*Lorenz-criterion*).

4. If two or more loans are merged, the measure of concentration must not decrease (*super-additivity*).

5. Consider a portfolio consisting of loans of equal size. The measure of concentration must not increase with an increase in the number of loans (*independence of loan quantity*).

6. Granting an additional loan of a relatively low amount does not increase the concentration measure (*irrelevance of small exposures*).

It can be proved that if an index satisfy properties 1 and 6 then it satisfies all six properties.

Common used concentration indexes are *Concentration Ratio*, *Lorenz Curve*, *Gini Coefficient* and the *Herfindahl-Hirschman Index* (HHI) which is used extensively particularly in the empirical literature. For this reason, later in Chapter 3, we use the HHI index to measure the concentration of the sample portfolios. The HHI index is defined as the sum of squared market shares, measured in fractions of the total portfolio, of each market participant,

$$HHI = \frac{\sum_{n=1}^{N} E_n^2}{(\sum_{n=1}^{N} E_n)^2} = \sum_{n=1}^{N} s_n^2.$$

The HHI index is a continuous measure with zero corresponding to the fully granular case (each participant has an infinitesimal share) and unity corresponding to monopoly (there is only one participant). Holding all else equal, the closer the HHI of a portfolio is to 1, the more concentrated the portfolio is. It can be shown that the HHI index satisfies the all six properties stated above.

However, neither the HHI index nor the rest of the mentioned indexes can incorporate the effects of obligor specific credit qualities, which are, for example, represented by obligor specific default probabilities. For these reasons, certain model-based methods for measuring concentration risks have been developed which can deal more explicitly with exposure distribution, credit quality and default dependencies. In the following chapters we will discuss some of these model-based approaches for the measurement of concentration risk.

The aim of this dissertation is to present a novel methodology based on wavelets to approximate the loss distribution function of a credit portfolio in presence of exposure concentration. We refer it as the Wavelet Approximation (WA) method. The WA method is especially suitable for small or name concentrated portfolios where the ASRF model tends to fail. In order to introduce the methodology in Chapter 2 we present a numerical Laplace transform inversion method based on the Haar wavelets especially suitable for stepped shape functions. Then, in Chapter 3, we approximate the loss function of a credit portfolio by a summation of Haar wavelet basis and calculate the VaR value once we have recovered the loss distribution function. Finally in Chapter 4 we make an extension of this new method for computing the ES measure and the risk contributions to the VaR and the expected shortfall. Furthermore, we present an improvement in the Gauss-Hermite quadrature that allows us to obtain, in an efficient and fast way when compared to conventional Monte Carlo methods, the risk measures and contributions.

Chapter 2

On the Haar Wavelets Method for Numerically Inverting the Laplace Transform

2.1 Introduction

This chapter is devoted to the development of a new method based on wavelets for numerically inverting the Laplace transform, which is especially suited for stepped shape functions. First of all, we review the Haar wavelets which underlies the approximation presented to carry out the inversion. Second, we present the most popular methods for inverting the Laplace transform and finally we introduce the Wavelet Approximation method.

2.2 Numerical Analysis of Wavelet Methods

Since the 1960's, multi-scale methods have been used in numerous areas of applied mathematics as diverse as signal analysis, statistics, computer aided geometric design, image processing and numerical analysis. The mathematical background underlying these methods was substantially reinforced with the emergence of wavelet basis in the 1980's.

Roughly speaking, multi-scale methods are based on approximations $(f_j)_{j\geq 0}$ to the data (or the unknown function) f of a given problem, at various resolution levels indexed by j. The corresponding scales of resolution h_j are usually chosen to be of order 2^{-j} and the approximation f_j should thus be viewed as a sketchy picture of f that cannot oscillate at a frequency higher than 2^j. As an example, if f is a univariate continuous function, one could choose for f_j the unique function such that $f_j(2^{-j}k) = f(2^{-j}k)$ and such that f_j is affine when restricted to $[\frac{k}{2^j}, \frac{k+1}{2^j}]$ for all $k \in \mathbb{Z}$. Formally, one obtains a multi-scale decomposition by expanding f into the sum of its coarsest approximation and additional details,

$$f = f_0 + \sum_{j=0}^{\infty} g_j,$$

where each $g_j = f_{j+1} - f_j$ represents the fluctuation of f between the two successive levels of resolution j and $j+1$.

In practice, these approximations and decompositions can be defined and implemented in various ways. In contrast with the one and unique Fourier transform, a *multi-scale transform* can be picked out among a versatile collection of mathematical tools. Furthermore, some of them can be implemented by means of fast algorithms, and therefore they are more appealing for numerical applications. In the context of numerical computations, one can be interested in further decomposing each fluctuation g_j into local contributions. For specific types of multi-resolution approximations, this task can be achieved using a *wavelet basis* for this purpose: one introduces an appropriate *mother* function ψ (in the case where f is a univariate function) that is well localized both in space and frequency, oscillates in the sense that $\int \psi = 0$, and allows us to expand g_j according to,

$$g_j = \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k},$$

where $d_{j,k}$ are scalar coefficients and each wavelet $\psi_{j,k}(x) = 2^{\frac{j}{2}} \psi(2^j x - k)$ contributes to the fluctuation of f at scale 2^{-j} in a neighborhood of size $2^{-j}|supp(\psi)|$ around the point $2^{-j}k$.

The word *wavelet* is used in mathematical analysis to denote a kind of orthonormal basis in L^2 with remarkable approximation properties. The theory of wavelets was developed by Y. Meyer, I. Daubechies, S. Mallat and others at the end of 1980's (see [Mey97], [Dau92], [Mal98], [Coh03] and [Wal02]).

Qualitatively, the difference between the usual sine wave and a wavelet may be described by the localization property: the sine wave is localized in frequency domain, but not in time domain, while a wavelet is localized both in frequency and time domain. By saying localized frequency we do not mean that the support of a wavelet is compact, we rather mean that the mass of oscillations of a wavelet is concentrated on a small interval. Clearly this is not the case for a sine wave. The Fourier orthonormal basis is composed of waves, while the aim of the theory of wavelets is to construct orthonormal basis composed of wavelets.

Besides the already discussed localization property of wavelets there are other remarkable features of this technique. Wavelets provide a useful tool in data compression and have excellent statistical properties in data smoothing.

Wavelets allow to simplify the description of a complicated function in terms of a small number of coefficients. Often the necessary coefficients needed are less than in the classical Fourier approximation.

Example 2.2.1. *Let $f(x)$ be of the form shown in Figure 2.1. The function is*

$$f(x) = \begin{cases} \sin(8\pi x), & x \in [0, \frac{1}{2}], \\ \sin(32\pi x), & x \in (\frac{1}{2}, 1], \\ 0, & otherwise. \end{cases}$$

The support of f is made of two intervals $[a,b]=[0,1/2]$ and $[c,d]=[1/2,1]$. On $[a,b]$ the frequency of oscillations of f is smaller than on $[c,d]$. When doing the Fourier expansion, we should include both frequencies, ω_1, the frequency of [a,b] and ω_2, the frequency of [c,d]. But since the sine waves have infinite support, we are forced to compensate the influence of ω_1 on $[c,d]$ and ω_2 on $[a,b]$ by adding a large number of higher frequency terms in the Fourier expansion. Using wavelets, only we need essentially only two pairs of time-frequency coefficients, $(\omega_1,[a,b])$ and $(\omega_2,[c,d])$.

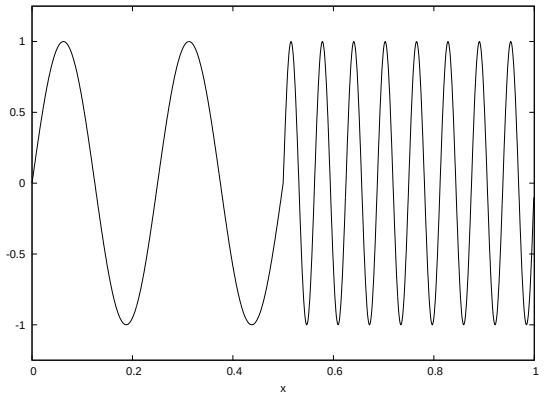

Figure 2.1: Two waves with different frequency.

2.2.1 The Haar Basis Wavelets System

Consider the space $L^2(\mathbb{R})=\{f : \int_{-\infty}^{+\infty}|f(x)|^2\,dx<\infty\}$. For simplicity, we can view this set as the functions $f(x)$ which get small in magnitude fast enough as x goes to plus and minus infinity. Let f be a function in $L^2(\mathbb{R})$. We can define piecewise constant approximations f_j of f at scale 2^{-j} by means of,

$$f_j(x)=2^j\int_{I_{j,k}} f(t)dt, \tag{2.1}$$

for all $x\in I_{j,k}, k\in\mathbb{Z}$,where $I_{j,k}=[\frac{k}{2^j},\frac{k+1}{2^j})$, i.e. f is approximated by its mean value on each interval $I_{j,k}$, $k\in\mathbb{Z}$. Let us make three simple comments on this particular choice for f_j:

1. We first remark that the choice of the mean value makes f_j the L^2-orthogonal projection of f onto the space,

$$V_j=\{f\in L^2 : f \textit{ is constant on } I_{j,k},\ k\in\mathbb{Z}\}.$$

Indeed, an orthogonal basis for V_j is given by the family

$$\phi_{j,k}(x)=2^{\frac{j}{2}}\chi_{I_{j,k}}(x)=2^{\frac{j}{2}}\phi(2^jx-k),\ k\in\mathbb{Z},$$

where $\phi(x) = \chi_{[0,1)}(x)$, and clearly f_j can be written as,

$$f_j = \sum_{k=-\infty}^{\infty} \langle f, \phi_{j,k} \rangle \phi_{j,k},$$

with the usual notation $\langle f, g \rangle = \int_{\mathbb{R}} f(t)\overline{g(t)}dt$. We will thus denote f_j by $\mathcal{P}_j f$ where $\mathcal{P}_j$ is the orthogonal projector onto V_j, i.e.,

$$\mathcal{P}_j : L^2(\mathbb{R}) \to V_j,$$

and we will also use the notation,

$$c_{j,k} = \langle f, \phi_{j,k} \rangle = 2^{\frac{j}{2}} \int_{I_{j,k}} f(t)dt, \tag{2.2}$$

for the normalized mean values which are the coordinates of $\mathcal{P}_j f$ in the basis $(\phi_{j,k})_{k \in \mathbb{Z}}$.

2. We also note that this approximation process is *local*: the value of $\mathcal{P}_j f$ in $I_{j,k}$ is only influenced by the value of f in the same interval. In particular, we can still use (2.1) to define $\mathcal{P}_j f$ when f is only locally integrable, or when f is only defined in a bounded interval such as $[0, 1]$ (in this case $\mathcal{P}_j f$ makes sense only for $j \geq 0$). As an example, we display in Figure 2.2 the function $f(x) = \sin(2\pi x)$ and its approximation $\mathcal{P}_4 f$ on $[0, 1]$.

3. Finally, since $V_j \subset V_{j+1}$, it is clear that $P_{j+1} f$ contains "more information" on f than the coarser approximation $\mathcal{P}_j f$, then the mean values on the intervals of size $\frac{1}{2^{j+1}}$ entirely determine those on the coarser intervals of double size $\frac{1}{2^j}$ just by taking their averages. More precisely, we have,

$$\mathcal{P}_j f_{|I_{j,k}} = \frac{\mathcal{P}_{j+1} f_{|I_{j+1,2k}} + \mathcal{P}_{j+1} f_{|I_{j+1,2k+1}}}{2}. \tag{2.3}$$

Let us consider the orthogonal complement W_j of V_j into V_{j+1}, that is $V_{j+1} = V_j \oplus W_j$. Then, we can also define the orthogonal projection $\mathcal{Q}_j f = \mathcal{P}_{j+1} f - \mathcal{P}_j f$ onto W_j. From (2.3), it is clear that $\mathcal{Q}_j f$ "oscillates" in the sense that,

$$\mathcal{Q}_j f_{|I_{j+1,2k}} = -\mathcal{Q}_j f_{|I_{j+1,2k+1}}. \tag{2.4}$$

As an example, Figure 2.3 shows the components $\mathcal{P}_3 f$ and $\mathcal{Q}_3 f$ on $[0, 1]$, for $f(x) = \sin(2\pi x)$. We have thus decomposed $\mathcal{P}_4 f$ of Figure 1.2.1 into the sum of the coarser approximation $\mathcal{P}_3 f$ and the oscillating component $\mathcal{Q}_3 f$.

The oscillatory property (2.4) allows us to expand $\mathcal{Q}_j f$ into,

$$\mathcal{Q}_j f = \sum_{k=-\infty}^{\infty} d_{j,k} \psi_{j,k},$$

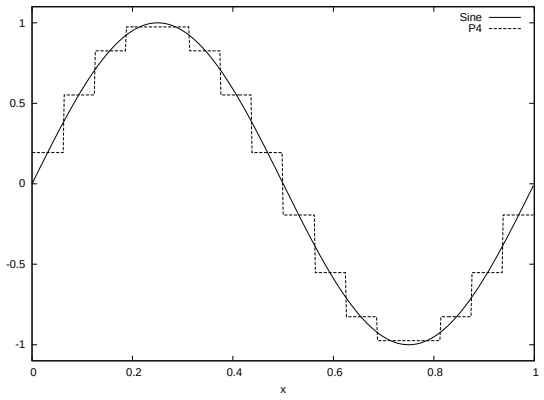

Figure 2.2: The function $f(x) = \sin(2\pi x)$ and its approximation $\mathcal{P}_4 f$.

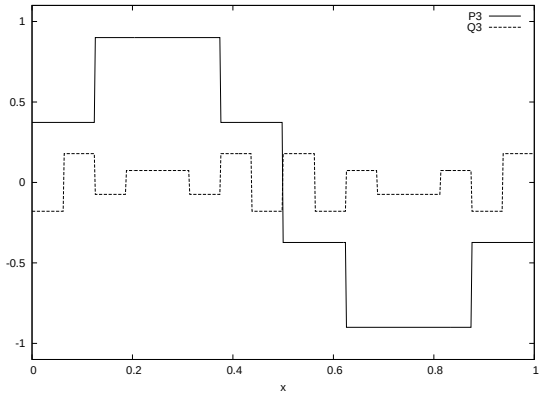

Figure 2.3: The approximation $\mathcal{P}_3 f$ and the fluctuation $\mathcal{Q}_3 f$ for $f(x) = \sin(2\pi x)$.

where $\psi_{j,k}(x) = 2^{\frac{j}{2}}\psi(2^j x - k)$ and $\psi(x) = \chi_{[0,\frac{1}{2})} - \chi_{[\frac{1}{2},1)}$.

Since the $\psi_{j,k},\ k \in \mathbb{Z}$ are also an orthonormal system, they constitute an orthonormal basis for W_j and we thus have,

$$d_{j,k} = \langle f, \psi_{j,k} \rangle.$$

We have thus re-expressed the "two-level" decomposition of $\mathcal{P}_{j+1} f$ into the coarser approximation $\mathcal{P}_j f$ and the additional fluctuations $\mathcal{Q}_j f$, according to,

$$\sum_{k=-\infty}^{\infty} c_{j+1,k}\phi_{j+1,k} = \sum_{k=-\infty}^{\infty} c_{j,k}\phi_{j,k} + \sum_{k=-\infty}^{\infty} d_{j,k}\psi_{j,k}. \tag{2.5}$$

This decomposition can be iterated for any arbitrary number of levels: if $j_0 < j_1$, we can rewrite the orthogonal decomposition,

$$\mathcal{P}_{j_1} f = \mathcal{P}_{j_0} f + \sum_{j=j_0}^{j_1-1} \mathcal{Q}_j f, \tag{2.6}$$

according to,

$$\sum_{k=-\infty}^{\infty} c_{j_1,k}\phi_{j_1,k} = \sum_{k=-\infty}^{\infty} c_{j_0,k}\phi_{j_0,k} + \sum_{j=j_0}^{j_1-1}\sum_{k=-\infty}^{\infty} d_{j,k}\psi_{j,k}. \tag{2.7}$$

Both (2.6) and (2.7) express and additive *multi-scale decomposition* of $\mathcal{P}_{j_1} f$ into a coarser approximation and a sequence of fluctuations at intermediate scales. Formula (2.7), however gives a local description of each contribution and should be viewed as an orthonormal change of basis in V_{j_1}: both $\{\phi_{j_1,k}\}_{k\in\mathbb{Z}}$ and $\{\phi_{j_0,k}\}_{k\in\mathbb{Z}} \cup \{\psi_{j,k}\}_{j_0\le j<j_1,k\in\mathbb{Z}}$ are orthonormal bases for V_{j_1}, and any function in V_{j_1} has thus a unique decomposition in each of these bases.

Note the different roles played by the functions ϕ and ψ: the former is used to characterize the approximation of a function at different scales, while the later is needed to represent the fluctuation between successive levels of approximation. In particular, we have $\int \psi = 0$, reflecting the oscillatory nature of these fluctuations. In the more general multi-resolution context, ϕ is called the *scaling* function and ψ the *mother* wavelet, in the sense that all the wavelets $\psi_{j,k}$ are generated form translations and dilations of ψ. We display in Figure 2.4 the functions $\phi_{2,3}$ and $\psi_{2,3}$.

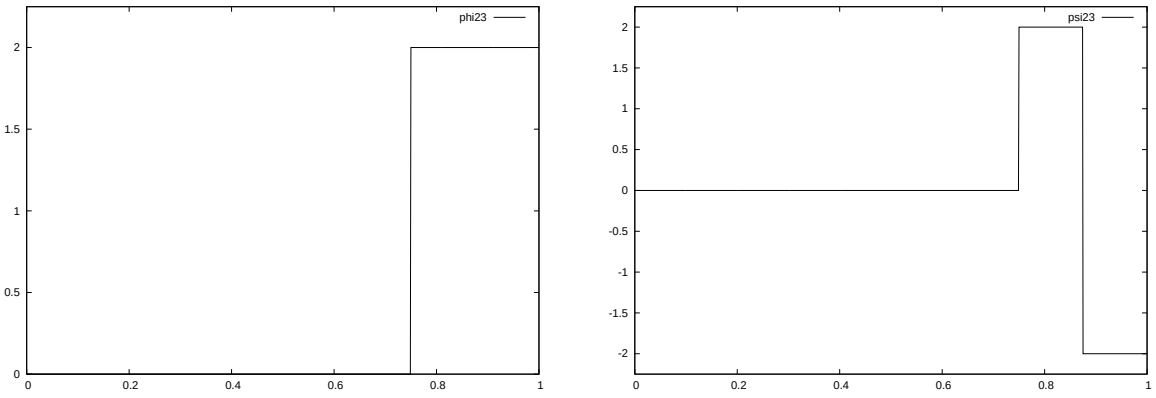

Figure 2.4: Scaling ($\phi_{2,3}$) and wavelet ($\psi_{2,3}$) functions.

Example 2.2.2. *Define the step function,*

$$f(x) = \begin{cases} -1, & x \in [-\frac{1}{2}, 0], \\ 1, & x \in (0, \frac{1}{2}], \\ 0, & \textit{otherwise.} \end{cases}$$

This function is poorly approximated by its Fourier series. The Fourier expansion for $f(x)$ *has the form,*

$$f(x) = \sum_{\substack{k=1 \\ k \text{ odd}}}^{\infty} \frac{4}{\pi k} \sin(2\pi k x) = \sum_{\substack{k=1 \\ k \text{ odd}}}^{\infty} b_k \varphi_k(x), \tag{2.8}$$

where $\varphi_k(x) = \sqrt{2}\sin(2\pi k x)$ *and* $b_k = \frac{2\sqrt{2}}{\pi k}$*. The left plot of Figure 2.5 shows this function together with the approximated Fourier series with* 5 *terms. The Fourier coefficients* b_k *decrease as* $\mathcal{O}(k^{-1})$ *which is a slow rate and one needs many terms of the Fourier expansion to approximate* f *with a good accuracy. The right plot of Figure 2.5 shows the step function* $f(x)$ *with the Fourier expansion using* 50 *terms in (2.8). If we include* 500 *terms in this Fourier expansion it would not look drastically different from what we already see in Figure 2.5. As it is well known, the Fourier approximation tends to keep the undesirable oscillations near the jump point and at the endpoints of the interval.*

Wavelets are more flexible in this situation. In fact, wavelet systems localize the jump putting small and extremely oscillating wavelets around it. This involves only one (or a small number) of coefficients, in contrast to the Fourier case. One of such wavelet system is given by the Haar basis with mother wavelet,

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

scaling function,

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

and their scaled and translated versions $\psi_{j,k} = 2^{\frac{j}{2}}\psi(2^j x - k)$ *and* $\phi_{j,k} = 2^{\frac{j}{2}}\phi(2^j x - k)$*. It is clear that with such a basis the step function in Figure 2.5 can be perfectly represented by two terms or coefficients,* $f(x) = \frac{\sqrt{2}}{2}\phi_{1,0}(x) - \frac{\sqrt{2}}{2}\phi_{1,-1}(x)$*, whereas using a Fourier series with* 50 *terms still produces wiggles in the reconstruction.*

Clearly, the union of the approximation spaces V_j is dense in $L^2(\mathbb{R})$, i.e.,

$$\lim_{j \to +\infty} \|f - \mathcal{P}_j f\|_{L^2} = 0, \tag{2.9}$$

for all $f \in L^2(\mathbb{R})$. Combining (2.9) with (2.7), we obtain that the orthonormal family $\{\phi_{j_0,k}\}_{k\in\mathbb{Z}} \cup \{\psi_{j,k}\}_{j\ge j_0, k\in\mathbb{Z}}$ is a complete orthonormal system of $L^2(\mathbb{R})$. Then, any function $f \in L^2(\mathbb{R})$ can be decomposed into,

$$f(x) = \sum_{k=-\infty}^{\infty} c_{j_0,k}\phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty}\sum_{k=-\infty}^{\infty} d_{j,k}\phi_{j,k}(x), \tag{2.10}$$

where the series converges in $L^2(\mathbb{R})$.

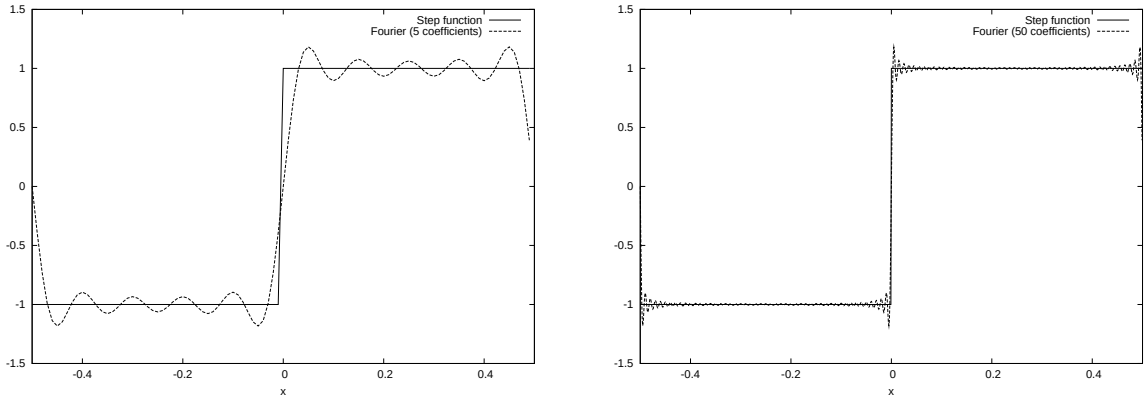

Figure 2.5: The step function and the Fourier series with 5 terms (left) and 50 terms (right).

Remark 2.2.1. *Of course we can apply the same method to decompose a function that is defined on a bounded interval. In particular, ϕ and the $\psi_{j,k}$'s for $j \geq 0$ and $0 \leq k < 2^j$ constitute an orthonormal basis for $L^2([0,1])$. This basis was introduced in* 1909 *by A. Haar, and is therefore referred as the Haar system.*

If the function f is defined on the whole $\mathbb{R}$, we can also let the coarsest level j_0 go to $-\infty$ in (2.10). It turns out that the coarse scale approximation vanishes in the L^2 sense, when passing to this limit, i.e., for any $f \in L^2(\mathbb{R})$, we have,

$$\lim_{j \to -\infty} \|\mathcal{P}_j f\|_{L^2} = 0. \tag{2.11}$$

In order to prove (2.11) we first suppose that $f = \chi_{[a,b]}$, for some $a < b$. Then, when j is negative and large enough so that $2^{-j} \geq \max\{|a|, |b|\}$, we have $c_{j,k} = 0$ if k is not 0 or -1 (the functions f and $\phi_{j,k}$ have disjoint supports). Thus, for this range of j, we have,

$$\|\mathcal{P}_j f\|_{L^2}^2 = |c_{j,0}|^2 + |c_{j,1}|^2 \leq 2\left(|a-b|2^{\frac{j}{2}}\right)^2,$$

which implies (2.11) for $f = \chi_{[a,b]}$ and thus for any finite linear combination of such a characteristic functions.

For an arbitrary $f \in L^2(\mathbb{R})$ and $\epsilon > 0$, there exists a function g which is of the type

$$g = \sum_{i=0}^{N} c_i \chi_{[a_i,b_i]},$$

such that,

$$\|f - g\|_{L^2} \leq \epsilon,$$

(finite linear combinations of characteristic functions of finite intervals are dense in $L^2(\mathbb{R})$). Since $\|\mathcal{P}_j g\|_{L^2}$ tends to zero as j goes to $-\infty$, we have $\|\mathcal{P}_j g\|_{L^2} \leq \epsilon$ for j large enough negatively, and therefore,

$$\|\mathcal{P}_j f\|_{L^2} \leq \|\mathcal{P}_j g\|_{L^2} + \|\mathcal{P}_j(f-g)\|_{L^2} \leq \|\mathcal{P}_j g\|_{L^2} + \|f-g\|_{L^2} \leq 2\epsilon.$$

Since ϵ is arbitrary, this proves (2.11) for a general function $f \in L^2(\mathbb{R})$.

An immediate consequence of (2.11) is that *the set* $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ *constitutes an orthonormal basis for* $L^2(\mathbb{R})$. This fact might appear as a paradox: a function f can be expanded in terms of basis functions that all have vanishing integral, even if $\int f \neq 0$. This is possible, simply because the convergence of $\mathcal{P}_j f$ to zero as j goes to $-\infty$ holds in L^2 *but not in* L^1, for a general function $f \in L^1 \cap L^2$. Indeed, according to the definition of $\mathcal{P}_j$ using (2.1), we clearly have $\int \mathcal{P}_j f = \int f$, so that $\mathcal{P}_j f$ cannot go to zero in L^1 as soon as f has a non-zero integral.

In the context of applications of multi-scale decompositions in numerical computations, one does not make so much use of the full wavelet basis $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ and of the fact that $\mathcal{P}_j f$ goes to zero in L^2 as j tends to $-\infty$. Indeed, most problems are confined to a bounded domain that limits the coarseness of the analysis scale. In our context, we will consider multi-scale decompositions of the type (2.10) for a fixed minimal scale level j_0, which quite often will be chosen to be $j_0 = 0$. More important for the numerician is the behavior of $\mathcal{P}_j f$ as j grows, in particular concerning approximation results dealing with the size of error $f - \mathcal{P}_j f$.

From a computational point of view, if we want to implement the decomposition of a function in the Haar system, we also need to limit the resolution level from above: we shall start out analysis from an approximation $f_{j_1} = \sum_{k=-\infty}^{\infty} c_{j_1,k}\phi_{j_1,k}$ of a function f at the finest resolution level j_1. We then have to deal with the following problem: how can we compute the coefficients $c_{j_0,k}$ and $d_{j,k}, j_0 \leq j < j_1$ from the coefficients $c_{j_1,k}$ at the finest scale?

Remark 2.2.2. *Note that this problem makes sense even if* f_{j_1} *is an approximation of* f *in* V_{j_1} *that differs form its projection* $\mathcal{P}_{j_1} f$*: we are then interested in decomposing this particular approximation into multi-scale components. This point is crucial, since in numerical analysis applications, the approximate solution of a problem is rarely the* L^2*-orthogonal projection of the true solution.*

To answer this question, we start from the "two level" decomposition (2.5). From (2.3) and the L^2 normalization of the functions $\phi_{j,k}$ and $\psi_{j,k}$, we derive the simple relations,

$$c_{j,k} = \frac{c_{j+1,2k} + c_{j+1,2k+1}}{\sqrt{2}}, \tag{2.12}$$

and,

$$d_{j,k} = \frac{c_{j+1,2k} - c_{j+1,2k+1}}{\sqrt{2}}. \tag{2.13}$$

Another way to derive relations (2.12) and (2.13) is to insert the identities $\phi_{j,k} = \frac{\phi_{j+1,2k}+\phi_{j+1,2k+1}}{\sqrt{2}}$ and $\psi_{j,k} = \frac{\psi_{j+1,2k}-\psi_{j+1,2k+1}}{\sqrt{2}}$ inside $c_{j,k} = \langle f, \phi_{j,k}\rangle$ and $d_{j,k} = \langle f, \psi_{j,k}\rangle$. These relations show the local aspect of the change of basis in (2.5), the vectors $(c_{j,k}, d_{j,k})$ and $(c_{j+1,2k}, d_{j+1,2k+1})$ are related by a simple orthogonal transformation. The inverse transform is given by,

$$c_{j+1,2k} = \frac{c_{j,k} + d_{j,k}}{\sqrt{2}}, \tag{2.14}$$

and,

$$c_{j+1,2k+1} = \frac{c_{j,k} - d_{j,k}}{\sqrt{2}}. \tag{2.15}$$

From these four relations, we can derive simple decomposition and reconstruction algorithms.

Decomposition algorithm:

1. Start from finest scale approximation coefficients $c_{j_1,k}$.
2. Compute the sequences $c_{j_1-1,k}$ and $d_{j_1-1,k}$ using (2.12) and (2.13).
3. Store the details $d_{j_1-1,k}$ and iterate the decomposition for the approximation coefficients $c_{j_1-1,k}$.
4. Iterate the decomposition from fine to coarse scales.
5. Stop the decomposition when the coarsest level coefficients $c_{j_0,k}$ and $d_{j_0,k}$ are reached.

Reconstruction algorithm:

1. Start from the coarsest level coefficients $c_{j_0,k}$ and $d_{j_0,k}$.
2. Compute the sequence $c_{j_0+1,k}$ using (2.14) and (2.15).
3. Iterate the reconstruction from coarse to fine scales.
4. Stop the reconstruction when the finest approximation coefficients $c_{j_1,k}$ are computed.

In practice, the function to be decomposed is defined (and approximated at the finest level) on a bounded domain. In turn, the sequence $c_{j,k}$ has finite length, allowing a computer to perform these algorithms. For example, if the function f is defined on the interval $[0, 1]$, these algorithms operate on 2^{j_1} coefficients.

The total number of operations for both algorithms is therefore given by $\sum_{j=j_0}^{j_1} 2^{j+1} = \mathcal{O}(N)$ where $N = 2^{j_1}$ is the length of the finest approximation sequence. This complexity is "optimal", in the sense that the number of required operations is of the same order as the number of computed values. A general linear transformation would require $\mathcal{O}(N^2)$ operations. From a linear algebra point of view, the Haar decomposition and reconstruction algorithm can thus be viewed as a smart factorization of a linear transform. A similar situation is encountered when implementing the discrete Fourier transform by the Fast Fourier Transform algorithm which has complexity $\mathcal{O}(N \log N)$.

2.2.2 Multiresolution Analysis

In a more general setting, we can define a structure for wavelets in $L^2(\mathbb{R})$ which is called a *Multi-Resolution Analysis* (MRA).

We start with a family of closed nested subspaces,

$$... \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset ...,$$

in $L^2(\mathbb{R})$ where,

$$\bigcap_{j\in\mathbb{Z}} V_j = \{0\}, \qquad \overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R}),$$

and,

$$f(x) \in V_j \Longleftrightarrow f(2x) \in V_{j+1}.$$

If these conditions are met, then there exists a function $\phi \in V_0$ such that $\{\phi_{j,k}\}_{k\in\mathbb{Z}}$ is an orthonormal basis of V_j, where,

$$\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k).$$

In other words, the function ϕ, called the *scaling function*, will generate an orthonormal basis for each V_j subspace.

Then we define W_j such that $V_{j+1} = V_j \oplus W_j$ as we have already seen. This says that W_j is the space of functions in V_{j+1} but not in V_j, and so, $L^2(\mathbb{R}) = \sum_j \oplus W_j$. Then (see [Dau92]) there exists a function $\psi \in W_0$ such that $\{\psi_{j,k}\}_{k\in\mathbb{Z}}$ is an orthonormal basis of W_j, and $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ is a wavelets basis of $L^2(\mathbb{R})$, where,

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k).$$

The function ψ is called the *mother function*, and the $\psi_{j,k}$ are the *wavelets functions*.

