Universitat Autònoma de Barcelona

Wavelet Approach in Computational Finance

Gemma COLLDEFORNS PAPIOL

Wavelet Approach in Computational Finance

Wavelet Approach in Computational Finance

PhD THESIS

to obtain the degree of Philosophy Doctor in Mathematics
of the Universitat Autònoma de Barcelona,

by

Gemma COLLDEFORNS PAPIOL

Barcelona, January 2018

This thesis was supervised by:
Dr. L. Ortiz Gracia (UB),
co-supervised by:
Prof. C. W. Oosterlee (TUDelft and CWI),
and tutored by:
Dr. A. Corral Cano (CRM and UAB)

This thesis was developed at:
Centre de Recerca Matemàtica
together with:
Department of Mathematics at the
Universitat Autònoma de Barcelona

This research was partially supported by La Caixa foundation in the framework of the training program on Collaborative Mathematics presented by the Centre de Recerca Matemàtica and by the grant FI-DGR 15 of AGAUR.

Wavelet Approach in Computational Finance
Dissertation at Universitat Autònoma de Barcelona

Reserve your right to think,
for even to think wrongly is better than not to think at all.

Hypatia of Alexandria

Acknowledgements

This dissertation concludes my PhD research at Centre de Recerca Matemàtica (CRM) and Universitat Autònoma de Barcelona (UAB) from October 2014 to February 2018. It is a real satisfaction to have succeeded on its completion and with a lot of pleasure I would like to acknowledge those who have contributed in one way or another to this thesis.

To begin with, I would like to thank Luis for his excellent supervision. It has been three and a half years working together, first at CRM and then at UB, having a nice guidance all through the thesis. I take this period of time as a great experience in which I have learned a lot from him about finance, mathematics and life. My most sincere gratitude also goes to Kees, who has kindly received me in all my visits at the Dutch mathematics and informatics research centre (CWI), and has been giving me nice advice in all the work done in site or in the distance. I really appreciate Kees support any time I got stuck, asking, when necessary, help to his acquaintances. That is why I also want to thank his retired and invaluable friends, Nico Temme and Piet Hemker, for their contributions when dealing with some integrals in the Chebyshev wavelets method development. I would like to emphasize that both Luis and Kees, apart from being exceptional in their work, are also really good people who made me feel comfortable, integrated and valued.

My PhD research has been carried out mainly at CRM where I had the chance to meet extraordinary people. I would like to thank the administration people of the centre for all the tasks they have developed and the friendly atmosphere that they provided. But most of all, I am grateful to all the researchers in there with whom I have shared seminars, talks and of course some non-job related activities which have been really useful to see things clearer later. It is really convenient in order to produce nice work to have an adequate, friendly and supportive environment. Specially, I would like to notice the great people I have shared office with, for listening and helping each other in all senses, my most sincere appreciation goes to Núria, Helena, Claudia, Victor, Marc; also to the other PhD students of the centre with whom I have shared daily live, Marina, Genis, Bernat, Alberto, Carmelo, Carles, Nestor, Anel; and to the ones that are already gone, Isa, Dani, Roberto and Vinnie. Furthermore, I would like to express my great gratitude to Stef because during his research stay at CRM we had very productive conversations and blackboard times about wavelets, COS and SWIFT.

During the last year of the PhD, because of Luis change of workplace, I have been researching from Universitat de Barcelona (UB). First, I would like to thank the research group "Risc en finances i assegurances" for providing me with an office. During my time at UB two people have been really important there. On the one hand Emma, who has shared with me lots of working time moments, and who has been by my side in good and bad moments, I really appreciate our nice friendship. And on the other hand, Alvaro, who is one of the most positive person I have ever met and because of that not only has he converted all my problems into challenges, but he has been very supportive at the finalization of the thesis too.

The PhD is enrolled in the UAB programme and because of that I would like to thank the UAB mathematics administration people for the tasks they have carried out.

Since I have been co-supervised by Kees in the Netherlands, I had the chance to visit CWI in some occasions. Nada, the secretary there, has been the nicest person one could be welcomed by, being exceptional at her job but also absolutely friendly. At CWI I had the opportunity to share work and also funny activities with the researchers. My most sincere thanks go to Kees students and collaborators Paco, Prashant, Qian, Ki Wai, Anton, and also to all of the scientific computing group, specially Sangeetika, who since the first day in the Netherlands helped me a lot to feel integrated, also Willem Jan, Krysztof, Nick and some others. I would like to specially mention Zaza, who deeply introduced me to the Dutch culture, and who has been a remarkable friend since the moment we meet in the Leiden conference.

I think it is quite important the people you meet in conferences, because you share PhD experiences and support each other, talking about how important or not are the speakers and coming to listening to your talks and getting interested in what you do. So, I express my gratitude to all of them too.

I would also like to thank the members of my defence committee for reading my dissertation and participating in my defence.

In addition, there are also the people who were already in my life but have made the PhD period easier. Although being 1200 km away, the support of my boyfriend, Francesc, has always been with me. He, as a good mathematician, has tried to help in all my weird or absurd mathematical questions, and he has cared to help as much as he has been able to. Nevertheless, he has been genuinely lucky to watch, mainly through Skype, the previews of all the presentations I had to give. However, the most important aspect of his contributions has been in the personal level, showing his faith in me and keeping me calm any moment I needed.

I am also grateful to my parents, who always believed in me, and provided me with all the necessary means to be here. Also to my brother and his partner; despite being lawyers, they have showed great interest on my research. Family support is always helpful and necessary.

Moreover, I would like to thank my friends who have been next to me during this time, providing support and sharing activities to disconnected from the working live. It would have not been the same without my friends from Vilanova nor without my friends from the mathematics degree.

Finally, I thank all sources that provided funding in order to carry out the PhD and its related activities: La Caixa, Agaur, CRM and the Spanish ministry projects.

Summary

In the computational finance world both derivatives pricing and risk management have attracted lots of interest amongst practitioners and academia. This PhD thesis aims to provide wavelets based techniques to enhance some of the methodologies used in the mentioned areas. Wavelets are families of functions that can be arbitrarily translated and dilated in order to generate orthogonal basis of $L^2(\mathbb{R})$. In relation to them, a collection of Fourier inversion methods has emerged; they are based on the approximation of functions by projecting on the wavelets basis such that the coefficients of the expansion are expressed by means of the Fourier transform of the function to approximate.

The SWIFT (Shannon wavelet inverse Fourier technique) method for pricing European-style options on one underlying asset was recently published and presented as an accurate, robust and highly efficient technique based on Shannon wavelets. One of the achievements of the thesis is the extension of the method to higher dimensions by pricing exotic option contracts, called rainbow options, whose payoff depends on multiple assets. The multidimensional extension inherits the properties of the one-dimensional method, being the exponential convergence one of them. Thanks to the nature of local Shannon wavelets basis, we do not need to rely on a-priori truncation of the integration range, we have an error bound estimate and we use fast Fourier transform (FFT) algorithms to speed up computations. We test the method for several examples comparing it with state-of-the-art methods found in the literature.

When managing the risk, regulators measure the risk exposure of a financial institution to determine the amount of capital that the institution must hold as a buffer against unexpected losses. The Basel Committee on Banking Supervision (BCBS) is the committee of the world's bank regulators. BCBS has recently set out the revised standards for minimum capital requirements for market risk, it has focused, among other things, on the two key areas of moving risk measures from Value-at-Risk (VaR) to Expected Shortfall (ES) and considering a comprehensive incorporation of the risk of market illiquidity. Another goal of this thesis is the presentation of a novel numerical method based on SWIFT to compute the VaR and ES of a given portfolio within the stochastic holding period framework to take into account liquidity issues. Two approaches are considered: the delta-gamma approximation, for modelling the change in value of the portfolio as a quadratic approximation of the change in value of the risk factors, and some of the state-of-the-art stochastic processes for driving the dynamics of the log-value change of the portfolio like the Merton jump-diffusion model and the Kou model.

Credit risk is the risk of losses from the obligor's failure to honour the contractual agreements and it is usually the main source of risk in a commercial bank. In this thesis, we also investigate the challenging problem of estimating credit risk measures of portfolios with exposure concentration under the multi-factor Gaussian and multi-factor t-copula models. It is well-known that Monte Carlo (MC) methods are highly demanding from the computational point of view in these situations. To overcome this issue, we present efficient and robust numerical techniques based on the Haar wavelets theory for recovering the cumulative distribu-

tion function of the loss variable from its characteristic function. The analysis of the approximation error and the results obtained in the numerical experiments section show a reliable and useful machinery for credit risk capital measurement purposes in line with Pillar II of the Basel Accords.

Resum (Catalan summary)

En el món de les finances computacionals, tant els preus de derivats com la gestió de riscos han atret molt d'interès entre els professionals i l'acadèmia. Aquesta tesi pretén proporcionar tècniques basades en ondetes a fi de millorar algunes de les metodologies utilitzades en aquestes àrees. Les ondetes són famílies de funcions que es poden traslladar i dilatar arbitràriament generant bases ortogonals de $L^2(\mathbb{R})$. En relació amb elles, ha sorgit una col·lecció de mètodes d'inversió de Fourier que es basen en l'aproximació de funcions a partir de la projecció en la base d'ondetes, de tal manera que els coeficients de l'expansió s'expressen mitjançant la transformada de Fourier de la funció a aproximar.

El mètode SWIFT (*Shannon wavelet inverse Fourier technique*) per valorar opcions d'estil europeu en un subjacent s'ha publicat i presentat recentment com una tècnica precisa, robusta i altament eficient basada en les ondetes de Shannon. Motivats per la valoració d'un tipus d'opcions exòtiques anomenades opcions arc iris, que depenen de diversos actius, un dels objectius de la tesi és l'extensió multidimensional del mètode SWIFT. Aquesta extensió hereta les propietats del mètode unidimensional i la convergència exponencial n'és una. Gràcies a la naturalesa de les ondetes de Shannon, no és necessari un truncament a priori del rang d'integració, tenim una estimació de l'error i fem ús d'algorismes de la transformada ràpida de Fourier per accelerar els càlculs. El mètode es prova a partir de diversos exemples comparant-lo amb mètodes recents en la literatura.

En l'àmbit de gestió de risc, són els reguladors els que s'encarreguen de determinar la quantitat de capital que les entitats financeres han de guardar per estar preparades per suportar pèrdues inesperades. És l'anomenat comitè de Basilea el que s'encarrega dels reguladors bancaris mundials. Recentment, el comitè ha revisat els estàndards establerts pel capital mínim per risc de mercat i, entre altres canvis, suggereix d'una banda moure l'estès ús de la mesura de risc VaR cap a la mesura ES, i l'altra banda, considerar la incorporació del risc derivat a causa de la no-liquiditat del mercat. És per això, que un segon objectiu de la tesi és presentar un nou mètode numèric basat en SWIFT per als càlculs del VaR (*Value-at-Risk*) i l'ES (*Expected Shortfall*) tenint en compte un horitzó de temps estocàstic per tenir en compte els problemes de liquiditat del mercat. Es consideren dos enfocaments diferents: l'aproximació delta-gamma que aproxima quadràticament el valor de la cartera en el canvi de valor dels factors de risc, i també l'aproximació del valor logarítmic de la cartera a partir de processos estocàstics coneguts com ara el Merton o el Kou.

El risc de crèdit és el risc de pèrdues per part del deutor en cas que no compleixi els acords i en general és la principal font de risc en un banc comercial. En aquesta tesi, investiguem el problema de calcular les mesures de risc de crèdit de les carteres sota els models multifactorials Gaussians i de t-còpula. És ben sabut que els mètodes de Montecarlo, que són els que generalment s'empren, són molt exigents des del punt de vista computacional en aquestes situacions. Per millorar aquests problemes, a la tesi presentem tècniques numèriques eficients i robustes basades en les ondetes de Haar per recuperar la funció de distribució acumulada de la variable de pèrdues a partir de la seva funció característica. L'anàlisi de l'error i els resul-

tats obtinguts en l'apartat d'experiments numèrics mostren una maquinària fiable i útil per al mesurament de capital de risc de crèdit d'acord amb el Pilar II dels Acords de Basilea.

Contents

Acknowledgements vii

Summary ix

Resum (Catalan summary) xi

List of most used abbreviations xv

1 Introduction 1
- 1.1 Objectives . . . 5
- 1.2 Outline . . . 7

2 Wavelet Theory 9
- 2.1 Introduction . . . 9
- 2.2 Mathematical overview . . . 10
- 2.3 Fourier analysis . . . 16
  - 2.3.1 Fourier series . . . 16
  - 2.3.2 Fourier and inverse Fourier transforms . . . 17
  - 2.3.3 Discrete Fourier transform . . . 18
- 2.4 Wavelets . . . 20
  - 2.4.1 Multiresolution analysis . . . 21
- 2.5 Examples of wavelet functions . . . 24
  - 2.5.1 Haar wavelets . . . 24
  - 2.5.2 Shannon wavelets . . . 25
  - 2.5.3 Chebyshev wavelets . . . 26
- 2.6 Inversion methods . . . 30
  - 2.6.1 Haar wavelet approximation (WA) . . . 31
  - 2.6.2 Shannon wavelet inverse Fourier technique (SWIFT) . . . 32
  - 2.6.3 Chebyshev wavelet approximation (CWA) . . . 34
- 2.7 Conclusions . . . 36

3 A Novel Two-Dimensional Technique for Pricing European Options 37
- 3.1 Introduction . . . 37
- 3.2 European option pricing framework . . . 39
  - 3.2.1 Lévy processes . . . 40
  - 3.2.2 Numerical methods . . . 44
  - 3.2.3 One-dimensional option pricing wavelet methods . . . 45
- 3.3 Two-dimensional SWIFT for option pricing . . . 49
  - 3.3.1 Multidimensional wavelets framework . . . 50
  - 3.3.2 European rainbow option pricing . . . 53
  - 3.3.3 Error analysis and parameters selection . . . 57
  - 3.3.4 Numerical experiments . . . 63

3.3.5 Strengths of 2D-SWIFT . . . 70
3.4 Conclusions . . . 74

4 Computation of Market Risk Measures with Stochastic Liquidity Horizon 77
4.1 Introduction . . . 77
4.2 Basic concepts in risk management . . . 80
4.2.1 Risk measures . . . 80
4.2.2 Standard methods for market risks . . . 83
4.3 Computation of market risk measures by SWIFT . . . 84
4.3.1 Risk measures by SWIFT . . . 84
4.3.2 Error analysis and selection of parameters . . . 87
4.3.3 Numerical examples . . . 92
4.4 Conclusions . . . 102

5 Quantifying Credit Portfolio Losses under Multi-Factor Models 105
5.1 Introduction . . . 105
5.2 The basics of credit risk management . . . 107
5.2.1 Credit portfolio losses . . . 107
5.2.2 Factor models and their characteristic functions . . . 108
5.3 Quantifying credit portfolio losses by WA method . . . 113
5.3.1 Efficient computation of characteristic functions . . . 113
5.3.2 Risk measures computation by WA . . . 118
5.3.3 Numerical examples . . . 120
5.4 Conclusions . . . 126

6 Conclusions and Outlook 129
6.1 Conclusions . . . 129
6.2 Outlook . . . 131

References 133

List of most used abbreviations

ASRF	Asymptotic Single Risk Factor
BCBS	Basel Committee on Banking Supervision
BM	Brownian Motion
BS	Black-Scholes
CDF	Cumulative Distribution Function
CEO	Chief Executive Officer
ChF	Characteristic Function
CPU	Central Processing Unit
CWA	Chebyshev Wavelet Approximation
DFT	Discrete Fourier Transform
DWT	Discrete Wavelet Transform
ES	Expected Shortfall
FFT	Fast Fourier Transform
FMC	Full Monte Carlo
FX	Foreign-eXchange
GBM	Geometric Brownian Motion
GPU	Graphics Processing Unit
HL	Hermite-Laguerre
IDFT	Inverse Discrete Fourier Transform
JD	Jump-Diffusion
LGD	Loss Given Default
MC	Monte Carlo
MJD	Merton Jump-Diffusion
MRA	MultiResolution Analysis
MSE	Mean Squared Error
NLTI	Numerical Laplace Transform Inversion
PDE	Partial Differential Equation
PDF	Probability Density Function
PIDE	Partial Integro-Differential Equation
PMC	Partial Monte Carlo
PRDC	Power-Reverse Dual Swaps
QTA	Quadratic Transform Approximation
RE	Relative Error
SDE	Stochastic Differential Equation
SLH	Stochastic Liquidity Horizon
SWIFT	Shannon Wavelet Inverse Fourier Technique
VaR	Value-at-Risk
WA	(Haar) Wavelet Approximation

CHAPTER 1

Introduction

Computational finance [27] is a cross-disciplinary field between financial mathematics and numerical analysis. It deals with problems of practical interest in finance such as trading, hedging and investment decisions, as well as facilitating the risk management of those decisions. Areas where computational finance techniques are employed include: investment banking, corporate strategic planning, securities trading, financial risk management, derivatives trading, and investment management. Over the last 20 years, the field of computational finance has expanded into virtually every area of finance. Not only has the demand for practitioners grown dramatically, but also many specialized companies have grown up to supply computational finance software and services.

In this thesis, we focus our efforts in two applications of computational finance: derivatives pricing (specifically option pricing) and risk management (concretely market risk and credit risk). Both areas arouse interest amongst practitioners and academia alike. The reason is that while they are essential for the good functioning of the financial sector, both areas are so challenging that there is still a lot of room for innovation and improvement; and the mathematics and computational challenges behind them attracts the interest of many academics.

Option Pricing

In finance, a derivative is essentially an instrument whose value depends on the values of others, called underlying assets. They serve several purposes, their main one is to allow the banks to hedge risks from some investment. This is, it allows financial institutions to keep the losses bounded in case some investment goes wrong. The main types of derivative contracts seen within financial markets include forward contracts, futures, swaps and options; see [42] for a complete description of the instruments. The type of derivative we focus on in this thesis are options, which are a type of contracts that give the holder the right, but not the obligation to exercise the contract. There are two types of options: a call option, which gives the holder the right to buy the underlying asset by a certain date for a certain price, and a put option, which gives the holder the right to sell the underlying asset by a certain date for a certain price. Depending on when they can be exercised we find, for example, European options, which can be exercised only on the expiration date itself; or American options, which can be exercised at any point before their expiration date. These options are usually referred as plain vanilla options, since they are the most basic or standard version of a financial instrument. On the opposite side, we find the options we refer to as exotic options. Those options have a much more complicated structure than the standard vanilla options and may depend on more variables than just the price of a stock or have some extra conditions concerning under which circumstances the investor will receive a certain payoff. In this work, we focus on a class of exotic options namely the multidimensional options, also called multicolour rainbow

options, whose payoff depends on multiple assets.

For any financial player trading in options, it is essential to understand how to price them. However, option pricing meaning to be able to calculate the right value of the option (or the fair price of the contract) before maturity, is one of the big challenges in the financial world.

In 1973, Black and Scholes published a paper [10] providing a model for valuing European options. This is considered the starting point of option pricing theory which since then has become a main area of mathematical finance and experienced a huge development. The Black-Scholes (BS) model is stated under certain assumptions such as the option must be European, risk-free rate and volatility of the underlying are known and constant, and returns on the underlying asset follows a Geometric Brownian Motion (GBM) dynamics. In the BS model, the price of the option is obtained as the solution of a partial differential equation (PDE) that defines the evolution of the option price over time. The solution of this PDE can alternatively be written as a conditional expectation by means of the Feynman-Kac theorem, originally from [44]. This theorem shows that the conditional expectation of the value of a contract payoff function under the risk-neutral measure is equivalent to the solution of a partial differential equation and under the previous assumption, this partial differential equation is the BS model and hence provides us with a closed formula for its solution.

However, the analytical formula for an option price is generally unavailable. Hence, there is a need to develop numerical methods to approximate their prices and develop efficient algorithms to implement them, so that they provide useful information in a market that changes rapidly. The existing numerical methods for option pricing can be classified into three major groups: Monte Carlo (MC) simulation [13], partial-(integro) differential equation (PIDE or PDE) methods [43], and Fourier inversion methods [41]. The main advantage of Monte Carlo simulation is that complex and exotic derivatives can be treated easily, furthermore, options on several assets can also be handled straightforwardly, nevertheless, the main drawback of the method is the computational cost. Alternative PIDE methods consist in deriving and then solving numerically the partial integro-differential equation, the advantage of these methods is that complex and exotic payoffs can be treated easily; the limitations are the slower speed in comparison to inversion methods. Another choice of methods, which are the ones we use in the thesis, are the Fourier inversion techniques. As previously mentioned, from the numerical pricing partial differential equation, the corresponding solution can also be found by means of the Feynman-Kac formula as a discounted expectation of the option value at final time, the called risk neutral valuation formula. This class of methods is known for its computational efficiency; however, they rely on the assumption that the characteristic function of the process is known in order to recover the probability density function inside the risk neutral valuation formula, and unfortunately it is not always available.

RISK MANAGEMENT

Banks, brokers, portfolio managers and even the entire financial industry are exposed to risks on a daily basis. Due to this fact and especially since the financial crisis, the importance of risk management has increased rapidly. The fundamentals of risk management are not to completely eliminate risks, but to manage them accordingly. As financial markets grow, there is an increasing need to manage risks appropriately. In [57], concepts, techniques and tools for quantitative risk management are presented.

We can distinguish between three risk categories, the boundaries of which are not always clearly defined, nor do they form an exhaustive list of the full range of possible risks affecting

a financial institution. In banking, one of the categories is market risk, the risk of a change in the value of a financial position due to changes in the value of the underlying component on which that position depends, such as stock and bond prices, exchange rates, commodity prices and so on. Another important category of risk is credit risk, the risk of not receiving promised repayments on outstanding investments such as loans and bonds, because of the default of the borrower. A further risk category that has received a lot of recent attention is operational risk, the risk of losses resulting from inadequate or failed internal processes, people and systems, or from external events.

A regulator measures the risk exposure of a financial institution in order to determine the amount of capital that the institution has to hold as a buffer against unexpected losses. The Basel Committee on Banking Supervision (BCBS) is a committee of the world's bank regulators that meets regularly in Basel, Switzerland; see [5–8] for some of its statements on market risk. In 1988, the committee published what has become known as Basel I: this is an agreement between the regulators on how the capital a bank is required on hold for credit risk should be calculated. Later, the Basel Committee published the 1996 Amendment which was implemented in 1998 and required banks to hold capital for market risk as well as credit risk. That Amendment calculates capital for the trading book using the Value-at-Risk (VaR) measure with a horizon of 10 days and a confidence level of 99%. Basel I has been followed by Basel II, Basel II.5, and Basel III. Basel II, which was generally implemented in 2007, uses VaR with a one-year time horizon and a 99.99% confidence level for calculating capital for credit risk and operational risk. Basel II.5, implemented in 2012, revised the way market risk capital is calculated (stressed VaR is introduced). Basel III is increasing the amount of capital that banks are required to hold and the proportion of that capital that must be equity. In May 2012, the Basel committee issued a discussion paper indicating that it is considering switching from VaR to Expected Shortfall for market risk. It has also focused on a comprehensive incorporation of the risk of market illiquidity. Introducing either stochastic volatility assumptions or stochastic jump processes into modelling of risk factors will help to overcome these shortcomings. Another message of paramount importance is that the time it takes to liquidate a risk position varies, depending on its transactions costs, the size of the risk position in the market, the trade execution strategy and market conditions; some studies suggest that, for some portfolios, this aspect of liquidity risk could also be addressed by extending the VaR risk measurement horizon.

In this thesis, we focus specifically on two of the categories of risk: market risk and credit risk.

Market risk refers to the sensitivity of an asset or portfolio to overall market price movements such as interest rates, inflation, equities, currency and property. While market risk cannot be completely removed by diversification, it can be reduced by hedging. See [26] for a great coverage of the topic. As it is standard in practice, the value of the portfolio is modelled as a function of time and a multidimensional vector of risk factors. Depending on the model, different risk factors may be used, some common ones are logarithmic prices of financial assets, yields and logarithmic exchange rates. A well-known method for measuring market risk is the delta-gamma approximation [60], which assumes that the change in value of the portfolio is a quadratic function of the change in value of the risk factors.

Value-at-Risk, developed in JP Morgan [62] in the market risk framework, is a commonly used measure of risk. VaR calculates an expected loss amount that may not be exceeded at

a specified confidence interval over a given holding period, assuming normal market conditions. In plain words, this can be interpreted as the higher the portfolio's VaR, the greater the possible losses that the portfolio may suffer, so its exposure to market risks is higher. There are commonly used methodologies to calculate VaR in the market risk framework. One of the early methods is historical simulation, see [49], which calculates VaR from a distribution of historical returns. It can only reflect asset sensitivities to events captured in the time horizon used, but it is easy to calculate and understand. The parametric approach (also known as Variance/Covariance), see [56], calculates VaR typically assuming that returns are normally distributed and estimates it directly from the standard deviation of the portfolio returns; the strengths of this method are the simplicity of the calculations and the fact that the data for the inputs is very easy to obtain, nevertheless, the assumption of normality is the biggest weakness of the method since it usually does not correspond with the reality. Monte Carlo simulation [35] can calculate VaR from a distribution constructed from random outcomes. It is very computationally intensive since thousands of scenarios need to be run for each constituent asset in order to get an accurate solution. Nevertheless, if the assets are modelled accurately, which is not an easy task, they would give the most accurate measure.

Whilst VaR remains an important metric for measuring market risk exposure, there are limitations with this measure. VaR is an attractive measure because it is easy to understand. However, it is not a coherent risk measure, as defined in [4], since it fails the sub-additivity condition. Which means that it does not take into account the benefits of diversification. Moreover, it does not give an indication about the severity of losses beyond the computed quantile. Regulators are increasingly recommending a broader range of risk metrics to evaluate risk exposure. There are other risk metrics which can give a more in-depth understanding of a situation of the portfolio. For example, a measure that deals with those problems is Expected Shortfall (ES), presented in [4], also known as conditional VaR or tail loss. It averages all the expected losses greater than VaR; if VaR is calculated at a 99% confidence level, ES averages the worst 1% expected losses, thus ES is very useful where there is a high exposure to derivatives as the distribution may be highly skewed. ES is computationally very involved and methods used for its computations are the same as the ones presented for the VaR. Both VaR and ES are risk measures dependent on two parameters: the confidence level and the risk horizon.

Credit risk is the risk of losses from the obligor's failure to honour the contractual agreements. It is usually the main source of risk in a commercial bank. See [11] as a good reference on the area. When computing credit risk, it is usual to fix a time horizon, and to consider a credit portfolio consisting of a set of obligors. Every obligor is then characterized by three parameters: the exposure at default, the loss given default and the probability of default. The loss function of the portfolio is defined as the sum of individual credit losses. And the individual credit loss is the product of the exposure at default by the loss given default by a default indicator. To evaluate the portfolio loss distribution, a key issue is to model the various dependence effects, common practice is to use a so-called factor model of asset correlations. A popular example of this model is the Vasicek one-factor Gaussian copula model (it forms the basis of the Basel II approach), also its multi-factor version which accounts for sector concentration, or different copula models like the t-copula, which is capable of introducing tail dependence in credit portfolios. Credit risk management is computed by means of risk measures of the loss distribution, using for example the respective VaR and ES of credit portfolios.

WAVELETS

As already mentioned, Fourier inversion methods are a possible approach for both option pricing and risk management. Carr and Madan [15] popularized these kind of techniques. Inspired by their method, extensive research in the Fourier domain has emerged. Fourier methods depend on the availability of an expression for the characteristic function of the stochastic processes modelling the underlying assets, but if the characteristic function is available, they are generally the fastest methods amongst finite difference and Monte Carlo methods.

A state-of-the-art Fourier pricing method is the COS method of [28]. It is a fast method based on a cosine expansion of the probability density function firstly used for European option pricing. A full range of extensions of the method can be found: a two-dimensional extension [74], Asian options pricing [83], pricing early-exercise and discrete barrier options [29] to name just a few.

The first method that used wavelet expansions in this field was the WA (wavelet approximation) method [54], where Haar wavelets were used to recover the distribution function of credit risk losses and compute risk measures. More works related with Haar wavelets (also with B-splines) are [64], in the credit risk area [63], in market risk [66] and in the option valuation framework [46, 47, 65].

Shannon wavelets, which are based on the cardinal sine (sinc) function have also been used in the Fourier inversion context, which resulted in the method known as SWIFT (Shannon wavelet inverse Fourier technique) [67]. The SWIFT method has been used in option pricing showing great results in terms of accuracy and speed. Extensions of the method have been done for pricing early-exercise and discrete barrier options in [51], for Asian options in [82], and some other covered in this thesis.

Let us clarify what wavelets are. Important books devoted to wavelets are [12, 19, 25, 50]. In a rough sense, a wavelet is a function that looks like a wave, travels for one or more periods and is non-zero only over a finite interval instead of propagating forever the way sines and cosines do. Overall, they are functions that can be arbitrarily translated and dilated in order to generate basis. Wavelets theory can be seen as an extension of Fourier analysis in the sense that they both aim to represent complicated functions using sums of simple ones. Although wavelets were already used implicitly in mathematics, they were only given the status of a consolidated scientific field around three decades ago. Up until today, their most important contribution has been made in signal processing, an area with applications that range from image compression to the detection of gravitational waves, passing through improving the quality of medical imaging. However, plenty of other areas can benefit from them, amongst them financial computing in which they have become a very active area of research.

Given a family of wavelets, we have that they form an orthonormal basis of $L^2(\mathbb{R})$, thus functions in this space can be approximated in the wavelets basis. More efficient computations can be made in the special case in which the scaled and shifted wavelets form a multiresolution analysis (MRA), it is based on the existence of a family of subspaces of $L^2(\mathbb{R})$ satisfying certain properties, which is called MRA. Previously mentioned wavelet methods, WA and SWIFT come from wavelets that form an MRA.

1.1. OBJECTIVES

This thesis tackles three different topics in the computational finance field approached from a common point of view: the use of wavelets to implement Fourier inversion methods. The

different goals that have been pursued throughout this thesis are the following ones, exposed chronologically in accordance as how they appear in the memory.

One idea we had in mind was the use of a different wavelet family instead of Haar and Shannon. The main feature of Haar wavelets is their compact support; in contrast, Shannon wavelets are characterized by their continuity. For this reason, Haar is really appropriate to approximate piecewise continuous functions and Shannon for continuous densities. We thought about exploring Chebyshev wavelets [3] as basis for the inversion methods; note that Chebyshev wavelets definition comes straight forward from the Chebyshev polynomials. Regardless their shape, we thought about these wavelets because of the notably relation with the COS method: if the COS method is the cosine expansion of a function $f(x)$, this expansion matches the Chebyshev series expansion of $f(\cos(t))$, see [28]. To this end, we try to work with an inversion Fourier method based on Chebyshev wavelets in option pricing.

Another focus of the research is on the extensions of the applications of the SWIFT method in the financial area. One of the targets consists of extending the Shannon wavelets Fourier inversion methods for pricing a class of multidimensional exotic options, namely multi-asset options. Analytical formulae to price multi-asset derivatives are only available for the simplest cases and hence numerical solutions are necessary. However, numerical approximations are very expensive to obtain, despite the efforts in the financial community to come up with optimal methods. Some examples of multidimensional options are exotic option contracts called multicolour rainbow options whose payoff depends on multiple assets. One well-developed multidimensional Fourier-based method is the multidimensional COS method presented in [74] and called 2D-COS when the dimension is two. However, it may exhibit problems in the vicinity of the integration boundaries because of the periodic behaviour of cosines. In the one-dimensional case, local wavelets bases overcome some of the problems of the well-known one-dimensional COS method [28]. We aim to extend the one-dimensional SWIFT method to a higher dimension to be able to price European-style financial contracts with a payoff depending on more than one asset. Providing an error analysis which facilitates the choice of the parameters of the new method and enhance the overall speed with an FFT algorithm.

A further application to consider for the SWIFT method is in the market risk area. The goal is to present efficient and robust numerical techniques to address the computation of the VaR and ES risk measures of a market portfolio. Moreover, following the idea introduced in [14], we consider that the holding period follows a certain positive stochastic process to account for liquidity risk. We also try to go a step further than [14] by considering more realistic models for the log-value of the portfolio. On the one hand, we propose the use of the delta-gamma approach [60], where it is assumed that the change in portfolio value is a quadratic function of the changes in the risk factors. On the other hand, we consider the Merton jump diffusion (MJD) model [58] and the Kou model [48] to drive the log-return on the portfolio value. Under any of these scenarios, the closed formulae to compute the risk measures within the Gaussian setting in [14] are not available any more. However, the characteristic function of the change in (log-)value of the portfolio is known in closed form for most of the interesting processes in finance, in particular for the two models mentioned above. We therefore recover the density function from its Fourier transform and then we calculate the VaR and the ES values. Among the methods available in the literature for Fourier inversion, we choose the SWIFT method because an important feature of the present method is that the scale of approximation is es-

timated a priori by means of the characteristic function, and this makes this method really convenient in practice, as opposed to the numerical methods presented in [66] in which the number of terms has to be obtained by trial and error.

Finally, we consider the application of wavelets methods in the credit risk problem. As previously stated, Haar wavelets have already been used in this area in [54, 63]. However, they are only used under the one-factor Heston model, which forms the basis of the Basel II approach. Our objective here is to investigate the challenging problem of estimating credit risk measures of portfolios with exposure concentration under the multi-factor Gaussian and multi-factor t-copula models. We point out the study of multi-factor models because they can account for sector concentration and also t-copula models because provide better empirical fit to the observed data. To the best of our knowledge, this is the first time that multi-factor t-copula models are considered outside the MC framework. We desire to develop numerical techniques to contribute to the efficient measurement of VaR and ES values for small or big portfolios in the presence of exposure concentration under high-dimensional models. It is worth remarking that small and/or concentrated portfolios are particularly challenging cases, since asymptotic methods usually work out well for large and diversified portfolios. We aim to estimate the risk measures in a procedure composed of two main parts. The first part is the numerical computation of the characteristic function associated to the portfolio loss variable. The second part of the procedure consists of a Fourier inversion to recover its cumulative distribution function. For this purpose, we thought about using WA method and also SWIFT, but because of the shape of the wavelet we expect to see a better performance with WA because here the functions to approximate are piecewise constant.

1.2. OUTLINE

The thesis is organized as follows. Chapter 2 provides mathematical background that will be used later. It contains a brief introduction into functional analysis on which wavelets theory is based; then, we present Fourier analysis which is closely related to wavelets; and finally, wavelet theory itself. Three wavelet families are introduced: Haar, Shannon and Chebyshev. Finally, inversion methods based on these types of wavelets are stated.

With all the mathematical and wavelet background discussed, Chapters 3, 4 and 5 form the core of the thesis. In these three chapters, new applications of the wavelet inversion methods are presented.

In Chapter 3, we introduce the option pricing topic and discuss three wavelet European option pricing methods based on Haar, Shannon and Chebyshev wavelets. The main section of the chapter is Section 3.3 in which two-dimensional SWIFT for option pricing is introduced: first the multi-dimensional wavelet framework is presented, then the 2D-SWIFT pricing formula derivation with a rigorous study of the error analysis of the method, and a full set of numerical examples to show its strengths in comparison with 2D-COS.

Chapter 4 is devoted to market risk. It starts discussing the basics in risk measurement such as the risk measures VaR and ES. The key section in the chapter, Section 4.3, is devoted to explaining the methodology to calculate the VaR and ES risk measures by SWIFT. It includes a complete error analysis as well as the way to select the parameters to control the numerical method, and a wide variety of examples.

Credit risk is addressed in Chapter 5. It starts with the formulation of the credit portfolio losses problem as well as the description of the Gaussian and t-copula models for dependence

among obligors. Then, the derivations of the characteristic functions for all the models are presented with their efficient evaluation. It tackles the computations of VaR and ES using WA and some numerical examples to support the efficiency of the method are presented.

Finally, in Chapter 6 we draw conclusions and results on the whole thesis. Moreover, we give some recommendations for future research.

CHAPTER 2

Wavelet Theory

This chapter aims to provide the reader with the necessary mathematical background so that the thesis is self-contained.

Here, we introduce the basic theory related to wavelets. We start with a brief introduction of basic concepts, mainly of functional analysis on which wavelets theory is based on; then we present Fourier analysis which is closely related to wavelets; and finally, wavelet theory itself.

This chapter is mainly based on the books [12], [19], [25], [50], [72], [73], and [80].

2.1. INTRODUCTION

Wavelets theory can be seen as an extension of Fourier analysis in the sense that they both aim to represent complicated functions using sum of simple ones, although wavelet theory appeared much later than Fourier analysis. In the thesis, we combine tools provided by these two areas. Let us present first a brief introduction of their history and relevance.

Fourier analysis was introduced by Joseph Fourier in the early 19th century as a mean to study the heat equation. Since then, it has become a main topic both in pure mathematics, leading to the development of harmonic analysis which is a main field in mathematics, and in applications including physics, signal processing, optics and several more. The main tools of Fourier Analysis are the Fourier transform and the Fourier series, which are introduced later. The relevance of these tools in numerous areas of technology and science comes from the fact that they provide convenient formulae to express and approximate the complicated functions that they often encounter.

On the other hand, the development of wavelets has been much more recent. Although wavelets were already used implicitly in mathematics, physics, signal processing and numerical analysis, they were only given the status of a consolidated scientific field around three decades ago. Since then, their importance has increased to the point that one of the fathers of the theory, Yves Meyer, was awarded the Abel price in 2017 (more details in [33]). Wavelets attracted considerable interest from the mathematical community and from members of several disciplines in which they had promising applications. The reason being that, they are very interesting mathematical objects while being relatively simple to apply to different areas. Up until today, their most important contribution has been made in signal processing, an area with applications that range from image compression to the detection of gravitational waves, passing through improving the quality of medical imaging. However, plenty of other areas can benefit from them, amongst them financial computing in which they have become a very active area of research.

Wavelet analysis and classical Fourier analysis can be thought as brothers in any field of interest. In the Fourier case the target is to measure the local frequency content of a signal,

while in the wavelet case one is comparing several magnifications of a signal with distinct resolutions. The building blocks of Fourier analysis are sines and cosines (waves). In wavelet analysis, the blocks are the wavelets, which are no longer just sinusoidal functions; instead they are functions that can be arbitrarily translated and dilated in order to generate bases. Analogous to Fourier analysis, there are also two important mathematical entities in wavelet analysis: the integral wavelet transform and the wavelet series.

This second chapter of the thesis intents to be an introduction to the topic we work with during the whole thesis development: wavelet inversion methods. We start defining Hilbert spaces and we show the formula that we can use to approximate any function in these spaces using orthonormal bases. After that, we move to Fourier analysis. We show that Fourier series can approximate any function in an appropriate Hilbert space since the functions used in the Fourier series expansions form an orthonormal basis of this space. We also introduce the Fourier transform and its discrete version, which are very relevant for the methodology we use in the thesis. In this section, the fast Fourier transform (FFT) is also introduced, which is an algorithm used in our methods in order to accelerate the calculations. Then, we present wavelets and multiresolution analysis which is a structure generated by a certain type of wavelets. Wavelets, like Fourier expansions, approximate functions using orthonormal basis using the expansion presented in the beginning. The main difference with Fourier expansion is that one can choose between several different wavelet bases. The most convenient wavelet will be chosen depending on the kind of function one wants to approximate. Therefore, next we present the wavelets used throughout the thesis: the Haar, the Shannon and the Chebyshev wavelets of which the first two form a multiresolution analysis. Finally, we introduce the Fourier inversion methods. For each family of wavelets presented, we show how to implement the method with special attention devoted to being able to apply fast Fourier transforms in the computations. Fourier inversion methods for the Haar and Shannon wavelets are already known in the literature so our implementation builds on the references. On the other side, to the best of our knowledge, there is no prior work done in trying to formulate a Fourier inversion method using Chebyshev wavelets. We tried several different ideas to formulate the method in an efficient way and in the last section we present our most promising methodology.

This chapter is organized as follows. Section 2.2 is the mathematical overview to the topic, it is divided in basics on probability theory, norms, inner products and orthogonality. Next, we have Section 2.3 where Fourier analysis is presented, it consist of three different sections: Fourier series, Fourier transforms, and discrete Fourier transforms, which includes the fast Fourier transform algorithm. Then, we move to the wavelet world in Section 2.4, there we also present multiresolution analysis as well as we talk about the associated wavelet and the wavelet transform. Examples of three wavelets families (Haar, Shannon and Chebyshev) are given in Section 2.5. Finally, in Section 2.6 we introduce the three Fourier inversion methods based on the previous wavelets which are used in the thesis. Brief conclusions are given in Section 2.7.

2.2. MATHEMATICAL OVERVIEW

This mathematical overview section provides preliminaries of probability theory and the basics to go into the wavelet world understanding all the concepts and theory that is used there. Mainly this first section is dedicated to probability theory, norms, inner products and orthog-

onality. We start with some basics on probability theory that we need later when applying the inversion methods.

Through the whole work $\mathbb{R}$ refers to the real line, $\mathbb{C}$ the complex plane and $\mathbb{Z}$ stands for the set of integers.

BASICS ON PROBABILITY THEORY

Definition 2.2.1. *A **σ-algebra** (also σ-field) on a set Ω is a collection $\mathscr{F}$ of subsets of Ω that includes the empty subset, is closed under complement, and is closed under countable unions and countable intersections. The pair $(\Omega, \mathscr{F})$ is called a **measurable space**.*

Definition 2.2.2. *Let $\mathscr{F}$ be a σ-field on Ω. A function $\mathbb{P} : \mathscr{F} \to [0,1]$ is a **probability measure** if the following conditions hold,*

1. $\mathbb{P}(\Omega) = 1$,
2. *if $A_1, A_2, \ldots$ are pairwise disjoint sets (i.e. $A_i \cap A_j = \emptyset$ for $i \neq j$) belonging to $\mathscr{F}$, then,*

$$\mathbb{P}(A_1 \cup A_2 \cup \cdots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \cdots. \tag{2.1}$$

*The triplet $(\Omega, \mathscr{F}, \mathbb{P})$ is called a **probability space**. The sets belonging to $\mathscr{F}$ are called **events**. An event A is said to occur almost surely (a.s.) whenever $\mathbb{P}(A) = 1$.*

Definition 2.2.3. *If $\mathscr{F}$ is a σ-field on Ω, then a function $\xi : \Omega \to \mathbb{R}$ is said to be **$\mathscr{F}$-measurable** if $\{\omega \in \Omega : \xi(\omega) \in B\}$ for every Borel set $B \in \mathscr{B}(\mathbb{R})$, being $\mathscr{B}(\mathbb{R})$ the Borel σ-algebra i.e. the σ-algebra generated by the open sets of $\mathbb{R}$. If $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space, then such a function ξ is called a **random variable**.*

Definition 2.2.4. *The **probability density function** (PDF) of a continuous random variable X is an integrable function $f(x)$ satisfying the following,*

1. *$f(x) \geq 0$, for all $x \in \mathbb{R}$.*
2. *$\int_{\mathbb{R}} f(x)\,dx = 1$.*
3. *If A is an interval, $\mathbb{P}(X \in A) = \int_A f(x)\,dx$.*

Definition 2.2.5. *The **cumulative distribution function** (CDF) of a continuous random variable X is defined as,*

$$F(x) = \int_{-\infty}^{x} f(t)\,dt, \tag{2.2}$$

for $-\infty < x < \infty$.

Definition 2.2.6. *Let X be a random variable with probability density function f, then the **expected value** is defined as the following Lebesgue integral,*

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\,dx. \tag{2.3}$$

It can be shown that the expected value of a measurable function of X, $g(X)$, where X has a probability density function $f(x)$, is given by,

$$\mathbb{E}\left[g(X)\right] = \int_{-\infty}^{\infty} g(x) f(x)\,dx. \tag{2.4}$$

Definition 2.2.7. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and $X : \Omega \to \mathbb{R}^n$ a random variable in it. Let (U, Σ) be a measurable space, and $Y : \Omega \to U$ another random variable. Let $g : U \to \mathbb{R}^n$ be a Σ-measurable function such that, for every Σ-measurable function $f : U \to \mathbb{R}^n$,*

$$\int g(Y) f(Y) d\mathbb{P} = \int X f(Y) d\mathbb{P}. \tag{2.5}$$

*The random variable $g(Y)$, denoted as $\mathbb{E}[X \mid Y]$, is a **conditional expectation** of X given Y.*

Proposition 2.2.8 (Law of iterated expectations)**.** *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and X and Y two random variables in it, such that the expected value of X, $\mathbb{E}[X]$, is defined. Then,*

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]], \tag{2.6}$$

i.e., the expected value of the conditional expected value of X given Y is the same as the expected value of X.

Definition 2.2.9. *Let X be a continuous random variable with probability density function f. Then, the **characteristic function** (ChF) of X is defined as,*

$$\mathbb{E}\left[e^{-i\omega X}\right] = \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx. \tag{2.7}$$

We see later that this expression is equivalent to the Fourier transform of f.

Unless otherwise stated, we assume all functions f defined on the real line $\mathbb{R}$ to be measurable. Now, we move on to the essential concepts to introduce Fourier analysis and wavelet theory.

NORMS

The notion of norm generalizes the classical notion of length or size of vectors in $\mathbb{R}^p$, $p \in \mathbb{N}$, to elements of a general vector space.

Definition 2.2.10. *Let V be a vector space over a field F, F being either complex $\mathbb{C}$ or real $\mathbb{R}$. A **norm** on V is a function $p : V \to \mathbb{R}$ that satisfies the following properties. For all $a \in F$ and all $u, v \in V$,*

1. *(Homogeneity) $p(av) = |a| p(v)$.*
2. *(Triangle inequality) $p(u + v) \leq p(u) + p(v)$.*
3. *(Positivity) $p(v) \geq 0$. $p(v) = 0$ only for $v = 0$.*

The norm of a vector $v \in V$ is usually denoted by $||v||_V$.

*A **normed vector space** is a vector space endowed with a norm.*

Definition 2.2.11. *A **Banach space** is a vector space over either $\mathbb{R}$ or $\mathbb{C}$ equipped with a norm which is complete with respect to that norm.*

The completeness condition is equivalent to that for all Cauchy sequences $\{v_n\}$ in V, there exist $v \in V$ such that $||v_n - v||_V \to 0$, as $n \to \infty$.

One example of Banach spaces are the L^p spaces, which are function spaces defined using a natural generalization of the p-norm for finite dimensional vector spaces.

Definition 2.2.12. *For each p, $1 \leq p < \infty$, let $\boldsymbol{L^p(S)}$ denote the class of measurable functions f on a measure space S such that,*

$$\int_S |f(x)|^p \, dx < \infty. \tag{2.8}$$

*Those spaces are Banach spaces with the $\boldsymbol{L^p(S)}$ **norm** defined by,*

$$\|f\|_p := \begin{cases} \left\{\int_{-\infty}^{\infty} |f(x)|^p \, dx\right\}^{\frac{1}{p}}, & \text{for } 1 \leq p < \infty, \\ \text{ess sup}_{x \in S} |f(x)|, & \text{for } p = \infty, \end{cases} \tag{2.9}$$

where ess sup stands for the essential supremum, which is the proper generalization to measurable functions of the maximum.

Note that functions in $L^p(S)$ are only defined up to a set of measure 0. Since any point $x \in S$ is a measure 0 set under the Lebesgue measure, one cannot talk about pointwise evaluation of the functions in this space. For this reason, we need to use the essential supremum instead of supremum. Two functions are considered equivalent in those spaces if and only if they are the same up to a set of measure 0.

Definition 2.2.13. *A **separable vector space** V is a vector space such that there exist a countable dense subset $\{f_n\}_{n=0}^{\infty}$, $f_n \in V$. This means that for all $g \in V$ and $\epsilon > 0$ there exists n such that $||f_n - g|| \leq \epsilon$.*

For all $p < \infty$, the L^p spaces are separable. This property is used later in order to approximate any function in those spaces using functions from the countable subset.

Some important inequalities that hold for the L^p spaces equipped with those norms are,

- The Minkowski inequality, which is an extension of the triangle inequality for function spaces,
$$\|f + g\|_p \leq \|f\|_p + \|g\|_p. \tag{2.10}$$
- The Hölder inequality,
$$\|fg\|_1 \leq \|f\|_p \|g\|_{p(p-1)^{-1}}, \tag{2.11}$$
where $p(p-1)^{-1}$ should be replaced by 1 when $p = \infty$.

  As a special case of Hölder inequality we have the Schwarz inequality,
$$\|fg\|_1 \leq \|f\|_2 \|g\|_2. \tag{2.12}$$

INNER PRODUCT

The standard inner product of two functions is a generalization to the dot product of two finite dimensional vectors.

Definition 2.2.14. *Let V be a vector space over a scalar field F, an **inner product** on V is defined as a map $\langle \cdot, \cdot \rangle : V \times V \to F$ such that the following properties hold. For $x, y \in V$ and $a, b \in \mathbb{C}$,*

1. *(Symmetry) $\langle y, x \rangle = \overline{\langle x, y \rangle}$, where the bar stands for the complex conjugate.*
2. *(Bilinearity) $\langle ax_1 + bx_2, y \rangle = a\langle x_1, y \rangle + b\langle x_2, y \rangle$.*

3. *(Positivity)* $\langle x, x\rangle \geq 0$, *where the case of equality holds only for* $x = 0$.

An ***inner vector space*** *is a vector space endowed with an inner product.*

Observe that a norm can be defined from an inner product. For $f \in V$,

$$\|f\|_V = \sqrt{\langle f, f\rangle}. \tag{2.13}$$

Definition 2.2.15. *A* ***Hilbert space*** H *is an inner product space that is also a complete metric space with respect to the norm induced by the inner product.*

Taking $p = 2$ in Definition 2.2.12 we can define an inner product on $L^2(S)$ by,

$$\langle f, g\rangle = \int_S f(x)\overline{g(x)}dx, \quad f, g \in L^2(S). \tag{2.14}$$

This product satisfies all the conditions to be an inner product. Observe that,

$$\langle f, f\rangle = \|f\|_2^2, \quad f \in L^2(S). \tag{2.15}$$

So, the standard norm of $L^2(S)$ is derived from this inner product. We know from before that $L^2(S)$ is closed with respect to this norm. Hence, endowed with this inner product, $L^2(S)$ becomes a Hilbert space.

The following definition describes a discrete analogue of L^2.

Definition 2.2.16. *The space* l^2 *is the set of all sequences* $X = \ldots, x_{-1}, x_0, x_1, \ldots$, $x_i \in \mathbb{C}$, *with* $\sum_{-\infty}^{\infty} |x_n|^2 < \infty$.

ORTHOGONALITY

An important part of the thesis consists of approximating functions in $L^2(\mathbb{R})$ using orthonormal basis. In the following we present concepts such as orthogonality and projection.

Definition 2.2.17. *Suppose* V *is an inner product space. The notion of orthogonality implies,*

- *The vectors* X *and* Y *are said to be* ***orthogonal vectors*** *if* $\langle X, Y\rangle = 0$.
- *Two subspaces* V_1 *and* V_2 *of* V *are said to be* ***orthogonal subspaces*** *if each vector in* V_1 *is orthogonal to each vector in* V_2.
- *Two functions* f *and* g *are said to be* ***orthogonal functions****, and we write* $f \perp g$, *when* $\langle f, g\rangle = 0$.
- *A set of functions is an* ***orthogonal set of functions*** *if and only if every distinct pair in the set is orthogonal.*

Definition 2.2.18. *Suppose* V_0 *is a finite dimensional subspace of an inner product space* V. *For any vector* $v \in V$, *the* ***orthogonal projection*** *of* v *onto* V_0 *is the unique vector* $v_0 \in V_0$ *that is closest to* v; *i.e.,*

$$\|v - v_0\| = \min_{\omega \in V_0} \|v - \omega\|. \tag{2.16}$$

To complete the theory, we also need to introduce the notion of orthonormality.

Definition 2.2.19 (Orthonormality)**.**

- *The collection of vectors e_i, $i = 1, \cdots, N$, is said to be an **orthonormal vector collection** if each e_i has unit length, $\|e_i\| = 1$, and e_i and e_j are orthogonal for $i \neq j$.*

- *A sequence of functions $\{f_n\}_{n\in\mathbb{Z}}$ is said to be an **orthonormal sequence of functions** if $\langle f_m, f_n\rangle = \delta_{m,n}$, where $\delta_{j,k}$ is the Kronecker delta defined by,*

$$\delta_{j,k} := \begin{cases} 1, & \text{for } j = k, \\ 0, & \text{for } j \neq k. \end{cases} \tag{2.17}$$

If $\{\phi_1, \phi_2, \phi_3, \ldots\}$ is any orthogonal set of non-zero functions, then a corresponding orthonormal set $\{\psi_1, \psi_2, \psi_3, \ldots\}$ can be constructed by "normalizing" each ϕ_k, that is,

$$\psi_k(x) = \frac{\phi_k(x)}{\|\phi_k\|}. \tag{2.18}$$

Now we are in a position to define a basis or Hilbert basis.

Definition 2.2.20. *Given a Hilbert space V, a **Hilbert basis** (or simply basis) for V is an orthonormal set of vectors, H, with the property that every vector in V can be written as an infinite linear combination of the vectors in the basis.*

Theorem 2.2.21. *Let $\{f_n\}_{n=1}^{\infty}$ be an orthonormal set on $L^2(S)$. Then the following conditions also characterize an orthonormal basis.*

1. *For each $g \in L^p(S)$,*

$$g = \sum_{n=0}^{\infty} \langle g, f_n\rangle f_n. \tag{2.19}$$

2. *For each $g \in L^p(S)$,*

$$\|g\|^2 = \sum_n |\langle g, f_n\rangle|^2, \tag{2.20}$$

which is known as Parseval's relation.

This corresponds to Theorem 2.4 of [80], where you can also find the proof. The theorem provides a general procedure to approximate functions on L^2. We first find an orthonormal basis (which is not unique), pick a subset of this basis and do the orthogonal projection on to it. As we increase the size of the subset we converge to the right function.

The values $\langle g, f_n\rangle$ are usually referred to as Fourier coefficients the reason for this will become apparent when we introduce the Fourier series later. Note that by the Parseval's relation, the Fourier coefficients of the expansion form an infinite dimensional vector which is an element of l^2.

2.3. FOURIER ANALYSIS

The beginning of Fourier Analysis is usually placed in 1807 when Joseph Fourier presented a memoir to the Institut de France. In his work, he claimed that any periodic function could be represented as a series of harmonically related sinusoids.

As exposed in the introduction the two main operations in Fourier analysis are the Fourier transforms and the Fourier series. They can be given a physical interpretation such as time-frequency analysis of signals (considered periodic functions of the time variable). However, they are also relevant from a mathematical point of view since they provide a tool to tackle a really wide range of problems in topics including partial differential equations and probability among others.

The basic goal of Fourier series is to take a periodic signal, and decompose it into its various frequency components which are represented by the sine and cosine.

The Fourier transform can be viewed as an extension of the Fourier series to general, non-periodic functions. Fourier transform takes a signal and express it in terms of the frequencies of the waves that make up that signal.

2.3.1. FOURIER SERIES

We turn now to the study of periodic functions (specifically of period $2a$ for $a \in \mathbb{R}$). For each p, $L^p(-a,a)$ denotes the Banach space of functions f satisfying $f(x+2a) = f(x)$ almost everywhere (a.e.) in $\mathbb{R}$ and $\|f\|_{L^p(-a,a)} < \infty$.

A Fourier series is a particular way of rewriting functions as the sum of simple waves (or series of trigonometric functions). More formally, it decomposes any periodic function into the sum of a (possibly infinite) set of simple oscillating functions, specifically complex exponentials (or, equivalently, sines and cosines).

Theorem 2.3.1. *Let $a \in \mathbb{R}$, the set of functions,*

$$\left\{ \frac{1}{\sqrt{2a}} e^{\frac{in\pi x}{a}};\ n = \ldots, -2, -1, 0, 1, 2, \ldots \right\}, \tag{2.21}$$

is an orthonormal basis for $L^2(-a,a)$.

Proof can be found in [12]. We can now introduce the complex form of Fourier series.

