

**Measuring the risk of financial portfolios with
nonlinear instruments and non-Gaussian risk
factors**

Roberto Bustreo

A Thesis submitted for the degree of Doctor of Philosophy

Department of Mathematical Sciences
Brunel University

October 2013

Abstract

The focus of my research has been computationally efficient means of computing measures of risk for portfolios of nonlinear financial instruments when the risk factors might be possibly non-Gaussian. In particular, the measures of risk chosen have been Value-at-Risk (VaR) and conditional Value-at-Risk (CVaR). I have studied the problem of computation of risk in two types of financial portfolios with nonlinear instruments which depend on possibly non-Gaussian risk factors:

1. Portfolios of European stock options when the stock return distribution may not be Gaussian;
2. Portfolios of sovereign bonds (which are nonlinear in the underlying risk factor, i.e. the short rate) when the risk factor may or may not be Gaussian.

Addressing both these problems need a wide array of mathematical tools both from the field of applied statistics (Delta-Gamma-Normal models, characteristic function inversion, probability conserving transformation) and systems theory (Vasicek stochastic differential equation model, Kalman filter). A new heuristic is proposed for addressing the first problem, while an empirical study is presented to support the use of filter-based models for addressing the second problem. In addition to presenting a discussion of these underlying mathematical tools, the dissertation also presents comprehensive numerical experiments in both cases, with simulated as well as real financial market data. Backtesting is used to confirm the validity of the proposed methods.

Different parts of the work presented in this dissertation has been submitted to peer reviewed journals in terms of two papers:

- Chapter 2: “*Measuring the risk of a nonlinear portfolio with fat tailed risk factors*”

through probability conserving transformation” submitted to IMA Journal of Management Mathematics;

- Chapter 3: “*Value-at-Risk for fixed income portfolios: A Kalman filtering approach*” submitted to IMA Journal of Management Mathematics.

List of Abbreviations

BIS	Bank of International Settlements
CDF	Cumulative Distribution Function
CIR	Cox-Ingersoll-Ross
CVaR	Conditional Value-at-Risk
DFT	Discrete Fourier Transform
DSP	Digital Signal Processing
ES	Expected Shortfall
EVT	Extreme Value Theory
FFT	Fast Fourier Transform
FOCUS	Financial and Operating Combined Uniform Single
GEV	Generalized Extreme Value
GPD	Generalized Pareto Distribution
HS	Historical Simulation
ICA	Independent Component Analysis
KF	Kalman Filter
IDFT	Inverse Discrete Fourier Transform
MC	Monte Carlo
MRAE	Mean of the Relative Absolute Error for the in-sample
MRAE*	Mean of the Relative Absolute Error for the out-of-sample
PCA	Principal Component Analysis
PDF	Probability Density Function
PFA	Prime Factor Algorithms
POT	Peak Over Threshold
RAE	Relative Absolute Error

SEC Securities Exchange Commission

UNCR Uniform Net Capital Rule

USD United States Dollar

VaR Value-at-Risk

VC Variance/Covariance

Contents

1	Introduction	1
1.1	Managing the risk for financial institutions	1
1.2	A short history of VaR	4
1.3	Four scenarios for a financial portfolio	7
1.4	Fitting the term structure and computing VaR for fixed income portfolios . .	10
1.5	Outline of the thesis	12
2	Mathematical tools for modeling the risk in stock portfolios	13
2.1	Introduction	13
2.2	Linear portfolio with normal returns	14
2.3	Historical simulation method and variance/covariance method	15
2.4	Linear portfolio with non-normal returns	16
2.4.1	Extreme value theory and risk measures	17
2.4.2	Modeling fat tails with Student's t distribution	22
2.4.3	Copula functions	25
2.5	Nonlinear portfolio with normal returns	28
2.5.1	Delta-Gamma-Normal model	28
2.5.2	Digital signal processing	34
2.5.3	Discrete Fourier transform (DFT)	35
2.5.4	The fast Fourier transform (FFT)	39
2.5.5	The characteristic function and the inversion integral	42
2.5.6	Characteristic function inversion using FFT	44
2.6	Nonlinear portfolio with non-normal returns	47
2.6.1	Probability conserving transformation	47

2.6.2	Delta-Gamma-Q for assets with non-normal returns	50
2.6.3	Nonparametric density estimation	51
2.7	Summary	54
3	Mathematical tools for modeling risk in fixed income portfolios	55
3.1	Introduction	55
3.2	Preliminary definitions for interest rate models	56
3.3	The Kalman filter	57
3.4	A linear system: The Vasicek model	58
3.4.1	Single factor Vasicek model	59
3.4.2	Two-factor Vasicek model	60
3.5	A nonlinear system: The CIR model	62
3.5.1	A single factor CIR model	62
3.5.2	Two-factor CIR model	63
3.6	Summary	65
4	Backtesting: Methodology	66
4.1	Introduction	66
4.2	The unconditional coverage test	67
4.3	The independence test	68
4.4	The conditional coverage test	69
4.5	Summary	69
5	Nonlinear stock portfolios with fat-tailed risk factors: Results	70
5.1	Introduction	70
5.2	Numerical experiments with simulated data	70
5.3	Numerical experiments with real data	79
5.3.1	Data	79
5.3.2	Test portfolio	80
5.3.3	Applying the Delta-Gamma-dP Normal method: Results	83
5.3.4	Applying the Delta-Gamma-Q method: Results	84
5.4	Summary	87

6 Nonlinear fixed income portfolios: Results	88
6.1 Introduction	88
6.2 Numerical experiments with simulated data	88
6.2.1 Simulation using the two-factor Vasicek model	92
6.2.2 Sensitivity study for the two-factor Vasicek model	95
6.2.3 Simulation using the two-factor CIR model	97
6.2.4 Sensitivity study for the two-factor CIR model	99
6.2.5 Comments on the simulation experiments	101
6.3 Numerical experiments with real data	102
6.3.1 Data	103
6.3.2 Results	103
6.4 Summary	106
7 Summary of contributions and future research	107
Appendix A Function Documentation	109

List of Figures

5.1	FTSE100 returns histogram for the period 10/07/2009 – 10/06/2011.	81
5.2	Graphical representation of portfolio composition in time.	82
6.1	A simulated interest rate r and its Kalman filtering prediction.	94

List of Tables

4.1	Non rejection regions for Kupiec's test.	68
5.1	95% confidence intervals for Monte Carlo 95% and 99% VaR.	74
5.2	Parameters range.	74
5.3	Values of 95% and 99% VaR for the four experiments reported.	75
5.4	Values of 95% and 99% CVaR over four combination of parameters.	75
5.5	Percentage errors for the four experiments carried out.	76
5.6	Net computational times for the computation of VaR and CVaR.	78
5.7	Summary statistics of the FTSE100 returns.	80
5.8	Portfolio composition in the different intervals.	82
5.9	Considered strike prices and maturities.	82
5.10	Descriptive statistics of the ΔS	83
5.11	Summary of test results for the Delta-Gamma-dP Normal model.	84
5.12	Summary of test results for the Delta-Gamma-Q modeling of real data.	87
6.1	Coefficients used for the simulation of the two-factor Vasicek model.	92
6.2	Features of the bonds providing the measurement values.	92
6.3	Features of the bonds included in the simulated portfolio.	93
6.4	Estimated parameters for the subset considered.	93
6.5	Summary of test results for the considered simulation.	94
6.6	New set of coefficients used for the simulation of the two-factor Vasicek model.	95
6.7	Estimated parameters for the new subset considered.	96
6.8	Summary of test results for the considered new simulation.	96
6.9	Coefficients used for the simulation of the two-factor CIR model.	97
6.10	Features of the bonds providing the measurement values.	97

6.11	Estimated parameters for the subset considered.	98
6.12	Summary of test results for the considered simulation.	98
6.13	New set of coefficients used for the simulation of the two-factor CIR model. .	99
6.14	Estimated parameters for the new subset considered.	100
6.15	Summary of test results for the considered new simulation.	100
6.16	Estimated parameters for two-factor Vasicek model.	104
6.17	Estimated parameters for two-factor CIR model.	104
6.18	Summary of test results for the experiments with real data.	105

Acknowledgements

I would like to thank my supervisor Dr Paresh Date for his support and for all the time he has dedicated to provide useful advice. He has proved a truly inspiring figure throughout my PhD programme, helping me to overcome some challenging times in my research.

I would also like to thank my family and all my friends for the continuous support during all these years.

Chapter 1

Introduction

This chapter provides motivation to address the problems discussed in the thesis. In the first section, basic concepts about measuring risk for financial portfolios are introduced. Next, a short history of Value-at-Risk is provided. This is followed by providing details on the existing modeling paradigm and the proposed model on the two type of portfolios considered.

1.1 Managing the risk for financial institutions

Financial institutions need to monitor and effectively manage market risk. Quantitative risk measures have become crucial management instruments for portfolio managers. Risk managers and market regulators have to control their risks or to appropriately allocate their capital. Value-at-Risk (VaR) has been chosen by the Basel Committee on Banking Supervision in Basel II as the standard risk measure for financial risk managers (see Basel Committee (2006) and Qian and Gerlach (2011) for details). It measures the worst expected loss under normal market conditions over a specific time interval at a given confidence interval. It may be seen as the “best of worst cases scenario” and it therefore systematically underestimates the potential losses associated with the specified level of probability.

To define VaR, consider a real-valued random variable \mathcal{X} on a probability space (Ω, \mathcal{A}, P) that expresses the random profit or loss of some asset or portfolio, its cumulative distribution function (cdf) $F_{\mathcal{X}}(x) = \Pr[\mathcal{X} \leq x]$ and set a confidence level $\alpha \in (0, 1)$. The quantity

$q_\alpha(\mathcal{X}) = \inf\{x | F_{\mathcal{X}}(x) \geq \alpha\}$ is the α -quantile of \mathcal{X} . Then VaR at a level α is given by:

$$VaR_\alpha = -q_{1-\alpha}(\mathcal{X}). \tag{1.1}$$

VaR is one number that encapsulates the risk of the entire portfolio of financial assets. It assesses the potential shortfall in a portfolio and can be valid for many different financial instruments to compute the total risk that is faced by an investor. Therefore senior managers are able to make decisions based on Value-at-Risk. The Basel Accord and the market risk amendments contain the international guidelines implemented by the national agencies,¹. Value-at-Risk has become the industry yardstick for measuring market risk mainly because it is endorsed in the Basel Accord and partly this is because of its insightful appeal.

For a given portfolio, VaR expresses the worst expected loss that a firm can experience over a given time interval under normal market conditions, at a given confidence level. It measures this risk by using statistical and simulation models designed to capture the volatility of assets in a financial portfolio. In statistical terms, it is a percentile of the distribution for changes in portfolio value.

Suppose a portfolio were to remain untraded for a certain period of time. Let us measure time in trading days, and let t_0 be the current time. The portfolio's current market value is $\Pi(t_0)$. Its market value $\Pi(t_1)$, in one trading day, is unknown. It is a random variable and like with any random variable, a probability distribution can be associated to this conditional to time t_0 . Furthermore, the portfolio's market risk can be measured using some parameter of that conditional distribution.

If we consider the portfolio variation $\Delta\Pi = \Pi(t_1) - \Pi(t_0)$ during the holding period, then Value-at-Risk is the solution to a nonlinear equation:

$$\Pr[\Delta\Pi \leq -VaR_\alpha] = 1 - \alpha, \tag{1.2}$$

where α is the confidence level. An alternative understanding is that in the long term we expect losses exceeding VaR with frequency $1 - \alpha$. For $\alpha = 99\%$, we assume losses exceeding

¹The Basel Accord and related documents are available from the Bank of International Settlements (www.bis.org).

VaR one out of every 100 days.

VaR has received criticism by Artzner et al. (1999), Acerbi and Tasche (2002) and Szegö (2002) for not being a coherent measure of risk. This is because of the fact that it does not generally fulfill the axiom of sub-additivity, apart from the case of linear portfolios with normally distributed risk factors and similar special cases. A coherent risk measure is defined in Artzner et al. (1999) as follows: let us consider a set V of real-valued random variables. A function $\rho: V \rightarrow \mathbb{R}$ is called a coherent risk measure if it is:

1. Translation invariant: $X \in V, a \in \mathbb{R} \Rightarrow \rho(X + a) = \rho(X) - a$;
2. Sub-additive: $X, Y, X + Y \in V \Rightarrow \rho(X + Y) \leq \rho(X) + \rho(Y)$;
3. Positively homogeneous: $X \in V, h > 0, hX \in V \Rightarrow \rho(hX) = h\rho(X)$;
4. Monotonous: $X, Y \in V, Y \geq X \Rightarrow \rho(Y) \leq \rho(X)$.

VaR is not a coherent risk measure because it does not generally fulfill the axiom of sub-additivity. VaR is sub-additive (total VaR is always less or equal than the sum of partial VaR's) in the case of normal distributed returns of a portfolio and under similar special cases, see Artzner et al. (1999). This property specifies the fact that a portfolio made of subportfolios will risk an amount which is at most the sum of the divided amounts risked by its subportfolios. This axiom captures the essence of how a risk measure should behave under the composition/addition of portfolios.

A sub-additive alternative to VaR is the conditional Value-at-Risk (CVaR), defined as the conditional expectation of loss for losses beyond the VaR level. CVaR and its minimization formula were first developed in Rockafellar and Uryasev (2000), where the authors demonstrated the numerical efficiency of their proposed procedure through several case studies, including portfolio optimization and option hedging. The conditional Value-at-Risk at level α can be defined as:

$$CVaR_\alpha = -\mathbb{E}[\mathcal{X} | \mathcal{X} \leq q_{1-\alpha}(\mathcal{X})]. \quad (1.3)$$

The Expected Shortfall (ES) at level α is closely related but not coincident to the notion of conditional Value-at-Risk, and it is defined (see, e.g. Acerbi (2002)) as:

$$ES_\alpha = -\frac{1}{1-\alpha} \int_0^{1-\alpha} q_l(\mathcal{X}) dl.$$

Conditional Value-at-Risk is not a coherent measure in general; it coincides with ES_α (and it is therefore coherent) only under suitable conditions such as continuity of the probability distribution function $F_{\mathcal{X}}(x)$. Computation of CVaR, in effect, requires computing VaR at several different confidence levels and hence is typically computationally more intensive.

The rest of this chapter is structured as follows. Section 1.2 includes a short history of Value-at-Risk. In section 1.3, the four possible scenarios for the computation of VaR and CVaR are listed. In section 1.4 some approaches for fitting the term structure of interest rates and some approaches to VaR computation for fixed income portfolios are provided.

1.2 A short history of VaR

The expression “Value-at-Risk” was not broadly used before the mid 1990s but its foundations lie further back in time (for a more substantial treatment refer to Holton (2002)). The origins of VaR lie in the risk-return trade-off and optimization paradigm introduced by Markowitz and others in 1950s. The spotlight on market risks and the outcomes of the co-movements in these risks are crucial to how VaR is computed. The drive for the use of VaR measures came from the crises that overwhelmed financial service firms over time and the regulatory reactions to these crises. The first regulatory capital requirements for banks were performed in the aftermath of the Great Depression and the bank crashes of the era, when the Securities Exchange Act launched the Securities Exchange Commission (SEC) and demanded banks to keep their borrowings under 2,000% of their equity capital (see, e.g. Damodaran (2007)).