For any function $f \in L^2(\mathbb{R})$ a projection map of $L^2(\mathbb{R})$ onto V_m,

$$\mathcal{P}_m : L^2(\mathbb{R}) \to V_m,$$

is defined by,

$$\mathcal{P}_m f(x) = \sum_{j=-\infty}^{m-1} \sum_{k=-\infty}^{k=+\infty} d_{j,k}\psi_{j,k}(x) = \sum_{k\in\mathbb{Z}} c_{m,k}\phi_{m,k}(x). \tag{2.16}$$

where,

$$d_{j,k} = \int_{-\infty}^{+\infty} f(x)\psi_{j,k}(x)dx,$$

are the wavelets coefficients and,

$$c_{m,k} = \int_{-\infty}^{+\infty} f(x)\phi_{m,k}(x)dx,$$

are the scaling coefficients. The first part in (2.16) is a truncated wavelets series. If j were allowed to go to infinity, we would have the full wavelets summation. The second

part in (2.16) gives an equivalent sum in terms of the scaling functions $\phi_{m,k}$. Considering higher m values, meaning that more terms are used, the truncated series representation of our function improves.

To develop our work, we have used the Haar system defined in the previous section. The unique thing about using wavelets as opposed to Fourier series is that the wavelets can be moved (by the k value), stretched or compressed (by the j value) to accurately represent a function local properties. Moreover, $\phi_{j,k}$ is nonzero only inside the interval $[\frac{k}{2^j}, \frac{k+1}{2^j})$. These facts will be used later to compute the VaR without the need of knowing the whole distribution of the loss function.

2.2.3 Behavior of Haar Coefficients Near Jump Discontinuities

Suppose that $f(x)$ is a function defined on $[0,1]$, with a jump discontinuity at $x_0 \in (0,1)$ and continuous everywhere else in $[0,1]$. The fact that the Haar functions $\psi_{j,k}(x)$ have good localization in time leads us to ask the question: do the Haar coefficients $d_{j,k}$ such that $x_0 \in I_{j,k}$ behave in a different way than the Haar coefficients such that $x_0 \notin I_{j,k}$? In particular, can we find the location of a jump discontinuity just by examining the absolute value of the Haar coefficients? We will give an affirmative answer and see that these two facts are closely related.

For simplicity, let us assume that the given function $f(x)$ is C^2 in the intervals $[0, x_0]$ and $[x_0, 1]$. This means that both $f'(x)$ and $f''(x)$ exist, are continuous functions, and hence are bounded on each of these intervals. We fix integers $j \geq 0$ and $0 \leq k \leq 2^j - 1$, and let $x_{j,k}$ be the midpoint of the interval $I_{j,k}$, that is, $x_{j,k} = \frac{k+1/2}{2^j}$. There are now two possibilities, either $x_0 \notin I_{j,k}$ or $x_0 \in I_{j,k}$.

Case 1: $x_0 \notin I_{j,k}$. If $x_0 \notin I_{j,k}$, then expanding $f(x)$ about $x_{j,k}$ by means of Taylor's formula, it follows that for all $x \in I_{j,k}$,

$$f(x) = f(x_{j,k}) + f'(x_{j,k})(x - x_{j,k}) + \frac{1}{2} f''(\xi_{j,k})(x - x_{j,k})^2,$$

where $\xi_{j,k}$ is some point in $I_{j,k}$. Now, using the fact that $\int_{I_{j,k}} \psi_{j,k}(x)dx = 0$,

$$\begin{aligned} d_{j,k} &= \int_{I_{j,k}} f(x)\psi_{j,k}(x)dx \\ &= f(x_{j,k}) \int_{I_{j,k}} \psi_{j,k}(x)dx + f'(x_{j,k}) \int_{I_{j,k}} \psi_{j,k}(x)(x - x_{j,k})dx + \\ &+ \frac{1}{2} \int_{I_{j,k}} \psi_{j,k}(x)(x - x_{j,k})^2 f''(\xi_{j,k})dx \\ &= -\frac{1}{4} f'(x_{j,k}) 2^{-3j/2} + r_{j,k}(x), \end{aligned}$$

where,

$$\begin{aligned}|r_{j,k}(x)| &= \frac{1}{2}\left|\int_{I_{j,k}} \psi_{j,k}(x)(x-x_{j,k})^2 f''(\xi_{j,k})dx\right| \\ &\leq \frac{1}{2}2^{j/2}\max_{x\in I_{j,k}}|f''(x)|\int_{I_{j,k}}(x-x_{j,k})^2dx \\ &= \frac{1}{24}\max_{x\in I_{j,k}}|f''(x)|\,2^{-5j/2}.\end{aligned}$$

If j is large, then $2^{-5j/2}$ will be very small compared with $2^{-3j/2}$, so we conclude that for large j,

$$|d_{j,k}| \simeq \frac{1}{4}|f'(x_{j,k})|\,2^{-3j/2}. \tag{2.17}$$

Case 2: $x_0 \in I_{j,k}$. If $x_0 \in I_{j,k}$, then either it is in $I^l_{j,k} = [\frac{k}{2^j}, \frac{k+1/2}{2^j})$ or it is in $I^r_{j,k} = [\frac{k+1/2}{2^j}, \frac{k+1}{2^j})$. Let us assume that $x_0 \in I^l_{j,k}$ (the other case is similar). Expanding $f(x)$ in a Taylor series about x_0, we have,

$$\begin{aligned}f(x) &= f(x_0^-) + f'(\xi^-)(x-x_0),\ x\in[0,x_0),\ \xi^-\in[x,x_0], \\ f(x) &= f(x_0^+) + f'(\xi^+)(x-x_0),\ x\in(x_0,1],\ \xi^+\in[x_0,x],\end{aligned}$$

Therefore,

$$\begin{aligned}d_{j,k} &= \int_{I_{j,k}} f(x)\psi_{j,k}(x)dx \\ &= 2^{j/2}\int_{2^{-j}k}^{x_0} f(x_0^-)dx + 2^{j/2}\int_{x_0}^{2^{-j}(k+1/2)} f(x_0^+)dx - \\ &- 2^{j/2}\int_{2^{-j}(k+1/2)}^{2^{-j}(k+1)} f(x_0^+)dx + \epsilon_{j,k},\end{aligned}$$

where,

$$\epsilon_{j,k} = \int_{2^{-j}k}^{x_0} f'(\xi^-)(x-x_0)\psi_{j,k}(x)dx + \int_{x_0}^{2^{-j}(k+1)} f'(\xi^+)(x-x_0)\psi_{j,k}(x)dx.$$

Thus,

$$\begin{aligned}|\epsilon_{j,k}| &\leq \max_{t\in I_{j,k}-\{x_0\}}|f'(t)|\int_{I_{j,k}}|x-x_0|\,|\psi_{j,k}(x)|\,dx \\ &\leq \max_{t\in I_{j,k}-\{x_0\}}|f'(t)|\,2^{j/2}\int_{I_{j,k}}|x-x_0|\,dx \\ &\leq \frac{1}{4}\max_{t\in I_{j,k}-\{x_0\}}|f'(t)|\,2^{-3j/2}.\end{aligned}$$

If j is large, then $2^{-3j/2}$ will be very small compared with $2^{-j/2}$, so we conclude that for large j,

$$|d_{j,k}| \simeq 2^{j/2} \left|x_0 - 2^{-j}k\right| \left|f(x_0^-) - f(x_0^+)\right|.$$

The quantity $|x_0 - 2^{-j}k|$ can in principle be small if x_0 is close to the left endpoint of $I_{j,k}^l$ and can even be zero. However, we can expect than in most cases, x_0 will be in the middle of $I_{j,k}^l$ so that $|x_0 - 2^{-j}k| \simeq \frac{1}{4}2^{-j}$. Thus, for large j,

$$|d_{j,k}| \simeq \frac{1}{4} \left|f(x_0^-) - f(x_0^+)\right| 2^{-j/2}. \tag{2.18}$$

Comparing (2.18) with (2.17), we see that the decay of $|d_{j,k}|$ for large j is considerably slower if $x_0 \in I_{j,k}$ than if $x_0 \notin I_{j,k}$. That is, large coefficients in the Haar expansion of a function $f(x)$ that persist for all scales suggest the presence of a jump discontinuity in the intervals $I_{j,k}$ corresponding to the large coefficients.

2.3 Numerical Transform Inversion

2.3.1 The Laplace Transform

Suppose that f is a real- or complex-valued function of the variable $x > 0$ and s is a real or complex parameter. We define the *Laplace transform* of f as,

$$\widetilde{f}(s) = \int_0^{+\infty} e^{-sx} f(x)dx = \lim_{\tau \to +\infty} \int_0^{\tau} e^{-sx} f(x)dx, \tag{2.19}$$

whenever the limit exists (as a finite number).

The parameter s belongs to some domain on the real line or in the complex plane. We will choose s appropriately so as to ensure the convergence of the Laplace integral (2.19). In a mathematical and technical sense, the domain of s is quite important. However, in a practical sense, the domain of s is routinely ignored.

Although the Laplace transform exists for many functions, there are some for which the integral (2.19) does not converge.

Example 2.3.1. *For the function* $f(x) = e^{x^2}$,

$$\lim_{\tau \to +\infty} \int_0^{\tau} e^{-sx} e^{x^2} dx = \lim_{\tau \to +\infty} \int_0^{\tau} e^{x^2 - sx} dx = +\infty,$$

for any choice of the variable s, *since the integrand grows without bound as* $\tau \to +\infty$.

The integral (2.19) is said to be *absolutely convergent* if,

$$\lim_{\tau \to +\infty} \int_0^{\tau} \left|e^{-sx} f(x)\right| dx,$$

exists. If $\widetilde{f}(s)$ converges absolutely, then,

$$\left| \int_{\tau}^{\tau'} e^{-sx} f(x) dx \right| \leq \int_{\tau}^{\tau'} \left| e^{-sx} f(x) \right| dx \to 0,$$

as $\tau \to +\infty$, for all $\tau' > \tau$. This then implies that $\widetilde{f}(s)$ also converges in the ordinary sense of (2.19).

Since we can compute the Laplace transform for some functions and not for others, such as e^{x^2}, we introduce a large class of functions that have a Laplace transform. For this purpose we make few restrictions on the functions we wish to consider.

Definition 2.3.1. *A function f is* piecewise continuous *on the interval $[0, +\infty)$ if,*

1. $\lim_{x \to 0^+} f(x) = f(0^+)$ *exists.*
2. *f is continuous on every finite interval $(0, b)$ except possibly at a finite number of points $\tau_1, \tau_2, \ldots, \tau_n$ in $(0, b)$ at which f has a jump discontinuity.*

The second consideration of our class of functions possessing a well-defined Laplace transform has to do with the growth rate of the functions. In the definition,

$$\widetilde{f}(s) = \int_0^{+\infty} e^{-sx} f(x) dx,$$

when we take $s > 0$ (or $\Re(s) > 0$), the integral will converge as long as f does not grow too fast. We have already seen by Example 2.3.1 that $f(x) = e^{x^2}$ grows too rapidly for our purposes. A suitable rate of growth can be made explicit.

Definition 2.3.2. *A function f has* exponential order α *if there exist constants $M > 0$ and α such that for some $x_0 \geq 0$,*

$$|f(x)| \leq M e^{\alpha x}, \quad x \geq x_0.$$

This large class of functions possesses a Laplace transform (see [Sch88] for a detailed proof of next theorem).

Theorem 2.3.1. *If f is piecewise continuous on $[0, +\infty)$ and of exponential order α, then the Laplace transform $\widetilde{f}(s)$ exists for $\Re(s) > \alpha$ and converges absolutely.*

We are now ready to derive and use the formula for the inverse Laplace Transform. We state the inverse transform as a theorem (see [Dyk01] for a detailed proof).

Theorem 2.3.2. *(Bromwich inversion integral) If the Laplace transform of $f(x)$ exists, then,*

$$f(x) = \lim_{k \to +\infty} \frac{1}{2\pi i} \int_{\sigma - ik}^{\sigma + ik} \widetilde{f}(s) e^{sx} ds, \quad x > 0, \tag{2.20}$$

where $|f(x)| \leq e^{\Sigma x}$ for some positive real number Σ and σ is any other real number such that $\sigma > \Sigma$.

The usual way of evaluating this integral is via the residues method taking a closed contour, often called the Bromwich contour. This contour consists of a portion of a circle of radius R, together with a straight line segment connecting the two points $\sigma - iR$ and $\sigma + iR$. The real number σ must be selected so that all the singularities of the function $\widetilde{f}(s)$ are to the left of this line. This follows from the conditions of Theorem 2.3.2. The integral $\int_C \widetilde{f}(s)e^{sx}ds$, where C is the Bromwich contour is evaluated using Cauchy's residue theorem, perhaps with the addition of one or two cuts. The integral itself is the sum of the integral over the curved portion, the integral along any cuts presents and $\int_{\sigma-iR}^{\sigma+iR} \widetilde{f}(s)e^{sx}$, and the whole is $2\pi i$ times the sum of the residues of $\widetilde{f}(s)e^{sx}$ inside C. The previous integral is made the subject of this formula, and as $R \to +\infty$ this integral becomes $f(x)$.

2.3.2 Laplace Transform Inversion

In this section we present some numerical algorithms in order to invert the Laplace transform. A natural starting point for the numerical inversion of Laplace transforms is the Bromwich inversion integral stated in Theorem 2.3.2. As it is usual in contour integrals, there is plenty of flexibility in choosing the contour provided that it is to the right of all singularities of $\widetilde{f}$. However, there is no need to be bothered by the complexities of complex variables. If we choose a specific contour and perform a change of variables, then we obtain an integral of a real valued function of a real variable. First, by making the substitution $s = \sigma + iu$ in (2.20), we obtain,

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{(\sigma+iu)x} \widetilde{f}(\sigma + iu)du.$$

Then, since $e^{(\sigma+iu)x} = e^{\sigma x}(\cos(ux)+i\sin(ux))$, $\sin(ux) = -\sin(-ux)$, $\cos(ux) = \cos(-ux)$, $\Im\left(\widetilde{f}(\sigma + iu)\right) = -\Im\left(\widetilde{f}(\sigma - iu)\right)$ and $\Re\left(\widetilde{f}(\sigma + iu)\right) = \Re\left(\widetilde{f}(\sigma - iu)\right)$, where $\Im(z)$ and $\Re(z)$ denote the imaginary and the real part of z respectively, we obtain,

$$f(x) = \frac{2e^{\sigma x}}{\pi} \int_0^{+\infty} \Re\left(\widetilde{f}(\sigma + iu)\right) \cos(ux)du, \tag{2.21}$$

and,

$$f(x) = \frac{-2e^{\sigma x}}{\pi} \int_0^{+\infty} \Im\left(\widetilde{f}(\sigma + iu)\right) \sin(ux)du.$$

Theorem 2.3.2 implies that $f(x)$ can be calculated from the transform $\widetilde{f}$ by performing a numerical integration quadrature. Since there are many numerical integration algorithms, there are many possible approaches to the numerical inversion via the Bromwich integral. In this context, the remaining goal is to exploit the special structure of the integrand in (2.21) in order to calculate the integral accurately and efficiently.

An approach that proves to be remarkably effective in this context is the trapezoidal rule. If we use a step of size h, the trapezoidal rule applied to (2.21) gives,

$$f(x) \simeq f_h(x) = \frac{he^{\sigma x}}{\pi} \Re\left(\widetilde{f}(\sigma)\right) + \frac{2he^{\sigma x}}{\pi} \sum_{k=1}^{+\infty} \Re\left(\widetilde{f}(\sigma + ikh)\right) \cos(khx). \tag{2.22}$$

Now, in order to use the approximation (2.22), we have to be aware of the different types of errors we can have when using numerical integration. First, we have the discretization error which comes from approximating f with f_h. The second type of error comes from the evaluation of the infinite sum in (2.22).

The following theorem (see [Aba00] for details) characterizes the first type of error.

Theorem 2.3.3. *Under regularity conditions, the discretization error for the trapezoidal rule approximation in (2.22) is,*

$$e_h(x) \equiv f_h(x) - f(x) = \sum_{\substack{k=-\infty \\ k\neq 0}}^{+\infty} f\left(x + \frac{2\pi k}{h}\right) e^{-2\pi k\sigma/h}. \tag{2.23}$$

Rather than just truncate the infinite sum in (2.22), [Aba00] applies a summation acceleration method. They apply Euler summation, after transforming the infinite sum in (2.22) into a nearly alternating series.

In order to control the round-off error, they introduce a parameter l which is a positive integer that often can be given the default value $l = 1$. They convert (2.22) into a nearly alternating series by taking $h = \pi/lx$ and $\sigma = A/2lx$. After algebraic manipulation, we get,

$$f_{A,l}(x) = \sum_{k=0}^{+\infty} (-1)^k a_k(x), \tag{2.24}$$

where,

$$a_k(x) = \frac{e^{A/2l}}{2lx} b_k(x),\ k \geq 0, \tag{2.25}$$

$$b_0(x) = \widetilde{f}\left(\frac{A}{2lx}\right) + 2\sum_{j=1}^{l} \Re\left(\widetilde{f}\left(\frac{A}{2lx} + \frac{ij\pi}{lx}\right) e^{ij\pi/l}\right),$$

and,

$$b_k(x) = 2\sum_{j=1}^{l} \Re\left(\widetilde{f}\left(\frac{A}{2lx} + \frac{ij\pi}{lx} + \frac{ik\pi}{x}\right) e^{ij\pi/l}\right),\ k \geq 1. \tag{2.26}$$

Also the aliasing error in (2.23) becomes,

$$e_h(x) \equiv e_{A,l}(x) = \sum_{j=1}^{+\infty} e^{-Aj} f\left((1 + 2jl)x\right). \tag{2.27}$$

If $|f(t)| \leq C$ for all $t \geq (1 + 2l)x$ then the discretization error in (2.27) is bounded by,

$$|e_{A,l}(x)| \leq \frac{Ce^{-A}}{1 - e^{-A}},$$

However, controlling this discretization error is not enough and [Aba00] also controls for round-off error. If we can compute the $b_k(x)$ in (2.26) with an error of about $10^{-\eta}$, then,

after multiplying by the pre-factor $\frac{e^{A/2l}}{2lx}$ in (2.25), the final error in $f_{A,l}(x)$ will be about $10^{-\eta}\frac{e^{A/2l}}{2lx}$. Since this round-off error estimate is increasing with A, while the discretization error e^{-A} (assuming that $C = 1$) is decreasing with A, the maximum of the two errors is minimized when the two errors are equal. Then, [Aba00] finds an appropriate value for the parameter A by means of solving the equation,

$$e^{-A} = 10^{-\eta}\frac{e^{A/2l}}{2lx},$$

which yields,

$$A = \left(\frac{2l}{2l+1}\right)(\eta \ln(10) + \ln(2lx)).$$

Let us now indicate how to apply Euler summation to approximate the infinite series in (2.24). Euler summation can be simply described as the weighted average of the last partial m sums by a binomial probability distribution with parameters m and $p = 1/2$. In particular, let s_n be the approximation $f_{A,l}(x)$ in (2.24) with the infinite series truncated to n terms,

$$s_n(x) = \sum_{k=0}^{n}(-1)^k a_k(x).$$

The Euler sum approximation applied to m terms after the initial partial sum with n terms $E(m, n)$ is defined as,

$$E(m,n)(x) = \sum_{k=0}^{m}\left(\begin{array}{c} m \\ k \end{array}\right) 2^{-m} s_{n+k}(x). \tag{2.28}$$

In order to reduce computatinal time, we rewrite (2.28) in the following form,

$$\begin{aligned} E(m,n)(x) &= \sum_{k=0}^{m}\left(\begin{array}{c} m \\ k \end{array}\right) 2^{-m} s_{n+k}(x) \\ &= \sum_{k=0}^{m}\left(\begin{array}{c} m \\ k \end{array}\right) 2^{-m}\left(s_n(x) + (s_{n+k}(x) - s_n(x))\right) \\ &= \sum_{k=0}^{m}\left(\begin{array}{c} m \\ k \end{array}\right) 2^{-m} s_n(x) + \sum_{k=1}^{m}\left(\begin{array}{c} m \\ k \end{array}\right) 2^{-m}\left(s_{n+k}(x) - s_n(x)\right) \\ &= s_n(x) + \sum_{k=1}^{m}\left(\begin{array}{c} m \\ k \end{array}\right) 2^{-m} s_{n+1:n+k}(x), \end{aligned}$$

where,

$$s_{i:j}(x) = \sum_{k=i}^{j}(-1)^k a_k(x),\ i < j.$$

As pointed out in [Aba00], the Bromwich inversion integral is not the only inversion formula and there are quite different numerical inversion algorithms. To illustrate this fact, we mention two more. A first example is the Post-Wider inversion formula, which involves differentiation instead of integration:

Theorem 2.3.4. *Under regularity conditions,*

$$f(x) = \lim_{n\to+\infty} \frac{(-1)^n}{n!} \left(\frac{n+1}{x}\right)^{n+1} \widetilde{f}^{(n)}((n+1)/x),$$

where $\widetilde{f}^{(n)}$ *is the* n^{th} *derivative of* $\widetilde{f}$*.*

However in general, theorem (2.3.4) does not provide a very practical means for inverting the Laplace transform, since repeated differentiation of $\widetilde{f}$ generally leads to very complicated expressions. Another case is the Laguerre series representation given in [Aba96] for inverting the Laplace transform, which is known to be an efficient method for smooth functions. However, if f is not smooth at just one value of x, then the Laguerre method has difficulties at any value of x.

Theorem 2.3.5. *Under regularity conditions,*

$$f(x) = \sum_{n=0}^{+\infty} q_n l_n(x), \quad x \geq 0, \tag{2.29}$$

where,

$$l_n(x) = e^{-x/2} L_n(x), \quad x \geq 0,$$
$$L_n(x) = \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \frac{(-x)^k}{k!}, \quad x \geq 0,$$

and the q_n *are obtained from the expansion,*

$$Q(z) \equiv \sum_{n=0}^{+\infty} q_n z^n = \frac{1}{1-z} \widetilde{f}\left(\frac{1+z}{2(1-z)}\right).$$

The functions L_n in Theorem 2.3.5 are the *Laguerre polynomials*, while l_n are the associated *Laguerre functions*. The scalars q_n in (2.29) are known as the *Laguerre coefficients* and $Q(z)$ is the *Laguerre generating function* (the generating function of the Laguerre coefficients). The Laguerre functions form and orthonormal basis for the space $L^2([0,+\infty))$ of square integrable functions on the nonnegative real line, so that the equality (2.29) is understood (in the sense of convergence in $L^2([0,+\infty))$ for any f in $L^2([0,+\infty))$.

In the context of numerical Laplace inversion, [Gao03] recovers the function f with a procedure based on wavelets. They consider $s = \beta + i\omega$ in expression (2.20), where ω is a real variable and β is a real constant that fulfills $f(x)e^{-\beta x} \in L^2(\mathbb{R})$, assuming that $f(x) = 0$ when $x < 0$. Then, the equation (2.19) can be rewritten as,

$$\widetilde{f}(\beta + i\omega) = \int_{-\infty}^{+\infty} e^{-\beta x} e^{-i\omega x} f(x) dx.$$

Defining,

$$h(x) = f(x)e^{-\beta x}, \text{ then } \hat{h}(\omega) = \widetilde{f}(\beta + i\omega), \tag{2.30}$$

where $\hat{h}$ denotes the Fourier transform of h. Now, $\hat{h}(\omega) \in L^2(\mathbb{R})$ since $h(x) \in L^2(\mathbb{R})$. According to the theory of MRA they expand $\hat{h}(\omega)$,

$$\hat{h}(\omega) = \sum_{k=-\infty}^{+\infty} c_{m,k}\phi_{m,k}(\omega) + \sum_{j=m}^{+\infty}\sum_{k=-\infty}^{+\infty} d_{j,k}\psi_{j,k}(\omega). \tag{2.31}$$

where,

$$c_{m,k} = \int_{-\infty}^{+\infty} \hat{h}(\omega)\phi_{m,k}(\omega),\ d_{j,k} = \int_{-\infty}^{+\infty} \hat{h}(\omega)\psi_{j,k}(\omega),$$

and $\phi_{m,k}(\omega) = 2^{m/2}\phi(2^m\omega - k)$, $\psi_{j,k}(\omega) = 2^{j/2}\psi(2^j\omega - k)$, being ϕ and ψ the scaling and wavelet functions respectively.

The next step consists in inverting the expression (2.31) by means of the Fourier inversion formula, yielding,

$$\begin{aligned} h(x) &= \frac{1}{2\pi}\sum_{k=-\infty}^{+\infty} c_{m,k}\int_{-\infty}^{+\infty}\phi_{m,k}(\omega)e^{i\omega x}d\omega + \frac{1}{2\pi}\sum_{j=m}^{+\infty}\sum_{k=-\infty}^{+\infty} d_{j,k}\int_{-\infty}^{+\infty}\psi_{j,k}(\omega)e^{i\omega x}d\omega \\ &= \frac{1}{2\pi 2^{m/2}}\hat{\phi}\left(-\frac{x}{2^m}\right)\sum_{k=-\infty}^{+\infty} c_{m,k}e^{ixk/2^m} + \frac{1}{2\pi}\sum_{j=m}^{+\infty} 2^{-j/2}\hat{\psi}\left(-\frac{x}{2^j}\right)\sum_{k=-\infty}^{+\infty} d_{j,k}e^{ixk/2^j}. \end{aligned} \tag{2.32}$$

The coefficients $c_{m,k}$ and $d_{j,k}$ involve inner products with wavelet basis and are usually difficult to calculate using ordinary numerical methods. It is possible to design wavelet functions which allow to obtain these coefficients more easily. The authors choose Coiflets wavelets to perform the approximation in (2.31). By the characteristics of Coiflets it can be shown that,

$$c_{m,k} \simeq 2^{-m/2}\hat{h}\left(\frac{k+M_1}{2^m}\right), \quad \text{with } M_1 = \int_{-\infty}^{+\infty}\omega\phi(\omega)d\omega. \tag{2.33}$$

If $\hat{h}$ is a smooth function, the approximation above converge for significantly large scale m. Now, using the expression (2.33) in (2.32) and again taking into account the theory of MRA, the expression (2.32) can be rewritten as,

$$h(x) = \lim_{m\to+\infty} h_m(x),$$

where,

$$h_m(x) = \frac{1}{2^{m+1}\pi}\hat{\phi}\left(-\frac{x}{2^m}\right)\sum_{k=-\infty}^{+\infty}\hat{h}\left(\frac{k+M_1}{2^m}\right)e^{ixk/2^m}.$$

Considering the expression (2.30), the formulae of Laplace inversion become,

$$\begin{aligned} f_m(x) &= \frac{e^{\beta x}}{2^{m+1}\pi}\hat{\phi}\left(-\frac{x}{2^m}\right)\sum_{k=-\infty}^{+\infty}\widetilde{f}\left(\beta + i\frac{M_1+k}{2^m}\right)e^{ixk/2^m}, \\ f(x) &= \lim_{m\to+\infty} f_m(x). \end{aligned}$$

One drawback of this approximation is that the wavelet approach involves an infinite product of complex series and the computation of the Fourier transform of some scaling functions. This can look intimidating for practical applications and may also take relatively long computational time. To address this problem, [Gao05] develops a simplified approach in which the Fourier transform of the scaling function is approximated by a single harmonic base in a certain domain. This approach yields a simple formula for Laplace inversion via summation of a weighted harmonic series as,

$$f(x) \simeq f_m(x) = \frac{1}{2^{m+1}\pi} \sum_{k=-\infty}^{+\infty} \widetilde{f}\left(\beta + i\frac{k}{2^m}\right) e^{x(\beta+ik/2^m)},$$

where $x \in [0, 2^m\pi]$ and β is a real constant which ensures $f(x)e^{-\beta x} \in L^2(\mathbb{R})$.

Based on operational matrices and Haar wavelets, [Chen01] presents a new method for performing numerical inversion of the Laplace transform where only matrix multiplications and ordinary algebraic operations are involved. However, the essential step in the method consists in expressing the Laplace transform in terms of $\frac{1}{s}$, which is impossible when we just know numerically the transform. Another drawback of this method is that the matrices become very big for larger scales.

2.3.3 Generating Functions (z-Transforms)

As it is also well known, generating functions, also called z-transforms, are closely related to Laplace transforms in the sense that we briefly review. Assume that $f(t)$ is a continuous function and we sample it at time intervals of T. We obtain the data,

$$f(0), f(T), f(2T), \ldots, f(nT), \ldots$$

Assume also that the Dirac impulse function at $t = T$ is denoted by $\delta(t-T)$. If we denote by f^* the sampled function, we can write,

$$f^*(t) = f(0)\delta(t) + f(T)\delta(t-T) + f(2T)\delta(t-2T) + \cdots = \sum_{n=0}^{+\infty} f(nT)\delta(t-nT).$$

If we now denote by $\mathcal{LT}(f^*)$ the Laplace transform of f^*, i.e. $\mathcal{LT}(f^*(t)) \equiv \widetilde{f}^*(s)$ then,

$$\widetilde{f}^*(s) = \mathcal{LT}(f^*(t)) = \sum_{n=0}^{+\infty} f(nT)\mathcal{LT}(\delta(t-nT)) = \sum_{n=0}^{+\infty} f(nT)e^{-nTs}.$$

Setting $z = e^{-sT}$ or equivalently $s = -(1/T)\ln(z)$, we can then define,

$$F(z) = \sum_{n=0}^{+\infty} f(nT)z^n.$$

This function $F(z)$ is called the *z-transform* of the discrete time signal function $f(nT)$, $F(z) = \mathcal{ZT}(f(t))$. In other words,

$$\mathcal{ZT}(f(t)) = F(z) = \widetilde{f}^*(s) = \widetilde{f}^*\left(-\frac{1}{T}\ln(z)\right) = \left[\mathcal{LT}\left(\sum_{n=0}^{+\infty} f(nT)\delta(t-nT)\right)\right]_{s=-\frac{1}{T}\ln(z)}.$$

2.4 The Wavelet Approximation Method

In this next section we present a novel numerical inversion of the Laplace transform based on Haar wavelets. This approximation fits especially well for the inversion of stepped form functions, which particularly arise when we deal with cumulative distributions of discrete random variables. Later on, this approximation will be used to recover the probability distribution function of the loss function in credit risk portfolios.

Let f be a function in $L^2([0,1])$. According to the theory of MRA and to the Remark 2.2.1 we have,

$$f(x) = \sum_{k=0}^{2^m-1} c_{m,k}\phi_{m,k}(x) + \sum_{j=m}^{+\infty}\sum_{k=0}^{2^j-1} d_{j,k}\psi_{j,k}(x), \tag{2.34}$$

or alternatively,

$$f(x) = \lim_{m\to+\infty} f_m(x), \quad \text{with} \quad f_m(x) = \sum_{k=0}^{2^m-1} c_{m,k}\phi_{m,k}(x), \tag{2.35}$$

where,

$$c_{m,k} = \int_{\frac{k}{2^m}}^{\frac{k+1}{2^m}} f(x)\phi_{m,k}(x)dx, \quad k = 0,\dots,2^m-1,$$

$$d_{j,k} = \int_{\frac{k}{2^m}}^{\frac{k+1}{2^m}} f(x)\psi_{j,k}(x)dx, \quad j \geq m, k = 0,\cdots,2^j-1,$$

and $\{\phi_{m,k}\}_{k=0,\dots,2^m-1} \cup \{\psi_{j,k}\}_{j\geq m,k=0,\cdots,2^j-1}$ is the Haar basis system in $L^2([0,1])$. In what follows, for simplicity we use expression (2.35) instead of expression (2.34) due to the fact that expression (2.34) uses two indexes for the detail coefficients.