Definition 2.3.2. *The **complex Fourier series** of a periodic function $f(x) \in L^2(-a,a)$ is given by,*

$$f(x) = \sum_{n=-\infty}^{\infty} \alpha_n e^{\frac{in\pi x}{a}}, \tag{2.22}$$

where the coefficients of the complex Fourier series are,

$$\alpha_n = \frac{1}{2a} \int_{-a}^{a} f(x) e^{\frac{-in\pi x}{a}} dx. \tag{2.23}$$

This expansion makes sense for any f in $L^2(-a,a)$ thanks to Theorem 2.3.1 and Theorem 2.2.21.

Due to the relation with the exponential, and sine and cosine functions it is easy to see that the real expression of the Fourier series is also well defined when f is a real function. This next version is the original Fourier series.

Definition 2.3.3. *The **Fourier series** of a periodic function $f(x) \in L^2(-a,a)$ of period $2a$ is given by,*

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{\pi n}{a}x\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{\pi n}{a}x\right), \tag{2.24}$$

where the coefficients of the Fourier series are,

$$a_0 = \tfrac{1}{2a}\int_{-a}^{a} f(x)dx,$$
$$a_n = \tfrac{1}{a}\int_{-a}^{a} f(x)\cos\left(\tfrac{\pi n}{a}x\right)dx,$$
$$b_n = \tfrac{1}{a}\int_{-a}^{a} f(x)\sin\left(\tfrac{\pi n}{a}x\right)dx.$$

2.3.2. FOURIER AND INVERSE FOURIER TRANSFORMS

The Fourier series showed us how to rewrite any periodic function into a sum of sinusoids. The Fourier Transform is the extension of this idea to non-periodic functions.

We start introducing the Fourier transform for functions in $L^1(\mathbb{R})$, and then this is extended to the space of functions in $L^2(\mathbb{R})$. This is done in this way to avoid convergence issues since it is not straightforward that we can define the Fourier transform on $L^2(\mathbb{R})$.

Definition 2.3.4. *The **Fourier transform** of a function $f \in L^1(\mathbb{R})$ is defined by,*

$$\hat{f}(\omega) := \int_{-\infty}^{\infty} e^{-i\omega x} f(x)dx. \tag{2.25}$$

Since $f \in L^1(\mathbb{R})$ the integral in (2.25) converges and is bounded,

$$|\hat{f}(\omega)| \leq \int_{-\infty}^{\infty} |f(x)|dx < +\infty. \tag{2.26}$$

One can verify that $\hat{f}$ is a continuous function of ω. If $\hat{f}$ is also integrable, the next theorem gives us the inverse Fourier transform.

Theorem 2.3.5. *If $f \in L^1(\mathbb{R})$ and $\hat{f} \in L^1(\mathbb{R})$ then the **inverse Fourier transform** of $\hat{f}$ is,*

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega)e^{i\omega x}d\omega. \tag{2.27}$$

The proof can be found in [50].

We move now to the extension in $L^2(\mathbb{R})$. If $f \in L^2(\mathbb{R})$ but $f \notin L^1(\mathbb{R})$ its Fourier transform cannot be calculated with the Fourier integral formula (2.25) because $f(x)e^{-i\omega x}$ is not integrable. The Fourier transform is defined as a limit using the Fourier transforms of functions in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.

Since $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is dense in $L^2(\mathbb{R})$, a family $\{f_n\}_{n\in\mathbb{Z}}$ of functions in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ that converges to f can be found,

$$\lim_{n\to+\infty} \|f - f_n\|_2 = 0. \tag{2.28}$$

Since $\{f_n\}_{n\in\mathbb{Z}}$ converges, it is a Cauchy sequence, which means that $\|f_p - f_q\|_2$ is arbitrarily small if p and q are large enough. Moreover, $f_n \in L^1(\mathbb{R})$, so its Fourier transform $\hat{f}_n$ is well defined. The Plancherel formula, states that,

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\hat{f}(\omega)|^2 d\omega, \tag{2.29}$$

proves that $\{\hat{f}_n\}_{n\in\mathbb{Z}}$ is also a Cauchy sequence because $\|\hat{f}_p - \hat{f}_q\|_2 = \sqrt{2\pi}\|f_p - f_q\|_2$ is arbitrarily small for large enough p and q. A Hilbert space is complete, which means that all Cauchy sequences converge to an element of the space. Thus, there exists $\hat{f} \in L^2(\mathbb{R})$ such that

$$\lim_{p\to+\infty} \|\hat{f} - \hat{f}_p\|_2 = 0. \tag{2.30}$$

By definition, $\hat{f}$ is the Fourier transform of f.

We conveniently do an abuse of notation and use (2.25) when a limiting process is understood.

If a function f is compactly supported, $f(x) = 0$ if $x < a$ or $x > b$, where $-\infty < a < b < \infty$, then its Fourier transform $\hat{f}(\xi)$ is well defined also for a complex ξ.

The Parseval Identity states the following relation.

Theorem 2.3.6. *For all $f, g \in L^2(\mathbb{R})$, the following relation holds,*

$$\langle f, g\rangle = \frac{1}{2\pi}\langle \hat{f}, \hat{g}\rangle. \tag{2.31}$$

In particular, $\|f\|_2 = (2\pi)^{-\frac{1}{2}}\|\hat{f}\|_2$.

2.3.3. DISCRETE FOURIER TRANSFORM

We denote by $f[n]$, $n \in \mathbb{N}$ the sample values or discrete signals. In practice, $f[n]$ is known over a finite domain, say $0 \le n < N$.

Definition 2.3.7. *The **discrete Fourier transform** (DFT) of f is*

$$\hat{f}[k] = \sum_{n=0}^{N-1} f[n] \exp\left(\frac{-i2\pi kn}{N}\right), \tag{2.32}$$

where exp *stands for the exponential function. And the **inverse discrete Fourier transform** (IDFT) formula,*

$$f[n] = \frac{1}{N}\sum_{k=0}^{N-1} \hat{f}[k] \exp\left(\frac{i2\pi kn}{N}\right). \tag{2.33}$$

The formula for the discrete Fourier transform is the analogous to the formula for the kth Fourier coefficient with the sum over n taking the place of the integral over x.

The computation of the discrete Fourier transform is equivalent to the following matrix computation. Let $y_i = f[i]$ and $\omega = \exp(\frac{-2\pi i}{N})$, then we have,

$$\hat{y} = (\bar{F}_N)(y), \tag{2.34}$$

where $y = (y_0, \cdots, y_{N-1})^T$, $\hat{y} = (\hat{y}_0, \cdots, \hat{y}_{N-1})^T$ and,

$$\bar{F}_N = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(N-1)} \\ \vdots & & \ddots & & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \cdots & \omega^{(N-1)^2} \end{pmatrix}. \tag{2.35}$$

THE FAST FOURIER TRANSFORM

The fast Fourier transform is an efficient algorithm for computing the discrete Fourier and inverse discrete Fourier transform. The computation of the discrete Fourier transform is equivalent to multiplying the sequence y (which is an $N \times 1$ column vector) by an $N \times N$ matrix $\bar{F}_N$, which is an operation that requires N^2 multiplications. However, the FFT algorithm takes advantage of the special form of the matrix $\bar{F}_N$ to reduce the number of multiplications to roughly $5N\log_2 N$ (i.e. to $\mathcal{O}(N\log_2 N)$) by reorganizing the calculations. The relative savings increase as N gets larger.

The algorithm goes as follows. When the frequency index is even, we group the terms n and $n+N/2$,

$$\hat{f}[2k] = \sum_{n=0}^{N/2-1} \big(f[n] + f[n+N/2]\big) \exp\left(\frac{-i2\pi kn}{N/2}\right). \tag{2.36}$$

When the frequency index is odd, the same grouping becomes,

$$\hat{f}[2k+1] = \sum_{n=0}^{N/2-1} \exp\left(\frac{-i2\pi n}{N}\right) \big(f[n] - f[n+N/2]\big) \exp\left(\frac{-i2\pi kn}{N/2}\right). \tag{2.37}$$

Equation (2.36) proves that even frequencies are obtained by calculating the DFT of the $N/2$ periodic signal $f_{\text{even}}[n] = f[n] + f[n+N/2]$.

Odd frequencies are derived from (2.37) by computing the Fourier transform of the $N/2$ periodic signal:

$$f_{\text{odd}}[n] = \exp\left(\frac{-i2\pi n}{N}\right) \big(f[n] - f[n+N/2]\big). \tag{2.38}$$

A DFT of size N may thus be calculated with two discrete Fourier transforms of size $N/2$ plus $\mathcal{O}(N)$ operations.

The IDFT of $\hat{f}$ is derived from the forward fast Fourier transform of its complex conjugate $\hat{f}^*$ by observing that,

$$f^*[n] = \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}^*[k] \exp\left(\frac{-i2\pi kn}{N}\right). \tag{2.39}$$

The computational complexity is obtained as follows. Let $C(N)$ be the number of elementary operations needed to compute a DFT with the FFT. Since f is complex, the calculation of f_{even} and f_{odd} requires N complex additions and $N/2$ complex multiplications. Let KN be the corresponding number of elementary operations. We have,

$$C(N) = 2C(N/2) + KN, \tag{2.40}$$

since the Fourier transform of a single point is equal to itself, $C(1) = 0$. With the change of variable $l = \log_2 N$ and the change of function $T(l) = \frac{C(N)}{N}$, we derive,

$$T(l) = T(l-1) + K. \tag{2.41}$$

Since $T(0) = 0$, we get $T(l) = Kl$ and therefore,

$$C(N) = KN\log_2(N). \tag{2.42}$$

Several variations of this fast algorithm exist. The goal is to minimize the constant K. The most efficient fast DFT to this date is the split-radix FFT algorithm, which is slightly more complicated than the procedure just described. However, it requires only $N\log_2 N$ real multiplications and $3N\log_2 N$ additions. When the input signal is real, there are half as many parameters to compute.

2.4. WAVELETS

One disadvantage of Fourier series is that its building blocks, sines and cosines are periodic waves. A different set of building blocks, called wavelets is designed to model different type of signals.

In a rough sense, a wavelet looks like a wave that travels for one or more periods and is non-zero only over a finite interval instead of propagating forever the way sines and cosines do. A wavelet can be translated forward or backward in time. It also can be stretched or compressed by scaling to obtain low and high frequency wavelets.

More mathematically, wavelets consist of a family of functions constructed from dilation and translation of a single function ψ in $L^2(\mathbb{R})$ called the wavelet. When the dilation parameter a and the translation parameter b vary continuously, we have the following family of **continuous wavelets**,

$$\psi_{a,b}(x) = |a|^{-\frac{1}{2}}\psi\left(\frac{x-b}{a}\right), \quad a,b \in \mathbb{R}, \quad a \neq 0. \tag{2.43}$$

We often refer to ψ as the mother wavelet and $\psi_{a,b}$ as the child wavelets.

Analogously as in Fourier analysis, the main tools provided by wavelets are the wavelet transform and their discrete version, to which we refer as wavelet series although in the literature they are also referred as discrete wavelet transform.

Definition 2.4.1. *Let f be a function in $L^2(\mathbb{R})$, its **continuous wavelet transform** is defined as,*

$$WT_\psi\{f\}(a,b) := |a|^{-\frac{1}{2}} \int_{\mathbb{R}} f(x)\overline{\psi_{a,b}\left(\frac{x-b}{a}\right)}\,dx. \tag{2.44}$$

During the thesis we only work with wavelet series. Therefore, we do not expand more on the continuous wavelet transform. For the interested reader we refer to reference [19].

The parameters a and b can be restricted to the following discrete values: $a = a_0^{-m}$ and $b = nb_0a_0^{-m}$, $a_0 > 1$, $b_0 > 0$ and $n, m \in \mathbb{Z}$. Then, we have the following family of **discrete wavelets**,

$$\psi_{m,n}(x) = |a_0|^{m/2}\psi(a_0^m x - nb_0). \tag{2.45}$$

In the same way, the discrete wavelet transform (DWT) is any wavelet transform for which the wavelets are discretely sampled. We can now define the wavelet series.

Definition 2.4.2. *Let f be a function in $L^2(\mathbb{R})$. The **wavelet series** is defined as,*

$$f(x) = \sum_{m\in\mathbb{Z}}\sum_{n\in\mathbb{Z}} \langle f, \psi_{m,n}\rangle \psi_{m,n}(x), \tag{2.46}$$

provided that the functions $\{\psi_{m,n} : m,n \in \mathbb{Z}\}$ form an orthonormal basis of $L^2(\mathbb{R})$.

We recall that if we have an orthonormal basis of $L^2(\mathbb{R})$ then we are able to reconstruct any function $f \in L^2(\mathbb{R})$ using the expansion from Theorem 2.2.21.

Observe that the last formula involves a double infinite sum, even though in practice we only use a finite number of terms, for each term we have to calculate an integral. The computation of the discrete wavelet transform can be made more efficient in the special case in which the scaled and shifted wavelets form a multiresolution analysis. In these cases, there exists an auxiliary function, the father wavelet ϕ (also referred as scaling function) in $L^2(\mathbb{R})$

and a_0 is an integer. The most common choice for a_0 and b_0 are $a_0 = 2$ and $b_0 = 1$ and from now on we use those values. Note that not every orthonormal discrete wavelet basis can be associated to a multiresolution analysis.

Multiresolution analysis does not just allow for significant computational savings with respect to the other orthonormal discrete wavelet basis, but in practice most constructions of the wavelet expansion make use of the multiresolution analysis which defines the mother wavelet via a scaling function as well. In the thesis, we mostly use wavelets defined via multiresolution analysis, we introduce it in detail below.

2.4.1. MULTIRESOLUTION ANALYSIS

As previously mentioned, there are two functions that play a primary role in multiresolution analysis, the scaling function ϕ and the wavelet ψ, both functions in $L^2(\mathbb{R})$. To emphasize the marriage involved in building this family, ϕ is sometimes called father wavelet and ψ mother wavelet.

Here we present a method for constructing orthonormal wavelets that is based on the existence of a family of subspaces of $L^2(\mathbb{R})$ satisfying certain properties. Such a family is called a multiresolution analysis. The main feature of this method is to describe mathematically the process of studying signals at different scales. It is traditional to define an MRA by specifying five properties that the family of subspaces must satisfy.

Most of the theory we present is due to Stéphane Mallat [50]. We follow multiresolution analysis as introduced in [12], unless otherwise stated the proof of the theorems can be found there.

Definition 2.4.3. *Let V_m, $m = \ldots, -2, -1, 0, 1, 2, \ldots$ be a sequence of subspaces of functions in $L^2(\mathbb{R})$. The collection of spaces $\{V_m, m \in \mathbb{Z}\}$ is called a **multiresolution analysis** (MRA) with scaling function ϕ (also known by father wavelet) if the following conditions hold,*

1. *(nested) $V_m \subset V_{m+1}$,*
2. *(density) $\overline{\cup V_m} = L^2(\mathbb{R})$,*
3. *(separation) $\cap V_m = \{0\}$,*
4. *(scaling) he function $f(x)$ belongs to V_m if and only if the function $f(2^{-m}x)$ belongs to V_0,*
5. *(orthonormal basis) the function ϕ belongs to V_0 and the set $\{\phi(x-k), k \in \mathbb{Z}\}$ is an orthonormal basis (using the $L^2(\mathbb{R})$ inner product) for V_0.*

The V_m's are called approximation spaces and different choices for ϕ yield different multiresolution analysis. And we refer to m as the scale.

Within the literature, the scaling function is sometimes introduced in an alternative weaker way. In the MRA framework the scaling function is an orthonormal basis of V_0. However, it could also be considered to be a Riesz basis, we see it in the next remark. Both ways are equivalent.

Remark 2.4.4. *Sometimes condition 5 (orthonormal basis of Definition 2.4.3) is relaxed by assuming that $\{\phi(x-k), k \in \mathbb{Z}\}$ is a Riesz basis for V_0. That is, for every $f \in V_0$ there exists a unique sequence $\{\alpha_n\}_{n\in\mathbb{Z}} \in l^2(\mathbb{Z})$ such that,*

$$f(x) = \sum_{n\in\mathbb{Z}} \alpha_n \phi(x-n), \tag{2.47}$$

with convergence in $L^2(\mathbb{R})$ *and,*

$$A\sum_{n\in\mathbb{Z}}|\alpha_n|^2 \le \|\sum_{n\in\mathbb{Z}}\alpha_n\phi(x-n)\|_2^2 \le B\sum_{n\in\mathbb{Z}}|\alpha_n|^2, \tag{2.48}$$

with $0 < A \le B < \infty$ *constants independent of* $f \in V_0$*. If this is the case we say that we have an MRA with a Riesz basis. The fact that this weaker version of condition 5 (orthonormal basis) is equivalent to the one announced above is proved in the reference [12], at the end of Section 2.1.*

Theorem 2.4.5. *Suppose* $\{V_m; m \in \mathbb{Z}\}$ *is a multiresolution analysis with scaling function* ϕ*. Then, for any* $m \in \mathbb{Z}$*, the set of functions,*

$$\{\phi_{m,k}(x) = 2^{m/2}\phi(2^m x - k); k \in \mathbb{Z}\}, \tag{2.49}$$

is an orthonormal basis for V_m*.*

We are now ready to state the scaling relation.

Theorem 2.4.6. *Suppose* $\{V_m; m \in \mathbb{Z}\}$ *is a multiresolution analysis with scaling function* ϕ*. Then the following relation holds,*

$$\phi(x) = \sum_{k\in\mathbb{Z}} p_k\phi(2x-k) \;\; \textit{where} \;\; p_k = 2\int_{-\infty}^{\infty}\phi(x)\overline{\phi(2x-k)}\,dx. \tag{2.50}$$

Moreover, we also have,

$$\phi(2^{m-1}x - l) = \sum_{k\in\mathbb{Z}} p_{k-2l}\phi(2^m x - k), \tag{2.51}$$

or, equivalently,

$$\phi_{m-1,l} = 2^{-1/2}\sum_k p_{k-2l}\phi_{m,k}, \tag{2.52}$$

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

THE ASSOCIATED WAVELET AND WAVELET SPACES

Theorem 2.4.7. *Suppose* $\{V_m; m \in \mathbb{Z}\}$ *is a multiresolution analysis with scaling function,*

$$\phi(x) = \sum_k p_k\phi(2x-k), \tag{2.53}$$

p_k *are the coefficients in Theorem 2.4.6. Let* W_m *be the span of* $\{\psi(2^m x - k); k \in \mathbb{Z}\}$*, where,*

$$\psi(x) = \sum_{k\in\mathbb{Z}} (-1)^k\overline{p_{1-k}}\phi(2x-k). \tag{2.54}$$

Then $W_m \subset V_{m+1}$ *is the orthogonal complement of* V_m *in* V_{m+1}*, known as the detailed space. Furthermore,* $\{\psi_{m,k}(x) := 2^{m/2}\psi(2^m x - k), k \in \mathbb{Z}\}$ *is an orthonormal basis for the* W_m *and* ψ *is known as the mother wavelet.*

Note that $\psi_{m,l} = 2^{m/2}\psi(2^m x - l)$ has the expansion,

$$\psi_{m,l} = 2^{-1/2}\sum_{k\in\mathbb{Z}} (-1)^k\overline{p_{1-k+2l}}\phi_{m+1,k}, \tag{2.55}$$

which follows from the definitions of ψ given in Theorem 2.4.7 by substituting $2^m x - l$ for x, multiplying both sides by $2^{m/2}$ and adjusting the summation index.

From Theorem 2.4.7, the set $\{\psi_{m-1,k}\}_{k\in\mathbb{Z}}$ is an orthonormal basis for the space W_{m-1}, which is the orthogonal complement of V_{m-1} in V_m (so $V_m = W_{m-1} \oplus V_{m-1}$). By successive orthogonal decompositions,

$$\begin{aligned} V_m &= W_{m-1} \oplus V_{m-1} \\ &= W_{m-1} \oplus W_{m-2} \oplus V_{m-2} \\ &\cdots \\ &= W_{m-1} \oplus W_{m-2} \oplus \cdots \oplus W_0 \oplus V_0. \end{aligned} \tag{2.56}$$

Since we have defined V_m for $m < 0$, we can keep going,

$$\begin{aligned} V_m &= W_{m-1} \oplus W_{m-2} \oplus \cdots \oplus W_0 \oplus W_{-1} \oplus V_{-1} \\ &\cdots \\ &= W_{m-1} \oplus W_{m-2} \oplus \cdots \oplus W_{-1} \oplus W_{-2} \cdots. \end{aligned} \tag{2.57}$$

The V_m are nested and the union of all the V_m is the space $L^2(\mathbb{R})$. Therefore, we can let $j \to \infty$ and obtain the following theorem.

Theorem 2.4.8. *Let $\{V_m; m \in \mathbb{Z}\}$ be a multiresolution analysis with scaling functions ϕ. Let W_m be the orthogonal complement of V_m in V_{m+1}. Then,*

$$L^2(\mathbb{R}) = \cdots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \cdots. \tag{2.58}$$

In particular, each $f \in L^2(\mathbb{R})$ can be uniquely expressed as a sum $\sum_{k=-\infty}^{\infty} \omega_k$, $\omega_k \in W_k$, where the ω_k's are mutually orthogonal. Equivalently, the set of all wavelets, $\{\psi_{m,k}\}_{m,k\in\mathbb{Z}}$, is an orthonormal basis for $L^2(\mathbb{R})$.

The infinite sum appearing in this theorem should be thought of as an approximation by finite sums. In other words, each $f \in L^2(\mathbb{R})$ can be approximated arbitrarily closed in the L^2-norm by finite sums of the form $\omega_{-m} + \omega_{1-m} + \cdots + \omega_{m-1} + \omega_m$ for suitable large m.

We have two bases that we can use to approximate functions in L^2: the one produced by the mother wavelet and the one produced by the father wavelet.

Proposition 2.4.9. *For any $f \in L^2(\mathbb{R})$, a projection map of $L^2(\mathbb{R})$ onto V_m, which we denote by $\mathscr{P}_m : L^2(\mathbb{R}) \to V_m$, is defined by,*

$$\mathscr{P}_m f(x) = \sum_{j=-\infty}^{m-1} \sum_{k=-\infty}^{k=+\infty} d_{j,k} \psi_{j,k}(x) = \sum_{k\in\mathbb{Z}} c_{m,k} \phi_{m,k}(x). \tag{2.59}$$

Here,

$$d_{j,k} = \int_{-\infty}^{+\infty} f(x) \psi_{j,k}(x) dx, \tag{2.60}$$

are the wavelet coefficients, and,

$$c_{m,k} = \int_{-\infty}^{\infty} f(x) \phi_{m,k} dx, \tag{2.61}$$

are the scaling coefficients.

Considering higher m values (i.e. when more terms are used), the truncated series representation of the function f improves since we know, by Theorem 2.2.21, that as m goes to infinity the expansion converges to the exact function. As opposed to Fourier series, a key fact regarding the use of wavelets is that wavelets can be moved (by means of the k value), stretched or compressed (by means of the m value) to accurately represent the local properties of a function.

2.5. EXAMPLES OF WAVELET FUNCTIONS

This section describes Haar, Shannon and Chebyshev wavelets which are the families we use throughout the thesis. As one can see in the development of the section, Haar and Shannon are defined through a scaling function and thus, they define a multiresolution analysis. However, Chebyshev is just introduced by the child wavelet, so we do not have a multiresolution structure in this case.

Despite the fact that we only work with three different wavelets families there exist a wide range of them: Haar, Daubechies, Symlets, Coiflets, Biorthogonal, Meyer, Gaussian, Morlet and so on. The reason why we work with the chosen ones is mainly related to the shape of the wavelet and the shape of the function we wish to recover, together with the easy treatment of the wavelet expression.

2.5.1. HAAR WAVELETS

In 1909 Alfréd Haar proposed the Haar sequence in [40]. Haar used these functions to give an example of an orthonormal system for $L^2([0,1])$. However, as previously mentioned, the study of wavelets, and even the term wavelet, did not come until much later.

The Haar function is recognised as the first known wavelet basis and it is the simplest example of an orthogonal wavelet. In addition, the Haar scaling function is a good example of a compactly supported function. A function has compact support if it is identically zero outside a finite interval.

The Haar scaling function (or equivalently father wavelet) is defined as,

$$\phi_H(x) = \begin{cases} 1, & \text{if } 0 \le x < 1, \\ 0, & \text{elsewhere.} \end{cases} \tag{2.62}$$

It has as Fourier transform the function,

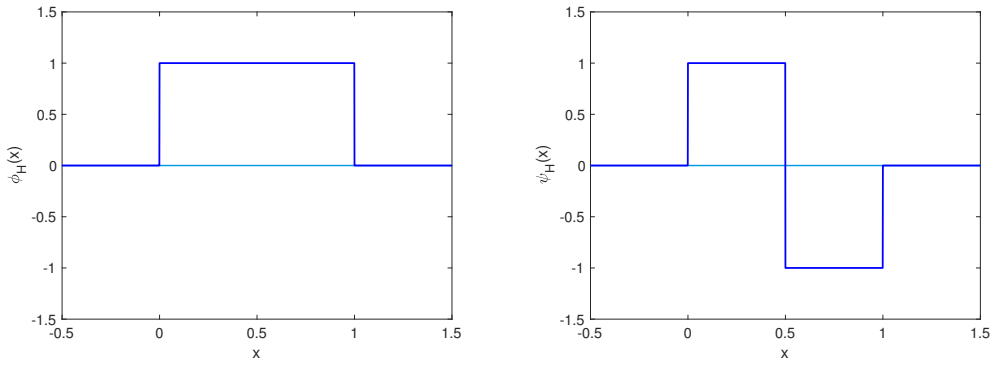
$$\hat{\phi}_H(\omega) = \frac{1 - e^{-i\omega}}{i\omega}. \tag{2.63}$$

The Haar mother wavelet function can be described as,

$$\psi_H(x) = \begin{cases} 1, & \text{for } 0 \le x < \frac{1}{2}, \\ -1, & \text{for } \frac{1}{2} \le x < 1, \\ 0, & \text{otherwise.} \end{cases} \tag{2.64}$$

The Haar scaling function and the Haar mother wavelet are shown in Figure 2.1.

Observe that by rescaling the father wavelet as $\phi_H\left(\frac{x-a}{b-a}\right)$ we get the father wavelet for an orthonormal system for the finite interval $[a,b]$, where $-\infty < a < b < \infty$ and $a, b \in \mathbb{R}$.



(a) Haar scaling function.

(b) Haar mother wavelet.

Figure 2.1: Haar wavelet family.

The drawback with the Haar decomposition is that both father and mother wavelet are discontinuous, and as a result, it provides only crude approximations to some continuous functions.

2.5.2. SHANNON WAVELETS

In some situations, we use wavelets in order to approximate very regular functions (usually density functions). Therefore, it makes sense to choose a wavelet that is smooth in the time-domain such as the Shannon wavelet. Shannon wavelets are extensively described by C. Cattani in [16], they form an orthonormal basis in $L^2(\mathbb{R})$. Unlike Haar wavelets, they can approximate functions with unbounded domain.

Shannon scaling function is given by the sinc function,

$$\phi_S(x) = \text{sinc}(x) = \begin{cases} \dfrac{\sin(\pi x)}{\pi x}, & \text{if } x \neq 0, \\ 1, & \text{if } x = 0. \end{cases} \tag{2.65}$$

It has as Fourier transform the function,

$$\hat{\phi}_S(\omega) = \text{rect}\left(\frac{\omega}{2\pi}\right), \tag{2.66}$$

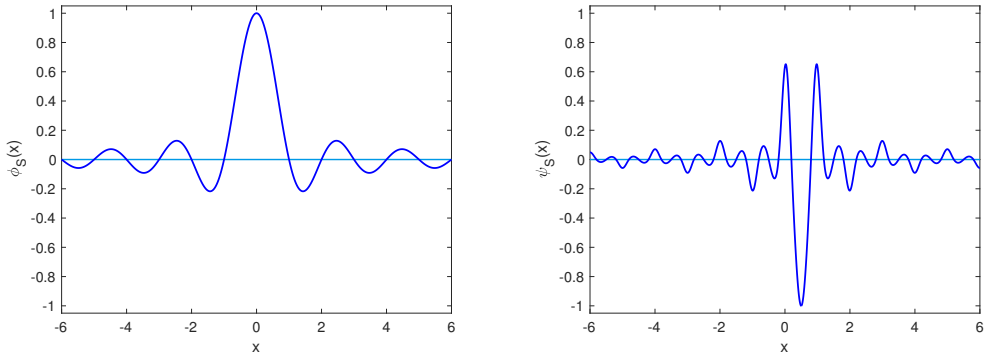
where rect is the rectangle function, defined as,

$$\text{rect}(x) = \begin{cases} 1, & \text{if} |x| < \dfrac{1}{2}, \\ 1/2, & \text{if} |x| = \dfrac{1}{2}, \\ 0, & \text{if} |x| > \dfrac{1}{2}. \end{cases} \tag{2.67}$$

The mother function is,

$$\psi_S(x) = \frac{\sin\left(\pi\left(x-\frac{1}{2}\right)\right) - \sin\left(2\pi\left(x-\frac{1}{2}\right)\right)}{\pi\left(x-\frac{1}{2}\right)}. \tag{2.68}$$

Plots of Shannon scaling function and Shannon mother wavelet are shown in Figure 2.2.



(a) Shannon scaling function. (b) Shannon mother wavelet.

Figure 2.2: Shannon wavelet family.

2.5.3. CHEBYSHEV WAVELETS

As a third wavelet basis, we consider Chebyshev wavelets. The reason to look at this wavelet family is because the COS method [28] (state-of-the-art within the class of Fourier inversion methods) is related with Chebyshev polynomials of first kind. COS method is the cosine expansion of a function $f(x)$ in x; this expansion matches the Chebyshev series expansion of $f(\arccos(t))$ in t.

Chebyshev wavelets, see [3] for a nice description, are derived from conventional Chebyshev polynomials. That is why we start by introducing the Chebyshev polynomials.

CHEBYSHEV POLYNOMIALS

In mathematics, the Chebyshev polynomials [18], named after Pafnuty Chebyshev, are a sequence of orthogonal polynomials which are related to de Moivre's formula and which can be defined recursively. One usually distinguishes between Chebyshev polynomials of the first kind which are denoted T_n and Chebyshev polynomials of the second kind which are denoted U_n.

Both Chebyshev polynomials T_n and U_n are polynomials of degree n and the sequence of Chebyshev polynomials of either kind composes a polynomial sequence. In the study of differential equations, they arise as the solution to the Chebyshev differential equations,

$$(1-x^2)\,y'' - x\,y' + n^2\,y = 0, \tag{2.69}$$

and,

$$(1-x^2)\,y'' - 3x\,y' + n(n+2)\,y = 0, \tag{2.70}$$

for the polynomials of the first and second kind, respectively.

As we mentioned, we are interested in seeing the behaviour of the wavelets related to first kind Chebyshev polynomials, that is why we concentrate only on the first kind ones. However, note that everything can be analogously developed for the second kind ones.

Chebyshev polynomials of first kind

Let us denote by $T_k(x)$, $k = 0, 1, \ldots$, the Chebyshev polynomials of first kind of degree k (see Figure 2.3), given by,

$$T_k(x) = \cos(k\theta), \tag{2.71}$$

in which $\theta = \arccos(x)$. So, $T_k(\cos(\theta)) = \cos(k\theta)$ or $T_k(x) = \cos(k \arccos(x))$. Alternatively, they can be expressed by the recursive relation,

- $T_0(x) = 1$,
- $T_1(x) = x$,
- $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$.

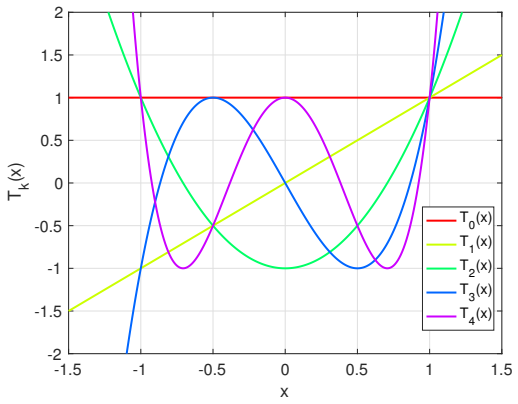


Figure 2.3: First kind Chebyshev polynomials.

Chebyshev polynomials of the first kind are orthogonal with respect to the weight function, on $(-1, 1)$,

$$\varpi\left(\frac{x+1}{2}\right) = \frac{1}{\sqrt{1-x^2}}. \tag{2.72}$$

The Fourier transform of the Chebyshev polynomials is presented in theorem 1 of [31].

Theorem 2.5.1. *Let $\hat{T}_k(\lambda)$ denote the finite Fourier transform of the Chebyshev polynomial $T_k(x)$, i.e.,*

$$\hat{T}_k(\lambda) = \int_{-1}^{1} e^{-i\lambda x} T_k(x)\, dx, \quad \lambda \in \mathbb{C}, \quad k = 0, 1, 2, \ldots, \tag{2.73}$$

where $T_k(x)$ denotes the Chebyshev polynomial

$$T_k(x) = \cos(k \arccos(x)), \quad -1 < x < 1, \quad k = 0, 1, 2, \ldots. \tag{2.74}$$

Then,

$$\hat{T}_k(0) = \begin{cases} 0, & k = 1, \\ \frac{(-1)^{k+1}-1}{k^2-1}, & k = 0, 2, 3, \ldots. \end{cases} \tag{2.75}$$

Furthermore,

$$\hat{T}_k(\lambda) = \sum_{n=1}^{k+1} \alpha_n^k \left[\frac{e^{i\lambda}}{(i\lambda)^n} + (-1)^{n+k} \frac{e^{-i\lambda}}{(i\lambda)^n} \right], \quad \lambda \in \mathbb{C} \setminus \{0\}, \quad k = 0, 1, 2, \ldots, \tag{2.76}$$

where the coefficients α_n^k are defined as follows:

$$\alpha_1^k = (-1)^k, \quad \alpha_2^k = (-1)^{k+1} k^2, \tag{2.77}$$

$$\alpha_n^k = (-1)^{k+n-1} 2^{n-2} k \sum_{l=1}^{k-n+2} \binom{n+l-3}{l-1} \prod_{j=l}^{n+l-3} (k-j), \quad n = 3, 4, \ldots, k+1. \tag{2.78}$$

Remark 2.5.2. *Some properties of Chebyshev polynomials are,*

- $\frac{d}{dx} T_n(x) = n U_{n-1}(x)$, $n = 1, 2, \ldots$.
- $\int U_n(x) dx = \frac{T_{n+1}(x)}{n+1}$.
- $\int T_n(x) dx = \frac{n T_{n+1}(x)}{n^2-1} - \frac{x T_n(x)}{n-1}$.

CHEBYSHEV WAVELETS

Unlike the other two wavelets we have presented before, the Chebyshev wavelets do not have multiresolution analysis structure. This will make their definition and approximating formula very different from the other two. Observe that we do not define them from the scaling function, which does not exist, but from the child wavelets (2.43) instead.

Chebyshev wavelets of first kind

Presented in [3], first kind Chebyshev wavelets are defined by $\psi^C_{m,n,k}(x) = \psi^C(m, n, k, x)$, so as one can note they have four arguments which are,

- m which is any non-negative integer representing the scale,
- $n = 1, 2, 3, \ldots, 2^{m-1}$,
- k is the degree of Chebyshev polynomial of first kind,
- x is the normalized time.

The Chebyshev wavelets of 1st kind are defined on the interval $[0, 1)$ by,

$$\psi^C_{m,n,k}(x) = \begin{cases} 2^{m/2} \tilde{T}_k(2^m x - 2n + 1), & \frac{n-1}{2^{m-1}} \le x < \frac{n}{2^{m-1}}, \\ 0, & \text{otherwise}, \end{cases} \tag{2.79}$$

where,

$$\tilde{T}_k(x) = \begin{cases} \frac{1}{\sqrt{\pi}}, & k = 0, \\ \sqrt{\frac{2}{\pi}} T_k(x), & k > 0, \end{cases} \tag{2.80}$$

and $k = 0, 1, \cdots, M-1$, $n = 1, 2, \cdots, 2^{m-1}$.

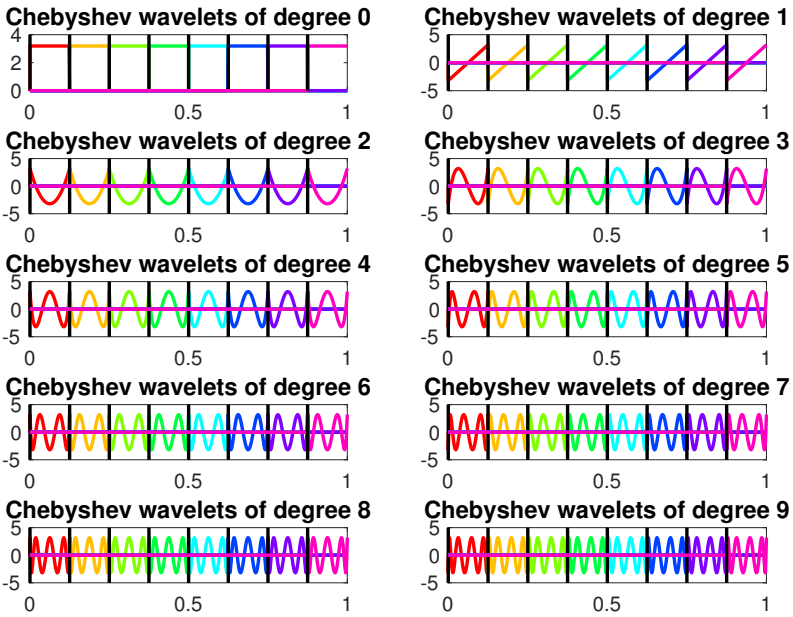


Figure 2.4: First kind Chebyshev wavelets for different degrees, with $m = 4$.

Chebyshev wavelets form an orthonormal basis of the space $L^2_\varpi([0,1))$ which is the space of weighted square integrable functions, where $\varpi(x) = \frac{1}{2\sqrt{x(1-x)}}$ as in (2.72). Let $g, f \in L^2_\varpi([0,1)$ then the inner product for this space is given by

$$\langle g, f\rangle = \int_0^1 g(x)\bar{f}(x)\varpi(x)dx. \tag{2.81}$$

This inner product has the same properties as the standard inner product presented previously in Definition 2.2.14. The weight functions for the Chebyshev wavelets are,

$$\varpi_m(x) = \begin{cases} \varpi_{1,m}(x), & 0 \le t < \frac{1}{2^{m-1}}, \\ \varpi_{2,m}(x), & \frac{1}{2^{m-1}} \le x < \frac{2}{2^{m-1}}, \\ \vdots & \vdots \\ \varpi_{2^{m-1},m}(x), & \frac{2^{m-1}-1}{2^{m-1}} \le x < 1, \end{cases} \tag{2.82}$$

where $\varpi_{n,m}(x) = \varpi(2^{m-1}x - n + 1)$.

Generalising to an interval $[a,b)$ Chebyshev wavelets are defined by,

$$\psi^C_{m,n,k}(x) = \begin{cases} 2^{m/2}\tilde{T}_k(2^m x - 2n + 1), & a + \frac{(b-a)(n-1)}{2^{m-1}} \le x < a + \frac{(b-a)n}{2^{m-1}}, \\ 0, & \text{otherwise.} \end{cases} \tag{2.83}$$

And they are orthonormal with respect to the weight function,

$$\varpi_m(x) = \begin{cases} \varpi_{1,m}(x), & a \le x < a + \frac{(b-a)}{2^{m-1}}, \\ \varpi_{2,m}(x), & a + \frac{(b-a)}{2^{m-1}} \le x < a + \frac{2(b-a)}{2^{m-1}}, \\ \vdots & \vdots \\ \varpi_{2^{m-1},m}(x), & a + \frac{(2^{m-1}-1)(b-a)}{2^{m-1}} \le x < b. \end{cases} \tag{2.84}$$

Proposition 2.5.3. *The Fourier transform of the Chebyshev wavelets of the first kind in a general interval $[a,b)$ is given by,*

$$\hat{\psi}^C_{m,n,k}(\omega) = 2^{-m/2}\sqrt{\frac{2}{\pi}} e^{-i\omega\left(\frac{2n-1}{2^m}\right)} \hat{T}_k\left(\frac{\omega}{2^m}\right). \tag{2.85}$$

Proof.

$$\begin{aligned} \hat{\psi}^C_{m,n,k}(\omega) &= \int_{\mathbb{R}} e^{-i\omega x} \psi^C_{m,k}(x)\, dx \\ &= 2^{m/2} \int_{\frac{n-1}{2^{m-1}}}^{\frac{n}{2^{m-1}}} e^{-i\omega x} \tilde{T}_k(2^m x - 2n + 1)\, dx. \end{aligned} \tag{2.86}$$

For $k > 0$,

$$\begin{aligned} \hat{\psi}^C_{m,n,k}(\omega) &= 2^{m/2}\sqrt{\frac{2}{\pi}} \int_{\frac{n-1}{2^{m-1}}}^{\frac{n}{2^{m-1}}} e^{-i\omega x} T_k(2^m x - 2n + 1)\, dx \\ &= 2^{m/2}\sqrt{\frac{2}{\pi}} \int_{\frac{n-1}{2^{m-1}}}^{\frac{n}{2^{m-1}}} e^{-i\omega x} \cos(k \arccos(2^m x - 2n + 1))\, dx. \end{aligned} \tag{2.87}$$

By using a change of variables $y = 2^m x - 2n + 1$, we end up with,

$$\begin{aligned} \hat{\psi}^C_{m,n,k}(\omega) &= 2^{-m/2}\sqrt{\frac{2}{\pi}} \int_{-1}^{1} e^{-i\omega\left(\frac{y+2n-1}{2^m}\right)} \cos(k \arccos(y))\, dy \\ &= 2^{-m/2}\sqrt{\frac{2}{\pi}} e^{-i\omega\left(\frac{2n-1}{2^m}\right)} \int_{-1}^{1} e^{-i\omega\frac{y}{2^m}} \cos(k \arccos(y))\, dy, \end{aligned} \tag{2.88}$$

which equals (2.85). □

2.6. INVERSION METHODS

This section contains the description of the Fourier inversion methods based on the previously mentioned wavelets.

Given an unknown function $f \in L^2(\mathbb{R})$ for which we know its Fourier transform $\hat{f}$, the Fourier inversion method consists of using the $\hat{f}$ to recover f. To make this statement more precise, for an orthonormal basis $\{\phi_n\}_{n=0}^{\infty}$, from (2.19) we know we can express f as,

$$f = \sum_{n=0}^{\infty} \langle f, \phi_n \rangle \phi_n. \tag{2.89}$$

Fourier inversion methods are based on choosing appropriate basis functions so that one can express the inner product in (2.89) in terms of $\hat{f}$.

As the reader can notice in this section, we consider for Haar and Shannon basis father wavelets (or scaling functions) rather than mother wavelets due to its simplicity.

2.6.1. HAAR WAVELET APPROXIMATION (WA)

The WA (wavelet approximation) method is introduced in [54] and is based on a wavelet expansion using the Haar scaling function from Section 2.5.1.

Let us consider a function $f \in L^2(\mathbb{R})$ and its Fourier transform $\hat{f}$ whenever it exists. We restrict ourselves to a bounded interval $[a,b]$, since $f \in L^2(\mathbb{R})$ we can expect that the mass in the tails of f decays to zero, so it can be well approximated in a finite interval $[a,b]$ by,

$$f^{[a,b]}(x) = \begin{cases} f(x), & \text{if } x \in [a,b], \\ 0, & \text{otherwise.} \end{cases} \tag{2.90}$$

Since Haar wavelets form an MRA, given any m, we can use the right hand side of formula (2.59) to approximate $f^{[a,b]}(x) \approx f_m^{[a,b]}(x)$ for all $x \in [a,b]$, where,

$$f(x) \approx f_m^{[a,b]}(x) := \sum_{k=0}^{2^m-1} c_{m,k}^H \phi_{m,k}^H \left(\frac{x-a}{b-a}\right), \tag{2.91}$$

with convergence in the $L^2(\mathbb{R})$-norm. Here $c_{m,k}^H$ are the wavelet coefficients and the Haar wavelet family $\{\phi_{m,k}^H\}_{k\in\mathbb{Z}}$ is defined by $\phi_{m,k}^H(x) := 2^{m/2}\phi_H(2^m x - k)$ (according to Theorem 2.4.5) and for a fixed scale m forms an orthonormal basis of V_m.

Observe that the sum goes from 0 to 2^m-1. The reason for this is that at scale m, we divide the interval $[a,b]$ in 2^m disjoint intervals of equal size such that the union of all of them is the full interval. For every $k = 0,\dots,2^m-1$, we get a child wavelet with support in one of these intervals.

Note that we will specify later on how to choose the bounded interval $[a,b]$ according to the application of the method. As we will see, usually a set of functions called cumulants is employed. In other situations, maybe the function to approximate is already defined in an interval. For the scale of approximation m, according to the specific application, a range of values is normally used based on the performance appreciated in numerical experiments.

COEFFICIENTS COMPUTATION

The main idea behind WA is to approximate $\hat{f}$ by $\hat{f}_m^{[a,b]}$ and to compute the wavelet coefficients $c_{m,k}^H$ by inverting the Fourier transform. Proceeding this way, as in [54], we have,

$$\begin{aligned} \hat{f}(\omega) &= \int_{\mathbb{R}} f(y) e^{-i\omega y} dy \approx \int_{\mathbb{R}} f_m^{[a,b]}(y) e^{-i\omega y} dy \\ &= \sum_{k=0}^{2^m-1} c_{m,k}^H \left(\int_{\mathbb{R}} \phi_{m,k}^H \left(\frac{y-a}{b-a}\right) e^{-i\omega y} dy \right). \end{aligned} \tag{2.92}$$

Introducing the change of variables $x = \frac{y-a}{b-a}$,

$$\begin{aligned} \hat{f}(\omega) &\approx (b-a) e^{-ia\omega} \sum_{k=0}^{2^m-1} c_{m,k}^H \int_{\mathbb{R}} e^{-i\omega(b-a)x} \phi_{m,k}^H(x) dx \\ &= (b-a) e^{-ia\omega} \sum_{k=0}^{2^m-1} c_{m,k}^H \hat{\phi}_{m,k}^H((b-a)\omega). \end{aligned} \tag{2.93}$$

Note,

$$\phi_{m,k}^H(\xi) = 2^{-\frac{m}{2}} \hat{\phi}_H\left(\frac{\xi}{2^m}\right) e^{-i\frac{k}{2^m}\xi}. \tag{2.94}$$

Considering the change of variables $z = e^{-i\frac{(b-a)}{2^m}\omega}$, we have that

$$\hat{f}\left(\frac{2^m}{b-a} i\log(z)\right) \approx 2^{-\frac{m}{2}}(b-a) z^{\frac{2^m a}{b-a}} \hat{\phi}_H\left(i\log(z)\right) \sum_{k=0}^{2^m-1} c_{m,k}^H z^k. \tag{2.95}$$

And defining,

$$P_m(z) := \sum_{k=0}^{2^m-1} c_{m,k}^H z^k, \qquad Q_m(z) := \frac{2^{m/2} z^{\frac{2^m a}{b-a}} \hat{f}\left(\frac{2^m}{b-a} i\log(z)\right)}{(b-a)\hat{\phi}_H(i\log(z))}, \tag{2.96}$$

we have that $P_m(z) \approx Q_m(z)$. Since P_m is a polynomial, it is analytic inside a disc of the complex plane $\{z \in \mathbb{C} : |z| < \rho\}$ for $\rho > 0$. Thus, by means of the Cauchy's integral formula we obtain the following expression for the coefficients,

$$c_{m,k}^H = \frac{1}{2\pi i} \int_\gamma \frac{P_m(z)}{z^{k+1}} dz, \quad k = 0, \cdots, 2^m - 1, \tag{2.97}$$

where γ denotes a circle of radius ρ, $\rho > 0$, about the origin. As pointed out in [64], we take $\rho = 0.9995$, as the optimal value to minimize the discretization and roundoff errors when computing the coefficients numerically.

Now, applying the change of variables $z = \rho e^{iu}$ and the approximation $P_m \approx Q_m$, we end up with the coefficients expression,

$$c_{m,k}^H \approx \frac{1}{2\pi\rho^k} \int_0^{2\pi} Q_m(\rho e^{iu}) e^{-iku} du. \tag{2.98}$$

Remark 2.6.1. *The integral in (2.98) is approximated by the trapezoidal rule and computed using the FFT. The approximation of the coefficients with the trapezoidal rule is,*

$$c_{m,k}^H \approx c_{m,k}^{H,*} := \frac{1}{\rho^k M} \sum_{n=0}^{M-1} Q_m(\rho e^{i\frac{2\pi}{M}n}) e^{-i2\pi\frac{k}{M}n}, \tag{2.99}$$

where $M = 2^m$. Note here that we have used the fact that the integrand for $n = 0$ is the same as $n = M$, and we have added them together. The use of the FFT in the sum expression (2.99) is straightforward.

2.6.2. SHANNON WAVELET INVERSE FOURIER TECHNIQUE (SWIFT)

Given a function $f \in L^2(\mathbb{R})$, we consider its expansion in terms of the Shannon scaling function (see Section 2.5.2) at the level of resolution m. The aim is to recover the coefficients of the approximation from the Fourier transform of the function f which is assumed to be known in closed-form. The SWIFT method introduced in [67] effectively achieves this purpose.

As mentioned previously, the Shannon scaling function forms a multiresolution analysis. Following the MRA theory (according to Proposition 2.4.9), a function $f \in L^2(\mathbb{R})$ can be approximated at the level of resolution m by,

$$f(x) \approx \mathscr{P}_m f(x) = \sum_{k\in\mathbb{Z}} c_{m,k}^S \phi_{m,k}^S(x), \tag{2.100}$$

where $\mathscr{P}_m f$ converges to f in $L^2(\mathbb{R})$, i.e. $\|f - \mathscr{P}_m f\|_2 \to 0$, when $m \to +\infty$. Here, the coefficients $c^S_{m,k}$ are defined in (2.61) and $\phi^S_{m,k}(x) := 2^{m/2}\phi_S(2^m x - k)$ from (2.49), where ϕ_S is defined in (2.65).

The infinite series in (2.100) is well-approximated (see Lemma 1 of [67] for details) by a finite summation without loss of considerable density mass, since $f \in L^2(\mathbb{R})$,

$$\mathscr{P}_m f(x) \approx f_m(x) := \sum_{k=k_1}^{k_2} c^S_{m,k}\phi^S_{m,k}(x), \tag{2.101}$$

for certain accurately chosen values k_1 and k_2.

COEFFICIENTS COMPUTATION

The next step is the computation of the coefficients in (2.101). Recalling (2.61) and (2.49), we have that,

$$c^S_{m,k} = \langle f, \phi^S_{m,k}\rangle = \int_{\mathbb{R}} f(x)\overline{\phi}^S_{m,k}(x)\,\mathrm{d}x = 2^{m/2}\int_{\mathbb{R}} f(x)\phi_S(2^m x - k)\mathrm{d}x. \tag{2.102}$$

Using the classical Vieta's formula [32], the cardinal sine can be expressed as the following infinite product,

$$\phi_S(x) = \mathrm{sinc}(x) = \prod_{j=1}^{+\infty}\cos\left(\frac{\pi x}{2^j}\right). \tag{2.103}$$

If we truncate the infinite product (2.103) into a finite product with a total of J terms, then, thanks to the cosine product-to-sum identity [71], we have,

$$\prod_{j=1}^{J}\cos\left(\frac{\pi x}{2^j}\right) = \frac{1}{2^{J-1}}\sum_{j=1}^{2^{J-1}}\cos\left(\frac{2j-1}{2^J}\pi x\right). \tag{2.104}$$

By (2.103) and (2.104) the sinc function can thus be approximated as,

$$\phi_S(x) = \mathrm{sinc}(x) \approx \mathrm{sinc}^*(x) := \frac{1}{2^{J-1}}\sum_{j=1}^{2^{J-1}}\cos\left(\frac{2j-1}{2^J}\pi x\right). \tag{2.105}$$

Replacing the function ϕ_S in (2.102) by the approximation (2.105) we obtain,

$$c^S_{m,k} \approx c^{S,*}_{m,k} = \frac{2^{m/2}}{2^{J-1}}\sum_{j=1}^{2^{J-1}}\int_{\mathbb{R}} f(x)\cos\left(\frac{2j-1}{2^J}\pi(2^m x - k)\right)\mathrm{d}x. \tag{2.106}$$

Next, by taking into account that $\Re\left(\hat{f}(\xi)\right) = \int_{\mathbb{R}} f(x)\cos(\xi x)\,\mathrm{d}x$ in (2.25), where $\Re(z)$ denotes the real part of z, and observing that,

$$\hat{f}(\xi)e^{ik\pi\frac{2j-1}{2^J}} = \int_{\mathbb{R}} e^{-i\left(\xi x - \frac{k\pi(2j-1)}{2^J}\right)}f(x)\mathrm{d}x, \tag{2.107}$$

we can simplify (2.106) to,

$$c^{S,*}_{m,k} = \frac{2^{m/2}}{2^{J-1}}\sum_{j=1}^{2^{J-1}}\Re\left[\hat{f}\left(\frac{(2j-1)\pi 2^m}{2^J}\right)e^{\frac{ik\pi(2j-1)}{2^J}}\right]. \tag{2.108}$$

Remark 2.6.2. *Formula (2.108) can be conveniently rearranged to compute the coefficients with the use of the FFT.*

PARAMETERS SELECTION

The set of parameters that are important in the method are the scale of approximation m, the Vieta's formula truncation value J, and the truncation parameters k_1 and k_2.

Usually k_1 and k_2 are also computed by the use of cumulants, but it depends on the application. We will discuss their choice in more detail later in Definition 3.2.9 in a specific application of the method.

The scale of approximation is selected according to [51], where a procedure to choose the scale parameter m based on band-limited functions is presented.

Definition 2.6.3. *A function f is called **band-limited** if there exists a positive constant $B < \infty$ such that its Fourier transform $\hat{f}$ is identically zero in $|\omega| > B\pi$, i.e.,*

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-B\pi}^{B\pi} \hat{f}(\omega) e^{i\omega x} d\omega. \tag{2.109}$$

We can state that if a function f is band-limited with $B < 2^m$, then there exists an ω such that for all $|\omega| > 2^m\pi$, $|\hat{f}(\omega)| < \text{tol}$, for a user defined $\text{tol} > 0$. As we will see, although the functions we work with are not band-limited, the modulus of their characteristic functions tend to zero for $\omega \to \infty$, so we can choose the scale of approximation such that the coefficients are negligibly small. Hence, we define a tolerance level tol a priori, and we search for the smallest value m, such that $|\hat{f}(\omega)| < \text{tol}$.

A strategy for choosing J follows from Lemma 2 of [67] which states the following.

Lemma 2.6.4. *If we define the absolute error $\varepsilon(x) := sinc(x) - \frac{1}{2^{J-1}} \sum_{j=1}^{2^{J-1}} \cos\left(\frac{2j-1}{2^J}\pi x\right)$, then,*

$$|\varepsilon(x)| \leq \frac{(\pi c)^2}{2^{2(J+1)} - (\pi c)^2}, \tag{2.110}$$

for $t \in [-c, c]$, where $c \in \mathbb{R}$, $c > 0$ and $J \geq \log_2(\pi c)$.

Together with what is stated in Theorem 1 of [67], we assume $F(x)$ to be the CDF of a random variable X and define $H(x) := F(-x) + 1 - F(x)$. Let $a > 0$ be a constant such that $H(a) < \epsilon$, for $\epsilon > 0$. The lemma implies that if we set $M_{m,k} := \max(|2^m a - k|, |2^m a + k|)$ and then we define $M_m := \max_{k_1 < k < k_2} M_{m,k}$, a good strategy to choose J is to take $J = \jmath := \lceil \log_2(\pi M_m) \rceil$, where $\lceil x \rceil$ denotes the smallest integer greater or equal than x. Although for each k different values of J could be chosen, we decide to choose a global one in order to benefit from the application of the FFT.