In subsequent decades, banks developed risk measures and control devices to certify that they met these capital requirements. In the early 1970s the advent of derivative markets and floating exchange rates increased the market risk. Hence the capital constraints were enhanced and expanded in the SEC’s Uniform Net Capital Rule (UNCR) that was circulated in 1975, which classified the financial assets that banks held into twelve classes, based upon risk, and entailed different capital requirements for each, spanning from 0% for short term

treasuries to 30% for equities. Banks were demanded to report on their capital computations in quarterly statements, named Financial and Operating Combined Uniform Single (FOCUS) reports.

In 1980 the SEC tied the capital requirements of financial service firms to the losses that would be incurred, with 95% confidence, over a thirty-day interval in numerous security classes; this marked the first regulatory measures using Value-at-Risk. Historical returns were used to calculate the possibility of losses. It was obvious that SEC was insisting on financial service organizations to start the process of estimating one month 95% VaR and hold sufficient capital to cover potential losses. Around that period there was a growing need for more complex and sensible risk control measures since trading portfolios of investment and commercial banks were becoming larger and more volatile. In internal documents, Ken Garbade at Banker's Trust, presented advanced measures of Value-at-Risk in 1986 for the firm's fixed income portfolios, based upon the covariance in yields on bonds of different maturities.

During the late 1980s, J.P.Morgan elaborated a company-wide VaR system, that modeled several hundred risk factors using a covariance matrix that was updated quarterly from historical data. Every day, trading units would give details by e-mail about their positions' deltas with respect to each of the risk factors. These were combined to express the whole portfolio's value as a linear polynomial of the risk factors. The portfolio's standard deviation value was computed. One of various VaR metrics employed was the one-day 95% USD VaR, which was assessed assuming that the portfolio's value was normally distributed. J.P.Morgan's first use of VaR was an attempt to replace a bulky system of notional market risk limits with a simple system of VaR limits.

From 1990, VaR numbers and P&L's list were brought together in a report for the 4:15 PM daily Treasury meeting in New York. Those documents, with comments from the Treasury group, were forwarded to Chairman Weatherstone. Till Guldimann was one of the architects of the new VaR measure; his career with J.P. Morgan had placed him to help developing and encourage the use of VaR within the company. During the mid 1980's, he was in charge for the asset/liability analysis of the firm. He developed, together with other professionals, notions that would be used in the VaR measure. Afterwards as chairman of

the market risk committee of the company, he promoted the VaR measure internally and in its next position he could promote the VaR measure outside the organization. Guldemann, in 1990, took responsibility for Global Research, supervising research activities to support marketing to institutional clients; here he managed an annual research conference for clients. Guldemann during the conference in 1993, whose theme was the risk management, gave the crucial address and arranged for a demonstration of J.P.Morgan's VaR system. The interest generated was significant: clients asked if they might buy or lease the system. After many catastrophic losses related to the use of derivatives and leverage between 1993 and 1995, ending with the collapse of Barings, the British investment bank, as a result of unauthorized trading in Nikkei futures and options by a young trader in Singapore, companies were ready for more sophisticated risk measures.

In 1995 J.P.Morgan provided free public access over the internet to data (variances and covariances across various security and asset classes) that was used internally to manage risk for almost a decade and allowed software makers to develop software to measure risk. It titled the service "RiskMetrics" and included an exhaustive technical document as well as a covariance matrix for several hundred key factors, that was updated daily. The expression Value-at-Risk was used to express the risk measure that emerged from the data: it found a ready audience with commercial and investment banks, and the regulators supervising them, who welcomed its intuitive appeal. The timing for RiskMetrics was optimal, given the time of worldwide concerns about derivatives and leverage. In the last decade, VaR has turned into the conventional measure of risk exposure in financial service companies and has even started to find recognition in non-financial service business.

Following the issuing of RiskMetrics and the extensive adoption of VaR measures, there was somewhat of a backlash against it. This has come to be called the "VaR debate". Criticisms followed three themes:

- that different VaR implementations produced inconsistent results;
- that, as a measure of risk, VaR is conceptually flawed;
- that widespread use of VaR entails systemic risks.

There are criticisms suggesting that VaR measures were conceptually flawed. One such critic is Taleb (1997): *"The condensation of complex factors naturally does not just affect the ac-*

curacy of the measure. Critics of VaR (including the author) argue that simplification could result in such distortions as to nullify the value of the measurement. Furthermore, it can lead to charlatanism: lulling an innocent investor or business manager into a false sense of security could be a serious breach of faith. Operators are dealing with unstable parameters, unlike those of the physical sciences, and risk measurement should not just be understood to be a vague and imprecise estimate. This approach can easily lead to distortions. The most nefarious effect of the VaR is that it has allowed people who have never had any exposure to market risks to express their opinion on the matter”.

Despite these criticisms, VaR remains an extremely popular and widely adapted measure of risk, due to its conceptual simplicity and regulatory acceptance. Furthermore, while Basel committee on banking supervision has recently proposed to replace VaR with CVaR, CVaR is far harder to backtest properly.

1.3 Four scenarios for a financial portfolio

In general, the computation of VaR or CVaR for a financial portfolio leads to one of the four possible scenarios:

1. The portfolio is linear in the underlying risk factors and the distribution of risk factors can be approximated by a Gaussian distribution. This is a case when the portfolio consists of stocks or stock futures, risk factors are stock returns only and the confidence level is $\alpha = 0.95$ or lower. For higher α values, distributional assumption of Gaussianity is inappropriate, as has been proven time and again in literature. Nevertheless, Gaussian approach remains popular due to its simplicity and is often used as a starting point or as benchmark; see, *e.g.* Albanese et al. (2004) and Alexander (2008).
2. The second scenario is the one which involves a linear portfolio with non-Gaussian risk factors. This means that the distribution of the portfolio net worth may be different from that of the underlying risk factors. Typically, this scenario requires a Monte Carlo simulation to compute the quantiles of the portfolio value distribution. There are two distinct choices of risk factor distributions followed in literature; the approach based on extreme value theory is followed in Embrechts et al. (1999) and Gilli and K ellezi (2006), among others. On the other hand, different thick tailed distributions which do

not belong to the class of extreme value distributions have also been used, including the use of mixture of normals in Zangari (1996a) and Duffie and Pan (1997), Student's t as outlined in Alexander (2008) and the use of mixture of elliptic distributions in Kamdem (2003). An alternative approach is provided by the copula-based models (see e.g. Nelsen (2006) for an introduction to copulas) in Rahman et al. (2011) and in Sak and Haksöz (2011), among others.

3. The third scenario is when the portfolio has Gaussian risk factors but the portfolio value is a nonlinear function of risk factors, e.g. due to presence of derivative instruments such as options. A very common way of dealing with this scenario is using a quadratic (or Delta-Gamma) approximation to the portfolio value, since quantiles of a quadratic form of jointly Gaussian random variables can be efficiently computed using Fourier inversion integral; see, e.g. Glasserman (2003) for a detailed treatment. Other approaches include Cornish-Fisher expansion as discussed in Jaschke (2001) and Zangari (1996c), Johnson transformation discussed in Zangari (1996b), Solomon-Stephens approximations discussed in Britten-Jones and Schaefer (1999), saddle point approximation outlined in Feuerverger and Wong (2000) and other Fourier inversion based methods described in Albanese et al. (2004) and Rouvinez (1997). Monte Carlo simulation with the quadratic approximation of the portfolio value (so-called partial Monte Carlo simulation) is computationally typically simpler than full Monte Carlo, as the computational load of pricing nonlinear instruments for given values of risk factors may be far greater than computing quadratic forms of risk factors; see Pritsker (1997) for an application of partial Monte Carlo. Comparisons of accuracy and efficiency among mentioned methodologies are carried out in Mina and Ulmer (1999) and in Castellacci and Siclari (2003). Different methods for dealing with Gaussian risk factors, nonlinear portfolio scenarios are compared in Britten-Jones and Schaefer (1999) and in Pichler and Selitsch (1999).
4. The last and the most general scenario involves a portfolio with nonlinear instruments which depend on non-Gaussian risk factors. Any financial portfolios involving derivative instruments on stocks (such as hedge funds or absolute return funds) fall under this category when the stock return distribution cannot be satisfactorily approximated with a Gaussian distribution. Unlike the normal distribution, the quadratic forms of fat-tailed distributions are not amenable to easy evaluation in general. Apart from

full Monte Carlo or partial Monte Carlo simulation (using quadratic approximation of nonlinear functions), few computationally cheap alternatives currently exist for this scenario; see Glasserman et al. (2000) and El-Jahel et al. (1999) for some approaches to this problem.

The financial crisis of 2008 has brought back to researchers and managers the fact that normality assumption for risk factors is not realistic. Gaussian distribution-based models are appealing because of their simple implementation, but they fail to explain real world risk factor characteristics such as fat-tailedness and skewness. In Nozari et al. (2010), Sheikh and Qiao (2010) and Stoyanov et al. (2011) authors highlight the fact that empirical research on financial returns carried out since 1950s leads to the need to consider several phenomena including fat tails, skewness and serial correlation. Therefore models that rely on the assumption of normal distributions need to be relaxed in order to avoid underestimating risk. Parsimonious Gaussian-based models are often employed by practitioners despite empirical evidence to the contrary. However, especially during periods of high turbulence in the financial markets, such as the one in the recent crisis, Gaussianity assumption can lead to severe underestimation of risk.

The focus in this thesis is to provide heuristics to calculate VaR and CVaR for the last scenario mentioned above, i.e. for nonlinear portfolios with non-normal risk factors. We consider two different cases: stock portfolios and fixed income portfolios.

In the case of stock portfolios, the proposed heuristic is computationally cheaper than a full or partial Monte Carlo with fat tailed distributions, yet has a potential to be far more accurate than normal distribution-based evaluation. The novelty of our heuristic lies in offering an alternative with a potentially intermediate level of complexity and accuracy between the two extremes: between carrying out a full Monte Carlo simulation by sampling from fat tailed distributions followed by multiple pricing function evaluations on one hand and evaluating a single, one dimensional integral assuming a normal distribution and a quadratic approximation of the portfolio value on the other hand. This approach rests on transforming the problem with non-Gaussian marginals into Gaussian ones via a probability conserving transformation, as used in Sornette et al. (2000). Unlike mean-variance based Gaussian approximation of marginal densities, probability conserving transformation retains the tail behavior which is crucial in the computation of risk measures. This transformation is then

followed by construction of a joint Gaussian density and construction of a quadratic form in the jointly Gaussian transformed variables. One can then use the Delta-Gamma-Normal method for nonlinear portfolios with Gaussian risk factors for the computation of risk measures. Further, one can re-use the existing code for the fast Fourier transform (FFT) based evaluation of CVaR using Delta-Gamma-Normal method, by using the proposed heuristic of transforming the risk factors into Gaussian factors.

In the case of fixed income portfolios, we concentrate on portfolios of sovereign (or default risk-free) zero coupon bonds, with the changes in the interest rate as the sole source of risk. We model the interest rate changes via Kalman filter, which provides a highly parsimonious and tractable model for Monte Carlo simulation using instantaneous short rate. Bond prices are nonlinear functions of risk factors in this case, and the risk factors are non-Gaussian when a non-Gaussian (and often more realistic) short rate model is used.

1.4 Fitting the term structure and computing VaR for fixed income portfolios

Several mathematical tools have been applied to fit the term structure of interest rates. We will provide a literature review of some of the mathematical tools here, although detailed discussion of Kalman filter is postponed to chapter 3. A first approach is using tools that smooth the yield term structure. In his pioneering work, McCulloch (1975) used cubic splines, followed by Vasicek and Fong (1982) that used exponential splines. Adams and Van Deventer (1994) more recently suggested the maximum smoothing method and Fisher et al. (1995) proposed a smoothing splines endowed with a penalty that reduced the effective number of parameters. A different approach by Litterman and Scheinkman (1991) involving principal component analysis (PCA) is able to explain up to 95% of bond yield variance employing only three factors. In this thesis we consider exponential affine term structure models, which is a class of models often employed to understand the dynamics of interest rates. Seminal work on this class has been provided by Vasicek (1977), Cox et al. (1985) where, respectively, Gaussian and non-Gaussian single factor models are proposed. As pointed out in Brigo and Mercurio (2006) there is evidence that the single factor models fail to explain accurately some features of the term structure and it became necessary to

extend it. A generalization of the single-factor models to higher dimensions have been presented e.g. in Chaplin and Sharp (1993), Chen (1995) and Duffie and Kan (1996), while Jegadeesh and Pennacchi (1996), Babbs and Nowman (1999) and De Rossi (2004) focused on Gaussian multi-factor models. Babbs and Nowman (1999) found that the two-factor model performed better for the term structure of interest rates in bond market in the US and other nine countries. In Date and Wang (2009) the authors state that two-factor Gaussian model performs better than a single-factor model and a three-factor model when comparing out-of-sample one-step ahead forecasting. A two factor model seems to offer a good compromise between the difficulty of calibrating a three factor model and poor accuracy of a one factor model. Hence a two factor model has been chosen for the current analysis.

Computation of Value-at-Risk for fixed income portfolios has attracted a lot of interest and several approaches are available in literature. In Smith (2009) the author calculates VaR for a three zero-coupon bond portfolio by focusing on yield volatilities and correlations. The *duration*, usually linking changes in yield to maturity on an single bond to variations in its market value, is combined to non-parallel yield curve shifts to provide VaR for each individual bond. Using all the individual bond VaR with the respective correlations for the changes in the yield between bonds one can obtain the VaR for the portfolio. Fixed-income Brazilian portfolios performances are assessed in Carvalho and Daumas (2010) computing several measures including VaR and CVaR, which are calculated assuming normal distribution of interest rate variations. The authors conclude that the choice criteria based on minimum VaR and CVaR achieves satisfactory results. In Darbha (2001) VaR for fixed income portfolios is computed through extreme value theory. The VaR estimates so obtained are then compared to estimates found using the variance/covariance method and the historical simulation method, concluding that the extreme value method provides the best VaR estimates. A methodology often employed to assess VaR for fixed income portfolios is scenario simulation from principal components of the yield curve. In Jamshidian and Zhu (1997) the fixed income bond prices are modeled using a small number of risk factors, the joint distribution is approximated using a multivariate discrete distribution and VaR is computed by selecting the appropriate quantile of the discrete cumulative distribution function. This implementation presents two shortcomings, mentioned in Gibson and Pritsker (2000): a poor choice of risk factors could lead to an underestimation of VaR and a discrete approximation to the distribution of risk factors could provide an inadequate VaR based on a

poorly modeled distribution of portfolio value. The author suggested two efficient solutions to overcome these issues: select risk factors able to explain changes in portfolio value and work with a continuous distribution of extracted risk factors. In Fiori and Iannotti (2007) authors apply PCA to Monte Carlo simulation considering the non-normality of historical observations. Another approach is suggested in Chen et al. (2007) where authors use independent component analysis (ICA), a tool utilized in sound engineering, to calculate VaR for foreign exchange rate portfolios.