Let us consider the Laplace transform of the function f whenever it exists,

$$\widetilde{f}(s) = \int_0^{+\infty} e^{-sx} f(x)dx,$$

where we can assume that $f(x) = 0$ for all $x \notin [0,1]$. The main idea behind the Wavelet Approximation method is to approximate $\widetilde{f}$ by $\widetilde{f_m}$ and then to compute the coefficients $c_{m,k}$ inverting the Laplace Transform. Proceeding this way,

$$\begin{aligned}
\widetilde{f}(s) &= \int_0^{+\infty} e^{-sx} f(x)dx \simeq \int_0^{+\infty} e^{-sx} f_m(x)dx = \int_0^{+\infty} e^{-sx}\left(\sum_{k=0}^{2^m-1} c_{m,k}\phi_{m,k}(x)\right)dx \\
&= \sum_{k=0}^{2^m-1} c_{m,k}\left(\int_0^{+\infty} e^{-sx}\phi_{m,k}(x)dx\right) = 2^{m/2}\sum_{k=0}^{2^m-1} c_{m,k}\left(\int_{\frac{k}{2^m}}^{\frac{k+1}{2^m}} e^{-sx}dx\right) \\
&= \frac{2^{m/2}}{s}\sum_{k=0}^{2^m-1} c_{m,k}e^{-s\frac{k}{2^m}}\left(1-e^{-s\frac{1}{2^m}}\right) = \frac{2^{m/2}}{s}\left(1-e^{-s\frac{1}{2^m}}\right)\sum_{k=0}^{2^m-1} c_{m,k}\mathcal{LT}\left(\delta\left(x-\frac{k}{2^m}\right)\right) \\
&= \frac{2^{m/2}}{s}\left(1-e^{-s\frac{1}{2^m}}\right)\mathcal{LT}\left(\sum_{k=0}^{2^m-1} c_{m,k}\,\delta\left(x-\frac{k}{2^m}\right)\right).
\end{aligned}$$

Finally, making the change of variable $z = e^{-s\frac{1}{2^m}}$,

$$\sum_{k=0}^{2^m-1} c_{m,k} z^k \simeq \overline{Q}_m(z),$$

where,

$$\overline{Q}_m(z) \equiv \frac{2^{m/2}\ln(z)\widetilde{f}(-2^m\ln(z))}{z-1}. \tag{2.36}$$

Here we note that $\overline{Q}_m(z)$ is not defined at $z = 0$. However, if we assume that the limit,

$$\lim_{z\to 0} \overline{Q}_m(z) = \overline{Q}_m^0 < +\infty,$$

then,

$$Q_m(z) \equiv \begin{cases} \overline{Q}_m(z), & \text{if } z \neq 0, \\ \overline{Q}_m^0, & \text{if } z = 0, \end{cases}$$

is analytic inside the disc of the complex plane $\{z \in \mathbb{C} : |z| < r\}$ for $r < 1$, since the singularity in $z = 0$ is avoidable. Then, given the generating function $Q_m(z)$, we can obtain expressions for the coefficients $c_{m,k}$ by means of the Cauchy's integral formula. This is,

$$\begin{aligned} c_{m,0} &\simeq Q_m(0), \\ c_{m,k} &\simeq \frac{1}{2\pi i}\int_\gamma \frac{Q_m(z)}{z^{k+1}}dz, \quad k = 1, ..., 2^m - 1 \quad (z \neq 0) \end{aligned}$$

where γ denotes a circle of radius r, $0 < r < 1$, about the origin.

Considering now the change of variable $z = re^{iu}$, $0 < r < 1$ we have,

$$\begin{aligned} c_{m,k} &\simeq \frac{1}{2\pi r^k}\int_0^{2\pi} \frac{Q_m(re^{iu})}{e^{iku}}du = \frac{1}{2\pi r^k}\int_0^{2\pi} \left[\Re(Q_m(re^{iu}))\cos(ku) + \Im(Q_m(re^{iu}))\sin(ku)\right] du \\ &= \frac{2}{\pi r^k}\int_0^{\pi} \Re(Q_m(re^{iu}))\cos(ku)du, \quad k = 1, ..., 2^m - 1, \end{aligned} \tag{2.37}$$

where again $\Re(z)$ and $\Im(z)$ denotes the real and imaginary part of z respectively.

The integral in (2.37) can be evaluated by means of the trapezoidal rule. So we can define,

$$\begin{aligned} I(r,k) &= \int_0^{\pi} \Re(Q_m(re^{iu}))\cos(ku)du, \\ I(r,k;h) &= \frac{h}{2}\left(Q_m(r) + (-1)^k Q_m(-r) + 2\sum_{j=1}^{m_T-1} \Re(Q_m(re^{ih_j}))\cos(kh_j)\right), \end{aligned}$$

where $h = \frac{\pi}{m_T}$ and $h_j = jh$ for all $j = 0, \ldots, m_T$. Then,

$$c_{m,k} \simeq \frac{2}{\pi r^k} I(r,k) \simeq \frac{2}{\pi r^k} I(r,k;h)$$
$$= \frac{1}{m_T r^k} \left(Q_m(r) + (-1)^k Q_m(-r) + 2 \sum_{j=1}^{m_T - 1} \Re(Q_m(re^{ih_j})) \cos(kh_j) \right), \quad k = 1, ..., 2^m - 1. \tag{2.38}$$

2.4.1 Error Analysis

We note that there are three sources of error in our procedure for computing the numerical Laplace transform inversion using Haar wavelets. The first one is the *truncation error*, which arises when approximating f by f_m at scale m. Following the expression (2.34),

$$\|f - f_m\|^2_{L^2([0,1])} = \left\| \sum_{j=m}^{+\infty} \sum_{k=0}^{2^j - 1} d_{j,k} \psi_{j,k} \right\|^2_{L^2([0,1])} = \sum_{j=m}^{+\infty} \sum_{k=0}^{2^j - 1} |d_{j,k}|^2, \tag{2.39}$$

since $\|\psi_{j,k}\|^2_{L^2([0,1])} = 1$. Then, the truncation error depends on the detail coefficients.

The second source of error is the *discretization error*, which is produced when approximating the integral $I(r,k)$ with $I(r,k;h)$ using the trapezoidal rule. We can apply the formula for the error of the compound trapezoidal rule considering,

$$q_{m,k}(u) = \Re(Q_m(re^{iu})) \cos(ku), \quad E^{m,k}_{T(h)} = I(r,k) - I(r,k;h),$$

and assuming that $q \in C^2([0,\pi])$. Then,

$$\left| E^{m,k}_{T(h)} \right| = \frac{\pi^3}{12 m_T^2} \left| q''_{m,k}(\mu) \right|, \quad \mu \in (0,\pi). \tag{2.40}$$

Finally we must consider the *round-off error*. If we can calculate the sum in expression (2.38) with a precision of about $10^{-\eta}$, then the round-off error after multiplying by the factor $\frac{1}{m_T r^k}$ is approximately $\frac{1}{m_T r^k} \cdot 10^{-\eta}$. Then, the round-off error decays to 0 when $r \nearrow 1$ and conversely, it becomes larger when $r \searrow 0$.

The Haar wavelets method is particularly suitable to invert stepped-form functions, as we illustrate in the following example,

Example 2.4.1. *Let $f(x) = \chi_{[x_0,1)}(x)$, $x_0 \in (0,1)$ be a step function defined in $[0,1]$. Then,*

$$\widetilde{f}(s) = \int_0^{+\infty} e^{-sx} \chi_{[x_0,1)}(x) dx = \frac{1}{s} \left(e^{-sx_0} - e^{-s} \right),$$

is the Laplace transform of f. In order to invert the Laplace transform using the wavelet approximation method, we need to compute the coefficients in expression (2.38). So,

$$Q_m(z) = \frac{2^{m/2} \ln(z) \widetilde{f}(-2^m \ln(z))}{z-1} = \frac{-2^{m/2} \ln(z) \frac{1}{2^m \ln(z)} \left(e^{2^m x_0 \ln(z)} - e^{2^m \ln(z)} \right)}{z-1}$$
$$= \frac{z^{2^m} - z^{2^m x_0}}{2^{m/2}(z-1)},$$

for $z \neq 0$, *and,*

$$c_{m,0} = \lim_{z \to 0} \frac{z^{2^m} - z^{2^m x_0}}{2^{m/2}(z-1)} = 0.$$

Now,

$$\begin{aligned} Q_m(re^{iu}) &= \frac{r^{2^m} e^{i2^m u} - r^{2^m x_0} e^{i2^m x_0 u}}{2^{m/2}(re^{iu} - 1)} \\ &= \frac{r^{2^m}\left(\cos(2^m u) + i\sin(2^m u)\right) - r^{2^m x_0}\left(\cos(2^m x_0 u) + i\sin(2^m x_0 u)\right)}{2^{m/2}(r\cos(u) + ir\sin(u) - 1)}, \end{aligned}$$

where $u \in [0, \pi]$, $0 < r < 1$, *and,*

$$\Re(Q_m(re^{iu})) = \frac{\Re(z_1)\cdot\Re(z_2) + \Im(z_1)\cdot\Im(z_2)}{(\Re(z_2))^2 + (\Im(z_2))^2}, \tag{2.41}$$

where,

$$\begin{aligned} z_1 &= r^{2^m}\left(\cos(2^m u) + i\sin(2^m u)\right) - r^{2^m x_0}\left(\cos(2^m x_0 u) + i\sin(2^m x_0 u)\right), \\ z_2 &= 2^{m/2}(r\cos(u) + ir\sin(u) - 1). \end{aligned}$$

This is,

$$\begin{aligned} \Re(z_1) &= r^{2^m}\cos(2^m u) - r^{2^m x_0}\cos(2^m x_0 u), \quad \Re(z_2) = 2^{m/2}(r\cos(u) - 1), \\ \Im(z_1) &= r^{2^m}\sin(2^m u) - r^{2^m x_0}\sin(2^m x_0 u), \quad \Im(z_2) = 2^{m/2} r\sin(u). \end{aligned}$$

Then, expression (2.41) casts into,

$$\begin{aligned} &\Re(Q_m(re^{iu})) = \\ &= \frac{r^{2^m+1}\cos(2^m u - u) - r^{2^m}\cos(2^m u) - r^{2^m x_0+1}\cos(2^m x_0 u - u) + r^{2^m x_0}\cos(2^m x_0 u)}{2^{m/2}(r^2 - 2r\cos(u) + 1)}. \end{aligned}$$

Now, we must choose an appropriate r *in order to control the discretization and the round-off errors. First of all, we consider the discretization error which can be estimated by means of expression (2.40). We note that* $q_{m,k} \in C^2([0,\pi])$ *since,*

$$0 < (r-1)^2 \leq r^2 - 2r\cos(u) + 1 \leq (r+1)^2, \quad \forall u \in [0, \pi].$$

So we have,

$$\begin{aligned} q'_{m,k}(u) &= \frac{d}{du}\Re(Q_m(re^{iu}))\cos(ku) - k\Re(Q_m(re^{iu}))\sin(ku), \\ q''_{m,k}(u) &= \frac{d^2}{du^2}\Re(Q_m(re^{iu}))\cos(ku) - 2k\frac{d}{du}\Re(Q_m(re^{iu}))\sin(ku) - k^2\Re(Q_m(re^{iu}))\cos(ku). \end{aligned}$$

and,

$$\left|Re(Q_m(re^{iu})\right| \leq \frac{r^{2^m+1} + r^{2^m} + r^{2^m x_0+1} + r^{2^m x_0}}{2^{m/2}(r-1)^2} \simeq r^{2^m x_0} + \mathcal{O}(r^{2^m x_0}), \tag{2.42}$$

where the last approximation holds for suitably small values of the parameter r. For sake of simplicity, we consider only the terms with smaller exponents in the parameter r for the expressions $\frac{d}{du}\Re(Q_m(re^{iu}))$ and $\frac{d^2}{du^2}\Re(Q_m(re^{iu}))$. Then,

$$\left|\frac{d}{du}\Re(Q_m(re^{iu}))\right| \leq \frac{2^{m/2}x_0 r^{2^m x_0} + A(r)}{(r-1)^4} \simeq r^{2^m x_0} + \mathcal{O}(r^{2^m x_0}), \tag{2.43}$$

and,

$$\left|\frac{d^2}{du^2}\Re(Q_m(re^{iu}))\right| \leq \frac{2^{m/2}x_0^2 r^{2^m x_0} + B(r)}{(r-1)^8} \simeq r^{2^m x_0} + \mathcal{O}(r^{2^m x_0}), \tag{2.44}$$

where $A(r)$ and $B(r)$ are polynomials in r with degree greater than $2^m x_0$, and the approximations in (2.43) and (2.44) hold for suitably small values of the parameter r. Finally, taking into account expressions (2.40),(2.42),(2.43) and (2.44) we have,

$$\left|E^{m,k}_{T(h)}\right| \leq \frac{\pi^3}{12m_T^2}\left(r^{2^m x_0} + 2kr^{2^m x_0} + k^2 r^{2^m x_0}\right) + \mathcal{O}(r^{2^m x_0}) = \frac{\pi^3}{12m_T^2}\left(k^2+2k+1\right)r^{2^m x_0} + \mathcal{O}(r^{2^m x_0}).$$

We note that $\left|E^{m,k}_{T(h)}\right| \to 0$ as $r \searrow 0$ while the round-off error is increasing in r as r approaches to zero. Then, the total error should be approximately minimized when the two estimates are equal. This leads to the equation,

$$\frac{1}{m_T r^k}\cdot 10^{-\eta} = \frac{\pi^3}{12m_T^2}\left(k^2+2k+1\right)r^{2^m x_0}.$$

After algebraic manipulation,

$$r_{m,k} = \left(\frac{12m_T \cdot 10^{-\eta}}{\pi^3(k^2+2k+1)}\right)^{\frac{1}{2^m x_0 + k}}, \quad k = 1, \ldots, 2^m - 1. \tag{2.45}$$

In what follows we consider $m_T = 2^m$. Let us assume that $x_0 = 0.5$ is the jump point for function f. Observe that if we take $m = 1$, the approximation error in expression (2.39) equals 0, due to the fact that $d_{j,k} = 0$ for $j \geq 1$ and $k = 0, \ldots, 2^j - 1$. Thus, the approximation is exact and the only remaining errors are the discretization and the round-off errors. If we assume that $\eta = 10^{-15}$, then according to (2.45) $r_{1,1} = 1.39108 \cdot 10^{-8}$ and we obtain $c_{1,0} = 0$, $c_{1,1} = 0.707107$. The plot of figure 2.6 shows the absolute error of this approximation.

Let us now consider $x_0 = 0.75$. In this case using $m = 1$ the approximation error is greater than zero, since the detail coefficient $d_{1,1}$ in expression (2.39) is not zero. However, taking $m = 2$, the detail coefficients become zero for $j \geq 2$ and $k = 0, \ldots, 2^j - 1$. Thus, the approximation at scale 2 is exact for the step function with the jump discontinuity at 0.75. If we assume that $\eta = 10^{-15}$, then we use $r_{2,1} = 1.4026 \cdot 10^{-4}$, $r_{2,2} = 7.0325 \cdot 10^{-4}$, $r_{2,3} = 2.14262 \cdot 10^{-3}$ and we obtain $c_{2,0} = 0$, $c_{2,1} = 0$, $c_{2,2} = 0$, $c_{2,3} = 0.5$.

Although this is not the case of the former examples, we observe that for large values of m, the computation of the coefficients can be intensive since the long sum in (2.38)

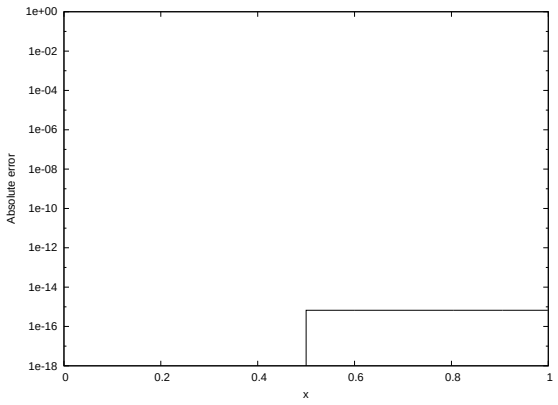

Figure 2.6: Absolute error of the approximation to the step function with $x_0 = 0.5$.

must be evaluated for $k = 1, \ldots, 2^m - 1$. *If we take* $m_T = 2k$ *instead of* $m_T = 2^m$ *this leads to the following expression,*

$$\begin{aligned} c_{m,k} &\simeq \frac{2}{\pi r^k} I(r,k;2k) \\ &= \frac{1}{2kr^k}\left(Q_m(r) + (-1)^k Q_m(-r) + 2\sum_{j=1}^{2k-1} \Re(Q_m(re^{ij\frac{\pi}{2k}}))\cos(kj\frac{\pi}{2k})\right) \\ &= \frac{1}{2kr^k}\left(Q_m(r) + (-1)^k Q_m(-r) + 2\sum_{j=1}^{k-1} (-1)^j \Re(Q_m(re^{ij\frac{\pi}{2k}}))\right), \quad k = 1, ..., 2^m - 1, \end{aligned} \tag{2.46}$$

so the computation time considerable reduces for large scale approximations. If we perform the numerical inversion by means of formula (2.46) then, assuming again that $\eta = 10^{-15}$, *we obtain* $r_{2,1} = 1.17944 \cdot 10^{-4}$, $r_{2,2} = 7.0325 \cdot 10^{-4}$, $r_{2,3} = 2.29242 \cdot 10^{-3}$ *and* $c_{2,0} = 0$, $c_{2,1} = 0$, $c_{2,2} = 0$, $c_{2,3} = 0.5$.

The left plot of figure 2.7 represents the absolute error of the approximation with $h = 2^m$, *while the right plot shows the absolute error with* $h = 2k$.

Observe that if x_0 *is not a dyadic rational number, that is,* $x_0 \neq \frac{p}{q}$ *with* $q = 2^t, t \in \mathbb{Z}$, *then for each scale of approximation, there exists only one detail coefficient which is nonzero. So, if we assume that* x_0 *is not a dyadic rational number, then taking into account expressions (2.18) and (2.39) the approximation error yields,*

$$\|f - f_m\|^2_{L^2([0,1])} = \sum_{j=m}^{+\infty}\sum_{k=0}^{2^j-1} |d_{j,k}|^2 = \sum_{j=m}^{+\infty} |d_{j,k_m}|^2 \simeq \frac{1}{16}\sum_{j=m}^{+\infty}\frac{1}{2^j} = \frac{1}{2^{m+3}},$$

since $d_{j,k} = 0$ *for* $j \geq m, k \neq k_m$ *and* $x_0 \in \left[\frac{k_m}{2^m}, \frac{k_m+1}{2^m}\right]$.

Remark 2.4.1. *For the applications that will be presented later, we have developed the numerical inversion method in the* $L^2([0,1])$ *space. However, the method can be easily*

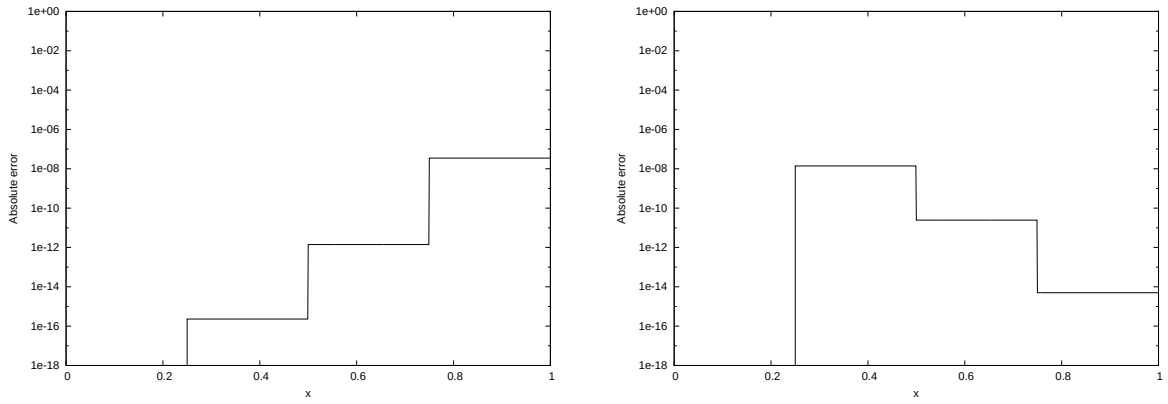

Figure 2.7: Absolute error of the approximation of the step function with $x_0 = 0.75$, fixed step $h = 2^m$ (left) and changing step $h = 2k$ (right).

extended to $L^2(\mathbb{R})$*, allowing the parameter* k *in expressions (2.34) and (2.35) to go to* $-\infty$ *and* $+\infty$*,*

$$f(x) = \sum_{k=-\infty}^{+\infty} c_{m,k}\phi_{m,k}(x) + \sum_{j=m}^{+\infty}\sum_{k=-\infty}^{+\infty} d_{j,k}\psi_{j,k}(x).$$

In this case we can expect that the coefficients converge to zero quickly, due to the fact that $f \in L^2(\mathbb{R})$*, so the parameter* k *can be truncated conveniently.*

Chapter 3

The WA Methodology to Quantify Losses

3.1 Introduction

This chapter is devoted to the application of the Wavelet Approximation method developed in Chapter 2 to compute the distribution of losses from default in a credit portfolio, under the one-factor model framework. This new method has been recently develop by [Mas11] and has been published in *Quantitative Finance*. Direct generation of loss distribution may require Monte Carlo simulation which is time consuming and is not effective. To overcome the computational complexity, in the literature a number of approaches have been undertaken based on assumptions imposed in the input parameters, the accepted level of tolerance and computational errors. A large effort has been dedicated to generation of loss distributions for credit portfolios. We will briefly review those methods applied to a one-factor Merton model framework for credit risk which are widespread used in the industry. The calculation of the tail probabilities of losses is a prerequisite to calculation of the VaR and the Expected Shortfall values. In the Merton model, latent variables drive default for each firm. Defaults are triggered when the latent variables fall below a threshold. The model typically assumes that the latent variables, normally distributed, are dependent on underlying risk factors reflecting the state of the economy (or the business cycle) as well as specific risk factors. Conditional on the systematic risk factor, default indicators random variables are independent and the probabilities of default known. This implies that, conditional on the systematic risk factor, the total loss from default in the credit portfolio is a weighted sum of independent Bernoulli random variables. The main problem here is that we want to estimate the distribution of the total losses from default when the systematic risk factor is not known, the unconditional loss distribution. An exact formula for its distribution is almost always unknown and numerical techniques are required to approximate it.

3.2 Methods for Computing the Loss Distribution

This Chapter focuses on the measurement of name concentration in loan portfolios. The model framework is the one-factor Merton model. We can decompose credit risk in loan portfolios into a systematic and an idiosyncratic component. Systematic risk represents the effect of unexpected changes in macroeconomic and financial market conditions on the performance of borrowers while idiosyncratic risk represents the effects of risks that are particular to individual borrowers.

We give an overview of the most recent research concerning analytical and numerical methods for calculating the tail of the credit portfolio loss distributions. Furthermore, we present our alternative method based on wavelets which is particularly suitable for small or concentrated portfolios due to the fact that the stepped form of the cumulative distribution of losses appears to be accentuated. The Wavelet Approximation method is accurate, fast and robust as we show in the numerical examples section.

3.2.1 The Granularity Adjustment

The Asymptotic Single Risk Factor model discussed in Chapter 1 assumes that portfolios are fully diversified with respect to individual borrowers, so that economic capital depends only on systematic risk. Thus, the IRB formula omits the contribution of the residual idiosyncratic risk to the required economic capital. This form of credit concentration risk is sometimes also called *lack of granularity*. In this section we discuss an approach how to assess a potential add-on to capital for the effect of lack of granularity in a general risk-factor framework. We follow the revised methodology developed in [Gor07], which is similar to the granularity adjustment that was included in the Second Consultative Paper of Basel II (see [BCBS01]).

The Granularity Adjustment has been developed as an extension of the ASRF model which underpins the IRB model. Let Y denote the systematic risk factor, U_n denote the loss rate in position n and consider our portfolio $\mathcal{L}$ of N risky loans as in Section 1.4,

$$\mathcal{L} = \sum_{n=1}^{N} s_n \cdot U_n.$$

When economic capital is measured as Value at Risk at the α^{th} percentile, we wish to estimate $\mathrm{VaR}_\alpha(\mathcal{L})$. The IRB formula, however, delivers the α^{th} percentile of the expected loss conditional to the systematic factor $\mathrm{VaR}_\alpha(\mathbb{E}(\mathcal{L}|Y))$. The difference, $\mathrm{VaR}_\alpha(\mathcal{L}) - \mathrm{VaR}_\alpha(\mathbb{E}(\mathcal{L}|Y))$ is the adjustment for the effect of undiversified idiosyncratic risk in the portfolio. If we define the functions $\mu(Y) = \mathbb{E}(\mathcal{L}|Y)$ and $\sigma^2(Y) = \mathbb{V}(\mathcal{L}|Y)$ as the conditional mean and variance of the portfolio loss respectively, and let ϕ be the probability density function of Y, then, for $\epsilon = 1$ the portfolio loss $\mathcal{L}$ is given by,

$$\mathcal{L} = \mathbb{E}(\mathcal{L}|Y) + \epsilon(\mathcal{L} - \mathbb{E}(\mathcal{L}|Y)).$$

The asymptotic expansion for the calculation of the α^{th} quantile of the portfolio loss as shown in [Wil01] is based on a second order Taylor expansion in powers of ϵ around the

conditional mean $\mathbb{E}(\mathcal{L}|Y)$. Evaluating the resulting formula at $\epsilon = 0$, we obtain,

$$\begin{aligned}\mathrm{VaR}_\alpha(\mathcal{L}) = \mathrm{VaR}_\alpha(\mathbb{E}(\mathcal{L}|Y)) + \frac{\partial}{\partial\epsilon}\mathrm{VaR}_\alpha\left(\mathbb{E}(\mathcal{L}|Y) + \epsilon(\mathcal{L} - \mathbb{E}(\mathcal{L}|Y))\right)|_{\epsilon=0} + \\ + \frac{1}{2}\frac{\partial^2}{\partial^2\epsilon}\mathrm{VaR}_\alpha\left(\mathbb{E}(\mathcal{L}|Y) + \epsilon(\mathcal{L} - \mathbb{E}(\mathcal{L}|Y))\right)|_{\epsilon=0} + \mathcal{O}(\epsilon^3).\end{aligned}$$

Hence, the granularity adjustment of the portfolio is given by,

$$\mathrm{GA} = \frac{\partial}{\partial\epsilon}\mathrm{VaR}_\alpha\left(\mu(Y) + \epsilon(\mathcal{L} - \mu(Y))\right)|_{\epsilon=0} + \frac{1}{2}\frac{\partial^2}{\partial^2\epsilon}\mathrm{VaR}_\alpha\left(\mu(Y) + \epsilon(\mathcal{L} - \mu(Y))\right)|_{\epsilon=0}.$$

Using Lemma 1 in [Gou00], it can be proved that the first derivative in the Taylor expansion of the quantile vanishes. So taking this fact into account, [Gou00] finally proves that the remaining second derivative in the Taylor expansion, and thus the granularity adjustment, can be expressed as,

$$\mathrm{GA} = \frac{-1}{2\phi(l_{1-\alpha}(Y))} \cdot \frac{\partial}{\partial y}\left(\frac{\sigma^2(y)\phi(y)}{\mu'(y)}\right)\Big|_{y=l_{1-\alpha}(Y)}.$$

As the GA is derived as an asymptotic approximation, the GA formula might not work well on small portfolios. GA overstates the effect of granularity, but is quite accurate for modest sized portfolios of about 200 obligors (for a low quality portfolio) or like 500 obligors (for and investment-grade portfolio).

3.2.2 Recursive Approximation

The Recursive Approximation was pioneered by [And03] in order to obtain the distribution of portfolio losses for single-tranche CDO sensitivity calculations. Consider the credit portfolio $\mathcal{L}$ with N obligors as before. Let $\mathcal{L}_k$ be the credit portfolio consisting of the first k obligors, this is $\mathcal{L}_k = E_1D_1 + \cdots + E_kD_k$, where we have assumed constant loss given default 100%. This random variable takes values in the set $\mathcal{L}_k \in \{0, \ldots, l_{\max,k}\}$, where $l_{\max,k} = \sum_{i=1}^k E_i$.

Given the systematic risk factor Y, the probability that the loss for the credit portfolio with the first k obligors is l will be denoted by $P^k(l|Y) \equiv \mathbb{P}(\mathcal{L}_k = l|Y = y)$. These probabilities satisfy the recursive relationship,

$$P^{k+1}(l|Y) = P^k(l - E_{k+1}|Y)P_{k+1}(y) + P^k(l|Y)(1 - P_{k+1}(y)), \tag{3.1}$$

since,

$$\begin{aligned}P^{k+1}(l|Y) &= \mathbb{P}(\mathcal{L}_{k+1} = l|Y = y) = \mathbb{P}(\mathcal{L}_{k+1} = l|(Y = y) \cap (D_{k+1} = 1))\mathbb{P}(D_{k+1} = 1) + \\ &+ \mathbb{P}(\mathcal{L}_{k+1} = l|(Y = y) \cap (D_{k+1} = 0))(1 - \mathbb{P}(D_{k+1} = 1)) \\ &= \mathbb{P}(\mathcal{L}_{k+1} = l - E_{k+1}|(Y = y) \cap (D_{k+1} = 1))\mathbb{P}(D_{k+1} = 1) + \\ &+ \mathbb{P}(\mathcal{L}_k = l|(Y = y) \cap (D_{k+1} = 0))(1 - \mathbb{P}(D_{k+1} = 1)) \\ &= \mathbb{P}(\mathcal{L}_k = l - E_{k+1}|Y = y)P_{k+1}(y) + \mathbb{P}(\mathcal{L}_k = l|Y = y)(1 - P_{k+1}(y)).\end{aligned}$$

The initial condition of this recursive relationship is $P^0(l|Y) = \delta_{0l}$, where $\delta_{0l} = 1$ if $l = 0$ and 0 otherwise. This initial condition and the recursive relationship (3.1) allow to find the final distribution of the total loss for the credit portfolio,

$$P^N(l|Y) = \mathbb{P}(\mathcal{L}_N = l|Y = y) = \mathbb{P}(\mathcal{L} = l|Y = y).$$

Now, to obtain the conditional tail probability for the loss of the credit portfolio $\mathbb{P}(\mathcal{L} > l|Y = y)$, we need to add the conditional tail probabilities over all the possible losses above l,

$$\mathbb{P}(\mathcal{L} > l|Y = y) = \sum_{\bar{l}>l}^{l_{\max}} \mathbb{P}(\mathcal{L} = \bar{l}|Y = y), \quad \text{where } l_{\max} = \sum_{i=1}^{N} E_i.$$

Finally, we can integrate numerically over the risk factor Y in order to evaluate the unconditional tail probability $\mathbb{P}(\mathcal{L} > l)$ for the loss of the credit portfolio.

[Gla07] implements the recursive approximation in the one-factor and multi-factor setting and conclude that the method performs well and give accurate estimates for the tail probability of the losses when the number of obligors is small, however, the quality of the approximations obtained with this method deteriorate quickly as the number of obligors increase.

3.2.3 Normal Approximation

The Normal Approximation is a direct application of the central limit theorem (CLT) and can be found in [Mar04]. We recall that the CLT roughly states that, under appropriate conditions, the distribution of the sum of a large number of independent random variables is normally distributed. When the portfolio is not sufficiently large for the law of large numbers to hold, or not very homogeneous, unsystematic risk arises. Then we need to take into account the variability of portfolio loss $\mathcal{L}$ conditional to the common factor Y. This can be easily approximated using the CLT. Conditional on the common factor Y, the portfolio loss $\mathcal{L}$ is normally distributed with mean $\mu(Y)$ and variance $\sigma^2(Y)$ such that,

$$\mu(Y) = \sum_{n=1}^{N} E_n \cdot P_n(y), \ \sigma^2(Y) = \sum_{n=1}^{N} E_n^2 \left(P_n(y)\right)\left(1 - P_n(y)\right),$$

where $P_n(y)$ are the conditional default probabilities, and we have assumed $L_n = 1$ for simplicity and clarity. Then, it follows that the conditional tail probability reads,

$$\mathbb{P}(\mathcal{L} > x|Y) = \Phi\left(\frac{\mu(Y) - x}{\sigma^2(Y)}\right).$$

The unconditional default probability can then be obtained by integrating over the factor Y,

$$\mathbb{P}(\mathcal{L} > x) = \mathbb{E}\left(\Phi\left(\frac{\mu(Y) - x}{\sigma^2(Y)}\right)\right) = \int_{\mathbb{R}} \Phi\left(\frac{\mu(y) - x}{\sigma^2(y)}\right)\phi(y)dy,$$

where ϕ is the PDF of the standard normal distribution.

If the number of assets in the portfolio is large, and the individual default probabilities are not too small, this is a good approximation. However, in general, it is quite a poor approximation in practice, because credit loss distributions by their nature are asymmetric (the downside is much more severe than the upside). In fact, the normal approximation is particularly bad at estimating the tail of the distribution, that is, the region of high losses. In [Luc99] there is a detailed numerical analysis about the quality of this approximation.