2.6.3. CHEBYSHEV WAVELET APPROXIMATION (CWA)

Recall that we focus on function approximations by first kind Chebyshev wavelets (2.5.3). As we explained in Section 2.5.3, $\{\psi_{m,n,k}\}_{n \in \mathbb{N}^+, k \in \mathbb{N}}$ form an orthonormal basis for $L^2_\omega([a,b])$. Hence, according to Theorem 2.2.21, any function in this space may be expanded as,

$$f(x) = \sum_{n=1}^{\infty} \sum_{k=0}^{\infty} c^C_{m,n,k} \psi^C_{m,n,k}(x), \tag{2.111}$$

where,

$$c^C_{m,n,k} = \langle f(x), \psi^C_{m,n,k}(x) \rangle_{\omega_m}. \tag{2.112}$$

The series (2.111) is truncated as

$$f(x) \approx \sum_{n=1}^{2^{m-1}} \sum_{k=0}^{M-1} c^C_{m,n,k} \psi^C_{m,n,k}(x). \tag{2.113}$$

The sum for the n index goes from 1 to 2^{m-1}. The reason for this is that the interval $[a,b)$ is divided in 2^{m-1} disjoint intervals of equal size such that the union of all of them is the full interval. Hence for any fixed $n \in \{1,...2^{m-1}\}$, the wavelets $\psi_{m,n,k}$ are supported in the n-th interval. On the other side, M is an arbitrary number, which refers to the degree of the Chebyshev polynomials used in the approximations. The higher M is, the more precise the approximation becomes.

Note that we have already defined the method by considering the function to approximate f in an interval $[a,b)$. If it was not the case, the function would of course need to be truncated, again by probably the use of cumulants. Also, the scale of approximation m needs to be defined. Note that it appears here a different parameter with respect to the other methods which is the maximum degree of the polynomials considered, M.

COEFFICIENTS COMPUTATION

Our goal is to get an expression for the coefficients in (2.112). We proceed by using the inverse Fourier transform and the Bessel function as follows.

Let us consider the inverse Fourier transform,

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(\omega) \mathrm{e}^{i\omega x} d\omega, \tag{2.114}$$

and place it inside (2.112):

$$\begin{aligned} c^C_{m,n,k} &= \langle f(x), \psi^C_{m,n,k}(x) \rangle_{\varpi_m} = \int_{\mathbb{R}} f(x) \psi^C_{m,k}(x) \varpi_m(x) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{f}(\omega) \mathrm{e}^{i\omega x} \psi^C_{m,n,k}(x) \varpi_m(x) d\omega dx. \end{aligned} \tag{2.115}$$

Using Fubini's theorem and replacing the weight expression (2.82) inside the previous integral (2.115) we end up with the following expression. For $k > 0$,

$$c^C_{m,n,k} = \frac{1}{2^{m/2}} \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{2\pi}} \int_0^1 \int_{\mathbb{R}} \hat{f}(\omega) e^{i\omega \left(\frac{\cos(y) + 2n - 1}{2^m} \right)} d\omega. \tag{2.116}$$

Using the Bessel function, defined by,

$$J_n(z) = \frac{i^{-n}}{\pi} \int_0^\pi e^{iz\cos(\theta)\cos(n\theta)} d\theta, \tag{2.117}$$

we obtain,

$$c^C_{m,n,k} = \frac{i^k}{2^{m/2}} \int_{\mathbb{R}} e^{i\omega \left(\frac{2n-1}{2^m} \right)} \hat{f}(\omega) J_k\left(\frac{\omega}{2^m} \right) d\omega. \tag{2.118}$$

For $k = 0$ the expression is the same except from a constant.

Remark 2.6.5. *The Bessel function is implemented efficiently in Matlab which is the software we use.*

Note that we may apply the FFT in (2.118) after a numerical integration. That is one of the reasons we choose this coefficients computation technique as the preferred one.

2.7. CONCLUSIONS

Through this chapter, we have introduced the mathematical foundations that lay underneath the research we present in the following chapters; to this end, we presented basic concepts of probability theory and functional analysis. After that, we introduced Fourier analysis and wavelet theory. Using the theory from the functional analysis section we have justified how both Fourier and wavelet analysis can be used to approximate functions in $L^2(\mathbb{R})$. Once the main concepts of wavelet theory were introduced, we presented the wavelet families that we will be using throughout the thesis: Haar, Shannon and Chebyshev wavelets. We closed the chapter presenting the Fourier inversion methods, and how to describe and implement them efficiently combining concepts introduced in Fourier and wavelet theory.

Haar and Shannon wavelet families have been widely used in the literature, and algorithms to implement Fourier inversion methods using them already exist. On the other hand, Chebyshev wavelets have been introduced recently and to the best of our knowledge there has been no previous work towards applying them to the Fourier inversion method. Hence, the computation of the coefficients based on the use of the Bessel function is novel. Nevertheless, note that approximation coefficients from the Chebyshev method have an extra parameter because of the double sum on the approximation, thus the approach does not look promising so far.

The inversion methods are based on the assumption that the Fourier transform of the function to approximate is known. Since we aim to propose methods to approximate functions in a competitive computational time, we emphasize that FFT algorithm in the coefficients computation of the three inversion methods can be applied.

CHAPTER 3

A Novel Two-Dimensional Technique for Pricing European Options

The main aim of this chapter is to introduce the two-dimensional SWIFT method, a Fourier inversion method based on Shannon wavelets used for pricing European rainbow options (developed through the scope of the thesis and published in the paper [22]).

With this objective in mind, this chapter starts by introducing the reader in the option pricing area, presenting then the collection of the wavelet methods used for one-dimensional European option pricing (based on Haar, Shannon and Chebyshev wavelets) and finalizing by putting forward the SWIFT multidimensional extension.

This chapter is based on the books [9], [42] and on the articles [22], [65] and [67].

3.1. INTRODUCTION

In the last 50 years, derivatives have become increasingly important in finance, options being actively traded on many exchanges through the world. A derivative can be defined as a financial instrument whose value depends on the values of others, called underlying assets. Option contracts are a type of derivatives that give the holder a right, but not the obligation to exercise the contract. Options are traded both on exchanges and in over-the-counter markets. There are two types of options: a call option, which gives the holder the right to buy the underlying asset by a certain date for a certain price, and a put option, which gives the holder the right to sell the underlying asset by a certain date for a certain price. Options are traded on a wide range of different underlying assets. The price in the contract is known as the exercise price or strike price, the date in the contract is known as the expiration date or maturity. There are two sides to every option contract: the investor who has bought the option is said to have taken a long position, and the investor who has sold the option has taken a short position.

We can distinguish different options according to when they can be exercised. The most basic ones, European options, can be exercised only on the expiration date itself. American options can be exercised at any time point before their expiration date. Exotic options differ in structure from common American or European options in terms of the underlying asset, or the calculation of how or when the investor receives a certain payoff.

The optionality in exercising the right written in the option should be emphasized. This is what distinguishes options from other contracts like forwards and futures, where the holder is obliged to buy or sell the underlying asset. Whereas it costs nothing to enter into a forward or futures contract, there is a cost to acquiring an option. Option pricing i.e., knowing the value of the option (or the fair price of the contract) before maturity, is one of the big challenges in the financial world. Option pricing theory has made vast strides since 1973, when Black

and Scholes published their ground-breaking paper providing a model for valuing dividend-protected European options. The Black-Scholes (BS) model makes certain assumptions: the option must be European, no dividends are paid out during the life of the option, efficiency of the markets, no transaction costs in buying the option, risk-free rate and volatility of the underlying are known and constant, and returns on the underlying follow predetermined dynamics called geometric Brownian motion (GBM).

In the BS model, the price of the option is obtained as the solution of a partial differential equation. The solution of this PDE can alternatively be written as a conditional expectation by means of the Feynman-Kac theorem. This theorem shows that the conditional expectation of the value of a contract payoff function under the risk-neutral measure is equivalent to the solution of a partial differential equation and under the previous assumption, this partial differential equation is the BS model and hence provides us with a closed formula for its solution. Moreover, it is also useful within the Fourier inversion framework as we will see later.

An analytical formula for the option price is generally unavailable. Hence, there is a need to develop numerical methods to approximate their prices and develop efficient algorithms to implement them, so that they provide useful information in a market that changes rapidly. The existing numerical methods for option pricing can be classified into three major groups: Monte Carlo (MC) simulation, partial-(integro) differential equation (PIDE or PDE) methods, and Fourier inversion methods. This thesis is focused on the subset of the Fourier inversion methods that make use of wavelets. The wavelet inversion methods we work with are based on knowing the Fourier transform of the probability density function corresponding to the underlying stochastic process. That is why we focus on the subset of asset models called exponential Lévy processes for which the characteristic function is available. Exponential Lévy models generalize the classical Black and Scholes set-up by allowing the stock prices to jump while preserving the independence and stationarity of returns.

On the pricing of one-dimensional European options, we study Fourier inversion methods based on Haar, Shannon and Chebyshev wavelets. From those methods, Shannon wavelets are the ones that give better performance and that is why the method is chosen to be extended to multidimensional option pricing. We focus on extending the Shannon wavelets Fourier inversion methods for pricing a class of multidimensional exotic options, known as rainbow options. Analytical formulae to price multi-asset derivatives are only available for the simplest cases and hence numerical solutions are necessary. However, numerical approximations are very expensive to obtain, despite the efforts in the financial community to come up with optimal methods. One of the most commonly used methods for pricing multidimensional options is Monte Carlo simulation. This method has the advantage of scaling linearly with the number of dimensions. However, convergence is slow and a large number of simulations is needed if accurate results are desired. On the other hand, the numerical complexity of both PDE and Fourier-based methods increase exponentially with dimension, a phenomenon known as curse of dimensionality. For this reason, and despite their drawbacks, Monte Carlo methods are the most commonly used methods when the dimension is larger than four or five, depending on the specific product. Nevertheless, in the cases in which the characteristic function of the multidimensional process is known, even for higher dimensions than one, Fourier methods may be efficiently implemented so their numerical complexity is lower than Monte Carlo simulation when the dimension is not very large.

One well-developed multidimensional Fourier-based method is the multidimensional COS

method presented in [74] and called 2D-COS when the dimension is two. However, it may exhibit problems in the vicinity of the integration boundaries because of the periodic behaviour of cosines. This problem becomes evident for long maturity options, where round-off errors may accumulate near domain boundaries. Also for short maturity options, typically governed by a highly peaked density function, many cosine terms may be needed for an accurate representation. In addition, an accurate integration interval is important to capture the whole mass of the recovered density, but the choice of the interval is entirely based on the cumulants, which sometimes are not easily available or do not provide a good truncation range.

In the one-dimensional case, local wavelets bases have been considered in [65, 67], which overcome some of the problems of the known one-dimensional COS method [28]. Wavelets give flexibility and enhance robustness when pricing long maturity options and heavy tailed asset processes. We aim to extend the one-dimensional SWIFT method to higher dimensions to be able to price European-style financial contracts with a payoff depending on more than one asset. For two-dimensional contracts we call it 2D-SWIFT. We provide an error analysis which facilitates the choice of the parameters of the new method and enhance the overall speed with a fast Fourier transform (FFT) algorithm. We test our method with examples and methods from the literature, like for instance the well-known 2D-COS method. We price basket (both geometric and arithmetic), spread, call-on-max, put-on-min, and correlation options. The numerical experiments show that the 2D-SWIFT is a very competitive pricing algorithm. The method benefits from the fact that we do not need to truncate the integration range a priori, the density coefficients can be computed with an FFT algorithm and the number of terms can be automatically calculated once settled up the precision in the approximation, achieving great precision with short running times.

This chapter is organized as follows. Section 3.2 is an introduction to the option pricing topic, it covers the basic concepts, as well as the Black-Scholes setting, the idea of Lévy processes, and the notion of the different approaches for option pricing. Inside it, the three one-dimensional option pricing methods based on wavelets are also presented: Haar, Shannon and Chebyshev. Then, we move to the main section of the chapter, Section 3.3, in which two-dimensional SWIFT for option pricing is shown: first, it is introduced the multidimensional wavelet framework, then the 2D-SWIFT pricing formula derivation with a rigorous study of the error analysis of the method is presented, and at the end, a full set of numerical examples to show the strength of the method is provided. Finally, Section 3.4 is a brief conclusion about the aspects treated in the chapter.

3.2. EUROPEAN OPTION PRICING FRAMEWORK

We start by formally introducing the European call and put options also called vanilla options because of their simplicity.

Definition 3.2.1. *A **European call option** with exercise price (or strike price) K and maturity time (exercise date) T on the underlying asset (with price S_t at time t) is a contract defined by the following clauses: the holder of the contract has, exactly at time T, the right to buy one unit (or one share) of the underlying stock at the price K from the writer of the option; the holder of the option is in no way obliged to buy the underlying stock.*

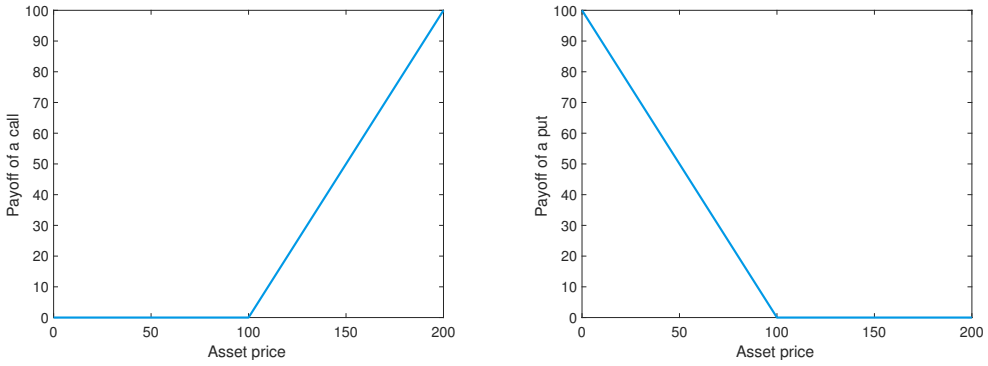
*A **European put option** is an option which in the same way gives the holder the right to sell a unit (or a share) of the underlying asset at a predetermined strike price at time T, with no obligation on selling.*

For convenience, we refer to the underlying asset as stock. The payoff of an option on the expiration date is determined by the price of the underlying asset. The payoffs for a European call option and European put options are as follows,

$$\begin{aligned} \text{Call option payoff} &= (S_T - K)^+ = \max(S_T - K, 0), \\ \text{Put option payoff} &= (K - S_T)^+ = \max(K - S_T, 0), \end{aligned} \tag{3.1}$$

where S_T is equal to the spot price of the underlying security at time T and K equals the strike price of the option.

Note that payoffs are never negative. We can plot them at expiration as function of the price of the underlying asset, see Figure 3.1.



(a) Payoff of a call. (b) Payoff of a put.

Figure 3.1: Vanilla options payoffs, $K = 100$.

As mentioned in the introduction, the answer to the question: "what is a fair price for the contract?" is one of the more widely studied problems in finance. The first model to answer this question is attributed to Black, Scholes and Merton, and we introduce it next.

3.2.1. LÉVY PROCESSES

BLACK-SCHOLES SETTING

The Black-Scholes formula (also called Black-Scholes-Merton) was the first widely used model for option pricing. It is used to calculate the theoretical value of European-style options using current stock prices, expected dividends, the option's strike price, expected interest rates, time to expiration and expected volatility. The formula, developed by three economists: Fischer Black, Myron Scholes and Robert Merton, is perhaps the world's most well-known option pricing model, and was introduced in 1973 in their paper, "The Pricing of Options and Corporate Liabilities" published in the Journal of Political Economy, [10]. Black passed away two years before Scholes and Merton were awarded the 1997 Nobel Prize in Economics for their work in finding a new method to determine the value of derivatives (the Nobel Prize is not given posthumously; however, the Nobel committee acknowledged Black's role in the Black-Scholes model).

The basic tool to understand continuous-time finance is the Brownian motion (also known as Wiener process). To introduce it, first let us define what is a stochastic process, which is a

mathematical object consisting of a (countable or uncountable) collection of random variables indexed by time.

Definition 3.2.2. *Suppose that* $(\Sigma, \mathscr{F}, \mathbb{P})$ *is a probability space, and that* $\mathscr{T}$ *is a subset of* $[0,\infty)$. *Suppose further that for each* $t \in \mathscr{T}$, *there is a random variable* $X_t : \Omega \to \mathbb{R}$ *defined on* $(\Sigma, \mathscr{F}, \mathbb{P})$. *A family of random variables indexed by* $\mathscr{T}$, $X = \{X_t, t \in \mathscr{T}\}$, *is called a* ***stochastic process****.*

A discrete time stochastic process is a countable collection of random variables indexed by the non-negative integers, $\mathscr{T} = \mathbb{N}$, *and a continuous time stochastic process is an uncountable collection of random variables indexed by the non-negative real numbers,* $\mathscr{T} = [0,\infty)$.

Definition 3.2.3. *Being* t *the time variable, a* ***Brownian motion*** *(BM) or Wiener process is a continuous-time stochastic process* $\{W_t, 0 \le t < \infty\}$ *such that,*

1. $W_0 = 0$.
2. W_t *is almost surely continuous.*
3. W_t *has independent increments.*
4. $W_t - W_s \sim \mathcal{N}(0, t-s)$ *for* $0 \le s \le t$,

where $\mathcal{N}(\mu, \sigma^2)$ *denotes the normal distribution with expected value* μ *and variance* σ^2. *The condition that it has independent increments means that if* $0 \le s_1 < t_1 \le s_2 < t_2$ *then* $W_{t_1} - W_{s_1}$ *and* $W_{t_2} - W_{s_2}$ *are independent random variables.*

BM is not appropriate for modelling stock prices; the most obvious reason is that it can take negative values. We introduce here a non-negative variation of BM called geometric Brownian motion, which is a popular model for generating stock prices and leads to the Black-Scholes model.

Definition 3.2.4. *A stochastic process* $\{S_t, t \in \mathscr{T}\}$ *is said to follow a* ***geometric Brownian motion*** *if it satisfies the following stochastic differential equation (SDE),*

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t, \tag{3.2}$$

where W_t *is a Brownian motion, and* μ, *known as the percentage drift, and* σ, *known as the percentage volatility, are constants.*

Note that for an arbitrary initial value S_0 the above SDE has the analytic solution,

$$S_t = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t\right). \tag{3.3}$$

Now we turn to the Black-Scholes formula. The basic idea behind this formula is an arbitrage equilibrium among three assets: stock, bond, and European call option. It is a risk-neutral valuation because investors in their model economy were implicitly assumed to be risk neutral and they are concerned only with maximizing profits. A more realistic, discrete-time formulation of option pricing (by using a binomial tree) was later proposed by William Sharpe [78], and formalized by Cox-Ross-Rubinstein in [24], which will not be discussed here.

Theorem 3.2.5 (Black-Scholes partial differential equation)**.** *Let $v(x,t)$ denote the price at time t of a derivative (such as a European call option) of stock when $S_t = x$. Then v must satisfy the partial differential equation,*

$$\frac{\partial v}{\partial t} + \frac{\partial v}{\partial x} r x + \frac{1}{2}\frac{\partial^2 v}{\partial x^2}\sigma^2 x^2 - r v = 0, \tag{3.4}$$

where, r is the interest rate and σ the volatility. With a final condition being the payoff of the option at the maturity time.

There exists a relationship between call and put options, which naturally arises because of the arbitrage conditions in financial markets equilibrium. It is called the put-call parity relation, if v_p is the price of the put and v_c the price of the call, then,

$$v_c + Ke^{-rT} = v_p + S_0, \tag{3.5}$$

at time $t = 0$. The formula shows that the value of a European call with a certain exercise price and exercise date can be deduced from the value of a European put with the same exercise price and exercise date, and viceversa.

It remains to introduce the Black-Scholes pricing formula. The approach adopted by Black and Scholes consisted in transforming the Black Scholes PDE into the heat equation [10], for which a solution is known. But we will alternatively present it by the application of the Feynman-Kac theorem. Feynman-Kac theorem establishes a link between parabolic partial differential equations and stochastic processes.

Theorem 3.2.6 (Feynman-Kac)**.** *Suppose that $x = x_t$ follows the process,*

$$dx = \mu(x,t)dt + \sigma(x,t)dW_t^{\mathbb{Q}}, \tag{3.6}$$

under the risk neutral probability measure $\mathbb{Q}$, with $W_t^{\mathbb{Q}}$ a Wiener process. And suppose that for all $x \in \mathbb{R}$ and $t \in [0,T]$ a sufficiently differential function $v = v(x,t)$ follows the partial differentiable equation given by,

$$\frac{\partial v}{\partial t} + \mu(x,t)\frac{\partial v}{\partial x} + \frac{1}{2}\sigma(x,t)^2\frac{\partial^2 v}{\partial x^2} - r(x,t)v(x,t) = 0, \tag{3.7}$$

with boundary condition $v(x,T)$, where μ, σ, r are known functions. Then, $v(x,t)$ can be written as,

$$v(x,t) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r(x,u)du} v(x,T) | \mathscr{F}_t\right], \tag{3.8}$$

where $\mathbb{E}^{\mathbb{Q}}$ represents the expected value with respect to the risk neutral measure $\mathbb{Q}$ and $\mathscr{F}_t$ is a filtration.

To apply the Feynman-Kac theorem to the Black-Scholes call price, one can note that the value $v_c = v(S_t,t)$ of a European call option written at time t with strike price K when interest rates are constant r follows the Black-Scholes PDE (3.4) with boundary condition $v(S_T,T) = (S_T - K)^+$. This PDE is the PDE in Equation (3.7) with $x = S_t$, $\mu(x,t) = rS_t$, and $\sigma(x,t) = \sigma S_t$. Hence, the Feynman-Kac theorem applies and the value of the European call is,

$$v_c = v(S_t,t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[(S_T - K)^+ | \mathscr{F}_t\right], \tag{3.9}$$

which can be evaluated by straightforward integration and gives the call price that we present in the next theorem. The put price can be obtained analogously or by the put-call parity relation in (3.5).

Theorem 3.2.7 (Black-Scholes option pricing formula)**.** *When the stock price S_t follows a GBM, the price of a European call option $v_c(S_t,t)$ with time to maturity $\Delta t = T - t$ (T is maturity date and t the initial date), strike price K, with constant volatility σ and constant rate r is given by,*

$$v_c(S_t,t) = S_t\Phi(d_1) - e^{-r\Delta t}K\Phi(d_2), \tag{3.10}$$

where

$$d_1 = \frac{\log\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\Delta t}{\sigma\sqrt{\Delta t}}, \tag{3.11}$$

$d_2 = d_1 - \sigma\sqrt{\Delta t}$, *and where* $\Phi(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-\frac{1}{2}t^2}\,dt$ *is the standard normal CDF. Similarly, the price of a European put option $v_p(S_t,t)$ on the same variables is,*

$$v_p(S_t,t) = e^{-r\Delta t}K\Phi(-d_2) - S_t\Phi(-d_1). \tag{3.12}$$

LÉVY PROCESSES

As mentioned in the introduction, the Black-Scholes model relies on a number of assumptions, one of them being that the underlying asset dynamics follows a GBM. However, later research has shown that GBM has serious drawbacks at describing the stocks market. Exponential Lévy models generalize the classical Black and Scholes set-up by allowing the stock prices to jump while preserving the independence and stationary of returns. This agrees with empirical data so it makes them better fitted to model the underlying dynamics. Moreover, we are especially interested in them since the characteristic function is known so we can apply Fourier inversion methods. Lévy processes provide a framework that can easily capture the empirical observations both under the real-world and under the risk-neutral measure. The reader is referred to [69] for indicative examples and also to [23] and [77] for description of the processes.

Definition 3.2.8. *A stochastic process $\mathscr{L} = \{\mathscr{L}_t : t \geq 0\}$ is said to be a **Lévy process** if it satisfies the following properties:*

1. $\mathscr{L}_0 = 0$ *almost surely.*
2. *For any $0 \leq t_1 < t_2 < \cdots < t_n < \infty$, $\mathscr{L}_{t_2} - \mathscr{L}_{t_1}, \mathscr{L}_{t_3} - \mathscr{L}_{t_2}, \ldots, \mathscr{L}_{t_n} - \mathscr{L}_{t_{n-1}}$ are independent (independence of increments).*
3. *For any $s < t$, $\mathscr{L}_t - \mathscr{L}_s$ is equal in distribution to $\mathscr{L}_{t-s}$ (stationary increments).*
4. *For any $\epsilon > 0$ and $t \geq 0$ it holds that $\lim_{h\to 0}\mathbb{P}(|\mathscr{L}_{t+h} - \mathscr{L}_t| > \epsilon) = 0$ (continuity in probability).*

Under the real-world measure, we model the asset price process as the exponential of a Lévy process, that is,

$$S_t = S_0 \exp \mathscr{L}_t, \quad 0 \leq t \leq T, \tag{3.13}$$

where $\{\mathscr{L}_t : t \geq 0\}$ is the Lévy process whose distribution has been estimated from the data set available for the particular asset.

By the Lévy-Khintchine formula, see [23], the characteristic function of the Lévy process $\mathscr{L}_t$ equals,

$$\hat{f}_{\mathscr{L}}(\omega, t) = \exp - t\Psi_{\mathscr{L}}(\omega), \tag{3.14}$$

where $\Psi_{\mathscr{L}}$ is the characteristic exponent defined as $\Psi_{\mathscr{L}}(\omega) := \log \mathbb{E}\left[\exp(i\omega\mathscr{L}_1)\right]$, and its existence is proved in [69].

3.2.2. NUMERICAL METHODS

As stated previously, the numerical methods used for option pricing can be classified in three major categories.

MONTE CARLO METHODS

A widely used method for pricing options is to use a Monte Carlo simulation. The main advantage of this method is that complex and exotic derivatives can be treated easily – which is very important in applications. Moreover, options on several assets can also be handled easily using Monte Carlo simulations. The main drawback of Monte Carlo methods is the slow computational speed due to the required large number of simulations in order to have accurate results.

PIDE/PDE METHODS

An alternative for pricing options is to derive and then solve numerically the partial integro-differential equation that the option price satisfies. Note that in their seminal paper Black and Scholes derive such a PDE for the price of a European option.

The advantage of PIDE/PDE methods is that complex and exotic payoffs can be treated easily; the limitations are the slower speed in comparison to inversion methods and the computational complexity when handling options on several assets.

FOURIER INVERSION METHODS

Another choice of methods are the Fourier inversion techniques. The pricing of European options in computational finance is governed by the numerical solution of partial differential, or partial integro-differential, equations. The corresponding solution, being the option value at time t, can be found by means of the Feynman-Kac formula as a discounted expectation of the option value at final time $t = T$, the so-called payoff function. Here we consider the risk neutral valuation formula,

$$v(x,t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[v(y,T)|x\right] = e^{-r(T-t)}\int_{\mathbb{R}} v(y,T)f(y|x)dy, \tag{3.15}$$

where v denotes the option pricing value, T is the maturity, t the initial date, $\mathbb{E}^{\mathbb{Q}}$ the expectation under the risk-neutral measure $\mathbb{Q}$, x and y are state variables at time t and T, respectively, $f(y|x)$ is the probability density of y given x and r is the risk-neutral interest rate.

Whereas f is typically not known, the characteristic function of the log-asset price is often available (sometimes in closed-form), that is, the Fourier transform of f. We represent the payoff as a function of the log-asset price, and denote the log-asset prices by,

$$x = \log(S_t/K) \quad \text{and} \quad y = \log(S_T/K), \tag{3.16}$$

with S_t the underlying price at time t and K the strike price. The payoff $v(y,T)$ for European options in log-asset space is then given by,

$$v(y,T) = \left[\alpha K(e^y - 1)\right]^+, \text{ with } \alpha = \begin{cases} 1, & \text{for a call,} \\ -1, & \text{for a put.} \end{cases} \tag{3.17}$$

From formula (3.17), the strategy to determine the price of the option using wavelet inversion methods consists of replacing the density function f by an approximation based on a wavelet expansion (i.e. the methods we have already presented in Section 2.6) where the coefficients are recovered from the Fourier transform.

In what follows, we introduce how to deal with option pricing in each of the wavelet inversion methods presented. From now on, we assume that the density functions are in $L^2(\mathbb{R})$. Since we can expect that the mass in the tails of the density function $f(y|x)$ rapidly decays to zero, we can apply the methods explained before in Section 2.6.

Now, let us introduce the concept of cumulants which helps us to determine finite domains in these Fourier techniques.

Definition 3.2.9. *Let X be a random variable and $\hat{f}_X$ its corresponding characteristic function. We define the **cumulant generating function** of X by,*

$$g(\omega) := \log \mathbb{E}\left[e^{\omega X}\right] = \log \hat{f}_X(-i\omega), \tag{3.18}$$

*and the **i^{th} cumulant** of X, denoted by κ_i, is given by the $i-th$ derivative at zero of $g(\omega)$,*

$$\kappa_i := g^{(i)}(0). \tag{3.19}$$

When numerically approximating integrals, truncation of integrals with infinite integration domain is required. A proposed way in [28] is to make use on the cumulants of the distribution. Let k_n denote the n-th cumulant of $y = \log(S_T/K)$ and let $L > 0$ be a scaling parameter, then,

$$[a,b] := \left[\kappa_1 - L\sqrt{\kappa_2 + \sqrt{\kappa_4}}, k_1 + L\sqrt{\kappa_2 + \sqrt{\kappa_4}}\right], \tag{3.20}$$

where L is suggested to be chosen in the range $[7.5, 10]$.

3.2.3. ONE-DIMENSIONAL OPTION PRICING WAVELET METHODS

We devote this section to introduce the option pricing methods based on the wavelets presented before in Section 2.5.

HAAR-WA OPTION PRICING

This option pricing method is original from [65]. We follow the steps given by the authors to introduce it. Let us start by truncating the infinity integration range to $[a,b] \subset \mathbb{R}$ without loosing significant accuracy. In order to carry out this truncation the cumulants approach of (3.20) is used. Using the WA expansion (2.91), the density function is approximated by,

$$f_m^{[a,b]}(y|x) = \sum_{k=0}^{2^m-1} c_{m,k}^H(x)\phi_{m,k}^H\left(\frac{y-a}{b-a}\right). \tag{3.21}$$

Note the $[a,b]$ because of the truncation of the interval. Then, it is used in the valuation formula,

$$\begin{aligned} v(x,t) &\approx v^{[a,b]}(x,t) = e^{-r(T-t)}\int_a^b v(y,T)f(y|x)dy, \\ v^{[a,b]}(x,t) &\approx v_m^{[a,b]}(x,t) = e^{-r(T-t)}\int_a^b v(y,T)f_m^{[a,b],}(y|x)dy, \\ v_m^{[a,b]}(x,t) &= v_{m,H}^{[a,b]}(x,t) := e^{-r(T-t)}\sum_{k=0}^{2^m-1} c_{m,k}^H V_{m,k}^\alpha, \end{aligned} \tag{3.22}$$

where,

$$V_{m,k}^\alpha := \int_a^b v(y,T)\phi_{m,k}^H\left(\frac{y-a}{b-a}\right)dy, \tag{3.23}$$

are called the payoff coefficients, and $c_{m,k}^H$ are the coefficients of the density approximation given by (2.97) and thus called density coefficients. Recall that we have $\alpha = 1$ for a call option and $\alpha = -1$ for a put option. The expression of the payoff coefficients is given in the next proposition. The proof can be found in [65].

Proposition 3.2.10. *Let us define* $\Delta_m = \frac{b-a}{2^m}$, $\beta_k = a + k\Delta_m$, $\gamma_k = \beta_k + \Delta_m$, $\delta_k = max(0,\beta_k)$ *and* $\zeta_k = min(0,\gamma_k)$. *Assuming that* $a < 0 < b$, *then the coefficients computed by the WA method are as follows,*

$$V_{m,k}^1 = \begin{cases} 2^{m/2}K\left(e^{\gamma_k - e^{\delta_k} + \delta_k - \gamma_k}\right), & \text{for } \gamma_k > 0, \\ 0, & \text{otherwise.} \end{cases}, \tag{3.24}$$

$$V_{m,k}^{-1} = \begin{cases} 2^{m/2}K\left(e^{\beta_k - e^{\zeta_k} + \zeta_k - \beta_k}\right), & \text{for } \beta_k < 0, \\ 0, & \text{otherwise.} \end{cases}, \tag{3.25}$$

for $k = 0,\dots,2^m-1$. *If* $a < b < 0$ *then,* $V_{m,k}^1 = 0$, $V_{m,k}^{-1} = 2^{m/2}K\left(e^{\beta_k} - e^{\gamma_k} + \gamma_k - \beta_k\right)$, *for* $k = 0,\dots,2^m-1$. *And conversely, if* $0 < a < b$ *then,* $V_{m,k}^{-1} = 0$, $V_{m,k}^1 = 2^{m/2}K\left(e^{\gamma_k} - e^{\beta_k} + \beta_k - \gamma_k\right)$ *for* $k = 0,\dots,2^m-1$.

There is a complete error analysis available in Section 4 of [65].

SWIFT OPTION PRICING

The SWIFT option pricing method is introduced in [67]. It starts by approximating the density function f in terms of a Shannon wavelets expansion, as presented in Section 2.6.2,

$$f(y|x) \approx \sum_{k\in\mathbb{Z}} c_{m,k}^S(x)\phi_{m,k}^S(y) \approx \sum_{k=k_1}^{k_2} c_{m,k}^S(x)\phi_{m,k}^S(y). \tag{3.26}$$

The approximation on the right hand side of (3.26) come from the fact of that the coefficients vanish when $k \to \infty$ if we assume that $\lim_{x\to\pm\infty} f(x) = 0$.

We set $I_m = \left[\frac{k_1}{2^m}, \frac{k_2}{2^m}\right]$, then the option pricing formula becomes,

$$v(x,t) \approx v_{m,S}(x,t) := e^{-r(T-t)}\sum_{k=k_1}^{k_2} c_{m,k}^S V_{m,k}^\alpha, \tag{3.27}$$

where $V_{m,k}$ are the payoff coefficients defined as,

$$V^{\alpha}_{m,k} := \int_{I_m} v(y,T)\phi^{S}_{m,k}(y)\,dy. \tag{3.28}$$

The truncation parameters k_1 and k_2, are defined by means of cumulants, see (3.20). If a and b are the parameters given by the cumulants, then

$$\frac{k_1}{2^m} \le a < b \le \frac{k_2}{2^m}, \tag{3.29}$$

which implies $[a,b] \subset I_m$.

Based on Proposition 2 of [67], we have the following expression for the payoff coefficients. Note that the proof can be found there as well.

Proposition 3.2.11. *Let us define* $\bar{k}_1 = \max(k_1, 0)$, $\bar{k}_2 = \min(k_2, 0)$. *The payoff coefficients for a European call option are computed as follows,*

$$V^{1}_{m,k} \approx V^{1,*}_{m,k} := \begin{cases} \frac{K2^{m/2}}{2^{J-1}} \sum_{j=1}^{2^{J-1}} \left[I_1\left(\frac{\bar{k}_1}{2^m}, \frac{k_2}{2^m}\right) - I_2\left(\frac{\bar{k}_1}{2^m}, \frac{k_2}{2^m}\right)\right], & \text{if } k_2 > 0, \\ 0, & \text{if } k_2 \le 0. \end{cases}, \tag{3.30}$$

and for the put option,

$$V^{-1}_{m,k} \approx V^{-1,*}_{m,k} := \begin{cases} \frac{K2^{m/2}}{2^{J-1}} \sum_{j=1}^{2^{J-1}} \left[I_1\left(\frac{k_1}{2^m}, \frac{\bar{k}_2}{2^m}\right) - I_2\left(\frac{k_1}{2^m}, \frac{\bar{k}_2}{2^m}\right)\right], & \text{if } k_1 < 0, \\ 0, & \text{if } k_1 \ge 0. \end{cases}, \tag{3.31}$$

where,

$$\begin{aligned} I_1(a,b) =& \frac{C_j 2^m}{1+(C_j 2^m)^2}\left[e^b \sin\left(C_j(2^m b - k)\right) - e^a \sin\left(C_j(2^m a - k)\right)\right. \\ &\left. + \frac{1}{C_j 2^m}\left(e^b \cos\left(C_j(2^m b - k)\right) - e^a \cos\left(C_j(2^m a - k)\right)\right)\right], \\ I_2(a,b) =& \frac{1}{2^m}\left(\sin\left(C_j(2^m b - k)\right) - \sin\left(C_j(2^m a - k)\right)\right), \end{aligned} \tag{3.32}$$

and $C_j = \frac{2j-1}{2^J}\pi$.

Note that by considering a constant J in (3.30) and (3.31), it is possible to apply the FFT to speed up the computation of payoff coefficients.

CHEBYSHEV-WA OPTION PRICING

For option pricing using Chebyshev wavelets we consider first kind Chebyshev wavelets presented in Section 2.5.3. The option pricing formula is,

$$v(x,t) = e^{-r\Delta t}\int_{\mathbb{R}} v(y,T) f(y \mid x)\,dy, \tag{3.33}$$

where $v(y,T)$ is the payoff function of the option. After truncating the integral and approximating the density by Chebyshev wavelets, we end up with,

$$v(x,t) \approx v^{[a,b]}_{m,C}(x,t) := e^{-r\Delta t}\sum_{n=1}^{2^{k-1}}\sum_{k=0}^{M-1} c^{C}_{m,n,k}\int_a^b v(y,T)\psi^{C}_{m,n,k}(y \mid x)\,dy. \tag{3.34}$$

We define the payoff coefficients as,

$$V_{m,n,k}(x) := \int_a^b \phi^C_{m,n,k}(y \mid x)\, v(y,T)\, dy. \tag{3.35}$$

Thus, we can rewrite the pricing formula as,

$$v^{[a,b]}_{m,C}(x,t) = e^{-r\Delta t} \sum_{n=1}^{2^{k-1}} \sum_{k=0}^{M-1} c^C_{m,n,k} V_{m,n,k}(x). \tag{3.36}$$

Let us show how to compute the payoff coefficients (3.35) for a put option (the call option would be analogous). Recall that for a put option, the payoff is $v(y,T) = [K(1-e^y)]^+$. Thus,

$$\begin{aligned} V_{m,n,k}(x) &= K \int_a^b \left(1-e^y\right)^+ \psi^C_{m,n,k}(y \mid x)\, dy \\ &= K \int_{[a,b]\cap(-\infty,0]} (1-e^y) \psi^C_{m,n,k}(y \mid x)\, dy. \end{aligned} \tag{3.37}$$

From this last expression, we consider three situations: $a<0<b$, $a<b<0$ and $0<a<b$, let I represent any of them. In each case, we need to compute the following two integrals,

$$\int_I \psi^C_{m,n,k}(y)\, dy, \qquad \int_I e^y \psi^C_{m,n,k}(y)\, dy. \tag{3.38}$$

Using the properties detailed in Remark 2.5.2, it is straightforward to compute the first integral. The difficulty comes from computing the second integral. After some algebraic manipulations we arrive to the following integral,

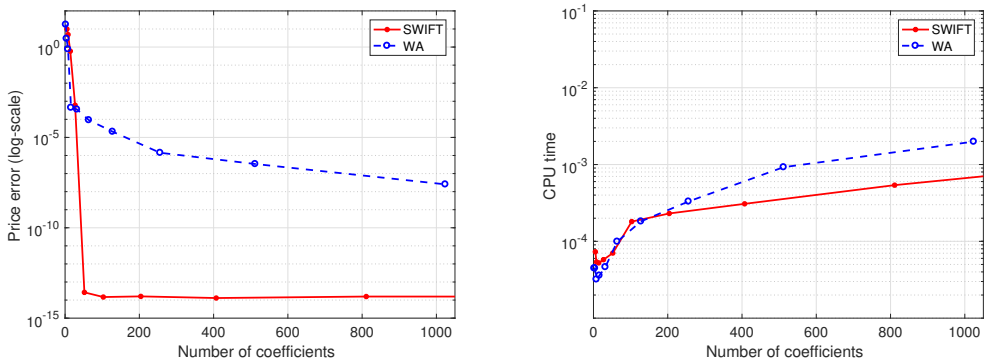
$$\int_{I'} e^{\frac{x}{2^k}} T_k(x)\, dx, \tag{3.39}$$

where I' is an arithmetic modification of the previous domain I. We tried to solve this integral using two different approaches: the trigonometric expression and the recursive formula. However, we did not arrive to any efficient form to compute the integral.

Summarizing, Chebysev wavelets in the context of option pricing present lots of difficulties in terms of computational efficiency. We could use numerical integration to approximate the integral but this would mean more errors and a significant increase in CPU time. Recall that the method to compute the density coefficients in (2.118) is computationally involved. That is why we do not consider Chebyshev wavelet method for option pricing any more since it cannot compete against WA or SWIFT.

NUMERICAL EXAMPLES

We briefly show the performance of WA and SWIFT in option pricing. We can observe that SWIFT achieves more accurate results even using fewer terms in the approximation. For this reason, we will extend it to the multidimensional case because it is the one that performs better in the option pricing problem, see more about its performance in [67].



(a) Errors. (b) CPU time.

Figure 3.2: Comparing WA and SWIFT for pricing European Call option with $S = 100$, $K = 120$, $r = 0.1$, $\sigma = 0.3$ and $T = 0.1$.

3.3. TWO-DIMENSIONAL SWIFT FOR OPTION PRICING

We extend the SWIFT method to higher dimensions and use it to price exotic option contracts. The multidimensional extension inherits the properties of the one-dimensional method, the exponential convergence being one of them.

In order to present the method, we first need to extend wavelet theory to the multidimensional case. It is presented for two dimensions, and it can be extended in the same way to any dimensions. However, we should keep in mind that due to the curse of dimensionality of this method (or any other method of this nature), wavelets can only approximate functions up to a reasonable number of dimensions.

The state-of-the-art method from the same nature is the 2D-COS method [74]. We use it as a benchmark, so let us present a brief introduction to the method. Note that bold letters denote vectors.

Remark 3.3.1. *The COS method is based on Fourier cosine series expansions. The COS method in one dimension [28] reads,*

$$v(x,t) = e^{-r\Delta t}\mathbb{E}^{\mathbb{Q}}\left[v(X_T,T)\right] \approx \frac{b-a}{2}e^{-r\Delta t}{\sum_{k=0}^{N-1}}' V_k(T)F_k(x), \tag{3.40}$$

where,

$$\begin{aligned} F_k(x) &= \frac{2}{b-a}\Re\left(\hat{f}\left(\frac{k\pi}{b-a}\right)e^{ik\pi\frac{x-a}{b-a}}\right), \\ V_k(T) &= \frac{2}{b-a}\int_a^b v(y,T)\cos\left(k\pi\frac{y-a}{b-a}\right)dy, \end{aligned} \tag{3.41}$$

and the apostrophe in the sum means a slightly modification for the $k=0$ coefficient expression. The parameters $[a,b]$ are the truncated domain determined by the cumulants. And N is the number of terms, without a rule on how to choose it.

The 2D-COS method [74] reads,

$$v(\boldsymbol{x},t)=e^{-r\Delta t}\mathbb{E}^{\mathbb{Q}}\left[v(\boldsymbol{X}_T,T)\right]\approx\frac{b_1-a_1}{2}\frac{b_2-a_2}{2}e^{-r\Delta t}\sum_{k_1=0}^{N_1-1}{}'\sum_{k_2=0}^{N_2-1}{}'V_{k_1,k_2}(T)F_{k_1,k_2}(\boldsymbol{x}),\tag{3.42}$$

where,

$$\begin{aligned}
F_{k_1,k_2}(\boldsymbol{x})&=\frac{F^{+}_{k_1,k_2}(\boldsymbol{x})+F^{-}_{k_1,k_2}(\boldsymbol{x})}{2},\\
F^{\pm}_{k_1,k_2}(\boldsymbol{x})&=\frac{2}{b_1-a_1}\frac{2}{b_2-a_2}\Re\left(\hat{f}\left(\frac{k_1\pi}{b_1-a_1,\pm\frac{k_2\pi}{b_2-a_2}}|\boldsymbol{x}\right)\exp\left(-ik_1\pi\frac{a_1}{b_1-a_1}\mp ik_2\pi\frac{a_2}{b_2-a_2}\right)\right),\\
V_{k_1,k_2}(T)&\approx\sum_{n_1=0}^{Q-1}\sum_{n_2=0}^{Q-1}g(y_1^{n_1},y_2^{n_2})\cos\left(k_1\pi\frac{2n_1+1}{2Q}\right)\cos\left(k_2\pi\frac{2n_2+1}{2Q}\frac{b_1-a_1}{Q}\frac{b_2-a_2}{Q}\right).
\end{aligned}\tag{3.43}$$

The parameters $[a_1,b_1]$, $[a_2,b_2]$ are the truncated domains in each dimension determined by the cumulants. N_1, N_2 are the same as N for each dimension, and Q is a parameter that appears from a numerical integration.

From now on, in order to simplify notation, we omit the measure $\mathbb{Q}$ in the expectation symbol $\mathbb{E}^{\mathbb{Q}}$, so we just use $\mathbb{E}$.

3.3.1. MULTIDIMENSIONAL WAVELETS FRAMEWORK

We introduce here some useful definitions for the two-dimensional wavelet framework which can be easily extended to more dimensions.

We work now in $L^2(\mathbb{R}^2)$, which is the vector space of measurable, square-integrable two-dimensional functions. For any two functions $h(x,y)$, $p(x,y)\in L^2(\mathbb{R}^2)$, their inner product in $L^2\left(\mathbb{R}^2\right)$ is defined by,

$$\left\langle h\left(x,y\right),p\left(x,y\right)\right\rangle:=\iint_{\mathbb{R}^2}h\left(x,y\right)\bar{p}\left(x,y\right)dxdy,\tag{3.44}$$

where $\bar{p}$ is the complex conjugate of p. The bivariate Fourier transform of a function $h(x,y)\in L^2(\mathbb{R}^2)$ is defined by,

$$\hat{h}(u_1,u_2):=\iint_{\mathbb{R}^2}h(x,y)e^{-i\left(u_1x+u_2y\right)}\,dxdy.\tag{3.45}$$

We introduce now the wavelet theory in two dimensions. Note that we base the discussion on the multiresolution analysis because as we have already mentioned, it is how Shannon wavelets are structured. Multiresolution analysis which defines general wavelet structures in L^2 spaces, can be generalized to any dimension $d\in\mathbb{N}$ (see [59] for details). We illustrate the two-dimensional case.

A multiresolution approximation of $L^2(\mathbb{R}^2)$ is a sequence of subspaces of $L^2(\mathbb{R}^2)$ which satisfies a straightforward two-dimensional extension of the properties of MRA presented in Definition 2.4.3. Let $\left(V_m^2\right)_{m\in\mathbb{Z}}$ be such a multiresolution approximation of $L^2(\mathbb{R}^2)$. One can show that there exists a unique two-dimensional scaling function $\Phi(x,y)$ whose dilation and translation give an orthonormal basis of each space V_m^2 (see [25, 50]).

We stay in the particular case of separable wavelet bases and separable multiresolutions, due to the direct connection to the one-dimensional case, and because it avoids mixing of information at two different scales. For such multiresolution approximations, each vector space V_m^2 is decomposed as a tensor product of two identical subspaces of $L^2(\mathbb{R})$ whose elements are products of functions dilated at the same scale,

$$V_m^2 = V_m \otimes V_m. \tag{3.46}$$

The sequence of vector spaces $\left(V_m^2\right)_{m\in\mathbb{Z}}$ forms a multiresolution approximation of $L^2(\mathbb{R}^2)$ if and only if $\{V_m, m \in \mathbb{Z}\}$ is a multiresolution analysis of $L^2(\mathbb{R})$. It is easy to see that the two-dimensional scaling function $\Phi(x,y)$ can be written as,

$$\Phi(x,y) := \phi(x)\phi(y), \tag{3.47}$$

where $\phi(x)$ is the one-dimensional scaling function of V_m.

As stated, $\{\phi_{m,k}; k \in \mathbb{Z}\}$ is an orthonormal basis of V_m. Since $V_m^2 = V_m \otimes V_m$, for $\boldsymbol{x} = (x_1, x_2)$ and $\boldsymbol{k} = (k_1, k_2)$, we have that,

$$\left\{\Phi_{m,\boldsymbol{k}}(\boldsymbol{x}) = \phi_{m,k_1}(x_1)\phi_{m,k_2}(x_2)\right\}_{\boldsymbol{k}\in\mathbb{Z}^2} \tag{3.48}$$

is an orthonormal basis of V_m^2.

Let W_m^2 be the detail space equal to the orthogonal complement of the lower resolution approximation space V_m^2 in V_{m+1}^2,

$$V_{m+1}^2 = V_m^2 \oplus W_m^2. \tag{3.49}$$

The following theorem builds a wavelet basis of each detail space W_m^2, to construct later a wavelet orthonormal basis of $L^2(\mathbb{R}^2)$. A separable orthonormal wavelet basis of $L^2(\mathbb{R}^2)$ is constructed with separable products of a scaling function ϕ and a wavelet ψ.

Theorem 3.3.2 (Theorem 7.24 of [50])**.** *Let ϕ be a scaling function and ψ be the corresponding wavelet generating a orthonormal wavelet basis of $L^2(\mathbb{R})$. We define the three wavelets,*

$$\Psi^1(\boldsymbol{x}) = \phi(x_1)\psi(x_2),\ \Psi^2(\boldsymbol{x}) = \psi(x_1)\phi(x_2),\ \Psi^3(\boldsymbol{x}) = \psi(x_1)\psi(x_2), \tag{3.50}$$

and denote, for $1 \le n \le 3$,

$$\Psi_{m,k}^n(\boldsymbol{x}) = 2^m\Psi^n\left(2^m x_1 - k_1, 2^m x_2 - k_2\right). \tag{3.51}$$

Then, the wavelet family,

$$\left\{\Psi_{m,\boldsymbol{k}}^1, \Psi_{m,\boldsymbol{k}}^2, \Psi_{m,\boldsymbol{k}}^3\right\}_{\boldsymbol{k}\in\mathbb{Z}^2} \tag{3.52}$$

forms an orthonormal basis of W_m^2. And,

$$\left\{\Psi_{m,\boldsymbol{k}}^1, \Psi_{m,\boldsymbol{k}}^2, \Psi_{m,\boldsymbol{k}}^3\right\}_{(m,\boldsymbol{k})\in\mathbb{Z}\times\mathbb{Z}^2} \tag{3.53}$$

forms an orthonormal basis of $L^2(\mathbb{R}^2)$.

For any function $f(x_1,x_2)\in L^2(\mathbb{R}^2)$, a projection map $\mathscr{P}_m : L^2(\mathbb{R}^2)\to V_m^2$ is defined as,

$$\begin{aligned}\mathscr{P}_m f(x_1,x_2) &:= \sum_{j=-\infty}^{m-1}\sum_{n=1,2,3}\sum_{k_1\in\mathbb{Z}}\sum_{k_2\in\mathbb{Z}} d^n_{j,k_1,k_2}\Psi^n_{j,k_1,k_2}(x_1,x_2)\\ &= \sum_{k_1\in\mathbb{Z}}\sum_{k_2\in\mathbb{Z}} c_{m,k_1,k_2}\Phi_{m,k_1,k_2}(x_1,x_2)\\ &= \sum_{k_1\in\mathbb{Z}}\sum_{k_2\in\mathbb{Z}} c_{m,k_1,k_2}\phi_{m,k_1}(x_1)\phi_{m,k_2}(x_2),\end{aligned} \tag{3.54}$$

where,

$$c_{m,k_1,k_2} := \iint_{\mathbb{R}^2} f(x_1,x_2)\Phi_{m,k_1,k_2}(x_1,x_2)\,dx_1\,dx_2 \tag{3.55}$$

are the scaling coefficients, and,

$$d^n_{j,k_1,k_2} := \iint_{\mathbb{R}^2} f(x_1,x_2)\Psi^n_{m,k_1,k_2}(x_1,x_2)\,dx_1\,dx_2 \tag{3.56}$$

are the wavelets coefficients. The convergence is in the $L^2(\mathbb{R}^2)$ norm, this is,

$$\|f-\mathscr{P}_m f\|_2 \underset{m\to+\infty}{\longrightarrow} 0. \tag{3.57}$$

We consider the Shannon father function rather than the mother wavelet, due to its tractability and simplicity. Hence, our wavelet bases are a set of Shannon scaling functions in the subspace V_m, given by, for $x\in\mathbb{R}$,

$$\phi_{m,k}(x) = 2^{m/2}\mathrm{sinc}\left(2^m x-k\right). \tag{3.58}$$

The two-dimensional Shannon scaling function is given by, see Figure 3.3,

$$\Phi_{m,k_1,k_2}(x_1,x_2) := \phi_{m,k_1}(x_1)\phi_{m,k_2}(x_2) = 2^m\mathrm{sinc}(2^m x_1-k_1)\mathrm{sinc}(2^m x_2-k_2). \tag{3.59}$$

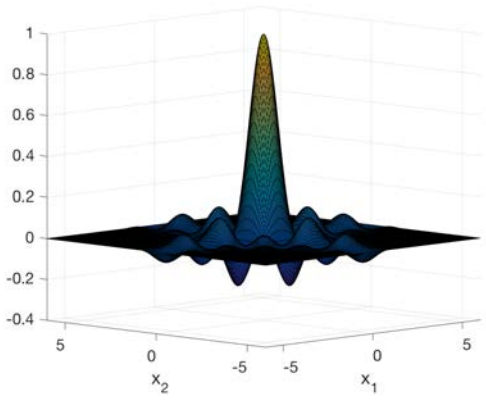


Figure 3.3: Two-dimensional Shannon wavelet $\Phi_{0,0,0}(x_1,x_2)$, with $(x_1,x_2)\in[-6,6]^2$.

Wavelets bases in higher dimensions

Separable orthonormal wavelet bases of $L^2(\mathbb{R}^p)$ can be constructed for any $p \geq 2$ with a procedure similar to the two-dimensional extension. Let ϕ be a scaling function and ψ a wavelet that yields an orthogonal basis of $L^2(\mathbb{R})$. Now, we have that $V_m^p = V_m \otimes \cdots \otimes V_m$ and W_m^p is the detail space, i.e. the orthogonal complement of the lower resolution approximation space.

We denote $\theta^0 := \phi$ and $\theta^1 := \psi$. For an integer $0 \leq \epsilon \leq 2^p$ written in binary form, $\epsilon = \epsilon_1 \cdots \epsilon_p$, we associate p-dimensional functions defined in $\boldsymbol{x} = (x_1, \cdots, x_p)$ by,

$$\psi^{\epsilon}(\boldsymbol{x}) := \theta^{\epsilon_1}(x_1) \cdots \theta^{\epsilon_n}(x_p), \tag{3.60}$$

For $\epsilon = 0$, we obtain a $p-$dimensional scaling function,

$$\psi^0(\boldsymbol{x}) = \phi(x_1) \cdots \phi(x_p). \tag{3.61}$$

Non-zero indexes ϵ correspond to $2^p - 1$ wavelets. At any scale 2^m and for $\boldsymbol{n} = (n_1, \cdots, n_p)$ we denote,

$$\psi_{m,\boldsymbol{n}}^{\epsilon}(\boldsymbol{x}) := 2^{pm/2} \psi^{\epsilon}(2^m x_1 - n_1, \cdots, 2^m x_p - n_p). \tag{3.62}$$

Theorem 3.3.3 (Theorem 7.25 in [50]). *The family obtained by dilating and translating the $2^p - 1$ wavelets for $\epsilon \neq 0$,*

$$\{\psi_{m,\boldsymbol{n}}^{\epsilon}\}_{1 \leq \epsilon < 2^p, \boldsymbol{n} \in \mathbb{Z}^p} \tag{3.63}$$

is an orthonormal basis of W_m^p. And the family,

$$\{\psi_{m,\boldsymbol{n}}^{\epsilon}\}_{1 \leq \epsilon < 2^p, (m,\boldsymbol{n}) \in \mathbb{Z}^{p+1}} \tag{3.64}$$

is an orthonormal basis of $L^2(\mathbb{R}^p)$.