An alternative way to the above mentioned works on VaR computation for fixed income portfolios is presented in this thesis. The instantaneous interest rate is modeled using a two-factor Vasicek model and a two-factor Cox-Ingersoll-Ross model. After calibration of the models, the changes in bond prices are simulated by using a Kalman filter. The performance of this methodology is demonstrated with calibration and backtesting on a real portfolio of US government bonds.

1.5 Outline of the thesis

The rest of this thesis is organized as follows:

- Chapter 2 describes the mathematical tools utilized and/or useful to understand the four envisaged scenarios for the first type of financial portfolio;
- Chapter 3 provides the mathematical tools we used for modeling the fixed income portfolios;
- Chapter 4 presents the backtesting tools, including tests for unconditional and conditional coverage in predicting tail losses;
- Chapters 5 and 6 report the results for nonlinear portfolios with fat-tailed risk factors and fixed income portfolios, respectively. Both these chapters contain numerical experiments with simulated data as well as with real financial data;
- Chapter 7 presents a summary of the work done and comments on the future research directions.

Chapter 2

Mathematical tools for modeling the risk in stock portfolios

2.1 Introduction

This chapter describes some models which are useful in understanding the behavior of the four possible scenarios mentioned in the previous chapter:

- A portfolio which is linear in the risk factors and the risk factors are multivariate normal;
- A portfolio which is linear in the risk factors, but the risk factors are non-normal. We consider the use of extreme value theory, Student's t distribution and copula functions for modeling non-normality;
- A portfolio which is a nonlinear function of risk factors and the risk factors are multivariate normal. In this case, we discuss the use of second order Taylor series expansion of the portfolio value function, which has a closed-form characteristic function, combined with a fast Fourier transform inversion. This method is also called the Delta-Gamma-Normal method;
- Portfolio which is a nonlinear function of risk factors, and the risk factors are non-normal. We investigate a new heuristic which we term Delta-Gamma-Q method, which combines the Delta-Gamma-Normal method along with the probability conserving

transformation, described later in subsection 2.6.1.

We will use ‘returns’ and ‘risk factors’ interchangeably where there is no risk of confusion. Each section of this chapter presents an introduction to the theory involved in the method applied. Section 2.2 mentions the multivariate normal setting to model linear portfolio with normal returns. Section 2.3 describes the historical simulation method and the variance/covariance method. Section 2.4 contains three methods to model linear portfolio with non-normal returns. Section 2.5 introduces a very common method for modeling nonlinear portfolio with normal returns. Section 2.6 presents our innovative method to model nonlinear portfolios with fat-tailed risk factors: the Delta-Gamma-Q method.

2.2 Linear portfolio with normal returns

Under the assumptions of a linear portfolio and normal returns, the relative changes vector in the market risk factors, \mathcal{F} , are distributed as a multivariate normal random variable with mean μ and covariance matrix C . Suppose now that the change in the portfolio’s value, as described in Albanese et al. (2004) and in Wiberg (2002), can be approximated by:

$$\Delta\Pi \approx \Xi + \Delta^T \mathcal{F},$$

where Ξ and Δ represent a constant and the proportion of the stocks in the portfolio, respectively. The difference $\Delta\Pi - \Xi$ is then a sum of normal random variables and it is therefore normal. Its mean and variance are given by:

$$\begin{aligned} \mu(\Delta\Pi - \Xi) &= \mathbb{E}[\Delta\Pi - \Xi] = \Delta^T \mu, \\ \text{var}(\Delta\Pi - \Xi) &= \mathbb{E}[\Delta^T \mathcal{F} \mathcal{F}^T \Delta] - \Delta^T \mu \mu^T \Delta = \Delta^T C \Delta. \end{aligned}$$

Hence, the VaR at level α for the considered portfolio is given by:

$$\text{VaR}_\alpha = \Phi^{-1}(\alpha) \sqrt{\Delta^T C \Delta} - \Xi - \Delta^T \mu. \quad (2.1)$$

where $\Phi^{-1}(\cdot)$ is the inverse cumulative distribution function (cdf) for a standard normal random variable; $\Phi^{-1}(\alpha)$ assumes values of 1.282, 1.645 and 2.326 for corresponding values of 90%, 95% and 99% *VaR*.

Both the assumptions, viz. *linearity* of portfolio and *normality* of returns are often violated in practice:

- Real world returns have fatter tails than normal distributions. The model may lead to too optimistic estimates of VaR since it underestimates the likelihood of extreme returns;
- If a portfolio contains derivatives such as options, the value of the portfolio is no longer a linear function of \mathcal{F} and a linear approximation may be very inadequate. An example of this will be provided later in chapter 5 while considering portfolios involving European call options.

2.3 Historical simulation method and variance/covariance method for finding the Value-at-Risk

Before exploring the remaining three scenarios for calculating VaR and CVaR in more details, we will first look at the two most common methods to find VaR which are most relevant to the first scenario described above. The historical simulation is a nonparametric procedure for computing Value-at-Risk where no specific assumptions about the distribution of risk factors are made. It considers the lower quantile of the distribution of the actual historical returns and assumes that history will repeat itself. Let us consider the time series of rates of return $R_n = \frac{\Pi_n - \Pi_{n-1}}{\Pi_{n-1}}$ at any time t_n , for $n = 1, 2, \dots, N_1$. Let \mathcal{R} be the random variable gathering all the computed rates of returns. Regardless of the distribution of \mathcal{R} , the following statements holds true:

$$\Pr[R_n > q_{1-\alpha}] = \alpha, \quad (2.2)$$

where $q_{1-\alpha}$ represents the $(1 - \alpha)$ quantile of the random variable \mathcal{R} , and α represents the level of confidence of VaR. We now assume that the lower quantiles of \mathcal{R} will remain constant in future samples of R_n for $n = N_1 + 1, N_1 + 2, \dots, N_2$. Hence, using the provided definition of R_n and rearranging terms in (2.2), we can write:

$$\Pr[\Pi_n - \Pi_{n-1} > q_{1-\alpha}\Pi_{n-1}] = \alpha,$$

that can be interpreted as “the probability of a change in portfolio value to be larger than the $\text{VaR}_\alpha = q_{1-\alpha}\Pi_{n-1}$ is α ”.

Similarly to the HS method, the variance/covariance method considers the lower quantile of the distribution of the actual historical returns and assumes that history will repeat itself. However VC methods assumes that rates of return are normally distributed: $\mathcal{R} \sim \mathcal{N}(\mu_{\mathcal{R}}, \sigma_{\mathcal{R}}^2)$. Hence:

$$\Pr [R_n > \mu_{\mathcal{R}} - q_{1-\alpha}^{\mathcal{N}}\sigma_{\mathcal{R}}] = \alpha, \quad (2.3)$$

where $q_{1-\alpha}^{\mathcal{N}}$ represents the $(1 - \alpha)$ quantile of the normal random variable \mathcal{R} (i.e. 1.645 and 2.326 for the 95% and 99% VaR, respectively), and α represents the level of confidence of VaR. Again, we assume that the lower quantiles of \mathcal{R} will remain constant in future samples of R_n for $n = N_1 + 1, N_1 + 2, \dots, N_2$. Hence, using the provided definition of R_n and rearranging terms in (2.3), we can write:

$$\Pr [\Pi_n - \Pi_{n-1} > q_{1-\alpha}^{\mathcal{N}}\Pi_{n-1}] = \alpha,$$

that can be interpreted as “the probability of a change in portfolio value to be larger than the $\text{VaR}_\alpha = q_{1-\alpha}^{\mathcal{N}}\Pi_{n-1}$ is α ”.

Computing unbiased VaR for the methods included in this section would require:

- Normality of risk factors and linearity for the VC method;
- Sufficient past data for the HS method.

One might notice that in portfolios of derivative instruments none of these requirements will be satisfied.

2.4 Linear portfolio with non-normal returns

This section contains three types of models that play an important role in describing linear portfolio with non-normal returns. Subsection 2.4.1 includes extreme value theory based models, subsection 2.4.2 describes modeling fat tails with Student’s t distributions and finally subsection 2.4.3 introduces copula functions.

2.4.1 Extreme value theory and risk measures

Modeling rare phenomena that sit outside the range of available observations can be handled by extreme value theory (EVT). It provides a concrete theoretical foundation on which we can build statistical models describing extreme events. There are two related methods to identify extremes in real data. The first method considers the maximum the variable takes in successive batches. These particular observations constitute the extreme events, also called block maxima. The second method concentrates on the realizations going beyond a given (high) threshold; when the realizations exceed u they constitute extreme events. The former method is traditionally used to analyze data with seasonality, e.g. hydrological data. The latter uses data more efficiently and seems, for this reason, to become the most chosen one in modern finance related applications. Several authors have used this approach including McNeil (1998), Embrechts et al. (1999), Gilli and K ellezi (2006) and Fern andez (2003).

The limit law for the block maxima, which we denote M_n with n the size of the subsample (block), is given by the following theorem (see Fisher and Tippett (1928) for further details):

Theorem 2.4.1. Let (\mathcal{X}_n) be a sequence of i.i.d. random variables. If there exist constants $b_n > 0$, $a_n \in \mathbb{R}$ and some non-degenerate distribution function \mathcal{H} such that:

$$\frac{M_n - a_n}{b_n} \xrightarrow{d} \mathcal{H},$$

then \mathcal{H} belongs to one of the three standard extreme value distributions (EVD):

Gumbel (type I):

$$\Lambda(x) = e^{-e^{-x}}, \quad x \in \mathbb{R};$$

Fr chet (type II):

$$\Phi_\alpha(x) = \begin{cases} 0 & x \leq 0 \\ e^{-x^{-\alpha}} & x > 0 \end{cases} \quad \alpha > 0;$$

Weibull (type III):

$$\Psi_\alpha(x) = \begin{cases} e^{-(-x)^\alpha} & x \leq 0 \\ 1 & x > 0 \end{cases} \quad \alpha > 0.$$

The Fréchet distribution has a polynomially decaying tail and therefore suits well heavy tailed distributions.

The EVD families can be generalized with the incorporation of the location ($\lambda \in \mathbb{R}$) and scale ($\delta > 0$) parameters, leading to:

$$\Lambda(x; \lambda, \delta) = \Lambda((x - \lambda)/\delta), \quad \Phi_\alpha(x; \lambda, \delta) = \Phi_\alpha((x - \lambda)/\delta), \quad \Psi_\alpha(x; \lambda, \delta) = \Psi_\alpha((x - \lambda)/\delta).$$

The mean, variance and mode of the EVD are (as reported in Fraga Alves and Neves (2011)), respectively :

$$\Lambda : \mathbb{E}[\mathcal{X}] = \gamma \text{ (Euler's constant)} \simeq 0.57721; \text{var}(\mathcal{X}) = \pi^2/6; \text{Mode}(\mathcal{X}) = 0;$$

$$\Phi_\alpha : \mathbb{E}[\mathcal{X}] = \Gamma(1 - 1/\alpha), \text{ for } \alpha > 1; \text{var}(\mathcal{X}) = \Gamma(1 - 2/\alpha), \text{ for } \alpha > 2; \\ \text{Mode}(\mathcal{X}) = (1 + 1/\alpha)^{-1/\alpha};$$

$$\Psi_\alpha : \mathbb{E}[\mathcal{X}] = -\Gamma(1 + 1/\alpha); \text{var}(\mathcal{X}) = \Gamma(1 + 2/\alpha) - \Gamma^2(1 + 1/\alpha); \\ \text{Mode}(\mathcal{X}) = -(1 - 1/\alpha)^{-1/\alpha}, \text{ for } \alpha > 1, \text{ and } \text{Mode}(\mathcal{X}) = 0, \text{ for } 0 < \alpha \leq 1;$$

where Γ denotes the gamma function $\Gamma(s) := \int_0^\infty x^{s-1} e^{-x} dx$, $s > 0$.

Jenkinson (1955) and von Mises (1954) proposed the following one-parameter representation:

$$H_\xi(x) = \begin{cases} e^{-(1+\xi x)^{-1/\xi}} & \text{if } \xi \neq 0 \\ e^{-e^{-x}} & \text{if } \xi = 0 \end{cases} \quad (2.4)$$

of these standard distributions, with x such that $1 + \xi x > 0$. This distribution, a generalization of the previous classification, is known as the generalized extreme value (GEV)

distribution. It is obtained by setting $\xi = \alpha^{-1}$ for the Fréchet distribution, $\xi = -\alpha^{-1}$ for the Weibull distribution and by interpreting the Gumbel distribution as the limit case for $\xi = 0$.

The mean, variance and mode of the GEV are (as reported in Fraga Alves and Neves (2011)), respectively:

$$\begin{aligned} \mathbb{E}[\mathcal{X}] &= -\frac{1}{\xi}[1 - \xi(1 - \xi)], & \text{for } \xi < 1; \\ \text{var}(\mathcal{X}) &= \frac{1}{\xi^2}[\Gamma(1 - 2\xi) - \Gamma^2(1 - \xi)], & \text{for } \xi < 1/2; \\ \text{Mode}(\mathcal{X}) &= \frac{1}{\xi}[(1 + \xi)^{-\xi} - 1], & \text{for } \xi \neq 0. \end{aligned}$$

Furthermore the skewness coefficients of the GEV distribution, defined as $skew_{G_\xi} := \mathbb{E}[(\mathcal{X} - \mathbb{E}[\mathcal{X}])^3]/(\text{var}(\mathcal{X}))^{3/2}$, is equal to 0 at $\xi_0 \simeq -2.8$. Moreover, $skew_{G_\xi} > 0$, for $\xi > \xi_0$, and $skew_{G_\xi} < 0$ for $\xi < \xi_0$. Additionally, for the Gumbel distribution, $skew_{G_0} \simeq 1.14$.

Since we do not generally know in advance the type of limiting distribution of the sample maxima, the generalized representation is particularly useful when maximum likelihood estimates have to be computed. Additionally the standard GEV defined in (2.4) is the limiting distribution of normalized extrema. Since we do not know the true returns' distribution in practice, therefore we do not have any idea about constants a_n and b_n , we use three parameter specification:

$$H_{\xi, \sigma, \mu}(x) = H_\xi\left(\frac{x - \mu}{\sigma}\right) \quad x \in \mathcal{D}, \quad \mathcal{D} = \begin{cases} (-\infty, \mu - \frac{\sigma}{\xi}) & \xi < 0 \\ (-\infty, \infty) & \xi = 0 \\ (\mu - \frac{\sigma}{\xi}, \infty) & \xi > 0 \end{cases} \quad (2.5)$$

of the GEV, which is the limiting distribution of the unnormalized maxima. The two additional parameters μ and σ are the location and the scale parameters representing the unknown norming constants.

There is another approach, called peak over threshold (POT) method, which considers the distribution of exceedances over a certain threshold. The problem is, given an unknown distribution F of a random variable \mathcal{X} , to estimate the distribution function F_u of values of x above a certain threshold u . The distribution function F_u is called the conditional excess distribution function and is defined as:

$$F_u(y) = \Pr[\mathcal{X} - u \leq y | \mathcal{X} > u], \quad 0 \leq y \leq x_F - u, \quad (2.6)$$

where \mathcal{X} is a random variable, u is a given threshold, $y = x - u$ are the excesses and $x_F \leq \infty$ is the right endpoint of F . F_u can be written as:

$$F_u(y) = \frac{F(u + y) - F(u)}{1 - F(u)} = \frac{F(x) - F(u)}{1 - F(u)}, \quad (2.7)$$

Given that in general very little observation are located in this area the estimation of the portion F_u could be difficult.