3.2.4 Saddle Point Approximation

The computation of the PDF of the sum of independent random variables can be facilitated by the use of the MGF, which is defined by $M_{\mathcal{X}}(s) = \mathbb{E}(e^{s\mathcal{X}})$. For a finite sequence of independent random variables $\mathcal{X}_i$ for $i = 1, \ldots, n$, with known analytic moment generating functions $M_{\mathcal{X}_i}$, the MGF of the sum $X = \sum_{i=1}^{n} \mathcal{X}_i$ is the product of MGF of $\mathcal{X}_i$,

$$M_{\mathcal{X}}(s) = \prod_{i=1}^{n} M_{\mathcal{X}_i}.$$

The moment generating function of the loss random variable $\mathcal{L} = \sum_{n=1}^{N} \mathcal{L}_n$ with density $f_{\mathcal{L}}$ is defined as the analytic function,

$$M_{\mathcal{L}}(s) = \mathbb{E}(e^{s\mathcal{L}}) = \int_0^{+\infty} e^{sl} f_{\mathcal{L}}(l) dl,$$

of a complex variable s, provided that the integral exists. Inverting the above formula gives the well known Bromwich integral,

$$f_{\mathcal{L}}(l) = \frac{1}{2\pi i} \int_{-i\infty}^{+i\infty} e^{-sl} M_{\mathcal{L}}(s) ds, \tag{3.2}$$

where the path of integration is the imaginary axis. Saddle Point approximation arises in this setting to give an accurate analytic approximation. Further details of this approximation can be found in [Jen95]. The Saddle Point method provides a tool to derive the PDF of such a sum of random variables by approximating the integral in the inversion formula (3.2).

Define the *cumulant generating function* (CGF) of $\mathcal{L}$ as $K_{\mathcal{L}}(s) = \ln(M_{\mathcal{L}}(s))$. Thus, we can write equation (3.2) as,

$$f_{\mathcal{L}}(l) = \frac{1}{2\pi i} \int_{-i\infty}^{+i\infty} e^{K_{\mathcal{L}}(s) - sl} ds. \tag{3.3}$$

The *saddle points* are the points at which the terms in the exponential are stationary, i.e. the points s for which,

$$\frac{\partial}{\partial s}(K_{\mathcal{L}}(s) - sl) = \frac{\partial K_{\mathcal{L}}(s)}{\partial s} - l = 0. \tag{3.4}$$

Having found these points, the Saddle Point approximation method consists in applying a second order Taylor expansion to the term $K_{\mathcal{L}}(s) - sl$ and then computing the resulting Gaussian integral. By equation (3.4) for a fixed value l the saddle point $\hat{l}$ satisfies,

$$\left.\frac{\partial K_{\mathcal{L}}(s)}{\partial s}\right|_{s=\hat{l}} = \frac{1}{M_{\mathcal{L}}(\hat{l})} \cdot \left.\frac{\partial M_{\mathcal{L}}(s)}{\partial s}\right|_{s=\hat{l}} = l. \tag{3.5}$$

It can be shown that for each l in the distribution of $\mathcal{L}$ there exists a unique $\hat{l}$ on the real axis that solves (3.5) since the cumulant generating function is convex. Note that for $\hat{l} = 0$ we have,

$$\left.\frac{\partial K_{\mathcal{L}}(s)}{\partial s}\right|_{s=0} = \frac{1}{M_{\mathcal{L}}(0)} \cdot \left.\frac{\partial M_{\mathcal{L}}(s)}{\partial s}\right|_{s=0} = \int_0^{+\infty} u \cdot f_{\mathcal{L}}(u)du = \mathbb{E}(\mathcal{L}),$$

implying that $\hat{l} = 0$ is the saddle point corresponding to the value $l = \mathbb{E}(\mathcal{L})$. If we think of l being a quantile of the distribution of $\mathcal{L}$ then, due to equation (3.5), we can relate quantities l of the distribution of $\mathcal{L}$ to $K_{\mathcal{L}}$ evaluated at the saddle point $\hat{l}$. This is an important feature for applications in credit risk. Moreover, this is also one reason to obtain analytical expressions for the portfolio VaR.

Applying a second order Taylor expansion to the term $K_{\mathcal{L}}(s) - sl$ around $\hat{l}$ in equation (3.3),

$$K_{\mathcal{L}}(s) - sl \simeq K_{\mathcal{L}}(\hat{l}) - l\hat{l} + \frac{1}{2}K''_{\mathcal{L}}(\hat{l})\left(s - \hat{l}\right)^2,$$

and then computing the resulting Gaussian integral we obtain,

$$f_{\mathcal{L}}(l) = \frac{1}{2\pi i}\int_{-i\infty}^{+i\infty} e^{K_{\mathcal{L}}(s)-sl}ds \simeq \frac{e^{K_{\mathcal{L}}(\hat{l})-l\hat{l}}}{2\pi i}\int_{-i\infty}^{+i\infty} e^{\frac{1}{2}(s-\hat{l})^2 K''_{\mathcal{L}}(\hat{l})}ds = \frac{e^{K_{\mathcal{L}}(\hat{l})-l\hat{l}}}{\sqrt{2\pi K''_{\mathcal{L}}(\hat{l})}}. \tag{3.6}$$

Here we use that the first term in the Taylor expansion vanishes at $\hat{l}$ due to the definition of the saddle point $\hat{l}$. Note that by applying equation (3.3) we can compute the tail probability as,

$$\begin{aligned}\mathbb{P}(\mathcal{L} > l) &= \int_l^{\infty} f_{\mathcal{L}}(u)du = \int_l^{\infty}\left(\frac{1}{2\pi i}\int_{-i\infty}^{+i\infty} e^{K_{\mathcal{L}}(s)-su}ds\right)du \\ &= \frac{1}{2\pi i}\int_{-i\infty,(0+)}^{+i\infty}\left[\frac{e^{K_{\mathcal{L}}(s)-su}}{-s}\right]_{u=l}^{\infty} ds = \frac{1}{2\pi i}\int_{-i\infty,(0+)}^{+i\infty}\frac{e^{K_{\mathcal{L}}(s)-sl}}{s}ds,\end{aligned}$$

where the integration path is the imaginary axis and the notation $(0+)$ denotes that the contour runs to the right of the origin to avoid the pole there. Thus, by applying formula (3.6) the tail probability of $\mathcal{L}$ can be recovered from the CGF by a contour integral of the form,

$$\mathbb{P}(\mathcal{L} > l) = \frac{1}{2\pi i}\int_{-i\infty,(0+)}^{+i\infty}\frac{e^{K_{\mathcal{L}}(s)-sl}}{s}ds \simeq \frac{e^{K_{\mathcal{L}}(\hat{l})-l\hat{l}}}{2\pi i}\int_{-i\infty,(0+)}^{+i\infty}\frac{e^{\frac{1}{2}(s-\hat{l})^2 K''_{\mathcal{L}}(\hat{l})}}{s}ds.$$

This can be further elaborated and we finally obtain,

$$\mathbb{P}(\mathcal{L} > l) \simeq \begin{cases} e^{K_{\mathcal{L}}(\hat{l}) - l\hat{l} + \frac{1}{2}\hat{l}^2 K''_{\mathcal{L}}(\hat{l})} \Phi\left(-\sqrt{\hat{l}^2 K''_{\mathcal{L}}(\hat{l})}\right), & \text{for } l > \mathbb{E}(\mathcal{L}), \\ \frac{1}{2}, & \text{for } l = \mathbb{E}(\mathcal{L}), \\ 1 - e^{K_{\mathcal{L}}(\hat{l}) - l\hat{l} + \frac{1}{2}\hat{l}^2 K''_{\mathcal{L}}(\hat{l})} \Phi\left(-\sqrt{\hat{l}^2 K''_{\mathcal{L}}(\hat{l})}\right), & \text{for } l < \mathbb{E}(\mathcal{L}), \end{cases} \tag{3.7}$$

where Φ denotes the cumulative normal distribution function.

Provided that one can calculate $K_{\mathcal{L}}(s)$, $K'_{\mathcal{L}}(s)$, $K''_{\mathcal{L}}(s)$ for any value of s, it is quite easy to use the Saddle Point method in practice. For example, to calculate the VaR at a confidence level α, i.e., the value l such that $\mathbb{P}(\mathcal{L} > l) = 1 - \alpha$, for a given value $\alpha \in (0, 1)$, we use (3.7) with l replaced by $K'_{\mathcal{L}}(\hat{l})$ and adjust $\hat{l}$ until the right-hand side of (3.7) becomes equal to $1 - \alpha$. This is a straightforward root-finding problem.

If all of the $\mathcal{L}_n$ are identically distributed, the relative errors of both approximations in (3.6) and (3.7) are known to be $\mathcal{O}(N^{-1})$. Higher order approximations of the density and the tail probability are given by [Dan87],

$$f_{\mathcal{L}}(l) = \frac{\phi(z_1)}{\sqrt{K''_{\mathcal{L}}(\hat{l})}} \left(\left\{ 1 + \left(-\frac{5K'''_{\mathcal{L}}(\hat{l})^2}{24K''_{\mathcal{L}}(\hat{l})^3} + \frac{K^{(4)}_{\mathcal{L}}(\hat{l})}{8K''_{\mathcal{L}}(\hat{l})^2} \right) \right\} + \mathcal{O}(N^{-2}) \right),$$

with ϕ the PDF of the standard normal distribution, and the Lugannani and Rice (see [Lug80]) formula,

$$\mathbb{P}(\mathcal{L} > l) = 1 - \phi(z_1) + \phi(z_1) \left(\frac{1}{z_2} - \frac{1}{z_1} + \mathcal{O}(N^{-3/2}) \right), \tag{3.8}$$

where $z_1 = sgn(\hat{l})\sqrt{2(l\hat{l} - K_{\mathcal{L}}(\hat{l}))}$ and $z_2 = \hat{l}\sqrt{K''_{\mathcal{L}}(\hat{l})}$.

The Saddle Point approximation is highly accurate in the tail of a distribution. The use of saddle point approximations in credit loss portfolio is pioneered by a series of papers by R. Martin, K. Thompson and C. Browne ([Mar01a], [Mar01b], [Mar01c]). They apply the Saddle Point approximation to the unconditional moment generating function, despite the fact that the individual loss variables $\mathcal{L}_i$ are not independent. A different approach has been proposed by [Hua07a] that consists in applying the Saddle Point approximation to the *conditional* moment generating function instead of the unconditional moment generating function, so that $\mathcal{L}$ is a weighted sum of independent Bernoulli random variables. This is the situation where the Saddle Point approximation works well. In this approximation, a uniform accuracy of density and tail probability for different levels of portfolio loss $\mathcal{L}$ is achieved at the expense of some extra computational cost, due to the fact that the saddle points need to be found for each realization of the factor Y. However, if there is a strong concentration in the portfolio, for instance, if the exposure is dominated by a few loans, a straightforward approximation may be insufficient.

Recall that in the one-factor model framework, if the systematic factor Y is fixed, default occurs independently because the only remaining uncertainty is the idiosyncratic

risk. The MGF conditional on Y is thus given by the product of each obligor's MGF as,

$$\bar{M}_{\mathcal{L}}(s;y) \equiv \mathbb{E}(e^{s\mathcal{L}} \mid Y=y) = \prod_{n=1}^{N} \mathbb{E}\left(e^{sE_nL_nD_n} \mid Y=y\right) = \prod_{n=1}^{N} \left[1-P_n(y)+P_n(y)e^{sE_nL_n}\right],$$

where $P_n(y)$ for $n=1,\ldots,N$ are the conditional default probabilities.

Notice that we are assuming non stochastic LGD. If we define $w_n = L_nE_n$ and the conditional CGF $\bar{K}_{\mathcal{L}}(s;y) = \ln \bar{M}_{\mathcal{L}}(s;y)$,

$$\begin{aligned}
\bar{K}_{\mathcal{L}}(s;y) &= \sum_{n=1}^{N} \ln\left(1-P_n(y)+P_n(y)e^{w_ns}\right), \\
\bar{K}'_{\mathcal{L}}(s;y) &= \sum_{n=1}^{N} \frac{w_nP_n(y)e^{w_ns}}{1-P_n(y)+P_n(y)e^{w_ns}}, \\
\bar{K}''_{\mathcal{L}}(s;y) &= \sum_{n=1}^{N} \frac{(1-P_n(y))w_n^2P_n(y)e^{w_ns}}{(1-P_n(y)+P_n(y)e^{w_ns})^2}, \\
\bar{K}'''_{\mathcal{L}}(s;y) &= \sum_{n=1}^{N} \left(\frac{(1-P_n(y))w_n^3P_n(y)e^{w_ns}}{(1-P_n(y)+P_n(y)e^{w_ns})^2} - \frac{2(1-P_n(y))w_n^3P_n^2(y)e^{2w_ns}}{(1-P_n(y)+P_n(y)e^{w_ns})^3}\right), \\
\bar{K}^{(4)}_{\mathcal{L}}(s;y) &= \sum_{n=1}^{N} \left(\frac{(1-P_n(y))w_n^4P_n(y)e^{w_ns}}{(1-P_n(y)+P_n(y)e^{w_ns})^2} - \frac{6(1-P_n(y))w_n^4P_n^2(y)e^{2w_ns}}{(1-P_n(y)+P_n(y)e^{w_ns})^3} + \right. \\
&\quad \left. + \frac{6(1-P_n(y))w_n^4P_n^3(y)e^{3w_ns}}{(1-P_n(y)+P_n(y)e^{w_ns})^4}\right).
\end{aligned} \tag{3.9}$$

With the above expressions available, we are able to calculate the conditional loss density and conditional tail probability by the Saddle Point approximation. Since $\bar{K}'_{\mathcal{L}}(s;y)$ is a monotonically increasing function of s, and it is bounded in the interval $[0, \sum_{n=1}^{N} w_n]$, the equation $\bar{K}'_{\mathcal{L}}(s;y) = l$ admits a unique solution $\hat{l}$ for $l \in [0, \sum_{n=1}^{N} w_n]$. So, once we have calculated the conditional tail probability $\mathbb{P}(\mathcal{L} > l \mid Y = y)$, we can integrate over the Y factor and the unconditional tail probability yields,

$$\mathbb{P}(\mathcal{L} > l) = \int_{\mathbb{R}} \mathbb{P}(\mathcal{L} > l \mid Y = y)\phi(y)dy.$$

3.3 The Wavelet Approximation Method for Computing Portfolio Losses

Let us consider the portfolio $\mathcal{L} = \sum_{n=1}^{N} \mathcal{L}_n$, with N obligors where $\mathcal{L}_n = E_n \cdot L_n \cdot D_n$ and let F be the cumulative distribution function of $\mathcal{L}$. Without loss of generality, we can assume $\sum_{n=1}^{N} E_n = 1$ and constant loss given default $L_n = 100\%$. To describe the obligor's default and its correlation structure, we use the one-factor Merton model with constant correlation ρ,

$$r_n = \sqrt{\rho}Y + \sqrt{1-\rho}\epsilon_n,$$

where Y and ϵ_n, $\forall n \leq N$ are i.i.d. standard normally distributed. Hence, the default probability of obligor n conditional on a specification $Y = y$ according to (1.10) can be written as,

$$P_n(y) \equiv \Phi\left(\frac{t_n - \sqrt{\rho}y}{\sqrt{1-\rho}}\right),$$

where $t_n = \Phi^{-1}(P_n)$, and Φ denotes the cummulative standard normal distribution function and Φ^{-1} its inverse.

Let us consider also,

$$F(x) = \begin{cases} \overline{F}(x), & \text{if } 0 \leqslant x \leqslant 1, \\ 1, & \text{if } x > 1, \end{cases}$$

for a certain $\overline{F}$ defined in $[0, 1]$.

Recall that the MGF conditional to Y is given by the product of each obligor's MGF,

$$\bar{\tilde{M}}_{\mathcal{L}}(s; y) = \prod_{n=1}^{N} \left[1 - P_n(y) + P_n(y)e^{-sE_n}\right].$$

Notice that we are assuming non stochastic LGD. Taking the expectation value of this conditional MGF yields the unconditional MGF,

$$\begin{aligned} \tilde{M}_{\mathcal{L}}(s) \equiv \mathbb{E}(e^{-s\mathcal{L}}) = \mathbb{E}(\mathbb{E}(e^{-s\mathcal{L}} \mid Y = y)) = \mathbb{E}(\bar{\tilde{M}}_{\mathcal{L}}(s; y)) = \mathbb{E}\left[\prod_{n=1}^{N} \left[1 - P_n(y) + P_n(y)e^{-sE_n}\right]\right] \\ = \int_{\mathbb{R}} \prod_{n=1}^{N} \left[1 - P_n(y) + P_n(y)e^{-sE_n}\right] \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy. \end{aligned} \tag{3.10}$$

Let us note a property regarding the CDF $\overline{F}$ previously defined. Since the loss can take only a (very big) finite number of discrete values (2^N at most), the PDF of the loss function is a sum of Dirac delta functions and then the CDF is a discontinuous function. Moreover, the stepped form of the CDF makes Haar wavelets a natural and very well-suited approximation procedure.

Since $\overline{F} \in L^2([0, 1])$, according to the theory of MRA we can approximate $\overline{F}$ in $[0, 1]$ by a summation of scaling functions,

$$\overline{F}(x) \simeq \overline{F}_m(x), \quad \overline{F}_m(x) = \sum_{k=0}^{2^m - 1} c_{m,k}\phi_{m,k}(x), \tag{3.11}$$

and,

$$\overline{F}(x) = \lim_{m \to +\infty} \overline{F}_m(x). \tag{3.12}$$

But if $f_{\mathcal{L}}$ is the probability density function of the loss function, then the unconditional MGF is also the Laplace transform of $f_{\mathcal{L}}$:

$$\tilde{M}_{\mathcal{L}}(s) \equiv \mathbb{E}(e^{-s\mathcal{L}}) = \int_0^{+\infty} e^{-sx} f_{\mathcal{L}}(x)dx = \widetilde{f_{\mathcal{L}}}(s). \tag{3.13}$$

Also, as we have noticed before,

$$f_{\mathcal{L}}(x) = \sum_{i=1}^{2^N} \mu_i \delta(x - x_i), \quad x_1, x_2, ..., x_{2^N} \in [0, 1]. \tag{3.14}$$

where $\delta(x - x_i)$ is the Dirac delta at x_i. Each Dirac delta can be thought as a density distribution of a unit of mass concentrated in the point x_i (i.e. $\int_0^{+\infty} g(x)\delta(x - x_i)dx = g(x_i)$, for every test function $g(x)$). Probabilistically, a distribution like (3.14) corresponds to a situation where only the scenarios $x_1, x_2, ..., x_{2^N}$ are feasible and with respective probabilities $\mu_1, \mu_2, ..., \mu_{2^N}$. Of course these probabilities must be positive and sum up to 1. This is,

$$\sum_{i=1}^{2^N} \mu_i = 1.$$

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 expression (3.13),

$$\tilde{M}_{\mathcal{L}}(s) = \int_0^{+\infty} e^{-sx} F'(x)dx = e^{-s} + s \int_0^1 e^{-sx} \overline{F}(x)dx. \tag{3.15}$$

So, $\left(\tilde{M}_{\mathcal{L}}(s) - e^{-s}\right)/s$ is the Laplace transform of $\overline{F}$. Then, we can apply the Wavelet Approximation method developed in Section 2.4 in order to invert the Laplace transform of $\overline{F}$. Following the formula (2.36) and making the change of variable $z = e^{-s\frac{1}{2^m}}$, we consider,

$$Q_m(z) = \frac{2^{m/2}\ln(z) \cdot \frac{\tilde{M}_{\mathcal{L}}(-2^m \ln(z)) - z^{2^m}}{-2^m \ln(z)}}{z - 1} = \frac{\tilde{M}_{\mathcal{L}}(-2^m \ln(z)) - z^{2^m}}{2^{m/2}(1 - z)}, \quad z \neq 0, \tag{3.16}$$

and,

$$Q(0) = \frac{\int_{\mathbb{R}} \prod_{n=1}^{N} [1 - P_n(y)] \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy}{2^{\frac{m}{2}}}.$$

Finally, for a fixed scale of approximation m, we can compute the coefficients of the approximation by,

$$c_{m,0} \simeq \frac{\int_{\mathbb{R}} \prod_{n=1}^{N} [1 - P_n(y)] \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy}{2^{\frac{m}{2}}},$$

and,

$$c_{m,k} \simeq \frac{2}{\pi r^k} \int_0^{\pi} \Re(Q_m(re^{iu})) \cos(ku)du, \quad k = 1, \ldots, 2^m - 1. \tag{3.17}$$

We can evaluate this integral by means of the ordinary trapezoidal rule with the required accuracy to obtain the coefficients. As a matter of fact, in the numerical examples section

we see that when computing VaR values at 99.9% confidence level, taking 2^m subintervals we converge towards the Monte Carlo result. Also is worth to mention that the MGF in the expression (3.10) is also accurately computed using a Gauss-Hermite quadrature formula with 20 nodes.

3.3.1 Computation of the Coefficients $c_{m,k}$

Let us explain in detail the computation of the coefficients in expression (3.17). The integral can be evaluated by means of the trapezoidal rule. So we consider,

$$I(r,k) = \int_0^{\pi} \Re(Q_m(re^{iu}))\cos(ku)du, \tag{3.18}$$

$$I(r,k;h) = \frac{h}{2}\left(Q_m(r) + (-1)^k Q_m(-r) + 2\sum_{j=1}^{m_T-1} \Re(Q_m(re^{ih_j}))\cos(kh_j)\right), \tag{3.19}$$

$$\tag{3.20}$$

where $h = \frac{\pi}{m_T}$ and $h_j = jh$ for all $j = 0, \ldots, m_T$. Then, the coefficients can be approximated by,

$$\begin{aligned} c_{m,k} &\simeq \frac{2}{\pi r^k} I(r,k) \simeq \frac{2}{\pi r^k} I(r,k;h) \\ &= \frac{1}{m_T r^k}\left(Q_m(r) + (-1)^k Q_m(-r) + 2\sum_{j=1}^{m_T-1} \Re(Q_m(re^{ih_j}))\cos(kh_j)\right), \quad k = 1, ..., 2^m - 1. \end{aligned} \tag{3.21}$$

Here, we must control the discretization and the round-off errors, the last one arising when we multiply by the prefactor $(m_T r^k)^{-1}$ in expression (3.21). If we can compute $I(r,k;h)$ with a precision of about $10^{-\eta}$, then the round-off error is approximately $(m_T r^k)^{-1} \cdot 10^{-\eta}$ after multiplying by the prefactor.

Considering the expression (3.16),

$$Q_m(re^{iu}) = \frac{\tilde{M}_{\mathcal{L}}(-2^m \ln(re^{iu})) - (re^{iu})^{2^m}}{2^{m/2}(1 - re^{iu})},$$

where $\tilde{M}_{\mathcal{L}}$ is the unconditional Moment Generating Function (3.10).

If we define $z_1 = \tilde{M}_{\mathcal{L}}(-2^m \ln(re^{iu})) - (re^{iu})^{2^m}$ and $z_2 = 2^{m/2}(1 - re^{iu})$ then we have,

$$\Re(Q_m(re^{iu})) = \frac{\Re(z_1)\Re(z_2) + \Im(z_1)\Im(z_2)}{(\Re(z_2))^2 + (\Im(z_2))^2}, \tag{3.22}$$

where,

$$\Re(z_1) = \Re(\tilde{M}_{\mathcal{L}}(-2^m \ln(re^{iu}))) - r^{2^m}\cos(2^m u), \quad \Re(z_2) = 2^{m/2}(1 - r\cos(u)),$$

and,

$$\Im(z_1) = \Im(\tilde{M}_{\mathcal{L}}(-2^m \ln(re^{iu}))) - r^{2^m}\sin(2^m u), \quad \Im(z_2) = -2^{m/2} r \sin(u).$$

Let us define $g_m(u) \equiv \Re(Q_m(re^{iu}))$, $N_{g_m(u)} \equiv \Re(z_1)\Re(z_2) + \Im(z_1)\Im(z_2)$ and $D_{g_m(u)} \equiv (\Re(z_2))^2 + (\Im(z_2))^2$. Now,

$$D_{g_m(u)} = 2^m(1 + r^2\cos^2(u) - 2r\cos(u)) + 2^m r^2 \sin^2(u) = 2^m(1 + r^2 - 2r\cos(u)).$$

Observe that,

$$0 < 2^m(1-r)^2 \leq D_{g_m(u)} \leq 2^m(1+r)^2, \quad \forall u \in [0,\pi], \forall r \in (0,1), \tag{3.23}$$

and,

$$\left|N_{g_m(u)}\right| \leq 2^{m/2}\left(\left|\Re(\tilde{M}_{\mathcal{L}}(-2^m\ln(re^{iu})))\right| + r^{2^m}\right)(1+r) + 2^{m/2}\left(\left|\Im(\tilde{M}_{\mathcal{L}}(-2^m\ln(re^{iu})))\right| + r^{2^m}\right)r. \tag{3.24}$$

Joining the expressions (3.23) and (3.24) we have,

$$|g_m(u)| \leq \frac{\left(\left|\Re(\tilde{M}_{\mathcal{L}}(-2^m\ln(re^{iu})))\right| + r^{2^m}\right)(1+r) + \left(\left|\Im(\tilde{M}_{\mathcal{L}}(-2^m\ln(re^{iu})))\right| + r^{2^m}\right)r}{2^{m/2}(1-r)^2}.$$

Moreover, if we consider expression (3.10),

$$\begin{aligned}
|\tilde{M}_{\mathcal{L}}(-2^m\ln(re^{iu}))| &\leq \int_{\mathbb{R}} \prod_{n=1}^{N} \left|1 - P_n(y) + P_n(y)e^{2^m E_n \ln(re^{iu})}\right| \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\
&\leq \int_{\mathbb{R}} \prod_{n=1}^{N} \left(|1 - P_n(y)| + |P_n(y)r^{2^m E_n}|\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\
&\leq \int_{\mathbb{R}} \prod_{n=1}^{N} (1 - P_n(y) + P_n(y)) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = 1.
\end{aligned}$$

Finally,

$$|g_m(u)| \leq \frac{\left(1 + r^{2^m}\right)(1+r) + \left(1 + r^{2^m}\right)r}{2^{m/2}(1-r)^2} = \frac{\left(1 + r^{2^m}\right)(1+2r)}{2^{m/2}(1-r)^2}.$$

Now, if we define $q_{m,k}(u) \equiv g_m(u)\cos(ku)$ and $E^{m,k}_{T(h)} = I(r,k) - I(r,k;h)$, then,

$$\left|E^{m,k}_{T(h)}\right| = \frac{\pi^3}{12m_T^2}\left|q''_{m,k}(\mu)\right|, \ \mu \in (0,\pi).$$

Taking into account that $q''_{m,k}(u) = -k^2 g_m(u)\cos(ku) - 2kg'_m(u)\sin(ku) + g''_m(u)\cos(ku)$, then,

$$\left|E^{m,k}_{T(h)}\right| \leq \frac{\pi^3}{12m_T^2}\left(\frac{\left(1 + r^{2^m}\right)(1+2r)}{2^{m/2}(1-r)^2}k^2 + 2|g'_m(\mu)|k + |g''_m(\mu)|\right).$$

Let us assume that $|g'_m(u)| \leq B_1(r)$ and $|g''_m(u)| \leq B_2(r)$ for all $u \in [0,\pi]$ such that the limits $\lim_{r\to 0} B_1(r)$ and $\lim_{r\to 0} B_2(r)$ exist. Then,

$$\left|E^{m,k}_{T(h)}\right| \leq \frac{\pi^3}{12m_T^2}\left(\frac{\left(1 + r^{2^m}\right)(1+2r)}{2^{m/2}(1-r)^2}k^2 + 2B_1(r)k + B_2(r)\right).$$

We recall that the parameter k ranges from 1 to $2^m - 1$, where m is the scale of the approximation. For small values of k, the discretization error can be very small taking the number of subintervals m_T suitably large. For large values of k and assuming r small, the expression above is approximately $\pi^3 k^2 / 12 m_T^2 2^{m/2}$. At this point, it is important to underline that the calculation of $\Re(Q_m(re^{iu_0}))$ for a given $u_0 \in [0, \pi]$ is computationally intensive, particularly for very big portfolios. So we must control the discretization error choosing m_T as small as possible. As we show in the numerical examples section, taking $m_T = 2^m$ we get convergence with MC results at high loss levels.

As mentioned before, the round-off error arises when multiplying the sum in expression (3.21) by the pre-factor $(m_T r^k)^{-1}$. With $m_T = 2^m$ the k of interest is $k = 2^m - 1$ which is the greatest value that this parameter can take (small values of k do not cause round-off error).

The left plot of Figure 3.1 represents the pre-factor for values of $r \geq 0.9$, while the right plot shows the pre-factor values for $r \geq 0.999$. We also display in Table 3.1 the pre-factor values $(m_T r^k)^{-1}$ for different values of r and scales $m = 8$, $m = 9$ and $m = 10$.

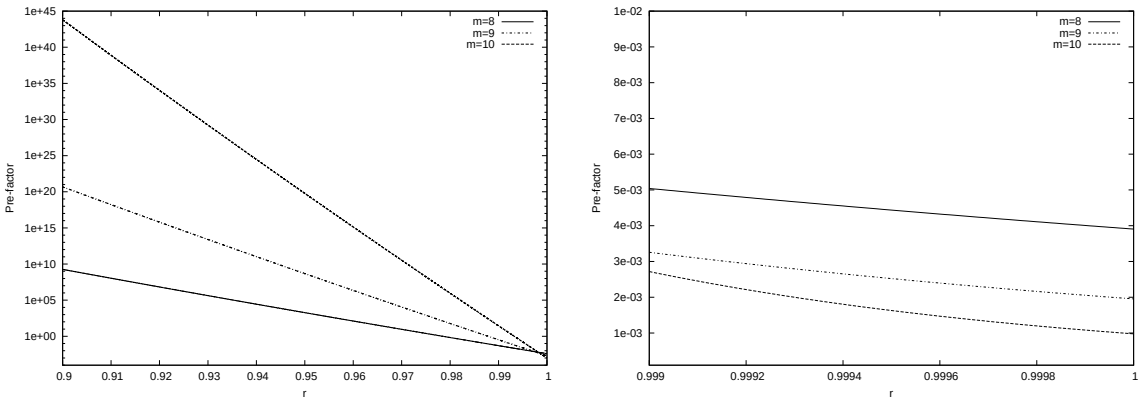

Figure 3.1: Pre-factor $(m_T r^k)^{-1}$ for $m_T = 2^m$, $k = 2^m - 1$ and scales $m = 8$, $m = 9$ and $m = 10$.

On the one hand, it is clear that we must concentrate in this second interval in order to get a reasonable round-off error and on the other hand, the discretization error grows when r is close to 1. For the numerical examples section, we have chosen $r = 0.9995$, the middle point of the interval $[0.999, 1)$.