For any function $f \in L^2(\mathbb{R}^p)$, a projection map of $L^2(\mathbb{R}^p)$ onto V_m^p, $\mathscr{P}_m : L^2(\mathbb{R}^p) \to V_m^p$, is defined as,

$$\begin{aligned} \mathscr{P}_m f(\boldsymbol{x}) :&= \sum_{j=-\infty}^{m-1} \sum_{0 < \epsilon \leq 2^p} \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} \cdots \sum_{k_p \in \mathbb{Z}} d_{j,k_1,\cdots,k_p}^{\epsilon} \psi_{j,k_1,\cdots,k_p}^{\epsilon}(\boldsymbol{x}) \\ &= \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} \cdots \sum_{k_p \in \mathbb{Z}} c_{m,k_1,\cdots,k_p} \phi_{m,k_1}(x_1) \phi_{m,k_2}(x_2) \cdots \phi_{m,k_p}(x_p), \end{aligned} \tag{3.65}$$

where $c_{m,k_1,\cdots,k_p}$ are the scaling coefficients and d_{j,k_1,k_2}^n the wavelets coefficients.

The p-dimensional Shannon scaling function is given by,

$$\Phi_{m,k_1,\cdots,k_p}(\boldsymbol{x}) := \phi_{m,k_1}(x_1) \cdots \phi_{m,k_p}(x_p) = 2^{\frac{pm}{2}} \operatorname{sinc}(2^m x_1 - k_1) \cdots \operatorname{sinc}(2^m x_p - k_p). \tag{3.66}$$

Further details about wavelets in high dimensions can be found in Section 7.7.4 of [50].

3.3.2. EUROPEAN RAINBOW OPTION PRICING

Motivation

We start this section by defining the two-dimensional pricing formula as a discounted expectation of the option value at expiration.

Similarly as in dimension one, assume $(\Omega, \mathscr{F}, P)$ is a probability space, $T > 0$ is the finite terminal time, and $\mathbb{F} = (\mathscr{F}_S)_{0 \leq S \leq T}$ is a filtration with the usual conditions. Then, the process

$\mathbf{X}_t = (X_t^1, X_t^2)$ denotes a two-dimensional stochastic process on the filtered probability space $(\Omega, \mathscr{F}, \mathbb{F}, P)$, representing the log-asset prices of the underlying asset. We assume that the bivariate characteristic function of the stochastic process is known.

The value of a European rainbow option with payoff function $g(\cdot)$, which depends on the underlying asset price, is given by the risk-neutral option valuation formula,

$$v(\mathbf{x}, t) = e^{-r\Delta t}\mathbb{E}\left[g(\mathbf{X}_T)\right] = e^{-r\Delta t}\iint_{\mathbb{R}^2} g(\mathbf{y}) f(\mathbf{y}|\mathbf{x})\, d\mathbf{y}, \tag{3.67}$$

where $\mathbf{x} = (x_1, x_2)$ is the current state, $f(y_1, y_2|x_1, x_2)$ is the underlying conditional density function, r is the risk-free rate and $\Delta t := T - t$ denotes time to expiration.

The strategy followed to determine the price of the multidimensional option follows the same idea as for the one-dimensional options. It consists of approximating the density function f in (3.67) by a finite combination of Shannon scaling functions and recovering the coefficients of the approximation from its Fourier transform.

Option pricing

We present the 2D-SWIFT formula to recover the density function f in expression (3.67) assuming that its Fourier transform $\hat{f}$ is known. The density function of the asset price process at terminal time T is usually not known, but often its characteristic function is known. The method is based on the approximation of the density function by a finite combination of Shannon wavelets in dimension two. Once the density has been recovered, we replace f by its approximation in formula (3.67) to get the final price of the financial contract. This two-step procedure is explained in detail first, while the general case for higher dimensions is briefly exposed later.

DERIVATION OF THE 2D-SWIFT METHOD

As mentioned before, the first step during the derivation of the 2D-SWIFT pricing formula is the recovery of the density function. We distinguish three main steps,

Step 1. We approximate the conditional density function by a finite combination of Shannon scaling functions,

$$f(\mathbf{y}|\mathbf{x}) \approx f_1(\mathbf{y}|\mathbf{x}) := \mathscr{P}_m f(\mathbf{y}|\mathbf{x}) = \sum_{k_1\in\mathbb{Z}}\sum_{k_2\in\mathbb{Z}} D_{m,k_1,k_2}(\mathbf{x})\Phi_{m,k_1,k_2}(\mathbf{y}), \tag{3.68}$$

where D_{m,k_1,k_2}, the density coefficients, are defined by,

$$D_{m,k_1,k_2}(\mathbf{x}) := \iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})\Phi_{m,k_1,k_2}(\mathbf{y})\, d\mathbf{y}. \tag{3.69}$$

Step 2. Next, we truncate the summation range such that $k_1 \in \{l_1, \ldots, u_1\}$ and $k_2 \in \{l_2, \ldots, u_2\}$, and thus, the density approximation becomes,

$$f_1(\mathbf{y}|\mathbf{x}) \approx f_2(\mathbf{y}|\mathbf{x}) := \sum_{k_1=l_1}^{u_1}\sum_{k_2=l_2}^{u_2} D_{m,k_1,k_2}(\mathbf{x})\Phi_{m,k_1,k_2}(\mathbf{y}). \tag{3.70}$$

Step 3. Using the complex exponential formula (3.74), we approximate the two-dimensional Shannon scaling function as,

$$\Phi_{m,k_1,k_2}(\mathbf{y}) \approx \Phi^*_{m,k_1,k_2}(\mathbf{y}) := \frac{2^m}{N^2}\sum_{j_1=1}^{N} e^{i\Omega_{j_1}(2^m y_1 - k_1)}\sum_{j_2=1}^{N} e^{i\Omega_{j_2}(2^m y_2 - k_2)}, \tag{3.71}$$

where $\Omega_j = -\pi + \frac{2j-1}{N}\pi$. For convenience, we use in both dimensions the same discretization for the sinc approximation, this is, N points.

Then, from (3.69) and (3.71), we approximate the density coefficients $D_{m,k_1,k_2}(\mathbf{x})$ by,

$$\begin{aligned} D^*_{m,k_1,k_2}(\mathbf{x}) :&= \iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})\Phi^*_{m,k_1,k_2}(\mathbf{y})\,d\mathbf{y} \\ &= \frac{2^m}{N^2}\sum_{j_1=1}^{N}\sum_{j_2=1}^{N} e^{-i\sum_{n=1,2}\Omega_{j_n}k_n}\iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})e^{i\sum_{n=1,2}\Omega_{j_n}2^m y_n}\,dy_1\,dy_2 \\ &= \frac{2^m}{N^2}\sum_{j_1=1}^{N}\sum_{j_2=1}^{N} e^{-i\sum_{n=1,2}\Omega_{j_n}k_n}\hat{f}\left(-\Omega_{j_1}2^m, -\Omega_{j_2}2^m|\mathbf{x}\right). \end{aligned} \tag{3.72}$$

Finally, from (3.70) and (3.72) we obtain the density approximation,

$$f_2(\mathbf{y}|\mathbf{x}) \approx f^*(\mathbf{y}|\mathbf{x}) := \sum_{k_1=l_1}^{u_1}\sum_{k_2=l_2}^{u_2} D^*_{m,k_1,k_2}(\mathbf{x})\Phi_{m,k_1,k_2}(\mathbf{y}). \tag{3.73}$$

Approximation of the sinc function

Central to the option pricing process, is a convenient approximation of the sinc function by means of a midpoint quadrature rule. For $t, x \in \mathbb{R}$,

$$\operatorname{sinc}(t) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\operatorname{sinc}}(x)e^{itx}\,dx = \frac{1}{2\pi}\int_{\mathbb{R}} \operatorname{rect}\left(\frac{x}{2\pi}\right)e^{itx}\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{itx}\,dx \approx \operatorname{sinc}^*_N(t) := \frac{1}{N}\sum_{j=1}^{N} e^{i\Omega_j t}, \tag{3.74}$$

where $\widehat{\operatorname{sinc}}$ represents the Fourier transform of sinc, N is the number of points in the midpoint quadrature and $\Omega_j := -\pi + \frac{2j-1}{N}\pi$.

Recall that the approximation presented in the one-dimensional case in (2.105) reads,

$$\operatorname{sinc}(t) \approx \operatorname{sinc}^*(t) := \frac{1}{2^{J-1}}\sum_{j=1}^{2^{J-1}} \cos\left(\frac{2j-1}{2^J}\pi t\right) = \frac{1}{2^{J-1}}\sum_{j=1}^{2^{J-1}} \cos\left(\omega_j t\right), \tag{3.75}$$

for $\omega_j = \frac{2j-1}{2^J}\pi$. It can be easily shown that the approximation (2.105) is a particular case of the approximation (3.74) for $N = 2^J$, indeed,

$$\begin{aligned} \operatorname{sinc}^*(t) &= \frac{2}{N}\sum_{j=1}^{N/2}\cos\left(\frac{2j-1}{N}\pi t\right) = \frac{1}{N}\sum_{j=1}^{N/2}\left(e^{i\left(\frac{2j-1}{N}\right)\pi t} + e^{-i\left(\frac{2j-1}{N}\right)\pi t}\right) = \frac{1}{N}\sum_{j=1}^{N/2} e^{i\left(\frac{2j-1}{N}\right)\pi t} \\ &+ \frac{1}{N}\sum_{j=1-N/2}^{0} e^{i\left(\frac{2j-1}{N}\right)\pi t} = \frac{1}{N}\sum_{j=1-N/2}^{N/2} e^{i\left(\frac{2j-1}{N}\right)\pi t} = \frac{1}{N}\sum_{j=1}^{N} e^{i\left(\frac{2j-N-1}{N}\right)\pi t} = \operatorname{sinc}^*_N(t). \end{aligned} \tag{3.76}$$

The complex exponential form (3.74) will facilitate the derivation of the density and payoff coefficients along the option pricing procedure presented in the next section.

The 2D-SWIFT formula for the approximation of $v(\mathbf{x}, t)$ in (3.67) is obtained first by truncating the integration range,

$$v(\mathbf{x}, t) \approx v_1(\mathbf{x}, t) := e^{-r\Delta t}\int_{a_1}^{b_1}\int_{a_2}^{b_2} g(\mathbf{y})f(\mathbf{y}|\mathbf{x})\,d\mathbf{y}, \tag{3.77}$$

for some values a_1, a_2, b_1, b_2, and then replacing the density $f(\mathbf{y}|\mathbf{x})$ by the approximation $f^*(\mathbf{y}|\mathbf{x})$ from (3.73). Finally, the 2D-SWIFT pricing formula is given by,

$$\begin{aligned} v_1(\mathbf{x},t) \approx v^*(\mathbf{x},t) &:= e^{-r\Delta t}\int_{a_1}^{b_1}\int_{a_2}^{b_2} g(\mathbf{y}) f^*(\mathbf{y}|\mathbf{x})\, d\mathbf{y} \\ &= e^{-r\Delta t}\sum_{k_1=l_1}^{u_1}\sum_{k_2=l_2}^{u_2} D^*_{m,k_1,k_2}(\mathbf{x})\int_{a_1}^{b_1}\int_{a_2}^{b_2} g(\mathbf{y})\Phi_{m,k_1,k_2}(\mathbf{y})\, d\mathbf{y} \\ &= e^{-r\Delta t}\sum_{k_1=l_1}^{u_1}\sum_{k_2=l_2}^{u_2} D^*_{m,k_1,k_2}(\mathbf{x})\, G_{m,k_1,k_2}, \end{aligned} \tag{3.78}$$

where we define the payoff coefficients as,

$$G_{m,k_1,k_2} := \int_{a_1}^{b_1}\int_{a_2}^{b_2} g(\mathbf{y})\Phi_{m,k_1,k_2}(\mathbf{y})\, d\mathbf{y}. \tag{3.79}$$

Remark 3.3.4. *An alternative method to get the density coefficients would be a straightforward extension of the one-dimensional method [67], which basically considers the approximation in (2.105) for the* sinc *function. We approximate the two-dimensional Shannon scaling function as,*

$$\Phi_{m,k_1,k_2}(\mathbf{y}) \approx \frac{2^m}{2^{2(J-1)}}\sum_{j_1=1}^{2^{J-1}}\sum_{j_2=1}^{2^{J-1}} \cos(\omega_{j_1}(2^m y_1 - k_1))\cos(\omega_{j_2}(2^m y_2 - k_2)), \tag{3.80}$$

where $\omega_j := \frac{2j-1}{2^J}\pi$. *Hence, the density coefficients can also be approximated by,*

$$D^{**}_{m,k_1,k_2}(\mathbf{x}) = \frac{2^m}{2^{2(J-1)}}\sum_{j_1=1}^{2^{J-1}}\sum_{j_2=1}^{2^{J-1}}\iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})\prod_{n=1,2}\cos(\omega_{j_n}(2^m y_n - k_n))\, d\mathbf{y}. \tag{3.81}$$

Using the goniometric relation $2\cos(\alpha)\cos(\beta) = \cos(\alpha+\beta)+\cos(\alpha-\beta)$, *we obtain from (3.81),*

$$\begin{aligned} D^{**}_{m,k_1,k_2}(\mathbf{x}) = \frac{2^{m+1}}{2^{2(J-1)}}\sum_{j_1=1}^{2^{J-1}}\sum_{j_2=1}^{2^{J-1}}\Bigg[&\iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})\cos\Big(\sum_{n=1,2}\omega_{j_n}(2^m y_n - k_n)\Big)\, d\mathbf{y} \\ &+ \iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})\cos\Big(\sum_{n=1,2}(-1)^{n+1}\omega_{j_n}(2^m y_n - k_n)\Big)\, d\mathbf{y}\Bigg]. \end{aligned} \tag{3.82}$$

Observe that $-\omega_{j_2} = \omega_{1-j_2}$. *Then,*

$$D^{**}_{m,k_1,k_2}(\mathbf{x}) = \frac{2^{m+1}}{2^{2(J-1)}}\sum_{j_1=1}^{2^{J-1}}\sum_{j_2=1-2^{J-1}}^{2^{J-1}}\iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})\cos\Big(\sum_{n=1,2}\omega_{j_n}(2^m y_n - k_n)\Big)\, d\mathbf{y}. \tag{3.83}$$

Since $\Re e^{-i\alpha} = \cos(\alpha)$, *we find,*

$$\begin{aligned} D^{**}_{m,k_1,k_2}(\mathbf{x}) &= \frac{2^{m+1}}{2^{2(J-1)}}\sum_{j_1=1}^{2^{J-1}}\sum_{j_2=1-2^{J-1}}^{2^{J-1}}\iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})e^{(-i\sum_{n=1,2}\omega_{j_n}(2^m y_n - k_n))}\, d\mathbf{y} \\ &= \frac{2^{m+1}}{2^{2(J-1)}}\sum_{j_1=1}^{2^{J-1}}\sum_{j_2=1-2^{J-1}}^{2^{J-1}} e^{(i\sum_{n=1,2}\omega_{j_n}k_n)}\iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})e^{(-i\sum_{n=1,2}\omega_{j_n}2^m y_n)}\, d\mathbf{y} \\ &= \frac{2^{m+1}}{2^{2(J-1)}}\sum_{j_1=1}^{2^{J-1}}\sum_{j_2=1-2^{J-1}}^{2^{J-1}} \hat{f}(\omega_{j_1}2^m, \omega_{j_2}2^m|\mathbf{x})e^{(i\sum_{n=1,2}\omega_{j_n}k_n)}, \end{aligned} \tag{3.84}$$

where $\hat{f}$ is the Fourier transform of f. For $N = 2^J$, we have $D^*_{m,k_1,k_2}(\mathbf{x}) = D^{**}_{m,k_1,k_2}(\mathbf{x})$. We note that Vieta's formula is more tedious to use and we will therefore consider the approximation (3.74) to compute the payoff coefficients in Section 3.3.4.

Remark 3.3.5. *In our search for an efficient method, formula (3.72) can be conveniently rearranged to get a new expression where an FFT algorithm can be applied. Using the fact that,*

$$-i\Omega_j k = \frac{-i2\pi k}{N}(j-1) - i\pi k\left(\frac{1}{N}-1\right), \tag{3.85}$$

we get,

$$D^*_{m,k_1,k_2}(\mathbf{x}) = \frac{2^m}{N^2} e^{-i\pi\left(\frac{1}{N}-1\right)(k_1+k_2)} \sum_{j_1=1}^{N}\sum_{j_2=1}^{N} \hat{f}\left(-\Omega_{j_1}2^m, -\Omega_{j_2}2^m \middle| \mathbf{x}\right) e^{\frac{-i2\pi k_2}{N}(j_2-1)} e^{\frac{-i2\pi k_1}{N}(j_1-1)}. \tag{3.86}$$

We apply N times an FFT for each sum in (3.86) and the computational complexity to perform this task is $\mathcal{O}(N^2\log_2(N))$.

GENERAL MULTIDIMENSIONAL SWIFT FORMULA

The 2D-SWIFT formula can be easily generalized to higher dimensions. It is clear, however, that if we choose the dimension d to be very large, the curse of dimensionality issue will occur and numerical techniques turn out to be useless in practical terms.

For $x \in \mathbb{R}^d$, the d-dimensional formula reads,

$$v^*(\mathbf{x},t) = e^{-r\Delta t} \sum_{k_1=l_1}^{u_1}\sum_{k_2=l_2}^{u_2}\cdots\sum_{k_d=l_d}^{u_d} D^*_{m,k_1,k_2,\cdots,k_d}(\mathbf{x}) G_{m,k_1,k_2,\cdots,k_d}, \tag{3.87}$$

where,

$$\begin{aligned} D^*_{m,k_1,k_2,\cdots,k_d}(\mathbf{x}) = &\frac{2^{\frac{dm}{2}}}{N^d} e^{-i\pi\left(\frac{1}{N}-1\right)(k_1+k_2+\cdots+k_d)} \sum_{j_1=l_1}^{u_1}\sum_{j_2=l_2}^{u_2}\cdots\sum_{j_d=l_d}^{u_d} \hat{f}(-\Omega_{j_1}2^m, -\Omega_{j_2}2^m, \cdots, -\Omega_{j_d}2^m) \\ &\cdot e^{\frac{-i2\pi k_d}{N}(j_d-1)}\cdots e^{\frac{-i2\pi k_2}{N}(j_2-1)} e^{\frac{-i2\pi k_1}{N}(j_1-1)}, \end{aligned} \tag{3.88}$$

and,

$$G_{m,k_1,k_2,\cdots,k_d} = \int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_d}^{b_d} g(\mathbf{y})\Phi_{m,k_1,k_2,\cdots,k_d}(\mathbf{y})\,d\mathbf{y}. \tag{3.89}$$

3.3.3. ERROR ANALYSIS AND PARAMETERS SELECTION

In this section, we present an error analysis of the 2D-SWIFT method and give a prescription on the selection of the parameters.

ERROR ANALYSIS

There are two main sources of error. The first one is the error due to the truncation of the integration range in (3.67), while the second one is the error caused by replacing the density function f in (3.77) by f^*. If we define,

$$\mathscr{E}_1 := |v(\mathbf{x},t) - v_1(\mathbf{x},t)|, \quad \text{and,} \quad \mathscr{E}_2 := |v_1(\mathbf{x},t) - v^*(\mathbf{x},t)|, \tag{3.90}$$

then the overall error $\mathscr{E} := |v(\mathbf{x},t) - v^*(\mathbf{x},t)|$ can be bounded by $\mathscr{E} \leq \mathscr{E}_1 + \mathscr{E}_2$. In what follows, we give a detailed analysis of the error. Let us start by considering,

$$\mathscr{E}_1 = |v(\mathbf{x},t) - v_1(\mathbf{x},t)| \leq \int_{\mathbb{R}\setminus[a_1,b_1]} \int_{\mathbb{R}\setminus[a_2,b_2]} |g(\mathbf{y}) f(\mathbf{y}|\mathbf{x})| \, d\mathbf{y}. \tag{3.91}$$

Since the mass in the tails of the density function f tends to zero at infinity, for every $\varepsilon_1 > 0$ there exist values $a_1, a_2, b_1, b_2 > 0$ such that,

$$\tau := \int_{\mathbb{R}\setminus[a_1,b_1]} \int_{\mathbb{R}\setminus[a_2,b_2]} f(\mathbf{y}|\mathbf{x}) \, d\mathbf{y} \leq \varepsilon_1. \tag{3.92}$$

If we assume that g is bounded in the domain of integration, then,

$$\mathscr{E}_1 \leq \tau \|g\|_\infty, \tag{3.93}$$

where $\|g\|_\infty$ is the infinity norm of g. This error can be made arbitrarily small by increasing the size of the truncation intervals. Observe that the assumption on the boundedness of function g is satisfied for put options. If we consider call options, then we may impose some assumptions on the decay rate of the density function f to have an estimation of the error. For sake of clarity and simplicity in the exposition, we stay within the assumption of bounded payoffs.

As mentioned in Section 3.3.2, there are three sources of error when approximating f by f^*. If we define,

(i) the projection error, given by,

$$\epsilon_p := \left| f(\mathbf{y}|\mathbf{x}) - f_1(\mathbf{y}|\mathbf{x}) \right| = \left| f(\mathbf{y}|\mathbf{x}) - \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} D_{m,k_1,k_2}(\mathbf{x}) \Phi_{m,k_1,k_2}(\mathbf{y}) \right|, \tag{3.94}$$

where D_{m,k_1,k_2} and Φ_{m,k_1,k_2} are in (3.69) and (3.59) respectively,

(ii) the sum truncation error,

$$\epsilon_t := |f_1(\mathbf{y}|\mathbf{x}) - f_2(\mathbf{y}|\mathbf{x})| = \left| \sum_{k_1 \notin \{l_1,\cdots,u_1\}} \sum_{k_2 \notin \{l_2,\cdots,u_2\}} D_{m,k_1,k_2}(\mathbf{x}) \Phi_{m,k_1,k_2}(\mathbf{y}) \right|, \tag{3.95}$$

(iii) and the coefficients approximation error,

$$\epsilon_c := |f_2(\mathbf{y}|\mathbf{x}) - f^*(\mathbf{y}|\mathbf{x})| = \left| \sum_{k_1=l_1}^{u_1} \sum_{k_2=l_2}^{u_2} \left(D_{m,k_1,k_2}(\mathbf{x}) - D^*_{m,k_1,k_2}(\mathbf{x}) \right) \Phi_{m,k_1,k_2}(\mathbf{y}) \right|, \tag{3.96}$$

then,

$$\begin{aligned} \mathscr{E}_2 &= \left| v_1(\mathbf{x},t) - v^*(\mathbf{x},t) \right| = \left| e^{-r\Delta t} \int_{a_1}^{b_1} \int_{a_2}^{b_2} g(\mathbf{y}) (f(\mathbf{y}|\mathbf{x}) - f^*(\mathbf{y}|\mathbf{x})) d\mathbf{y} \right| \\ &\leq e^{-r\Delta t} |b_1 - a_1| |b_2 - a_2| \, \|g\|_\infty (\epsilon_p + \epsilon_t + \epsilon_c), \end{aligned} \tag{3.97}$$

and,

$$\mathscr{E} \leq \left(\tau + e^{-r\Delta t} |b_1 - a_1| |b_2 - a_2| (\epsilon_p + \epsilon_t + \epsilon_c) \right) \|g\|_\infty. \tag{3.98}$$

The projection error in one dimension is studied in [51], and we use a similar procedure to give an estimation of that error for the two-dimensional case.

Lemma 3.3.6. *Consider an MRA generated by Shannon scaling function Φ defined in (3.47). With the same notation as before,*

$$\epsilon_p \leq K(2^m\pi, 2^m\pi), \tag{3.99}$$

where,

$$K(v_1, v_2) := \frac{1}{4\pi^2}\int_{|\omega_1|>v_1}\int_{|\omega_2|>v_2} |\hat{f}(\boldsymbol{\omega})|d\boldsymbol{\omega}, \tag{3.100}$$

and $\hat{f}$ is the bivariate Fourier transform of f defined in (3.45).

Proof. From (3.68), (3.69) and (3.94) we write,

$$\epsilon_p = \left| f(\mathbf{y}|\mathbf{x}) - \mathscr{P}_m f(\mathbf{y}|\mathbf{x}) \right|, \tag{3.101}$$

where,

$$\mathscr{P}_m f(\mathbf{y}|\mathbf{x}) = \sum_{k_1\in\mathbb{Z}}\sum_{k_2\in\mathbb{Z}} D_{m,k_1,k_2}(\mathbf{x})\Phi_{m,k_1,k_2}(\mathbf{y}), \tag{3.102}$$

and,

$$D_{m,k_1,k_2}(\mathbf{x}) := \iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})\Phi_{m,k_1,k_2}(\mathbf{y})\, d\mathbf{y}. \tag{3.103}$$

By Parseval's identity,

$$D_{m,k_1,k_2}(\mathbf{x}) = \frac{1}{4\pi^2}\iint_{\mathbb{R}^2} \hat{f}(\boldsymbol{\omega})\hat{\Phi}_{m,k_1,k_2}(\boldsymbol{\omega})d\boldsymbol{\omega}, \tag{3.104}$$

where $\boldsymbol{\omega} = (\omega_1, \omega_2)$ and $\hat{f}$ and $\hat{\Phi}_{m,k_1,k_2}$ denote the Fourier transform of f and Φ_{m,k_1,k_2} respectively. We observe that by (3.59) we can write,

$$\hat{\Phi}_{m,k_1,k_2} = \iint_{\mathbb{R}^2} \Phi_{m,k_1,k_2}(\mathbf{y})e^{-i(w_1y_1+w_2y_2)}\, d\mathbf{y} = \hat{\phi}_{m,k_1}(w_1)\hat{\phi}_{m,k_2}(w_2), \tag{3.105}$$

where, as pointed out in [67],

$$\hat{\phi}_{m,k}(w) = \frac{e^{-i\frac{k}{2^m}w}}{2^{m/2}}\text{rect}\left(\frac{w}{2^{m+1}\pi}\right), \tag{3.106}$$

and rect is the rectangle function, defined in (2.67). Now by (3.104), (3.105) and (3.106),

$$D_{m,k_1,k_2}(\mathbf{x}) = \frac{1}{4\pi^2}\frac{1}{2^m}\int\int_{\mathscr{C}} \hat{f}(\boldsymbol{\omega})e^{i\frac{k_1}{2^m}w_1}e^{i\frac{k_2}{2^m}w_2}d\boldsymbol{\omega}, \tag{3.107}$$

where $\mathscr{C} := [-2^m\pi, 2^m\pi]\times[-2^m\pi, 2^m\pi]$. If we replace this last expression of coefficients in (3.102) and interchange the summation and integration we get,

$$\mathscr{P}_m f(\mathbf{y}|\mathbf{x}) = \frac{1}{4\pi^2}\frac{1}{2^m}\int\int_{\mathscr{C}} \hat{f}(\boldsymbol{\omega})\left[\sum_{k_1\in\mathbb{Z}}\sum_{k_2\in\mathbb{Z}}\Phi_{m,k_1,k_2}(\mathbf{y})e^{i\frac{k_1}{2^m}w_1}e^{i\frac{k_2}{2^m}w_2}\right]d\boldsymbol{\omega}, \tag{3.108}$$

If we take into account that by Theorem 1.2.1 of [79],

$$\sum_{k\in\mathbb{Z}}\phi_{m,k}(y)e^{i\frac{k}{2^m}w} = 2^{\frac{m}{2}}e^{iwy}, \quad \text{when} \quad w\in(-2^m\pi, 2^m\pi), \tag{3.109}$$

then the projection $\mathscr{P}_m f$ can be written as,

$$\mathscr{P}_m f(\mathbf{y}|\mathbf{x}) = \frac{1}{4\pi^2}\int\int_{\mathscr{C}} \hat{f}(\boldsymbol{\omega}) e^{i\boldsymbol{\omega}'\mathbf{y}} d\boldsymbol{\omega}. \tag{3.110}$$

Finally, by (3.110) and the definition of the inverse Fourier transform[1] of f, we have,

$$\epsilon_p = \left| f(\mathbf{y}|\mathbf{x}) - \mathscr{P}_m f(\mathbf{y}|\mathbf{x}) \right| = \frac{1}{4\pi^2}\left| \int\int_{\mathbb{R}^2\setminus\mathscr{C}} \hat{f}(\boldsymbol{\omega}) e^{i\boldsymbol{\omega}'\mathbf{y}} d\boldsymbol{\omega} \right| \leq \frac{1}{4\pi^2}\int\int_{\mathbb{R}^2\setminus\mathscr{C}} |\hat{f}(\boldsymbol{\omega})| d\boldsymbol{\omega}, \tag{3.112}$$

and this concludes the proof. □

The sum truncation error ϵ_t depends on the size of the scaling coefficients D_{m,k_1,k_2}, since from (3.95) we have,

$$\epsilon_t \leq 2^m \cdot \sum_{k_1 \notin \{l_1,\cdots,u_1\}} \sum_{k_2 \notin \{l_2,\cdots,u_2\}} \left| D_{m,k_1,k_2}(\mathbf{x}) \right|, \tag{3.113}$$

and by Lemma 3.3.6,

$$\left| f(\mathbf{y}|\mathbf{x}) - \sum_{k_1\in\mathbb{Z}} \sum_{k_2\in\mathbb{Z}} D_{m,k_1,k_2}(\mathbf{x}) \Phi_{m,k_1,k_2}(\mathbf{y}) \right| \leq K(2^m\pi, 2^m\pi). \tag{3.114}$$

In particular, if we evaluate expression (3.114) in $\mathbf{y} = \left(\frac{k_1}{2^m}, \frac{k_2}{2^m}\right)$ then,

$$\left| f\left(\frac{k_1}{2^m}, \frac{k_2}{2^m}\,\middle|\,\mathbf{x}\right) - 2^m D_{m,k_1,k_2}(\mathbf{x}) \right| \leq K(2^m\pi, 2^m\pi), \tag{3.115}$$

which means that the coefficients D_{m,k_1,k_2} are very well approximated by,

$$D_{m,k_1,k_2}(\mathbf{x}) \approx \frac{1}{2^m} f\left(\frac{k_1}{2^m}, \frac{k_2}{2^m}\,\middle|\,\mathbf{x}\right), \tag{3.116}$$

when $|\hat{f}(\boldsymbol{\omega})|$ in (3.100) decays very fast, as typically happens with the densities considered in this work. Finally, if we assume that the density f tends to zero at minus and plus infinity, then the error ϵ_t can be neglected for sufficiently big truncation values l_1, l_2, u_1, u_2 in (3.70). In Section 3.3.3 we give a detailed explanation on how to select these values.

The last error we need to estimate is ϵ_c defined in (3.96). For this purpose, we present the following two lemmas.

Lemma 3.3.7 (Lemma 2 of [67]). *Define the absolute error* $\mathscr{E}_V(t) := sinc(t) - sinc^*(t)$. *Then,*

$$|\mathscr{E}_V(t)| \leq \frac{(\pi c)^2}{2^{2(J+1)} - (\pi c)^2}, \tag{3.117}$$

for $t \in [-c, c]$, *where* $c \in \mathbb{R}, c > 0$ *and* $J \geq \log_2(\pi c)$.

[1] The two-dimensional inverse Fourier transform f of $\hat{f}$ is by definition,

$$f(\mathbf{y}|\mathbf{x}) = \frac{1}{4\pi^2}\iint_{\mathbb{R}^2} \hat{f}(\boldsymbol{\omega}) e^{i\boldsymbol{\omega}'\mathbf{y}} d\boldsymbol{\omega}. \tag{3.111}$$

Lemma 3.3.8. *Define the error* $\bar{\mathscr{E}}_V(t_1,t_2) := sinc(t_1)sinc(t_2) - sinc^*(t_1)sinc^*(t_2)$. *Then,*

$$\left|\bar{\mathscr{E}}_V(t_1,t_2)\right| \leq |\mathscr{E}_V(t_1)| + |\mathscr{E}_V(t_2)|. \tag{3.118}$$

Proof. We observe that,

$$\begin{aligned}\bar{\mathscr{E}}_V(t_1,t_2) &= \text{sinc}(t_1)\text{sinc}(t_2) - \text{sinc}^*(t_1)\text{sinc}^*(t_2) - \text{sinc}^*(t_1)\text{sinc}(t_2) + \text{sinc}^*(t_1)\text{sinc}(t_2)\\ &= \text{sinc}(t_2)[\text{sinc}(t_1) - \text{sinc}^*(t_1)] + \text{sinc}^*(t_1)[\text{sinc}(t_2) - \text{sinc}^*(t_2)]\\ &= \text{sinc}(t_2)\mathscr{E}_V(t_1) + \text{sinc}^*(t_1)\mathscr{E}_V(t_1).\end{aligned} \tag{3.119}$$

The proof concludes by noting that $|\text{sinc}(t_2)| = |\text{sinc}^*(t_1)| = 1$ for all $t_1, t_2 \in \mathbb{R}$. □

Theorem 3.3.9. *Let* $F(\mathbf{x})$ *be the distribution function of a two-dimensional random variable* $\boldsymbol{X}$ *and define* $H(\mathbf{x}) := F(-\mathbf{x}) + 1 - F(\mathbf{x})$. *Let* $c_1, c_2 > 0$ *be constants such that* $H(c_1,c_2) < \epsilon$, *for* $\epsilon > 0$, *and let* $c := \max(c_1,c_2)$. *Define,*

$$M^1_{m,k_1} := \max\left(|2^m c - k_1|, |2^m c + k_1|\right), \quad M^2_{m,k_2} := \max\left(|2^m c - k_2|, |2^m c + k_2|\right), \tag{3.120}$$

$$M_{m,k_1,k_2} := \max\left(M^1_{m,k_1}, M^2_{m,k_2}\right), \tag{3.121}$$

and consider $J \geq \log_2\left(\pi M_{m,k_1,k_2}\right)$. *Then,*

$$|D_{m,k_1,k_2}(\mathbf{x}) - D^*_{m,k_1,k_2}(\mathbf{x})| \leq 2^m\left(2\epsilon + 4c\|f(\cdot\,|\,\mathbf{x})\|_2 \frac{\left(\pi M_{m,k_1,k_2}\right)^2}{2^{2(J+1)} - \left(\pi M_{m,k_1,k_2}\right)^2}\right), \tag{3.122}$$

and $\lim_{J\to+\infty} D^*_{m,k_1,k_2}(\mathbf{x}) = D_{m,k_1,k_2}(\mathbf{x})$.

Proof. From (3.59), (3.69), (3.71), (3.72) and taking into account the equivalence (3.76), we write,

$$\begin{aligned}&\left|D_{m,k_1,k_2}(\mathbf{x}) - D^*_{m,k_1,k_2}(\mathbf{x})\right|\\ &= 2^m\left|\iint_{\mathbb{R}^2} f(\mathbf{y}|\mathbf{x})\left[\text{sinc}(2^m y_1 - k_1)\text{sinc}(2^m y_2 - k_2) \;-\; \text{sinc}^*(2^m y_1 - k_1)\text{sinc}^*(2^m y_2 - k_2)\right] d\mathbf{y}\right|\\ &\leq 2^m\left[\int\int_{\mathscr{D}^c} f(\mathbf{y}|\mathbf{x})\left|\text{sinc}(2^m y_1 - k_1)\text{sinc}(2^m y_2 - k_2) \;-\; \text{sinc}^*(2^m y_1 - k_1)\text{sinc}^*(2^m y_2 - k_2)\right| d\mathbf{y}\right.\\ &+\left.\left|\int\int_{\mathscr{D}} f(\mathbf{y}|\mathbf{x})\left(\text{sinc}(2^m y_1 - k_1)\text{sinc}(2^m y_2 - k_2) \;-\text{sinc}^*(2^m y_1 - k_1)\text{sinc}^*(2^m y_2 - k_2)\right) d\mathbf{y}\right|\right|.\end{aligned} \tag{3.123}$$

where $\mathscr{D} := [-c_1,c_1]\cup[-c_2,c_2]$ and $\mathscr{D}^c := \mathbb{R}^2\backslash\mathscr{D}$.

Since the mass in the tails of the density f tends to zero when c_1 and c_2 tend to infinity, for all $\epsilon > 0$ there exist $c_1, c_2 > 0$ such that $H(c_1,c_2) < \epsilon$. Further,

$$\left|\text{sinc}(2^m y_1 - k_1)\text{sinc}(2^m y_2 - k_2) - \text{sinc}^*(2^m y_1 - k_1)\text{sinc}^*(2^m y_2 - k_2)\right| \leq 2, \tag{3.124}$$

for all $\mathbf{y} \in \mathbb{R}^2$, and therefore the first integral at the right hand-side of inequality (3.123) is bounded by 2ϵ.

Now if we apply the Cauchy-Schwarz inequality to the second integral then,

$$\begin{aligned}
&\left|\int\int_{\mathscr{D}} f(\mathbf{y}|\mathbf{x})\left(\operatorname{sinc}(2^m y_1-k_1)\operatorname{sinc}(2^m y_2-k_2)-\operatorname{sinc}^*(2^m y_1-k_1)\operatorname{sinc}^*(2^m y_2-k_2)\right)d\mathbf{y}\right| \\
&\leq \|f(\cdot|\mathbf{x})\|_2\left(\int\int_{\mathscr{D}}\left(\bar{\mathscr{E}}_V(2^m y_1-k_1,2^m y_2-k_2)\right)^2 d\mathbf{y}\right)^{\frac{1}{2}} \\
&\leq \|f(\cdot|\mathbf{x})\|_2\left(\int\int_{\mathscr{D}}\left(\mathscr{E}_V(2^m y_1-k_1)+\mathscr{E}_V(2^m y_2-k_2)\right)^2 d\mathbf{y}\right)^{\frac{1}{2}},
\end{aligned} \tag{3.125}$$

where the last inequality is satisfied by Lemma 3.3.8. We observe that if $-c \leq y_i \leq c$, then $-2^m c-k_i \leq 2^m y_i-k_i \leq 2^m c-k_i$ and therefore, $2^m y_i-k_i \in [-M^i_{m,k_i}, M^i_{m,k_i}]$, where by definition $c=\max(c_1,c_2)$ and $i=1,2$. We note that,

$$[-M^i_{m,k_i}, M^i_{m,k_i}] \subset [-M_{m,k_1,k_2}, M_{m,k_1,k_2}], \quad i=1,2, \tag{3.126}$$

then by Lemma 3.3.7, the integral at the right hand-side of the second inequality in (3.125) is bounded by,

$$2c\cdot\frac{2(\pi M_{m,k_1,k_2})^2}{2^{2(J+1)}-(\pi M_{m,k_1,k_2})^2}, \tag{3.127}$$

when $J \geq \log_2\left(\pi M_{m,k_1,k_2}\right)$.

Finally, by (3.123), (3.125) and (3.127) we end up with the error estimate,

$$|D_{m,k_1,k_2}(\mathbf{x})-D^*_{m,k_1,k_2}(\mathbf{x})| \leq 2^m\left(2\epsilon+4c\|f(\cdot\,|\,\mathbf{x})\|_2\frac{\left(\pi M_{m,k_1,k_2}\right)^2}{2^{2(J+1)}-\left(\pi M_{m,k_1,k_2}\right)^2}\right). \tag{3.128}$$

□

Finally, from (3.96) and Theorem 3.3.9,

$$\begin{aligned}
\epsilon_c &\leq 2^m(u_1-l_1+1)(u_2-l_2+1)\left|D_{m,k_1,k_2}(\mathbf{x})-D^*_{m,k_1,k_2}(\mathbf{x})\right| \\
&\leq 2^{2m}(u_1-l_1+1)(u_2-l_2+1)\left(2\epsilon+4c\|f(\cdot\,|\,\mathbf{x})\|_2\frac{\left(\pi M_{m,k_1,k_2}\right)^2}{2^{2(J+1)}-\left(\pi M_{m,k_1,k_2}\right)^2}\right).
\end{aligned} \tag{3.129}$$

PARAMETERS SELECTION

The parameters of the method are the scale of approximation m in (3.68), the range of coefficients $l_1 \leq k_1 \leq u_1$, $l_2 \leq k_2 \leq u_2$ in (3.70), the number of terms $N=2^J$ to approximate the cardinal sine function in (3.71), and the truncation of the integration domain $[a_1,b_1]\times[a_2,b_2]$ in expression (3.77).

Concerning the scale of approximation m, we know from Lemma 3.3.6 that the approximation error decreases with m. Further, this error is even smaller when the modulus of the characteristic function decays rapidly. We propose an adaptive computation of m by following this strategy. We select m such that given a tolerance $\varepsilon_2>0$,

$$|\hat{f}(-2^m\pi,-2^m\pi)|+|\hat{f}(-2^m\pi,2^m\pi)|+|\hat{f}(2^m\pi,-2^m\pi)|+|\hat{f}(2^m\pi,2^m\pi)|<\varepsilon_2. \tag{3.130}$$

An initial guess of the integration domain $[a_1,b_1]\times[a_2,b_2]$ is given (for example) by the $n-$th cumulant $\kappa_{i,n}$ of $X_t^i := \log S_t^i$ and the parameter $L=10$, like for the 2D-COS method in [74],

$$\begin{aligned}[a_1,b_1]\times[a_2,b_2] = &\left[\kappa_{1,1}-L\sqrt{\kappa_{2,1}+\sqrt{\kappa_{4,1}}},\,\kappa_{1,1}+L\sqrt{\kappa_{2,1}+\sqrt{\kappa_{4,1}}}\right]\\ &\times\left[\kappa_{1,2}-L\sqrt{\kappa_{2,2}+\sqrt{\kappa_{4,2}}},\,\kappa_{1,2}+L\sqrt{\kappa_{2,2}+\sqrt{\kappa_{4,2}}}\right].\end{aligned} \tag{3.131}$$

We compute the coefficients D^*_{m,k_1,k_2} by means of an FFT algorithm, with k_i ranging from $l_i=\lfloor 2^m a_i\rfloor$ to $u_i=\lceil 2^m b_i\rceil$ and $i=1,2$. It is worth underlining that this a-priori truncation facilitates the application of an FFT algorithm as mentioned in Remark 3.3.5 to get the stated computational complexity. Then, in order to know whether the initial truncated range is accurate or not, we measure the size of the density coefficients at the boundaries. For this purpose, we compute $D^*_{m,l_1,0}, D^*_{m,u_1,0}, D^*_{m,0,l_2}, D^*_{m,0,u_2}$, since as we have seen in (3.116) the size of those coefficients is closely related to the value of the density at the boundary points. If necessary, we can compute extra coefficients until the desired precision is reached. Moreover, we can also calculate the volume underneath the surface represented by the density f as a byproduct and verify that the volume is close to 1. Considering the partition of the domain $[a_1,b_1]\times[a_2,b_2]$ given by points of the form $\left(\frac{h_1}{2^m},\frac{h_2}{2^m}\right)$ for $h_1,h_2\in\mathbb{Z}$, the two-dimensional composite trapezoidal rule gives us the following approximation,

$$\begin{aligned}\mathrm{V}(f^*) := \iint_{\mathbb{R}^2} f^*(\mathbf{y}|\mathbf{x})\,d\mathbf{y} \approx \mathscr{S} := \frac{1}{2^{m+2}}\Bigg[&D^*_{m,l_1,l_2}+D^*_{m,u_1,l_2}+D^*_{m,l_1,u_2}+D^*_{m,u_1,u_2}\\ &+2\sum_{i=1}^{K_1-1}\left(D^*_{m,l_1+i,l_2}+D^*_{m,l_1+i,u_2}\right)+2\sum_{j=1}^{K_2-1}\left(D^*_{m,l_1,l_2+j}+D^*_{m,u_1,l_2+j}\right)\\ &+4\sum_{j=1}^{K_2-1}\left(\sum_{i=1}^{K_1-1}D^*_{m,l_1+i,l_2+j}\right)\Bigg],\end{aligned} \tag{3.132}$$

with $K_i=u_i-l_i+1$. For the sake of simplicity, we will consider $a=\min(a_1,a_2)$ and $b=\max(b_1,b_2)$, and work with a, b instead of a_1, a_2, b_1 and b_2, being this selection more conservative.

Finally, the selection of J is related to the error studied in Theorem 3.3.9. Although a different J can be selected for each pair (k_1,k_2), we prefer to consider a constant J, defined here by $J=\lceil\log_2(\pi\max_{k_1,k_2}M_{m,k_1,k_2})\rceil$ where $\lceil\cdot\rceil$ stands for the ceiling function. The reason is that, in practice, the computationally most involved part in (3.72) is the evaluation of $\hat{f}$ at the grid points. Those values can be computed only once and used in the FFT algorithm mentioned in Remark 3.3.5.

3.3.4. NUMERICAL EXPERIMENTS

We present a wide variety of examples to test the 2D-SWIFT method for pricing European rainbow options. We consider arithmetic basket call options, geometric basket put option, spread options, call-on-max and put-on-min options, and correlation options.

For simplicity, and without loss of generality, in all examples we assume $t=0$. The asset price is modelled by either correlated geometric Brownian motions or by Merton's jumps-

diffusion (JD) process (but any other Lévy process or process with known characteristic function could be used). The cumulants for these dynamics are well-known and we therefore consider them as our initial guess in (3.131). It is worth remarking that this selection of the interval appears to be accurate in all the examples considered in this work, and we do not need to compute extra coefficients. The programs in this chapter have been coded in MATLAB and run on a Dell Vostro 320 with Intel Core 2 Duo E7500 2.93GHz processor and 4GB RAM.

Under GBM dynamics the risk-neutral asset prices evolve according to the following dynamics,

$$dS_t^i = rS_t^i dt + \sigma_i S_t^i dW_t^i, \quad i = 1,2, \tag{3.133}$$

with correlation $dW_t^i dW_t^j = \rho_{ij} dt$, r the risk-free rate, and σ_i the volatility of asset i. We switch to the log-process $X_t^i := \log S_t^i$,

$$dX_t^i = (r - \frac{1}{2}\sigma_i^2)dt + \sigma_i dW_t^i. \tag{3.134}$$

The log-asset prices at time t given the current state at $t = 0$ are bivariate normally distributed, i.e.,

$$\mathbf{X}_t \sim \mathcal{N}(\mathbf{X}_0 + \boldsymbol{\mu}\Delta t, \boldsymbol{\Sigma}), \tag{3.135}$$

with $\mu_i = r - \frac{1}{2}\sigma_i^2$ and covariance matrix $\Sigma_{ij} = \sigma_i \sigma_j \rho_{ij} \Delta t$. The Fourier transform function (defined in (3.45)) reads $\hat{f}(\mathbf{u}|\mathbf{x}) = e^{-i\mathbf{x}'\mathbf{u}} \hat{f}_{\text{Lévy}}(-\mathbf{u})$, with,

$$\hat{f}_{\text{Lévy}}(\mathbf{u}) = \exp(i\boldsymbol{\mu}'\Delta t\mathbf{u} - \frac{1}{2}\mathbf{u}'\boldsymbol{\Sigma}\mathbf{u}). \tag{3.136}$$

Under a Merton's jump-diffusion process the asset prices follow the dynamics described by the stochastic differential equation,

$$dS_t^i = (r - \lambda\kappa_i)S_t^i dt + \sigma_i S_t^i dW_t^i + J_i S_t^i dq_t, \quad i = 1,2, \tag{3.137}$$

with $\kappa_i := \mathbb{E}\left[e^{J_i} - 1\right]$, q_t a Poisson process with mean arrival date λ, and $\mathbf{J} = (J_1, J_2)$ bivariate normally distributed jump sizes with mean $\boldsymbol{\mu}^J = [\mu_1^J, \mu_2^J]'$ and covariance matrix $\Sigma_{ij}^J = \sigma_i^J \sigma_j^J \rho_{ij}^J$. The log-processes $X_t^i := \log S_t^i$ read,

$$dX_t^i = (r - \lambda\kappa_i - \frac{1}{2}\sigma_i^2)dt + \sigma_i dW_t^i + J_i dq_t. \tag{3.138}$$

The Fourier transform function is $\hat{f}(\mathbf{u}|\mathbf{x}) = e^{-i\mathbf{x}'\mathbf{u}} \hat{f}_{\text{Lévy}}(-\mathbf{u})$, with,

$$\hat{f}_{\text{Lévy}}(\mathbf{u}) = \exp\left(i\boldsymbol{\mu}'\Delta t\mathbf{u} - \frac{1}{2}\mathbf{u}'\boldsymbol{\Sigma}\mathbf{u}\right)\exp\left(\lambda\Delta t\left(\exp\left(i\boldsymbol{\mu}'^J\mathbf{u}\frac{1}{2}\mathbf{u}'\boldsymbol{\Sigma}^J\mathbf{u}\right) - 1\right)\right), \tag{3.139}$$

where $\mu_i = (r - \lambda\kappa_i - \frac{1}{2}\sigma_i^2)\Delta t$, $\Sigma_{ij} = \sigma_i\sigma_j\rho_{ij}\Delta t$ and $\kappa_i = e^{\mu_i^J + \frac{1}{2}(\sigma_i^J)^2} - 1$.

BASKET OPTIONS ON TWO UNDERLYINGS

We present the pricing of arithmetic basket call options driven by two-dimensional correlated GBM. Table 3.1 contains the payoff of arithmetic and geometric basket options on two underlyings.

In the arithmetic case, the payoff coefficients (3.79) are given by,

	Arithmetic	Geometric
Call	$g(y_1,y_2)=\left(\frac{1}{2}e^{y_1}+\frac{1}{2}e^{y_2}-K\right)^+$	$g(y_1,y_2)=(\sqrt{e^{y_1}}\sqrt{e^{y_2}}-K)^+$
Put	$g(y_1,y_2)=\left(K-(\frac{1}{2}e^{y_1}+\frac{1}{2}e^{y_2})\right)^+$	$g(y_1,y_2)=(K-\sqrt{e^{y_1}}\sqrt{e^{y_2}})^+$

Table 3.1: Payoffs for basket options on two underlyings.

$$G_{m,k_1,k_2}=\iint_{\mathbb{R}^2}\left(\frac{1}{2}e^{y_1}+\frac{1}{2}e^{y_2}-K\right)^+\Phi_{m,k_1,k_2}(\mathbf{y})\,d\mathbf{y}. \tag{3.140}$$

If we truncate the integration range, use the exponential approximation (3.74) of the sinc function and interchange sums with integrals, then $G_{m,k_1,k_2}\approx G^*_{m,k_1,k_2}$, where,

$$G^*_{m,k_1,k_2}:=\frac{2^m}{N^2}\sum_{j_2=1}^{N}\sum_{j_1=1}^{N}e^{-i\Omega_{j_1}k_1}e^{-i\Omega_{j_2}k_2}\int_a^b\int_a^b\left(\frac{1}{2}e^{y_1}+\frac{1}{2}e^{y_2}-K\right)^+e^{i\Omega_{j_1}2^m y_1}e^{i\Omega_{j_2}2^m y_2}\,dy_1\,dy_2, \tag{3.141}$$

and $\Omega_{j_1}=-\pi+\frac{2j_1-1}{N}\pi,\Omega_{j_2}=-\pi+\frac{2j_2-1}{N}\pi$. The double integral in (3.141) cannot be solved analytically, and we therefore apply numerical integration. We approximate it by a midpoint quadrature with Q terms, obtaining,

$$G^*_{m,k_1,k_2}\approx\frac{2^m(b-a)^2}{N^2Q^2}\sum_{j_2=1}^{N}\sum_{j_1=1}^{N}\sum_{l_2=1}^{Q}\sum_{l_1=1}^{Q}\left(\frac{1}{2}e^{\eta_{l_1}}+\frac{1}{2}e^{\eta_{l_2}}-K\right)^+e^{i\Omega_{j_1}2^m\eta_{l_1}}e^{i\Omega_{j_2}2^m\eta_{l_2}}e^{-i\Omega_{j_1}k_1}e^{-i\Omega_{j_2}k_2}, \tag{3.142}$$

where $\eta_{l_i}=a+\frac{b-a}{2Q}(2l_i-1)$. After rearranging terms, we end up with the following formula,

$$\begin{aligned}G^*_{m,k_1,k_2}\approx&\frac{2^m(b-a)^2}{N^2Q^2}e^{2i2^m\pi\left(1+\frac{1}{N}\right)\left(\frac{b-a}{2Q}-a\right)}e^{i\pi\left(1+\frac{1}{N}\right)(k_1+k_2)}\sum_{j_2=1}^{N}\sum_{j_1=1}^{N}e^{i2^m\frac{2\pi}{N}\left(a+\frac{b-a}{2Q}\right)(j_1+j_2)}\\&\cdot\sum_{l_2=1}^{M}\sum_{l_1=1}^{M}X(l_1,l_2)e^{\frac{i2\pi}{M}j_1(l_1-1)}e^{\frac{i2\pi}{M}j_2(l_2-1)}e^{\frac{-i2\pi}{N}k_1(j_1-1)}e^{\frac{-i2\pi}{N}k_2(j_2-1)},\end{aligned} \tag{3.143}$$

for $M:=\frac{NQ}{2^m(b-a)}$. In order to ensure that $M\in\mathbb{Z}$, we choose Q of the form $Q=2^n(b-a)$, for $n\geq m$, and,

$$X(l_1,l_2):=\begin{cases}\left(\frac{1}{2}e^{\eta_{l_1}}+\frac{1}{2}e^{\eta_{l_2}}-K\right)^+e^{-\frac{i2^m\pi(b-a)\left(1+\frac{1}{N}\right)(l_1+l_2)}{Q}}, & \text{if } l_1\leq Q \text{ and } l_2\leq Q,\\ 0, & \text{otherwise.}\end{cases} \tag{3.144}$$

Clearly we can apply a combination of FFT and inverse FFT algorithms to compute the sums efficiently in (3.143). Figure 3.4 shows the log-scale plots of the errors (left) and CPU time (right), where n is chosen such that $Q=2^n(b-a)$ and m is the level of resolution. We take the parameters and reference price from [74]. We observe that 2D-SWIFT converges exponentially.

It is worth mentioning that the price of a geometric basket option under GBM equals the Black-Scholes price of the corresponding European option with initial price $\hat{S}_0=\sqrt{S_0^1}\sqrt{S_0^2}$,

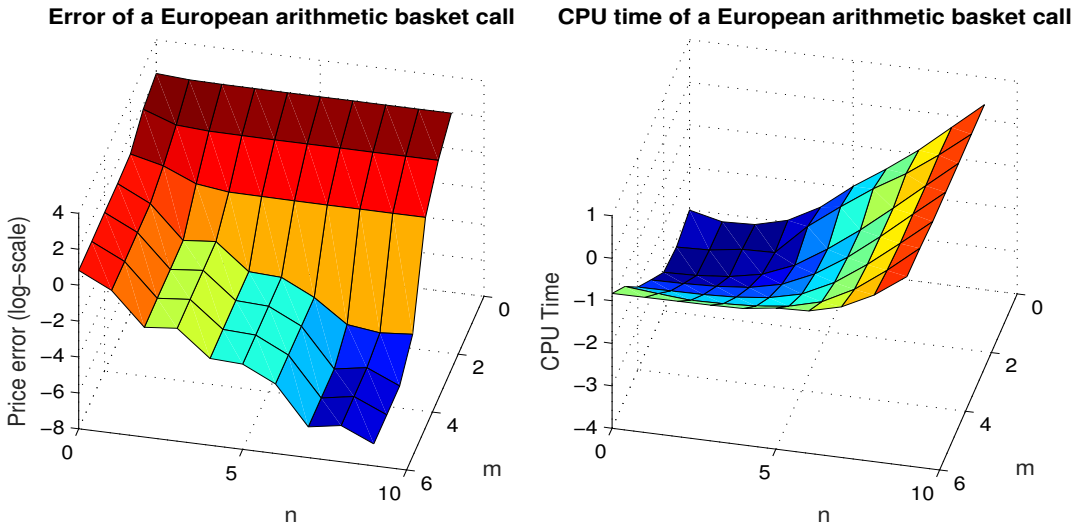


Figure 3.4: Error (left) and CPU time (right) in seconds of a two-dimensional arithmetic basket call option under the GBM dynamics. The parameter values are $S_0 = (90, 110)$, $r = 0.04$, $\sigma_1 = 0.2$, $\sigma_2 = 0.3$, $\rho = 0.25$, $T = 1$, $K = 100$ and $L = 10$. Reference: 10.173230.

volatility $\hat{\sigma}$ and dividend rate $\hat{\delta}$, where,

$$\hat{\sigma} = \frac{1}{2}\sqrt{\sum_{i,j} \sigma_i \sigma_j \rho_{ij}} \quad \text{and} \quad \hat{\delta} = \frac{1}{2}\sum_i \left(\delta_i + \frac{1}{2}\sigma_i^2\right) - \frac{1}{2}\hat{\sigma}^2. \tag{3.145}$$

So, we can perform a consistency check and compare the results of our method with the analytical option values.