EVT can then prove very useful as it offers us with a powerful result about the conditional excess distribution function which is stated in the following theorem (for further details see Pickands (1975) and Balkema and De Haan (1974)):

Theorem 2.4.2. For a large class¹ of underlying distribution functions F , the conditional excess distribution function $F_u(y)$, for u large, is well approximated by

$$F_u(y) \approx G_{\xi, \sigma}(y), \quad u \rightarrow \infty,$$

where

$$G_{\xi, \sigma} = \begin{cases} 1 - \left(1 + \frac{\xi}{\sigma}y\right)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1 - e^{-y/\sigma} & \text{if } \xi = 0 \end{cases} \quad (2.8)$$

for $y \in [0, (x_F - u)]$ if $\xi \geq 0$ and $y \in [0, -\frac{\sigma}{\xi}]$ if $\xi < 0$. $G_{\xi, \sigma}$ is the so called generalized Pareto distribution (GPD).

¹Please see chapter 1 of Kotz and Nadarajah (2000) for exact conditions defining the class of distributions in each case.

If x is defined as $x = u + y$, the GPD can also be expressed as a function of x , i.e. $G_{\xi, \sigma} = 1 - (1 + \xi(x - u)/\sigma)^{-1/\xi}$.

The parameter ξ is the shape parameter (or tail index) and gives an indication of the tail's heaviness, the larger ξ the heavier the tail. Only distributions with shape parameter $\xi \geq 0$ are suited to model financial return distributions because in general, one cannot establish an upper bound for financial losses.

Taking a GPD function for the tail distribution, the analytical expression for VaR_α and ES_α can be defined as a function of GPD parameters. Obtaining $F(x)$ from equation (2.7):

$$F(x) = (1 - F(u))F_u(y) + F(u),$$

and replacing F_u by the GPD and $F(u)$ by the estimate $(n - N_u)/n$, where n is the total number of observations and N_u the number of observations above the threshold u , we obtain:

$$\hat{F}(x) = \frac{N_u}{n} \left(1 - \left(1 + \frac{\hat{\xi}}{\hat{\sigma}}(x - u) \right)^{-1/\hat{\xi}} \right) + \left(1 - \frac{N_u}{n} \right),$$

which simplifies to:

$$\hat{F}(x) = 1 - \frac{N_u}{n} \left(1 + \frac{\hat{\xi}}{\hat{\sigma}}(x - u) \right)^{-1/\hat{\xi}}. \quad (2.9)$$

Inverting equation (2.9) for a given probability α gives

$$\widehat{VaR}_\alpha = u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{n}{N_u} \alpha \right)^{-\hat{\xi}} - 1 \right). \quad (2.10)$$

Expected shortfall can be written as

$$\widehat{CVaR}_\alpha = \widehat{VaR}_\alpha + \mathbb{E} \left[\mathcal{X} - \widehat{VaR}_\alpha | \mathcal{X} > \widehat{VaR}_\alpha \right],$$

where the second term on the right is the expected value of the exceedances over the threshold VaR_α . We can use the known result about the mean excess function for the GPD with

parameter $\xi < 1$:

$$e(z) = \mathbb{E}[\mathcal{X} - z | \mathcal{X} > z] = \frac{\sigma + \xi z}{1 - \xi}, \quad \text{with } \sigma + \xi z > 0. \quad (2.11)$$

This function gives the average of the excesses of \mathcal{X} over varying values of a threshold z . Another important result concerning the existence of moments is that if \mathcal{X} follows a GPD then, for all integers r such that $r < 1/\xi$, the r first moments exist (see Embrechts et al. (1999) for details).

Likewise, given the definition $ES_\alpha = \mathbb{E}[\mathcal{X} | \mathcal{X} > VaR_\alpha]$ for the expected shortfall and using expression (2.11), for $z = VaR_\alpha - u$ and \mathcal{X} representing the excesses y over u we obtain:

$$\widehat{CVaR}_\alpha = \widehat{VaR}_\alpha + \frac{\hat{\sigma} + \hat{\xi}(\widehat{VaR}_\alpha - u)}{1 - \hat{\xi}} = \frac{\widehat{VaR}_\alpha}{1 - \hat{\xi}} + \frac{\hat{\sigma} - \hat{\xi}u}{1 - \hat{\xi}}. \quad (2.12)$$

A couple of limitations about POT methods is reported in McNeil (1998): “...However, we note there are some practical problems with the use of threshold methods for unconditional inference about financial return series. The asymptotical properties of tail index estimators have not been extensively investigated for dependent samples of extreme values from dependent processes and estimates may be heavily biased unless the number of upper order statistics (or threshold) is chosen very carefully.”, “...In addition to the problem of threshold choice, there is also the related unresolved problem of how to give reasonable errors and confidence intervals for point estimates of tail indices or quantile based on dependent data.”.

2.4.2 Modeling fat tails with Student’s t distribution

The normal distribution is the most widespread used model of daily changes in market variables. In practice, several studies (Wilson (1993), Wilson (1998), Zangari (1996a), Venkataraman (1997), Duffie and Pan (1997) and Hull and White (1998)) show that distributions such as returns in equity, foreign exchanges, and commodity markets, are frequently asymmetric with fat tails. Assuming normality seems therefore rather inappropriate. In particular, Duffie and Pan (1997) and Hull and White (1998) show that daily changes in many variables, such as S&P 500, NASDAQ, NYSE All Share, and particularly exchange rates, exhibit significant amounts of positive kurtosis and negative skewness: these variables reveal fat tails and are typically skewed to the left. The practical implications is that neg-

ative returns are more probable than a normal distribution would expect when taking into account the kurtosis and skewness. In order to capture these features we consider a portfolio under the assumption of risk factor returns having a Student's t distribution. Different approaches involve skewed Student's t distribution (Ngoussou (2004)), or mixture of normal (Hull and White (1998), Wang (2001), and Zangari (1996a)) or Student's t distributions (Kamdem (2003)).

Skewness measures the asymmetry of a random variable and is found using its third central moment. It determines the degree of difference between positive deviations from the mean and negative deviation from the mean. Kurtosis measures how fat are the tails of a distribution and is found from its fourth central moment. It is very sensitive to extremely large market moves. Let \mathcal{X} be a random variable with probability distribution function $f_{\mathcal{X}}(x)$, then its k -th central moment is the quantity $\mu_k = \mathbb{E}[(\mathcal{X} - \mathbb{E}(\mathcal{X}))^k] = \int_{\mathbb{R}} (x - \mu)^k f_{\mathcal{X}}(x) dx$.

The first few central moments are:

$$\mu_1 = \mathbb{E}(\mathcal{X}), \quad \mu_2 = \mathbb{E}[(\mathcal{X} - \mu)^2], \quad \mu_3 = \mathbb{E}[(\mathcal{X} - \mu)^3], \quad \mu_4 = \mathbb{E}[(\mathcal{X} - \mu)^4].$$

Mean, variance, skewness and kurtosis are defined as:

$$\begin{aligned} \mu &= \mathbb{E}(\mathcal{X}) = \mu_1; \\ \sigma^2 &= \mathbb{E}[(\mathcal{X} - \mu)^2] = \mu_2 - \mu_1^2; \\ \tau &= \sigma^{-3} \mathbb{E}[(\mathcal{X} - \mu)^3] = \sigma^{-3} (\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3); \\ k^* &= \sigma^{-4} \mathbb{E}[(\mathcal{X} - \mu)^4] = \sigma^{-4} (\mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4). \end{aligned}$$

Any normally distributed random variable has kurtosis of three and skewness of zero.

Probability density functions are usually defined by three parameters: a location parameter, a scale parameter and a shape parameter. A very common measure of location parameter is the mean; the variability of the probability density function (pdf) is provided by the scale parameter, and the most commonly used is variance or standard deviation. Kurtosis and/or skewness, modeled through the shape parameter, influences how the variations are distributed around the location parameter. If data present heavy tails, the Value-at-Risk calculated using normal assumption can present huge differences from the one assuming Student's t distribution.

Probability density function of noncentral Student's t distribution, as discussed in Ngoussou

(2004), Kanto (2004) and Rozga and Arnerić (2009), is given by:

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu\sigma^2}} \left(1 + \frac{(x-\mu)^2/\sigma^2}{\nu}\right)^{-(\nu+1)/2}, \quad x \in \mathbb{R}, \quad (2.13)$$

where μ is the location parameter, σ is the scale parameter and ν is the shape parameter (degrees of freedom), and $\Gamma(\cdot)$ is the gamma function. Standard Student's t distribution assumes that $\mu = 0$, $\sigma^2 = 1$, with $\nu \in \mathbb{N}_0$. However, degrees of freedom can be estimated as non-integer, relating to kurtosis. The t distribution has tails heavier than the normal distribution when $\nu > 4$. Kurtosis is undefined for $\nu \leq 4$. Thus, a more leptokurtic empirical distribution would reflect in a smaller number of degrees of freedom.

The second and fourth central moments of function (2.13) are defined as:

$$\mu_2 = \mathbb{E}[(\mathcal{X} - \mu)^2] = \frac{\sigma^2\nu}{\nu - 2}, \quad (2.14)$$

$$\mu_4 = \mathbb{E}[(\mathcal{X} - \mu)^4] = \frac{3\sigma^4\nu^2}{(\nu - 2)(\nu - 4)}, \quad (2.15)$$

with standardized kurtosis:

$$k = \frac{\mu_4}{\mu_2^2} - 3 = \frac{6}{\nu - 4}. \quad (2.16)$$

Hence in order to obtain consistent estimators we could apply the method of moments, giving:

$$\hat{\nu} = 4 + \frac{6}{\hat{k}}, \quad (2.17)$$

$$\hat{\sigma}^2 = \left(\frac{3 + \hat{k}}{3 + 2\hat{k}}\right)\hat{\mu}_2. \quad (2.18)$$

where the sample variance is biased estimator of σ^2 .

Therefore the following correction factor must be introduced to achieve an unbiased estimator for the standard deviation:

$$\sqrt{\frac{3 + \hat{k}}{3 + 2\hat{k}}} = \sqrt{\frac{\hat{\nu} - 2}{\hat{\nu}}}.$$

Practice suggests that kurtosis is often larger than six, heading to estimation of non-integer degrees of freedom between four and five. However, kurtosis will depend on volatility persistence.

When the assumption of Student's t distribution is introduced VaR_α can be calculated as:

$$t_\alpha^\nu \cdot \sigma \cdot \sqrt{\frac{\nu-2}{\nu}} - \mu, \quad (2.19)$$

where t_α^ν is the critical value of t distribution depending on the given probability and estimated degrees of freedom, μ and σ are respectively the estimated mean and the estimated standard deviation, while $\sqrt{\frac{\nu-2}{\nu}}$ is the correction factor for unbiased standard deviation from sample.

Despite providing a more accurate representation than models assuming Gaussian distributed returns, the Student's t VaR models could present a possible source of error since the return distribution is assumed to be symmetric. To overcome this shortcoming there are skewed versions of the Student's t distribution, but the nonlinear transformations that support these distributions exclude an easy parametric linear VaR rule. On the other hand skewness can be included using mixture of normal or Student's t distributions; the resulting parametric linear VaR analytical formula has an implicit solution that can be solved using numerical methods (see Alexander (2008) for further details).

In general, changing the portfolio composition means varying the portfolio distribution in an unpredictable way. If weights of the portfolio changed and the returns were normal, then the portfolio returns would still be jointly normal and the VaR computation would involve a simple matrix-vector multiplication. If, however, portfolio weights changed and returns were not jointly normal, e.g. were Student's t distributed, the distribution parameters would need to be re-calibrated.

2.4.3 Copula functions

In general, the knowledge of the joint distribution of a set of random variables allows to obtain both the marginal distributions and their correlation structure. On the other hand,

given marginal distributions, a *copula function* is a function which maps the marginals into a joint distribution function. Copula function is a method that models the dependence among different random variables through a specified function. Often financial researchers possess more information about marginal distributions rather than their joint distribution. Hence the approach provided by copula functions is very useful, especially in cases where marginal distributions are non-Gaussian, since the dependence can be parameterized independently from the parameterization of the marginals.

A copula is a function which links a multivariate distribution function to its one-dimensional marginal distribution functions. The concept of copula was introduced in Sklar (1973) in the theorem which bears his name. A more detailed treatment of copulas can be found in Nelsen (2006). Here we limit our attention to the bivariate case, but it is easy to generalize to the m -dimensional case. Let us consider two continuous random variables \mathcal{X}_1 and \mathcal{X}_2 with arbitrary cumulative distribution function (cdf) $F_i = \Pr[\mathcal{X}_i \leq x_i]$, $i = 1, 2$, and joint distribution $F(x_1, x_2) = \Pr[\mathcal{X}_1 \leq x_1, \mathcal{X}_2 \leq x_2]$. Every pair $(x_1, x_2) \in [-\infty, +\infty]^2$ can be mapped into $[0, 1]^2$, using the respective cdf. A (two-dimensional) copula is a function $C : [0, 1]^2 \rightarrow [0, 1]$ such that:

1. $C(0, x) = C(x, 0) = 0$ and $C(1, x) = C(x, 1) = x$ for all $x \in [0, 1]$;

2. C is 2-increasing: for $x'_1, x''_1, x'_2, x''_2 \in [0, 1]$ with $x'_1 \leq x''_1$ and $x'_2 \leq x''_2$,

$$V_C([x'_1, x''_1] \times [x'_2, x''_2]) = C(x''_1, x''_2) - C(x'_1, x''_2) - C(x''_1, x'_2) + C(x'_1, x'_2) \geq 0.$$

The function V_C is the C -volume of the rectangle $[x'_1, x''_1] \times [x'_2, x''_2]$. A copula C induces a probability measure on $[0, 1]^2$ via $V_C([0, u_1] \times [0, u_2]) = C(u_1, u_2)$. The definitions so far introduced are linked together by the following theorem:

Sklar's Theorem. Let F be a two-dimensional distribution function with marginal distribution functions F_1 and F_2 . Then there exist a copula C such that $F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$. Conversely, for any distribution functions F_1 and F_2 and any copula C , the function F defined above is a two-dimensional distribution function with marginals F_1 and F_2 . Furthermore, if F_1 and F_2 are continuous, C is unique.

Given a joint distribution F with continuous marginals F_1 and F_2 it is possible to obtain the corresponding copula $C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2))$, where F_i^{-1} is the cadlag inverse of F_i , given by $F_i^{-1}(w) = \sup\{x | F_i(x) \leq w\}$. Furthermore, if \mathcal{X}_1 and \mathcal{X}_2 are continuous

random variables as defined earlier in this section, then C is the joint distribution function for the random variables $\mathcal{U}_1 = F_1(\mathcal{X}_1)$ and $\mathcal{U}_2 = F_2(\mathcal{X}_2)$.