Proceeding this way we have to find an expression for the computation of $\Re(\tilde{M}_{\mathcal{L}}(-2^m \ln(re^{iu})))$ and $\Im(\tilde{M}_{\mathcal{L}}(-2^m \ln(re^{iu})))$. Introducing the change of variable $y = \sqrt{2}x$ in (3.10) we obtain,

$$\tilde{M}_{\mathcal{L}}(s) = \int_{\mathbb{R}} \prod_{n=1}^{N} \left[1 - P_n(\sqrt{2}x) + P_n(\sqrt{2}x) e^{-sE_n}\right] \frac{1}{\sqrt{\pi}} e^{-x^2} dx.$$

We can approximate this integral by means of a Gauss-Hermite formula,

$$\tilde{M}_{\mathcal{L}}(s) \simeq \sum_{j=1}^{l/2} a_j \left(\hat{M}_{\mathcal{L}}(s; x_j^-) + \hat{M}_{\mathcal{L}}(s; x_j^+)\right), \tag{3.25}$$

	Scale		
r	$m = 8$	$m = 9$	$m = 10$
0.9000	$1.8194 \cdot 10^{9}$	$4.7077 \cdot 10^{20}$	$6.3040 \cdot 10^{43}$
0.9100	$1.0869 \cdot 10^{8}$	$1.6618 \cdot 10^{18}$	$7.7688 \cdot 10^{38}$
0.9200	$6.6968 \cdot 10^{6}$	$6.2395 \cdot 10^{15}$	$1.0833 \cdot 10^{34}$
0.9300	$4.2522 \cdot 10^{5}$	$2.4885 \cdot 10^{13}$	$1.7047 \cdot 10^{29}$
0.9400	$2.7807 \cdot 10^{4}$	$1.0529 \cdot 10^{11}$	$3.0193 \cdot 10^{24}$
0.9500	$1.8717 \cdot 10^{3}$	$4.7203 \cdot 10^{8}$	$6.0042 \cdot 10^{19}$
0.9600	$1.2960 \cdot 10^{2}$	$2.2394 \cdot 10^{6}$	$1.3373 \cdot 10^{15}$
0.9700	$9.2550 \cdot 10^{0}$	$1.1230 \cdot 10^{4}$	$3.3283 \cdot 10^{10}$
0.9800	$6.7470 \cdot 10^{-1}$	$5.9457 \cdot 10^{1}$	$9.2347 \cdot 10^{5}$
0.9900	$5.0674 \cdot 10^{-2}$	$3.3201 \cdot 10^{-1}$	$2.8503 \cdot 10^{1}$
0.9990	$5.0415 \cdot 10^{-3}$	$3.2566 \cdot 10^{-3}$	$2.7177 \cdot 10^{-3}$
0.9991	$4.9145 \cdot 10^{-3}$	$3.0942 \cdot 10^{-3}$	$2.4532 \cdot 10^{-3}$
0.9993	$4.6699 \cdot 10^{-3}$	$2.7934 \cdot 10^{-3}$	$1.9990 \cdot 10^{-3}$
0.9995	$4.4376 \cdot 10^{-3}$	$2.5219 \cdot 10^{-3}$	$1.6289 \cdot 10^{-3}$
0.9997	$4.2169 \cdot 10^{-3}$	$2.2768 \cdot 10^{-3}$	$1.3274 \cdot 10^{-3}$
0.9999	$4.0072 \cdot 10^{-3}$	$2.0555 \cdot 10^{-3}$	$1.0818 \cdot 10^{-3}$

Table 3.1: Pre-factor $(m_T r^k)^{-1}$ for $m_T = 2^m$, $k = 2^m - 1$ and scales $m = 8$, $m = 9$ and $m = 10$.

where $\hat{M}_{\mathcal{L}}(s;x) = \frac{1}{\sqrt{\pi}} \prod_{n=1}^{N} \left[1 - P_n(\sqrt{2}x) + P_n(\sqrt{2}x)e^{-sE_n}\right]$, $x_j^- = -x_j^+$ and a_j are the nodes and weights of the quadrature respectively. This is, we can compute,

$$\Re(\tilde{M}_{\mathcal{L}}(-2^m \ln(re^{iu}))) \simeq \sum_{j=1}^{l/2} a_j \left(\Re(\hat{M}_{\mathcal{L}}(-2^m \ln(re^{iu}); x_j^-)) + \Re(\hat{M}_{\mathcal{L}}(-2^m \ln(re^{iu}); x_j^+))\right), \tag{3.26}$$

and,

$$\Im(\tilde{M}_{\mathcal{L}}(-2^m \ln(re^{iu}))) \simeq \sum_{j=1}^{l/2} a_j \left(\Im(\hat{M}_{\mathcal{L}}(-2^m \ln(re^{iu}); x_j^-)) + \Im(\hat{M}_{\mathcal{L}}(-2^m \ln(re^{iu}); x_j^+))\right). \tag{3.27}$$

We notice that,

$$\hat{M}_{\mathcal{L}}(-2^m \ln(re^{iu}); x) = \frac{1}{\sqrt{\pi}} \prod_{n=1}^{N} \left[1 - P_n(\sqrt{2}x) + P_n(\sqrt{2}x) r^{2^m E_n} (\cos(2^m E_n u) + i \sin(2^m E_n u))\right].$$

So using polar coordinates this expression casts into,

$$\hat{M}_{\mathcal{L}}(-2^m \ln(re^{iu}); x) = \frac{1}{\sqrt{\pi}} \prod_{n=1}^{N} (R_n)_{\theta_n} = \frac{1}{\sqrt{\pi}} \left(\prod_{n=1}^{N} R_n\right)_{\sum_{n=1}^{N} \theta_n},$$

where $R_n = |z_n|$, $\theta_n = \arctan\left(\frac{\Im(z_n)}{\Re(z_n)}\right)$ and,

$$z_n = 1 - P_n(\sqrt{2}x) + P_n(\sqrt{2}x) r^{2^m E_n} (\cos(2^m E_n u) + i \sin(2^m E_n u)).$$

Finally, expanding these expressions we conclude,

$$R_n = \sqrt{(1-P_n(\sqrt{2}x))^2 + r^{2^{m+1}E_n}P_n^2(\sqrt{2}x) + 2P_n(\sqrt{2}x)(1-P_n(\sqrt{2}x))r^{2^m E_n}\cos(2^m E_n u)}, \tag{3.28}$$

$$\theta_n = \arctan\left(\frac{r^{2^m E_n}P_n(\sqrt{2}x)\sin(2^m E_n u)}{1-P_n(\sqrt{2}x)+P_n(\sqrt{2}x)r^{2^m E_n}\cos(2^m E_n u)}\right). \tag{3.29}$$

$$\Re(\hat{M}_{\mathcal{L}}(-2^m\ln(re^{iu});x)) = \frac{1}{\sqrt{\pi}}\left(\prod_{n=1}^{N}R_n\right)\cos\left(\sum_{n=1}^{N}\theta_n\right), \tag{3.30}$$

$$\Im(\hat{M}_{\mathcal{L}}(-2^m\ln(re^{iu});x)) = \frac{1}{\sqrt{\pi}}\left(\prod_{n=1}^{N}R_n\right)\sin\left(\sum_{n=1}^{N}\theta_n\right). \tag{3.31}$$

3.3.2 The Computation of VaR

Taking into account (2.2) it can be easily proved that,

$$0 \leq c_{m,k} \leq 2^{-\frac{m}{2}}, \quad k = 0, 1, ..., 2^m - 1,$$

and,

$$0 \leq c_{m,0} \leq c_{m,1} \leq ... \leq c_{m,2^m-1}.$$

Considering an approximation in a level of resolution m, VaR can now be quickly computed with m coefficients due to the compact support of the basis functions. Observe that due to the approximation (3.11) we have,

$$\overline{F}(\mathrm{VaR}_\alpha) \simeq \overline{F}_m(\mathrm{VaR}_\alpha) = 2^{\frac{m}{2}}\cdot c_{m,\overline{k}},$$

for a certain $\overline{k} \in \{0, 1, ..., 2^m - 1\}$. Thus, we can simply start searching VaR_α by means of the following simple iterative procedure: first we compute $\overline{F}_m(\frac{2^{m-1}}{2^m})$. If $\overline{F}_m(\frac{2^{m-1}}{2^m}) > \alpha$ then we compute $\overline{F}_m(\frac{2^{m-1}-2^{m-2}}{2^m})$, otherwise we compute $\overline{F}_m(\frac{2^{m-1}+2^{m-2}}{2^m})$, and so on. This algorithm finishes after m steps storing the $\overline{k}$ value such that $\overline{F}_m(\frac{\overline{k}}{2^m})$ is the closest value to α in our m resolution approximation. It is worth to highlight that there is almost any difference between calculating m coefficients and the whole distribution, i.e. 2^m coefficients.

In fact, due to the stepped shape of the Haar wavelets approximation, $\overline{F}_m(\xi) = \overline{F}_m(\frac{\overline{k}}{2^m})$, for all $\xi \in \left[\frac{\overline{k}}{2^m}, \frac{\overline{k}+1}{2^m}\right)$. In what follows let us take, $\mathrm{VaR}_\alpha^{W(m)} = \frac{2\overline{k}+1}{2^{m+1}}$, the middle point of this interval, as the VaR value computed by means of this wavelet algorithm at scale m.

Let us also consider the relative error at scale m defined by,

$$\mathrm{RE}(\alpha, m) = \frac{\mathrm{VaR}_\alpha^{W(m)} - \mathrm{VaR}_\alpha}{\mathrm{VaR}_\alpha}.$$

Assuming that VaR_α is well approximated by a value, VaR_α^M, obtained by means of a Monte Carlo method that will be taken as benchmark, as is common in this kind of studies, we will use the estimation,

$$\mathrm{RE}(\alpha, m) \simeq \overline{\mathrm{RE}}(\alpha, m), \quad \text{where} \quad \overline{\mathrm{RE}}(\alpha, m) = \frac{\mathrm{VaR}_\alpha^{W(m)} - \mathrm{VaR}_\alpha^M}{\mathrm{VaR}_\alpha^M}.$$

Observe that we consider the relative error without applying absolute value in order to maintain the economical sense, since a negative $\overline{\mathrm{RE}}$ indicates an underestimation of the risk measure, while a positive value results in a over-estimation of the risk.

3.4 Numerical Examples and Discussions

In this section we present a comparative study to calculate VaR using the Wavelet Approximation, the Saddle Point, the ASRF and the Monte Carlo method. As it is well known, MC has a strong dependence between the size of the portfolio and the computational time. When the size increases, MC becomes a very time consuming method.

Real situations in financial companies show the existence of strong concentrations in their credit portfolios, while Basel II formulae to calculate VaR are supported under unrealistic hypothesis, such as infinite number of obligors with small exposures. For these reasons, we also test our methodology with small and concentrated portfolios.

The Saddle Point method presented in Section 3.2.4 is known to be an accurate and fast approximation for the tail of the loss distribution. Under the conditional framework, the approach works out well, however, this method can fail if we introduce severe name concentration in the credit portfolio. To show this fact, we present the following two examples,

Portfolio 3.1. *We consider* $N = 100$ *obligors, with* $P_n = 1\%$, $E_n = 1/N$, $n = 1, \ldots, N$, $\rho = 0.2$ *and we assume constant loss given default* 100%.

Portfolio 3.2. *We consider* $N = 102$ *obligors, with* $P_n = 0.1\%$, $E_n = 1, n = 1, \ldots, 100$, $E_{101} = E_{102} = 20$, $\rho = 0.3$ *and we assume constant loss given default* 100%.

In both cases we normalize the exposures in order to meet the condition $\sum_{n=1}^N E_n = 1$ and we consider the Luganani and Rice formula (3.8) to calculate the conditional tail of the loss function $\mathbb{P}(\mathcal{L} > l|Y = y)$. Finally, we integrate over the risk factor Y, using a Gauss-Legendre quadrature of 100 points, to obtain the unconditional tail of the loss function $\mathbb{P}(\mathcal{L} > l)$.

Note that Portfolio 3.1 is totally homogeneous and so the saddle points can be calculated very quickly without the need of applying a numerical root finding algorithm. Considering $\bar{K}'_{\mathcal{L}}(s; y)$ in expression (3.9) and assuming that $E_n = E$ and $P_n(y) = P(y)$ for all $n = 1, \ldots, N$ then,

$$\begin{aligned}\bar{K}'_{\mathcal{L}}(s; y) &= \sum_{n=1}^N \frac{w_n P_n(y) e^{w_n s}}{1 - P_n(y) + P_n(y) e^{w_n s}} = \sum_{n=1}^N \frac{E \cdot P(y) e^{Es}}{1 - P(y) + P(y) e^{Es}} \\ &= \frac{N \cdot E \cdot P(y) e^{Es}}{1 - P(y) + P(y) e^{Es}} = \frac{P(y) e^{Es}}{1 - P(y) + P(y) e^{Es}},\end{aligned}$$

since $E = 1/N$. Finally,

$$\frac{P(y)e^{Es}}{1 - P(y) + P(y)e^{Es}} = l,$$

yields a unique solution for the saddle point,

$$\hat{l} = N \cdot \ln\left(\frac{l(1 - P(y))}{P(y)(1 - l)}\right).$$

The left plot of Figure 3.2 shows the tail of the loss distribution for the first portfolio, while the right plot represents the loss distribution for the second one. We have also represented a Monte Carlo approximation using five million scenarios which serve us as a benchmark, and finally the Asymptotic Single Risk Factor model. The ASRF model underestimates the risk in both portfolios. Although the first one is totally homogeneous, we recall that this method works out well for very big portfolios. The Saddle Point method gives good approximations in the first case, however it can not deal with the strong exposure concentration of the second portfolio where almost the 30% of the exposure belongs to only two obligors. Consequently it underestimates the risk at 99.9% loss level, while overestimates the risk at 99.99% loss level.

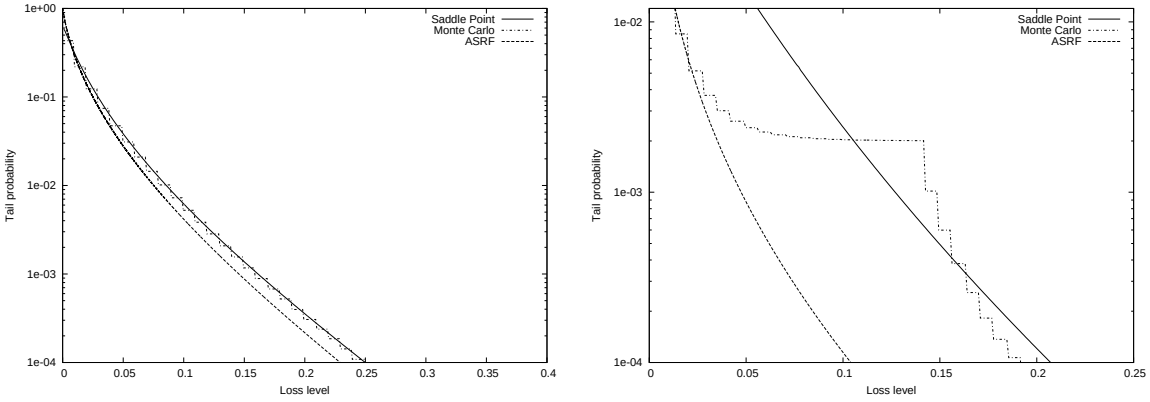

Figure 3.2: Tail probability approximation of a homogeneous portfolio (left) and a heterogeneous portfolio (right).

For comparison in Table 3.2 we give the VaR values computed for Portfolio 3.2 using the Saddle Point method, the Monte Carlo method and the ASRF model at loss level 99.9%. We also provide the same result with the Wavelet Approximation method at scale $m = 10$, where we take $l = 20$ nodes of Gauss-Hermite quadrature to evaluate expression (3.25) and $m_T = 2^m$ subintervals for the trapezoidal rule to compute the coefficients in (3.21).

We see that the Wavelet Approximation method is capable to deal with severe exposure concentration at a high loss level. Furthermore, the WA is fast, robust and accurate as we show with the following examples. We consider six portfolios ranging from 10 to 10000 obligors as described in Table 3.3.

Method	$\text{VaR}_{0.999}$	Relative Error
Monte Carlo	0.1500	
ASRF	0.0474	-68.39%
Saddle Point	0.1270	-15.37%
Wavelet Approximation	0.1490	-0.69%

Table 3.2: Computation of VaR for Portfolio 3.2 with severe exposure concentration. Relative errors are computed with respect to Monte Carlo results.

Portfolio	N	P_n	E_n	ρ	HHI	$\frac{1}{N}$
P1	100	0.21%	$\frac{C}{n}$	0.15	0.0608	0.0100
P2	1000	1.00%	$\frac{C}{n}$	0.15	0.0293	0.0010
P3	1000	0.30%	$\frac{C}{n}$	0.15	0.0293	0.0010
P4	10000	1.00%	$\frac{C}{n}$	0.15	0.0172	0.0001
P5	20	1.00%	$\frac{1}{N}$	0.5	0.0500	0.0500
P6	10	0.21%	$\frac{C}{n}$	0.5	0.1806	0.1000

Table 3.3: Portfolios selected for the numerical examples. In each case, C is a constant such that $\sum_{n=1}^{N} E_n = 1$.

In order to consider concentrated portfolios, we have taken $E_n = \frac{C}{n}$ (where C is a constant such that $\sum_{n=1}^{N} E_n = 1$), except for Portfolio P5 which is totally diversified. We provide also the Herfindahl-Hirschman index discussed in Section 1.5 for quantifying exposure concentration. This index can take values from $\frac{1}{N}$ to 1 (this latter value corresponds to a portfolio with only 1 obligor). Well-diversified portfolios with a very large number of very small exposures have a HHI value close to $\frac{1}{N}$, whereas heavily concentrated portfolios can have a considerably higher HHI value. We note that P5 is a small and completely diversified portfolio while P6 is a small but strongly concentrated one. The correlation parameter ρ, which measures the degree of the obligor's exposure to the systematic risk factor, and the probabilities of default P_n have been taken as representative examples. The potential loss increases when considering higher probabilities of default. This fact will be shown when comparing the VaR value of Portfolios P2 and P3 where the remaining parameters are unchanged.

The main numerical results are displayed in Table 3.4. We have computed[1] the VaR value at 99.9% confidence level with the WA method at scales 8, 9 and 10 and also the VaR value using the MC method with 5×10^6 random scenarios, which serve us as a benchmark. For WA we take again $l = 20$ nodes of Gauss-Hermite quadrature and $m_T = 2^m$ subintervals for the trapezoidal rule to compute the coefficients in (3.21). Plots at scales 9 and 10 corresponding to Portfolios P1, P2, P3 and P4 can be seen in Figures 3.3,

[1]Computations have been carried out sequentially in a personal computer Dell Vostro 320 under GNU/Linux OS, Intel CPU Core 2 E7500, 2.93GHz, 4GB RAM and using the gcc compiler with optimization level 2.

3.4, 3.5 and 3.6 respectively.

Portfolio	$\mathrm{VaR}^{W(8)}_{0.999}$	$\overline{\mathrm{RE}}(0.999,8)$	$\mathrm{VaR}^{W(9)}_{0.999}$	$\overline{\mathrm{RE}}(0.999,9)$	$\mathrm{VaR}^{W(10)}_{0.999}$	$\overline{\mathrm{RE}}(0.999,10)$	$\mathrm{VaR}^{M}_{0.999}$
P1	0.1934	−0.19%	0.1963	1.32%	0.1938	0.06%	0.1937
P2	0.1934	1.01%	0.1924	0.50%	0.1919	0.25%	0.1914
P3	0.1426	1.46%	0.1416	0.77%	0.1411	0.42%	0.1405
P4	0.1621	0.24%	0.1611	−0.36%	0.1616	−0.06%	0.1617

Table 3.4: Results of 99.9% VaR computation using the Wavelet Approximation at scales 8, 9 and 10 and the Monte Carlo simulations with 5×10^6 random scenarios. $\overline{\mathrm{RE}}$ estimations are also provided.

In Table 3.5 we provide as well the computational time in seconds in both for MC method and WA method at scales 8, 9 and 10. WA at scale 10 give us very accurate results in a short computational time when compared with Monte Carlo. $\overline{\mathrm{RE}}$ estimations for Portfolios P1, P2 and P4 at scale 8 are less or equal than 1%, displaying already very accurate and fast computed approximations (in particular for P2 and P4 which are big portfolios), while Portfolios P2, P3 and P4 are very well approximated at scale 9, with computational time needs reducing to 3.6, 3.6 and 36.1 seconds respectively.

In order to assess the robustness of the WA method we increased up to 100 the number of nodes in the Gauss-Hermite quadrature. With this setting we computed VaR values at 99.9%, 99.99% and 99.999% confidence levels at scale 10 for Portfolios P1, P2, P3 and P4. The results obtained in all cases were coincident with the errors with respect to MC using only 20 nodes.

Portfolio	$\mathrm{VaR}^{W(8)}_{0.999}$	$\mathrm{VaR}^{W(9)}_{0.999}$	$\mathrm{VaR}^{W(10)}_{0.999}$	$\mathrm{VaR}^{M}_{0.999}$
P1	0.2	0.4	0.7	58.3
P2	1.8	3.6	7.2	571.6
P3	1.8	3.6	7.2	567.6
P4	18.2	36.1	72.4	1379.1

Table 3.5: Computational time (in seconds) for Portfolios P1, P2, P3 and P4 needed to compute the VaR value at 99.9% confidence level.

We have also assessed the robustness of the WA method doubling the number of subintervals used in the trapezoidal rule of integration, to compute the wavelet coefficients at scale 10 for Portfolios P1, P2, P3 and P4. Numerical results are presented in Table 3.6. We get the same VaR values using $m_T = 2^m$ or $m_T = 2^{m+1}$ subintervals.

We have shown the suitability of the WA method to deal with concentration in portfolios of considerable size like P1, P2, P3 and P4. It is remarkable how the Haar wavelets are naturally capable of detecting jumps in the cumulative distribution function, making the approximation very precise both at small and high loss levels. To demonstrate that this interesting property still remains in small portfolios, we consider P5 and P6 plotted in figure 3.7, both at scale 10. Since P5 and P6 are very small portfolios, the non-smooth features appear accentuated. However the MC and WA plots are indistinguishable one

from the other, showing again the fast convergence of the WA method towards MC. The ASRF model systematically underestimates the risk in all these sample portfolios.

2^{10}				
Portfolio	$\mathrm{VaR}^{W(10)}_{0.9999}$	$\overline{\mathrm{RE}}(0.9999, 10)$	$\mathrm{VaR}^{W(10)}_{0.99999}$	$\overline{\mathrm{RE}}(0.99999, 10)$
P1	0.2251	-0.07%	0.2935	-1.70%
P2	0.2622	-0.46%	0.3325	-1.80%
P3	0.1812	-0.10%	0.2290	-1.88%
P4	0.2261	-0.25%	0.2935	-1.30%
2^{11}				
Portfolio	$\mathrm{VaR}^{W(10)}_{0.9999}$	$\overline{\mathrm{RE}}(0.9999, 10)$	$\mathrm{VaR}^{W(10)}_{0.99999}$	$\overline{\mathrm{RE}}(0.99999, 10)$
P1	0.2251	-0.07%	0.2935	-1.70%
P2	0.2622	-0.46%	0.3325	-1.80%
P3	0.1812	-0.10%	0.2290	-1.88%
P4	0.2261	-0.25%	0.2935	-1.30%
MC				
Portfolio	$\mathrm{VaR}^{M}_{0.9999}$		$\mathrm{VaR}^{M}_{0.99999}$	
P1	0.2253		0.2985	
P2	0.2634		0.3386	
P3	0.1813		0.2334	
P4	0.2267		0.2973	

Table 3.6: VaR values at 99.99% and 99.999% confidence levels for Portfolios P1, P2, P3 and P4 using 2^m and 2^{m+1} subintervals (with WA approximation at scale $m = 10$) for the trapezoidal rule. MC with 5×10^6 random scenarios and $\overline{\mathrm{RE}}$ estimations are also provided.

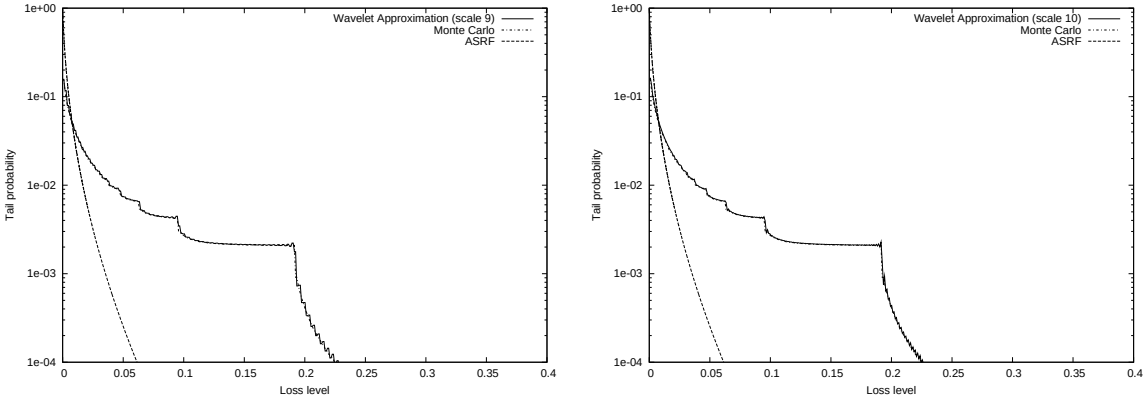

Figure 3.3: Tail probability approximation of Portfolio P1 at WA approximation scales $m = 9$ and $m = 10$.

In summary, we have presented a numerical approximation to the loss function based on Haar wavelets in order to avoid the computational effort of running Monte Carlo simulations. First of all we approximate the discontinuous distribution of the loss function by a finite summation of Haar scaling functions, and then we calculate the coefficients

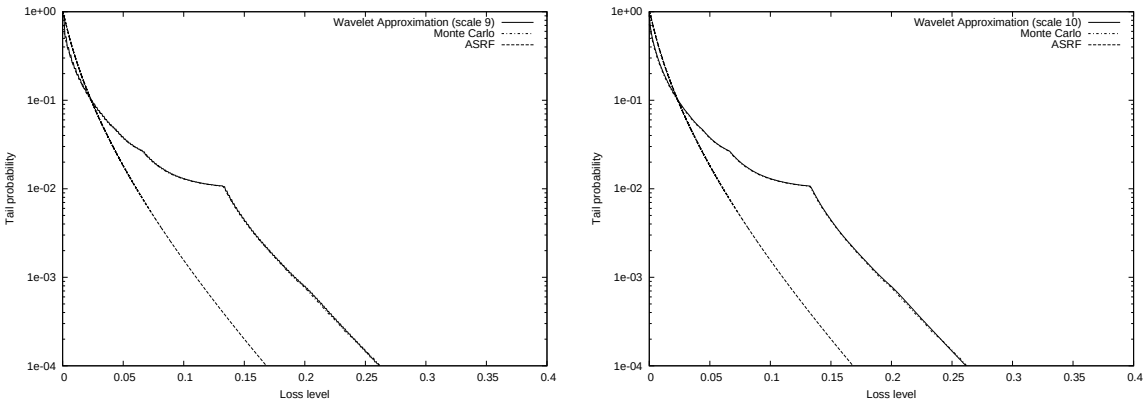

Figure 3.4: Tail probability approximation of Portfolio P2 at WA approximation scales $m = 9$ and $m = 10$.

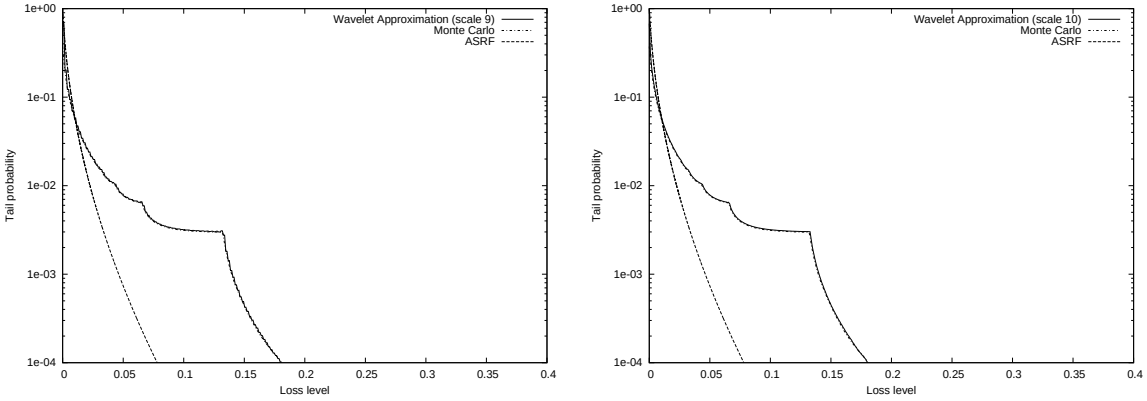

Figure 3.5: Tail probability approximation of Portfolio P3 at WA approximation scales $m = 9$ and $m = 10$.

of the approximation by inverting its Laplace transform. Due to the compact support property of the Haar system, only a few coefficients are needed for the VaR computation.

We have shown the performance of the numerical approximation in six sample portfolios. These results, among other simulations, show that the method is applicable and very accurate to different sized portfolios needing also of short time computations. Moreover, the Wavelet Approximation is robust since the method is very stable under changes in the parameters of the model. The stepped form of the approximated distribution makes the Haar wavelets natural and very suitable for the approximation.

We also remark that the algorithm is valid for continuous cumulative distribution functions, and that it can be used in other financial models without making conceptual changes in the development. For instance, we can easily introduce stochastic loss given default (just changing a bit the unconditional moment generating function) and to consider the multi-factor Merton model as the model framework as well.

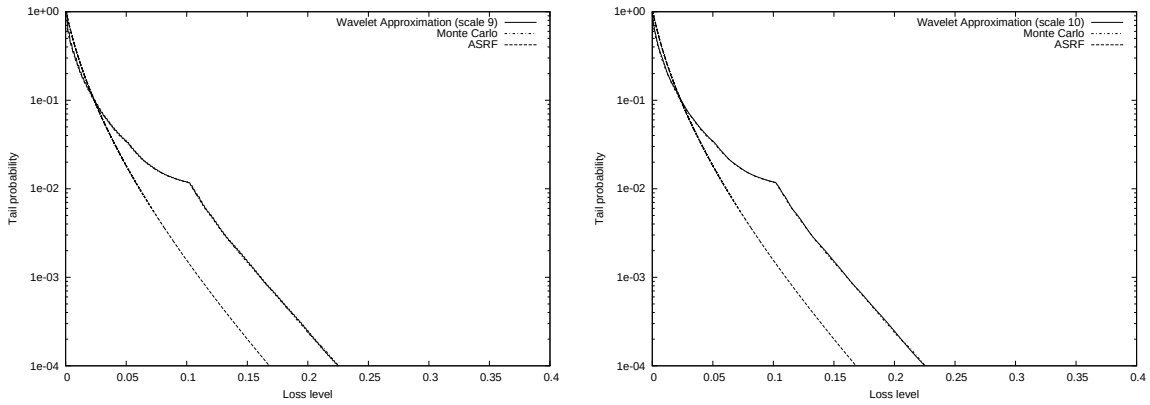

Figure 3.6: Tail probability approximation of Portfolio P4 at WA approximation scales $m = 9$ and $m = 10$.

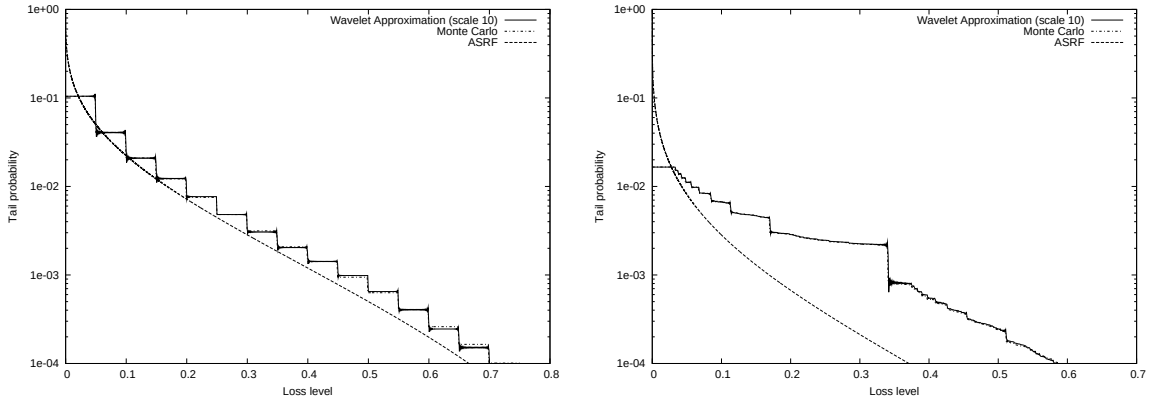

Figure 3.7: Tail probability approximation of Portfolio P5 (left) and P6 (right) at scale $m = 10$.

We have also compared the Wavelet Approximation with the Saddle Point method for a diversified portfolio and a concentrated portfolio. The SP method is fast and accurate to calculate the VaR in the first portfolio, while clearly underestimate the risk in the concentrated one. It has been also shown that the ASRF model, which underpins the Basel II formulae, is not capable to deal with realistic portfolios which often entail name concentration.

Chapter 4

Credit Risk Contributions with the Wavelet Approximation Method

4.1 Introduction

This Chapter is devoted to the computation of the ES as an alternative measure to VaR, as well as to the derivation and accurately calculation of the marginal contributions to the VaR and the ES under the one-factor Merton model framework. This work has been recently developed by [Ort11] and submitted to the *Journal of Computational Finance*.

Both VaR and ES can be decomposed as a sum of sensitivities (see [Tas00a]). These sensitivities, which are commonly named risk contributions, can be understood as the marginal impact on the risk of the total portfolio and are very important for loan pricing or asset allocation, to cite two examples.