Now, we price a geometric put option driven by two-dimensional correlated GBM. When pricing any option with 2D-SWIFT we need first to compute payoff coefficients, the payoff coefficients in this kind of options can be obtained analytically.

In Figure 3.5, we compare the error and the CPU time (expressed in seconds) between 2D-COS and 2D-SWIFT methods when using the same number of coefficients. We observe exponential convergence of the 2D-SWIFT method as well as 2D-COS in the left-side graph, and that the CPU time for 2D-SWIFT is higher than 2D-COS, although 2D-SWIFT is a very competitive method.

Depending on the type of option, the double integral in (3.140) cannot be solved analytically, and we therefore need to apply a numerical quadrature scheme. In this example we also solve the integral numerically in order to have an insight in the behaviour of the error according to the number of terms used in the integral approximation.

In Figure 3.6 the surfaces of the errors (left) and CPU time (right) are shown, where n is chosen such that $Q = 2^n(b-a)$ (recall that Q is the number of points in the midpoint quadrature for the double integral), and m is the level of resolution. The parameter values are the same as in Figure 3.5, and we present the absolute errors in Table 3.2.

From Figure 3.6 and Table 3.2 we observe the convergence of the method with respect to the scale of approximation m.

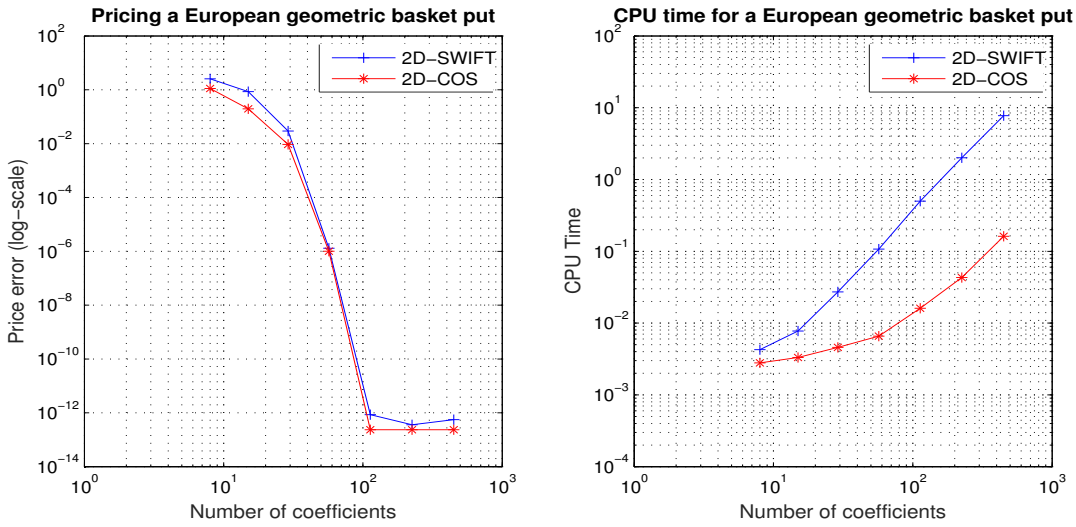


Figure 3.5: Log-error and CPU time in seconds of a two-dimensional geometric basket put option under the GBM dynamics. The parameter values are $S_0 = (90, 110), r = 0.04, \sigma_1 = 0.2, \sigma_2 = 0.3, \rho = 0.25, K = 100, L = 10$ and $T = 1$.

		m						
		0	1	2	3	4	5	6
	1	5.53e−01	2.09e+00	7.12e−01	3.68e−01	3.68e−01	3.68e−01	3.68e−01
	2	8.71e−01	1.05e+00	1.09e−01	5.34e−02	5.34e−02	5.34e−02	5.34e−02
	3	7.92e−01	1.04e+00	1.68e−02	1.84e−02	1.84e−02	1.84e−02	1.84e−02
	4	7.95e−01	1.01e+00	2.79e−02	6.14e−03	6.14e−03	6.14e−03	6.14e−03
n	5	7.98e−01	9.97e−01	3.26e−02	6.95e−04	6.96e−04	6.96e−04	6.96e−04
	6	7.98e−01	9.96e−01	3.27e−02	4.77e−04	4.79e−04	4.79e−04	4.79e−04
	7	7.98e−01	9.95e−01	3.33e−02	1.41e−04	1.40e−04	1.40e−04	1.40e−04
	8	7.98e−01	9.95e−01	3.31e−02	1.46e−05	1.33e−05	1.33e−05	1.33e−05
	9	7.98e−01	9.95e−01	3.31e−02	2.58e−06	4.12e−06	4.12e−06	4.12e−06

Table 3.2: Absolut errors for a two-dimensional geometric basket put option under GBM dynamics. The reference price is 6.696961159991261.

SPREAD OPTIONS

Next we price a two-dimensional European spread call option with strike K on two assets driven by a two-dimensional correlated GBM. Because of their generic nature, spread options are used in markets like the fixed income markets, the currency and foreign exchange markets, the commodity futures markets and the energy markets. The payoff function for the European call spread is given by,

$$g(y_1, y_2) = \max(e^{y_1} - e^{y_2} - K, 0). \tag{3.146}$$

We distinguish two cases, when $K = 0$ (this case is known as the exchange of assets) the payoff coefficients in (3.79) can be obtained analytically, while for $K > 0$ we use the midpoint quadrature, as presented in the previous example.

We report in Figure 3.7 the error and CPU time for the 2D-SWIFT method when the underlying processes follow GBM dynamics. We observe exponential convergence and competitive CPU time for 2D-SWIFT method. We take as reference price the result of the 2D-SWIFT

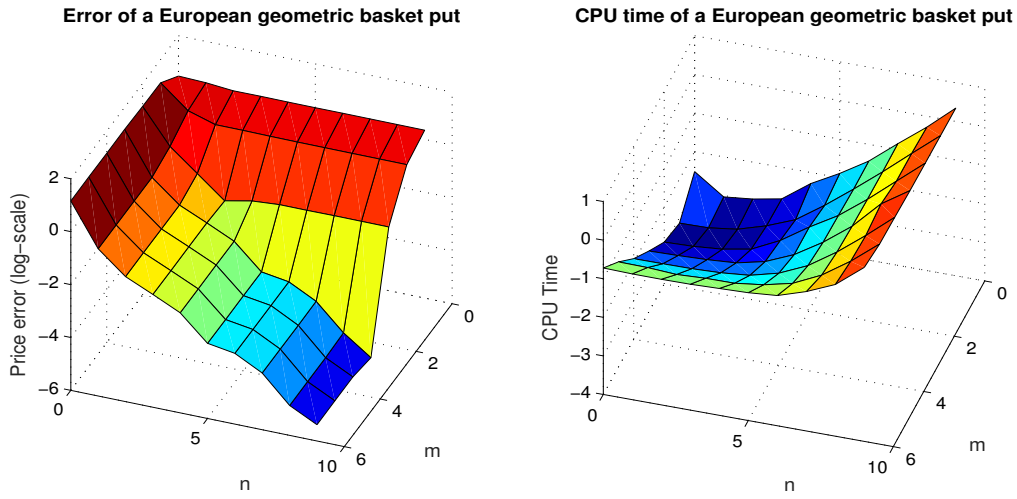


Figure 3.6: Error and CPU time in seconds (both measured in log-scale) of a two-dimensional geometric basket put option under the GBM dynamics.

method with $m = 12$ and $n = 12$.

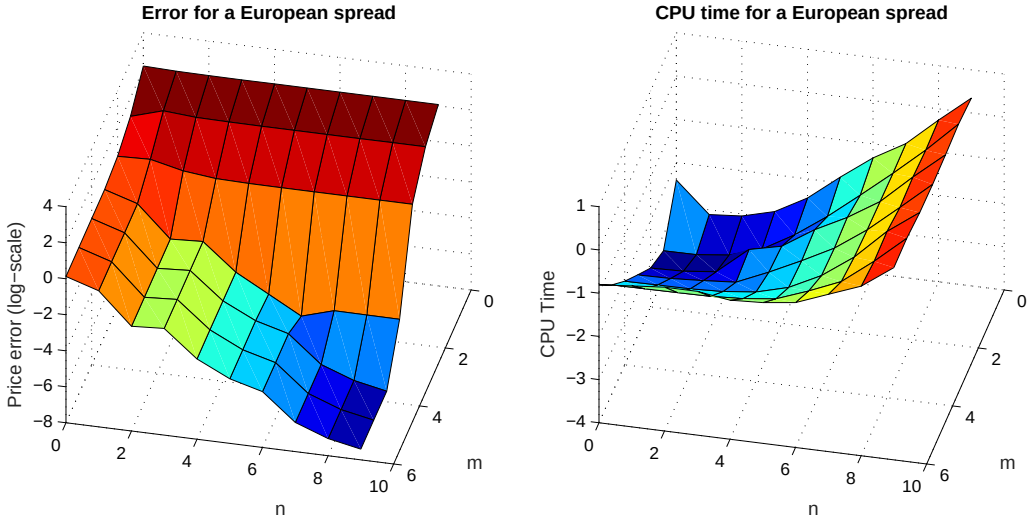


Figure 3.7: Error (left) and CPU time in seconds (right) corresponding to the pricing of a two-dimensional European call spread under the GBM dynamics. The parameters are $S_0 = (90, 110)$, $r = 0.04$, $\sigma_1 = 0.2$, $\sigma_2 = 0.3$, $\rho = 0.25$, $K = 20$, $L = 10$ and $T = 1$. Reference: 1.352591908717933.

As previously mentioned, exchange of assets is the name given to spread options when $K = 0$. If we assume that the processes follow GBM dynamics, we can use the Margrabe formula [52] as the reference price to compare the error and the CPU time between 2D-COS and 2D-SWIFT methods. We use the parameters from problem number 6 of the BENCHOP project [81]. The results are shown in Figure 3.8, where we observe again exponential convergence and competitive CPU time for the 2D-SWIFT method.

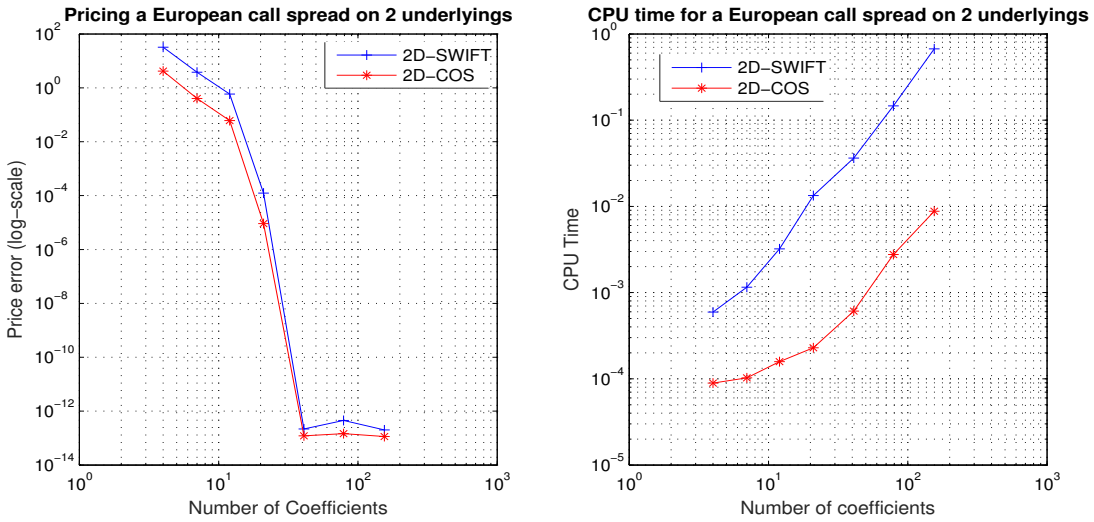


Figure 3.8: Error (left) and CPU time in seconds (right) corresponding to the pricing of a two-dimensional European call spread under the GBM dynamics. The parameters are $S_0 = (100,90)$, $r = 0.03, \sigma_1 = \sigma_2 = 0.15, \rho = 0.5, K = 0, L = 10$ and $T = 1$.

OPTIONS ON THE MINIMUM OR THE MAXIMUM OF TWO RISKY ASSETS

Here we consider a two-dimensional European option either on the minimum or on the maximum of two assets. Table 3.3 shows the corresponding payoffs.

	on minimum	on maximum
Call	$g(y_1,y_2) = (\min(e^{y_1}, e^{y_2}) - K)^+$	$g(y_1,y_2) = (\max(e^{y_1}, e^{y_2}) - K)^+$
Put	$g(y_1,y_2) = (K - \min(e^{y_1}, e^{y_2}))^+$	$g(y_1,y_2) = (K - \max(e^{y_1}, e^{y_2}))^+$

Table 3.3: Payoffs for options on the minimum/maximum of two assets.

For the payoffs described in Table 3.3, the double integral of the payoff coefficients in (3.79) can be solved analytically. In Figure 3.9 we present the error of some of these options driven by different dynamics and compare them with the 2D-COS method. Again, the 2D-SWIFT method converges exponentially.

CORRELATION OPTIONS

A correlation option is an extension of the plain vanilla European call to two dimensions. Its payoff is the product of two European calls with different strikes. Similar to spread options, correlation options allow the purchaser to speculate on how asset prices will move together, as the option requires both assets to move in the same direction in order to have at maturity time a non-zero value. The payoff of a correlation option is given by,

$$g(y_1,y_2) = \left(e^{y_1} - K_1\right)^+ \left(e^{y_2} - K_2\right)^+. \tag{3.147}$$

The payoff coefficients in this case are computed analytically. We present in Table 3.4 the relative error when pricing with the 2D-SWIFT method under GBM and JD dynamics. The parameters corresponding to the GBM process are $S_0 = (90,100)$, $r = 0.04, \sigma_1 = 0.2, \sigma_2 = 0.3, \rho =$

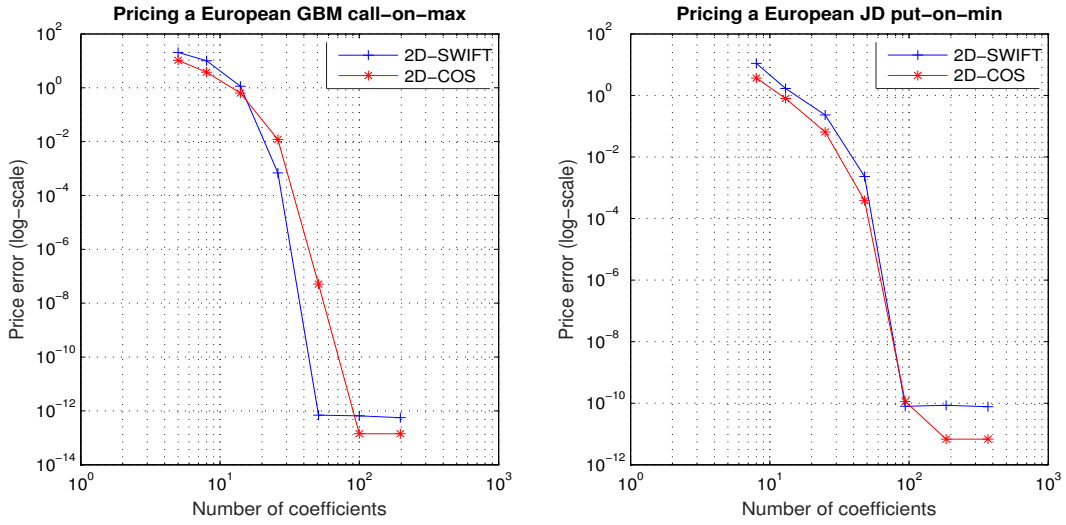


Figure 3.9: Pricing errors by means of 2D-SWIFT and 2D-COS methods under different dynamics. The left plot corresponds to a call-on-max driven by GBM dynamics, with parameters $S_0 = (40,40)$, $r = 0.048799$, $\sigma_1 = 0.2$, $\sigma_2 = 0.3$, $\rho = 0.5$, $K = 40$, $L = 10$ and $T = 7/12$. The right plot stands for a put-on-min driven by JD dynamics, with parameters $S_0 = (90,110)$, $r = 0.05$, $\sigma = (0.12,0.15)$, $\rho = 0.3$, $K = 100$, $\lambda = 0.6$, $\mu_J = (-0.1,0.1)$, $\sigma_J = (0.17,0.13)$, $\rho_J = 0.2$, $L = 10$ and $T = 1$.

$0.25, T = 1, K_1 = 90, K_2 = 110, L = 10$. The parameters for the JD dynamics are the same as for GBM with the jump component $\lambda = 0.6, \mu_J = (-0.1,0.1), \sigma_J = (0.17,0.13)$ and correlation $\rho_J = -0.2$. The reference value for the GBM case is computed by means of Monte Carlo simulation with one million paths and the 95% confidence interval is given. Regarding the JD dynamics, the reference price is given by 2D-SWIFT with scale of approximation $m = 10$. We select the scale of approximation m by means of formula (3.130) where we set $\epsilon_2 = 1.0e-04$.

Dynamics	m	Relative error
GBM	3	$1.9e-03$
JD	3	$8.5e-07$

Table 3.4: Reference price for GBM: 204.2355, 95% confidence interval: [203.0214, 205.4495], price given by 2D-SWIFT: 204.6172. Reference price for JD: 212.9888744552966.

With this example we can see how correlation options behave in the presence of jumps. We study the dependence on the parameter λ, i.e. the mean arrival rate in the Merton's jump-diffusion model. We see in Figure 3.10 the evolution of the option value according to this parameter. The rest of parameters of the model is the same as before. We can clearly observe the increasing value of the option when λ is increased.

3.3.5. STRENGTHS OF 2D-SWIFT

For the cases shown in the numerical experiments section, the convergence is similar for 2D-SWIFT and 2D-COS. The difference is that 2D-SWIFT is computationally a bit slower than

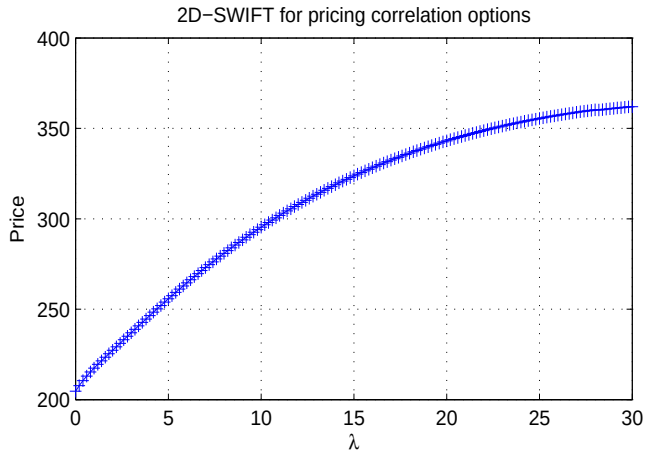


Figure 3.10: Correlation option prices with respect to the parameter λ under JD dynamics.

2D-COS, although it is still very competitive. It is worth mentioning that 2D-SWIFT could be speeded up by running in different threads (i.e. in parallel) the density and payoff coefficients. In this section we present the strengths of 2D-SWIFT with respect to 2D-COS, so that we can appreciate the advantages of using the new method presented in this work.

INTEGRATION RANGE AND SCALE OF APPROXIMATION

The 2D-COS method has a strong dependence on the integration range based on the cumulants in expression (3.131). On the contrary, as explained in Section 3.3.3, the 2D-SWIFT method considers the integration range as an initial guess, being able to adapt it if necessary.

Once the integration range has been calculated, the accuracy of the 2D-COS method depends on the number of coefficients used in the approximation of the density. If the number of terms in the expansion is not properly chosen, the 2D-COS method does not perform well. The larger the interval the more terms we should consider, although it is not a-priori clear how many coefficients should be used. In regard to the 2D-SWIFT method, the scale of approximation m is a-priori fixed with the help of formula (3.130) and the number of coefficients is determined automatically. In Figure 3.11 we have fixed the number of terms employed for the 2D-COS approximation and m for 2D-SWIFT, and we have changed the size of the integration range by modifying the parameter L in both methods. As we can observe, the approximation deteriorates for the 2D-COS while it remains very accurate for the 2D-SWIFT method, showing that 2D-SWIFT is not sensitive with respect to this parameter.

BEHAVIOUR FOR EXTREME MATURITIES

Small maturities

Small maturity options are important in high-frequency trading, also short-term binary options are well-known in the markets. The density function for small maturities is highly peaked. Thus, the characteristic function is very smooth with fat tails as we see in Figure 3.12.

In these situations, the length of the interval goes to zero, and the scale m increases when T tends to 0, because of the shape of the density. It can be seen numerically that the interval length goes to zero quicker than the scale goes to infinity. Thus, the number of coefficients

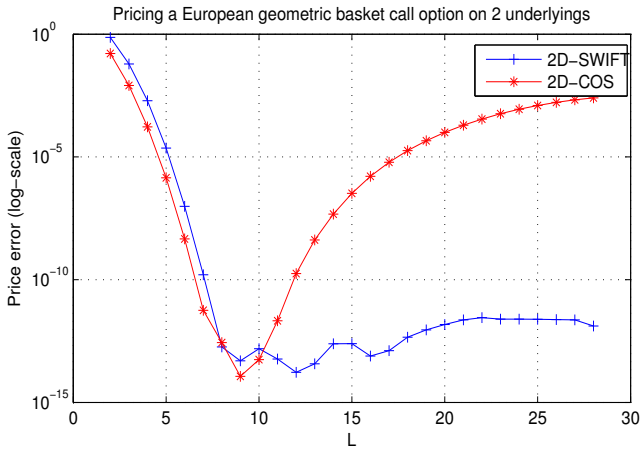


Figure 3.11: Absolute errors corresponding to the 2D-COS (red) and 2D-SWIFT (blue).

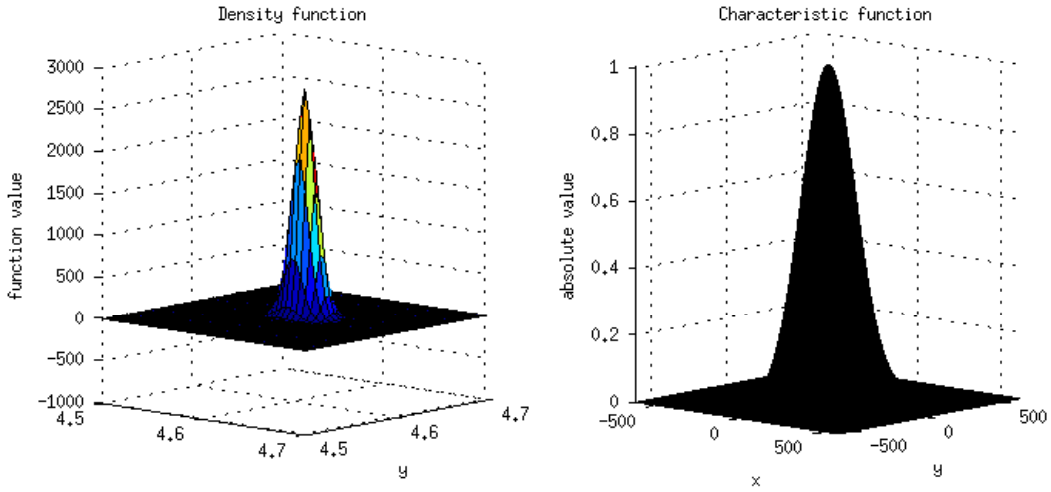


Figure 3.12: Density (left) and characteristic function (right) for GBM dynamics when $T = 0.001$.

needed by the 2D-SWIFT method tends to be very small, providing advantage in the use of the method for high dimensions. For example, when pricing with 2D-SWIFT a geometric call option on two assets driven by GBM dynamics with parameters $S_0 = (100, 100)$, $r = 0.1$, $\sigma_1 = 0.2$, $\sigma_2 = 0.3$, $\rho = 0.2$, $T = 0.001$ and $K = 100$, using a scale of $m = 6$ with just 6 coefficients we obtain an error of 3.32e−03.

Long maturities

Long maturity options are present in insurance markets. For example in [61], insurance contracts of a call spread up to 50 years are considered. Moreover, in recent years the long-dated FX option's market has grown considerably. Currently, some traded and liquid long-dated FX hybrid products are Power-Reverse Dual-Currency swaps (PRDC) as well as vanilla or exotic long-dated products such as barrier options. In the case of large maturities, the density func-

tion has fat tails and the respective characteristic function is very peaked as shown in Figure 3.13.

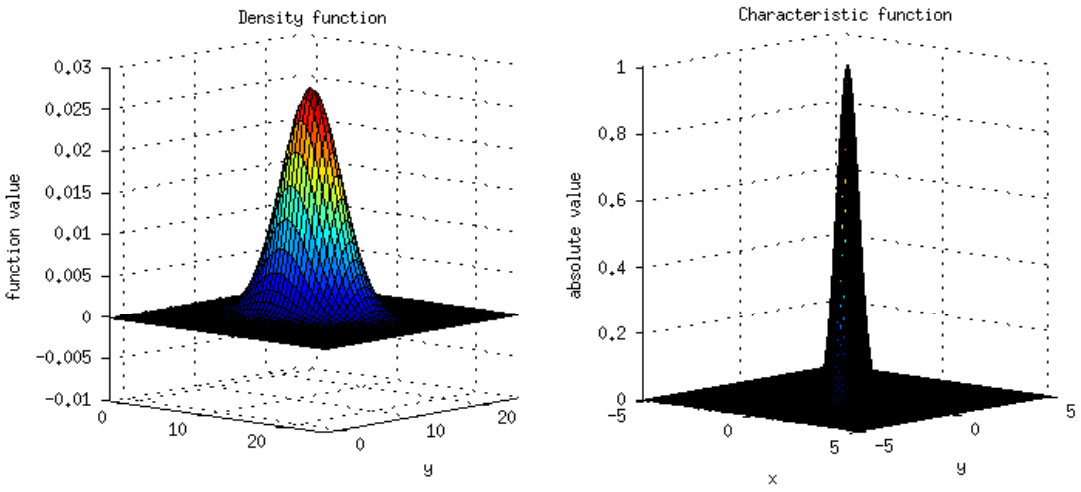


Figure 3.13: Density (left) and characteristic function (right) for GBM dynamics when $T = 100$.

Thus, the interval length increases and the scale value decreases when T takes large values. As we show in Figure 3.14, the choice of L when working with large maturities is again a problem. If L is set around 10 as suggested in [74] for moderate maturities, then the results are not accurate.

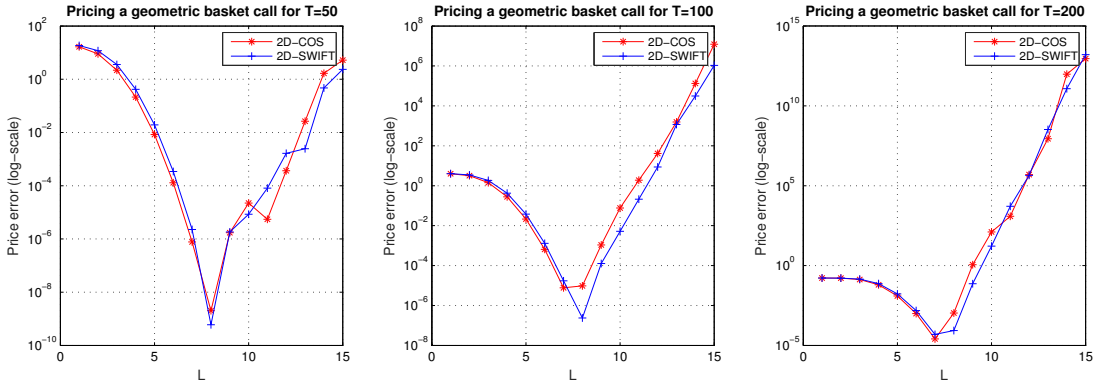


Figure 3.14: Absolute errors when pricing by means of the 2D-COS (with 200 terms) and 2D-SWIFT methods (at scale $m = 1$) a European geometric basket call option on two underlyings with parameters $S_0 = (100, 100)$, $r = 0.5$, $\sigma = (0.4, 0.4)$, $\rho = 0.2$, $K = 90$. The reference price is 20.189651798215621.

As studied in [65], when dealing with long maturities roundoff errors appear for payoff options that are unbounded. We see in Figure 3.15 the behaviour of the payoff functions that we consider and the domains where they can grow rapidly. The payoff functions for put options are bounded. However, for calls and spreads roundoff errors may appear. Due to the local nature of Shannon wavelets, we can remove part of the sum in the final pricing formula

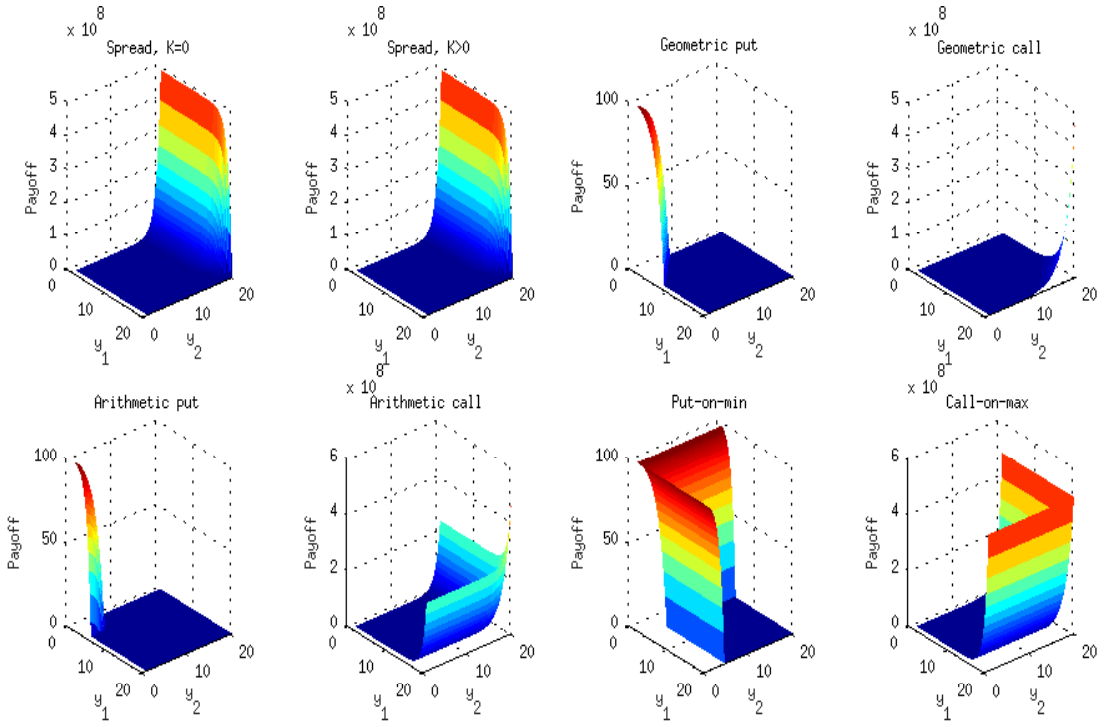


Figure 3.15: Some two-dimensional payoff functions.

(3.78) to avoid roundoff errors by eliminating some coefficients. We note that this is possible because each coefficient is only relevant for a very small interval within the integration range, while in the case of the 2D-COS method, all coefficients intervene in the approximation along the whole integration range. Table 3.5 confirms the quality of the 2D-SWIFT method. We use 50 (respectively 100) terms in the 2D-COS expansion, and the same number of coefficients for 2D-SWIFT, corresponding to the scale $m = 0$ (respectively $m = 1$). Although a direct implementation of both methods gives inaccurate results, 2D-SWIFT performs much better when we remove some terms at the boundaries.

Method	Error ($m = 0$)	Range	Error ($m = 1$)	Range
2D-COS	1.18e+01	–	4.48e−07	–
2D-SWIFT	1.59e+01	$-14 \le k_1, k_2 \le 37$	5.99e−07	$-28 \le k_1, k_2 \le 73$
2D-SWIFT (terms removed)	5.40e−02	$-14 \le k_1, k_2 \le 24$	4.60e−07	$-28 \le k_1, k_2 \le 65$

Table 3.5: Absolute errors when pricing a geometric basket call option with parameters $S_0 = (100, 100)$, $r = 0.1$, $\sigma = (0.25, 0.25)$, $\rho = 0.8$, $K = 120$, $T = 100$, $L = 10$. Reference price is 73.156120362425582.

3.4. CONCLUSIONS

We started the chapter with a brief introduction of the mathematical concepts behind option pricing. Then we introduced wavelets based Fourier inversion methods for option pricing and

carried out a numerical study of the SWIFT and WA methods. It has been shown that SWIFT method outperforms WA in accuracy, whereas they are similar in CPU time.

While those two methods can be found in the literature, we decided to try to make use of Chebyshev wavelets, which come from Chebyshev polynomials to invert a characteristic function in an analogous way as it is done with SWIFT and Shannon wavelets. To the best of our knowledge no previous attempts to use this wavelets in option pricing have been done. In the previous chapter we presented a technique for the computation of the density coefficients making use of the Bessel function. In this one, we attempted to price European put options. We were expecting to find the payoff coefficients in a simple form, but this has not been the case. Since we were not satisfied with the complexity behind approximating both density and payoff coefficients using these wavelets, we decided to abandon this approach.

After that, we have presented the two-dimensional SWIFT method which is motivated by the problem of multidimensional option pricing of European rainbow options.

First, we presented the multiresolution analysis framework in two and higher dimensions for separable spaces as well as the Shannon wavelets representation. We also presented pricing formulas for European options in two and higher dimensions using a more convenient approximation than the one previously followed in the literature for the cardinal sine function. We give a complete error analysis of the new method and provide a prescription on how to select the parameters appearing in the method according to the precision required.

We tested with a wide variety of numerical examples the efficiency of the method in the two-dimensional case for different kinds of European rainbow options of assets driven by different dynamics like GBM or JD. Basket options, spread options, options on the minimum or the maximum of two risky assets and correlation options were considered. We compared 2D-SWIFT results with the state-of-the-art 2D-COS method, with closed-form solutions when available, with Monte Carlo simulation or with 2D-SWIFT with a large scale of approximation. Finally, we presented the strengths of the 2D-SWIFT machinery, which includes the domain truncation issue, the calculation of the scale of approximation and the number of coefficients used as well as the results when dealing with extreme maturities. As it has been mentioned, extending the method to high dimensions, more than 4 or 5 depending on the specific kind of product, will not be useful because the curse of dimensionality appears.

We have shown that 2D-SWIFT inherits the strengths of the one-dimensional technique presented in [67] for European-style options, like for instance, the a-priori knowledge of the approximation scale.

CHAPTER 4

Computation of Market Risk Measures with Stochastic Liquidity Horizon

An investor or institution involved in a certain financial market may experience losses due to factors related to its performance. The quantitative study of those possible losses is known as market risk measurement.

Much of the regulation originated from the Basel Committee of Banking Supervision. This committee has recently set out the revised standards for minimum capital requirements for market risk. The Committee has focused, among other things, on the two key areas of moving from Value-at-Risk (VaR) to Expected Shortfall (ES) and considering a comprehensive incorporation of the risk of market illiquidity by extending the risk measurement horizon. In this chapter we present the application of the SWIFT method to compute the VaR and ES of a given portfolio within a stochastic holding period framework. Two approaches are considered, the delta-gamma approximation, for modelling the change in value of the portfolio as a quadratic approximation of the change in value of the risk factors, and some of the state-of-the-art stochastic processes for driving the dynamics of the log-value change of the portfolio like the Merton jump-diffusion model and the Kou model.

This chapter is based on the article [20].

4.1. INTRODUCTION

Financial risk refers to the possibility of suffering financial losses (or gains) because of unpredictable changes in the underlying risk factors. In this chapter we are concerned with the measurement of one particular form of financial risk known as market risk, or the risk of loss arising from unexpected changes in market prices such as security prices or market rates. Within market risk, we find different categories such as interest-rate risk, equity risk, exchange rate risk, commodity price risk amongst others. This classification is based on the different underlying risk factors which can be for example the interest rate or a stock price.

Market risk is one of the three main categories of financial risk, the other two being credit risk which concerns with the risk of loss arising from the failure of a counterpart to make a promised payment which we will cover in the next chapter; and operational risk which deal with the risk of loss arising from the failures of internal systems or from people who operate in them. The three risk categories boundaries are not always clearly defined and they are not a complete list of all the possible risks a financial institution is exposed to. Some risk concepts are present in most risk categories, like liquidity and model risk. Liquidity risk can be roughly defined as the risk that an investment may not be bought or sold quickly enough to prevent or minimize the loss because of the absence of a buyer/seller on the market to perform this

operation. On the other hand, model risk refers to the risk caused by the use of a misspecified or inappropriate model for measuring the risk. Since the pioneering work of Harry Markowitz in the 1950s [53] the field of risk management has experienced great developments, to the extend that it is now considered a distinct sub-field of the theory of finance.

A bank regulator is a financial person in charge of measuring the risk exposure of a financial institution and determines the amount of capital that an institution has to hold in order to cover against unexpected losses. The Basel Committee on Banking Supervision (BCBS) is a committee of the world's bank regulators that meets regularly in Basel, Switzerland. In 1988 it published what is now known as Basel I. This is an agreement between the regulators on how the capital a bank is required to hold for credit risk should be calculated. Later the Basel Committee published the 1996 Amendment which was implemented in 1998 and required banks to hold capital for market risk as well as credit risk. That Amendment calculates capital for the trading book using the Value-at-Risk measure with a horizon of 10 days and a confidence level of 0.99%. Basel I has been followed by Basel II, Basel II.5, and Basel III.

The BCBS states in the consultative documents [6, 7] that *"the financial crisis exposed material weaknesses in the overall design of the framework for capitalising trading activities. The level of capital required against trading book exposure proved insufficient to absorb losses"*. In this document, the Basel Committee initiated a comprehensive review of the trading book regime, starting with an assessment of what went wrong. They established a revised standard for minimum capital requirements for market risk in [8].

The Committee has focused, amongst other things, on two key areas of moving from VaR to ES and considering a comprehensive incorporation of the risk of market illiquidity. In regard to the first issue, a number of weaknesses have been identified with using VaR for determining regulatory capital requirements, including its inability to capture the risk in the tail. For this reason, the Committee has considered alternative risk metrics like, in particular, the ES, which measures the riskiness of a position by considering both the size and the likelihood of losses above a certain confidence level. The second issue relies on the importance of incorporating the risk of market illiquidity as a key consideration in banks' regulatory capital requirements for trading portfolios. The assumption that trading book risk positions were liquid, i.e., that banks could exit or hedge these positions over a ten-day horizon proved to be false during the recent crisis. As liquidity conditions deteriorated during the crisis, banks were forced to hold risk positions for much longer than originally expected and incurred large losses due to fluctuations in liquidity premia and associated changes in market prices.

In its deliberations on revising the prudential regime for trading activities, the Committee has drawn lessons both from the academic literature (see [5]) and banks' current and emerging risk management practices. One of the important messages from the academic literature on risk measurement in the trading book is that there are limitations of VaR models that rely on the use of continuous stochastic processes with only deterministic volatility assumptions. Introducing either stochastic volatility assumptions or stochastic jump process into modelling of risk factors will help to overcome these shortcomings. Another message of paramount importance is that the time it takes to liquidate a risk position varies, depending on its transactions costs, the size of the risk position in the market, the trade execution strategy and market conditions. Some studies suggest that, for some portfolios, this aspect of liquidity risk could also be addressed by extending the VaR risk measurement horizon. These findings are in accordance with those derived from a survey of industry practices in risk management

for the trading book carried out by the Committee. As for the length of the holding period, that poll reveals that for day-to-day risk management the use of one-day VaR is universal among the banks surveyed. However, for internal capital adequacy and strategic risk management, banks are generally moving beyond short-horizon models (e.g. one-day and 10-day VaR). It is now acknowledged that, to determine the level of capital necessary to remain in business after sustaining a large loss, risk must be assessed over a longer period. Shorter horizons do not address the liquidity risk for all exposures and do not capture tail events that are important for capital adequacy. Further, almost all banks' VaR models capture non-linearities at a local level (i.e. small price changes) for much of their market risk exposure, but many banks' VaR models fail to capture non-linearity at a global level (i.e., large price changes). A common weakness in the capture of non-linearity is the use of scaling of one-day VaR to estimate exposures at longer holding periods. Such scaling only captures local non-linearity in the range of one-day price changes and can underestimate non-linear exposure over longer horizons. The Committee has agreed that the differentiation of market liquidity across the trading book will be based on the concept of *liquidity horizons*[1]. It proposes that banks' trading book exposures be assigned to a small number of liquidity horizon categories ranging from ten days to one year. The shortest liquidity horizon (most liquid exposures) is in line with the current 10-day VaR treatment in the trading book. The longest liquidity horizon (least liquid exposures) matches the banking book horizon at one year.

The estimation of the ES for several trading desks and taking into account different liquidity horizons is computationally very involved. In this chapter we present efficient and robust numerical techniques to address the aforementioned challenges. We compute the VaR and ES risk measures of a market portfolio and we assume that the holding period follows a certain positive stochastic process to account for liquidity risk. We will therefore measure the risk in the situation where the holding period is the liquidity horizon, and we will use these two terms interchangeably throughout the thesis. While the regulatory capital calculation is based on a series of increasing deterministic liquidity horizons for different assets in the trading book, our approach does not distinguish between asset classes and is therefore more suitable for an internal risk management assessment. To our knowledge, this idea was first introduced in [14] as a proposal to open a research effort in stochastic holding period models for risk measures. In that paper, the authors assume that the log-return on the portfolio value is normally distributed, which facilitates the calculation of the risk measures. Within this work, we go a step further by considering more realistic models for the log-value of the portfolio. On the one hand, we propose the use of the delta-gamma approach [60], where it is assumed that the change in portfolio value is a quadratic function of the changes in the risk factors. On the other hand, we consider the Merton jump diffusion (MJD) model [58] and the Kou model [48] to drive the log-return on the portfolio value. Under any of these scenarios, the closed formulae to compute the risk measures within the Gaussian setting in [14] are not available anymore. However, the characteristic function of the change in (log-)value of the portfolio is known in closed form for most of the interesting processes in finance, in particular for the two models mentioned above. We therefore recover the density function from its Fourier transform and then we calculate the VaR and the ES values. Among the methods available in the literature for Fourier inversion, we choose the SWIFT method (see Section 2.6.2 of

[1] The definition of liquidity horizon given in [8] is: *"the time required to exit or hedge a risk position without materially affecting market prices in stressed market conditions."*

Chapter 2) originally developed in [67] for option pricing, where the density function is approximated by a finite combination of Shannon wavelets. A Haar wavelets-based procedure as well as a cosine series expansion have been previously used in the literature, in [66], to recover the density function within the delta-gamma approach with the multivariate Gaussian model for the individual risk factors. The most important feature of the present method is that the scale of approximation is estimated a priori by means of the characteristic function, and this makes this method a real applicable one in practice, as opposed to the aforementioned numerical methods based on trial and error.

The layout of the chapter is as follows. We define the basics in risk measurement such as the risk measures and introduce the concept of stochastic liquidity horizon (SLH) in Section 4.2. Section 4.3 is devoted to explaining the methodology to calculate the VaR and ES risk measures, together with a complete error analysis as well as the way to select the parameters of the numerical method, and a wide variety of examples. Finally, brief conclusions are given in Section 4.4.

4.2. BASIC CONCEPTS IN RISK MANAGEMENT

The uncertainty about future states of the world is represented as $(\Omega, \mathscr{F}, \mathbb{P})$, a probability space on which the random variables we introduce throughout the chapter will be defined. Consider a given portfolio such as a collection of stocks or bonds, a book of derivatives, a collection of risky loans and so on. We denote the value of this portfolio at time t by $\Pi(t)$. Given a time horizon Δt, for example 1 day or 10 days, the change in value of the portfolio over the period $[t, t+\Delta t]$ is given by,

$$\Delta\Pi := \Pi(t+\Delta t) - \Pi(t). \tag{4.1}$$

Practitioners in risk management are often concerned with the so-called profit and-loss (P&L) distribution; which is the distribution of the change in value, $\Delta\Pi$. In market risk management we often work with financial models where the calendar time is measured in years, and interest rates and volatilities are quoted on an annualized basis.

As it is standard in risk-management practice, the value $\Pi(t)$ is modelled as a function of time and a d-dimensional vector of risk factors. Depending on the model, different risk factors may be used, some common ones are logarithmic prices of financial assets, yields and logarithmic exchange rates. In this work we consider two different approaches. The first one is the well-known delta-gamma approximation [60], which assumes that the change in value of the portfolio is a quadratic function of the change in value of the risk factors. The second approach consists of assuming that the value of the portfolio follows a certain stochastic process and we are therefore interested in measuring the change in the log-value of the portfolio rather than in the value itself.

4.2.1. RISK MEASURES

There exist different approaches to measuring the risk of a financial position. One of them is given by the risk measures. Most modern measures of the risk in a portfolio are statistical quantiles describing the loss distribution of the portfolio over some predetermined horizon; Value-at-Risk and Expected Shortfall are examples of them.

Definition 4.2.1. *A **risk measure** ϱ is a function mapping a distribution of losses (or gains) $\mathscr{G}$*

to $\mathbb{R}$*, that is,*

$$\varrho : \mathscr{G} \to \mathbb{R}. \tag{4.2}$$

Definition 4.2.2. *A risk measure* ϱ *is a* ***coherent risk measure*** *if it satisfies the properties of normalization, monotonicity, sub-additivity, positive homogeneity, and translational invariance, which are,*

1. *Normalization:* $\varrho(0) = 0$.
2. *Monotonicity: For* $\mathscr{L}_1, \mathscr{L}_2 \in \mathscr{G}$ *such that* $\mathscr{L}_1 \leq \mathscr{L}_2$ *almost surely we have* $\varrho(\mathscr{L}_1) \leq \varrho(\mathscr{L}_2)$.
3. *Sub-additivity: For all* $\mathscr{L}_1, \mathscr{L}_2 \in \mathscr{G}$ *we have* $\varrho(\mathscr{L}_1 + \mathscr{L}_2) \leq \varrho(\mathscr{L}_1) + \varrho(\mathscr{L}_2)$.
4. *Positive homogeneity: For all* $\mathscr{L} \in \mathscr{G}$ *and every* $\lambda > 0$ *we have* $\varrho(\lambda\mathscr{L}) = \lambda\varrho(\mathscr{L})$.
5. *Translational invariance: For all* $\mathscr{L} \in \mathscr{G}$ *and every* $l \in \mathbb{R}$ *we have* $\varrho(\mathscr{L} + l) = \varrho(\mathscr{L}) + l$.

Note that normalization means that the risk of holding no assets is zero. Monotonicity implies that positions that lead to higher losses in every state of the world require more risk capital. Sub-additivity implies that risk is reduced by diversification; the use of non-sub-additive risk measures may lead to optimal portfolios that are very concentrated and this is considered a high-risk practice in economics standards. Positive homogeneity exposes that the risk of a position is proportional to its size. And translation invariance states that by adding or subtracting a deterministic quantity to a position, we alter our capital requirements by exactly that amount. In general risk measures do not need to satisfy all the conditions stated in coherence. Indeed, the notions of sub-additivity and positive homogeneity can be replaced by the notion of convexity:

- Convexity: If $\mathscr{L}_1, \mathscr{L}_2 \in \mathscr{G}$ and $\lambda \in [0,1]$ then $\varrho(\lambda\mathscr{L}_1 + (1-\lambda)\mathscr{L}_2) \leq \lambda\varrho(\mathscr{L}_1) + (1-\lambda)\varrho(\mathscr{L}_2)$.

We now present the two risk measures we are interested in.

VALUE-AT-RISK

Value-at-Risk is probably the most widely used risk measure in institutions and has also made its way into the Basel II capital-adequacy framework. Its birth dates back to 1993, when at JPMorgan the famous Weatherstone 4.15 report asked for a one-day, one-page summary of the bank's market risk to be delivered to the CEO in the late afternoon (hence the 4.15), and that is when VaR came up.

Let $f_{\Delta\Pi}$ be the probability density function (PDF) of $\Delta\Pi$ and $F_{\Delta\Pi}$ its cumulative distribution function (CDF). Assuming we short the portfolio, we have that the right tail of $f_{\Delta\Pi}$ represents losses.

Definition 4.2.3. *Given some confidence level* $\alpha \in (0,1)$*, the VaR measures the risk of holding the portfolio during the period* Δt *is given by the smallest number* l *such that the probability that the loss* $\Delta\Pi$ *exceeds* l *is no larger than* $1-\alpha$*. Formally,*

$$VaR(\alpha) := \inf\{l \in \mathbb{R} : \mathbb{P}(\Delta\Pi > l) \leq 1-\alpha\} = \inf\{l \in \mathbb{R} : F_{\Delta\Pi}(l) \geq \alpha\}. \tag{4.3}$$

In probabilistic terms, VaR is a quantile of the loss distribution. Note that by its very definition the VaR at confidence level α does not provide any information about the severity of the losses which occur with a probability less than $1-\alpha$. Clearly, this is a drawback of VaR as a risk measure.

VaR is an attractive measure because it is easy to understand. However, it is not a coherent risk measure since it fails to satisfy the sub-additivity condition. Hence, it does not take into account the benefits of diversification. Moreover, as pointed out above it does not give indication about the severity of losses beyond the computed quantile. A measure that deals with these drawbacks is the ES. Whereas VaR answers the question "How bad can things get?", the ES answers "If things do get bad, how much can the company expect to lose?". ES was suggested by P. Artzner *et al.* (1999) in [4], where they define a coherent risk measure and show that the standard VaR is not but ES is. Proof of these facts can be found in Chapter 6 of [57].

EXPECTED SHORTFALL

Definition 4.2.4. *For a portfolio value change $\Delta\Pi$ with $\mathbb{E}[|\Delta\Pi|] < \infty$ and distribution function $F_{\Delta\Pi}$, the Expected Shortfall at confidence level $\alpha \in (0,1)$ is defined as,*

$$ES(\alpha) := \frac{1}{1-\alpha}\int_\alpha^1 q_u(F_{\Delta\Pi})\,du, \tag{4.4}$$

where $q_u(F_{\Delta\Pi})$ is the quantile function of $F_{\Delta\Pi}$.

ES is related to VaR by,

$$\mathrm{ES}(\alpha) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}(u)\,du. \tag{4.5}$$

Instead of fixing a particular level α, we average VaR over all levels $u \geq \alpha$ and thus, as pointed out in [57], we "look further into the tail" of the loss distribution. Obviously, $\mathrm{ES}(\alpha)$ depends only on the distribution of $\Delta\Pi$, and $\mathrm{ES}(\alpha) \geq \mathrm{VaR}(\alpha)$. For continuous loss distributions a more intuitive expression can be obtained that shows that ES can be interpreted as the expected loss that is incurred in the event that VaR is exceeded. For an integrable loss $\Delta\Pi$ with continuous distribution function $F_{\Delta\Pi}$ and for any $\alpha \in (0,1)$ we have,

$$\mathrm{ES}(\alpha) = \mathbb{E}[\,\Delta\Pi \mid \Delta\Pi \geq \mathrm{VaR}(\alpha)\,], \tag{4.6}$$

or, in integral form,

$$\mathrm{ES}(\alpha) = \frac{1}{1-\alpha}\int_{\mathrm{VaR}(\alpha)}^{+\infty} x f_{\Delta\Pi}(x)\,dx. \tag{4.7}$$

Example 4.2.5 (VaR and ES for normal and t-distributions)**.** *Suppose that the portfolio change distribution $F_{\Delta\Pi}$ is normal with mean μ and variance σ^2. Fix $\alpha \in (0,1)$. Then,*

$$VaR(\alpha) = \mu + \sigma\Phi^{-1}(\alpha) \qquad \text{and} \qquad ES(\alpha) = \mu + \sigma\frac{\varphi\left(\Phi^{-1(\alpha)}\right)}{1-\alpha}, \tag{4.8}$$

where φ is the density of the standard normal distribution and Φ denotes the standard normal distribution function, thus, $\Phi^{-1}(\alpha)$ is the α-quantile of Φ.

Next, suppose that $\Delta\Pi$ is such that $\frac{\Delta\Pi-\mu}{\sigma} =: \tilde{L}$ has a standard t-distribution with ν degrees of freedom. For $\nu > 2$, we get,

$$VaR(\alpha) = \mu + \sigma\, t_\nu^{-1}(\alpha) \quad \text{and} \quad ES(\alpha) = \mu + \sigma ES_{\tilde{L}}(\alpha), \tag{4.9}$$

$$for \quad ES_{\tilde{L}}(\alpha) = \frac{g_\nu(t_\nu^{-1}(\alpha))}{1-\alpha}\left(\frac{\nu + (t_\nu^{-1}(\alpha))^2}{\nu - 1}\right), \tag{4.10}$$

where t_ν denotes the distribution and g_ν the density of the standard t.

Proofs can be found in [57].

As we have seen, VaR and ES are comprised of two parameters: the time horizon (also known by holding period) and the confidence level. Typical values for α are $\alpha = 0.95$ or $\alpha = 0.99$. In market risk management the time horizon is usually 1 or 10 days. In this work, however, we go further and consider the holding period stochastic; we build upon the work in [14]: in order to take into account liquidity issues we assume that the time horizon follows a certain stochastic process $\{H(t)\}_{t\geq 0}$ where $H(t)$ is a positive random variable associated to the liquidity horizon at time $t \geq 0$. Thus, we are interested in measuring the change in value of the portfolio within the stochastic liquidity horizon framework.

4.2.2. STANDARD METHODS FOR MARKET RISKS

As pointed out before, the future increments in the value of the portfolio or loss function are random from the perspective of the present. However, we need the probability distribution of the loss function for a fixed, generally short, period of time, in order to evaluate its market risk exposure. While its exact distribution is not available, a good approximation of it is essential since a risk measure will only give us a valid result for the market risk if applied to a distribution that is close to the real distribution of the loss function. There are several methods to construct this approximation, the two most commonly used are historical simulation and model building approach.

Historical simulation: It consists of using past data to estimate what will happen in the future. Instead of estimating the loss distribution under some explicit parametric model, the historical simulation method can be thought of as estimating the distribution of the loss under the empirical distribution of data. It is easy to implement, nevertheless, the success of the approach is highly dependent on the ability to collect sufficient quantiles of relevant, synchronized data for all risk factors.

Model-building approach: This is the main alternative to historical simulation. It consists of assuming a certain relation between the loss function and its different risk factors (stocks, bonds, derivatives that are part of the portfolio). A linear model may be used to approximate this relation. For instance in a portfolio based entirely on stocks it is clear that there is a linear relation between the increments in the price of its stocks and the increment in the value of the portfolio. On the other hand, for more complex portfolios including derivatives, the relation between the risk factors and the loss function is not linear anymore. A linear model could still be used as an approximation, but a better approximation can be obtained including also a quadratic term. These are called quadratic models, and the coefficients in this approximation are usually determined by considering the quadratic model in terms of the Taylor expansion of the loss function; more details on that are given in the examples section. Once the model is set up, the loss function is a deterministic function of the risk factors, which is where the randomness of the model comes from. While in the linear case, the distribution of the model may be obtained analytically, in the quadratic case this will generally not be the case. The distribution then needs to be approximated, and Monte Carlo is the most common method to do that. Several realizations of the risk factors are done and the loss function is calculated for each of them using the explicit model. With that we obtain an empirical approximation of

the loss function. The main drawback of the model building approach is that any results that are obtained will only be as good as the model we use, and the adequacy of a model may be hard to evaluate a priori.