A large number of copulas are present in the literature. A few copula functions which are important from a practical point of view are listed below:

- Product copula:

$$C(u_1, u_2) = u_1 u_2;$$

- Farlie-Gumbel-Morgenstern copula:

$$C(u_1, u_2; \theta) = u_1 u_2 (1 + \theta(1 - u_1)(1 - u_2)),$$

where the dependence parameter $\theta \in [-1, 1]$;

- Gaussian copula:

$$C(u_1, u_2; \rho) = \Phi_G(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho),$$

where Φ is the cdf of the standard normal distribution, Φ_G is the standard bivariate normal distribution with correlation parameter $\rho \in (-1, 1)$;

- Clayton copula:

$$C(u_1, u_2; \theta) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta},$$

where the dependence parameter $\theta \in (0, \infty)$;

- Frank copula: $C(u_1, u_2; \theta) = -\theta^{-1} \ln \left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1} \right)$,

where the dependence parameter $\theta \in \mathbb{R}$.

Despite popularity in certain (especially bivariate) applications, copulas have attracted criticism for their use in finance as well. Some critical opinions have been expressed in Mikosch (2006), among them the author states that “*Copulas do not contribute to a better understanding of multivariate extremes*”, “*Various copula models are mostly chosen because they are mathematically convenient; the rationale for their applications is murky*” and “*There is little statistical theory for copulas. Sensitivity studies of estimation procedures and goodness-of-fit tests for copulas are unknown. It is unclear whether a good fit of the copula of the data yields a good fit to the distribution of the data*”.

2.5 Nonlinear portfolio with normal returns

Evaluation of VaR for nonlinear functions of random risk factors can often be done efficiently using characteristic function methods, which in turn rely on discrete Fourier transforms (DFT) and fast Fourier transforms (FFT). Hence this section is organized as follows: firstly, in subsection 2.5.1 we introduce the Delta-Gamma-Normal approximation and use of FFT in order to deal with nonlinear portfolios with normal returns. Secondly, we provide the theoretical background to FFT: subsection 2.5.2 explains the fundamentals of digital signal processing which are relevant for this thesis; subsections 2.5.3 and 2.5.4 outline DFT and FFT algorithms respectively; subsection 2.5.5 explains the characteristic function and the inversion integral; subsection 2.5.6 regards the characteristic function inversion using FFT and lists some numerical integration algorithms along with their respective approximation errors.

2.5.1 Delta-Gamma-Normal model

In order to manage risk one needs in the first place to measure market risk, particularly the risk of significant losses. For substantial portfolios involving complex financial securities, this can be a major challenge. To achieve a good risk assessment there is an administrative side – building a precise, centralized database of a company’s positions with a leg on each side of several markets and asset classes – and a statistical and computational sides. As mentioned in Glasserman (2003) processes to evaluate market risk must specifically answer two questions:

- which statistical model correctly and opportunely describes individual risk sources of risk movement and co-movements of multiple sources of risk having an effect on a portfolio?
- How will the value of a portfolio vary in reply to variations in the underlying sources of risk?

The first question involves the joint distribution of changes in various risk factors – interest rates, stocks, exchange rates, and commodity prices that can affect a portfolio. The second question requires a mapping from risk factors to portfolio value. Once both components are provided, the distribution of the portfolio’s profit and loss is basically determined, and so it is any risk measure that reviews this distribution.

Answering these two questions unavoidably implies a trade-off between computational complexity and modeling accuracy. For example, modeling market prices with a multivariate normal distribution could be inadequate in terms of modeling the tail losses; however, it is still extensively used due to its computational simplicity. We will concentrate on the computational problems raised by the first, employing the model described in Glasserman (2003).

Let us recall the model described in Glasserman (2003) for computation of Value-at-Risk with normally distributed stock returns, but for a portfolio which depends nonlinearly on stock prices. We introduce the following notation:

$$\begin{aligned}
 \mathcal{S} &= \text{vector of } m \text{ market prices and rates;} \\
 \Delta t &= \text{risk-measurement horizon;} \\
 \Delta \mathcal{S} &= \text{change in } \mathcal{S} \text{ over interval } \Delta t; \\
 \Pi(\mathcal{S}, t) &= \text{portfolio value at time } t \text{ and market prices } \mathcal{S}; \\
 \mathcal{L} &= \text{loss over interval } \Delta t = -\Delta \Pi = \Pi(\mathcal{S}, t) - \Pi(\mathcal{S} + \Delta \mathcal{S}, t + \Delta t); \\
 F_{\mathcal{L}}(x) &= \Pr[\mathcal{L} < x], \text{ the distribution of } \mathcal{L}.
 \end{aligned}$$

The time interval Δt is typically very short, while the number m of risk factors could reach thousands. Regulators and rating agencies demand measurements over a two-week horizon, which is often understood as the time that might be required in the case of an adverse market move to unwind complicated positions.

Some implicit simplifying assumptions have been made. Only the net loss over the horizon Δt is under consideration, for example paying no attention to the maximum and minimum portfolio value with the horizon. The dynamics of the market prices are also neglected, including all the details about the evolution of \mathcal{S} in the vector of changes $\Delta \mathcal{S}$. Furthermore the assumption that the composition of the portfolio does not change is made, but clearly the value of its components may vary in reaction to the market movement $\Delta \mathcal{S}$ and the passage of time Δt , which may bring assets closer to expiry or maturity.

We recall that the Value-at-Risk of a portfolio is a quantile of its loss distribution over

a fixed horizon Δt : $\Pr[\mathcal{L} > x_\alpha] = \alpha$ and, for simplicity, is assumed throughout that the pdf of loss, $F_{\mathcal{L}}$, is continuous, so that such a point exists. A quantile supplies a simple system of summarizing information about a distribution's tail, and this specific value is often interpreted as a realistic worst-case loss level.

There are many ways to calculate or to approximate Value-at-Risk, each representing some trade-off between realism and tractability. This compromise depends on the portfolio's complexity and on the accuracy needed. The two main approaches employed to compute measures of risk such as VaR and CVaR for non-linear portfolios are the full Monte Carlo simulation and the partial Monte Carlo simulation. The full Monte Carlo simulation involves computation of the portfolio value, as a function of risk factors, for a large number of possible values of the risk factors as sampled from their multivariate distribution. Partial Monte Carlo simulation involves the same computation, apart from the fact that the portfolio value function is approximated by its Taylor series expansion truncated up to second order term. The partial Monte Carlo model is particularly appealing in the Gaussian case, since inverting the characteristic function of its probability distribution provides an approximation comparable with the one obtained through full Monte Carlo simulation, but with a very competitive computational time. Next, we introduce the partial Monte Carlo simulation.

A portfolio that includes options depends nonlinearly on the underlying asset prices and many fixed-income securities have a nonlinear dependence on interest rates. The change in portfolio value $\Delta\Pi(\mathcal{S}, t)$ is defined as $\Pi(\mathcal{S} + \Delta\mathcal{S}, t + \Delta t) - \Pi(\mathcal{S}, t)$. Some nonlinearity can be captured using Taylor expansion until the quadratic term:

$$\Delta\Pi(\mathcal{S}, t) \approx \Theta\Delta t + \delta^T\Delta\mathcal{S} + \frac{1}{2}\Delta\mathcal{S}^T\Gamma\Delta\mathcal{S}, \quad (2.20)$$

where the vector $\Delta\mathcal{S}$ denotes the change of the underlying value, while the scalar $\Theta = \frac{\partial\Pi(\mathcal{S}, t)}{\partial t}$, the vector $\delta = \frac{\partial\Pi(\mathcal{S}, t)}{\partial\mathcal{S}}$ and the matrix $\Gamma = \frac{\partial^2\Pi(\mathcal{S}, t)}{\partial\mathcal{S}^2}$ represent the sensitivities of the portfolio value to time and to the changes in the stock prices at time t . Given the portfolio Π , all the sensitivities Θ , δ and Γ are assumed to be given exogenously for the purpose of this paper. For individual derivative instruments such as options, δ values are provided by commercial economic data providers such as Bloomberg. Depending on the nature of

derivative instruments in the portfolio, these sensitivity values may be found using a variety of methods from market data and from commonly used risk neutral models, including finite differences, implied volatilities, Malliavin calculus (for use of Malliavin calculus in computing sensitivities of certain exotic options, see e.g. Montero and Kohatsu-Higa (2003)), etc; see, e.g. Glasserman (2003) and Hull (2009). Sensitivities δ and Γ are regularly computed for hedging purposes by single trading desks and can be combined (at the end of the day, for example).

There are two ways to compute the change in portfolio value $\Delta\Pi$ through formula (2.20) either directly using Monte Carlo simulation or employing FFT, which is described next. Assuming $\Delta\mathcal{S} \sim \mathcal{N}(0, \Sigma_{\mathcal{S}})$, then we can write:

$$\Delta\mathcal{S} = CZ \quad \text{with} \quad CC^T = \Sigma_{\mathcal{S}},$$

where $Z \sim \mathcal{N}(0, I)$ and C is a square root of $\Sigma_{\mathcal{S}}$. Square root matrices are not unique and the exact choice of C is decided as follows. We can re-write (2.20) in terms of Z :

$$\mathcal{L} \approx a - (C^T \delta)^T Z - \frac{1}{2} Z^T (C^T \Gamma C) Z, \quad (2.21)$$

with $a = -\Theta\Delta t$ deterministic.

Let us choose the matrix C to diagonalize the quadratic term in (2.21) in the following way: choose \hat{C} so that $\hat{C}\hat{C}^T = \Sigma_{\mathcal{S}}$, such as the one found by Cholesky factorization. Being the matrix $-\frac{1}{2}\hat{C}^T\Gamma\hat{C}$ symmetric, it admits the representation

$$-\frac{1}{2}\hat{C}^T\Gamma\hat{C} = U\Lambda U^T,$$

where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$ and U is an orthogonal matrix ($UU^T = I$). Columns of the matrix U are then eigenvectors of $-\frac{1}{2}\hat{C}^T\Gamma\hat{C}$. The λ_j are eigenvalues of both this matrix and $-\frac{1}{2}\Gamma\Sigma_{\mathcal{S}}$. Fixing $C = \hat{C}U$, one can observe that

$$CC^T = \hat{C}UU^T\hat{C}^T = \Sigma_{\mathcal{S}}$$

and

$$-\frac{1}{2}C^T \Gamma C = -\frac{1}{2}U^T(\hat{C}^T \Gamma \hat{C})U = U^T(U \Lambda U^T)U = \Lambda.$$

Then by setting $\mathbf{b} = -C^T \delta$ the approximation (2.21) becomes

$$\mathcal{L} \approx a + \mathbf{b}^T \mathcal{Z} + \mathcal{Z}^T \Lambda \mathcal{Z} = a + \sum_{i=1}^m (b_i \mathcal{Z}_i + \lambda_i \mathcal{Z}_i^2) \equiv \mathcal{Q}. \quad (2.22)$$

Since we approximated $\Pr[\mathcal{L} > x] \approx \Pr[\mathcal{Q} > x]$, now we have to derive the distribution of \mathcal{Q} .

The distribution of \mathcal{Q} is determined using its moment generating function and characteristic function. The moment generating function is finite in a neighborhood of the origin and, considering the independence of the summands in (2.22), can be factorized as:

$$\mathbb{E}[e^{\theta \mathcal{Q}}] = e^{a\theta} \prod_{i=1}^m \mathbb{E}[e^{\theta(b_i \mathcal{Z}_i + \lambda_i \mathcal{Z}_i^2)}] \equiv e^{a\theta} \prod_{i=1}^m e^{\psi_i(\theta)}.$$

If $\lambda_i = 0$, then the moment generating function² produces $\psi_j(\theta) = b_j^2 \theta^2 / 2$. Otherwise, we write:

$$b_i \mathcal{Z}_i + \lambda_i \mathcal{Z}_i^2 = \lambda_i \left(\mathcal{Z}_i + \frac{b_i}{2\lambda_i} \right)^2 - \frac{b_i^2}{4\lambda_i},$$

a linear transformation of a noncentral chi-square random variable³. Using the equation (29.6) of Johnson et al. (1994), we obtain:

$$\mathbb{E}[\exp(\theta(\mathcal{Z}_i + c)^2)] = (1 - 2\theta)^{-1/2} \exp\left(\frac{\theta c^2}{1 - 2\theta}\right),$$

for $\theta < 1/2$, one can arrive at the expression (see Glasserman (2003)):

$$\psi(\theta) \equiv a\theta + \sum_{i=1}^m \psi_i(\theta) = a\theta + \frac{1}{2} \sum_{i=1}^m \left(\frac{\theta^2 b_i^2}{1 - 2\theta \lambda_i} - \log(1 - 2\theta \lambda_i) \right), \quad (2.23)$$

²The moment generating function for a random variable $\mathcal{X} \sim \mathcal{N}(\mu, \Sigma)$, with \mathcal{X} d -dimensional, is given by $\mathbb{E}[\exp(\theta^T \mathcal{X})] = \exp(\mu^T \theta + \frac{1}{2} \theta^T \Sigma \theta)$, for all $\theta \in \mathbb{R}^d$.

³For $\nu \in \mathbb{N}_0$ and constants d_1, d_2, \dots, d_ν , the distribution of

$$\sum_{i=1}^{\nu} (\mathcal{Z}_i + d_i)^2$$

is noncentral chi-square with ν degrees of freedom and noncentral parameter $\lambda = \sum_{i=1}^{\nu} d_i^2$. This representation explains the term ‘‘noncentral’’.

where $\psi(\theta) = \log \mathbb{E}[\exp(\theta \mathcal{Q})]$ is the cumulant generating function of \mathcal{Q} ; the equation (2.23) holds for all θ such that $\max_i \theta \lambda_i < 1/2$.

Let us consider a generic continuous random variable \mathcal{X} with cumulative distribution function (cdf) $F_{\mathcal{X}}(x) = \Pr[\mathcal{X} \leq x]$. Its characteristic function, $\varphi_{\mathcal{X}} : \mathbb{R} \rightarrow \mathbb{C}$, is defined as:

$$\varphi_{\mathcal{X}}(t) = \mathbb{E}[e^{jt\mathcal{X}}] = \int_{-\infty}^{\infty} e^{jtx} dF_{\mathcal{X}}(x),$$

where $j = \sqrt{-1}$.

The characteristic function of the approximation \mathcal{Q} as expressed by (2.22) is reported in several articles (see, e.g. Feuerverger and Wong (2000) for further details) and is given by:

$$\varphi_{\mathcal{Q}}(t) = \mathbb{E}[e^{jt\mathcal{Q}}] = e^{\psi(jt)} = e^{jta} \prod_{i=1}^m \frac{1}{\sqrt{1 - 2j\lambda_i t}} \exp\left(-\frac{1}{2} \sum_{i=1}^m \frac{b_i^2 t^2}{1 - 2j\lambda_i t}\right). \quad (2.24)$$

The loss distribution can be found using numerical inversion of the characteristic function. The formula we employed for the computation of the numerical inversion of the characteristic function is the following:

$$f_{\mathcal{Q}}(q) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_{\mathcal{Q}}(t) e^{-jtq} dt.$$

More discussion on characteristic functions is postponed until subsection 2.5.5. The Delta-Gamma-Normal algorithm combines the Delta-Gamma method with the characteristic function inversion using a numerically efficient procedure called Fast Fourier Transform (FFT). Therefore once one computes coefficients a, b_i, λ_i , for $i = 1, 2, \dots, m$, as explained in subsection 2.5.1 it is possible to use formula (2.24) to calculate $\varphi_{\mathcal{Q}}$ for a N -length grid of r_n values in order to obtain $x(n)$ values according to formula (2.54) given later in subsection 2.5.6. To obtain more accurate integration one can include one of the numerical integration rules mentioned in subsection 2.5.6. Inverting these quantities using FFT will provide the information necessary to compute both VaR and CVaR.