In practice, each risk contribution is usually computed by means of MC calculated as the expected value of the loss distribution conditioned on a rare event, the VaR value, which represents an extreme loss for the credit portfolio. The usual Plain Monte Carlo presents practical inconveniences due to the large number of simulations required to get the rare events. Although in this context of MC simulation, [Gla05] develops efficient methods based on importance sampling to calculate VaR and ES contributions in the multi-factor Merton model, the computational effort is still very important. For this reason, analytical or fast numerical techniques are always welcome. One of such analytical techniques for VaRC computations is the Saddle Point method pioneered by Martin *et al.* ([Mar01a] and [Mar01b]), which has been introduced in Section 3.2.4 to calculate the VaR value. They apply the approximation to the unconditional Moment Generating Function (MGF) and obtain accurate results at very small tail probabilities. This method is known to perform well with big portfolios at high loss levels. [Hua07a] computes the risk measures and contributions implementing a higher order Saddle Point method in the Merton model and apply the approximation to the conditional MGF instead of the unconditional MGF, where the Saddle Point works better, with an extra computational time. [Hua07b] presents a comparative study for the calculation of VaR and VaRC with the SP method, MC with importance sampling (IS) and the normal approximation (NA) method. They conclude that there is not a perfect method that prevails among the others and the choice is a trade-

off between speed, accuracy and robustness. NA is an accurate method and the fastest one but is not capable of handling with exposure concentration. IS is the most robust method but is highly demanding from a computational point of view when estimating the VaRC. The SP method preserves a good balance between speed and accuracy and is better than normal approximation to deal with exposure concentration. However, if the loss distribution is not smooth due to exceptional exposure concentration, a straightforward implementation of SP may be insufficient, and an adaptive SP should be employed instead. Alternatively, [Tak08] addresses the problem of calculating the marginal contributions using a numerical Laplace transform inversion of the MGF in the multi-sector setting and provide precise results in big size portfolios. In this Chapter we extend the work undertaken with the estimation of the VaR value with the WA method in Chapter 3 and we develop a new methodology for the computation of the ES, the VaR contributions and the ES contributions. We recall that this methodology approximate the credit loss cumulative distribution function (CDF) by a finite combination of Haar wavelets basis functions in order to invert the Laplace transform of the unconditional MFG. It was tested under the one-factor Merton model, showing accurate and fast results for a wide range of portfolios at very high loss levels. We will show that WA can get very accurate results even in presence of extremely exposure concentration when computing the risk measures and contributions, where a straightforward implementation of SP would fail. The key point for the calculation of the VaR, ES, VaRC and ESC is how to evaluate the coefficients of the wavelet expansion and their derivatives with respect to exposures. We have explained in detail the calculation of the coefficients in Chapter 3. Now, we get the same accuracy with fewer nodes by means of truncating the integration variable in the Gauss-Hermite quadrature representing the business cycle, obtaining also a proportional reduction in the computational time. We point out that although we apply the WA method at scale ten in order to maintain accurate results in all the sample portfolios, the WA performs very well with smaller scales with a considerable reduction in the computational effort (essentially the CPU time is divided by two when moving from scale m to $m-1$).

In the following we consider, as in Chapter 3, a portfolio $\mathcal{L} = \sum_{n=1}^{N} \mathcal{L}_n$, with N obligors where $\mathcal{L}_n = E_n \cdot L_n \cdot D_n$. Without loss of generality, we can assume $\sum_{n=1}^{N} E_n = 1$ and constant loss given default $L_n = 100\%$. To describe the obligor's default and its correlation structure, we follow the one-factor Merton model with constant correlation ρ,

$$r_n = \sqrt{\rho} Y + \sqrt{1-\rho}\epsilon_n,$$

where Y and ϵ_n, $\forall n \leq N$ are i.i.d. standard normally distributed. Hence, the default probability of obligor n conditional on a specification $Y = y$ can be written as,

$$P_n(y) \equiv \Phi\left(\frac{t_n - \sqrt{\rho}y}{\sqrt{1-\rho}}\right),$$

where $t_n = \Phi^{-1}(P_n)$, and Φ denotes the cummulative standard normal distribution function and Φ^{-1} its inverse.

4.2 Haar Wavelets Approach

Let F be the cumulative distribution function of $\mathcal{L}$. It is crucial for risk measurement and risk management to allocate the risk to individual obligors. So, after we have selected the risk measure and following (1.2) or (1.3) we must consider the partial derivative of the measure with respect to the exposures E_n. Then, for a given $E = (E_1, \ldots, E_N)$ let us consider,

$$F(E,x) = \begin{cases} \overline{F}(E,x), & \text{if } 0 \leqslant x \leqslant 1, \\ 1, & \text{if } x > 1, \end{cases}$$

for a certain $\overline{F}$ defined in $[0,1]$. Following the work in Chapter 3, we approximate the distribution function by a finite sum of Haar wavelets basis functions with convergence in $L^2([0,1])$,

$$\overline{F}(E,x) \simeq \overline{F}_m(E,x), \quad \overline{F}_m(E,x) = \sum_{k=0}^{2^m-1} c_{m,k}(E)\phi_{m,k}(x), \tag{4.1}$$

where m is the scale of the approximation and k is the translation parameter.

The unconditional moment generating function,

$$\tilde{M}_{\mathcal{L}}(E,s) \equiv \mathbb{E}(e^{-s\mathcal{L}}) = \mathbb{E}(\mathbb{E}(e^{-s\mathcal{L}} \mid Y)) = \int_{\mathbb{R}} \prod_{n=1}^{N} \left[1 - P_n(y) + P_n(y)e^{-sE_n}\right] \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy, \tag{4.2}$$

is also the Laplace transform of the density function $f_{\mathcal{L}}$ of $\mathcal{L}$,

$$\tilde{M}_{\mathcal{L}}(E,s) \equiv \mathbb{E}(e^{-s\mathcal{L}}) = \int_0^{+\infty} e^{-sx} f_{\mathcal{L}}(E,x)dx. \tag{4.3}$$

Integrating by parts the expression (4.3) and using the approximation in (4.1), the coefficients $c_{m,k}$ are recovered by the Laplace transform inversion, giving us the expressions,

$$c_{m,0} \simeq \frac{\int_{\mathbb{R}} \prod_{n=1}^{N} [1 - P_n(y)] \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy}{2^{\frac{m}{2}}},$$

and,

$$c_{m,k}(E) \simeq \frac{2}{\pi r^k} \int_0^{\pi} \Re(Q_m(E, re^{iu})) \cos(ku)du, \quad k = 1, ..., 2^m - 1, \tag{4.4}$$

where,

$$Q_m(E,z) \equiv \sum_{k=0}^{2^m-1} c_{m,k}(E)z^k \simeq \frac{\tilde{M}_{\mathcal{L}}(E, -2^m \ln(z)) - z^{2^m}}{2^{\frac{m}{2}}(1-z)}. \tag{4.5}$$

In Chapter 3 the coefficients $c_{m,k}$ are accurately computed by means of the ordinary trapezoidal rule and the MGF is evaluated using Gauss-Hermite formulae with 20 nodes. The results show a fast convergence towards the value obtained by means of a Monte Carlo method with five million sampling scenarios, which serves as a benchmark. However, the

speed of the algorithm is highly conditioned by the number of times that we have to evaluate the MGF, i.e., the number of subintervals in the trapezoidal rule, the number of nodes used in Gauss-Hermite integration and the size of the portfolio. It has been proved that 2^m subintervals are enough to integrate (4.4) using the trapezoidal rule, getting accurate VaR values at 99.9%, 99.99% and 99.999% confidence levels. Later on, we present an improved version of the algorithm which again reduces significantly the computational cost when a large number of nodes of GH are required, or when the portfolio contains a big number of obligors.

4.3 Credit Risk Measures and Contributions

In this section we recall the calculation of the VaR value and present a numerical formula for the estimation of the Expected Shortfall using the Wavelet Approximation method presented previously. We also apply this method to compute the marginal contribution of the obligors to the total risk at portfolio level.

4.3.1 VaR and Expected Shortfall

Let us consider a portfolio with exposures $E = (E_1, \dots, E_N)$ and let $\alpha \in (0, 1)$ be a given confidence level. The α-quantile of the loss distribution of $\mathcal{L}$ in this context, is called Value at Risk. This is,

$$\mathrm{VaR}_\alpha(E) = \inf\{l \in \mathbb{R} : \mathbb{P}(\mathcal{L} \leq l) \geq \alpha\} = \inf\{l \in \mathbb{R} : F(E, l) \geq \alpha\},$$

where again we underline the dependence on the exposures. This is the measure chosen in the Basel II Accord for the computation of capital requirement, meaning that a bank, managing its risks according to Basel II, must reserve capital by an amount of $\mathrm{VaR}_\alpha(E)$ to cover extreme losses.

It has been extensively detailed in Chapter 3 that considering a wavelet approximation in a level of resolution m, the VaR value at confidence level α calculated by the WA method is of the form $\mathrm{VaR}_\alpha^{W(m)}(E) = \frac{2\overline{k}+1}{2^{m+1}}$ for a certain $\overline{k} \in \{0, 1, ..., 2^m - 1\}$, where $\overline{k}$ is such that $\overline{F}_m\left(E, \mathrm{VaR}_\alpha^{W(m)}(E)\right) \simeq \alpha$.

A crucial property for a coherent risk measure is the sub-additivity condition. As mentioned previously, the VaR measure fails to satisfy this condition although the measure is widely used in practice. This can be explained by the fact that when distributions are normal, or close to normal, it can be shown that VaR and ES are quite close and behave similarly. However, as soon as a distribution is characterized by a long tail behavior, the similarity between VaR and ES does not hold anymore. In this case, employing the VaR measure may lead to a considerable underestimation of risk.

By definition, the Expected Shortfall at confidence level α is given by,

$$\mathrm{ES}_\alpha(E) = \frac{1}{1-\alpha} \int_{VaR_\alpha(E)}^{+\infty} x f_{\mathcal{L}}(E, x) dx.$$

Then, integrating by parts and using the approximation in (4.1) we have,

$$\mathrm{ES}_\alpha(E) = \frac{1}{1-\alpha}\left(1 - \alpha\mathrm{VaR}_\alpha(E) - \int_{VaR_\alpha(E)}^{1} \overline{F}(E,x)dx\right) \simeq \mathrm{ES}_\alpha^{W(m)}(E), \tag{4.6}$$

where,

$$\mathrm{ES}_\alpha^{W(m)}(E) \equiv \frac{1}{1-\alpha}\left(1 - \alpha\mathrm{VaR}_\alpha^{W(m)}(E) - \frac{1}{2^{\frac{m}{2}+1}}c_{m,\overline{k}}(E) - \frac{1}{2^{\frac{m}{2}}}\sum_{k=\overline{k}+1}^{2^m-1} c_{m,k}(E)\right).$$

For sake of clarity, the explicit dependence of all ES and VaR expressions that appear in next sections with respect to E from now on will be dropped.

4.3.2 VaR Contributions and Expected Shortfall Contributions

Let us consider how to decompose the total risk into individual transactions. We carry out the allocation principle given by the partial derivative of the risk measure with respect to the exposure of an obligor. In this way following [Tas00a] we define the risk contribution of obligor i to the VaR value at confidence level α by,

$$\mathrm{VaRC}_{\alpha,i} \equiv E_i \cdot \frac{\partial \mathrm{VaR}_\alpha}{\partial E_i},$$

which satisfies the additivity condition, $\sum_{i=1}^{N} \mathrm{VaRC}_{\alpha,i} = \mathrm{VaR}_\alpha$.

Taking into account that $\overline{F}(E, \mathrm{VaR}_\alpha) = \alpha$ then,

$$\mathrm{VaRC}_{\alpha,i} \equiv E_i \cdot \frac{\partial \mathrm{VaR}_\alpha}{\partial E_i} = -E_i \cdot \frac{\frac{\partial \overline{F}(E,\mathrm{VaR}_\alpha)}{\partial E_i}}{\left.\frac{\overline{F}(E,x)}{\partial x}\right|_{x=\mathrm{VaR}_\alpha}} = -E_i \cdot \frac{\frac{\partial \overline{F}(E,\mathrm{VaR}_\alpha)}{\partial E_i}}{f_{\mathcal{L}}(E,\mathrm{VaR}_\alpha)}. \tag{4.7}$$

Now according to (4.1), $\overline{F}(E,x) \simeq \sum_{k=0}^{2^m-1} c_{m,k}(E)\phi_{m,k}(x)$. Taking partial derivatives with respect to E_i we get,

$$\frac{\partial \overline{F}(E,x)}{\partial E_i} \simeq \sum_{k=0}^{2^m-1} \frac{\partial c_{m,k}(E)}{\partial E_i}\phi_{m,k}(x), \tag{4.8}$$

and evaluating this expression in VaR_α and using the approximation to the VaR value given by $\mathrm{VaR}_\alpha^{W(m)}$, we have,

$$\frac{\partial \overline{F}(E,\mathrm{VaR}_\alpha)}{\partial E_i} \simeq \sum_{k=0}^{2^m-1} \frac{\partial c_{m,k}(E)}{\partial E_i}\phi_{m,k}(\mathrm{VaR}_\alpha) \simeq \sum_{k=0}^{2^m-1} \frac{\partial c_{m,k}(E)}{\partial E_i}\phi_{m,k}(\mathrm{VaR}_\alpha^{W(m)}) = 2^{\frac{m}{2}}\frac{\partial c_{m,\overline{k}}(E)}{\partial E_i}, \tag{4.9}$$

due to the fact that,

$$\phi_{m,k}(x) = \begin{cases} 2^{\frac{m}{2}}, & \text{if } \frac{k}{2^m} \le x < \frac{k+1}{2^m}, \\ 0, & \text{elsewhere.} \end{cases}$$

Finally, considering the expressions (4.7) and (4.9) we obtain,

$$\mathrm{VaRC}_{\alpha,i} \simeq \mathrm{VaRC}_{\alpha,i}^{W(m)}, \tag{4.10}$$

where $\mathrm{VaRC}_{\alpha,i}^{W(m)} \equiv \mathcal{C} \cdot E_i \cdot \frac{\partial c_{m,\overline{k}}(E)}{\partial E_i}$ and $\mathcal{C}$ is a constant such that $\sum_{i=1}^{N} \mathrm{VaRC}_{\alpha,i}^{W(m)} = \mathrm{VaR}_{\alpha}^{W(m)}$.

In a similar way we define the ESC for the obligor i at confidence level α as,

$$\mathrm{ESC}_{\alpha,i} \equiv E_i \cdot \frac{\partial \mathrm{ES}_{\alpha}}{\partial E_i},$$

satisfying the additivity condition $\sum_{i=1}^{N} \mathrm{ESC}_{\alpha,i} = \mathrm{ES}_{\alpha}$. For the computation of these expected shortfall contributions we have taken the derivative of the expected shortfall expression in (4.6) with respect to E_i and used the approximation (4.8). That is,

$$\begin{aligned}\mathrm{ESC}_{\alpha,i} \equiv E_i \cdot \frac{\partial \mathrm{ES}_{\alpha}}{\partial E_i} &= E_i \cdot \frac{1}{1-\alpha}\left(-\alpha \frac{\partial \mathrm{VaR}_{\alpha}}{\partial E_i} + \frac{\partial \mathrm{VaR}_{\alpha}}{\partial E_i}\overline{F}(E, \mathrm{VaR}_{\alpha}) - \int_{\mathrm{VaR}_{\alpha}}^{1} \frac{\partial \overline{F}(E,x)}{\partial E_i} dx\right) \\ &\simeq \mathrm{ESC}_{\alpha,i}^{W(m)},\end{aligned} \tag{4.11}$$

where,

$$\mathrm{ESC}_{\alpha,i}^{W(m)} \equiv -E_i \cdot \frac{1}{2^{\frac{m}{2}}} \cdot \frac{1}{1-\alpha} \cdot \left(\frac{1}{2}\frac{\partial c_{m,\overline{k}}(E)}{\partial E_i} + \sum_{k=\overline{k}+1}^{2^m-1} \frac{\partial c_{m,k}(E)}{\partial E_i}\right).$$

Later, in the numerical examples section, we will test the accuracy of the WA method for the calculation of VARC and ESC by means of Monte Carlo estimates, which will serve us as a benchmark. Under appropriate conditions the marginal VaR contribution at confidence level α of the obligor i is,

$$\mathrm{VaRC}_{\alpha,i} = \mathbb{E}(\mathcal{L}_i | \mathcal{L} = \mathrm{VaR}_{\alpha}), \tag{4.12}$$

and the marginal contribution at confidence level α to the expected shortfall is,

$$\mathrm{ESC}_{\alpha,i} = \mathbb{E}(\mathcal{L}_i | \mathcal{L} \geq \mathrm{VaR}_{\alpha}). \tag{4.13}$$

Thus, in both cases, the marginal risk contributions are conditional expectations of the individual loss random variables, conditioned on rare values of the portfolio loss $\mathcal{L}$. We note that the expressions (4.12) and (4.13) decompose the total risk,

$$\sum_{i=1}^{N} \mathbb{E}(\mathcal{L}_i | \mathcal{L} = \mathrm{VaR}_{\alpha}) = \mathbb{E}\left(\sum_{i=1}^{N} \mathcal{L}_i | \mathcal{L} = \mathrm{VaR}_{\alpha}\right) = \mathbb{E}(\mathcal{L} | \mathcal{L} = \mathrm{VaR}_{\alpha}) = \mathrm{VaR}_{\alpha},$$

and,

$$\sum_{i=1}^{N} \mathbb{E}(\mathcal{L}_i | \mathcal{L} \geq \mathrm{VaR}_{\alpha}) = \mathbb{E}\left(\sum_{i=1}^{N} \mathcal{L}_i | \mathcal{L} \geq \mathrm{VaR}_{\alpha}\right) = \mathbb{E}(\mathcal{L} | \mathcal{L} \geq \mathrm{VaR}_{\alpha}) = \mathrm{ES}_{\alpha}.$$

We estimate risk contributions by means of Monte Carlo simulations in two steps. First, we compute the VaR value through an ordinary Monte Carlo simulation and then we use the estimated VaR in (4.12) and (4.13). In this second step, for a given loss level l, we consider the problem of estimating $\mathbb{E}(\mathcal{L}_i|\mathcal{L} = l)$ and $\mathbb{E}(\mathcal{L}_i|\mathcal{L} \geq l)$. Both cases can be treated together considering $C_i = \mathbb{E}(\mathcal{L}_i|\mathcal{L} \in \mathcal{A})$ where $\mathcal{A} = l$ or $\mathcal{A} = [l, \infty)$. For each sampled scenario we proceed as follows,

1. Generate the systemic factor Y.

2. For each obligor generate the idiosyncratic component $\epsilon_i, \quad i = 1, ..., N$.

3. Finally, generate the default indicators D_i and the individual losses $\mathcal{L}_i, \quad i = 1, ..., N$.

These steps are repeated till K independent scenarios are generated. Let $\mathcal{L}^{(k)} = \mathcal{L}_1^{(k)} + ... + \mathcal{L}_N^{(k)}$ be the total portfolio loss on the k-th replication. To estimate the risk contributions C_i, we use,

$$\hat{C}_i = \frac{\sum_{k=1}^{K} \mathcal{L}_i^{(k)} \chi_{\{\mathcal{L}^{(k)} \in \mathcal{A}\}}}{\sum_{k=1}^{K} \chi_{\{\mathcal{L}^{(k)} \in \mathcal{A}\}}}.$$

To measure the variability of this estimator we use the following proposition (since $\hat{C}_i$ is a ratio estimator, we can not use a simple standard deviation to measure its precision. For a detailed explanation of the MC method for risk contributions see [Gla05]).

Proposition 4.3.1. *Suppose* $\mathbb{P}(L \in \mathcal{A}) > 0$ *and let*

$$\hat{\sigma}_i^2 = \frac{K \sum_{k=1}^{K} (\mathcal{L}_i^{(k)} - \hat{C}_i)^2 \chi_{\{\mathcal{L}^{(k)} \in \mathcal{A}\}}}{\left(\sum_{k=1}^{K} \chi_{\{\mathcal{L}^{(k)} \in \mathcal{A}\}} \right)^2},$$

taking the ratio to be zero whenever the denominator is zero. Then the distribution of $\frac{\hat{C}_i - C_i}{\hat{\sigma}_i / \sqrt{K}}$ *converges to the standard normal and* $(\hat{C}_i - z_{\delta/2}\hat{\sigma}_i/\sqrt{K}, \hat{C}_i + z_{\delta/2}\hat{\sigma}_i/\sqrt{K})$ *is an asymptotically valid* $1 - \delta$ *confidence interval for* C_i*, with* $\Phi(z_{\delta/2}) = 1 - \delta/2$.

4.4 An Adaptive Gauss-Hermite Integration Formula

An important issue regarding the computation of the coefficients $c_{m,k}$ in (4.4) and $\frac{\partial c_{m,k}(E_i)}{\partial E_i}$ in (4.8) is the way to compute $\Re(Q(E, re^{iu}))$ and $\frac{\partial \Re(Q(E, re^{iu}))}{\partial E_i}$ for a fixed loss level u. In this section we present, in addition, a simplification in the GH formula that considerable reduces the computational effort.

4.4.1 Fast computation of $c_{m,k}(E)$

We have explained in detail the calculation of the coefficients $c_{m,k}$ in Section 3.3.1. There are three key points that essentially determine the computational complexity of the Wavelet Approximation method. The first one is the portfolio size, N, which is fixed. The second one is the number of times that the MGF must be evaluated (2^m+1) and the last one the number of nodes (l) to be used in the GH quadrature every time the MGF is computed at a loss level u. Let us show in the following paragraphs an interesting fact regarding this last point.

We remark that, in practice, financial institutions tend to calibrate the parameter of probability of default of a company from their rating systems and they get in this way several pools. This means that portfolio obligors are classified in rating categories, being the parameter P_n identical for all the obligors inside the same group, and distinct among different groups.

For sake of simplicity, and without loss of generality, let us assume that we have a unique rating category over the whole portfolio. This is $P_n = P$ for all $n = 1, ..., N$, and consequently, $P_n(y) = \overline{P}(y)$ for all $n = 1, ..., N$. Considering (3.28) and (3.29) we observe that under this hypothesis,

$$\lim_{\overline{P}\to 0} R_n = 1, \quad \lim_{\overline{P}\to 0} \theta_n = 0,$$

and,

$$\lim_{\overline{P}\to 1} R_n = r^{2^m E_n}, \quad \lim_{\overline{P}\to 1} \theta_n = 2^m E_n u.$$

Then, given a tolerance ϵ, we can simplify (3.30) and (3.31) by doing,

$$\Re(\hat{M}_{\mathcal{L}}(E, -2^m \ln(re^{iu}); x)) \simeq \frac{1}{\sqrt{\pi}}, \quad \Im(\hat{M}_{\mathcal{L}}(E, -2^m \ln(re^{iu}); x)) \simeq 0,$$

for all $x \geq x_1$ where x_1 is such that,

$$\overline{P}(\sqrt{2}x_1) < \epsilon, \tag{4.14}$$

this is, when $y > \frac{\Phi^{-1}(P) - \sqrt{1-\rho}\Phi^{-1}(\epsilon)}{\sqrt{\rho}}$.

In a similar way,

$$\Re(\hat{M}_{\mathcal{L}}(E, -2^m \ln(re^{iu}); x)) \simeq \frac{1}{\sqrt{\pi}} r^{2^m} \cos(2^m u),$$

and,

$$\Im(\hat{M}_{\mathcal{L}}(E, -2^m \ln(re^{iu}); x)) \simeq \frac{1}{\sqrt{\pi}} r^{2^m} \sin(2^m u),$$

for all $x \leq x_2$ where x_2 is such that $\overline{P}(\sqrt{2}x_2) > 1 - \epsilon$, i.e. when $y < \frac{\Phi^{-1}(P) - \sqrt{1-\rho}\Phi^{-1}(1-\epsilon)}{\sqrt{\rho}}$.

Taking into account these facts we have that (3.26) and (3.27) can be computed as,

$$\begin{aligned}&\Re(\tilde{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu}))) \simeq \\ &\simeq \sum_{j=1}^{l/2-n_2} a_j \Re(\hat{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu});x_j^-)) + \sum_{j=1}^{l/2-n_1} a_j \Re(\hat{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu});x_j^+)) + \\ &+ \frac{1}{\sqrt{\pi}}(n_1 + n_2 r^{2^m}\cos(2^m u)),\end{aligned} \tag{4.15}$$

and,

$$\begin{aligned}&\Im(\tilde{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu}))) \simeq \\ &\simeq \sum_{j=1}^{l/2-n_2} a_j \Im(\hat{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu});x_j^-)) + \sum_{j=1}^{l/2-n_1} a_j \Im(\hat{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu});x_j^+)) + \\ &+ \frac{1}{\sqrt{\pi}} n_2 r^{2^m}\sin(2^m u),\end{aligned} \tag{4.16}$$

where n_1 (respectively n_2) is the number of nodes in the Gauss-Hermite quadrature greater (respectively smaller) or equal than x_1 (respectively x_2). Thus, given a tolerance ϵ we have truncated the left and right tails of the integration variable representing the business cycle making use of the limit behavior instead of continuing with the quadrature. In the section devoted to numerical examples we show the large amount of computation time saved this way, while the accuracy remains the same.

4.4.2 Fast computation of $\frac{\partial c_{m,k}(E)}{\partial E_n}$

Let us explain the computations of the partial derivatives of the coefficients of the WA with respect to the exposures. Again in this case we perform a truncation of the integration variable in a similar way as in the previous section.

Taking the derivative of (4.4) with respect to E_n we get,

$$\frac{\partial c_{m,k}(E)}{\partial E_n} \simeq \frac{2}{\pi r^k}\int_0^{\pi} \frac{\partial \Re(Q_m(E,re^{iu}))}{\partial E_n}\cos(ku)du.$$

Then taking into account (3.22) we have,

$$\frac{\partial \Re(Q_m(E,re^{iu}))}{\partial E_n} = \frac{\frac{\partial \Re(\tilde{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu})))}{\partial E_n}\Re(z_2) + \frac{\partial \Im(\tilde{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu})))}{\partial E_n}\Im(z_2)}{(\Re(z_2))^2 + (\Im(z_2))^2},$$

and again using the Gauss-Hermite quadrature we have,

$$\begin{aligned}\frac{\partial \Re(\tilde{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu})))}{\partial E_n} \simeq \sum_{j=1}^{l/2} a_j \Bigg(&\frac{\partial \Re(\hat{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu});x_j^-))}{\partial E_n} + \\ &+ \frac{\partial \Re(\hat{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu});x_j^+))}{\partial E_n}\Bigg),\end{aligned} \tag{4.17}$$

and,

$$\frac{\partial \Im(\tilde{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu})))}{\partial E_n} \simeq \sum_{j=1}^{l/2} a_j \left(\frac{\partial \Im(\hat{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu});x_j^-))}{\partial E_n} + \right. \\ \left. + \frac{\partial \Im(\hat{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu});x_j^+))}{\partial E_n} \right). \tag{4.18}$$

Now following steps similar to the end of Section 3.3.1 one obtains,

$$\frac{\partial \Re(\hat{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu});x))}{\partial E_n} = \frac{1}{\sqrt{\pi}} \left(\prod_{\underline{n}=1}^{N} R_{\underline{n}} \right) \left[\frac{1}{R_n} \frac{\partial R_n}{\partial E_n} \cos\left(\sum_{\underline{n}=1}^{N} \theta_{\underline{n}} \right) - \frac{\partial \theta_n}{\partial E_n} \sin\left(\sum_{\underline{n}=1}^{N} \theta_{\underline{n}} \right) \right],$$

$$\frac{\partial \Im(\hat{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu});x))}{\partial E_n} = \frac{1}{\sqrt{\pi}} \left(\prod_{\underline{n}=1}^{N} R_{\underline{n}} \right) \left[\frac{1}{R_n} \frac{\partial R_n}{\partial E_n} \sin\left(\sum_{\underline{n}=1}^{N} \theta_{\underline{n}} \right) + \frac{\partial \theta_n}{\partial E_n} \cos\left(\sum_{\underline{n}=1}^{N} \theta_{\underline{n}} \right) \right],$$

$$\frac{\partial R_n}{\partial E_n} = \\ \frac{2^m r^{2^m E_n}}{R_n} \Big(r^{2^m E_n} P_n^2(\sqrt{2}x) \ln r + P_n(\sqrt{2}x)(1 - P_n(\sqrt{2}x)) \left(\ln r \cos(2^m E_n u) - u \sin(2^m E_n u) \right) \Big),$$

$$\frac{\partial \theta_n}{\partial E_n} = \\ \frac{2^m r^{2^m E_n} P_n(\sqrt{2}x) \Big(\theta_n^D (\ln r \sin(2^m E_n u) + \cos(2^m E_n u) u) - \theta_n^N (\ln r \cos(2^m E_n u) - u \sin(2^m E_n u)) \Big)}{(\theta_n^N)^2 + (\theta_n^D)^2},$$

with,

$$\theta_n^N = r^{2^m E_n} P_n(\sqrt{2}x) \sin(2^m E_n u),$$
$$\theta_n^D = 1 - P_n(\sqrt{2}x) + P_n(\sqrt{2}x) r^{2^m E_n} \cos(2^m E_n u).$$

Again we try to reduce the computational effort in the Gauss-Hermite quadratures (4.17) and (4.18). To this end we take into account that,

$$\lim_{\overline{P}\to 0} \frac{\partial R_n}{\partial E_n} = 0, \quad \lim_{\overline{P}\to 0} \frac{\partial \theta_n}{\partial E_n} = 0,$$

and,

$$\lim_{\overline{P}\to 1} \frac{\partial R_n}{\partial E_n} = 2^m r^{2^m E_n} \ln r, \quad \lim_{\overline{P}\to 1} \frac{\partial \theta_n}{\partial E_n} = 2^m u.$$

Then, given a tolerance $\overline{\epsilon}$ we have,

$$\frac{\partial \Re(\hat{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu});x))}{\partial E_n} \simeq 0, \qquad \frac{\partial \Im(\hat{M}_{\mathcal{L}}(E,-2^m\ln(re^{iu});x))}{\partial E_n} \simeq 0,$$

for all $x \geq \overline{x}_1$ where $\overline{x}_1$ is such that,

$$\overline{P}(\sqrt{2}\overline{x}_1) < \overline{\epsilon}, \tag{4.19}$$

this is, $y > \frac{\Phi^{-1}(P)-\sqrt{1-\rho}\Phi^{-1}(\overline{\epsilon})}{\sqrt{\rho}}$.

Also,

$$\frac{\partial \Re(\hat{M}_{\mathcal{L}}(E, -2^m \ln(re^{iu}); x))}{\partial E_n} \simeq \frac{1}{\sqrt{\pi}} 2^m r^{2^m} (\ln r \cos(2^m u) - u \sin(2^m u)),$$

$$\frac{\partial \Im(\hat{M}_{\mathcal{L}}(E, -2^m \ln(re^{iu}); x))}{\partial E_n} \simeq \frac{1}{\sqrt{\pi}} 2^m r^{2^m} (\ln r \sin(2^m u) + u \cos(2^m u)),$$

for all $x \leq \overline{x}_2$ where $\overline{x}_2$ is such that $\overline{P}(\sqrt{2}\overline{x}_2) < \overline{\epsilon}$, i.e. $y < \frac{\Phi^{-1}(P)-\sqrt{1-\rho}\Phi^{-1}(\overline{\epsilon})}{\sqrt{\rho}}$.

Taking into account these facts, (4.17) and (4.18) can be computed by means of,

$$\begin{aligned}
&\frac{\partial \Re(\tilde{M}_{\mathcal{L}}(E, -2^m \ln(re^{iu})))}{\partial E_n} \simeq \\
&\simeq \sum_{j=1}^{l/2-\overline{n}_2} a_j \frac{\partial \Re(\hat{M}_{\mathcal{L}}(E, -2^m \ln(re^{iu}); x_j^-))}{\partial E_n} + \sum_{j=1}^{l/2-\overline{n}_1} a_j \frac{\partial \Re(\hat{M}_{\mathcal{L}}(E, -2^m \ln(re^{iu}); x_j^+))}{\partial E_n} + \\
&+ \frac{1}{\sqrt{\pi}} \overline{n}_2 2^m r^{2^m} (\ln r \cos(2^m u) - u \sin(2^m u)),
\end{aligned} \tag{4.20}$$

and,

$$\begin{aligned}
&\frac{\partial \Im(\tilde{M}_{\mathcal{L}}(E, -2^m \ln(re^{iu})))}{\partial E_n} \simeq \\
&\simeq \sum_{j=1}^{l/2-\overline{n}_2} a_j \frac{\partial \Im(\hat{M}_{\mathcal{L}}(E, -2^m \ln(re^{iu}); x_j^-))}{\partial E_n} + \sum_{j=1}^{l/2-\overline{n}_1} a_j \frac{\partial \Im(\hat{M}_{\mathcal{L}}(E, -2^m \ln(re^{iu}); x_j^+))}{\partial E_n} + \\
&+ \frac{1}{\sqrt{\pi}} \overline{n}_2 2^m r^{2^m} (\ln r \sin(2^m u) + u \cos(2^m u)),
\end{aligned} \tag{4.21}$$

where $\overline{n}_1$ (respectively $\overline{n}_2$) is the number of nodes in the Gauss-Hermite quadrature greater (respectively smaller) or equal than $\overline{x}_1$ (respectively $\overline{x}_2$). Again, given a tolerance $\overline{\epsilon}$, we have truncated the left and right tails of the quadrature in the variable of the business cycle and used the limit behavior instead.