Since analytical expressions are generally not available, Monte Carlo simulation is what is often used to compute the risk measures, the main drawback being the computational effort. From this point of view, the situation worsens when we consider a stochastic liquidity horizon $H(t)$, since an extra source of randomness is introduced and must be simulated as well. For this reason, there is an increasing interest in looking for alternative and more efficient methods. Here we propose to use the SWIFT method introduced before in Section 2.6.2, because as it has been shown for option pricing, that the SWIFT method gives us an accurate and extremely fast recovery of the density function and we gave insight on how to select the parameters appearing in the numerical method. All these features make our proposal efficient, robust and reliable for practical implementation. Note that an efficient numerical method to compute the VaR and ES with a deterministic holding period based on Haar wavelets using WA method of Section 2.6.1 was proposed in [66]. Since densities here are also regular, we expect to have better performance with SWIFT than with WA the same way as it happened with option pricing. Let us mention that this method has been developed under the delta-gamma approach.

In our work we consider two different approaches inside the model-building setting. The first one is the well-known delta-gamma approximation [60], which assumes that the change in value of the portfolio is a quadratic function of the change in value of the risk factors. Recently, an efficient numerical method to compute the VaR and ES with a deterministic holding period was proposed (see [66] for details). Within the present context of stochastic liquidity horizon, the change in value of the portfolio under the delta-gamma approach is defined as $\Delta\Pi := \Pi(t + H(t)) - \Pi(t)$. To the best of our knowledge, this is the first time that the delta-gamma approach is considered with a stochastic holding period. The second approach consists of assuming that the value of the portfolio follows a certain stochastic process and we are therefore interested in measuring the change in the log-value of the portfolio rather than in the value itself. Then, we define $X := \log(\Pi(t + H(t))) - \log(\Pi(t))$. Let $f_{\Delta\Pi}$ (respectively f_X) be the PDF of $\Delta\Pi$ (respectively X) and $F_{\Delta\Pi}$ (respectively F_X) its CDF. If we assume that we short the portfolio, then the right tail of f_X represents losses.

Note that all definitions are exactly the same under a stochastic time horizon but Δt must be replaced by the process $H(t)$.

4.3. COMPUTATION OF MARKET RISK MEASURES BY SWIFT

4.3.1. RISK MEASURES BY SWIFT

In this section we present formulae to calculate the VaR and ES risk measures. The strategy that we follow consists of recovering the density function of the change in value (or log-value) of a certain portfolio from its Fourier transform, which is known in many situations of interest. To simplify the notation, we assume in the present section that f is the unknown density function while $\hat{f}$ is its known Fourier transform. Let us assume that f is well approximated at scale of resolution m in a finite interval $[a, b] \subset \mathbb{R}$. We define $k_1 := \lfloor 2^m a \rfloor$ and $k_2 := \lceil 2^m b \rceil$, where $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x, and $\lceil x \rceil$ denotes the smallest integer greater than or equal to x. In Section 4.3.2 we give an explanation on the selection of the scale and the interval of approximation as well as the algorithm to get the VaR and ES

values.

VALUE-AT-RISK

From the definition of VaR in (4.3) we have to compute the value $l_\alpha := \text{VaR}(\alpha)$, with $a < l_\alpha < b$ such that $I_{\alpha=0}$, where,

$$I := \int_{-\infty}^{l_\alpha} f(x)dx. \tag{4.11}$$

We truncate the infinite integration domain $(-\infty, l_\alpha]$ in (4.11) into a finite domain $[a, l_\alpha]$,

$$I_1 := \int_{a}^{l_\alpha} f(x)dx, \tag{4.12}$$

and we replace f in (4.12) by its approximation f_m^* (note that in order to make everything clearer we avoid the superscript S in the notation because it is clear everything is from the method SWIFT) in terms of Shannon scaling wavelets as in (2.101), this is,

$$I_m^* := \int_a^{l_\alpha} f_m^*(x)dx = \sum_{k=k_1}^{k_2} c_{m,k}^* \int_a^{l_\alpha} \varphi_{m,k}(x)dx = 2^{m/2} \sum_{k=k_1}^{k_2} c_{m,k}^* \int_a^{l_\alpha} \text{sinc}(2^m x - k)dx. \tag{4.13}$$

Now, if we make the change of variables $y = 2^m x - k$ and replace sinc by its approximation sinc^* in (2.105), we end up with the expression,

$$\begin{aligned} I_m^*(J) :&= \frac{1}{2^{m/2}} \frac{1}{2^{J-1}} \sum_{k=k_1}^{k_2} c_{m,k}^* \sum_{j=1}^{2^{J-1}} \int_{2^m a-k}^{2^m l_\alpha - k} \cos\left(\frac{2j-1}{2^J}\pi y\right) dy \\ &= \frac{2}{\pi} \frac{1}{2^{m/2}} \sum_{k=k_1}^{k_2} c_{m,k}^* \sum_{j=1}^{2^{J-1}} \frac{1}{2j-1} \left[\sin\left(\frac{2j-1}{2^J}\pi(2^m l_\alpha - k)\right) - \sin\left(\frac{2j-1}{2^J}\pi(2^m a - k)\right)\right]. \end{aligned} \tag{4.14}$$

Finally, we use a root-finding method to determine the value l_α such that $I_m^*(J) - \alpha = 0$ and we call it $\text{VaR}^*(\alpha)$.

Remark 4.3.1. *In the following we show an interesting relation between the area underneath f and the computation of the density coefficients $c_{m,k}^*$ in (2.108). Using this relation we present an alternative method to calculate the VaR. Although this method is extremely fast in terms of CPU time, we decided to perform all the numerical experiments with the method presented above since it is more accurate and keeps a good balance between CPU time and accuracy.*
We define,

$$I_1(m,h) := \int_{-\infty}^{h/2^m} f(x)dx, \tag{4.15}$$

where $h \in \mathbb{Z}$, and we truncate the infinite integration domain $(-\infty, h/2^m]$ into a finite domain $[k_1/2^m, h/2^m]$,

$$I_1(m,h) \approx I_2(m,h) := \int_{\frac{k_1}{2^m}}^{\frac{h}{2^m}} f(x)dx. \tag{4.16}$$

If we apply the trapezoidal rule with step $1/2^m$, then we end up with the following formula,

$$I_2(m,h) \approx S_1(m,h) := \frac{1}{2^{m+1}} \left[2 \sum_{k=k_1+1}^{h-1} f\left(\frac{k}{2^m}\right) + f\left(\frac{k_1}{2^m}\right) + f\left(\frac{h}{2^m}\right)\right]. \tag{4.17}$$

From (2.101) we have that $f\left(\frac{l}{2^m}\right) \approx 2^{m/2} c_{m,l}$*, for all* $l \in \mathbb{Z}$*, and therefore, applying this to the expression (4.17) gives us,*

$$\begin{aligned} S_1(m,h) \approx S_2(m,h) &:= \frac{1}{2^{m+1}} \left[2^{m/2+1} \sum_{k=k_1+1}^{h-1} c_{m,k} + 2^{m/2} c_{m,k_1} + 2^{m/2} c_{m,h} \right] \\ &= \frac{1}{2^{m/2}} \left[\frac{c_{m,k_1}}{2} + \sum_{k=k_1+1}^{h-1} c_{m,k} + \frac{c_{m,h}}{2} \right]. \end{aligned} \tag{4.18}$$

Finally, the coefficients $c_{m,k}$ *in (4.18) are approximated by* $c^*_{m,k}$ *in expression (2.108). Then,*

$$S_2(m,h) \approx S_3(m,h) := \frac{1}{2^{m/2}} \left[\frac{c^*_{m,k_1}}{2} + \sum_{k=k_1+1}^{h-1} c^*_{m,k} + \frac{c^*_{m,h}}{2} \right]. \tag{4.19}$$

Then the VaR can be calculated just by adding density coefficients until the value of $S_3(m,h)$ *reaches (or is approximately equal to) the confidence level* α*. We add terms until the condition* $S_3(m,h^*) \le \alpha \le S_3(m,h^*+1)$ *is satisfied for a certain* $h^* \in \mathbb{Z}$ *and we select the VaR value as the midpoint of the interval* $[\frac{h^*}{2^m}, \frac{h^*+1}{2^m}]$*, this is,*

$$\widetilde{VaR}(\alpha) := \frac{2h^*+1}{2^{m+1}}. \tag{4.20}$$

EXPECTED SHORTFALL

The ES can be determined once we obtain the VaR value as detailed in Section 4.3.1. From (4.7), we have to compute the integral,

$$\mathrm{ES}(\alpha) = \frac{1}{1-\alpha} \int_{l_\alpha}^{+\infty} x f(x)\, dx. \tag{4.21}$$

We replace l_α by the VaR computed in the previous section, and we define $l^*_\alpha := \mathrm{VaR}^*(\alpha)$. Now, we focus on the calculation of the following integral,

$$\mathrm{ES}_1(\alpha) := \frac{1}{1-\alpha} \int_{l^*_\alpha}^{+\infty} x f(x)\, dx, \tag{4.22}$$

and truncate the infinite integration domain $[l^*_\alpha, +\infty)$ into the finite domain $[l^*_\alpha, b]$, this gives us,

$$\mathrm{ES}_2(\alpha) := \frac{1}{1-\alpha} \int_{l^*_\alpha}^{b} x f(x)\, dx. \tag{4.23}$$

The last step consists of replacing the function f in (4.23) by its approximation f^*_m, and making the change of variables $y = 2^m x - k$, we obtain,

$$\begin{aligned} \mathrm{ES}^*_m(\alpha) &:= \frac{1}{1-\alpha} \int_{l^*_\alpha}^{b} x f^*_m(x)\, dx = \frac{2^{m/2}}{1-\alpha} \sum_{k=k_1}^{k_2} c^*_{m,k} \int_{l^*_\alpha}^{b} x \frac{\sin\left((2^m x - k)\pi\right)}{(2^m x - k)\pi} dx \\ &= \frac{2^{\frac{m}{2}}}{(1-\alpha)\, 2^{2m}} \sum_{k=k_1}^{k_2} c^*_{m,k} \int_{2^m l^*_\alpha - k}^{2^m b - k} (y+k) \frac{\sin(\pi y)}{\pi y} dy \\ &= \frac{1}{(1-\alpha)\, 2^{\frac{3}{2}m}} \sum_{k=k_1}^{k_2} c^*_{m,k} \left[\frac{1}{\pi} \int_{2^m l^*_\alpha - k}^{2^m b - k} \sin(\pi y)\, dy + k \int_{2^m l^*_\alpha - k}^{2^m b - k} \mathrm{sinc}(y)\, dy \right]. \end{aligned} \tag{4.24}$$

The first integral at the right hand side in (4.24) is solved analytically,

$$\int_{2^m l_\alpha^* - k}^{2^m b - k} \sin(\pi y)\, dt = \frac{1}{\pi}\left[\cos\left(\pi\left(2^m l_\alpha^* - k\right)\right) - \cos\left(\pi\left(2^m b - k\right)\right)\right], \tag{4.25}$$

while for the second integral $\mathscr{I} := \int_{2^m l_\alpha^* - k}^{2^m b - k} \operatorname{sinc}(y)\, dy$ we use the formula in (2.105) to approximate the cardinal sine function and $\mathscr{I}$ can be replaced by,

$$\begin{aligned}\mathscr{I}_1 :&= \frac{1}{2^{J-1}} \int_{2^m l_\alpha^* - k}^{2^m b - k} \sum_{j=1}^{2^{J-1}} \cos\left(\frac{2j-1}{2^J}\pi y\right) dy = \frac{1}{2^{J-1}} \sum_{j=1}^{2^{J-1}} \int_{2^m l_\alpha^* - k}^{2^m b - k} \cos\left(\frac{2j-1}{2^J}\pi y\right) dy \\ &= \frac{2}{\pi} \sum_{j=1}^{2^{J-1}} \frac{1}{(2j-1)} \left[\sin\left(\frac{2j-1}{2^J}\pi\left(2^m b - k\right)\right) - \sin\left(\frac{2j-1}{2^J}\pi\left(2^m l_\alpha^* - k\right)\right)\right].\end{aligned} \tag{4.26}$$

Finally, by (4.24), (4.25) and (4.26), the Expected Shortfall ES(α) can be calculated with the formula,

$$\begin{aligned}\mathrm{ES}^*(\alpha) :&= \frac{1}{(1-\alpha)\, 2^{\frac{3}{2}m}} \sum_{k=k_1}^{k_2} c_{m,k}^* \left[\frac{1}{\pi^2}\left(\cos\left(\pi\left(2^m l_\alpha^* - k\right)\right) - \cos\left(\pi\left(2^m b - k\right)\right)\right)\right. \\ &\left. + \frac{2}{\pi} k \sum_{j=1}^{2^{J-1}} \frac{1}{(2j-1)} \left(\sin\left(\frac{2j-1}{2^J}\pi\left(2^m b - k\right)\right) - \sin\left(\frac{2j-1}{2^J}\pi\left(2^m l_\alpha^* - k\right)\right)\right)\right].\end{aligned} \tag{4.27}$$

Remark 4.3.2. *Note that we can speed up the evaluation of (4.14) and (4.27) by means of a discrete sine transform.*

4.3.2. ERROR ANALYSIS AND SELECTION OF PARAMETERS

In this section we perform an error analysis on the SWIFT method when it is used to calculate the risk measures, and explain how to determine the value of the parameters that play a role in the numerical method.

ERROR ESTIMATION IN THE COMPUTATION OF VAR

Let us define $\mathscr{E} := |I - I_m^*(J)|$, $\mathscr{E}_1 := |I - I_1|$, $\mathscr{E}_2 := |I_1 - I_m^*|$ and $\mathscr{E}_3 := |I_m^* - I_m^*(J)|$. Then, the overall error when approximating I in (4.11) by $I_m^*(J)$ in (4.14) can be bounded by,

$$\mathscr{E} \leq \mathscr{E}_1 + \mathscr{E}_2 + \mathscr{E}_3. \tag{4.28}$$

From (4.11) and (4.12) we have,

$$\mathscr{E}_1 \leq \int_{-\infty}^{a} f(x)dx. \tag{4.29}$$

We can make this integral arbitrarily small by selecting a appropriately, since f is a density function.

We define the projection error, denoted by ϵ_p, as,

$$\epsilon_p := \left|f(x) - \mathscr{P}_m f(x)\right| = \left|f(x) - \sum_{k\in\mathbb{Z}} c_{m,k}\varphi_{m,k}(x)\right|. \tag{4.30}$$

We also define the truncation error, denoted by ϵ_t, as,

$$\epsilon_t := |\mathscr{P}_m f(x) - f_m(x)| = |\sum_{k \notin \{k_1,\dots,k_2\}} c_{m,k}\varphi_{m,k}(x)| \,. \tag{4.31}$$

We denote by ϵ_c the error arising from using approximated coefficients $c^*_{m,k}$ instead of the exact ones $c_{m,k}$. We have,

$$\epsilon_c := |f_m(x) - f^*_m(x)| = |\sum_{k=k_1}^{k_2} (c_{m,k} - c^*_{m,k})\varphi_{m,k}(x)|. \tag{4.32}$$

Then, we have,

$$|f(x) - f^*_m(x)| \le \epsilon_p + \epsilon_t + \epsilon_c \,, \tag{4.33}$$

and,

$$\mathscr{E}_2 \le \int_a^{l_\alpha} |f(x) - f^*_m(x)| dx \le (l_\alpha - a)(\epsilon_p + \epsilon_t + \epsilon_c) \le (b-a)(\epsilon_p + \epsilon_t + \epsilon_c). \tag{4.34}$$

First, we consider the projection error ϵ_p. The projection $\mathscr{P}_m f$ can be written as [51],

$$\mathscr{P}_m f(x) = \frac{1}{2\pi}\int_{-2^m\pi}^{2^m\pi} \hat{f}(\xi) e^{i\xi x} \mathrm{d}\xi \,. \tag{4.35}$$

By definition of the inverse Fourier transform of f, we have,

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(\xi) e^{i\xi x} \mathrm{d}\xi \,. \tag{4.36}$$

Let,

$$K(v) = \frac{1}{2\pi}\int_{|\xi|>v} |\hat{f}(\xi)| \mathrm{d}\xi \,, \tag{4.37}$$

then,

$$\epsilon_p \le K(2^m\pi) \,. \tag{4.38}$$

Next, we consider the truncation error ϵ_t. We observe that,

$$\epsilon_t = |\mathscr{P}_m f(x) - f_m(x)| \le 2^{m/2} \sum_{k \notin \{k_1,\dots,k_2\}} |c_{m,k}| \,. \tag{4.39}$$

since $|\varphi_{m,k}(x)| \le 2^{m/2}$. The following theorem allows us to give an estimate of the size of the coefficients $c_{m,k}$ in terms of the rate of decay of the density function f.

Theorem 4.3.3 (Theorem 1.3.2 of [79])**.** *Let f be defined on $\mathbb{R}$, and let $\hat{f}$ be its Fourier transform such that for some positive constant d,*

$$|\hat{f}(y)| = \mathscr{O}\left(e^{-d|y|}\right), \quad |y| \to \pm\infty. \tag{4.40}$$

Then, as $h \to 0$,

$$\frac{1}{h}\int_{\mathbb{R}} f(t)S(j,h)(t)dt - f(jh) = \mathscr{O}\left(e^{-\frac{\pi d}{h}}\right), \tag{4.41}$$

where $S(j,h)(t) := sinc\left(\frac{x}{h} - j\right)$.

If we consider $h = \frac{1}{2^m}$, then by Theorem 4.3.3, the terms $|c_{m,k}|$ can be well approximated by $\frac{1}{2^{m/2}} f\left(\frac{k}{2^m}\right)$ provided that $|\hat{f}|$ decays like in (4.40). As pointed out in [51], this rate of decay is typically encountered in most of the interesting processes in finance, like for instance the GBM, MJD and Kou models, to name just a few. Then, we can assume a certain rate of decay for f to conclude that the series in (4.39) is a convergent series of terms which decrease very fast in value when k goes to minus and plus infinity.

Finally, we consider ϵ_c. The coefficients $c_{m,k}$ are to be calculated by means of Vieta's formula and the numerical error can be estimated as,

$$\epsilon_c \leq \sum_{k=k_1}^{k_2} |c_{m,k} - c^*_{m,k}||\varphi_{m,k}(x)| \leq 2^{m/2} \sum_{k=k_1}^{k_2} |c_{m,k} - c^*_{m,k}| \,. \tag{4.42}$$

The coefficients approximation error is studied in Theorem 1 of [67] and we stated it previously in Theorem 3.3.9. If we define $\mathscr{A} := \max(|a|,|b|)$ and assume that $H(\mathscr{A}) < \epsilon$, then we can apply Theorem 3.3.9 with $J \geq \log_2(\pi M_m)$, where $M_m := \max_{k_1<k<k_2} M_{m,k}$. Finally,

$$\epsilon_c \leq 2^{m/2} \sum_{k=k_1}^{k_2} |c_{m,k} - c^*_{m,k}| \leq 2^m (k_2 - k_1 + 1) \left(2\epsilon + \sqrt{2\mathscr{A}} \|f\|_2 \frac{(\pi M_m)^2}{2^{2(J+1)} - (\pi M_m)^2}\right). \tag{4.43}$$

Next we consider,

$$\mathscr{E}_3 := |I^*_m - I^*_m(J)| \leq 2^{m/2} \sum_{k=k_1}^{k_2} |c^*_{m,k}| \int_a^{l_\alpha} |\mathrm{sinc}(2^m x - k) - \mathrm{sinc}^*(2^m x - k)| dx. \tag{4.44}$$

If we make the change of variables $y = 2^m x - k$, then,

$$\mathscr{E}_3 \leq \frac{1}{2^{m/2}} \sum_{k=k_1}^{k_2} |c^*_{m,k}| \int_{2^m a - k}^{2^m l_\alpha - k} |\mathrm{sinc}(y) - \mathrm{sinc}^*(y)| dy. \tag{4.45}$$

We observe from (2.106) that $|c^*_{m,k}| \leq 2^{m/2}$. Further, we will use the following lemma to get an upper bound of the integral in (4.45).

Lemma 4.3.4 (Lemma 2 of [67])**.** *Define the absolute error $\mathscr{E}_V(t) := sinc(t) - sinc^*(t)$. Then,*

$$|\mathscr{E}_V(t)| \leq \frac{(\pi c)^2}{2^{2(J+1)} - (\pi c)^2}, \tag{4.46}$$

for $t \in [-c, c]$, where $c \in \mathbb{R}, c > 0$ and $J \geq \log_2(\pi c)$.

Since $-2^m \mathscr{A} - k \leq 2^m a - k \leq y \leq 2^m l_\alpha - k \leq 2^m \mathscr{A} - k$, then from (4.45) and Lemma 4.3.4 we have,

$$\int_{2^m a - k}^{2^m l_\alpha - k} |\mathrm{sinc}(y) - \mathrm{sinc}^*(y)| dy \leq \int_{-M_m}^{M_m} |\mathrm{sinc}(y) - \mathrm{sinc}^*(y)| dy \leq \frac{2\pi^2 (M_m)^3}{2^{2(J+1)} - (\pi M_m)^2}. \tag{4.47}$$

Finally,

$$\mathscr{E}_3 \leq (k_2 - k_1 + 1) \frac{2\pi^2 (M_m)^3}{2^{2(J+1)} - (\pi M_m)^2}. \tag{4.48}$$

ERROR ESTIMATION IN THE COMPUTATION OF ES

Let us define $\bar{\mathscr{E}} := |\mathrm{ES}_1(\alpha) - \mathrm{ES}^*(\alpha)|$, $\bar{\mathscr{E}}_1 := |\mathrm{ES}_1(\alpha) - \mathrm{ES}_2(\alpha)|$, $\bar{\mathscr{E}}_2 := |\mathrm{ES}_2(\alpha) - \mathrm{ES}^*_m(\alpha)|$ and $\bar{\mathscr{E}}_3 := |\mathrm{ES}^*_m(\alpha) - \mathrm{ES}^*(\alpha)|$. Then, the overall error when approximating $\mathrm{ES}_1(\alpha)$ in (4.22) by $\mathrm{ES}^*(\alpha)$ in (4.27) is bounded by,

$$\bar{\mathscr{E}} \leq \bar{\mathscr{E}}_1 + \bar{\mathscr{E}}_2 + \bar{\mathscr{E}}_3. \tag{4.49}$$

The error $\bar{\mathscr{E}}_1$ can be bounded following an analogous argument as in Section 4.3.2. Let us study $\bar{\mathscr{E}}_2$ in detail,

$$\bar{\mathscr{E}}_2 = |\mathrm{ES}_2(\alpha) - \mathrm{ES}^*_m(\alpha)| = \left| \frac{1}{1-\alpha} \int_{l^*_\alpha}^{b} x \left(f(x) - f^*_m(x) \right) dx \right| \leq \frac{1}{1-\alpha} \int_{l^*_\alpha}^{b} x \left| f(x) - f^*_m(x) \right| dx. \tag{4.50}$$

Then, from (4.33) we have,

$$\bar{\mathscr{E}}_2 \leq \left[\frac{1}{1-\alpha} \int_{l^*_\alpha}^{b} x\, dx \right] \left(\epsilon_p + \epsilon_t + \epsilon_c \right) = \frac{b^2 - (l^*_\alpha)^2}{2(1-\alpha)} \left(\epsilon_p + \epsilon_t + \epsilon_c \right). \tag{4.51}$$

Regarding $\bar{\mathscr{E}}_3$, we observe that the source of error is the replacement of sinc by sinc* in $\mathscr{I} := \int_{2^m l^*_\alpha - k}^{2^m b - k} \mathrm{sinc}(y) dy$. Thus, we can consider a similar argument as in the last part of Section 4.3.2 to get the same bound given in (4.47) for the error $|\mathscr{I} - \mathscr{I}_1|$.

CHOICE OF m, J AND THE TRUNCATION INTERVAL $[a, b]$

Looking at expressions (4.14) and (4.27) for computing the VaR and ES values respectively, we can observe that the parameters m and J must be selected before we carry out the approximation. As we have shown in Section 4.3.2, the projection error (4.30) is bounded by,

$$\epsilon_p \leq K(2^m \pi), \tag{4.52}$$

where,

$$K(v) = \frac{1}{2\pi} \int_{|\xi| > v} |\hat{f}(\xi)| \mathrm{d}\xi. \tag{4.53}$$

In our setting, the characteristic function $\hat{f}$ is known in closed form and we can therefore calculate the value of m that makes the projection error smaller than a certain tolerance ϵ_m. In general, the integral in (4.53) cannot be solved analytically and we compute the value of m that satisfies,

$$\frac{1}{2\pi} \left(|\hat{f}(-2^m \pi)| + |\hat{f}(2^m \pi)| \right) \leq \epsilon_m. \tag{4.54}$$

Note that a more accurate method can be used to compute the integral in (4.53) based on numerical integration. Moreover, we can make a more conservative selection of the scale m, when computing the ES, by considering the error amplifying factor $1/(1-\alpha)$ in (4.51), this is,

$$\frac{1}{1-\alpha} \cdot \frac{1}{2\pi} \left(|\hat{f}(-2^m \pi)| + |\hat{f}(2^m \pi)| \right) \leq \epsilon_m. \tag{4.55}$$

By selecting m as in (4.55), we typically get a higher scale leading to more accurate results at the cost of extra computational time. We therefore use (4.54) to compute only the VaR and we use (4.55) when we desire both, the VaR and the ES values (since the computation of ES implicitly involves the computation of VaR).

Once the scale of approximation m is known, we provide a strategy to determine the interval $[a,b]$. At the beginning of Section 4.3.1 we assumed that f is well approximated at scale of resolution m in a finite interval $[a,b] \subset \mathbb{R}$, and then we defined $k_1 := \lfloor 2^m a \rfloor$ and $k_2 := \lceil 2^m b \rceil$. Thus, the determination of an appropriate truncation interval is an important issue. We use the cumulants to determine an initial guess for the domain $[a,b]$.

Next we set the parameter J. Although a different J can be selected for each k, we prefer to consider a constant J, defined here by $J := \lceil \log_2(\pi M_m) \rceil$ (in accordance with the error analysis performed before), where $M_m := \max_{k_1<k<k_2} M_{m,k}$. The reason is that, in practice, the computationally most involved part in (2.108) is the evaluation of $\hat{f}$ at the grid points. Those values can be computed only once and used by the FFT algorithm, as follows. From expression (2.108),

$$c^*_{m,k} = \frac{2^{m/2}}{2^{J-1}} \sum_{j=1}^{2^{J-1}} \Re\left[\hat{f}\left(\frac{(2j-1)\pi 2^m}{2^J}\right) e^{\frac{ik\pi(2j-1)}{2^J}}\right] = \frac{2^{m/2}}{2^{J-1}} \Re\left[e^{\frac{ik\pi}{2^J}} \sum_{j=0}^{2^{J-1}-1} \hat{f}\left(\frac{(2j+1)\pi 2^m}{2^J}\right) e^{\frac{2\pi ikj}{2^J}}\right]. \tag{4.56}$$

Finally, we assume that $\hat{f}\left(\frac{(2j+1)\pi 2^m}{2^J}\right) = 0$, from 2^{J-1} to $2^J - 1$, so that the last equality in (4.56) is equivalent to,

$$c^*_{m,k} = \frac{2^{m/2}}{2^{J-1}} \Re\left[e^{\frac{ik\pi}{2^J}} \sum_{j=0}^{2^J-1} \hat{f}\left(\frac{(2j+1)\pi 2^m}{2^J}\right) e^{\frac{2\pi ikj}{2^J}}\right], \tag{4.57}$$

and therefore the FFT algorithm can be applied to compute the density coefficients $c^*_{m,k}$. The error analysis reveals that the same J is used to compute the risk measures in (4.14) and (4.27). As pointed out in Remark 4.3.2, this choice of $J = J$ allows the acceleration of the evaluation of these risk measures by means of a discrete sine transform.

As stated in Remark 4.3.1 and shown in Section 4.3.2, we have that $f\left(\frac{l}{2^m}\right) \approx 2^{m/2} c_{m,l}$, for all $l \in \mathbb{Z}$ provided that the modulus of the Fourier transform $\hat{f}$ of f decays sufficiently fast. We can therefore control the quality of the truncated interval by evaluating the density f at a and b, since $f(a) \approx 2^{m/2} c_{m,k_1}$ and $f(b) \approx 2^{m/2} c_{m,k_2}$. We summarize the overall process in Algorithm 1. It is worth remarking that with all the parameters fixed beforehand, this methodology is reliable and directly applicable in practice.

1: Select the value m such that a certain accuracy ϵ_m is reached, according to (4.54);
2: Determine $[a,b]$ by means of the cumulants;
3: Set $k_1 = \lfloor a/2^m \rfloor$ and $k_2 = \lceil b/2^m \rceil$;
4: Set $J = J$ where $J := \lceil \log_2(\pi M_m) \rceil$, $M_m := \max_{k_1<k<k_2} M_{m,k}$, $M_{m,k} := \max(|2^m \mathscr{A} - k|, |2^m \mathscr{A} + k|)$ and $\mathscr{A} := \max(|a|,|b|)$;
5: Compute the density coefficients $c^*_{m,k}$ with the inversion formula (2.108) (use FFT optionally);
6: Given a tolerance ϵ_z, use a root-finding method to determine the VaR value $\mathrm{VaR}^*(\alpha)$ such that $I^*_m(J) - \alpha = 0$, where $I^*_m(J)$ is taken from (4.14);
7: Calculate the ES value using (4.27) and the VaR value computed in the former step;

Algorithm 1: Algorithm to calculate $\mathrm{VaR}^*(\alpha)$ and $\mathrm{ES}^*(\alpha)$.

4.3.3. NUMERICAL EXAMPLES

In this section, we present a wide variety of numerical examples to illustrate the accuracy, speed and robustness of SWIFT method when it is used to compute the risk measures VaR and ES within the stochastic liquidity horizon framework. The programs in this chapter have been coded in MATLAB and run under Linux OS on a laptop with Intel Core i7-5500U 2.40 GHz processor and 7.7 GB of memory.

We divide this section into five subsections. The first two are devoted to the delta-gamma approach, where we measure the change in value of the portfolio as a quadratic function of the change in value of the risk factors. We assume first that the change in risk factors follows a Gaussian distribution and we move on to a more challenging problem by assuming that these changes are driven by the heavy-tailed t-distribution. In next subsections we consider the log-value change when the dynamics of the portfolio is driven by the Geometric Brownian motion, the Merton jump diffusion model and the Kou model. The root-finding method that we pick for all the numerical examples is the bisection algorithm with the stopping criterion $\epsilon_z = 1.0e-06$. Regarding the liquidity horizon process, we consider the type of distributions used in [14]. We select the Bernoulli distribution for the delta-gamma approach under the Gaussian model, the exponential distribution for the delta-gamma approach under the t-distribution, the exponential distribution for the GBM, the generalised Pareto distribution for the Merton jump diffusion model and the inverse gamma distribution for the Kou model. See Table 4.1 and Figure 4.1 for a complete definition of the last three distributions and the selection of parameters in each case. We run Monte Carlo simulations as a benchmark, with one million scenarios for the risk factors and one hundred scenarios for the SLH. We consider $\epsilon_m = 1.0e{-}02$. The reason for this choice of ϵ_m is that we can expect, in the deterministic case, at most two or three digits of accuracy due to the slow convergence of MC methods. The stochastic case is less encouraging with MC simulation, since there is an additional source of randomness when considering a stochastic holding period, and these two or three digits of accuracy are not guaranteed any more.

Distribution	**Parameters**	**PDF**
Exponential	λ	$f(x) = \lambda e^{-\lambda x}$
Generalised Pareto	k, σ, θ	$f(x) = \begin{cases} \left(\frac{1}{\sigma}\right)\left(1 + k\frac{x-\theta}{\sigma}\right)^{-1-\frac{1}{k}}, & \text{for } \theta \le x \text{ when } k > 0, \\ & \text{for } \theta \le x \le \theta - \frac{\sigma}{k} \text{ when } k < 0. \\ \frac{1}{\theta} e^{-\frac{x-\theta}{\sigma}}, & \text{for } \theta \le x \text{ when } k = 0. \end{cases}$
Inverse gamma	α, β	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} e^{-\frac{\beta}{x}}$

Table 4.1: Continuous distributions considered for the stochastic liquidity horizon.

DELTA-GAMMA WITH MULTIVARIATE GAUSSIAN MODEL FOR THE INDIVIDUAL RISK FACTORS

The delta-gamma method is a well-known approach used in market risk problems (see for instance [60]). It is based on the assumption that the change in portfolio value is a quadratic function of the changes in the risk factors, typically assumed normally distributed.

We therefore consider p risk factors $S(t) = (S_1(t), \cdots, S_p(t))^T$ at time t. We define $\Delta S = S(t + \Delta t) - S(t)$ as the change in value of the risk factors during the time interval $[t, t + \Delta t]$.

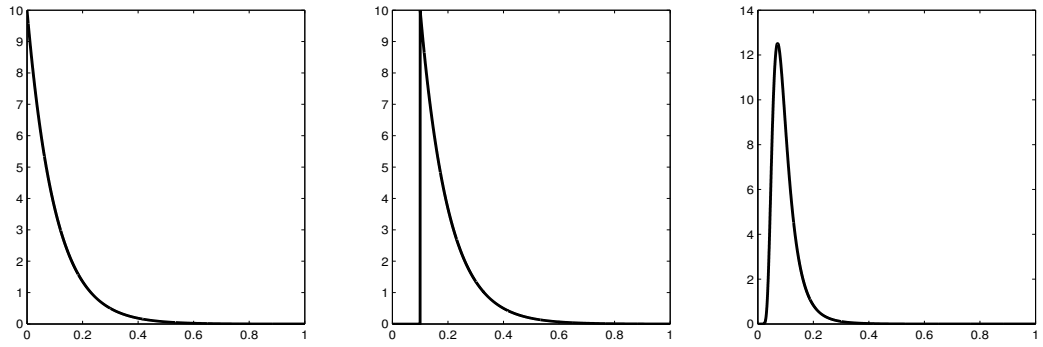


Figure 4.1: Exponential density with $\lambda = 10$ (left plot), generalised Pareto with $k = 0$, $\sigma = 0.1$, $\theta = 0.1$ (central plot) and inverse gamma with $\alpha = 6$, $\beta = 0.5$ (right plot).

Then, the change in value $\Delta\Pi := \Pi(t + H(t)) - \Pi(t)$ defined in Section 4.2 is approximated by,

$$\Delta\Pi \approx \Delta\Pi_\gamma := \Theta\Delta t + \delta^T \Delta S + \frac{1}{2}\Delta S^T \Gamma \Delta S, \tag{4.58}$$

where $\Theta = \frac{\partial\Pi}{\partial t}$, $\delta_i = \frac{\partial\Pi}{\partial S_i}$ and $\Gamma_{i,j} = \frac{\partial^2\Pi}{\partial S_i \partial S_j}$ are the Greeks evaluated at time t, and the random variable $H(t)$ is the constant Δt in the deterministic case. If we assume that ΔS follows a normal distribution, then the following proposition gives us the characteristic function of $\Delta\Pi_\gamma$.

Proposition 4.3.5 (Theorem 3.2a.2 of [55])**.** *Assume that $\Delta S \sim \mathcal{N}(0, \Sigma)$ for some positive definite matrix Σ. Let $\lambda_1, \cdots, \lambda_p$ be the eigenvalues of $\Sigma\Gamma$, and let Λ be the diagonal matrix with these eigenvalues on the diagonal. There is a matrix C satisfying $CC^T = \Sigma$ and $C^T\Gamma C = \Lambda$. Let $d = C^T\delta$. Then, the characteristic function corresponding to $f_{\Delta\Pi}$ is given by,*

$$\hat{f}_{\Delta\Pi_\gamma}(u) = \mathbb{E}\left[e^{-iu\Delta\Pi_\gamma}\right] = exp\left(-iu\Theta\Delta t - \frac{u^2}{2}\sum_{j=1}^{p}\frac{d_j^2}{1+i\lambda_j u}\right)\prod_{j=1}^{p}\left(1+i\lambda_j u\right)^{-\frac{1}{2}}, \tag{4.59}$$

where $u \in \mathbb{R}$.

Without loss of generality, we restrict ourselves to the univariate case $p = 1$, since the procedure to recover the density function from its characteristic function is the same. In that case the characteristic function reads,

$$\hat{f}_{\Delta\Pi_\gamma}(u;\Delta t) = \exp\left(-iu\Theta\Delta t - \frac{d_1^2}{2}\cdot\frac{u^2}{1+i\lambda_1 u}\right)(1+i\lambda_1 u)^{-\frac{1}{2}}. \tag{4.60}$$

Let us first perform a consistency check for the SWIFT method by taking a base portfolio from [66] and considering three different deterministic holding periods $\Delta t = 1/365$, $10/365$, $30/365$. Note that we assume 365 days per year instead of trading days per year. Our portfolio is made of one short European call and half a short European put with maturity 60 days. The underlying asset at time t is 100 with volatility level $\sigma = 0.3$, interest rate 0.1 and strike price 101 for each option. Following similar steps as in [66] we select,

$$\Sigma = \left(S(t)\sigma\sqrt{\Delta t}\right)^2,\ \Gamma = \sum_{i=1}^{n} x_i \frac{\partial^2 v_i}{\partial S^2},\ C = S(t)\sigma\sqrt{\Delta t},\ \delta = \sum_{i=1}^{n} x_i \frac{\partial v_i}{\partial S}, \tag{4.61}$$

where n represents the number of assets in the portfolio, x_i is the amount of asset i and v_i the value of asset i. Finally,

$$[a,b] = [\kappa_1 - L\sqrt{\kappa_2}, \kappa_1 + L\sqrt{\kappa_2}], \tag{4.62}$$

where κ_1 and κ_2 stand for the first and second cumulants[2] respectively,

$$\begin{aligned} \kappa_1 &= \frac{1}{2}\mathrm{tr}\,(\Gamma\Sigma) + \Theta\Delta t, \\ \kappa_2 &= \frac{1}{2}\mathrm{tr}\left((\Gamma\Sigma)^2\right) + \delta^T\Sigma\delta. \end{aligned} \tag{4.63}$$

The Greeks are computed using the Black-Scholes formula and $L = 10$ in all the numerical examples hereinafter. As stated in Section 4.3.2, the coefficients $|c_{m,k}|$ can be well approximated by $\frac{1}{2^{m/2}} f\left(\frac{k}{2^m}\right)$. We can therefore assess the suitability of the selected interval $[a,b]$ by evaluating the value of the density at the extremes of the interval, this is, $f\left(\frac{k_1}{2^m}\right) \approx 2^{m/2} c_{m,k_1}$ and $f\left(\frac{k_2}{2^m}\right) \approx 2^{m/2} c_{m,k_2}$ since these two coefficients have been already calculated. As pointed out in [66], when $p = 1$ we know the shape of the density and the number of modes (Corollary 1 of [66]). Further, from Corollary 1 and Corollary 2 of [66], we know that if $\Gamma_{1,1} > 0$ then the density function is supported on $[\mathfrak{B}, +\infty)$, with $\mathfrak{B} := -\frac{d_1^2}{2\lambda_1} + \Theta\Delta t$, while when $\Gamma_{1,1} < 0$ then the density function is supported on $(-\infty, \mathfrak{B}]$. Thus, we select the interval $[\mathfrak{B}, \kappa_1 + L\sqrt{\kappa_2}]$ in the first case and $[\kappa_1 - L\sqrt{\kappa_2}, \mathfrak{B}]$ in the second case.

We present the results of our numerical experiments in Table 4.2. The benchmark solution is calculated by means of Partial Monte Carlo (PMC) simulation, which basically means that we simulate ΔS and use the approximation in (4.58). For sake of completeness we give the results corresponding to Full Monte Carlo (FMC) simulation where we evaluate the portfolio with the Black-Scholes formula for each new value of the underlying risk factor. We show the scale of approximation calculated with the error formula (4.54). We observe that the absolute error reported is in accordance with ϵ_m. It is worth remarking that for the cases $\Delta t = 10/365$ and $\Delta t = 30/365$ the VaR computation is somewhat more challenging than in the case $\Delta t = 1/365$, due to the asymptotic behaviour of the densities at $\mathfrak{B}$ (see Figure 4.2 and the details provided in Section 4.4 and Section 5 of [66]). In these two extreme cases the bisection method does not work properly, since the condition $g(a)\cdot g(b) < 0$ with $g(l_\alpha) = I_m^*(J) - \alpha$ is not satisfied. When this situation occurs, we set $\mathrm{VaR}^*(\alpha) = \mathfrak{B}$.

Δt	PMC	FMC	SWIFT	Absolute error	m
1/365	0.9024	0.8792	0.9038	$1.4e-03$	2
10/365	1.7044	1.5430	1.7050	$5.3e-04$	6
30/365	3.0434	2.8148	3.0439	$5.0e-04$	6

Table 4.2: Absolute errors with respect to Partial Monte Carlo simulation, corresponding to the computation of VaR with deterministic holding period Δt and $\alpha = 0.99$.

Next, we study the performance of the SWIFT method within the SLH framework. We assume that $H(t)$ follows the Bernoulli distribution and we distinguish the three different

[2] Look at Theorem 3.3.2 of [55] for details.

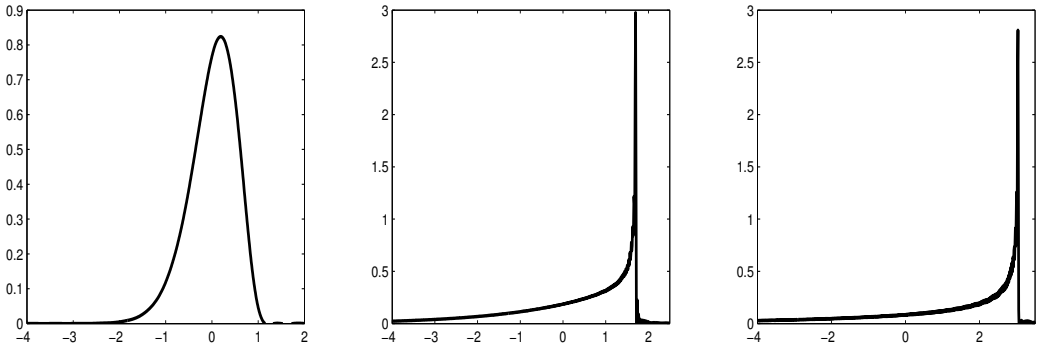


Figure 4.2: Density plots for $\Delta t = 1/365$ (left), $\Delta t = 10/365$ (central) and $\Delta t = 30/365$ (right).

cases represented in Table 4.3, where $H(t)$ takes the value $h_1 = 10/365$ with probability p and $h_2 = 30/365$ with probability $1-p$. The only different issue within the SLH framework with respect to the deterministic case is the need for computing the appropriate characteristic function to be used by the SWIFT method. Since we know the characteristic function in the deterministic case, we apply the law of iterated expectations (Proposition 2.2.8),

$$\hat{f}_{\Delta\Pi_\gamma}(u) = \mathbb{E}\left[e^{-iu\Delta\Pi_\gamma}\right] = \mathbb{E}\left[\mathbb{E}\left[e^{-iu\Delta\Pi_\gamma}|H(t)\right]\right] = p\hat{f}_{\Delta\Pi_\gamma}(u;h_1) + (1-p)\hat{f}_{\Delta\Pi_\gamma}(u;h_2). \tag{4.64}$$

The interval of approximation in this case is calculated as,

$$[a,b] = [\min\{a_{h_1}, a_{h_2}\}, \max\{b_{h_1}, b_{h_2}\}], \tag{4.65}$$

where $[a_{h_1}, b_{h_1}]$ and $[a_{h_2}, b_{h_2}]$ correspond to the intervals calculated in the deterministic case with $h_1 = 10/365$ and $h_2 = 30/365$, respectively. We observe that again, the absolute error is in accordance with ϵ_m and the scale calculated with formula (4.54) is $m = 5$ in all three cases.

$H(t)$	PMC	SWIFT	Absolute error	m
$\mathbb{P}(h_1) = 0.25, \mathbb{P}(h_2) = 0.75$	3.0430	3.0439	$8.3e-04$	5
$\mathbb{P}(h_1) = 0.5, \mathbb{P}(h_2) = 0.5$	3.0415	3.0425	$1.0e-03$	5
$\mathbb{P}(h_1) = 0.75, \mathbb{P}(h_2) = 0.25$	3.0312	3.0327	$1.5e-03$	5

Table 4.3: Absolute errors with respect to Partial Monte Carlo simulation, corresponding to the computation of VaR with stochastic holding period driven by a Bernoulli distribution and $\alpha = 0.99$.

DELTA-GAMMA APPROACH WITH MULTIVARIATE t-DISTRIBUTION FOR THE RISK FACTORS

Next we consider the delta-gamma approach presented in [36] where the underlying risk factors are heavy-tailed distributed by means of the t-distribution. We build upon the work exposed in [36] for the deterministic case and we extend that approach to the stochastic liquidity horizon framework. A t-distribution is characterized by the number of degrees of freedom ν. The tails of its density decay at a polynomial rate of $x^{-\nu}$, so the parameter ν determines the

heaviness of the tail and the number of finite moments. Let t_ν be the univariate t-distribution with ν degrees of freedom, which has density,

$$f_{t_\nu}(x) = \frac{\Gamma\left(\frac{1}{2}(\nu+1)\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{1}{2}\nu\right)}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2}, \quad -\infty < x < \infty, \tag{4.66}$$

where $\Gamma(\cdot)$ denotes the gamma function. The multivariate t-distribution has density,

$$f_{\nu,\Sigma}(x) = \frac{\Gamma\left(\frac{1}{2}(p+\nu)\right)}{(\nu\pi)^{p/2}\Gamma\left(\frac{1}{2}\nu\right)|\Sigma|^{1/2}}\left(1+\frac{1}{\nu}x^T\Sigma^{-1}x\right)^{-\frac{1}{2}(p+\nu)}, \quad x\in\mathbb{R}^p, \tag{4.67}$$

where Σ is a symmetric, positive definite matrix. If $\nu > 2$, then $\nu\Sigma/(\nu-2)$ is the covariance matrix of $f_{\nu,\Sigma}$. The multivariate $t_{\nu,\Sigma}$ density (4.67) belongs to the class of scale mixtures of normals. Thus, it has representation as the distribution of the product of a multivariate normal random vector and a univariate random variable independent of the normal. If $(X_1,\cdots,X_p)$ has density $f_{\nu,\Sigma}$, then,

$$(X_1,\cdots,X_p) \sim \frac{(\xi_1,\cdots,\xi_p)}{\sqrt{Y/\nu}}, \tag{4.68}$$

where $\xi = (\xi_1,\cdots,\xi_p) \sim \mathcal{N}(0,\Sigma)$, $Y \sim \chi^2_\nu$ (chi-square with ν degrees of freedom), and ξ and Y are independent.

If we assume within this section that ΔS in (4.58) follows the multivariate t-distribution in (4.67), then $\Delta S \sim \frac{\xi}{\sqrt{Y/\nu}}$, thanks to the ratio representation (4.68), and ΔS is therefore modelled as a scale mixture of normals. For the sake of simplicity in the exposition, we give a brief summary of the methodology developed in [36], where the authors define the loss $L = -\Delta\Pi$ and,

$$L \approx a_0 + a^T\Delta S + \Delta S^T A\Delta S \equiv a_0 + \mathcal{Q}, \tag{4.69}$$

with $a_0 = -\Theta\Delta t$, $a = -\delta$ and $A = -\frac{1}{2}\Gamma$.

By defining,

$$\mathcal{Q}_x := \left(\frac{Y}{\nu}\right)(\mathcal{Q}-x), \tag{4.70}$$

and observing that $\mathbb{P}(\mathcal{Q}\le x) = \mathbb{P}(\mathcal{Q}_x \le 0) \equiv F_x(0)$, we can compute $\mathbb{P}(\mathcal{Q}\le x)$ by finding the characteristic function of $\mathcal{Q}_x$ and then we invert it to find $\mathbb{P}(\mathcal{Q}_x \le 0)$. The following theorem gives us the characteristic function of $\mathcal{Q}_x$.

Theorem 4.3.6 (Theorem 3.1 of [36]). *Let $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_p$ be the eigenvalues of ΣA and let Λ be the diagonal matrix with these eigenvalues on the diagonal. There is a matrix C satisfying $CC^T = \Sigma$ and $C^TAC = \Lambda$. Let $b = a^TC$. Then $\mathbb{P}(\mathcal{Q}\le x) = F_x(0)$, where the distribution F_x has moment generating function,*

$$\phi_x(\theta) = \left(1+\frac{2\theta x}{\nu} - \sum_{j=1}^{p}\frac{\theta^2 b^2/\nu}{1-2\theta\lambda_j}\right)^{-\nu/2}\prod_{j=1}^{p}\frac{1}{\sqrt{1-2\theta\lambda_j}}. \tag{4.71}$$

The characteristic function of $\mathcal{Q}_x$ is given by $\hat{f}_{\mathcal{Q}_x} = \mathbb{E}[\exp(-i\omega\mathcal{Q}_x)] = \phi_x(-i\omega)$.

We can easily obtain a closed-form expression for the Fourier transform $\hat{F}_{\mathcal{Q}_x}$ of $F_{\mathcal{Q}_x}$ from $\hat{f}_{\mathcal{Q}_x}$ by integrating by parts. If we use the SWIFT method to recover $F_{\mathcal{Q}_x}$ from $\hat{F}_{\mathcal{Q}_x}$, and we take into account that,

$$F_{\Delta\Pi}(y) = 1 - F_L(-y) = 1 - F_{\mathcal{Q}}(-y - a_0) = 1 - F_{\mathcal{Q}_{(-y-a_0)}}(0), \tag{4.72}$$

then we can compute the VaR value by applying a bisection method to find l such that,

$$1 - \alpha - F_{\mathcal{Q}_{-(l+a_0)}}(0) = 0. \tag{4.73}$$

Whilst in Section 4.3.3 only one Fourier inversion was performed, it is worth remarking that, in this case, each step in the bisection method involves a Fourier inversion by means of the SWIFT method. This is clearly a challenge in terms of computation and we can tackle this problem only with a very efficient numerical method to avoid the propagation of the error.

As in the previous example section and without loss of generality, we restrict ourselves to the univariate case $p = 1$. We use first a deterministic holding period $\Delta t = 1/365$. Our portfolio is the same as before, made of one short European call and half a short European put with maturity 60 days. The underlying asset at time t is 100 with volatility level $\sigma = 0.3$, interest rate 0.1 and strike price 101 for each option. Note that the scale parameter m is recalculated at each bisection step following the tolerance error $\epsilon_m = 1.0e-02$ set at the beginning. We present the results of the numerical experiments in Table 4.4 for values of $\nu \in \{3,5,7\}$. The benchmark solution is calculated by means of PMC with one million scenarios as mentioned previously. To show the accuracy of the proposed numerical method, we have added some extra numerical experiments in Table 4.5 with ten million scenarios for PMC to be used as a benchmark for the SWIFT method with $\epsilon_m = 1.0e-03$ and $\epsilon_m = 1.0e-04$, respectively.

ν	PMC (10^6)	SWIFT	Absolute error ($\epsilon_m = 10^{-2}$)
3	0.9438	1.0201	$7.6e-02$
5	0.9562	1.0119	$5.6e-02$
7	0.9456	0.9678	$2.2e-02$

Table 4.4: Absolute errors with respect to Partial Monte Carlo simulation, corresponding to the computation of VaR with deterministic holding period $\Delta t = 1/365$, $\alpha = 0.99$ and $\nu = 3,5,7$.

ν	PMC (10^7)	Absolute error ($\epsilon_m = 10^{-3}$)	Absolute error ($\epsilon_m = 10^{-4}$)
3	0.9438	$5.2e-03$	$2.7e-04$
5	0.9548	$2.9e-03$	$1.6e-04$
7	0.9447	$3.8e-03$	$2.6e-04$

Table 4.5: Absolute errors with respect to Partial Monte Carlo simulation, corresponding to the computation of VaR with deterministic holding period $\Delta t = 1/365$, $\alpha = 0.99$ and $\nu = 3,5,7$.

Next we study the performance of the SWIFT method within the SLH framework. We consider the same portfolio as in the deterministic case and we assume that $H(t)$ follows an exponential distribution with parameter $\lambda = 10$. Since the holding period is not deterministic,

we cannot apply directly the bisection method in (4.73). To circumvent this problem, we condition on a realization of H and use the law of iterated expectations (Proposition 2.2.8),

$$\begin{aligned} F_{\mathcal{Q}_{-(l+\Theta H(t))}}(0) &= \mathbb{P}(\mathcal{Q}_{-l-\Theta H(t)} \leq 0) = \mathbb{E}\left[\mathbb{1}_{\{\mathcal{Q}_{-(l+\Theta H(t))}\leq 0\}}\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\{\mathcal{Q}_{-(l+\Theta H(t))}\leq 0\}} \,\middle|\, H(t)=h\right]\right] \\ &= \int_{\mathbb{R}} \mathbb{E}\left[\mathbb{1}_{\{\mathcal{Q}_{-(l+\Theta h)}\leq 0\}}\right] f_H(h)\,dh = \int_{\mathbb{R}} F_{\mathcal{Q}_{-(l+\Theta h)}}(0) f_H(h)\,dh, \end{aligned} \tag{4.74}$$

where $f_H(h) = 10\exp(-10h)$ is the probability density function of the stochastic holding period. Now we can apply the bisection method to find l such that,

$$1 - \alpha - F_{\mathcal{Q}_{-(l+\Theta H(t))}}(0) = 0. \tag{4.75}$$

The integral at the right hand side of (4.74) must be calculated at each iteration of the bisection method. The integral is evaluated by means of the trapezoidal rule and the infinite integration domain is replaced by the finite domain $[0, h^*]$, where h^* is such that $F_H(h^*) < 1.e-06$, being F_H the CDF of $H(t)$. Note that for each quadrature point h, we calculate $F_{\mathcal{Q}_{-(l+\Theta h)}}(0)$ as in the deterministic case. The results are shown in Table 4.6. We run one hundred times the PMC method and we consider the average as the benchmark solution, giving also a 95% confidence interval. It is worth remarking that Monte Carlo simulation is extremely demanding in terms of computing due to the heavy-tailed distribution combined with the stochastic holding period.

ν	PMC	95% CI	SWIFT	Absolute error
3	10.1776	[8.0220, 13.5035]	10.0803	$9.7e-02$
5	9.8275	[7.4039, 13.2777]	9.5490	$2.8e-01$
7	9.7135	[7.6791, 12.5971]	9.5047	$2.1e-01$

Table 4.6: Absolute errors with respect to Partial Monte Carlo simulation, corresponding to the computation of VaR with stochastic holding period driven by an exponential distribution with $\lambda = 10$ and $\alpha = 0.99$.

GEOMETRIC BROWNIAN MOTION

In this section we consider that our portfolio $\Pi(t)$ follows GBM dynamics,

$$d\Pi(t) = \mu\Pi(t)\,dt + \sigma\Pi(t)\,dW_t, \tag{4.76}$$

where, as usual, μ represents the drift and σ the volatility. As it was stated in Section 4.2, we are interested in measuring the change in the log-value of the portfolio rather than in the value itself. Thus, we consider $X = \log(\Pi(t+H(t))) - \log(\Pi(t))$, where $H(t)$ is the constant value Δt in the deterministic case. It is well known that X follows a normal distribution with mean $\left(\mu - \frac{\sigma^2}{2}\right)\Delta t$ and variance $\sigma^2\Delta t$, its characteristic function reads,

$$\hat{f}_X(u;\Delta t) = \exp\left(-i\left(\mu - \frac{\sigma^2}{2}\right)\Delta t u - \frac{\sigma^2\Delta t}{2}u^2\right). \tag{4.77}$$

The first two cumulants are $\kappa_1 = \left(\mu - \frac{\sigma^2}{2}\right)\Delta t$ and $\kappa_2 = \sigma^2 \Delta t$, and we determine the interval of approximation $[a,b]$ following the rule-of-thumb (4.62). In this particular case there are closed form solutions for VaR and ES values (see Example 4.2.5),

$$\text{VaR}(\alpha) = \left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}\Phi^{-1}(\alpha), \quad \text{ES}(\alpha) = \left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \frac{\sigma\sqrt{\Delta t}\phi(\Phi^{-1}(\alpha))}{1-\alpha}, \tag{4.78}$$

where ϕ stands for the PDF of a normal standard and Φ is its CDF, and we use them as the benchmark solution. We present the numerical experiments in Table 4.7. We consider three different confidence levels α, where $\alpha = 0.99$ is the traditional regulatory confidence level to measure the VaR while $\alpha = 0.975$ has become the new regulatory confidence level to measure the ES. Since we aim at computing both risk measures, we use the error formula (4.55) to estimate the parameter m. We see again that the ES error is in accordance with the fixed tolerance ϵ_m. We observe that the VaR error is extremely small due to the fact that formula (4.55) is equivalent to use formula (4.54) with a tolerance error of $(1-\alpha)\epsilon_m$. To show the power of approximation of SWIFT method, we note that when $\epsilon_m = 1.e-05$ and $\alpha = 0.99$ then $m = 7$ and the ES error is $5.1e-06$.