Note that there are two reasons for truncating Taylor's series after two terms while approximating portfolio loss. First reason is the computational simplicity of computing quantiles of a quadratic form of a normal distribution (as outlined above). Secondly, widely reported

empirical experience supports the use of second order approximation as adequate reflection of changes in portfolio value, especially over short time horizons as considered here. In Glasserman et al. (2000), the authors relax the assumption that the risk factors are normally distributed and demonstrate that the result can be extended to risk factors which have a multivariate t distribution. In this thesis, we go a lot further in generalizing Delta-Gamma-Normal method by using possibly non-parametric distributions. The key to our method is a probability conserving transformation, which is described later in subsection 2.6.1.

2.5.2 Digital signal processing

First, we will have a look at some mathematical tools relevant in the characteristic function inversion mentioned earlier. Signal processing concerns with the science of analyzing time-varying natural processes (for further details, refer to digital signal processing textbooks such Lynn and Fuerst (2000) and Lyons (2004)). Signal processing is partitioned into two classes, analog signal processing and digital signal processing. A waveform that is continuous in time and can assume a continuous range of amplitude values is also called analog; on the other hand, a signal whose independent time variable is quantized, that means that the value of the signal is known only at discrete instants in time, is called discrete or digital. The representation of a discrete-time signal is not a continuous waveform but it is given by a sequence of values. Not only the time is quantized, indeed a discrete-time signal quantizes the signal amplitude, too.

For example a continuous cosine wave with a peak amplitude of 1 at a frequency f_0 is described by the equation

$$x(t) = \cos(2\pi f_0 t). \quad (2.25)$$

the frequency f_0 is measured in hertz (Hz). With t in equation (2.25) representing time in seconds, the factor $f_0 t$ has dimensions of cycles, and the complete $2\pi f_0 t$ term is an angle measured in radians. Sampling it once every t_s seconds and representing it as a sequence of discrete values, would give us its discrete version.

The discrete sequence $x(n)$ can be expressed by

$$x(n) = \cos(2\pi f_0 n t_s). \quad (2.26)$$

Equations (2.25) and (2.26) describe what are also referred to as time-domain signals because both the independent variables, the continuous time t and the discrete time nt_s , are measure of time.

Given that an analog signal have to be represented by a set of equally-spaced samples, then Shannon's sampling theorem states how often it should be sampled:

Shannon's Theorem. An analog signal containing components up to some maximum frequency f_1 Hz may be completely represented by regularly-spaced samples, provided that the sampling rate is at least $2f_1$ samples per second.

Using the minimum rate indicated by the theorem, the sampling interval t_s is clearly:

$$t_s = \frac{1}{2f_1}.$$

Conversely, if we have a digital signal with sampling interval t_s , the maximum analog frequency which can be represented is:

$$f_1 = \frac{1}{2t_s} \text{ Hz, or } \omega_1 = 2\pi f_1 \text{ radians/sec.}$$

2.5.3 Discrete Fourier transform (DFT)

Aperiodic signals and data with a finite number of nonzero sample values are very frequent – for example, the price of the copper or the temperature record.

The discrete Fourier transform (DFT) of a general signal $x(n)$, defined over the set $0 \leq n \leq N - 1$, is given by:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N} = \sum_{n=0}^{N-1} x(n)W_N^{kn}, \quad (2.27)$$

where $j = \sqrt{-1}$, $W_N = e^{-j2\pi/N}$, and the spectral coefficients $X(k)$ are calculated for $0 \leq k \leq N - 1$.

Using Euler's relationship $e^{-j\phi} = \cos(\phi) - j \sin(\phi)$, the equation (2.27) can be expressed in

the rectangular form:

$$X(k) = \sum_{n=0}^{N-1} x(n)[\cos(2\pi nk/N) - j \sin(2\pi nk/N)]. \quad (2.28)$$

The real and imaginary parts of the their spectrum are then given by:

$$X_{real}(k) = \text{Re}(X(k)) = \sum_{n=0}^{N-1} x(n) \cos(2\pi kn/N) \quad (2.29)$$

and

$$X_{imag}(k) = \text{Im}(X(k)) = - \sum_{n=0}^{N-1} x(n) \sin(2\pi kn/N). \quad (2.30)$$

Thus, an arbitrary DFT output value, $X(k)$, can be represented by

$$X(k) = X_{real}(k) + jX_{imag}(k) = |X(k)| \text{ at an angle } X_{\phi}(k), \quad (2.31)$$

where $|X(k)|$ is the magnitude of $X(k)$ and it is defined as

$$|X(k)| = \sqrt{X_{real}^2(k) + X_{imag}^2(k)},$$

while, the phase angle of $X(k)$, $X_{\phi}(k)$ is

$$X_{\phi}(k) = \arctan \left(\frac{X_{imag}(k)}{X_{real}(k)} \right).$$

We are going to consider signals $x(n)$ s that are real-valued, the real part of their corresponding $X(k)$ must be even while the imaginary part must be odd. Furthermore the magnitude and phase of $X(k)$ must be even and odd respectively, and $X(k) = \bar{X}(-k)$, where the bar denotes the complex conjugate. The DFT can therefore be completely defined by just half the total set of spectral coefficients, given the symmetry of the spectrum. As a result the DFT of a real, N -valued, signal bears two real coefficients, $X(0)$ and $X(N/2)$, together with $N/2 - 1$ distinct complex coefficients (the other $N/2 - 1$ complex coefficients are their complex conjugates). Two real and $N/2 - 1$ complex coefficients are for that reason needed to define the whole spectrum.

If $x(n)$ is even, i.e. $x(n) = x(-n)$, then its spectrum contains only cosines terms, therefore the imaginary part of $X(k)$ is zero; equally, when $x(n)$ is odd, its spectrum contains only sine terms and the real part of $X(k)$ is zero.

The inverse DFT, or IDFT, which allows us to retrieve the signal from its spectrum, is given by:

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) W_N^{-kn}, \quad (2.32)$$

where the values of $x(n)$ are calculated for $0 \leq n \leq N - 1$.

Both DFT and IDFT represent a finite-length sequence as one period of a periodic sequence: calculating additional values of $X(k)$ outside the range $0 \leq k \leq N - 1$ in equation (2.27) will bear a periodic spectral sequence. Likewise, using equation (2.32) to calculate additional values of $x(n)$ outside the range $0 \leq n \leq N - 1$ generates a periodic version of the signal. In fact the DFT considers an aperiodic signal $x(n)$ to be periodic for the purposes of computation. It can be noticed that the only difference between the DFT and the IDFT is the scaling factor of $1/N$, and a sign change in the exponent: this is a direct consequence between time and frequency domains. The DFT supplies us with N distinct spectral coefficients $X(k)$ for a signal with N distinct sample values; conversely, the IDFT redevelops the N signal values from the N spectral coefficients.

An aperiodic signal $x(n)$, defined for $0 \leq n \leq N - 1$, has DFT coefficients $X(k)$, defined for $0 \leq k \leq N - 1$. Then the sequence $x(n)$ transforms into the sequence $X(k)$; conversely $X(k)$ inverse transforms into $x(n)$. We use a double-headed arrow to denote the relationship between a signal and its spectrum: $x(n) \leftrightarrow X(k)$.

Since both $x(n)$ and $X(k)$ are N -periodic:

$$\begin{aligned} x(n) &= x(n + N) \quad \text{for all } n, \\ X(k) &= X(k + N) \quad \text{for all } k. \end{aligned}$$

Important properties of the DFT are:

1) *linearity*

if $x_1(n) \leftrightarrow X_1(k)$ and $x_2(n) \leftrightarrow X_2(k)$

then $Ax_1(n) + Bx_2(n) \leftrightarrow AX_1(k) + BX_2(k)$;

2) *time-shifting*

if $x(n) \leftrightarrow X(k)$

then $x(n - n_0) \leftrightarrow X(k)e^{-j2\pi kn_0/N} = X(k)W_n^{kn_0}$;

3) *convolution*

if $x_1(n) \leftrightarrow X_1(k)$ and $x_2(n) \leftrightarrow X_2(k)$

then $\sum_{m=0}^{N-1} x_1(n)x_2(m - n) \leftrightarrow X_1(k)X_2(k)$.

The sampling theorem result holds in the frequency domain; it could be stated as follow:
“The continuous spectrum of a signal with limited duration T_0 seconds may be completely represented by regularly-spaced frequency-domain samples, provided the samples are spaced not more than $1/T_0$ Hz apart”.

An aperiodic digital signal with N finite sample values having duration $T_0 = Nt_s$ seconds, where t_s is the sampling interval in the time domain, has a spectrum that can be fully represented by frequency-domain samples spaced $1/Nt_s$ Hz, or $2\pi/Nt_s$ radians per second, apart. Sampling must be done at intervals of $2\pi/N$ (or less).

The spectrum of a digital signal is always periodic, with period 2π : this is a consequence of sampling. Since one period is clearly sufficient to define it, sampling at the minimum interval of $2\pi/N$ provides us $\frac{2\pi}{2\pi/N} = N$ frequency-domain samples.

2.5.4 The fast Fourier transform (FFT)

Despite being the most straightforward procedure to reveal the frequency content of a time-domain sequence, DFT is computationally extremely inefficient. The execution of equations (2.29)-(2.30) involves a huge number of floating point operations; both equations (2.29) and (2.30) need $2N^2$ multiplications. Even in the presence of an even or odd signal still N^2 are required. In all of the mentioned cases there are analogous numbers of integer multiplications and floating-point additions/subtractions to be carried out. As a general rule, we may expect the calculation to be approximately proportional to N^2 . One can easily realize that as the number of points in the DFT is over 1,000, the amount of necessary computation become excessive. Highly cost-effective algorithms for computing the DFT were first developed in the 1960s. Collectively known as fast Fourier transforms (FFTs), they all linked to the fact that the standard DFT involves redundant computation.

The DFT of an N -length signal is given by

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N} = \sum_{n=0}^{N-1} x(n)W_N^{kn},$$

where $W_N = e^{-j2\pi/N}$, and $X(k)$ is calculated for $0 \leq k \leq N - 1$. It results that identical values of $x(n)W_N^{kn}$ are calculated many times as the computation proceeds. This is given by the periodic nature of W_N^{kn} , that originates a limited number of distinct values. A similar behavior applies to the IDFT equation (2.32). FFT algorithms aim to eliminate this redundancy.

There is a collection of FFT algorithms with dissimilar characteristics, benefits, and weaknesses. An algorithm which is appropriate for programming in high-level language on a general-purpose computer may not be the best for special-purpose DSP hardware. All of these different algorithms share a general approach, they split each DFT into a number of successively smaller, and easier to compute, DFTs.

This section will highlight why the most popular FFT algorithm (called the radix-2 FFT) is a cut above the classical DFT algorithm. The radix-2 FFT algorithm is a very effective procedure to carry out DFTs under the constraint that the number of points in the trans-

form (commonly referred to the DFT size) is $N = 2^k$, where k is some positive integer. The needed number of complex multiplications for an N -point DFT is N^2 , while for an N -point FFT is roughly $\frac{N}{2} \log_2 N$: a considerable reduction from the N^2 complex multiplications required by DFTs, particularly for large N .

When using the radix-2 FFT, if the length of the series $x(n)$ is not an integral power of two, we have two options: we could add enough zero-valued samples at the end of the data sequence to reach the number of points of the next largest radix-2 FFT (this technique is called zero-padding); alternatively one could get rid of enough data samples in order to match that particular integral power of two. The second approach is not suggested since ignoring data samples corrupt the resultant frequency-domain resolution.

Let us return to the equation for an N -point DFT:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N} = \sum_{n=0}^{N-1} x(n)W_N^{kn}, \quad 0 \leq k \leq N-1. \quad (2.33)$$

A derivation of the FFT progresses with the division of the input data sequence $x(n)$ in equation (2.33) into two parts: its even elements, indexed by $n = 2r$, and its odd elements indexed by $n = 2r + 1$ as:

$$X(k) = \sum_{r=0}^{N/2-1} x(2r)W_N^{2rk} + \sum_{r=0}^{N/2-1} x(2r+1)W_N^{(2r+1)k} \quad (2.34)$$

$$= \sum_{r=0}^{N/2-1} x(2r)(W_N^2)^{rk} + W_N^k \sum_{r=0}^{N/2-1} x(2r+1)(W_N^2)^{rk}. \quad (2.35)$$

Because $W_N^2 = e^{-j2\pi 2/N} = e^{-j2\pi/(N/2)} = W_{N/2}$ we may therefore write:

$$X(k) = \sum_{r=0}^{N/2-1} x(2r)W_{N/2}^{rk} + W_N^k \sum_{r=0}^{N/2-1} x(2r+1)W_{N/2}^{rk} \quad (2.36)$$

$$= A(k) + W_N^k B(k). \quad (2.37)$$

The original N -point DFT is now expressed in terms of two $N/2$ -point DFTs, $A(k)$ and $B(k)$. $A(k)$ is the transform of the even-indexed points in $x(n)$, and $B(k)$ is the transform of the odd-indexed points. $B(k)$ must be multiplied by the term W_N^k before adding it to $A(k)$. The necessary computation is reduced in (2.36) relative to (2.33) because the $W_{N/2}^{rk}$ terms in the two summations of (2.36) are exactly the same.

There is a further advantage in dividing the N -point DFT into two parts because the upper half of the DFT outputs is simple to calculate. Consider the $X(k + N/2)$ output. If we plug $k + N/2$ in for k in (2.36), then

$$X(k + N/2) = \sum_{r=0}^{N/2-1} x(2r)W_{N/2}^{r(k+N/2)} + W_N^{(k+N/2)} \sum_{r=0}^{N/2-1} x(2r+1)W_{N/2}^{r(k+N/2)}. \quad (2.38)$$

The phase angle expressions inside the summations assumes the shape:

$$W_{N/2}^{r(k+N/2)} = W_{N/2}^{rk} W_{N/2}^{rN/2} = W_{N/2}^{rk} (e^{-j2\pi r 2N/2N}) = W_{N/2}^{rk} \cdot 1 = W_{N/2}^{rk}, \quad (2.39)$$

for any integer r . Taking a look at the so-called twiddle factor, in front of the second summation in (2.38), it can be simplified as

$$W_N^{(k+N/2)} = W_N^k W_N^{N/2} = W_N^k (e^{-j2\pi N/2N}) = W_N^k (-1) = -W_N^k. \quad (2.40)$$

Then equation (2.38) can be represented as:

$$X(k + N/2) = \sum_{r=0}^{N/2-1} x(2r)W_{N/2}^{rk} - W_N^k \sum_{r=0}^{N/2-1} x(2r+1)W_{N/2}^{rk}. \quad (2.41)$$

Consequently, comparing (2.36) to (2.41) one can note that there is no need to perform any sine or cosine multiplications to get $X(k + N/2)$. Simply changing the sign of the twiddle factor W_N^k and using the results of the two summations from $X(k)$ allows to obtain $X(k + N/2)$. Now k goes from 0 to $N/2 - 1$ in (2.36) so, for an N -point DFT, an $N/2$ -point DFT is performed to get the first $N/2$ outputs and use those to calculate the last $N/2$

outputs.