This method is specially suitable for very big portfolios or when the GH quadrature would require a considerable number of nodes. The improvements are shown in the next section devoted to numerical examples.

4.5 Numerical Examples and Discussions

We test[1] the methodology for the computation of ES, VaRC and ESC developed in the previous sections considering four different portfolios. All them with exposure concentrations and ranging from 10 to 10000 obligors. The common set of parameters used for the Wavelet Approximation method are: scale $m = 10$, 2^m intervals for the trapezoidal quadrature in the coefficients formula (4.4) and $r = 0.9995$ (as in Section 3.4).

Portfolio 4.1. *This portfolio has* $N = 10000$ *obligors with* $\rho = 0.15$, $P_n = 0.01$ *and* $E_n = \frac{1}{n}$ *for* $n = 1, ..., N$, *as Portfolio P4 in Section 3.4.*

Portfolio 4.2. *This portfolio has* $N = 1001$ *obligors, with* $E_n = 1$ *for* $n = 1, ..., 1000$, *and one obligor with* $E_{1001} = 100$. $P_n = 0.0033$ *for all the obligors and* $\rho = 0.2$ *as in [Hua07b].*

Portfolio 4.3. *This portfolio has* $N = 100$ *obligors, all them with* $P_n = 0.01$, $\rho = 0.5$ *and exposures,*

$$E_n = \begin{cases} 1, & n = 1, ..., 20, \\ 4, & n = 21, ..., 40, \\ 9, & n = 41, ..., 60, \\ 16, & n = 61, ..., 80, \\ 25, & n = 81, ..., 100, \end{cases}$$

as in [Gla05].

Portfolio 4.4. *This portfolio has* $N = 10$ *obligors, all them with* $\rho = 0.5$, $P_n = 0.0021$ *and* $E_n = \frac{1}{n}$ *for* $n = 1, ..., N$ *as Portfolio P6 in Chapter 3.*

For practical purposes and without loss of generality, in all cases we normalize dividing E_n by $\sum_{n=1}^{N} E_n$ to meet the condition $\sum_{n=1}^{N} E_n = 1$.

From now on, let VaR_α^M, ES_α^M be the VaR and ES values computed by means of a Plain Monte Carlo simulation with 5 million scenarios and $\mathrm{VaRC}_{\alpha,i}^M$, $\mathrm{ESC}_{\alpha,i}^M$ the VaRC and ESC values computed by means of a Plain Monte Carlo simulation with 100 million scenarios to be used as a benchmark.

In what follows let us denote by $\mathrm{VaR}_\alpha^{W(m)}(l)$, $\mathrm{ES}_\alpha^{W(m)}(l)$,$\mathrm{VaRC}_{\alpha,i}^{W(m)}(l)$, $\mathrm{ESC}_{\alpha,i}^{W(m)}(l)$ the result of the Wavelet Approximation method for computing risk measures and contributions, remarking that, we use a Gauss-Hermite quadrature with l nodes for the integrals (3.26) and (3.27) in the case of VaR and ES, and also for the integrals (4.17) and (4.18) in the case of VaRC and ESC. Analogously, $\mathrm{VaR}_\alpha^{W(m)}(l, \epsilon)$, $\mathrm{ES}_\alpha^{W(m)}(l, \epsilon)$, $\mathrm{VaRC}_{\alpha,i}^{W(m)}(l, \bar{\epsilon})$, $\mathrm{ESC}_{\alpha,i}^{W(m)}(l, \bar{\epsilon})$ denote Wavelet Approximation results when computing risk measures and risk contributions considering l-nodes in the Gauss-Hermite quadrature, but using the expressions (4.14), (4.15) and (4.16) in the case of VaR and ES, and (4.19), (4.20), (4.21) in the case of VaRC and ESC.

Let VaR_α^A and $\mathrm{VaRC}_{\alpha,i}^A$ be the VaR and VaR contributions evaluated by the Asymptotic Single Risk Factor (ASRF) model (further details about this method can be found in

[1]Computations have been carried out sequentially in a personal computer Dell Vostro 320 under GNU/Linux OS, Intel CPU Core 2 E7500, 2.93GHz, 4GB RAM and using the gcc compiler with optimization level 2.

Section 1.4 and in [Lut09]). Table 4.1 presents the very high confidence VaR values for portfolio 4.1 computed by means of Monte Carlo, ASRF and the Wavelet Approximation method. The number of negative $(l/2 - n_2)$ and positive $(l/2 - n_1)$ nodes where the conditional MGF is evaluated are also specified. Observe that with the approximations $\mathrm{VaR}_\alpha^{W(10)}(20, 4{\cdot}10^{-1})$ and $\mathrm{VaR}_\alpha^{W(10)}(20, 6{\cdot}10^{-1})$ the MFG is not evaluated in positive nodes, due to the fact that the conditional default probabilities (4.14) are extremely small at this points. The relative error presented for the VaR value at 99.999% confidence level is about -1% when using the WA method. Figure 4.1 represents the loss distribution for the WA method with a 20-nodes GH formula and $\epsilon = 6 \cdot 10^{-1}$, which only requires the evaluation of the MGF in the first 7 negative nodes to achieve a high precision. We also provide the result for the ASRF method and Monte Carlo (which, as always, give us the benchmark). The estimation of VaR by means of the $\mathrm{VaR}_\alpha^{W(10)}(20, 6 \cdot 10^{-1})$ approximation requires 25.3 seconds of CPU time, while the $\mathrm{VaR}_\alpha^{W(10)}(20)$ approximation needs of 71.5 seconds. This is, the implementation of the asymptotic truncation of Sections 4.4.1 and 4.4.2 represents an important improvement. It is also worth to underline that $\mathrm{VaR}_{0.9999}^{W(9)}(20, 6 \cdot 10^{-1})$ and $\mathrm{VaR}_{0.9999}^{W(8)}(20, 6 \cdot 10^{-1})$ give also very accurate results (with relative errors equal to -0.47% and -0.90% respectively) and computation times of 12.7 and 6.4 seconds respectively. We also want to remark that the ASRF method clearly underestimates the risk due to the presence of name concentration.

The Expected Shortfall calculated at several confidence levels with $\mathrm{ES}_\alpha^{W(10)}(20, 6{\cdot}10^{-1})$ is presented in Table 4.2 (we omit the computational time because there is almost no difference between VaR and ES calculation in terms of computational effort). As the results show again, a high precision is achieved in terms of the relative error. Figures 4.2 and 4.3 represent the contributions to the expected shortfall at 99% and 99.9% confidence levels respectively by means of the WA method with a 20-nodes GH formula and $\bar{\epsilon} = 10^{-4}$ (we present the ESC instead of the VaRC due to the robustness of plain Monte Carlo simulation for the first measure). For sake of clarity in the plots, we have represented only the 250 biggest and smallest risk contributions. The convergence towards Monte Carlo is clear and the sum of the risk contributions shown in Table 4.3 are very close to the ES values given in Table 4.2. $\mathrm{ESC}_{\alpha,n}^{W(10)}(20, 10^{-4})$ takes 622.5 seconds of CPU to evaluate the partial derivative of the MGF in 14 nodes, while the $\mathrm{ESC}_{\alpha,n}^{W(10)}(20)$ method needs of 671.8 seconds.

Method	$l/2 - n_2$	$l/2 - n_1$	$\alpha = 0.9999$	$\alpha = 0.99999$
VaR_α^M			0.2267	0.2973
VaR_α^A			0.1683 (-25.76%)	0.2322 (-21.91%)
$\mathrm{VaR}_\alpha^{W(10)}(20)$	10	10	0.2261 (-0.25%)	0.2935 (-1.30%)
$\mathrm{VaR}_\alpha^{W(10)}(20, 4 \cdot 10^{-1})$	9	0	0.2261 (-0.25%)	0.2944 (-0.97%)
$\mathrm{VaR}_\alpha^{W(10)}(20, 6 \cdot 10^{-1})$	7	0	0.2261 (-0.25%)	0.2944 (-0.97%)

Table 4.1: VaR values at 99.99% and 99.999% confidence levels for Portfolio 4.1. Errors relative to Monte Carlo are shown in parenthesis.

Next we consider Portfolio 4.2 which is a well diversified portfolio where a big exposure

Method	$l/2-n_2$	$l/2-n_1$	$\alpha=0.99$	$\alpha=0.999$	$\alpha=0.9999$
ES_α^M			0.1290	0.1895	0.2553
$\mathrm{ES}_\alpha^{W(10)}(20)$	10	10	0.1290 (−0.02%)	0.1895 (−0.01%)	0.2556 (0.12%)
$\mathrm{ES}_\alpha^{W(10)}(20, 4\cdot 10^{-1})$	9	0	0.1289 (−0.12%)	0.1895 (−0.01%)	0.2556 (0.12%)
$\mathrm{ES}_\alpha^{W(10)}(20, 6\cdot 10^{-1})$	7	0	0.1289 (−0.11%)	0.1896 (0.00%)	0.2559 (0.25%)

Table 4.2: ES values at 99%, 99.9% and 99.99% confidence levels for Portfolio 4.1. Errors relative to Monte Carlo are shown in parenthesis.

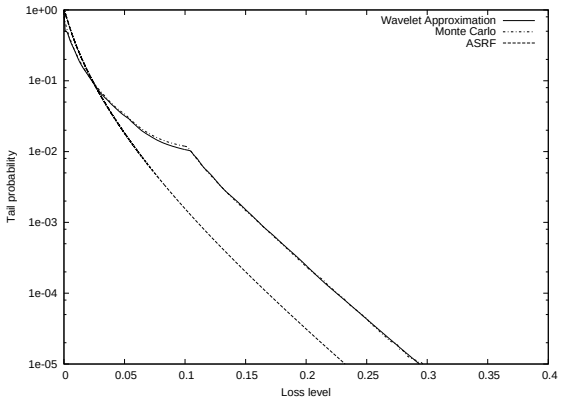

Figure 4.1: Tail probability approximation of Portfolio 4.1 with WA using a 20-nodes GH formula and $\epsilon = 6\cdot 10^{-1}$.

(representing about 9% of the total portfolio exposure) has been added. So it presents exposure concentration. As pointed out in [Hua07b], a straightforward Saddle Point approximation fails for all the quantiles preceding the point of non smoothness. However, WA method is capable to deal with this problem as we illustrate in Figure 4.4.

We also note that we obtain the same accuracy with the evaluation of only 15 nodes of the asymptotic formulae (which needs only 5.6 seconds of CPU time) and using the full one with 64 nodes (which needs of 23.4 seconds). The risk contributions to the 99.9% ES are also provided in Table 4.6 using different $\bar{\epsilon}$. We present only the biggest and the smallest risk contributions. Relative errors are almost identical with 33 and 64 nodes and the computation times are 78.4 and 103.5 seconds. It is important to mention that, in practice, we only need to compute the contributions for two different exposures. However, we have performed the calculations for the whole portfolio in order to have an idea of the computational effort for a portfolio of this size.

Let us consider now the Portfolio 4.3. The plot of Figure 4.5 shows the portfolio loss distribution approximated with the WA method with a 64-nodes GH formula and $\epsilon = 10^{-1}$. Table 4.8 contains different scenarios changing the parameter ϵ. It is remarkable the high precision achieved when using $\epsilon = 10^{-1}$, which means that the conditional MGF has been just evaluated in 12 negative nodes. The $\mathrm{VaR}_\alpha^{W(10)}(64)$ approximation needs 2.3 seconds while the approximation $\mathrm{VaR}_\alpha^{W(10)}(64, 10^{-1})$ needs only 0.6 seconds.

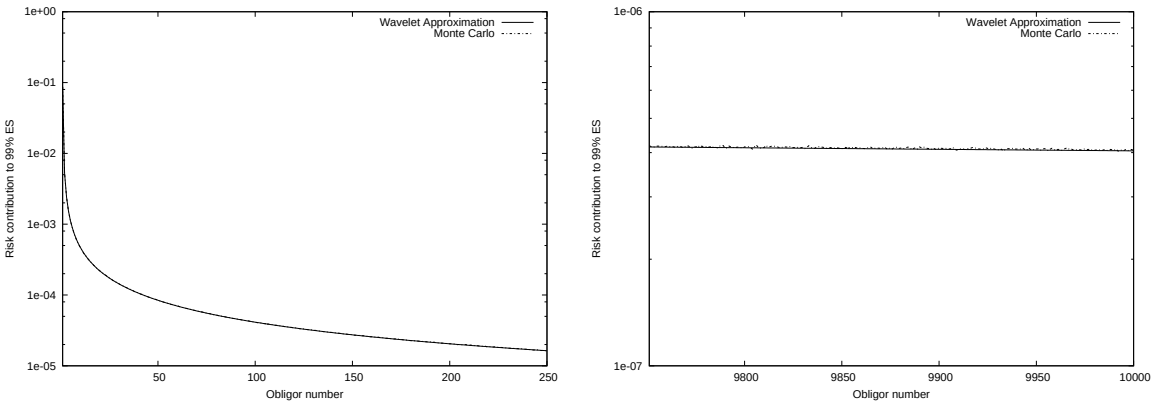

Figure 4.2: Risk contributions to the ES at 99% confidence level for the Portfolio 4.1.

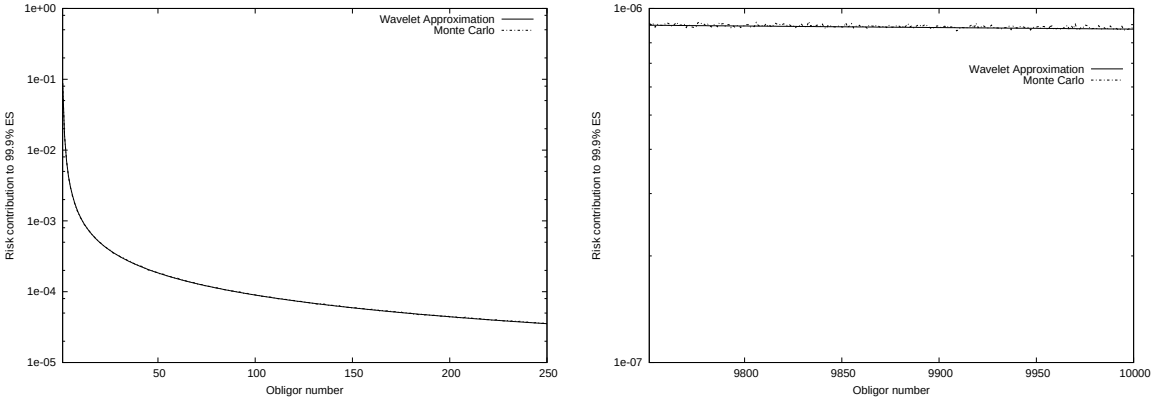

Figure 4.3: Risk contributions to the ES at 99.9% confidence level for Portfolio 4.1.

The 99.9% VaR contributions are presented in Table 4.9 and plotted in Figure 4.6. To compute the VaR contributions by means of Monte Carlo simulation we have considered $\mathcal{A} = (\mathrm{VaR}_\alpha^M - 5 \cdot 10^{-4}, \mathrm{VaR}_\alpha^M + 5 \cdot 10^{-4})$ instead of $\mathcal{A} = \mathrm{VaR}_\alpha^M$, due to the fact that VaR is a rare event. Moreover, we have generated 99% confidence intervals for the risk contributions as detailed in Proposition 4.3.1. The risk contributions calculated by the WA method using a 64-nodes GH formula and $\bar{\epsilon} = 10^{-4}$ lie in the 99% MC confidence intervals, showing again the excellent accuracy of the method. The ES contributions at 99.9% and 99.99% confidence levels are presented in Tables 4.10 and 4.11 respectively (and plotted in Figure 4.7). The $\mathrm{ESC}_{\alpha,n}^{W(10)}(64, 10^{-4})$ approximation shows very accurate results except for the 20 smallest exposures at confidence level 99.99% which considerably underestimates the risk. The $\mathrm{ESC}_{\alpha,n}^{W(10)}(64)$ method takes 8 seconds for the computations while $\mathrm{ESC}_{\alpha,n}^{W(10)}(64, 10^{-4})$ needs of 4.5 seconds. Like in the previous example, we have calculated all the risk contributions.

In order to highlight the power of the WA method regarding the computational time,

Method	$l/2-\overline{n}_2$	$l/2-\overline{n}_1$	$\alpha=0.99$	$\alpha=0.999$
$\sum_{n=1}^{N}\mathrm{ESC}_{\alpha,n}^{M}$			0.1290	0.1892
$\sum_{n=1}^{N}\mathrm{ESC}_{\alpha,n}^{W(10)}(20)$	10	10	0.1293	0.1891
$\sum_{n=1}^{N}\mathrm{ESC}_{\alpha,n}^{W(10)}(20,10^{-4})$	10	4	0.1293	0.1890

Table 4.3: Comparison of the total ES contributions at 99% and 99.9% confidence levels for Portfolio 4.1.

Method	CPU time
$\mathrm{VaR}_{\alpha}^{W(10)}(20,6\cdot10^{-1})$	25.3
$\mathrm{VaR}_{\alpha}^{W(9)}(20,6\cdot10^{-1})$	12.7
$\mathrm{VaR}_{\alpha}^{W(8)}(20,6\cdot10^{-1})$	6.4
$\mathrm{ESC}_{\alpha,n}^{W(10)}(20,10^{-4})$	622.5

Table 4.4: Computational time (in seconds) for Portfolio 4.1.

we have shown the most relevant results in Tables 4.4, 4.7 and 4.12. Roughly speaking, the WA approximation is more than 100 times faster than plain MC simulation.

Method	$l/2-n_2$	$l/2-n_1$	$\alpha=0.999$	$\alpha=0.9999$
VaR_α^M			0.1077	0.1532
VaR_α^A			0.0679 (−37.00%)	0.1195 (−21.99%)
$\mathrm{VaR}_\alpha^{W(10)}(64)$	32	32	0.1079 (0.18%)	0.1538 (0.41%)
$\mathrm{VaR}_\alpha^{W(10)}(64,10^{-8})$	32	13	0.1079 (0.18%)	0.1538 (0.41%)
$\mathrm{VaR}_\alpha^{W(10)}(64,10^{-4})$	30	3	0.1079 (0.18%)	0.1538 (0.41%)
$\mathrm{VaR}_\alpha^{W(10)}(64,10^{-2})$	25	0	0.1079 (0.18%)	0.1538 (0.41%)
$\mathrm{VaR}_\alpha^{W(10)}(64,5\cdot10^{-1})$	15	0	0.1079 (0.18%)	0.1538 (0.41%)
ES_α^M			0.1274	0.1809
$\mathrm{ES}_\alpha^{W(10)}(64)$	32	32	0.1273 (−0.02%)	0.1810 (0.09%)
$\mathrm{ES}_\alpha^{W(10)}(64,10^{-8})$	32	13	0.1273 (−0.02%)	0.1810 (0.09%)
$\mathrm{ES}_\alpha^{W(10)}(64,10^{-4})$	30	3	0.1273 (−0.02%)	0.1810 (0.09%)
$\mathrm{ES}_\alpha^{W(10)}(64,10^{-2})$	25	0	0.1273 (−0.02%)	0.1810 (0.09%)
$\mathrm{ES}_\alpha^{W(10)}(64,5\cdot10^{-1})$	15	0	0.1273 (−0.02%)	0.1810 (0.09%)

Table 4.5: VaR and ES values at 99.9% and 99.99% confidence levels for Portfolio 4.2. Errors relative to Monte Carlo are shown in parenthesis.

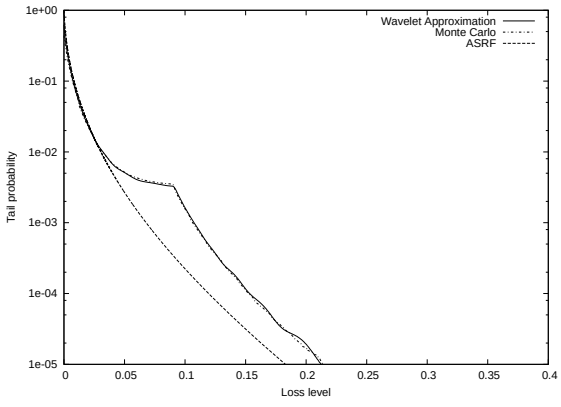

Figure 4.4: Tail probability approximation for Portfolio 4.2 with WA using a 64-nodes GH quadrature and $\epsilon=5\cdot10^{-1}$.

Method	$l/2-\overline{n}_2$	$l/2-\overline{n}_1$	$\alpha=0.999$
$\mathrm{ESC}^{M}_{\alpha,1001}$			0.082016
$\mathrm{ESC}^{W(10)}_{\alpha,1001}(64)$	32	32	0.081075 (−1.15%)
$\mathrm{ESC}^{W(10)}_{\alpha,1001}(64,10^{-8})$	32	13	0.081075 (−1.15%)
$\mathrm{ESC}^{W(10)}_{\alpha,1001}(64,10^{-6})$	32	9	0.081075 (−1.15%)
$\mathrm{ESC}^{W(10)}_{\alpha,1001}(64,10^{-4})$	30	3	0.081075 (−1.15%)
$\mathrm{ESC}^{M}_{\alpha,1}$			0.000045
$\mathrm{ESC}^{W(10)}_{\alpha,1}(64)$	32	32	0.000046 (2.04%)
$\mathrm{ESC}^{W(10)}_{\alpha,1}(64,10^{-8})$	32	32	0.000046 (2.04%)
$\mathrm{ESC}^{W(10)}_{\alpha,1}(64,10^{-6})$	32	9	0.000046 (2.04%)
$\mathrm{ESC}^{W(10)}_{\alpha,1}(64,10^{-4})$	30	3	0.000046 (1.92%)
$\sum_{n=1}^{N}\mathrm{ESC}^{M}_{\alpha,n}$			0.1274
$\sum_{n=1}^{N}\mathrm{ESC}^{W(10)}_{\alpha,n}(64)$	32	32	0.1274 (−0.01%)
$\sum_{n=1}^{N}\mathrm{ESC}^{W(10)}_{\alpha,n}(64,10^{-8})$	32	32	0.1274 (−0.01%)
$\sum_{n=1}^{N}\mathrm{ESC}^{W(10)}_{\alpha,n}(64,10^{-6})$	32	9	0.1274 (−0.01%)
$\sum_{n=1}^{N}\mathrm{ESC}^{W(10)}_{\alpha,n}(64,10^{-4})$	30	3	0.1273 (−0.05%)

Table 4.6: ES contributions at 99.9% confidence level for Portfolio 4.2. Errors relative to Monte Carlo are shown in parenthesis.

Method	CPU time
$\mathrm{VaR}^{W(10)}_{\alpha}(64,5\cdot10^{-1})$	5.6
$\mathrm{ESC}^{W(10)}_{\alpha,n}(64,10^{-4})$	78.4

Table 4.7: Computational time (in seconds) for Portfolio 4.2.

Method	$l/2-n_2$	$l/2-n_1$	$\alpha=0.999$	$\alpha=0.9999$
$\mathrm{VaR}^{M}_{\alpha}$			0.4350	0.6859
$\mathrm{VaR}^{A}_{\alpha}$			0.4209 (−3.25%)	0.6661 (−2.89%)
$\mathrm{VaR}^{W(10)}_{\alpha}(64)$	32	32	0.4341 (−0.21%)	0.6870 (0.16%)
$\mathrm{VaR}^{W(10)}_{\alpha}(64,10^{-8})$	22	6	0.4341 (−0.21%)	0.6870 (0.16%)
$\mathrm{VaR}^{W(10)}_{\alpha}(64,10^{-4})$	17	1	0.4341 (−0.21%)	0.6870 (0.16%)
$\mathrm{VaR}^{W(10)}_{\alpha}(64,10^{-2})$	14	0	0.4341 (−0.21%)	0.6870 (0.16%)
$\mathrm{VaR}^{W(10)}_{\alpha}(64,10^{-1})$	12	0	0.4341 (−0.21%)	0.6870 (0.16%)
ES^{M}_{α}			0.5445	0.7576
$\mathrm{ES}^{W(10)}_{\alpha}(64)$	32	32	0.5449 (0.08%)	0.7621 (0.59%)
$\mathrm{ES}^{W(10)}_{\alpha}(64,10^{-8})$	22	6	0.5449 (0.08%)	0.7621 (0.59%)
$\mathrm{ES}^{W(10)}_{\alpha}(64,10^{-4})$	17	1	0.5449 (0.08%)	0.7621 (0.59%)
$\mathrm{ES}^{W(10)}_{\alpha}(64,10^{-2})$	14	0	0.5449 (0.08%)	0.7621 (0.59%)
$\mathrm{ES}^{W(10)}_{\alpha}(64,10^{-1})$	12	0	0.5450 (0.09%)	0.7624 (0.63%)

Table 4.8: VaR and ES values at 99.9% and 99.99% confidence levels for Portfolio 4.3. Errors relative to Monte Carlo are shown in parenthesis.

	$l/2-\overline{n}_2$	$l/2-\overline{n}_1$	$j_1=1, j_2=20$	$j_1=21, j_2=40$	$j_1=41, j_2=60$	$j_1=61, j_2=80$	$j_1=81, j_2=100$
MC 99% CI for $\mathrm{VaRC}^{M}_{0.999,n}$			$(-0.000329, 0.001073)$	$(0.000807, 0.002209)$	$(0.002710, 0.004112)$	$(0.005585, 0.006987)$	$(0.009473, 0.010875)$
$\frac{\sum_{j=j_1}^{j_2} \mathrm{VaRC}^{A}_{0.999,j}}{20}$			0.000383	0.001530	0.003443	0.006121	0.009565
$\frac{\sum_{j=j_1}^{j_2} \mathrm{VaRC}^{W(m)}_{0.999,j}(64)}{20}$	32	32	0.000364	0.001472	0.003435	0.006229	0.010203
$\frac{\sum_{j=j_1}^{j_2} \mathrm{VaRC}^{W(m)}_{0.999,j}(64,10^{-8})}{20}$	22	6	0.000364	0.001472	0.003435	0.006229	0.010203
$\frac{\sum_{j=j_1}^{j_2} \mathrm{VaRC}^{W(m)}_{0.999,j}(64,10^{-6})}{20}$	20	4	0.000364	0.001472	0.003435	0.006229	0.010203
$\frac{\sum_{j=j_1}^{j_2} \mathrm{VaRC}^{W(m)}_{0.999,j}(64,10^{-4})}{20}$	17	1	0.000364	0.001475	0.003442	0.006226	0.010197

Table 4.9: VaR contributions at 99.9% confidence level for Portfolio 4.3. The values presented for the VaR contributions have been averaged over groups of obligors with the same exposure.

	$l/2-\overline{n}_2$	$l/2-\overline{n}_1$	$j_1=1, j_2=20$	$j_1=21, j_2=40$	$j_1=41, j_2=60$	$j_1=61, j_2=80$	$j_1=81, j_2=100$	$\sum$
$\frac{\sum_{j=j_1}^{j_2} \mathrm{ESC}^{M}_{0.999,j}}{20}$			0.000466	0.001883	0.004316	0.007861	0.012677	0.5441
$\frac{\sum_{j=j_1}^{j_2} \mathrm{ESC}^{W(10)}_{0.999,j}(64)}{20}$	32	32	0.000466 (−0.06%)	0.001884 (0.04%)	0.004315 (−0.03%)	0.007867 (0.08%)	0.012696 (0.15%)	0.5446 (0.09%)
$\frac{\sum_{j=j_1}^{j_2} \mathrm{ESC}^{W(10)}_{0.999,j}(64,10^{-8})}{20}$	22	6	0.000466 (−0.06%)	0.001884 (0.04%)	0.004315 (−0.03%)	0.007867 (0.08%)	0.012696 (0.15%)	0.5446 (0.09%)
$\frac{\sum_{j=j_1}^{j_2} \mathrm{ESC}^{W(10)}_{0.999,j}(64,10^{-6})}{20}$	20	4	0.000466 (−0.06%)	0.001884 (0.04%)	0.004315 (−0.03%)	0.007867 (0.08%)	0.012696 (0.15%)	0.5446 (0.09%)
$\frac{\sum_{j=j_1}^{j_2} \mathrm{ESC}^{W(10)}_{0.999,j}(64,10^{-4})}{20}$	17	1	0.000460 (−1.38%)	0.001884 (0.04%)	0.004315 (−0.03%)	0.007867 (0.08%)	0.012696 (0.15%)	0.5444 (0.07%)

Table 4.10: ES contributions at 99.9% confidence level for Portfolio 4.3. The values presented for the ES contributions have been averaged over groups of obligors with the same exposure. The last column represents the sum of all the credit risk contributions. Errors relative to Monte Carlo are shown in parenthesis.

	$l/2-\overline{n}_2$	$l/2-\overline{n}_1$	$j_1=1, j_2=20$	$j_1=21, j_2=40$	$j_1=41, j_2=60$	$j_1=61, j_2=80$	$j_1=81, j_2=100$	$\sum$
$\frac{\sum_{j=j_1}^{j_2} \mathrm{ESC}^{M}_{0.9999,j}}{20}$			0.000662	0.002667	0.006094	0.011034	0.017705	0.7632
$\frac{\sum_{j=j_1}^{j_2} \mathrm{ESC}^{W(10)}_{0.9999,j}(64)}{20}$	32	32	0.000658 (−0.51%)	0.002657 (−0.38%)	0.006067 (−0.45%)	0.011009 (−0.23%)	0.017643 (−0.35%)	0.7607 (−0.33%)
$\frac{\sum_{j=j_1}^{j_2} \mathrm{ESC}^{W(10)}_{0.9999,j}(64,10^{-8})}{20}$	22	6	0.000658 (−0.51%)	0.002657 (−0.38%)	0.006067 (−0.45%)	0.011009 (−0.23%)	0.017643 (−0.35%)	0.7607 (−0.33%)
$\frac{\sum_{j=j_1}^{j_2} \mathrm{ESC}^{W(10)}_{0.9999,j}(64,10^{-6})}{20}$	20	4	0.000658 (−0.51%)	0.002657 (−0.38%)	0.006067 (−0.45%)	0.011009 (−0.23%)	0.017643 (−0.35%)	0.7607 (−0.33%)
$\frac{\sum_{j=j_1}^{j_2} \mathrm{ESC}^{W(10)}_{0.9999,j}(64,10^{-4})}{20}$	17	1	0.000622 (−5.93%)	0.002657 (−0.38%)	0.006067 (−0.45%)	0.011009 (−0.23%)	0.017643 (−0.35%)	0.7600 (−0.43%)

Table 4.11: ES contributions at 99.99% confidence level for Portfolio 4.3. The values presented for the ES contributions have been averaged over groups of obligors with the same exposure. The last column represents the sum of all the credit risk contributions. Errors relative to Monte Carlo are shown in parenthesis.

Method	CPU time
$\mathrm{VaR}_{\alpha}^{W(10)}(64, 10^{-1})$	0.6
$\mathrm{ESC}_{\alpha,n}^{W(10)}(64, 10^{-4})$	4.5

Table 4.12: Computational time (in seconds) for Portfolio 4.3.

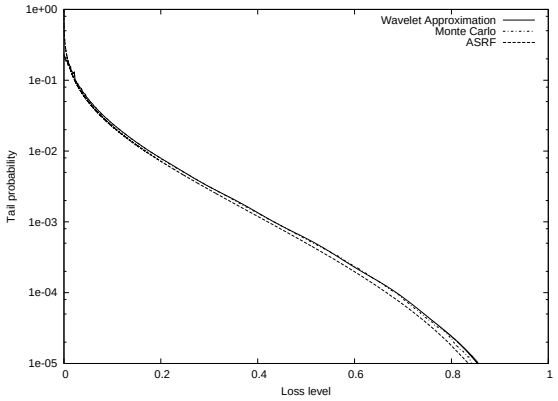

Figure 4.5: Tail probability approximation for Portfolio 4.3 using WA with a 64-nodes GH formula and $\epsilon = 10^{-1}$.

Finally we present the results for Portfolio 4.4. The computation of the 99.99% ES and the 99.99% ESC have been carried out by means of the plain WA method with a 20-nodes GH formula, since this portfolio is extremely small and the computational effort is tinny to consider other improvements. The results are presented in Tables 4.13 and 4.14. With this example we just want to remark that the WA method is very versatile and it can also deal with very small portfolios.