		VaR			ES		
α	m	Exact	SWIFT	Absolute Error	Exact	SWIFT	Absolute error
0.95	6	0.0430	0.0430	$1.5e-07$	0.0539	0.0594	$5.4e-03$
0.975	6	0.5012	0.0512	$2.8e-07$	0.0611	0.0720	$1.1e-02$
0.99	6	0.0608	0.0608	$1.9e-07$	0.0697	0.0969	$2.7e-02$

Table 4.7: Absolute errors for VaR and ES values with respect to the exact formula (4.78) when $\Pi(t)$ follows a GBM dynamics with $\mu = 0.1$, $\sigma = 0.5$ and $\Delta t = 1/365$.

Next we consider the GBM dynamics in combination with an exponential random variable to drive the SHL. In this case,

$$\hat{f}_X(u) = \mathbb{E}\left[e^{-iuX}\right] = \mathbb{E}\left[\mathbb{E}\left[e^{-iuX} \,\middle|\, H(t)\right]\right] = \mathbb{E}\left[\hat{f}_X(u;h)\right] = \int_{\mathbb{R}^+} \hat{f}_X(u;h) f_H(h)\, dh, \tag{4.79}$$

where $f_H(h) = \lambda \exp(-\lambda h)$. The integral in (4.79) can be solved analytically yielding,

$$\hat{f}_X(u) = \frac{-\lambda}{\lambda + i\left(\mu - \frac{\sigma^2}{2}\right)u + \frac{1}{2}\sigma^2 u^2}. \tag{4.80}$$

We use the expression (4.80) to calculate the scale of approximation m. We can also use the same expression to compute the cumulants. However, if we observe the cumulants in the deterministic case, we realize that they are increasing functions of Δt. We therefore consider the union of the two intervals corresponding to the minimum ($h_1 = 0$) and maximum holding period, where the maximum is determined by h_2 such that $F_H(h_2) < 1.e-06$, and F_H stands for the CDF of $H(t)$,

$$[a,b] = [\min\{a_{h_1}, a_{h_2}\}, \max\{b_{h_1}, b_{h_2}\}], \tag{4.81}$$

where $a_{h_1} = b_{h_1} = 0$, $a_{h_2} = \left(\mu - \frac{\sigma^2}{2}\right) h_2 - L\sigma\sqrt{h_2}$ and $b_{h_2} = \left(\mu - \frac{\sigma^2}{2}\right) h_2 + L\sigma\sqrt{h_2}$. We consider two benchmark solutions in this case. First of all, we run MC simulations like previously with the initial value of the portfolio $V_0 = 100$. The results are presented in Table 4.8, where in general the absolute error is of order $1.e-02$.

		VaR			ES		
α	m	MC	SWIFT	Absolute Error	MC	SWIFT	Absolute error
0.95	7	0.2331	0.2533	$2.0e-02$	0.3378	0.3638	$2.6e-02$
0.975	7	0.3064	0.3300	$2.4e-02$	0.4103	0.4404	$3.0e-02$
0.99	8	0.4005	0.4313	$3.1e-02$	0.5049	0.5418	$3.7e-02$

Table 4.8: Absolute errors for VaR and ES values with respect to MC simulation when $\Pi(t)$ follows a GBM dynamics with $\mu = 0.1$, $\sigma = 0.5$. The SLH is driven by an exponential distribution with $\lambda = 10$.

However, in order to really assess the accuracy of the SWIFT method, we use a second benchmark solution. We consider the numerical formulae employed in [14] for the VaR value,

$$\int_0^\infty \Phi\left(\frac{\mathrm{VaR}(\alpha) - \left(\mu - \frac{\sigma^2}{2}\right) h}{\sigma\sqrt{h}}\right) f_H(h)\,dh = \alpha, \tag{4.82}$$

and ES value,

$$\mathrm{ES}(\alpha) = \frac{1}{1-\alpha}\int_0^\infty \left(\left(\mu - \frac{\sigma^2}{2}\right) h\Phi\left(\frac{\left(\mu - \frac{\sigma^2}{2}\right) h - \mathrm{VaR}(\alpha)}{\sigma\sqrt{h}}\right) + \sigma\sqrt{h}\phi\left(\frac{\mathrm{VaR}(\alpha) - \left(\mu - \frac{\sigma^2}{2}\right) h}{\sigma\sqrt{h}}\right)\right) f_H(h)\,dh. \tag{4.83}$$

We use the MATLAB function `integral` to numerically solve the integrals in (4.82) and (4.83), and the MATLAB function `fzero` as a root-finding method in (4.82). The results are presented in Table 4.9, where we can observe very accurate results for VaR as well as for ES values.

		VaR			ES		
α	m	Reference	SWIFT	Absolute Error	Reference	SWIFT	Absolute error
0.95	7	0.2533	0.2533	$1.2e-06$	0.3639	0.3638	$6.2e-05$
0.975	7	0.3300	0.3300	$3.2e-06$	0.4405	0.4404	$1.3e-04$
0.99	8	0.4313	0.4313	$8.5e-06$	0.5418	0.5418	$5.5e-05$

Table 4.9: Absolute errors for VaR and ES values with respect to the formulae (4.82) and (4.83) when $\Pi(t)$ follows a GBM dynamics with $\mu = 0.1$, $\sigma = 0.5$. The SLH is driven by an exponential distribution with $\lambda = 10$.

MERTON JUMP-DIFFUSION MODEL

As pointed out in [23], jump-diffusion models assume that the evolution of prices is given by a diffusion process, punctuated by jumps at random intervals. Here the jumps represent rare events like crashes and large drawdowns. Such an evolution can be represented by modelling the log-price as a Lévy process with a nonzero Gaussian component and a jump part, which

is a compound Poisson process with finitely many jumps in every time interval. Examples of such models are the MJD model with Gaussian jumps [58] and the Kou model with double exponential jumps [48]. In this section we consider the MJD model for driving the dynamics of the value of the portfolio $\Pi(t)$, whilst the next section is devoted to the Kou model.

The process $\Pi(t)$ is assumed to follow the stochastic differential equation,

$$d\Pi(t) = (\mu - \lambda\kappa)\Pi(t)dt + \sigma\Pi(t)dW_t + (e^J - 1)\Pi(t)dq_t, \tag{4.84}$$

where q_t represents a Poisson process with mean arrival rate λ, J has normally distributed jumps with mean μ_J and standard deviation σ_J, $\kappa = \mathbb{E}\left[e^J - 1\right]$ and W_t is a standard Brownian motion process. The characteristic function of $X = \log(\Pi(t + H(t))) - \log(\Pi(t))$ in the deterministic case when $H(t)$ is Δt reads,

$$\hat{f}_X(u;\Delta t) = \exp\left(-i\left(\mu - \lambda\kappa - \frac{\sigma^2}{2}\right)\Delta t u - \frac{\sigma^2\Delta t}{2}u^2 + \lambda\Delta t\left(e^{-i\mu_J u - \frac{\sigma_J^2 u^2}{2}} - 1\right)\right), \tag{4.85}$$

and we use the cumulants,

$$\begin{aligned}
\kappa_1 &= \mu_J\lambda\Delta t + \left(\mu - \lambda\kappa - \frac{\sigma^2}{2}\right)\Delta t,\\
\kappa_2 &= \left(\sigma^2 + \lambda\left(\mu_J^2 + \sigma_J^2\right)\right)\Delta t,\\
\kappa_4 &= \left(\mu_J^4 + 6\sigma_J^2\mu_J^2 + 3\sigma_J^4\right)\lambda\Delta t.
\end{aligned} \tag{4.86}$$

To be more precise, the initial guess that we use for the MJD and Kou models to determine the truncation interval is,

$$[a,b] = \left[\kappa_1 - L\sqrt{\kappa_2 + \sqrt{\kappa_4}}, \kappa_1 + L\sqrt{\kappa_2 + \sqrt{\kappa_4}}\right]. \tag{4.87}$$

The benchmark solution method is the MC method with $\Pi(0) = 100$, since there are not closed form solutions in this case to compute the risk measures. We present the results in Table 4.10, where we illustrate both, the deterministic case with $\Delta t = 5/365$, and the stochastic case with the SLH driven by a generalised Pareto distribution. We chose the regulatory level $\alpha = 0.975$. The characteristic function within the stochastic case is obtained numerically by solving the integral in (4.79) by means of the trapezoidal rule. The truncation of the integration domain and the determination of the interval $[a,b]$ follows an entirely analogous process as in the GBM example. We observe that the SWIFT method performs well working at a low scale $m = 3$ in the challenging stochastic case. We measure the CPU time in seconds of the overall process, including the VaR as well as the ES computation. This measurement reveals the efficiency of the methodology capable to accurately estimate the risk measures in less than 0.1 second. We note that the stochastic case is more involved, since in this case, the characteristic function is obtained numerically.

KOU MODEL

Finally, we consider that the portfolio $\Pi(t)$ follows the dynamics of the Kou model. This model is also called the double exponential jump-diffusion model and it can reproduce the leptokurtic feature of the return distribution, which has semi-heavy (exponential) tails. Given a process $\Pi(t)$, it is modelled as,

$$d\Pi(t) = (\mu - \lambda\kappa)\Pi(t)dt + \sigma\Pi(t)dW_t + \left(e^J - 1\right)d\varsigma_t, \tag{4.88}$$

		VaR			ES			
$H(t)$	m	MC	SWIFT	Abs. Error	MC	SWIFT	Abs. error	CPU time
5/365	5	0.1171	0.1172	$8.2e-05$	0.1601	0.1555	$4.6e-03$	0.02
SLH	3	0.4726	0.4730	$3.7e-04$	0.6081	0.5557	$5.2e-02$	0.08

Table 4.10: Absolute errors for VaR and ES values with respect to MC simulation when $\Pi(t)$ follows a MJD dynamics with $\mu = 0.1$, $\sigma = 0.5$, $\lambda = 0.6$, $\mu_J = 0.1$, $\sigma_J = 0.2$. In the deterministic case $\Delta t = 5/365$. The SLH is driven by a generalised Pareto distribution with $k = 0$, $\sigma = 0.1$, $\theta = 0.1$. The confidence level is $\alpha = 0.975$. The CPU time is expressed in seconds.

where ς_t is a Poisson process with mean arrival rate λ and J has double exponentially distributed jumps with density,

$$f_J(y) = p\eta_1 \mathrm{e}^{-\eta_1 y}\mathbb{1}_{y\geq 0} + q\eta_2 \mathrm{e}^{\eta_2 y}\mathbb{1}_{y<0}, \tag{4.89}$$

with $\eta_1 > 1$, $\eta_2 > 0$ governing the decay of the tails and $p, q \geq 0, p+q = 1$, with p representing the probability of an upward jump, and $\kappa = \mathbb{E}\left[e^J - 1\right]$. The characteristic function of $X = \log(V(t+H(t))) - \log(V(t))$ in the deterministic case when $H(t)$ is Δt reads,

$$\hat{f}_X(u) = \exp\left(-i\left(\mu - \lambda\kappa - \frac{\sigma^2}{2}\right)\Delta t u - \frac{\sigma^2 \Delta t}{2}u^2 - \mathrm{i}\lambda\left(\frac{p}{\eta_1 + \mathrm{i}u} + \frac{q}{\eta_2 - \mathrm{i}u}\right)\Delta t u\right), \tag{4.90}$$

and the cumulants are,

$$\begin{aligned}
\kappa_1 &= \left(\mu - \lambda\kappa - \frac{\sigma^2}{2}\right)\Delta t + \left(\frac{p}{\eta_1} + \frac{q}{\eta_2}\right)\lambda\Delta t, \\
\kappa_2 &= \left(\frac{p}{\eta_1^2} - \frac{q}{\eta_2^2}\right)\lambda\Delta t + \sigma^2\Delta t, \\
\kappa_4 &= \left(\frac{p}{\eta_1^4} - \frac{q}{\eta_2^4}\right)\lambda\Delta t.
\end{aligned} \tag{4.91}$$

In Table 4.11 we present the VaR and ES values computed at the regulatory level $\alpha = 0.975$. In the deterministic case, we consider $\Delta t = 5/365$ and the inverse gamma distribution is used for governing the SLH dynamics. The benchmark solution is by the MC technique and $V_0 = 100$. We observe that the absolute errors are in line with those obtained in the former sections when comparing with MC.

4.4. CONCLUSIONS

In this chapter we presented a numerical method to efficiently calculate VaR and ES values within a stochastic liquidity horizon framework. We therefore focus on two aspects underlined as key regulatory changes by the Basel Committee of Banking Supervision, like moving from VaR to ES and considering the incorporation of the risk of market illiquidity. The estimation of the risk measures with a stochastic holding period appears to be particularly challenging in terms of computational power.

		VaR			ES		
$H(t)$	m	MC	SWIFT	Absolute Error	MC	SWIFT	Absolute error
5/365	5	0.1109	0.1110	$1.3e-04$	0.2225	0.1670	$5.6e-02$
SLH	4	0.3262	0.3335	$7.4e-03$	0.8129	0.8058	$7.1e-03$

Table 4.11: Absolute errors for VaR and ES values with respect to MC simulation when $\Pi(t)$ follows a Kou dynamics with $\mu = 0.1$, $\sigma = 0.5$, $\lambda = 0.6$, $\eta_1 = 1.5$, $\eta_2 = 1.8$, $p = 0.5$. In the deterministic case $\Delta t = 5/365$. The SLH is driven by an inverse gamma distribution with $\alpha = 6$, $\beta = 0.5$. The confidence level is $\alpha = 0.975$.

For the aforementioned reasons, we employ the SWIFT method, which recovers the density function of the change in value of a certain portfolio from its characteristic function. The density function is approximated by a finite expansion in terms of Shannon wavelets and the coefficients of the approximation are readily obtained by a Fourier transform inversion. This method relies on the availability of the characteristic function, which is known in closed form for many interesting processes in finance. We consider the well-known delta-gamma approach for modelling the change in value of the portfolio under normal and t-distributed risk factors, as well as the GBM, MJD and Kou models for the log-value change of the portfolio, where these two last models incorporate a jump component in the dynamics. As for the dynamics of the SLH, we consider the Bernoulli distribution, the exponential, the generalised Pareto and the inverse gamma distributions in combination with the delta-gamma approach, and the GBM, MJD and Kou dynamics. We carry out a detailed error analysis and we provide a prescription on how to select the parameters of the numerical method, making this technique more robust, reliable and applicable in practice.

CHAPTER 5

Quantifying Credit Portfolio Losses under Multi-Factor Models

Credit risk refers to the risk coming from a borrower who may not repay a loan and that the lender may lose the principal of the loan or the interest associated with it.

In this chapter, we investigate the challenges of estimating credit risk measures of portfolios with exposure concentration under the multi-factor Gaussian and multi-factor t-copula models. We present efficient and robust numerical techniques based on the Haar wavelets theory for recovering the cumulative distribution function of the loss variable from its characteristic function.

This chapter is based on the article [21].

5.1. Introduction

Financial institutions need to evaluate and manage the risk arising from its business activities. Credit risk is the risk of lose from the obligor's failure to honour the contractual agreements. It is usually the main source of risk in a commercial bank. Banks are subject to *regulatory capital* requirements under Basel Accords and they are forced to keep aside a cushion to absorb potential losses in the future. The capital needed in order to remain solvent at a certain confidence level is called *economic capital*. The basic regulatory risk measure in credit risk is Value-at-Risk (VaR) and it is a quantile of the loss distribution computed at 99.9% confidence level with a one-year time horizon. Although it is still the regulatory measure, the VaR value has mainly two drawbacks that may impede a robust credit risk measurement. One of these two disadvantages is that VaR is not sub-additive and contradicts the idea of diversification (see Section 4.2.1 in Chapter 4). The second is that VaR gives no indication about the severity of losses beyond the computed quantile. This is the reason why the Expected Shortfall (ES) might be used in place of the VaR value for internal risk capital assessment (i.e., for economic capital calculation) as well as what happens with market risk.

The Vasicek model forms the basis of the Basel II approach. It is a Gaussian one-factor model where default events are driven by a latent common factor that is assumed to follow a Gaussian distribution, also called the *Asymptotic Single Risk Factor* (ASRF) model. 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 analytical asymptotic approximation for the loss distribution and VaR value. This approximation works well for a large number of small exposures but can underestimate risks in the presence of exposure concentrations (see [63]). Concentration risks in credit portfolios arise from an unequal distribution of loans to single borrowers (*exposure* or *name concentration*) or different industry or regional

sectors (*sector* or *country concentration*). While regulatory capital is estimated by means of the ASRF model under Pillar I, economic capital takes into account concentration risks and is calculated under Pillar II.

Monte Carlo simulation either with one-factor or multi-factor models (to account for sector concentration or for modelling complicated correlation structures) is a standard method for measuring the risk of a credit portfolio. However, this method is time-consuming when the size of the portfolio increases. Computations can become unworkable in many situations, taking also into account that financial companies have to rebalance their credit portfolios frequently. On top of that, when using MC methods the variance is always an issue when estimating the risk measures at high confidence levels. For the aforementioned reasons, numerical methods are appealing in this field. Different techniques can be found in the literature for estimating the risk with multi-factor Gaussian copula models, like MC methods in [37], Hermite approximations in [68], where the main application is for large loan or mortgage portfolios (3000 loans), a hierarchical factor model in [30] where closed-form solutions are derived under the assumption that the number of sectors in the portfolio is large, and an extension of the granularity adjustment technique to a new dimension is developed in [70].

As pointed out in [45], some works suggested that default events driven by t-distributed random variables provide better empirical fit to the observed data. This is the called t-copula model, where default events are expressed as the ratio of a normal and a scaled chi-square random variable. The bivariate version of this last type of models is tackled with simulation in [17, 75] and a complicated multi-factor version in [45].

In this chapter, we develop numerical techniques to contribute to the efficient measurement of VaR and ES values for small or big portfolios in the presence of exposure concentration under high-dimensional models. It is worth remarking that small and/or concentrated portfolios are particularly challenging cases, since asymptotic methods usually work out well for large and diversified portfolios. We model the dependence among obligors by means of multi-factor Gaussian copula and multi-factor t-copula models. To the best of our knowledge, this is the first time that multi-factor t-copula models are considered outside the MC framework. We estimate the risk measures in a procedure composed of two main parts. The first part is the numerical computation of the characteristic function associated to the portfolio loss variable. We tackle this part with different techniques depending on the underlying model. For the (bivariate) t-copula model we perform a double integration with Gauss-Hermite and generalized Gauss-Laguerre quadrature, while the multi-factor Gaussian model is treated with the quadratic transform approximation (QTA) method put forward in [39], where the authors calculate the price of a collateralized debt obligation. We derive the characteristic function for the most challenging model, this is, the multi-factor t-copula, by conditioning on the chi-square random variable of the model and applying the QTA method at each discretization point of the resulting one-dimensional integral. This last model is by far the most involved in terms of computing effort. Once the characteristic function for the loss variable has been obtained, then the second part of the procedure consists of a Fourier inversion to recover its CDF. For this purpose, we use the WA method in Section 2.6.1 based on Haar wavelets developed in [54] for the one-factor Gaussian copula model. We also use the SWIFT method from Section 2.6.2 but, as expected we see a better performance with WA because the functions to approximate are piecewise constant. Furthermore, we have shown that this method outperforms the well-known numerical Laplace transform inversion method

[1, 2] used in risk management in [38, 39] in terms of efficiency and robustness. The numerical experiments carried out in this work show the high accuracy and speed of the method. Another point of importance is the robustness of the wavelet approach. We show how the scale of approximation (this is, the number of terms used to approximate the CDF) is related to the absolute error of the method. All these features make the proposed methodology an efficient and reliable machinery to be used in practice.

The outline of this chapter is as follows. We start with the basics of credit risk management in Section 5.2 where we present the credit risk portfolio problem and its formulation and we give a brief overview on Gaussian and t-copula models for dependence among obligors. Section 5.3 is made of three subsections which include methodology for the efficient evaluation of characteristic functions, the formulae for risk measures computation by the WA method and numerical examples to see the behaviour of the methods. Finally, in Section 5.4 a short conclusion is given.

5.2. THE BASICS OF CREDIT RISK MANAGEMENT

This section is devoted to present the credit portfolio losses problem on the one hand, and on the other hand the models we use to drive the dependence along with their characteristic functions.

5.2.1. CREDIT PORTFOLIO LOSSES

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

Consider a credit portfolio consisting of N obligors. Every obligor n can be characterized by three parameters: the *exposure at default* E_n, the *loss given default* LGD_n and the *probability of default* P_n. The **loss function** is then defined as,

$$L = \sum_{n=1}^{N} L_n, \tag{5.1}$$

with L_n being the individual credit loss of every obligor given by $L_n = E_n \cdot LGD_n \cdot \mathbb{1}_{D_n}$.

The best way to explain each of these parameters is by looking at them via a simple example. Assume a bank lends a certain amount of money (say 10 million Euro) to a company. Also suppose that the company has already taken the full amount and is not supposed to start returning the money before the time horizon T. Then 10 million would be the bank's E_n. On the other hand, the loss given default LGD_n is the percentage of the E_n that the bank would lose in case of default. This quantity takes into account that the bank may be able to recover some of the money, even in the case of default, via collaterals included in the loan (in the contract it may be written that in case of default the bank will get some of the assets of the company which account for a part of the total money lend). Finally the probability of default P_n is the probability that default will happen during the time horizon and is defined as $P_n = \mathbb{P}(D_n)$, where D_n is the event of default taking the value 1 in case of default or 0 otherwise. Note that E_n, LGD_n, P_n and all quantities derived from those three are measured with respect to a specified time horizon. Changing the time horizon would imply a change of those values as well.

We will assume, without loss of generality that $LGD_n = 1$ for all n i.e., the loss given default is 100%. We assume that these three parameters are estimated from empirical default data. If the reader is interested in learning more about how to estimate these values, we refer to the first chapter of [11]. Note that, for instance, in order to estimate the probability of default we can use the scores provided by the rating agencies. Also, regulatory frameworks force the banks to use their calibration methods to approximate the three quantities, so the assumption that we have the values for the three quantities already given is reasonable.

Even assuming known E_n, P_n and LGD_n for all n, depending on the correlation structures between different obligors, evaluating the loss distribution can become computationally expensive. However, for financial institutions, it is essential having an accurate, and fast approximation of the loss function in order to be able to compute accurate risk measures so they know the amount of capital they need to hold as a cushion in case of losses.

5.2.2. FACTOR MODELS AND THEIR CHARACTERISTIC FUNCTIONS

To evaluate the portfolio loss distribution, a key issue is to model the various dependency effects. Direct modelling of the pairwise correlations is impractical since a bank's credit portfolio can easily contain tens of thousands of obligors. Common practice to reduce the computational complexity is therefore to use a so-called factor model of asset correlations. In a factor model different being the individual credit loss, L_i and L_j, are independent conditional on some common factors, which can represent the state of the economy, different industries, geographical regions and so on.

Firstly, we build upon the work in [54] where the one-factor Gaussian copula model was considered as the model framework. After that, we move a step forward by considering, on the one hand, multi-factor models to account for sector concentration and, on the other hand, by selecting different copula models like the t-copula, which is capable of introducing tail dependence in credit portfolios. As a result, we compute the risk measures by means of both, the multi-factor Gaussian copula model as well as the multi-factor t-copula model. Copulas are simply the joint distribution functions of random vectors with standard uniform marginal distributions. Their value in statistics is that they provide a way of understanding how marginal distributions of single risks are coupled together to form joint distributions of groups of risks, that is, they provide a way of understanding the idea of statistical dependence. For the sake of completeness, we give in Remark 5.2.1 a review of the dependence structure given by the Gaussian and t-copula models. For further details on dependence and copulas see [57, 75].

As mentioned in Section 5.1, one of the two steps involved in the estimation of risk measures is the efficient computation of the characteristic function associated to the loss random variable L in (5.1). We consider the one-factor Gaussian copula model treated in [54, 63], the multi-factor Gaussian copula model, which is the extension from one factor to several factors, the t-copula and multi-factor t-copula models. We tackle the one-factor and multi-factor models separately for methodological reasons that will be explained later on.

THE ONE-FACTOR GAUSSIAN COPULA MODEL

The one-factor Gaussian copula model (also known as Vasicek model) belongs to the class of structural models and it is a one period default model, i.e., loss only occurs when an obligor defaults in a fixed time horizon. Based on Merton's firm-value model (1974), to describe the obligor's default and its correlation structure, we assign to each obligor a random variable called firm-value. The firm-value (or, more precisely, the asset value log-return) W_n of obligor

n at time T is represented by a common, standard normally distributed factor Y component (the state of the world or business cycle, usually called systematic factor) and an idiosyncratic noise component Z_n,

$$W_n = \sqrt{\rho_n} Y + \sqrt{1-\rho_n} Z_n, \tag{5.2}$$

where Y and Z_n, $\forall n \leq N$ are i.i.d. standard normally distributed and $\rho_1, \cdots, \rho_N \in (0,1)$ are correlation parameters calibrated to market data. In case that $\rho_n = \rho$ for all n, the parameter ρ is called the common asset correlation. Using the factor structure (5.2), obligors become independent conditional on Y.

In the Merton's model, obligor n defaults when its firm-value falls below the threshold level c_n, defined by $c_n := \Phi^{-1}(P_n)$, where $\Phi^{-1}(x)$ denotes the inverse of the standard normal CDF. We can therefore define $D_n := \{W_n < c_n\}$ and the probability of default of obligor n conditional to a realization $Y = y$ is given by,

$$p_n(y) := \mathbb{P}(D_n | Y = y) = \Phi\left(\frac{c_n - \sqrt{\rho_n} y}{\sqrt{1-\rho_n}}\right). \tag{5.3}$$

Consequently, the conditional probability of default depends on the systematic factor, reflecting the fact that the business cycle affects the possibility of an obligor's default.

By the law of iterated expectations (Proposition 2.2.8), the characteristic function Ψ_L of L is therefore given by,

$$\begin{aligned}
\Psi_L(\omega) &= \mathbb{E}\left[\mathbb{E}\left[e^{-i\omega L} \,\middle|\, Y \right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\exp\left(-i\omega \sum_{n=1}^{N} E_n \mathbb{1}_{\{W_n < c_n\}}\right) \,\middle|\, Y \right]\right] \\
&= \mathbb{E}\left[\prod_{n=1}^{N} \mathbb{E}\left[e^{-i\omega E_n \mathbb{1}_{\{W_n < c_n\}}} \,\middle|\, Y \right]\right] \\
&= \mathbb{E}\left[\prod_{n=1}^{N} \vartheta_n(\omega, y)\right] \\
&= \int_R f_Y(y) \prod_{n=1}^{N} \vartheta_n(\omega, y) dy,
\end{aligned} \tag{5.4}$$

where

$$\vartheta_n(\omega, y) := e^{-i\omega E_n} p_n(y) + 1 - p_n(y), \tag{5.5}$$

and f_Y is the standard normal probability density function of Y,

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}. \tag{5.6}$$

MULTI-FACTOR GAUSSIAN COPULA

We start this section by focussing our attention to the concept of copulas and then the introduction of the multi-factor model is straightforward.

Remark 5.2.1 (Gaussian and t-copula models)**.** *Let us give attention to the case in which default dependence is modelled as a multivariate Gaussian process. With the same notation as in Section 5.2.1, the unconditional probability of default P_n of obligor n becomes,*

$$\mathbb{P}(D_n) = \mathbb{P}\left(W_n < \Phi^{-1}(P_n)\right),$$

and the joint default probability is given by,

$$\mathbb{P}\left(\mathbb{1}_{D_1} = 1, \ldots, \mathbb{1}_{D_N} = 1\right) = \mathbb{P}\left(W_1 < \Phi^{-1}(P_1), \ldots, W_N < \Phi^{-1}(P_N)\right). \tag{5.7}$$

Let $(u_1, \ldots, u_N) = (P_1, \ldots, P_N)$ be a vector in $[0,1]^N$, choose a dependence structure described by correlation matrix Γ. Then, the unique Gaussian copula C_Γ associated with $(W_1, \ldots, W_N)$ is,

$$C_\Gamma(u_1, \ldots, u_N) = \Phi_\Gamma\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_N)\right) = \mathbb{P}\left(W_1 < \Phi^{-1}(u_1), \ldots, W_N < \Phi^{-1}(u_N)\right), \tag{5.8}$$

for any $(u_1, \ldots, u_N) \in [0,1]^N$, where Φ_Γ is the multivariate standard Gaussian distribution function with correlation matrix Γ.

In the same way, we can extract a copula from the multivariate normal distribution, we can extract an implicit copula from any other distribution with continuous marginal distribution functions. For example, the (N-dimensional) t-copula takes the form,

$$C_{\nu,\Gamma}(u_1, \ldots, u_N) = \Phi_{\nu,\Gamma}\left(\Phi_\nu^{-1}(u_1), \ldots, \Phi_\nu^{-1}(u_N)\right) = \mathbb{P}\left(W_1 < \Phi_\nu^{-1}(u_1), \ldots, W_N < \Phi_\nu^{-1}(u_N)\right), \tag{5.9}$$

for any $(u_1, \ldots, u_N) \in [0,1]^N$, where $\Phi_{\nu,\Gamma}$ is the multivariate t-distribution function with correlation matrix Γ, and Φ_ν^{-1} denotes the inverse of the distribution function of a standard univariate t-distribution with ν degrees of freedom.

Multi-factor models aim at modelling complicated correlation structures. The d-factor Gaussian copula model assumes that the correlation among W_n is introduced by a $d \times 1$ vector $\boldsymbol{Y} = [Y_1, Y_2, \ldots, Y_d]^T$ of independent standard normal random variables, representing the systematic risk factors,

$$W_n = \boldsymbol{a_n}^T \boldsymbol{Y} + b_n Z_n, \quad n = 1, \cdots, N. \tag{5.10}$$

Here, $\boldsymbol{a_n} = [a_{n1}, a_{n2}, \ldots, a_{nd}]^T$ is a $d \times 1$ vector of real constants satisfying $\boldsymbol{a_n}^T \boldsymbol{a_n} < 1$; it represents the vector of factor loading coefficients of the common factors, and Z_n are $\mathcal{N}(0,1)$ random variables representing the idiosyncratic risks, independent of each other and independent of $\boldsymbol{Y}$. The constant b_n, being the factor loading of the idiosyncratic risk factor, is chosen so that W_n has unit variance, i.e., $b_n = \sqrt{1 - (a_{n1}^2 + a_{n2}^2 + \cdots + a_{nd}^2)}$, which ensures that W_n are $\mathcal{N}(0,1)$.

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

$$\begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_N \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{N1} \end{bmatrix} Y_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{N2} \end{bmatrix} Y_2 + \cdots + \begin{bmatrix} a_{1d} \\ a_{2d} \\ \vdots \\ a_{Nd} \end{bmatrix} Y_d + \begin{bmatrix} b_1 Z_1 \\ b_2 Z_2 \\ \vdots \\ b_N Z_N \end{bmatrix}. \tag{5.11}$$

While each Z_n represents the idiosyncratic factor affecting only obligor n, each Y_j, for $j = 1, \ldots, d$, may affect all (or a certain group of) obligors. Although the factors Y_j are sometimes

given economic interpretations (as industry or regional risk factors, for example), the key role of the factors Y_j is that they allow us to model a complicated correlation structure in a non-homogeneous portfolio. More detailed information is available in [39].

Here, the probability of default of obligor n conditional on the realization $\boldsymbol{y} \in \mathbb{R}^d$ of systematic risk factor $\boldsymbol{Y}$ is defined as,

$$p_n(\boldsymbol{y}) := \mathbb{P}(W_n < \Phi^{-1}(P_n) | \boldsymbol{Y} = \boldsymbol{y}) = \Phi\left(\frac{\Phi^{-1}(P_n) - \boldsymbol{a_n^T y}}{b_n}\right). \tag{5.12}$$

Given a realization of the systematic risk factor $\boldsymbol{Y}$, defaults are independent. Then,

$$\mathbb{E}\left[\mathrm{e}^{-i\omega L} \,\middle|\, \boldsymbol{Y}\right] = \prod_{n=1}^{N} \mathbb{E}\left[e^{-i\omega E_n \mathbb{1}_{\{W_n < c_n\}}} \,\middle|\, \boldsymbol{Y}\right] = \prod_{n=1}^{N} \left(1 + p_n(\boldsymbol{y})\left(\mathrm{e}^{-i\omega E_n} - 1\right)\right), \tag{5.13}$$

where $c_n = \Phi^{-1}(P_n)$ and $\Phi^{-1}(\cdot)$ denotes, as usual, the inverse of the standard normal CDF. By the law of iterated expectations, the characteristic function Ψ_L of L is therefore given by,

$$\Psi_L(\omega) = \mathbb{E}\left[e^{-i\omega L}\right] = \mathbb{E}\left[\mathbb{E}\left[e^{-i\omega L} \,\middle|\, \boldsymbol{Y}\right]\right] = \mathbb{E}\left[\prod_{n=1}^{N} \vartheta_n(\omega, \boldsymbol{y})\right] = \int_{\mathbb{R}^d} f_{\boldsymbol{Y}}(\boldsymbol{y}) \prod_{n=1}^{N} \vartheta_n(\omega, \boldsymbol{y}) d\boldsymbol{y}, \tag{5.14}$$

where $\vartheta_n(\omega, \boldsymbol{y}) := 1 + p_n(\boldsymbol{y})(e^{-i\omega E_n} - 1)$ and $f_{\boldsymbol{Y}}$ is the d-dimensional standard Gaussian density.

ONE-FACTOR t-COPULA WITH ν DEGREES OF FREEDOM AND t-DISTRIBUTED MARGINS

As pointed out in [75], while Gaussian copulas do not exhibit tail dependence, this is, the defaults do not occur simultaneously, t-copula models admit tail dependence, with fewer degrees of freedom producing stronger dependence. We assume that latent random variables $W_1, \cdots, W_N$ are generated by,

$$W_n = \sqrt{\frac{\nu}{V}}(\sqrt{\rho_n} Y + \sqrt{1 - \rho_n} Z_n), \tag{5.15}$$

where $Z_1, \cdots, Z_N, Y \sim \mathcal{N}(0,1)$, $V \sim \chi^2(\nu)$ and $Z_1, \cdots, Z_N$, Y and V are mutually independent. Again, $\rho_1, \cdots, \rho_N \in (0,1)$ are correlation parameters calibrated to market data. Systematic risk factor Y may be interpreted as an underlying risk driver or economic factor, with each realization describing a scenario of the economy. Random variables $Z_1, \cdots, Z_N$ represent idiosyncratic, or obligor specific, risk. Scaling the model in (5.2) by the factor $\sqrt{\nu/V}$ transforms standard Gaussian random variables into t-distributed random variables with ν degrees of freedom.

The probability of default of obligor n conditional on $Y = y$ and $V = v$ is then given by,

$$\begin{aligned} p_n(y, v) &:= \mathbb{P}\left(W_n < \Phi_\nu^{-1}(P_n) \middle| Y = y, V = v\right) \\ &= \mathbb{P}\left(\sqrt{\frac{\nu}{V}}\left(\sqrt{\rho_n} Y + \sqrt{1 - \rho_n} Z_n\right) < \Phi_\nu^{-1}(P_n) \middle| Y = y, V = v\right) \\ &= \mathbb{P}\left(Z_n < \frac{\sqrt{v/\nu}\Phi_\nu^{-1}(P_n) - \sqrt{\rho_n} y}{\sqrt{1 - \rho_n}}\right) = \Phi\left(\frac{\sqrt{v/\nu}\Phi_\nu^{-1}(P_n) - \sqrt{\rho_n} y}{\sqrt{1 - \rho_n}}\right). \end{aligned} \tag{5.16}$$

As pointed out in [76] and because of the conditional independence, the characteristic function of the loss L now reads,

$$\begin{aligned}\Psi_L(\omega) &= \mathbb{E}\left[e^{-i\omega L}\right] = \mathbb{E}\left[\mathbb{E}\left[e^{-i\omega L}\,\middle|\, Y, V\right]\right] = \mathbb{E}\left[\prod_{n=1}^{N}\mathbb{E}\left[e^{-i\omega E_n \mathbb{1}_{\{W_n<c_n\}}}\,\middle|\, Y, V\right]\right] \\ &= \mathbb{E}\left[\prod_{n=1}^{N}\vartheta_n(\omega, y, v)\right] = \int_{\mathbb{R}\times[0,+\infty)} f_Y(y) f_V(v) \prod_{n=1}^{N}\vartheta_n(\omega, y, v)\, dy\, dv,\end{aligned} \tag{5.17}$$

where the threshold level is defined by $c_n = \Phi_v^{-1}(P_n)$ and,

$$\vartheta_n(\omega, y, v) := 1 + p_n(y, v)(e^{-i\omega E_n} - 1), \tag{5.18}$$

f_Y is the standard normal probability density function of Y, and f_V is the chi-square PDF with ν degrees of freedom, i.e.,

$$f_V(v) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} v^{\frac{\nu}{2}-1} e^{-v/2}. \tag{5.19}$$

Note that we have replaced the joint density of Y and V in (5.17) by the product of the marginal densities because they are independent random variables.

MULTI-FACTOR t-COPULA MODEL

In this section, we consider a multi-factor model with dependence among obligors driven by a t-copula. We assume that latent random variables $W_1, \cdots, W_N$ are generated by,

$$W_n = \sqrt{\frac{\nu}{V}}\left(\boldsymbol{a_n}^T \boldsymbol{Y} + b_n Z_n\right), \tag{5.20}$$

where $\boldsymbol{Y}, Z_n, \boldsymbol{a_n}$ and b_n are defined and follow the same hypothesis as in Section 5.2.2, with $V \sim \chi^2(\nu)$ as in Section 5.2.2. Here the probability of default of obligor n conditional on $\boldsymbol{Y} = \boldsymbol{y}$ and $V = v$ is given by,

$$p_n(\boldsymbol{y}, v) = \Phi\left(\frac{\sqrt{v/\nu}\Phi_\nu^{-1}(P_n) - \boldsymbol{a_n}^T \boldsymbol{y}}{b_n}\right), \tag{5.21}$$

the conditional expectation reads,

$$\mathbb{E}\left[e^{-i\omega L}\,\middle|\, \boldsymbol{Y}, V\right] = \prod_{n=1}^{N}\left(1 + p_n(\boldsymbol{y}, v)\left(\mathrm{e}^{-i\omega E_n}\right)\right), \tag{5.22}$$

and thus,

$$\Psi_L(\omega) = \int_{\mathbb{R}^d\times[0,+\infty)} f_{\boldsymbol{Y}}(\boldsymbol{y}) f_V(v) \prod_{n=1}^{N}\vartheta_n(\omega, \boldsymbol{y}, v)\, d\boldsymbol{y}\, dv, \tag{5.23}$$

where $\vartheta_n(\omega, \boldsymbol{y}, v) := 1 + p_n(\boldsymbol{y}, v)\left(\mathrm{e}^{-i\omega E_n} - 1\right)$, $f_{\boldsymbol{Y}}$ is the d-dimensional standard Gaussian density and f_V is the chi-square PDF with ν degrees of freedom.

5.3. QUANTIFYING CREDIT PORTFOLIO LOSSES BY WA METHOD

5.3.1. EFFICIENT COMPUTATION OF CHARACTERISTIC FUNCTIONS

The evaluation of the characteristic function (5.4), which is a particular case of (5.14) for $d = 1$, involves the computation of an integral that we solve efficiently as in [63] by means of numerical quadrature in the $d = 1$ case. The evaluation of the characteristic function (5.17) in a certain point w, which is a particular case of (5.23) for $d = 1$, involves the computation of a double integral that we solve efficiently in this section 5.3.1 by means of numerical quadrature for $d = 1$. However, looking at the expressions of the characteristic functions (5.14) and (5.23), corresponding to the multi-factor Gaussian and t-copula models respectively, we realize that a direct attempt of solving the d- and $(d+1)$-dimensional integrals, respectively, at fixed points w is not affordable with numerical integration. For these challenging tasks, we rely on the QTA method put forward in [39] for computing the Laplace transform of the portfolio loss within the multi-factor Gaussian copula model. We introduce the QTA method and use it to calculate the characteristic function (5.14), and we will show how we can benefit from the QTA approach for the challenging case of the multi-factor t-copula by conditioning on the V factor.

ONE-FACTOR GAUSSIAN COPULA

If we replace ϑ_n and f_Y in expression (5.4) by their corresponding expressions (5.5) and (5.6) we obtain the integral,

$$\Psi_L(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{y^2}{2}} \prod_{n=1}^{N} \left((e^{-i\omega E_n} - 1) p_n(y) + 1 \right) dy. \tag{5.24}$$

The one-factor Gaussian copula model was already treated in [54, 63] and the characteristic function integral is solved by a Gauss-Hermite quadrature. Alternatively, it can also be considered the QTA method presented for the multi-factor Gaussian copula for dimension $d = 1$ in a similar way as it is presented for the one-factor t-copula model later.

We thus solve the integral (5.24) by Gauss-Hermite quadrature. For sake of completeness we briefly review the method. Gauss-Hermite quadrature is a particular type of Gaussian quadrature for approximating the value of integrals of the form,

$$\int_{\mathbb{R}} e^{-x^2} f(x)\, dx. \tag{5.25}$$

In this case,

$$\int_{\mathbb{R}} e^{-x^2} f(x)\, dx \approx \sum_{i=1}^{n_h} w_i f(x_i), \tag{5.26}$$

where n_h is the number of sample points used. The $x_i, i = 1, 2, \cdots, n_h$, are the roots of the Hermite polynomial of degree n_h, $H_{n_h}(x)$, and the associated weights w_i are given by

$$w_i = \frac{2^{n_h - 1} n_h! \sqrt{\pi}}{{n_h}^2 [H_{n_h - 1}(x_i)]^2}. \tag{5.27}$$

Using this quadrature,

$$\begin{aligned}\Psi_L(\omega) &= \frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} e^{-x^2}\prod_{n=1}^{N}\Big((e^{-i\omega E_n}-1)p_n(\sqrt{(2)}x)+1\Big)dx \\ &= \frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} e^{-x^2}\vartheta(x)dx \\ &\approx \frac{1}{\sqrt{\pi}}\sum_{i=1}^{N_h} w_i\vartheta(x_i),\end{aligned} \tag{5.28}$$

where $\vartheta(x)=\prod_{n=1}^{N}\big((e^{-i\omega E_n}-1)p_n(\sqrt{(2)}x)+1\big)$ and n_h is the number of sample points of the quadrature.

MULTI-FACTOR GAUSSIAN COPULA

We start this section by enunciating the proposition which forms the basis of the QTA method.

Proposition 5.3.1 (Proposition 1 of [39]). *Let $\boldsymbol{Z}$ be a $d\times 1$ vector of independent standard normal variables. For any scalar $c\in\mathbb{C}$, vector $\boldsymbol{g}\in\mathbb{C}^d$, and matrix $\boldsymbol{H}\in\mathbb{C}^{d\times d}$ for which $\Re(\boldsymbol{H})$ is negative-semidefinite, it follows that,*

$$\mathbb{E}\left[e^{c+\boldsymbol{g}^T\boldsymbol{Z}+\boldsymbol{Z}^T\boldsymbol{H}\boldsymbol{Z}}\right]=\frac{1}{\sqrt{\det(\boldsymbol{I}-2\boldsymbol{H})}}e^{c+\boldsymbol{g}^T(\boldsymbol{I}-2\boldsymbol{H})^{-1}\boldsymbol{g}/2}, \tag{5.29}$$

where $\Re(z)$ denotes the real part of z.

First of all, let us define the mapping $s\mapsto g_n(s)$, for $n=1,\cdots,N$, as,

$$g_n(w,s):=1+\left(\mathrm{e}^{-i\omega E_n}-1\right)\Phi\left(\frac{\Phi^{-1}(P_n)+s\sqrt{\boldsymbol{a_n}^T\boldsymbol{a_n}}}{b_n}\right),\quad s\in\mathbb{R}. \tag{5.30}$$

Thus, one can rewrite the conditional expectation (5.13) as follows,

$$\mathbb{E}\Big[\,e^{-i\omega L}\,\Big|\,\boldsymbol{Y}\Big]=\prod_{n=1}^{N}g_n(w,S_n)=\mathrm{e}^{\sum_{n=1}^{N}\log g_n(w,S_n)},\quad\text{where } S_n:=-\frac{\boldsymbol{a_n}^T\boldsymbol{Y}}{\sqrt{\boldsymbol{a_n}^T\boldsymbol{a_n}}}. \tag{5.31}$$

If we approximate $\log g_n(w,S_n)$ by a quadratic function of S_n,

$$\log g_n(w,S_n)\approx\alpha_n(w)+\beta_n(w)S_n+\gamma_n(w)S_n^2, \tag{5.32}$$

where the scalars $\alpha_n(w),\beta_n(w)$ and $\gamma_n(w)$ are complex-valued, then one can use Proposition 5.3.1 to (5.31) and obtain a closed-form approximation for Ψ_L in (5.14),

$$\Psi_L(\omega)=\mathbb{E}\left[\mathrm{e}^{\sum_{n=1}^{N}\log g_n(w,S_n)}\right]\approx\mathbb{E}\left[\mathrm{e}^{\sum_{n=1}^{N}(\alpha_n(w)+\beta_n(w)S_n+\gamma_n(w)S_n^2)}\right]=\mathbb{E}\left[\mathrm{e}^{c(w)+\boldsymbol{g}^T(w)\boldsymbol{Y}+\boldsymbol{Y}^T\boldsymbol{H}(w)\boldsymbol{Y}}\right]. \tag{5.33}$$

The last equality follows from the fact that the S_n's are linear in $\boldsymbol{Y}$. The scalar c, vector $\boldsymbol{g}$, and matrix $\boldsymbol{H}$ are given explicitly by,

$$c(w)=\sum_{n=1}^{N}\alpha_n(w),\quad \boldsymbol{g}(w)=-\sum_{n=1}^{N}\frac{\beta_n(w)\boldsymbol{a_n}}{\sqrt{\boldsymbol{a_n}^T\boldsymbol{a_n}}},\quad\text{and}\quad \boldsymbol{H}(w)=\sum_{n=1}^{N}\gamma_n(w)\frac{\boldsymbol{a_n}\boldsymbol{a_n}^T}{\boldsymbol{a_n}^T\boldsymbol{a_n}}. \tag{5.34}$$

In order to obtain the coefficients $(\alpha_n(w),\beta_n(w),\gamma_n(w))$ we make use of the weighted least-squares method. We will go into the details in the numerical examples section.

ONE-FACTOR t-COPULA WITH ν DEGREES OF FREEDOM AND t-DISTRIBUTED MARGINS

If we replace ϑ_n, f_Y and f_V in expression (5.17) by their corresponding expressions (5.18), (5.6) and (5.19), we obtain the iterated integral,

$$\Psi_L(\omega) = \frac{1}{\sqrt{2\pi}2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \int_0^{+\infty} v^{\frac{\nu}{2}-1} e^{-\frac{v}{2}} \left(\int_{\mathbb{R}} e^{-\frac{1}{2}y^2} \prod_{n=1}^{N} \left[\left(e^{-i\omega E_n} - 1 \right) p_n(y,v) + 1 \right] dy \right) dv. \tag{5.35}$$

We propose in this section a numerical integration method to evaluate $\Psi_L(\omega)$ in (5.35) at a fixed point w. Later we show how the one-dimensional version of the QTA method can be applied by conditioning on the V factor. We will compare the efficiency of both methods in the numerical experiments section and show that in this case, numerical integration is superior to the QTA method.

Numerical integration

We solve the inner integral in (5.35) by Gauss-Hermite quadrature and the outer integral by generalized Gauss-Laguerre quadrature. For sake of completeness we review the generalized Gauss-Laguerre method here, Gauss-Hermite is already introduced in (5.26). Generalized Gauss-Laguerre quadrature is an extension of the Gauss-Laguerre quadrature method for approximating the value of integrals over $\mathbb{R}^+$ with integrands of the form $x^\alpha e^{-x} f(x)$ for some real number $\alpha > -1$. Precisely,

$$\int_0^{+\infty} x^\alpha e^{-x} f(x)\, dx. \tag{5.36}$$

This allows us to efficiently evaluate such integrals for polynomial or smooth $f(x)$ even when α is not an integer. In this case,

$$\int_0^{+\infty} x^\alpha e^{-x} f(x)\, dx \approx \sum_{i=1}^{n_l} w_i f(x_i), \tag{5.37}$$

where x_i is the i-th root of Laguerre polynomial of degree $n_l, \mathbb{L}_{n_l}(x)$, and the weight w_i is given by $w_i = \frac{x_i}{(n_l+1)^2[\mathbb{L}_{n_l+1}(x_i)]^2}$. Finally, if we perform the change of variables $y = \sqrt{2}x$ in (5.35), and apply the quadrature rules (5.26) to the inner integral and (5.36) to the outer integral, then,

$$\begin{aligned} \Psi_L(\omega) &= \frac{1}{\sqrt{\pi}2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \int_0^{+\infty} v^{\frac{\nu}{2}-1} e^{-\frac{v}{2}} \left(\int_{\mathbb{R}} e^{-x^2} \prod_{n=1}^{N} \vartheta_n(\omega, \sqrt{2}x, v) dx \right) dv \\ &\approx \frac{1}{\sqrt{\pi}2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \sum_{i=1}^{n_l} w_i^l \left(\sum_{j=1}^{n_h} w_j^h \tilde{\vartheta}(w, x_j^h, v_i^l) \right), \end{aligned} \tag{5.38}$$

where $\tilde{\vartheta}(w,x,v) = \prod_{n=1}^{N} \vartheta_n(w, \sqrt{2}x, v)$, and w_j^h, x_j^h are the weights and sample points in Gauss-Hermite quadrature and w_i^l, v_i^l are the weights and sample points in the generalized Gauss-Laguerre quadrature. We observe that the computational complexity of evaluating the numerical formula (5.38) for a fixed value w is $\mathcal{O}(n_l \cdot n_h)$.

One-dimensional QTA method

Let us start by presenting the one-dimensional version of formula (5.29) in Proposition 5.3.1.

Corollary 5.3.2. *Let Z be a standard normal random variable. For any scalars $c, g, h \in \mathbb{C}$, for which $\Re(h) \leq 0$, we have,*

$$\mathbb{E}\left[e^{c+gZ+hZ^2}\right] = \frac{1}{\sqrt{(1-2h)}} e^{c+\frac{g^2}{2(1-2h)}}. \tag{5.39}$$

Proof. Follows immediately from Proposition 5.3.1 by considering $d = 1$. □

The key idea in this section is conditioning on the factor V in (5.15) and applying Corollary 5.3.2 for each fixed value v of V. Looking at expression (5.17),

$$\Psi_L(\omega) = \int_{\mathbb{R}\times[0,+\infty)} f_Y(y) f_V(v) \prod_{n=1}^{N} \vartheta_n(\omega, y, v)\, dy\, dv, \tag{5.40}$$

we observe that we can write,

$$\Psi_L(\omega) = \int_0^{+\infty} f_V(v) \mathbb{E}\left[\mathbb{E}\left[e^{-i\omega L(v)} \,\middle|\, Y \right]\right] dv = \int_0^{+\infty} f_V(v) \mathbb{E}\left[e^{\sum_{n=1}^{N} \log g_n(\omega, Y, v)}\right] dv, \tag{5.41}$$

where $L(v)$ emphasizes the fact that V takes a fixed value v and,

$$g_n(\omega, y, v) := 1 + \left(\mathrm{e}^{-i\omega E_n} - 1\right) \Phi\left(\frac{\sqrt{\frac{v}{\nu}}\Phi_\nu^{-1}(P_n) - \sqrt{\rho_n}\, y}{\sqrt{1-\rho_n}} \right), \quad \text{for } (y, v) \in \mathbb{R} \times [0, +\infty). \tag{5.42}$$

We can compute $\mathbb{E}\left[e^{\sum_{n=1}^{N} \log g_n(\omega, Y, v)}\right]$ using a similar approximation as previously,

$$\log g_n(\omega, Y, v) \approx \alpha_n(\omega, v) + \beta_n(\omega, v) Y + \gamma_n(\omega, v) Y^2, \tag{5.43}$$

and applying Corollary 5.3.2,

$$\begin{aligned} \mathbb{E}\left[e^{\sum_{n=1}^{N} \log g_n(\omega, Y, v)}\right] &\approx \mathbb{E}\left[\mathrm{e}^{c(\omega,v)+g(\omega,v)Y+h(\omega,v)Y^2}\right] \\ &= \frac{1}{\sqrt{1-2h(\omega,v)}} e^{c(\omega,v)+\frac{g^2(\omega,v)}{1-2h(\omega,v)}}, \end{aligned} \tag{5.44}$$

for $\Re\left(h(\omega, v)\right) \leq 0$ and,

$$c(\omega, v) = \sum_{n=1}^{N} \alpha_n(\omega, v), \quad g(\omega, v) = \sum_{n=1}^{N} \beta_n(\omega, v), \quad \text{and} \quad h(\omega, v) = \sum_{n=1}^{N} \gamma_n(\omega, v). \tag{5.45}$$

Hence, from (5.41) and (5.44) we end up with,

$$\Psi_L(\omega) \approx \int_0^{+\infty} f_V(v) \frac{\mathrm{e}^{c(\omega,v)+\frac{g^2(\omega,v)}{2(1-2h(\omega,v))}}}{\sqrt{1-2h(\omega,v)}} dv = \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \int_0^{+\infty} v^{\frac{\nu}{2}-1} \mathrm{e}^{-\frac{v}{2}} \frac{\mathrm{e}^{c(\omega,v)+\frac{g^2(\omega,v)}{2(1-2h(\omega,v))}}}{\sqrt{1-2h(\omega,v)}} dv. \tag{5.46}$$

We solve the integral in (5.46) by means of generalized Gauss-Laguerre quadrature. It is worth remarking the dependence on v for the coefficients $c(\omega, v), g(\omega, v)$ and $h(\omega, v)$. As a consequence, this method involves N quadratic approximations of the type in (5.43) for each discretization point considered in the quadrature rule.