If we assume that the transform length N is an integer power of 2, it follows that $N/2$ is even. Therefore we can take the decomposition further, by breaking each $N/2$ -point subsequence down into two shorter, $N/4$ -point subsequences. The process can continue until, in the limit, we are left with a series of 2-point subsequences, each of which requires a very simple 2-point DFT. A complete decomposition of this type gives rise to one of the commonly-used radix-2, decimation-in-time, FFT algorithms. A generalization of radix-2 algorithms is given by prime factor algorithms (PFA) which is applicable for any transform length of the form $N = 2^p 3^q 5^r$, where $p, q, r \in \mathbb{N}$ (see, e.g. Temperton (1992), Takahashi and Kanada (2000) and Goedecker (1997)).

2.5.5 The characteristic function and the inversion integral

We will review the idea of characteristic function, mentioned earlier in subsection 2.5.1, in more details here. The characteristic function of a continuous random variable \mathcal{X} , can be expressed as:

$$\begin{aligned}\varphi_{\mathcal{X}}(t) &= \int_{-\infty}^{\infty} e^{jtx} dF_{\mathcal{X}}(x) \\ &= \int_{-\infty}^{\infty} \cos(tx) dF_{\mathcal{X}}(x) + j \int_{-\infty}^{\infty} \sin(tx) dF_{\mathcal{X}}(x) \\ &= \mathbb{E}[\cos(t\mathcal{X})] + j \mathbb{E}[\sin(t\mathcal{X})].\end{aligned}$$

Since $\cos(-\alpha) = \cos(\alpha)$ and $\sin(-\alpha) = -\sin(\alpha)$, then:

$$\varphi_{\mathcal{X}}(-t) = \bar{\varphi}_{\mathcal{X}}(t). \quad (2.42)$$

Applying (2.42) and using the fact that $z + \bar{z} = 2\text{Re}(z)$, one obtains:

$$\varphi_{\mathcal{X}}(t) + \varphi_{\mathcal{X}}(-t) = \varphi_{\mathcal{X}}(t) + \bar{\varphi}_{\mathcal{X}}(t) = 2\text{Re}(\varphi_{\mathcal{X}}(t)).$$

Because $|e^{jz}| = 1$, then:

$$|\varphi_{\mathcal{X}}(t)| = \left| \int_{-\infty}^{\infty} e^{jtx} dF_{\mathcal{X}}(x) \right| \leq \int_{-\infty}^{\infty} |e^{jtx}| dF_{\mathcal{X}}(x) = \int_{-\infty}^{\infty} dF_{\mathcal{X}}(x) = 1,$$

so the characteristic function always exist. The uniqueness theorem, proved in Lévy (1924), asserts that given two random variables \mathcal{X} and \mathcal{Y} :

$$\mathcal{X} \stackrel{d}{=} \mathcal{Y} \quad \Leftrightarrow \quad \varphi_{\mathcal{X}} = \varphi_{\mathcal{Y}}.$$

In order to obtain the pdf or the cdf of a continuous random variable \mathcal{X} using the characteristic function one can make use of the inversion integral. There are different versions of it, some of which are reported below:

let \mathcal{X} be a continuous random variable with pdf $f_{\mathcal{X}}(x)$ and $\int_{-\infty}^{\infty} |\varphi_{\mathcal{X}}(t)| dt < \infty$, then:

$$f_{\mathcal{X}}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_{\mathcal{X}}(t) e^{-jtx} dt. \quad (2.43)$$

Now, making the substitution $u = -t$, and using properties: $z + \bar{z} = 2\text{Re}(z)$, $e^{\bar{z}} = \overline{e^z}$, $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$, and (2.42) we obtain:

$$\int_{-\infty}^0 \varphi_{\mathcal{X}}(t) e^{-jtx} dt = \int_0^{\infty} \varphi_{\mathcal{X}}(-u) e^{jux} du = \int_0^{\infty} \overline{\varphi_{\mathcal{X}}(u)} e^{-jux} du = \int_0^{\infty} \overline{\varphi_{\mathcal{X}}(u) e^{-jux}} du;$$

hence equation (2.43) can be rewritten as:

$$\begin{aligned} f_{\mathcal{X}}(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_{\mathcal{X}}(t) e^{-jtx} dt = \frac{1}{2\pi} \left(\int_{-\infty}^0 \varphi_{\mathcal{X}}(t) e^{-jtx} dt + \int_0^{\infty} \varphi_{\mathcal{X}}(t) e^{-jtx} dt \right) \\ &= \frac{1}{2\pi} \left(\int_0^{\infty} \overline{\varphi_{\mathcal{X}}(u) e^{-jux}} du + \int_0^{\infty} \varphi_{\mathcal{X}}(t) e^{-jtx} dt \right) \\ &= \frac{1}{\pi} \int_0^{\infty} \text{Re}(\varphi_{\mathcal{X}}(t) e^{-jtx}) dt. \end{aligned}$$

An expression for the inversion formula derived in Gil-Pelaez (1951) is:

$$F_{\mathcal{X}}(x) = \frac{1}{2} + \frac{1}{2\pi} \int_0^{\infty} \frac{\varphi_{\mathcal{X}}(-t) e^{jtx} - \varphi_{\mathcal{X}}(t) e^{-jtx}}{jt} dt, \quad (2.44)$$

or, alternatively, the one reported in Wendel (1961) can be used:

$$F_{\mathcal{X}}(x) = \frac{1}{2} - \frac{1}{\pi} \int_0^{\infty} \frac{\text{Im}(\varphi_{\mathcal{X}}(t) e^{-jtx})}{t} dt, \quad (2.45)$$

another formulation is used in Glasserman (2003):

$$\Pr[\mathcal{X} \leq x] - \Pr[\mathcal{X} \leq x - y] = \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left(\frac{e^{jty} - 1}{jt} \varphi_{\mathcal{X}}(t) e^{-jtx} \right) dt. \quad (2.46)$$

2.5.6 Characteristic function inversion using FFT

The treatment of the discrete version of Fourier transform (DFT) below follows Section 2.5.3.

We recall that the DFT of a signal $x(n)$ is given by:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N} = \sum_{n=0}^{N-1} x(n) W_N^{kn}, \quad k = 0, 1, \dots, N-1. \quad (2.47)$$

Therefore given the vector $\mathbf{x} = (x_0, x_1, \dots, x_{N-1})$, we obtain the vector $\mathbf{X} = (X_0, X_1, \dots, X_{N-1}) = \mathcal{F}(\mathbf{x})$. Conversely, the vector \mathbf{x} can be obtained from the IDFT:

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j2\pi kn/N} = \frac{1}{N} \sum_{k=0}^{N-1} X(k) W_N^{-kn}, \quad n = 0, 1, \dots, N-1, \quad (2.48)$$

which can be written in a more compact way $\mathbf{x} = \mathcal{F}^{-1}(\mathbf{X})$.

The function \mathcal{F}^{-1} is the inverse of \mathcal{F} because given a vector $\mathbf{g} = (g_0, g_1, \dots, g_{N-1})$, $\mathbf{g} = \mathcal{F}^{-1}(\mathcal{F}(\mathbf{g}))$. Let $G = \mathcal{F}(\mathbf{g})$ and $W_N = e^{j2\pi/N}$, using (2.47) and (2.48) we get:

$$\frac{1}{N} \sum_{k=0}^{N-1} G(k) W_N^{-kn} = \frac{1}{N} \sum_{k=0}^{N-1} \left(\sum_{m=0}^{N-1} g(m) W_N^{km} \right) W_N^{-kn} = \frac{1}{N} \sum_{m=0}^{N-1} g(m) \left(\sum_{k=0}^{N-1} (W_N^{m-n})^k \right).$$

Having (as reported in Paoletta (2007)):

$$\sum_{l=0}^{N-1} (W_T^k)^l = \begin{cases} N & \text{if } k \in d\mathbb{Z} = \{0, d, -d, 2d, -2d, \dots\} \\ 0 & \text{otherwise} \end{cases} \quad (2.49)$$

if $m = n$ the inner sum is N , otherwise is 0, so the n -th component of $\mathcal{F}^{-1}(G) = g(n)$ for all $n = 0, 1, \dots, N-1$. Therefore $\mathcal{F}^{-1}(\mathcal{F}(\mathbf{g})) = \mathbf{g}$.

The characteristic inversion using the FFT algorithm given a continuous random variable \mathcal{X} , with pdf $f_{\mathcal{X}}$ and characteristic function $\varphi_{\mathcal{X}}$, for $l, u \in \mathbb{R}$ and $N \in \mathbb{N}$, can be carried out

- as proposed in Paoletta (2007) - recognizing that:

$$\varphi_{\mathcal{X}}(t) = \int_{-\infty}^{\infty} f_{\mathcal{X}}(x)e^{jtx} dx \approx \int_l^u f_{\mathcal{X}}(x)e^{jtx} dx \approx \sum_{k=0}^{N-1} f_{\mathcal{X}}(x_k)\Delta x e^{jtx_k} = \sum_{k=0}^{N-1} X(k)e^{jtx_k}, \quad (2.50)$$

where $x_k = l + k\Delta x$, with $k = 0, 1, \dots, N-1$ and $\Delta x = (u-l)/N$. This procedure creates an even spaced grid of x 's values spanning between a lower (l) and a upper (u) bounds, with an even increment between consecutive quantities of Δx . The last equality holds once we set $X(k) = f_{\mathcal{X}}(x_k)\Delta x$.

Both approximations in (2.50) can be made arbitrarily accurate by choosing magnitudes of l and u and N large enough. The first approximation is valid because, as an improper integral, $\varphi_{\mathcal{X}}(x)$ exists, and we assume that $\lim_{x \rightarrow -\infty} f_{\mathcal{X}}(x) = \lim_{x \rightarrow \infty} f_{\mathcal{X}}(x) = 0$; while the second approximation is valid from the definition of the Riemann integral. Assuming that $\varphi_{\mathcal{X}}(r)$ is available for any r , we want to retrieve the quantities $X(k)$ therefore, dividing both sides of (2.50) by e^{jrl} :

$$\varphi_{\mathcal{X}}(r)e^{-jrl} \approx \sum_{k=0}^{N-1} X(k)e^{jrk\Delta x}. \quad (2.51)$$

Equating $r k \Delta x$ with $2\pi k n / N$ gives:

$$\sum_{k=0}^{N-1} X(k)e^{jrk\Delta x} = \sum_{k=0}^{N-1} X(k)e^{j2\pi kn/N}. \quad (2.52)$$

One can recognize that equation (2.52) is basically equal to equation (2.48), apart from the coefficient $1/N$. Therefore we obtain:

$$\sum_{k=0}^{N-1} X(k)e^{j2\pi kn/N} = Nx(n).$$

Now, let us create a N -length grid of r -values:

$$r_n = \frac{2\pi n}{N\Delta x}, \quad \text{where } n = -\frac{N}{2}, -\frac{N}{2} + 1, \dots, \frac{N}{2} - 1. \quad (2.53)$$

It follows from (2.51), (2.52) and (3.17) that:

$$x(n) \approx \frac{1}{N}\varphi_{\mathcal{X}}(r_n)e^{-jr_n l}. \quad (2.54)$$

Since $\mathbf{X} = \mathcal{F}(\mathbf{x})$ applying the DFT, via FFT (that considerably reduces the number of multiplications from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \log_2 N)$), to $x(n)$ gives $X(k)$ that provides quantities $f_{\mathcal{X}}(x_k)\Delta x$. Dividing these values by Δx we are able to obtain the pdf of \mathcal{X} . An appropriate choice of the parameters $l, u, \Delta x$ and N is essential to get a desired trade-off between accuracy and computational time.

An improved evaluation of the integral is possible using one of the quadrature formulae, see e.g. Carr et al. (1999), Cerný (2004) and Menn and Rachev (2006). A detailed theoretical treatment of numerical integration can be found in Kythe and Schäferkötter (2004) or Phillips and Taylor (1996). Hereafter composite trapezoidal, Simpson's and mid-point rules are listed along with their respective approximation errors. Let us define $h = (b - a)/N$, where a and b represent the lower and upper limits of integration, respectively, and N is the number of even subintervals in which the interval $[a, b]$ is split.

The composite trapezoidal rule is given by:

$$\begin{aligned} \int_a^b f(x)dx &\approx h \left[\frac{f(a) + f(b)}{2} + \sum_{k=1}^{N-1} f\left(a + k \frac{b-a}{N}\right) \right] \\ &\approx \frac{h}{2} \left(f(x_0) + 2f(x_1) + 2f(x_2) + \dots + 2f(x_{N-1}) + f(x_N) \right), \end{aligned}$$

where $x_k = a + kh$, $k = 0, 1, \dots, N$.

For a twice continuously differentiable function the approximation error committed by the composite trapezoidal rule is:

$$-\frac{(b-a)^3}{12N^2} f^{(2)}(\xi),$$

for some $\xi \in (a, b)$.

The composite Simpson's rule is given by:

$$\begin{aligned} \int_a^b f(x)dx &\approx \frac{h}{3} \left[f(x_0) + 2 \sum_{k=1}^{N/2-1} f(x_{2k}) + 4 \sum_{k=1}^{N/2} f(x_{2k-1}) + f(x_N) \right] \\ &\approx \frac{h}{3} \left[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \dots + 4f(x_{N-1}) + f(x_N) \right], \end{aligned}$$

where $x_k = a + kh$, $k = 0, 1, \dots, N$.

For a four times continuously differentiable function the approximation error of the composite Simpson's rule is bounded by:

$$\frac{h^4}{180}(b-a) \max_{\xi \in [a,b]} |f^{(4)}(\xi)|.$$

The composite mid-point rule is given by:

$$\int_a^b f(x)dx \approx h[f(x_1) + f(x_3) + \dots + f(x_{N-3}) + f(x_{N-1})],$$

where $x_k = a + kh$, $k = 1, 3, \dots, N-1$.

For a twice continuously differentiable function the approximation error of the composite mid-point rule is:

$$\frac{f^{(2)}(\xi)}{24} \frac{(b-a)^3}{N^2},$$

for some $\xi \in (a, b)$.

We will use the composite trapezoidal rule to produce more accurate approximations for $x(n)$ in equation (2.54).

2.6 Nonlinear portfolio with non-normal returns

This section presents a new method to efficiently compute VaR and CVaR combining the Delta-Gamma-Normal model with a probability conserving transformation. It is organized as follows: subsection 2.6.1 introduces the probability conserving transformation and subsection 2.6.2 discusses the new Delta-Gamma-Q method for assets with non-normal returns, which is the principal contribution of this chapter. Subsection 2.6.3 provides some important features of the non parametric density estimation, for a more detailed treatment one can refer e.g. to Bowman and Azzalini (1997), Fan and Gijbels (1996) and references within.