Method	$\alpha = 0.9999$
ES_{α}^{M}	0.6833
$\mathrm{ES}_{\alpha}^{W(10)}(20)$	0.6814 (-0.29%)

Table 4.13: ES at 99.99% confidence level for Portfolio 4.4. Relative error to Monte Carlo is shown in parenthesis.

This Chapter extends a previous work undertaken in Chapter 3. It is based on a Haar wavelet approximation to the cumulative distribution of the loss function with the computation of the ES as an alternative coherent risk measure to VaR. Moreover, a detailed procedure for the calculation of the risk contributions to the VaR and the ES in a credit portfolio is provided. The risk contributions are known to be very computationally intensive to be estimated by means of MC because they are the expected value of the individual loss conditioned on a rare event. Therefore, analytical or fast numerical

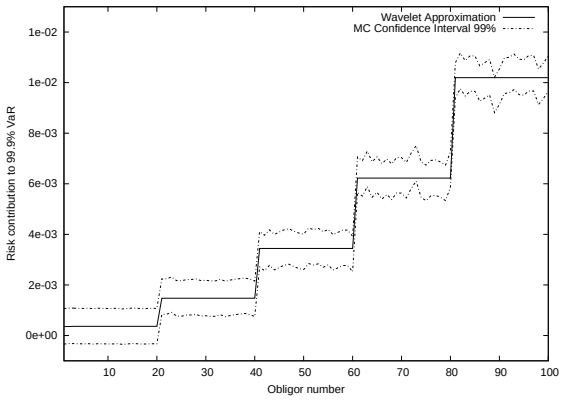

Figure 4.6: VaR contributions at 99.9% confidence level for Portfolio 4.3 using the WA method and GH integration formulas with 64 nodes and $\bar{\epsilon} = 10^{-4}$.

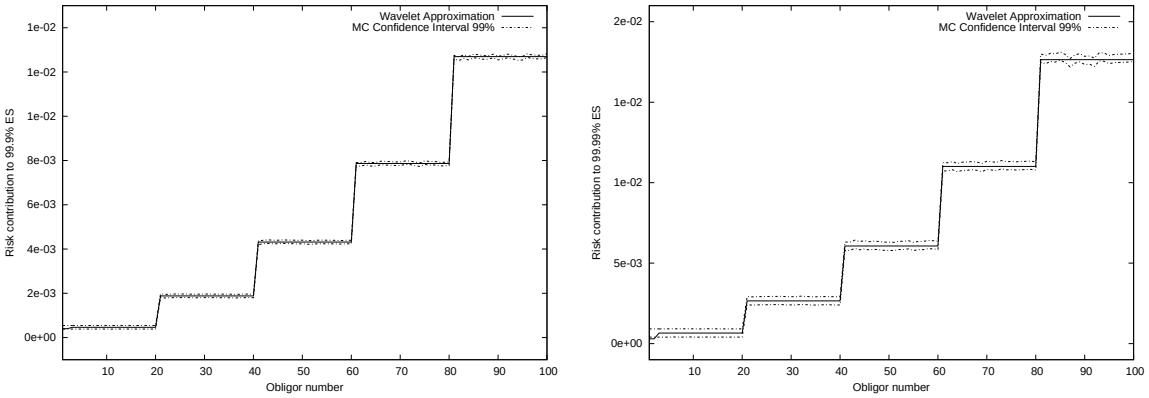

Figure 4.7: Expected Shortfall contributions at 99.9% (left) and 99.99% (right) confidence levels for Portfolio 4.3 using the WA method and GH integration formulas with 64 nodes and $\bar{\epsilon} = 10^{-4}$.

methods are welcome to overcome this problem. The model framework is the well known one-factor Merton model, a one-period default model, which is the basis of the Basel II Accord.

There are technical points taken into account that contribute to a considerable improvement of the WA method. We avoid the evaluation of the MFG in all the nodes of the Gauss-Hermite formulas by means of using its asymptotic behavior. Proceeding this way the speed of the WA method increases while accuracy remains. These improvements are also applied to the computation of risk contributions to VaR and ES, although the impact in the speed of the algorithm is much more relevant for risk measures than for risk contributions.

This new methodology has been tested in a wide sized variety of portfolios, all them with exposure concentration, where the Asymptotic Single Risk Factor Model fails due to the name concentration. The results presented show that the Wavelet Approximation

Obligor	$\mathrm{ESC}^{W(10)}_{0.9999,n}(20)$	$\mathrm{ESC}^{M}_{0.9999,n}$	Relative error
1	0.340472	0.338706	0.52%
2	0.128672	0.128885	−0.17%
3	0.059568	0.059426	0.24%
4	0.041619	0.042070	−1.07%
5	0.028044	0.028805	−2.64%
6	0.023233	0.022825	1.79%
7	0.019335	0.019348	−0.06%
8	0.016606	0.015935	4.21%
9	0.014527	0.014614	−0.60%
10	0.012971	0.012227	6.08%
$\sum_{n=1}^{N} \mathrm{ESC}^{W(10)}_{0.9999,n}(20)$	0.6850	0.6828	0.32%

Table 4.14: ES contributions at 99.99% confidence level for Portfolio 4.4 with the WA method.

method is highly competitive in terms of robustness, speed and accuracy being a very suitable method to measure and manage the risks that arise in credit portfolios of financial companies.

Chapter 5

The WA Extension to Merton Model with Several Factors

The Asymptotic Single Risk Factor model for the computation of regulatory capital that underlies Basell II, does not provide an appropriate quantitative framework for analyzing sector concentration risk, due to the fact that it is based on a credit portfolio model which is only applicable under the assumptions that the bank portfolios are perfectly fine-grained and there is only a single source of systematic risk. The simplicity of the model ensures its analytical tractability. However it makes it impossible to model risk concentrations in a realistic way. Neither is name concentration captured, as we have demonstrated in Chapters 3 and 4, nor is it possible to define sector concentration in this one-factor model.

To account for sector concentration we make use of the WA method under the multi-factor Merton model presented in Section 1.3.2, which involves the computation of high dimensional integrals. To do this, we introduce the *low discrepancy sequences* and quasi-Monte Carlo methods (QMC).

5.1 Low Discrepancy Sequences

Discrepancy is a measure of deviation from uniformity of a sequence of real numbers. In particular, the discrepancy of n points $x_1, \ldots, x_n \in [0,1]^d, d \geq 1$, is defined by,

$$D_n^{(d)} = \sup_E \left| \frac{A(E;n)}{n} - \lambda(E) \right|,$$

where the supremum is taken over all the subsets of $[0,1]^d$ of the form $E = [0,t_1) \times \cdots \times [0,t_d), 0 \leq t_j \leq 1, 1 \leq j \leq d, \lambda$ denotes the Lebesgue measure, and $A(E;n)$ denotes the number of the x_j that are contained in E.

A sequence $x_1, x_2, \ldots$ of points in $[0,1]^d$ is a low discrepancy sequence if,

$$D_n^{(d)} \leq c(d) \frac{(\log n)^d}{n}, \quad \forall n > 1,$$

where the constant $c(d)$ depends only on the dimension d. A detailed analysis of low discrepancy sequences can be found in [Nie92] and in references therein. The author gives a general method for constructing (t, d)-sequences, which are low discrepancy sequences such that the discrepancy of the first n points is given by,

$$D_n^{(d)} \leq c(t, d, b)\frac{(\log n)^d}{n} + \mathcal{O}\left(\frac{(\log n)^{d-1}}{n}\right),$$

where $b \geq 2$ is an integer parameter, upon which the sequence depends, $c(t, d, b) \simeq b^t/d! \cdot (b/2\log b)^d$ and $t \geq 0$. The value $t = 0$ is desirable.

The generalized Faure sequence (see [Tez95]) is a $(0, d)$-sequence and is obtained as follows. For a prime number $b \geq d$ and $n = 0, 1, \ldots$, consider the base b representation of n, i.e.,

$$n = \sum_{i=0}^{+\infty} a_i(n)b^i,$$

where $a_i(n) \in [0, b)$ are integers. The j-th coordinate of the point x_n is then given by,

$$x_n^{(j)} = \sum_{k=0}^{+\infty} x_{n,k}^{(j)} b^{-k-1}, \quad 1 \leq j \leq d,$$

where,

$$x_{n,k}^{(j)} = \sum_{s=0}^{+\infty} c_{k,s}^{(j)} a_s(n).$$

The matrix $C^{(j)} = \left(c_{k,s}^{(j)}\right)$ is called the generator matrix of the sequence and is given by $C^{(j)} = A^{(j)}P^{j-1}$, where $A^{(j)}$ is a nonsingular lower triangular matrix and P^{j-1} denotes the $j-1$ power of the Pascal matrix[1], $1 \leq j \leq d$.

Definition 5.1.1. *If f is sufficiently differentiable, then the variation of f on $[0, 1]^d$ in the sense of Hardy and Krause is,*

$$V(f) = \sum_{k=1}^{d} \sum_{1 \leq i_1 < \cdots < i_k \leq d} V^{(k)}(f; i_1, \ldots, i_k),$$

where,

$$V^{(k)}(f; i_1, \ldots, i_k) = \int_0^1 \cdots \int_0^1 \left|\frac{\partial^k f}{\partial x_{i_1} \cdots x_{i_k}}\right| dx_{i_1} \cdots x_{i_k},$$

is the restriction of f to the k-dimensional face $\{(x_1, \ldots, x_d) \in [0, 1]^d : x_j = 1, j \neq i_1, \ldots, i_k\}$.

In order to conclude, we state the Koksma-Hlawka inequality which establishes the relationship between low discrepancy sequences and multivariate integration (see [Nie92]

[1]Pascal matrix is an infinite matrix containing the binomial coefficients.

for details). If f is a real function defined in $[0,1]^d$, of bounded variation, $V(f)$, in the sense of Hardy and Krause, then for any sequence $x_1, \ldots, x_n \in [0,1)^d$ we have,

$$\left| \int_{[0,1]^d} f(x)\, dx - \frac{1}{n} \sum_{i=1}^{n} f(x_i) \right| \leq V(f) D_n^{(d)}. \tag{5.1}$$

5.2 The WA method for the Multi-Factor Merton Model

Let us consider the distribution of losses from default for a credit portfolio over a fixed time horizon $[0,T]$, which again is usually one year. Let us consider the total loss from default given by $\mathcal{L} = \sum_{n=1}^{N} \mathcal{L}_i$, where N is the number of obligors or firms to which the portfolio is exposed, $\mathcal{L}_i = E_n \cdot D_n$, E_n are the exposures and D_n the default indicators. Note that, without loss of generality, we are assuming constant loss given default 100%.

The indicator of default for the obligor n is assumed to be represented in terms of a latent variable r_n and a threshold T_n, $D_n = \chi_{\{r_n < T_n\}}$, where the latent variables can be represented as,

$$r_n = a_{n,1} X_1 + \cdots + a_{n,d} X_d + b_n \epsilon_n, \quad n = 1, \ldots, N, \tag{5.2}$$

with $X = (X_1, \ldots, X_d)$ representing systematic risk factors, $\epsilon_n, n = 1, \ldots, N$, idiosyncratic factors and $a_{n,1}, \ldots, a_{n,d}, b_n, n = 1, \ldots, N$ their corresponding factor loadings. Although the common factors $X_1, \ldots, X_d$ may be correlated, we can transform them into independent random variables by a factorization of their covariance matrix. So we assume that $X_1, \ldots, X_d, \epsilon_1, \ldots, \epsilon_N$ are independent standard normal random variables and,

$$a_{n,1}^2 + \cdots + a_{n,d}^2 + b_n^2 = 1, \quad n = 1, \ldots, N.$$

This implies that $r_1, \ldots, r_N$ are standard normal random variables with correlation,

$$cov(r_n, r_m) = \sum_{i=1}^{N} a_{n,i} a_{m,i}, \quad n \neq m.$$

Let us define $T_n = \Phi^{-1}(P_n), n = 1, \ldots, N$, where P_n are the default probabilities and Φ is the cumulative distribution function of the normal standard distribution. Now, conditional on a realization of the systematic risk factors $x = (x_1, \ldots, x_d)$, the default indicators $D_1, \ldots, D_n$ are independent and, using (5.2), the conditional probability of default for the obligor n is given by,

$$\begin{aligned} P_n(x) &= \mathbb{P}(D_n = 1 | X = x) = \mathbb{P}(r_n < T_n | X = x) = \mathbb{P}\left(\epsilon_n < \frac{T_n - \sum_{k=1}^{d} a_{n,k} x_k}{b_n} \right) \\ &= \Phi\left(\frac{T_n - \sum_{k=1}^{d} a_{n,k} x_k}{b_n} \right), \quad n = 1, \ldots, N. \end{aligned}$$

In order to avoid Monte Carlo simulation for the entire portfolio, we consider the WA approximation for credit portfolios developed in Section 3.3, using the multi-factor setting instead of the one-factor model. The only difference now, is that the approximation (3.10) for computing the unconditional moment generating function $\tilde{M}_{\mathcal{L}}(s)$ is a d-dimensional integral and it would be too computationally intensive to apply the Gauss-Hermite method used for the one dimensional case. Concretely, we have to calculate,

$$\tilde{M}_{\mathcal{L}}(s) = \int_{\mathbb{R}^d} \prod_{n=1}^{N} \left[1 - P_n(x) + P_n(x)e^{-sE_n}\right] \frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{\|x\|^2}{2}}\, dx, \tag{5.3}$$

where $\|\cdot\|$ represents the euclidean norm in $\mathbb{R}^d$.

We observe that although Monte Carlo methods are desirable in high dimension, a large number n of integrand evaluations can be required since the expected error decreases as $n^{-\frac{1}{2}}$. Instead, we propose quasi-Monte Carlo methods, which evaluate the integrand at deterministic points in contrast with MC methods, which evaluate the integrand at random points. The deterministic points belong to *low discrepancy* sequences which, roughly speaking, are uniformly spread as we have seen in the previous Section.

The Koksma-Hlawka inequality (5.1) states that low discrepancy sequences yields a worst case error for multivariate integration bounded by a multiple of $(\log n)^d/n$, where n is the number of evaluations and d is the dimension of the integrand. For d fixed and n large, the error $(\log n)^d/n$ beats the MC error $n^{-\frac{1}{2}}$, but for n fixed and d large, the $(\log n)^d/n$ factor looks ominous. Then, it was believed that QMC methods should not be used for high dimensional problems ($d = 12$ was considered high). However, several authors have tested QMC methods for high dimensional problems arising in finance and, although the results are empirical, QMC methods consistently beat MC methods ([Pas95],[Pap96]).

High dimensional problems in market risk management have been address using the generalized Faure sequence by [Pap99], which concludes that QMC retains its superior performance for modest n an large d often used in risk management. For this reason, we test our WA approximation by means of computing the multidimensional integral (5.3) with both, the generalized Faure sequence and a Monte Carlo method. As we did with the one-factor model, a plain Monte Carlo simulation for the entire portfolio with 5 million scenarios will serve us as a benchmark.

Making the change of variable $x_j = \Phi^{-1}(t_j), j = 1, \ldots, d$, expression (5.3) reads,

$$\tilde{M}_{\mathcal{L}}(s) = \int_{[0,1]^d} \prod_{n=1}^{N} \left[1 - \tilde{P}_n(t) + \tilde{P}_n(t)e^{-sE_n}\right] dt,$$

where,

$$\tilde{P}_n(t) = \Phi\left(\frac{T_n - \sum_{k=1}^{d} a_{n,k}\Phi^{-1}(t_k)}{b_n}\right), \quad n = 1, \ldots, N.$$

Finally, $\tilde{M}_{\mathcal{L}}(s)$ can be approximated by,

$$\tilde{M}_{\mathcal{L}}^{(l)}(s) = \frac{1}{l}\sum_{i=1}^{l} \prod_{n=1}^{N} \left[1 - \tilde{P}_n(t_i) + \tilde{P}_n(t_i)e^{-sE_n}\right], \tag{5.4}$$

where,

$$\tilde{P}_n(t_i) = \Phi\left(\frac{T_n - \sum_{k=1}^{d} a_{n,k}\Phi^{-1}(t_{i,k})}{b_n}\right), \quad n = 1, \ldots, N,$$

and $t_i = (t_{i,1}, \ldots, t_{i,d}) \in [0,1]^d, i = 1, \ldots, l$ are l consecutive terms of the generalized Faure sequence.

Let us consider the following test[2] portfolio:

Portfolio 5.1. *This portfolio has $N = 100$ obligors, with $d = 25$, $P_n = 0.1$, $a_{n,k} = \frac{1}{\sqrt{26}}, b_n = \frac{1}{\sqrt{26}}, k = 1, \ldots, d, n = 1, \ldots, N$ and exposures,*

$$E_n = \begin{cases} 25, & n = 1, ..., 20, \\ 16, & n = 21, ..., 40, \\ 9, & n = 41, ..., 60, \\ 4, & n = 61, ..., 80, \\ 1, & n = 81, ..., 100. \end{cases}$$

The set of parameters used for the Wavelet Approximation method are: scale $m = 10$, 2^m intervals for the trapezoidal quadrature in the coefficients formula (4.4) and $r = 0.9995$. The left plot of Figure 5.1 shows the WA approximation with a MC method with $l = 20000$ to compute (5.4), while the right plot of Figure 5.1 represents the WA approximation with a QMC method with the generalized Faure sequence and $l = 20000$ to compute (5.4). For this example, the QMC approach seems to fit better than the MC method in the tail of the distribution. Anyway, this behavior must be further investigated with different sized portfolios and with other low discrepancy sequences.

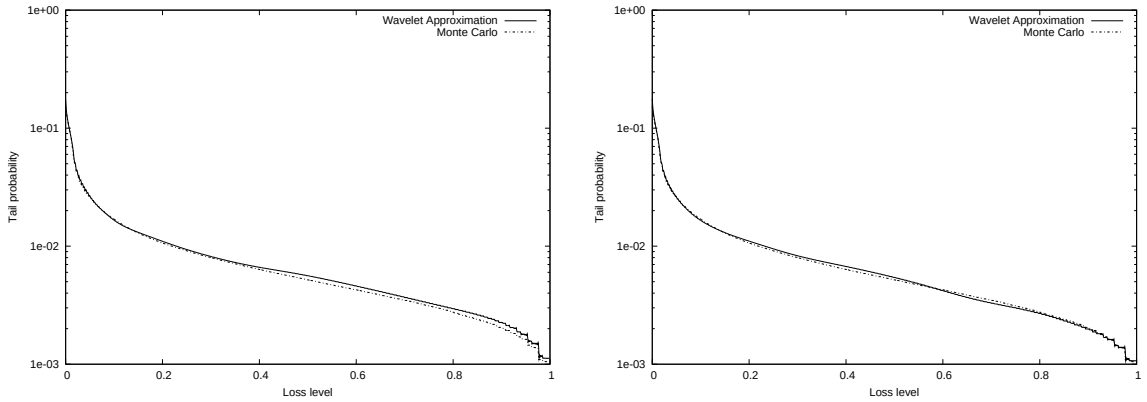

Figure 5.1: Tail probability approximation of Portfolio 5.1 with the WA method using a MC method (left) and a QMC method (right).

The risk measures and contributions can also be calculated once we have computed the loss distribution by means of the WA approximation with several factors. Although

[2]Computations have been carried out sequentially in a personal computer Dell Vostro 320 under GNU/Linux OS, Intel CPU Core 2 E7500, 2.93GHz, 4GB RAM and using the gcc compiler with optimization level 2.

the approximation of multidimensional integrals with MC or QMC methods can be time consuming, the direct estimation of the risk contributions is much more computationally intensive with plain Monte Carlo simulation. Moreover, in contrast with MC simulation, the Wavelet Approximation allows us to have an explicit expression of the cumulative distribution function. An adaptive method over the unit cube should also be investigated, comparing the convergence of the adaptive method with the convergence of the method used directly to solve the problem.

Chapter 6

Conclusions and Future Research

6.1 Conclusions

In this dissertation we have investigated the credit risk measurement of a credit portfolio under the asset-value model framework.

We have designed and implemented a new numerical method for inverting the Laplace transform based on Haar wavelets basis. This method is particularly well suited for stepped-shape functions, often arising in discrete probability models. Haar wavelets are capable to approximate a step function with just two coefficients while Fourier series with lot of terms still produces poor approximations at the jump point.

This inversion method have been successfully applied to the calculation of the risk measures VaR and Expected Shortfall in a credit portfolio under the one-factor Merton model. Many numerical tests have been performed, showing that the Wavelet Approximation is capable to deal with concentrated or small portfolios at high loss levels, where the ASRF model tends to fail. Moreover, the indirect Saddle Point approach has been also implemented and compared with the Wavelet Approximation. The results show that WA is very accurate and fast even when the credit portfolio contains a high degree of name concentration, while the Saddle Point underestimates the risk at certain loss levels and overestimates the risk at other loss levels. It is important to highlight that the WA approach computes the entire distribution of losses without extra computational time, while Saddle Point only calculates the quantile of the loss distribution function at the desired confidence level.

Furthermore, we have accurately computed the risk contributions to the VaR and the risk contributions to the ES with the WA method. The small relative errors obtained with respect to Monte Carlo methods in a wide variety of credit portfolios confirms the power of our method. An adaptive integration method for the Gauss-Hermite quadrature has been also implemented, with an important reduction of the amount of time needed when calculating the risk measures and contributions. VaR and ES can be calculated for very big portfolios with a personal computer in a few seconds, with a relative error less than 1%.

The WA method has been extended to the multi-factor case by means of a quasi-Monte Carlo method for computing high dimensional integrals, showing again accurate

results when benchmarking with a plain Monte Carlo method. However, more tests must be provided in order to assess whether QMC methods outperform MC methods for high dimensional integrals.

6.2 Future Research

Further extensions of the Wavelet Approximation method can be investigated in order to accommodate stochastic LGD instead of constant LGD. Moreover, empirical analysis have shown that the distribution of the LGD often exhibits a bimodal shape. This characteristic is intuitively reasonable when considering two possible developments of a defaulted loan. First, the obligor may recover and continues the contractual repayment covenants. This results in a very small loss amount basically driven by administrative costs. Secondly, the obligor may not recover which results generally in a higher loss amount. In accordance with the bimodal shape that the empirical results exhibit, the LGD modeling can be carried out by means of a bimodal model, considering a mixture of two beta distributions. Bimodal distributions are constructed by mixing canonical distributions under some certain assumptions. Since the LGD is in general a value between zero and one, it is reasonable to use distributions with a bounded support, e.g., beta distributions on the unit interval. The generation of the density function of a random variable following a multi-mode distribution resulting from a mixture of random variables is just the convex combination of the densities of these random variables.

Empirical findings have revealed certain dependence between the PD and the LGD parameters as highlighted by [Gie06] among others authors, so this fact should also be integrated in the WA method. It is now well understood that LGD is positively correlated to the default rate, in other words, LGD is high when the default rate is high, which suggests that there is also systematic risk in LGD, just like in the default rate parameter. Then, it would be worth to account for the influence of the systematic factors on the LGD somehow.

We should include in our models the effect that financial distress, initially affecting only a single obligor or a small group of obligors, can spread to a large part of the portfolio or even to the whole economy. Interactions which can be due to direct business links between firms, such as borrower-lender relationship, can provide a channel for the spread of financial distress within an economic system or a portfolio. We refer to this source of default correlation risk as default contagion, as pointed out in Chapter 1.

One drawback in the Merton model is that default can occur only at the maturity T of the bond and then, it is limited to the default-only mode. Regarding this point, it would be worth to incorporate credit migrations that allow us to observe how the probability of default of an obligor evolves. Credit migration can also be accommodated in a Merton-type model as noticed by [Lut09]. We can consider a firm which has been assigned to some rating category at time zero. The time horizon is fixed to T, and we assume that the transition probabilities of a firm are available for all rating grades (transition matrices can be estimated form empirical default data). The transition probability denotes the probability that the firm belongs to a certain rating class at the time of maturity T. In particular, one rating class among these rating grades denotes the default probability PD

of the firm. This important issue should be also investigated.

Finally, compared to a model with normally distributed asset returns as the Merton model, using a distribution with fatter tails as, for example, the Variance Gamma (VG) distribution, leads to an increase in the economic capital of the portfolio. We should assess the impact in the economic capital buffer substituting the normally distributed factors in the Merton model by VG distributed ones, or more generally speaking, by a Lévy model.

Bibliography

[Aba96] J. Abate, G.L. Choudhury and W. Whitt (1996). On the Laguerre method for numerically inverting Laplace transforms. *Journal on Computing*, v. 8, 413-427.

[Aba00] J. Abate, G.L. Choudhury and W. Whitt (2000). An introduction to numerical transform inversion and its application to probability models. *Computational Probability*, ed. W.K. Grassman, Kluwer, Norwell, MA.

[Ahn03] J. Ahn, S. Kang and Y. Kwon (2003). A flexible inverse Laplace transform algorithm and its application. *Computing*, v. 71, n. 2, 115-131.

[Alb07] H. Albrecher, S. Ladoucette and W. Shoutens (2007). A generic one-factor Lévy model for pricing synthetic CDOs. *Advances in Mathematical Finance*, R. J. Elliott et al. (eds.), Birkhäuser.

[And03] L. Andersen, J. Sidenius and S. Basu (2003). All your hedges in one basket. *RISK*, November, 67-72.

[Art97] P. Artzner, F. Delbaen, J.M. Eber, D. Heath (1997). Thinking coherently. *RISK*, November, 68-71.

[Art99] P. Artzner, F. Delbaen, J.M. Eber, D. Heath (1999). Coherent measures of risk. *Mathematical Finance*, v. 9, n. 3, 203-228.

[BCBS01] Basel Committee on Bank Supervision (2001). The new Basel capital accord, second consultative document. *Bank for International Settlements.*

[BCBS04] Basel Committee on Bank Supervision (2004). International convergence of capital measurement and capital standards, a revised framework. *Bank for International Settlements.*

[Bla73] F. Black, M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, v. 81, n. 3, 637-654.

[Blu03] C. Bluhm, L. Overbeck, C. Wagner (2003). Credit risk modeling. *Chapman and Hall/CRC.*

[Chen01] C.-F. Chen, C.-H. Chen, J.-L. Wu (2001). Numerical inversion of Laplace transform using Haar wavelet operational matrices. *IEEE transactions on circuits and systems* v. 48, n. 1, 120-122.

[Coh03] A. Cohen (2003). Numerical analysis of wavelet methods. *Elsevier*.

[Dan87] H. Daniels (1987). Tail probability approximations. *International Statistical Review*, v. 55, n. 1, 37-48.

[Dau92] I. Daubechies (1992). Ten lectures on wavelets. *CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM*.

[Dyk01] P. Dyke (2001). An introduction to Laplace transforms and Fourier series. *Springer*.

[Gao03] H. Gao, J. Wang and W. Zhou (2003). Computation of the Laplace inverse transform by application of the wavelet theory. *Communications in Numerical Methods in Engineering*, v. 19, n. 12, 959-975.

[Gao05] H. Gao, J. Wang (2005). A simplified formula of Laplace inversion based on wavelet theory. *Communications in Numerical Methods in Engineering*, v. 21, n. 10, 527-530.

[Gie06] G. Giese (2006). A saddle for complex credit portfolio models. *RISK*, july, 84-89.

[Gla05] P. Glasserman and J. Li (2005). Importance sampling for credit portfolios. *Management Science*, v. 51, n. 11, 1643-1656.

[Gla07] P. Glasserman and J. Ruiz-Mata (2007). Computing the credit loss distribution in the Gaussian copula model: a comparison of methods. *Journal of Credit Risk*, v. 2, n. 4, 33-66.

[Gor03] M. Gordy (2003). A risk-factor model foundation for ratings-based bank capital rules. *Journal of Financial Intermediation*, v. 12, n. 3, 199-232.

[Gou00] C. Gouriéroux, J.P. Laurent and O. Scaillet (2000). Sensitivity analysis of values at risk. *Journal of empirical finance*, v. 7, n. 3-4, 225-245.

[Gor07] M. Gordy and E. Lütkebohmert (2007). Granularity adjustment for Basel II. *Discussion paper series 2: banking and financial studies, n. 01/2007, Deutsche Bundesbank*.

[Hav09] E. Haven, X. Liu, C. Ma and L. Shen (2009). Revealing the implied risk-neutral MGF from options: The wavelet method. *Journal of Economics Dynamics & Control*, v. 33, n. 3, 692-709.

[Hoo82] F. de Hoog, J. Knight and A. Stokes (1982). An improved method for numerical inversion of Laplace transforms. *SIAM Journal on Scientific and Statistical Computing*, v. 3, n. 3, 357-366.

[Hua07a] X. Huang, C. W. Oosterlee and J. A. M. van der Weide (2007). Higher-order saddlepoint approximations in the Vasicek portfolio credit loss model. *Journal of Computational Finance*, v. 11, n. 1, 93-113.

[Hua07b] X. Huang, C. W. Oosterlee and M. A. M. Mesters (2007). Computation of VaR and VaR contribution in the Vasicek portfolio credit loss model. *Journal of Credit Risk*, v. 3, n. 3, 75-96.

[Jen95] J. L. Jensen (1995). Saddlepoint approximations. *Oxford statatistical science series.*

[Luc99] A. Lucas, P. Klaassen, P. Spreij and S. Staetmans (1999). An analytic approach to credit risk of large corporate bond and loan portfolios. *Journal of Banking and Finance*, v. 25, n. 9, 1635-1664.

[Lug80] R. Lugannani and S. Rice (1980). Saddlepoint approximations for the distribution of the sum of independent random variables. *Advances in applied probability*, v. 12, n. 2, 475-490.

[Lut09] E. Lütkebohmert (2009). Concentration risk in credit portfolios. *Springer.*

[Mal98] S. Mallat (1998). A wavelet tour of signal processing. *Academic Press, New York.*

[Mar01a] R. Martin, K. Thompson and C Browne (2001). Taking to the saddle. *RISK*, June, 91-94.

[Mar01b] R. Martin, K. Thompson and C Browne (2001). How dependent are defaults. *RISK*, July, 87-90.

[Mar01c] R. Martin, K. Thompson and C Browne (2001). VaR: who contributes and how much?. *RISK*, August, 99-102.

[Mar04] R. Martin (2004). Credit portfolio modeling handbook. *Credit Suisse.*

[Mar06] R. Martin and R. Ordovás (2006). An indirect view from the saddle. *RISK*, October, 94-99.

[Mas11] J. J. Masdemont and L. Ortiz-Gracia (2011). Haar wavelets-based approach for quantifying credit portfolio losses. *Quantitative Finance*, DOI: 10.1080/14697688.2011.595731.

[Mer74] R. Merton (1974). On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance*, v. 29, n. 2, 449-470.

[Mey97] Y. Meyer (1997). Wavelets and fast numerical algorithms. *Handbook of Numerical Analysis, v. 5., Elsevier.*

[Nie92] H. Niederreiter (1992). Random number generation and quasi-Monte Carlo methods. *CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM*, n. 63.

[Ort11] L. Ortiz-Gracia and J. J. Masdemont (2011). Credit risk contributions under the Vasicek one-factor model: a fast wavelet expansion approximation. Submitted to *Journal of Computational Finance*, available online at *www.crm.cat.*

[Pap99] A. Papageorgiou (1999). Deterministic simulation for risk management. *The Journal of Portfolio Management*, v. 25, n. 5, 122-127.

[Pap96] A. Papageorgiou and J. F. Traub (1996). Beating Monte Carlo. *RISK*, September, 63-65.

[Pas95] S. H. Paskov, J. F. Traub (1995). Faster valuation of financial derivatives. *The Journal of Portfolio Management*, v. 22, n. 1, 113-123.

[Sch88] J. L. Schiff (1988). The Laplace transform: theory and applications. *Springer*.

[Tak08] Y. Takano and J. Hashiba (2008). A novel methodology for credit portfolio analysis: numerical approximation approach. Available online at *www.defaultrisk.com*.

[Tas00a] D. Tasche (2000). Risk contributions and performance measurement. Available online at *http://www-m4.ma.tum.de/pers/tasche/*.

[Tas00b] D. Tasche (2000). Conditional expectation as quantile derivative. Available online at *http://www-m4.ma.tum.de/pers/tasche/*.

[Tez95] S. Tezuka (1995). Uniform random numbers: theory and practice. *Kluwer Academic Publishers, Boston*.

[Wal02] D. F. Walnut (2002). An introduction to wavelet analysis. *Birkhäuser*.

[Wil01] T. Wilde (2001). Probing granularity. *RISK*, August, 103-106.