MULTI-FACTOR t-COPULA MODEL

In a similar manner as previously, we define,

$$g_n(w,x,v) := 1 + \left(e^{-i\omega E_n} - 1\right)\Phi\left(\frac{\sqrt{\frac{v}{v}}\Phi_v^{-1}(P_n) + x\sqrt{\boldsymbol{a_n}^T\boldsymbol{a_n}}}{b_n}\right), \text{ for } (x,v) \in \mathbb{R}\times[0,+\infty). \tag{5.47}$$

With this notation we can rewrite,

$$\prod_{n=1}^{N}\vartheta_n(\omega,\boldsymbol{Y},v) = \prod_{n=1}^{N} g_n(w,X_n,v)) = e^{\sum_{n=1}^{N}\log g_n(w,X_n,v)}, \text{ for } X_n = -\frac{\boldsymbol{a_n}^T\boldsymbol{Y}}{\sqrt{\boldsymbol{a_n}^T\boldsymbol{a_n}}}. \tag{5.48}$$

For each n, we approximate $\log g_n(w,X_n,v)$ by a quadratic function in X_n,

$$\log g_n(w,X_n,v) \approx \alpha_n(w,v) + \beta_n(w,v)X_n + \gamma_n(w,v)X_n^2, \tag{5.49}$$

and then,

$$\begin{aligned}\sum_{n=1}^{N}\log g_n(w,X_n,v) &\approx \sum_{n=1}^{N}\alpha_n(w,v) + \sum_{n=1}^{N}\beta_n(w,v)X_n + \sum_{n=1}^{N}\gamma_n(w,v)X_n^2 \\ &= c(w,v) + \boldsymbol{g}^T(w,v)\boldsymbol{Y} + \boldsymbol{Y}^T\boldsymbol{H}(w,v)\boldsymbol{Y},\end{aligned} \tag{5.50}$$

where,

$$c(w,v) = \sum_{n=1}^{N}\alpha_n(w,v), \quad \boldsymbol{g}(w,v) = -\sum_{n=1}^{N}\frac{\beta_n(w,v)\boldsymbol{a_n}}{\sqrt{\boldsymbol{a_n}^T\boldsymbol{a_n}}}, \text{ and } \boldsymbol{H}(w,v) = \sum_{n=1}^{N}\gamma_n(w,v)\frac{\boldsymbol{a_n}\boldsymbol{a_n}^T}{\boldsymbol{a_n}^T\boldsymbol{a_n}}. \tag{5.51}$$

If we follow a similar procedure as before and we apply Proposition 5.3.1, then the characteristic function in (5.23) can be approximated as,

$$\begin{aligned}\Psi_L(\omega) &\approx \int_0^{+\infty} f_V(v)\mathbb{E}\left[e^{c(w,v)+\boldsymbol{g}^T(w,v)\boldsymbol{Y}+\boldsymbol{Y}^T\boldsymbol{H}(w,v)\boldsymbol{Y}}\right]dv \\ &= \frac{1}{2^{\frac{v}{2}}\Gamma\left(\frac{v}{2}\right)}\int_0^{+\infty} v^{\frac{v}{2}-1}e^{-\frac{v}{2}}\frac{1}{\sqrt{\det(\boldsymbol{I}-2\boldsymbol{H}(w,v))}}e^{c(w,v)+\boldsymbol{g}^T(w,v)(\boldsymbol{I}-2\boldsymbol{H}(w,v))^{-1}\boldsymbol{g}(w,v)/2}dv,\end{aligned} \tag{5.52}$$

where $c(w,v)\in\mathbb{C}$, $\boldsymbol{g}(w,v)\in\mathbb{C}^d$ is a vector, $\boldsymbol{H}(w,v)\in\mathbb{C}^{d\times d}$ is a matrix, $\boldsymbol{I}$ is the d-dimensional identity matrix and $f_V(v)$ is the PDF of a chi-square distribution with v degrees of freedom. We want to underline that in order to apply Proposition 5.3.1 we need $\Re(\boldsymbol{H})$ to be negative-semidefinite. In the numerical examples section, we will explain how to proceed when this condition is not satisfied. We compute the integral in expression (5.52) by generalized Gauss-Laguerre quadrature.

5.3.2. RISK MEASURES COMPUTATION BY WA

Central to the methodology presented in this work is an efficient method to carry out the inversion of characteristic functions. We have seen in Section 5.3.1 that characteristic functions of the loss L are known (numerically) for multi-factor Gaussian and t-copula models. We devote the present section to the numerical method selected to perform the inversion step. Let f_L and F_L be the PDF and CDF of L, respectively. Without loss of generality we assume that the sum of the exposures E_n equals one, this is, $\sum_{n=1}^{N} E_n = 1$. Then,

$$F_L(x) = \begin{cases} F^c(x), & \text{if } 0 \le x \le 1, \\ 0, & \text{if } x > 1, \end{cases} \tag{5.53}$$

for a certain F^c defined in $[0,1]$. Recall that Ψ_L is the Fourier transform of the density f_L and then,

$$\Psi_L(\omega) = \int_{\mathbb{R}} e^{-i\omega l} f_L(l)\, dl = \int_{\mathbb{R}} e^{-i\omega l} F_L'(l)\, dl, \tag{5.54}$$

where F_L' is the derivative of distribution function F_L in the context of generalized functions. If we integrate by parts,

$$\Psi_L(\omega) = e^{-i\omega} + i\omega \int_0^1 e^{-i\omega l} F^c(l)\, dl, \tag{5.55}$$

and then the Fourier transform of F^c is given by,

$$\hat{F}^c(w) = \frac{\Psi_L(\omega) - e^{-i\omega}}{i\omega}. \tag{5.56}$$

We aim at recovering F^c from its Fourier transform $\hat{F}^c$. For this purpose, we use the metod initially developed in [54] for Laplace transform inversion and further extended in [64] for Fourier transform inversion, where numerical errors are studied in detail as well. The method used is the WA presented in Section 2.6.1 and based on Haar wavelets. Here we do not need a truncation range because we are in the interval $[0,1]$. We recall in Section 5.3.2 how to calculate the VaR (as in [54]) and the ES (as in [63]).

In the numerical examples Section 5.3.3, we show the popular method [1, 2] for numerical Laplace transform inversion (NLTI), since it is used in our main reference paper [39] and it was also used in [38] for credit portfolio losses. We make a comparison against the WA method in terms of accuracy and speed.

The choice of Haar basis seems natural since the CDF of L is a piecewise constant function. In [65] the authors show how the WA method compares to the COS method [28] for Fourier inversion in the framework of credit portfolio losses. It is shown that COS produces oscillations in the tail leading to non-reliable risk measures. We also show in the numerical examples Section 5.3.3, the similar oscillating behaviour of SWIFT method.

COMPUTATION OF CREDIT RISK MEASURES BY WA

Let $\alpha \in (0,1)$ be a given confidence level (usually the α of interest is close to 1). The α-quantile of the loss distribution of L in this context is the VaR value,

$$\text{VaR}_L(\alpha) = \inf\{l \in \mathbb{R} : \mathbb{P}(L \le l)\} \ge \alpha = \inf\{l \in \mathbb{R} : F_L(l) \ge \alpha\}. \tag{5.57}$$

A bank that manages its risks according to Basel II must reserve capital by an amount of $\text{VaR}_L(\alpha)$ to cover potential extreme losses. According to [54], the coefficients $c_{m,k}$ of the WA approximation of F^C satisfy,

$$0 \leq c_{m,k} \leq 2^{-m/2}, \quad k = 0, \cdots, 2^m - 1, \tag{5.58}$$

and,

$$0 \leq c_{m,0} \leq c_{m,1} \leq \cdots \leq c_{m,2^m-1}. \tag{5.59}$$

Then, looking at (2.91), we conclude that,

$$F^c\left(\frac{k}{2^m}\right) \approx 2^{m/2} c_{m,k}, \tag{5.60}$$

for $k = 0, \cdots, 2^m - 1$. Since $F^c(\text{VaR}_L(\alpha)) \geq \alpha$ we therefore look for $\bar{k}$ such that $\text{VaR}_L(\alpha) \in \left[\frac{\bar{k}}{2^m}, \frac{\bar{k}+1}{2^m}\right]$ and $\frac{\alpha}{2^{m/2}} \in \left[c_{m,\bar{k}}, c_{m,\bar{k}+1}\right]$. Thus, we can simply start searching for the VaR value by means of the following iterative procedure. First we compute $F_m^c\left(\frac{2^{m-1}}{2^m}\right)$. If $F_m^c\left(\frac{2^{m-1}}{2^m}\right) > \alpha$ then we compute $F_m^c\left(\frac{2^{m-1}-2^{m-2}}{2^m}\right)$, otherwise we compute $F_m^c\left(\frac{2^{m-1}+2^{m-2}}{2^m}\right)$, and so on. This algorithm finishes after (at most) m steps storing the $\bar{k}$ value such that $F_m^c\left(\frac{\bar{k}}{2^m}\right)$ is the closest value to α in our m resolution approximation that satisfies $F^c\left(\frac{\bar{k}}{2^m}\right) \geq \alpha$. Subsequently, we approximate the VaR value by the midpoint of the interval $\left[\frac{\bar{k}}{2^m}, \frac{\bar{k}+1}{2^m}\right]$, this is,

$$\text{VaR}_L(\alpha) \approx l_\alpha^m := \frac{2\bar{k}+1}{2^{m+1}}. \tag{5.61}$$

By definition, the ES at confidence level α is given by,

$$\text{ES}_L(\alpha) = \frac{1}{1-\alpha} \int_{l_\alpha}^{+\infty} x f_L(x) dx, \tag{5.62}$$

where $l_\alpha := \text{VaR}_L(\alpha)$. From [63], if we integrate by parts the integral in (5.62) and we use the expansion in (2.91), we obtain the following formula for computing the ES,

$$\text{ES}_L(\alpha) \approx \frac{1}{1-\alpha}\left(1 - \alpha l_\alpha - \frac{1}{2^{\frac{m}{2}+1}} c_{m,\bar{k}} - \frac{1}{2^{m/2}} \sum_{k=\bar{k}+1}^{2^m-1} c_{m,k}\right), \tag{5.63}$$

where l_α is replaced by the VaR value l_α^m calculated at scale m in (5.61).

Remark 5.3.3. *The VaR value can be obtained by calculating (at most) m coefficients $c_{m,k}$ following the algorithm explained in Section 5.3.2. Taking into account that the computation of each coefficient in (2.99) involves $M = 2^m$ terms, the overall complexity for obtaining the VaR is (at most) $2^m \cdot m$ operations. If we apply the FFT algorithm, as explained in Section 2.6.1, we get all the coefficients $c_{m,k}$ at once (ranging from $k = 0$ until $k = 2^m - 1$) with $2^m \cdot log_2(2^m) = 2^m \cdot m$ operations, and we therefore need the same computational effort. The computational savings when using the FFT comes into play when we desire to compute the ES as well, since in that case we need to compute many more extra coefficients (for $\bar{k} + 1 \leq k \leq 2^m - 1$).*

5.3.3. NUMERICAL EXAMPLES

In this section, we show the performance of the WA method for a wide variety of portfolios when the firm-value is driven by the three different models presented in Section 5.2.2. The test portfolios are described in Table 5.1. The matrix $A = (a_{ij})$ for $i = 1,\dots,N$, $j = 1,\dots,d$ contains the factor loadings associated to the models (5.10) and (5.20), that have been generated by simulation following a uniform distribution $\mathscr{U}(a,b)$. As pointed out in [39], the approximation of Proposition 5.3.1 works out well for moderate correlation among obligors, this is, when $\|A\|$ is small, where $\|A\| := \max_j \sum_k |a_{jk}|$. A remedy for the case of strong correlation is presented in [39]. The correlation parameter ρ in (5.15) is assumed to be constant for all the obligors in the portfolio. The dimension $d = 1$ refers to the model in (5.15). Observe that all test portfolios considered show name concentration and C is computed such that $\sum_{n=1}^{N} E_n = 1$. In all experiments of this section, we run the MC simulation with 10^6 scenarios for the systematic risk factors to estimate the VaR and ES risk measures at the regulatory 99.9% confidence level and write down those results as a reference. In regard to this point, we underline that for the reasons mentioned in the previous section, this is not a robust reference to compare with. A robust benchmark is to consider the average of a large sample of VaR and ES estimates with (for instance) 10^6 scenarios each rather than a unique estimate. We dismiss this last strategy because the computation time is in general unaffordable for the models presented in this work applied to portfolios P2 and P4. The wavelet-based approximation F_m^c in (2.91) converges to the true distribution function F^c with the L^2-norm when m increases, since F^c belongs to $L^2([0,1])$. We therefore assume that the true solution is given by the WA method at scale $m = 10$ and use it as our benchmark. The programs in this chapter have been coded in MATLAB and run under macOS Sierra on a laptop with processor 2 GHz intel core i5 and memory 8 GB 1867 MHz LPDDR3.

The QTA approximation presented in Section 5.3.1 relies on the quadratic approximation of function g in (5.30) as well as on Proposition 5.3.1. This proposition can be applied when $\Re(\boldsymbol{H})$ is negative-semidefinite and, as pointed out in [39], this is accomplished when $\gamma_n(w)$ in (5.34) satisfies $\Re(\gamma_n(w)) \le 0$ for all $n = 1,\dots,N$. We carry out the computation of $\alpha_n(w), \beta_n(w), \gamma_n(w)$ by means of weighted least-squares and perform a linear approximation when $\Re(\gamma_n(w))$ is strictly positive. We look for $\alpha_n(w)$, $\beta_n(w)$ and $\gamma_n(w)$ that solve the minimization problem,

$$\min \sum_{\lambda\in\Lambda} \omega(\lambda) \left|\log\big(g_n(w,\lambda)\big) - \alpha_n(w) - \beta_n(w)\lambda - \gamma_n(w)\lambda^2\right|^2, \quad \Re(\gamma_n(w)) \le 0. \tag{5.64}$$

The summands represent the approximation errors at certain grid points $\lambda \in \Lambda$, where $\omega(\lambda)$ represents the penalty weight for the errors. We assume that Λ and ω are the same for all n and that $\sum_{\lambda\in\Lambda} \omega(\lambda) = 1$. The set of grid points Λ and the weight ω should be chosen to reflect the fact that variable S_n in (5.32) is a Gaussian distribution. For our numerical examples, we choose $\omega(\lambda)$ as exponentially decreasing in $|\lambda|^2$, specifically, we choose ω to be the probability density function of a standard Gaussian normalized over Λ. While the grid points in [39] are chosen evenly between -3 and 3, we select the range -7 and 7, since our numerical experiments are more accurate within this interval. The advantage of using the least-squares method to determine $\alpha_n(v)$, $\beta_n(v)$ and $\gamma_n(v)$ is that the optimization problem has a unique closed-form solution.

Portfolio	N	$P_n(\%)$	E_n	ρ/A	d
P1	100	0.21	C/n	0.15	1
P2	1000	1.00	C/n	0.15	1
P3	100	0.21	C/n	$\mathcal{U}(0,0.2)$	8
P4	1000	1.00	C/n	$\mathcal{U}(0,0.1)$	6
P5	100	0.21	C/n	$\mathcal{U}(0,0.1)$	5

Table 5.1: Test portfolios.

ONE-FACTOR GAUSSIAN COPULA

In this first section of the examples, we test the WA method against other two possible competitors. These two methods are on one hand the numerical Laplace transform inversion (NLTI), and on the other hand the SWIFT method from Section 2.6.2.

A popular method for Laplace transform inversion was presented in [1, 2] and used in the context of credit portfolio losses in [38, 39]. We provide a brief description of the method and compare it with the WA method.

Let f be a density function and $\tilde{f}$ its Laplace transform, given by,

$$\tilde{f}(s) = \int_0^{+\infty} e^{-st} f(t)\,dt. \tag{5.65}$$

According to the Bromwich inversion theorem (see for instance Appendix B of [38]),

$$f(t) = \frac{1}{2\pi i} \int_{b-i\infty}^{b+i\infty} e^{st} \tilde{f}(s)\,ds, \tag{5.66}$$

where b is a real number to the right of the singularities of $\tilde{f}$. By performing numerical integration, applying a change of variables and using the technique of Euler summation (as in [1]) to accelerate the convergence of the approximation (see Appendix B of [38]) the integral in (5.66) is approximated by the partial sum,

$$E(\bar{m}, n) = s_n(t) + \sum_{k=1}^{\bar{m}} \binom{\bar{m}}{k} 2^{-\bar{m}} s_{n+1:n+k}(t), \tag{5.67}$$

where,

$$s_{i:j}(t) = \sum_{k=i}^{j} (-1)^k a_k(t), \quad i < j, \qquad s_n(t) = s_{0:n}(t), \tag{5.68}$$

with,

$$a_k(t) = \frac{e^{A/2l}}{2lt} b_k(t), \quad k \geq 0, \tag{5.69}$$

where,

$$b_0(t) = \tilde{f}\left(\frac{A}{2lt}\right) + 2\sum_{j=1}^{l} \mathrm{Re}\left[\tilde{f}\left(\frac{A}{2lt} + \frac{ij\pi}{lt}\right) e^{ij\pi/l}\right], \tag{5.70}$$

and

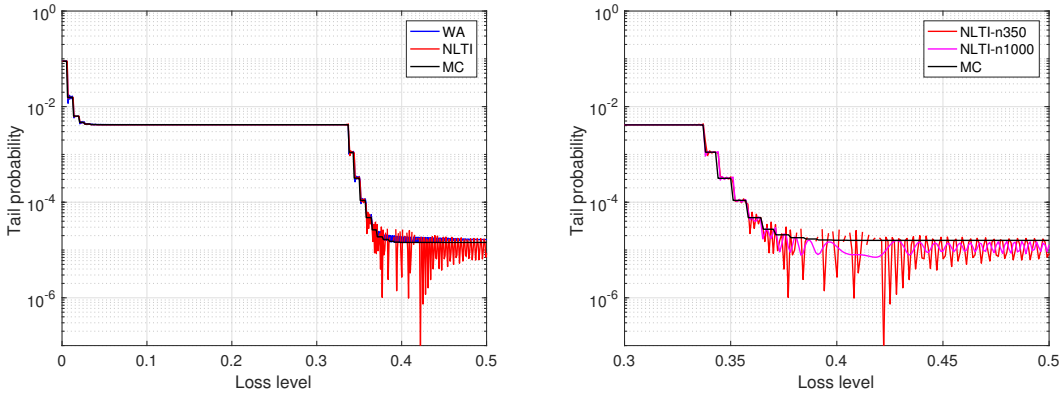
$$b_k(t) = 2\sum_{j=1}^{l} \mathrm{Re}\left[\tilde{f}\left(\frac{A}{2lt} + \frac{ij\pi}{lt} + \frac{ik\pi}{t}\right) e^{ij\pi/l}\right], \quad k \geq 1. \tag{5.71}$$

The parameters l, $\bar{m}$ and A are chosen to control the round-off and discretization errors of the approximation. Usually l is chosen equal to 1, so we consider $l = 1$. The parameter $\bar{m}$ controls the round-off error in an inversely proportional way because the functions $b_k(t)$ in (5.71) are computed with a round-off error of $10^{-\bar{m}}$. In our examples we take $\bar{m} = 8$ to achieve a round-off error of no more than 10^{-8}. The parameter A controls the discretization error of the approximation and according to [38] the following relation provides efficient control of the parameters,

$$A = \left(\frac{2l}{2l+1}\right)\left(\bar{m}\log(10) + \log(2lt)\right). \tag{5.72}$$

The remaining parameter n is the size of the partial sums.

We compare the WA approximation presented in Section 2.6.1 with the NLTI method. For this purpose, we plot the tail probabilities computed by both methods and we perform a MC simulation with 10^7 scenarios for the systematic factor to get our benchmark distribution. The model considered is the simple one-factor Gaussian copula model (5.2) since this comparison is related only to the inversion step. The test portfolio has $N = 50$ obligors and shows severe exposure concentration $E_1 = E_2 = 50, E_n = 1, n = 3, \ldots, N$. We consider $P_n = 0.21\%, \rho_n = 0.15, n = 1, \ldots, N$. The corresponding characteristic function is calculated as in [54], this is, following expression (5.14) for the one-dimensional case and solving the integral by means of Gauss-Hermite quadrature with 20 nodes. We plot the results in Figure 5.1. The left side plot of Figure 5.1 shows the tail probabilities obtained with the WA, NLTI and MC



(a) WA ($m = 10$), NLTI ($n = 350$) and MC.

(b) NLTI ($n = 350$), NLTI ($n = 1000$) and MC.

Figure 5.1: Tail probabilities.

methods. The number of terms employed in NLTI is $n = 350$, while WA has been run at scale of approximation $m = 10$. We have made this choice of n in order to compare the accuracy of both methods at the same cost of CPU time. We observe that while WA is capable to accurately approximate the benchmark solution, the NLTI method shows an oscillatory behaviour further in the tail, which may have an important negative impact in the computation of the ES. We zoom in to show in the right side plot of Figure 5.1 the approximation carried out by the NLTI method when considering more terms in the expansion. In particular, we represent the approximation with $n = 1000$ and observe that heavy oscillations remain in the tail. In addi-

tion to these oscillations, we underline the fact that there is not a prescription on the selection of the suitable parameter n, and adjusting its value may become a matter of trial and error. Regarding the selection made of the scale parameter $m = 10$ for the WA method it is worth remarking that, if we assume that coefficients $c_{m,k}$ in (2.98) were computed exactly, the true VaR value $\text{VaR}_L(\alpha)$ would lie within the interval $\left[\frac{\bar{k}}{2^m}, \frac{\bar{k}+1}{2^m}\right]$ and would differ at most $\frac{1}{2^{m+1}}$ from the computed VaR value $l_\alpha^m = \dfrac{2\bar{k}+1}{2^{m+1}}$ following the algorithm explained in Section 5.3.2, since $|\text{VaR}_L(\alpha) - l_\alpha^m| \leq \frac{1}{2^{m+1}}$. For $m = 10$ we have $\frac{1}{2^{m+1}} \approx 5 \cdot 10^{-4}$, and we therefore expect (at most) three digits accuracy in terms of absolute error with respect to the true solution. We take as a reference in this case a MC simulation with 10^7 scenarios. The error is roughly proportional to $\frac{1}{\sqrt{10^7}} \approx 3 \cdot 10^{-4}$, since this error is affected by the variance $\sqrt{\alpha(1-\alpha)}$ and the true density f_L evaluated at the true quantile (see [34] for details), $\frac{\sqrt{\alpha(1-\alpha)}}{f_L(\text{VaR}_L(\alpha))\sqrt{10^7}}$.

Clearly, SWIFT can be used in an analogous way as WA recovering the CDF F^C. We show the performance of both wavelet inversion methods in Figure 5.2. There, we show tail probability plots for portfolio P1 using WA, SWIFT and MC. We can appreciate that SWIFT method is not as accurate as WA; it oscillates quite a lot further in the tail; this is a similar behaviour observed with the COS method, see Section 5.4 of [65]. Perceptibly, SWIFT as well as COS are really accurate methods when working with continuous functions, however, in the portfolio credit risk problem the recovered function is piecewise constant and this feature is what makes Haar wavelets to perform much better.

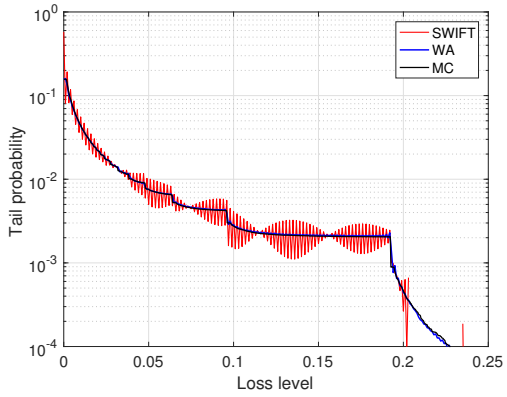
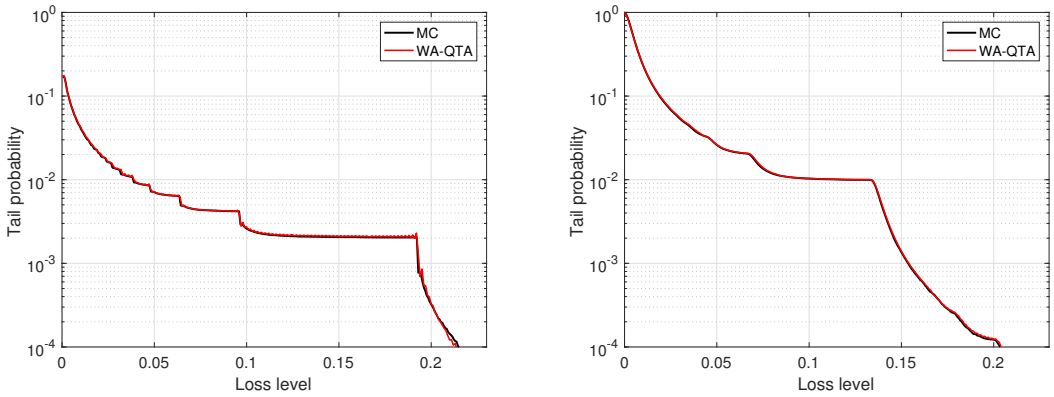


Figure 5.2: Tail probabilities for portfolio P1.

With the evidences of this section, the quality of the WA method becomes clear. Let us now see how it behaves under the different models and when computing the risk measures.

MULTI-FACTOR GAUSSIAN COPULA

In this section we select portfolios P3 and P4 from Table 5.1 for the experiments. Figure 5.3 shows the tail probabilities compared to MC simulation. We name WA-QTA the numerical method employed, where WA stands for the inversion methodology based on wavelets and QTA refers to the quadratic transform approximation for obtaining the characteristic function, as explained in Section 5.3.1. We present the VaR and ES values obtained in Table 5.2 and the



(a) Portfolio P3. (b) Portfolio P4.

Figure 5.3: Tail probabilities.

corresponding CPU times in Table 5.3. The risk measures can be obtained within less than 1% relative error at scale $m = 9$ in about one second for P3 and in about seven seconds for P4, where the dimensions are 8 and 6 respectively. We can get the risk measures in half of those CPU times at scale $m = 8$ keeping accurate results for VaR, although the accuracy worsens for ES.

	Portfolio P3		Portfolio P4	
Method	VaR	ES	VaR	ES
MC	0.1928	0.2029	0.1540	0.1718
WA-QTA ($m = 10$)	0.1948	0.2034	0.1558	0.1720
WA-QTA ($m = 9$)	0.1963 (0.8%)	0.1990 (2.2%)	0.1572 (0.9%)	0.1781 (3.6%)
WA-QTA ($m = 8$)	0.2012 (3.3%)	0.2212 (8.7%)	0.1582 (1.6%)	0.2383 (38.5%)

Table 5.2: VaR and ES values at 99.9% confidence level for portfolios P3 and P4. The relative errors at scales $m = 8,9$ with respect to the WA-QTA method at scale $m = 10$ are shown in parenthesis.

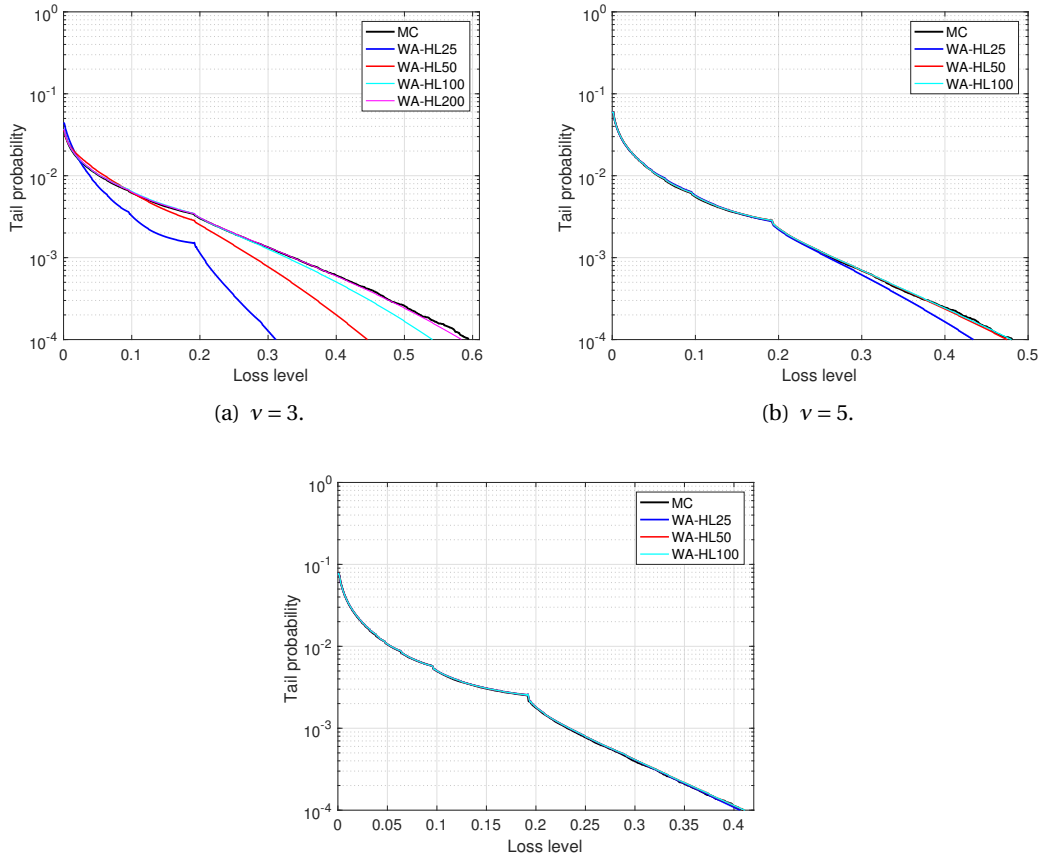
Method	Portfolio P3	Portfolio P4
WA-QTA ($m = 10$)	1.4	16.2
WA-QTA ($m = 9$)	0.7	7.2
WA-QTA ($m = 8$)	0.4	3.6

Table 5.3: CPU time measured in seconds for obtaining the VaR and ES values at 99.9% confidence level for portfolios P3 and P4.

ONE-FACTOR t-COPULA WITH ν DEGREES OF FREEDOM AND t-DISTRIBUTED MARGINS

In this section we select portfolios P1 and P2 from Table 5.1 for the experiments. We use the name WA-HLXX for the numerical method presented in Section 5.3.1, where H stands for Her-

mite, L stands for Laguerre and XX specifies the number of nodes considered for the generalized Gauss-Laguerre quadrature (n_l). The number of nodes for the Gauss-Hermite quadrature is fixed to $n_h = 20$ for all experiments, as in [54, 63] for the one-factor Gaussian copula model. Figure 5.4 shows the tail probabilities of portfolio P1 compared to MC simulation. We observe that the accuracy at high loss levels depends on the parameter ν, where a small ν requires a large n_l and conversely. We present in Table 5.4 results for portfolios P1 and P2 for $\nu = 5$ and



(a) $\nu = 3$.

(b) $\nu = 5$.

(c) $\nu = 7$.

Figure 5.4: Tail probabilities for portfolio P1.

$n_l = 50$, and the corresponding CPU time is shown in Table 5.5. It is worth remarking that, as mentioned previously, the theoretical absolute error estimate when computing the VaR value l_α^m at scale m is $1/2^{m+1}$. We observe that the theoretical error is in good agreement with the order of the empirical error at scale $m = 9$ when we compare with the benchmark at scale $m = 10$ since,

$$|l_\alpha^{10} - l_\alpha^9| \leq |\mathrm{VaR}_L(\alpha) - l_\alpha^{10}| + |\mathrm{VaR}_L(\alpha) - l_\alpha^9| \leq \frac{1}{2^{11}} + \frac{1}{2^{10}} \approx 0.001, \tag{5.73}$$

yielding a relative error of $0.001/26.81 \approx 0.4\%$ for P1 and $0.001/0.3970 \approx 0.3\%$ for P2. The CPU time is very competitive also for big portfolios like P2. Portfolios of this size are particularly

difficult to handle by MC simulation.

	Portfolio P1		Portfolio P2	
Method	VaR	ES	VaR	ES
MC	0.2666	0.3562	0.3971	0.4817
WA-HL50 ($m = 10$)	0.2681	0.3569	0.3970	0.4913
WA-HL50 ($m = 9$)	0.2686 (0.2%)	0.3559 (0.3%)	0.3994 (0.6%)	0.5040 (2.6%)
WA-HL50 ($m = 8$)	0.2715 (1.3%)	0.3596 (0.8%)	0.4004 (0.9%)	0.5305 (8.0%)

Table 5.4: VaR and ES values at 99.9% confidence level for portfolios P1 and P2. The relative errors at scales $m = 8,9$ with respect to the WA-HL50 method at scale $m = 10$ are shown in parenthesis.

Method	Portfolio P1	Portfolio P2
WA-HL50 ($m = 10$)	3.1	24.1
WA-HL50 ($m = 9$)	1.8	13.5
WA-HL50 ($m = 8$)	0.7	4.6

Table 5.5: CPU time measured in seconds for obtaining the VaR and ES values at 99.9% confidence level for portfolios P1 and P2.

In order to check the accuracy of the QTA method, we run this approximation at scale $m = 10$ for portfolio P1 using 50 nodes of generalized Gauss-Laguerre quadrature for the integral in (5.46). The results obtained are 0.2671 for the VaR and 0.3540 for the ES. The CPU time is 85.2 seconds. We get accurate results in comparison with the WA-HL50 method at the same scale of approximation, although the computational effort in this case is almost 30 times higher.

MULTI-FACTOR t-COPULA MODEL

We consider in this last section the most challenging case among the three models presented in this work. We perform our experiments with portfolio P5 of dimension 5, we assume that $\nu = 7$ and consider 25 nodes of integration for the generalized Gauss-Laguerre quadrature in (5.52). We plot tail probabilities in Figure 5.5, where WAm-QTALj stands for WA-QTA method at scale of approximation m and using j nodes of generalized Gauss-Laguerre quadrature. We plot as usual the MC results as a reference and we observe highly accurate VaR values at scales $m = 8,9$ with respect to the benchmark scale $m = 10$ both at 99.9% and 99.99% levels. These results are confirmed looking at VaR values in Table 5.6. The ES values are very accurate at the regulatory level 99.9% and less accurate (but still competitive) values are obtained further in the tail at level 99.99%. It is worth mentioning that this high confidence level is particularly very demanding when computed by MC simulation. The CPU times are presented in Table 5.7. We highlight the impressive CPU times (measured in seconds) in particular at scales $m = 8,9$.

5.4. CONCLUSIONS

In this chapter we have investigated two highly efficient numerical methods to obtain the VaR and ES values for portfolios with exposure concentration under multi-factor Gaussian and

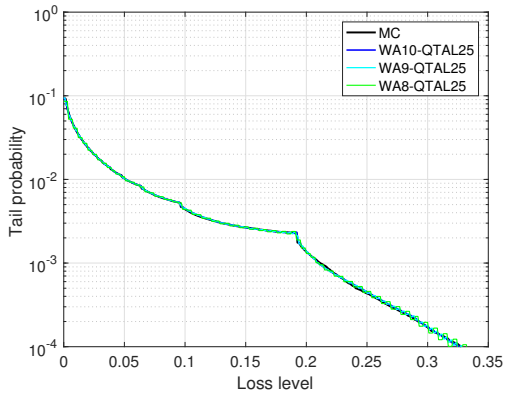


Figure 5.5: Tail probabilities for portfolio P5.

	$\alpha = 0.999$		$\alpha = 0.9999$	
Method	VaR	ES	VaR	ES
MC	0.2105	0.2613	0.3359	0.3886
WA-QTA ($m = 10$)	0.2114	0.2595	0.3267	0.3697
WA-QTA ($m = 9$)	0.2139 (1.2%)	0.2579 (0.6%)	0.3271 (0.2%)	0.3513 (5.0%)
WA-QTA ($m = 8$)	0.2168 (2.5%)	0.2651 (2.1%)	0.3301 (1.1%)	0.4177 (13.0%)

Table 5.6: VaR and ES values for portfolio P5. The relative errors at scales $m = 8,9$ with respect to the WA-QTA method at scale $m = 10$ are shown in parenthesis.

Method	$\alpha = 0.999$	$\alpha = 0.9999$
WA-QTA ($m = 10$)	25.3	23.9
WA-QTA ($m = 9$)	8.6	8.1
WA-QTA ($m = 8$)	5.5	4.0

Table 5.7: CPU time measured in seconds for obtaining the VaR and ES values for portfolio P5.

t-copula models. It is well-known that MC methods are highly demanding from the computational point of view when dealing with big sized and high-dimensional models for the estimation of VaR and ES values at high confidence levels.

These two methods are called WA-HL and WA-QTA and they are composed of two main parts. The first part is the numerical computation of the characteristic function associated to the portfolio loss variable. We tackle this part by different techniques depending on the underlying model. For the (bivariate) t-copula model we perform a double integration with Gauss-Hermite and generalized Gauss-Laguerre quadrature and call this method WA-HL, while the multi-factor Gaussian model is treated with the QTA method put forward in [39] and called WA-QTA. The characteristic function for the most challenging model, this is, the multi-factor t-copula model, is derived by conditioning on the chi-square random variable of the model and computed by applying the QTA method at each discretization point of the resulting one-dimensional integral. This last model is by far the most involved in terms of computing effort.

Once the characteristic function for the loss variable has been obtained then the second part of the procedure comes into play. This second step is the Fourier inversion method called WA, which is based on Haar wavelets and it was developed in [54, 63] to recover the CDF of the loss variable. We have improved the efficiency of the WA method by computing the coefficients of the expansion by means of an FFT algorithm and we have shown that this method outperforms the NLTI inversion method in terms of efficiency and robustness.

The overall CPU time of the numerical methods employed is impressive taking into account their size and dimension as well as the confidence levels considered. This may be the first time that a multi-factor t-copula model is considered outside the MC framework.

CHAPTER 6

Conclusions and Outlook

6.1. Conclusions

In this dissertation we have presented wavelets based numerical methods to deal with some multi-factorial problems in the computational finance area. The work is based on two wavelet methods: WA proposed in [54] and SWIFT from [67]. On one hand, we have multidimensionally extended SWIFT for option pricing. Moreover, we have applied it in the market risk area to get risk measures when featuring the liquidity problem. On the other hand, we have extended WA in the credit risk measurement to work under challenging models. We also evaluated the performance of Chebyshev wavelets [3]. In the following paragraphs we briefly summarize the most important findings in the thesis.

In Chapter 2 a mathematical background is presented to support the research in the following chapters. Using functional analysis theory, we have justified how wavelet methods can be used to approximate functions in $L^2(\mathbb{R})$. An overview on the wavelet families used throughout the thesis (Haar, Shannon and Chebyshev) is given together with their respective inversion methods.

Chapter 3 is devoted to option pricing. On one hand, it has been attempted to obtain a successful Chebyshev wavelet inversion method for pricing simple vanilla options. We could not get the payoff coefficient into a simple, hence neither competitive, expression. Since we were not satisfied with the complexity behind approximating both density and payoff coefficients using these wavelets, we decided to give up this approach considering that this type of wavelets do not seem to present any clear benefit when compared to other methods in the literature.

On the other hand, the most important contribution of the chapter is the presentation of the 2D-SWIFT method motivated by two-dimensional option pricing of European rainbow options. We have derived the corresponding pricing formulas for European options in two and higher dimensions using a more convenient approximation than the one previously followed in the literature for the cardinal sine function. A complete error analysis of the new method together with a prescription on how to select the parameters appearing in the method according to the precision required is provided. The efficiency of the method in the two-dimensional case has been tested for different types of European rainbow options of assets driven by different dynamics like GBM or JD. Basket options, spread options, options on the minimum or the maximum of two risky assets and correlation options are considered. We compared 2D-SWIFT results with the 2D-COS method, with closed-form solutions when available, with Monte Carlo simulation or with 2D-SWIFT with a large scale of approximation. The strengths of the 2D-SWIFT machinery include the domain truncation issue, the calculation of the scale of approximation and the number of coefficients used as well as the results

when dealing with extreme maturities. However, extending the method to higher dimensions, more than 4 or 5 depending on the specific kind of product, will not be useful because the curse of dimensionality appears. We have shown that 2D-SWIFT inherits the strengths of the one-dimensional technique presented in [67] for European-style options, like for instance, the a-priori knowledge of the approximation scale.

In Chapter 4 we have presented a numerical method to efficiently calculate VaR and ES values within a stochastic liquidity horizon framework. We have focussed on two aspects underlined as key regulatory changes by the Basel Committee of Banking Supervision, which are moving from VaR to ES and considering the incorporation of the risk of market illiquidity. The estimation of the risk measures with a stochastic holding period appears to be particularly challenging in terms of computational power. For the aforementioned reasons, we employ the SWIFT method, which recovers the density function of the change in value of a certain portfolio from its characteristic function. We have considered the well-known delta-gamma approach for modelling the change in value of the portfolio under normal and t-distributed risk factors, as well as the GBM, MJD and Kou models for the log-value change of the portfolio, where these two last models incorporate a jump component in the dynamics. As for the dynamics of the SLH, we considered the Bernoulli distribution, the exponential, the generalised Pareto and the inverse gamma in combination with the delta-gamma approach, and the GBM, MJD and Kou dynamics. We have carried out a detailed error analysis and we have provided a prescription on how to select the parameters of the numerical method, making this technique more robust, reliable and applicable in practice.

In Chapter 5 highly efficient numerical methods based on Haar wavelets to obtain the VaR and ES values for credit risk portfolios with exposure concentration under multi-factor Gaussian and t-copula models have been investigated. These methods are composed of two main parts. The first part is the numerical computation of the characteristic function associated to the portfolio loss variable; this part has been tackled by different techniques depending on the underlying model. For the (bivariate) t-copula model we have performed a double integration with Gauss-Hermite and generalized Gauss-Laguerre quadrature and called this method WA-HL, while the multi-factor Gaussian model has been treated with the QTA method put forward in [39] and called WA-QTA. The characteristic function for the most challenging model, this is, the multi-factor t-copula, has been derived by conditioning on the chi-square random variable of the model and computed by applying the QTA method at each discretization point of the resulting one-dimensional integral. This last model is by far the most involved in terms of computing effort. The second part is the Fourier inversion method where we have mainly used WA, from [54, 63], in order to recover the CDF of the loss variable. We have improved the efficiency of the WA method with respect to the original version [54] by computing the coefficients of the expansion by means of an FFT algorithm. We have seen that SWIFT in this situation is not as appropriate as WA and we have shown that this method outperforms the NLTI inversion method in terms of efficiency and robustness. The overall CPU time employed by WA is impressive taking into account the size of the portfolios and the dimension of the models as well as the confidence levels considered. This may be the first time that multi-factor t-copula model is considered outside the MC framework.

6.2. OUTLOOK

Research is a never-ending activity, and natural extensions of what has been presented can easily be approached.

After extending the SWIFT method, the fact that 2D-SWIFT inherits the strengths of the one-dimensional technique, opens the door to future applications of 2D-SWIFT to other financial contracts in two dimensions. For instance, previous work in [51] shows the valuation of one-dimensional barrier and early-exercise options using SWIFT, addressing important issues like boundary errors within the recursion backwards in time. Similar benefits are expected in two dimensions as well. Further, either multidimensional path dependent options pricing or Heston's stochastic volatility dynamics seem plausible to be considered.

When using SWIFT for market risk measurement, the calibration of the parameters of the models employed, as well as the stochastic liquidity horizon, with real market data would be interesting further work. Another extension of the market risk work would be the consideration of different risk factors having different liquidity horizons in line with the regulatory rules of the Basel Committee on Banking Supervision. The modelling of dependence between the stochastic holding period and log-returns of assets is a topic of interest in practice for future developments.

The credit risk research opens the door to calculate the risk contributions to the VaR and ES risk measures under the same model assumptions. Risk contributions are really time consuming when approached with MC, the standard methodology used, thus the benefit that we could gain by using WA could be really notorious.

References

[1] ABATE, J., CHOUDHURY, G. L., AND WHITT, W. An introduction to numerical transform inversion and its application to probability models. *Computational Probability* (2000), 257–323.

[2] ABATE, J., AND WHITT, W. Numerical inversion of Laplace transforms of probability distributions. *ORSA Journal on Computing 7*, 1 (1995), 36–43.

[3] ADIBI, H., AND ASSARI, P. Chebyshev wavelet method for numerical solution of Fredholm integral equations of the first kind. *Mathematical Problems in Engineering 2010* (2010).

[4] ARTZNER, P., DELBAEN, F., EBER, J., AND HEATH, D. Coherent measures of risk. *Mathematical Finance 9*, 3 (1999), 203–228.

[5] BASEL COMMITTEE ON BANKING SUPERVISION. Messages from the academic literature on risk measurement for the trading book. *Bank for International Settlements* (2011).

[6] BASEL COMMITTEE ON BANKING SUPERVISION. Fundamental review of the trading book. *Bank for International Settlements* (2012).

[7] BASEL COMMITTEE ON BANKING SUPERVISION. Fundamental review of the trading book: a revised market risk framework. *Bank for International Settlements* (2013).

[8] BASEL COMMITTEE ON BANKING SUPERVISION. Minimum capital requirements for market risk. *Bank for International Settlements* (2016).

[9] BJÖRK, T. *Arbitrage theory in continuous time.* Oxford University Press, 2009.

[10] BLACK, F., AND SCHOLES, M. The pricing of options and corporate liabilities. *Journal of Political Economy 81*, 3 (1973), 637–654.

[11] BLUHM, C., OVERBECK, L., AND WAGNER, C. *Introduction to credit risk modeling.* CRC Press, 2016.

[12] BOGGESS, A., AND NARCOWICH, F. J. *Wavelets with Fourier analysis.* Prentice Hall, 2001.

[13] BOYLE, P. P. Options: a Monte Carlo approach. *Journal of Financial Economics 4*, 3 (1977), 323–338.

[14] BRIGO, D., AND NODIO, C. A random holding period approach for liquidity-inclusive risk management. in innovations in quantitative risk management. *Springer Proceeding in Mathematics & Statistics 99* (2015).

[15] CARR, P., AND MADAN, D. Option valuation using the fast Fourier transform. *Journal of Computational Finance 2*, 4 (1999), 61–73.

[16] CATTANI, C. Shannon wavelets theory. *Mathematical Problems in Engineering* (2008).

[17] CHAN, J., AND KROESE, D. Efficient estimation of large portfolio loss probabilities in t-copula models. *European Journal of Operational Research 205* (2010), 361–367.

[18] CHEBYSHEV, P. Collected works. *Moscow: Izdatelstvo Akad. Nauk SSSR* (1955).

[19] CHUI, C. K. *An introduction to wavelets*. Academic Press, 1992.

[20] COLLDEFORNS-PAPIOL, G., AND ORTIZ-GRACIA, L. Computation of market risk measures with stochastic liquidity horizon. *Submitted for publication* (2017).

[21] COLLDEFORNS-PAPIOL, G., ORTIZ-GRACIA, L., AND OOSTERLEE, C. W. Quantifying credit portfolio losses under multi-factor models. *Submitted for publication* (2017).

[22] COLLDEFORNS-PAPIOL, G., ORTIZ-GRACIA, L., AND OOSTERLEE, C. W. Two-dimensional Shannon wavelet inverse Fourier technique for pricing European options. *Applied Numerical Mathematics 117* (2017), 115–138.

[23] CONT, R., AND TANKOV, P. *Financial modelling with jump processes*. Chapman & Hall / CRC Press, 2004.

[24] COX, J. C., ROSS, S. A., AND RUBINSTEIN, M. Option pricing: a simplified approach. *Journal of Financial Economics 7*, 3 (1979), 229–263.

[25] DAUBECHIES, I. *Ten lectures on wavelets*. Society for Industrial and Applied Mathematics, 1992.

[26] DOWD, K. *Measuring market risk*. John Wiley & Sons, 2007.

[27] DUAN, J., HÄRDLE, W. K., AND GENTLE, J. E. *Handbook of computational finance*. Springer Science & Business Media, 2011.

[28] FANG, F., AND OOSTERLEE, C. W. A novel option pricing method based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing 31*, 2 (2008), 826–848.

[29] FANG, F., AND OOSTERLEE, C. W. Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions. *Numerische Mathematik 114*, 1 (2009), 27–62.

[30] FOK, P.-W., YAN, X., AND YAO, G. Analysis of credit portfolio risk using hierarchical multifactor models. *Journal of Credit Risk 10*, 4 (2014), 45–70.

[31] FOKAS, A. S., AND SMITHEMAN, S. A. The Fourier transforms of the Chebyshev and Legendre polynomials. *ArXiv preprint arXiv:1211.4943* (2012).

[32] GEARHART, W. B., AND SHULTZ, H. S. The function sin(x)/x. *The College Mathematics Journal 21*, 2 (1990), 90–99.

[33] GHOSE, T. French mathematician Yves Meyer wins top prize for wavelet theory. *Live Science* (2017).

[34] GLASSERMAN, P. *Monte Carlo methods in financial engineering*, vol. 53. Springer Science & Business Media, 2013.

[35] GLASSERMAN, P., HEIDELBERGER, P., AND SHAHABUDDIN, P. *Efficient Monte Carlo methods for value-at-risk*. IBM Thomas J. Watson Research Division, 2000.

[36] GLASSERMAN, P., HEIDELBERGER, P., AND SHAHABUDDIN, P. Portfolio value-at-risk with heavy-tailed risk factors. *Mathematical Finance 12*, 3 (2002), 239–269.

[37] GLASSERMAN, P., KANG, W., AND SHAHABUDDIN, P. Large deviations in multifactor portfolio credit risk. *Mathematical Finance 17*, 3 (2007), 345–379.

[38] GLASSERMAN, P., AND RUIZ-MATA, J. Computing the credit loss distribution in the Gaussian copula model: a comparison of methods. *Journal of Credit Risk 2*, 4 (2006), 33–66.

[39] GLASSERMAN, P., AND SUCHINTABANDID, S. Quadratic transform approximation for CDO pricing in multifactor models. *SIAM Journal on Financial Mathematics 3*, 1 (2012), 137–162.

[40] HAAR, A. Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen 69*, 3 (1910), 331–371.

[41] HIRSA, A. *Computational methods in finance*. CRC Press, 2012.

[42] HULL, J. *Options, futures, and other derivatives*. Pearson, 2015.

[43] IN 'T HOUT, K. *Numerical partial differential equations in finance explained: an introduction to computational finance*. Palgrave Macmillan UK, 2017.

[44] KAC, M. On distributions of certain Wiener functionals. *Transactions of the American Mathematical Society 65*, 1 (1949), 1–13.

[45] KANG, W., AND SHAHABUDDIN, P. Fast simulation for multifactor portfolio credit risk in the t-copula model. *Proceedings of the 37th Conference on Winter Simulation* (2005), 1859–1868.

[46] KIRKBY, J. L. Efficient option pricing by frame duality with the fast Fourier transform. *SIAM Journal on Financial Mathematics 6*, 1 (2015), 713–747.

[47] KIRKBY, J. L. Robust barrier option pricing by frame projection under exponential Lévy dynamics. *Applied Mathematical Finance* (2017), 1–50.

[48] KOU, S. G. A jump-diffusion model for option pricing. *Management Science 48*, 8 (2002), 1086–1101.

[49] LINSMEIER, T. J., AND PEARSON, N. D. Risk measurement: an introduction to value at risk. *Department of Accountancy and Department of Finance. Technical Report 96-04. University of Illinois at Urbana-Champaign* (1996), 1–45.

[50] MALLAT, S. *A wavelet tour of signal processing*. Academic Press, 2009.

[51] MAREE, S. C., ORTIZ-GRACIA, L., AND OOSTERLEE, C. W. Pricing early-exercise and discrete barrier options by Shannon wavelet expansions. *Numerische Mathematik 136*, 4 (2017), 1035–1070.

[52] MARGRABE, W. The value of an option to exchange one asset for another. *Journal of Finance 33* (1978), 177–186.

[53] MARKOWITZ, H. Portfolio selection. *The Journal of Finance 7*, 1 (1952), 77–91.

[54] MASDEMONT, J. J., AND ORTIZ-GRACIA, L. Haar wavelets–based approach for quantifying credit portfolio losses. *Quantitative Finance 14*, 9 (2014), 1587–1595.

[55] MATHAI, A. M., AND PROVOST, S. B. *Quadratic forms in random variables*, vol. 126 of *Statistics: Textbooks and Monographs*. Dekker, 1992. Theory and Applications.

[56] MAUSSER, H., AND ROSEN, D. Beyond VaR: from measuring risk to managing risk. *Computational Intelligence for Financial Engineering, 1999. (CIFEr) Proceedings of the IEEE/IAFE 1999 Conference* (1999), 163–178.

[57] MCNEIL, A. J., FREY, R., AND EMBRECHTS, P. *Quantitative risk management: concepts, techniques and tools*. Princeton University Press, 2015.

[58] MERTON, R. C. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics 3* (1976), 125–144.

[59] MEYER, Y. *Wavelets and operators*. Cambridge University Press, 1993.

[60] MINA, J., AND ULMER, A. Delta-gamma four ways. *RiskMetrics Group LLC* (1999).

[61] MOLLER, T. On valuation and risk management at the interface of insurance and finance. *British Actuarial Journal 8*, IV (2002), 787–827.

[62] MORGAN, J., ET AL. Riskmetrics technical document. *New York* (1996).

[63] ORTIZ-GRACIA, L., AND MASDEMONT, J. J. Credit risk contributions under the Vasicek one-factor model: a fast wavelet expansion approximation. *Journal of Computational Finance 17*, 4 (2014), 59–97.

[64] ORTIZ-GRACIA, L., AND MASDEMONT, J. J. Peaks and jumps reconstruction with B-splines scaling functions. *Journal of Computational and Applied Mathematics 272* (2014), 258–272.

[65] ORTIZ-GRACIA, L., AND OOSTERLEE, C. W. Robust pricing of European options with wavelets and the characteristic function. *SIAM Journal on Scientific Computing 35*, 5 (2013), B1055–B1084.

[66] ORTIZ-GRACIA, L., AND OOSTERLEE, C. W. Efficient VaR and expected shortfall computations for nonlinear portfolios within the delta-gamma approach. *Applied Mathematics and Computation 244* (2014), 16–31.

[67] ORTIZ-GRACIA, L., AND OOSTERLEE, C. W. A highly efficient Shannon wavelet inverse Fourier technique for pricing European options. *SIAM Journal on Scientific Computing 38*, 1 (2016), B18–B143.

[68] OWEN, A., BRYERS, J., AND BUET-GOLFOUSE, F. Hermite approximations in credit portfolio modeling with probability of default-loss given default correlation. *Journal of Credit Risk 11*, 3 (2015), 1–20.

[69] PAPAPANTOLEON, A. An introduction to Lévy processes with applications in finance. *Lecture notes* (2008).

[70] PYKHTIN, M. Portfolio credit risk multi-factor adjustment. *Risk Magazine 17*, 3 (2004), 85–90.

[71] QUINE, B. M., AND ABRAROV, S. M. Application of the spectrally integrated Voigt function to line-by-line radiative transfer modelling. *Journal of Quantitative Spectroscopy & Radiative Transfer 244* (2013), 37–48.

[72] REED, M., AND SIMON, B. *Methods of modern mathematical physics I: functional analysis*. Academic Press New York, 1972.

[73] ROSS, S. M. *A first course in probability*. Pearson Education International, 2009.

[74] RUITER, M. J., AND OOSTERLEE, C. W. Two-dimensional Fourier cosine series expansion method for pricing financial options. *SIAM Journal on Scientific Computing 34*, 5 (2012), B642–B671.

[75] RUTKOWSKI, M., AND TARCA, S. Regulatory capital modeling for credit risk. *International Journal of Theoretical and Applied Finance 18*, 5 (2015).

[76] SCHLOEGL, L., AND O'KANE, D. A note on the large homogeneous portfolio approximation with the student-t copula. *Finance and Stochastics 9*, 4 (2005), 577–584.

[77] SCHOUTENS, W. *Lévy processes in finance: pricing financial derivatives*. Hohn Wiley & Sons Ltd., 2003.

[78] SHARPE, W. F., ALEXANDER, G. J., AND BAILEY, J. V. *Investments*, vol. 6. Prentice-Hall Upper Saddle River, NJ, 1999.

[79] STENGER, F. *Handbook of sinc numerical methods*. Chapman & Hall / CRC Numerical Analysis and Scientific Computing. CRC Press, Boca Raton, FL, 2011.

[80] TESCHL, G. Topics in real and functional analysis. *Preprint* (2014).

[81] VON SYDOW, L., ET AL. Benchop - the benchmarking project in option pricing. *International Journal of Computer Mathematics 92* (2015), 2361–2379.

[82] WAGNER, E. I. On the application of Shannon wavelet inverse Fourier techniques: an extension to Asian option valuation and European option pricing under the SABR model. *Master thesis TU Delft* (2017).

[83] ZHANG, B., AND OOSTERLEE, C. W. Efficient pricing of European-style Asian options under exponential Lévy processes based on Fourier cosine expansions. *SIAM Journal on Financial Mathematics 4*, 1 (2013), 399–426.