2.6.1 Probability conserving transformation

The idea of probability conserving transformation has been used in the context of analysis of portfolio strategies in Sornette et al. (2000). We employ this idea in the current context of

risk computation. Suppose that we have a set of dependent random variables $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_m$ with arbitrary marginal cdfs F_i , $i = 1, 2, \dots, m$. For each \mathcal{X}_i , samples $x_j^{(i)}$, $j = 1, 2, \dots, n$ drawn from $F_i(x)$ are available. In the present context, \mathcal{X} might represent a vector of asset returns or any other risk factors. Denoting the standard normal cdf by $\Phi(y)$, define a function of \mathcal{X}_i by

$$\psi(\mathcal{X}_i) := \Phi^{-1}F_i(\mathcal{X}_i),$$

and let $y_j^{(i)} = \psi(x_j^{(i)})$. If we envisage $y_j^{(i)}$ to be a sample of a random variable \mathcal{Y}_i with standard normal distribution, then it is clear that $\Phi(y_j^{(i)}) = F_i(x_j^{(i)})$. ψ is the so-called *probability conserving transformation*. We can then find the sample covariance matrix of \mathcal{Y} using $y_j^{(i)}$. We don't know *a priori* the joint distribution of the vector of transformed vector-valued random variable, \mathcal{Y} . However we can introduce an approximation using a standard result from information theory (Rao (2001)): conditioned only on the knowledge of the covariance matrix, the best representation of a multivariate distribution is the Gaussian distribution, in the sense that it maximizes entropy, see e.g. Cover and Thomas (1991) for details. Therefore, conditioned on the sole knowledge of the covariance matrix $\Sigma_{\mathcal{Y}}$, the best approximated parametric representation of the multivariate distribution of \mathcal{Y} is given by:

$$f_{\mathcal{Y}}(\mathbf{y}) = \frac{1}{(2\pi)^{m/2} \sqrt{\det(\Sigma_{\mathcal{Y}})}} \exp\left(-\frac{1}{2}\mathbf{y}^\top \Sigma_{\mathcal{Y}}^{-1} \mathbf{y}\right).$$

Unlike a mean-variance based normal approximation of a multivariate density, the proposed approximation heuristic preserves the fat-tailed nature of the original vector of risk factors \mathcal{X} . We use this method in the context of the VaR computation as follows.

If $F_i(x_j^{(i)})$ are not normal or if \mathcal{X}_i are not jointly normal, the VaR computation cannot be reduced to a computation of a one dimensional integral as outlined in the previous section, even after Delta-Gamma approximation of the nonlinearity. To remedy this, we carry out a nonlinear transformation to normal distributions, as proposed in Sornette et al. (2000): let ϕ be the standard normal pdf and Φ be the standard normal cdf. Let $f(u)$ be a density, $F(u)$ be its corresponding cdf and define the function

$$\Phi(w) = F(u), \quad i.e. \quad \frac{1}{\sqrt{2\pi}} \int_0^w e^{-\frac{z^2}{2}} dz = \int_0^u f(z) dz,$$

where the random variable $\mathcal{W} \sim \mathcal{N}(0, 1)$ by construction. It is therefore possible to map each value x_j into a new variable y_j :

$$y_j = \psi(x_j) = \Phi^{-1}(F(x_j)) = \sqrt{2} \operatorname{erf}^{-1}(2F(x_j) - 1),$$

where the error function operator, erf , is defined as:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

Referring back to notation introduced previously in this section since historic data is available (i.e. $x_j^{(i)}$, with $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$ where m is the number of factors and n is the number of data available for each factor), it is possible to compute $y_j^{(i)}$ for all i and j . The covariance matrix of the vector variable \mathcal{Y} is defined as $\Sigma_{\mathcal{Y}} = \mathbb{E}[\mathbf{y}\mathbf{y}^T]$, where each element $[\Sigma_{\mathcal{Y}}]_{ab}$ is obtained by sample average approximation:

$$[\Sigma_{\mathcal{Y}}]_{ab} = \frac{1}{n} \sum_{l=1}^n y_l^{(a)} y_l^{(b)}.$$

We need to find δ and Γ in terms of the transformed variables \mathcal{Y}_i (and their samples $y_j^{(i)}$), which will require finding the derivative $\frac{dx_i}{dy_i}$. Having $\Phi(\mathcal{Y}_i) = F_i(\mathcal{X}_i)$, we can write $\mathcal{X}_i = F_i^{-1}(\Phi(\mathcal{Y}_i))$, where \mathcal{Y}_i is a standard normal random variable and \mathcal{X}_i is a random variable with cdf F_i . Since Φ and F_i are both continuous with densities ϕ and f_i , respectively, it follows that the functional $\mathcal{X}_i = F_i^{-1}(\Phi(\mathcal{Y}_i))$ is continuously differentiable such that:

$$\frac{dx_i}{dy_i} = \frac{\phi(y_i)}{f_i(F_i^{-1}(\Phi(y_i)))}. \quad (2.55)$$

Now, remembering that $\Phi(\mathcal{Y}_i) = F_i(\mathcal{X}_i)$, equation (2.55) can be simplified as:

$$\frac{dx_i}{dy_i} = \frac{\phi(y_i)}{f_i(x_i)}. \quad (2.56)$$

Nonparametric estimation of distribution of \mathcal{X} allows to achieve maximum flexibility in obtaining quantities $F_i(x_j^{(i)})$ and $f_i(x_j^{(i)})$.

2.6.2 Delta-Gamma-Q for assets with non-normal returns

Here, we bring together the material in subsections 2.5.1 and 2.6.1 to propose a novel heuristic for risk computation for a nonlinear portfolio with possibly non-Gaussian risk factors. Given a series of changes in risk factors (e.g. asset prices) $\Delta\mathcal{S}_i$, we start by computing their probability conserving transformation $\Delta\mathcal{Y}_i$:

$$\Delta\mathcal{Y}_i = \Phi^{-1}(F_i(\Delta\mathcal{S}_i)), \quad i = 1, 2, \dots, m. \quad (2.57)$$

Each $\Delta\mathcal{Y}_i \sim \mathcal{N}(0, 1)$, and the vector $\Delta\mathcal{Y} \sim \mathcal{N}(0, \Sigma_{\mathcal{Y}})$, with $\text{diag}(\Sigma_{\mathcal{Y}}) = (1, 1, \dots, 1)$. We emphasize the fact that the matrix $\Sigma_{\mathcal{Y}}$ is generally not an identity matrix and depends on sample correlations between the elements of $\Delta\mathcal{Y}$. Since $\Delta\mathcal{Y}_i$ have been represented jointly normal, VaR and CVaR computations can be estimated using an adapted version of (2.20):

$$\Delta\Pi(\mathcal{S}, t) \approx \Theta_{\mathcal{Y}}\Delta t + \delta_{\mathcal{Y}}^T\Delta\mathcal{Y} + \frac{1}{2}\Delta\mathcal{Y}^T\Gamma_{\mathcal{Y}}\Delta\mathcal{Y}, \quad (2.58)$$

with $\Delta\mathcal{Y}$ derived from $\Delta\mathcal{S}$ using (2.57). Equation (2.58) can be converted to an one-dimensional integral only provided we can find sensitivities $\delta_{\mathcal{Y}}$ and $\Gamma_{\mathcal{Y}}$. Now, note that:

$$\Theta = \Theta_{\mathcal{Y}}, \quad (2.59)$$

$$[\delta_{\mathcal{Y}}]_i = \frac{\partial\Pi}{\partial\mathcal{Y}_i} = \frac{\partial\Pi}{\partial\mathcal{S}_i} \frac{d\mathcal{S}_i}{d\mathcal{Y}_i} = [\delta]_i \frac{d\mathcal{S}_i}{d\mathcal{Y}_i}, \quad \text{and} \quad (2.60)$$

$$[\Gamma_{\mathcal{Y}}]_{ij} = \frac{\partial}{\partial\mathcal{Y}_j} \left(\frac{\partial\Pi}{\partial\mathcal{Y}_i} \right) = \frac{\partial}{\partial\mathcal{S}_j} \left(\frac{\partial\Pi}{\partial\mathcal{S}_i} \frac{d\mathcal{S}_i}{d\mathcal{Y}_i} \right) \frac{d\mathcal{S}_i}{d\mathcal{Y}_j} = [\Gamma]_{ij} \left(\frac{d\mathcal{S}_i}{d\mathcal{Y}_i} \right) \left(\frac{d\mathcal{S}_i}{d\mathcal{Y}_j} \right), \quad (2.61)$$

where δ and Γ are sensitivities under normal conditions, while the derivatives $d\mathcal{S}_k/d\mathcal{Y}_k$ are computed using formula (2.55). Equation (2.58) has got the same structure as of equation (2.20). Hence we can apply the characteristic function inversion using the fast Fourier transform to find approximate VaR and CVaR of the portfolio using Delta-Gamma approximation. We will call this new method as Delta-Gamma-Q method for assets with non-normal returns. The name for the novel method expresses a mix between the Delta-Gamma-Normal modeling and the approximation \mathcal{Q} of portfolio losses \mathcal{L} provided by equation (2.22).

We summarize the steps for computation of VaR and CVaR using Delta-Gamma-Q method below:

- Find option sensitivities δ , Γ and Θ assuming normality for risk-factors;
- Apply nonparametric estimation to the distribution of the relevant risk-factors \mathcal{S} ;
- Map $\Delta\mathcal{S}$ into $\Delta\mathcal{Y}$ through probability conserving transformation;
- Determine the covariance matrix $\Sigma_{\mathcal{Y}}$;
- Find for each risk-factor the coefficient that expresses its non-normality by applying equation (2.55) and by averaging values;
- Evaluate the coefficients $\delta_{\mathcal{Y}}$ and $\Gamma_{\mathcal{Y}}$ using equations (2.60) and (2.61) respectively;
- Calculate characteristic function coefficients $a_{\mathcal{Y}} = -\Theta_{\mathcal{Y}}\Delta t$, $\mathbf{b}_{\mathcal{Y}} = -C_{\mathcal{Y}}^T\delta_{\mathcal{Y}}$ and $\Lambda_{\mathcal{Y}} = -\frac{1}{2}C_{\mathcal{Y}}^T\Gamma_{\mathcal{Y}}C_{\mathcal{Y}}$, where $C_{\mathcal{Y}}$ is a square root of $\Sigma_{\mathcal{Y}}$ as described in Section 2.5.1;
- Compute VaR and CVaR, either by evaluating the integral in equation (2.50) directly with appropriate limits or by using FFT.

To re-emphasize the point of this exercise, we are trying to re-gain the simplicity of computing the quantiles of a quadratic form for a Gaussian distribution, while still preserving the tail information in the marginal risk factor distributions. Approximating a distribution by a normal distribution via probability preserving transformation can lead to far better quantile estimates than a mean-variance based approximation, as the experiments in the subsequent chapter show.

Note that our VaR evaluation is static, i.e. we are computing VaR over a single time-step. As such, we are using nonparametric density estimation to model the distribution of risk factors and will then map them into normally distributed risk factors. The underlying stochastic process which generates the said nonparametric distribution is of no direct relevance in our current framework. Linking the evolution through time of the risk factors (e.g. as a Lévy process) to the evolution of quantiles of their multivariate nonlinear function (such as the portfolio net worth) is an interesting topic which is outside the scope of this paper and is a topic of current research.

2.6.3 Nonparametric density estimation

The notion of probability density function is crucial in statistical analysis and nonparametric density estimation provides a good approximation to the probability density function given

a sample drawn from a generic distribution. In parametric estimation the density is assumed to come from a certain family of distributions (i.e. Gaussian, binomial, Poisson, etc.) and the parameters are estimated using different statistical methods. In nonparametric estimation no assumption about the kind of distribution from which the samples are drawn is made.

Having a large number of observations available $\{x_1, x_2, \dots, x_n\}$ on an unidimensional random variable \mathcal{X} having density $f_{\mathcal{X}}(x)$, the simplest way to represent its distributional shape is using a histogram. The range of the observed values is divided in a number n of intervals and the number of samples that fall into each interval provides the frequency for that interval. Denoting x the generic point at which the density $f_{\mathcal{X}}(x)$ has to be assessed, then the histogram can be described as:

$$\bar{f}_{\mathcal{X}}(x) = \sum_{i=1}^n \mathbb{I}_{(x-\bar{x}_i, h)},$$

where $x_i, i = 1, 2, \dots, n$ are the sample data, $\mathbb{I}(z, h)$ is the indicator function of the interval $[-h, h]$ and \bar{x}_i is the center of the interval that contains x_i . The function $\bar{f}_{\mathcal{X}}$ needs to be scaled to integrate to one to be considered a density. The use of a histogram as a density estimate has three flaws:

1. Replacement of x_i with the central point of the interval \bar{x}_i cause loss of information;
2. The density function describing a random variable is usually assumed smooth and the histogram is not a smooth estimator;
3. The choice of the interval width h affects the behavior of the estimator.

The kernel density estimator provides a solution to the first two problems. The kernel density estimate of $f_{\mathcal{X}}(x)$ is given by:

$$\hat{f}_{\mathcal{X}}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right),$$

where the kernel $K(\cdot)$ is a function centered over each observation and that smoothly redistribute the probability mass at each datum point around its vicinity, and h is the smoothing parameter or bandwidth. The kernel $K(\cdot)$ is a generally chosen to be a continuous unimodal

probability density symmetric around zero such that:

$$\begin{aligned}\int_{\mathbb{R}} K(z) dz &= 1, \\ \int_{\mathbb{R}} zK(z) dz &= 0, \\ \int_{\mathbb{R}} z^2K(z) dz &> 0.\end{aligned}$$

Popular choices of $K(\cdot)$ are:

- Gaussian: $K(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}$;
- Uniform: $K(z) = \frac{1}{2}\mathbb{I}_{|z|\leq 1}$;
- Triangular: $K(z) = (1 - |z|)\mathbb{I}_{|z|\leq 1}$;
- Epanechnikov: $K(z) = \frac{3}{4}(1 - z^2)\mathbb{I}_{|z|\leq 1}$.

The choice of the bandwidth h affects the kernel density estimates, hence has to be done carefully. Choosing a small h leads to estimates that displays variation associated with individual observation while choosing a large h tends to provide an obscured structure of the original data.

Some basic statistical properties of the kernel density estimators are:

$$\begin{aligned}\mathbb{E}[\hat{f}_{\mathcal{X}}(x)] &= f_{\mathcal{X}}(x) + \mathcal{O}(h^2); \\ \text{var}(\hat{f}_{\mathcal{X}}(x)) &\approx \frac{f_{\mathcal{X}}(x)}{nh} \int_{\mathbb{R}} K^2(z) dz.\end{aligned}$$

Therefore the bias of $\hat{f}_{\mathcal{X}}(x)$ decreases as h get smaller, while its variance tends to zero as nh tends to infinity. There is a trade-off between undersmoothing (if the bandwidth is too small, the variance becomes large) and oversmoothing (if the bandwidth is too large, the bias becomes large). There is still much debate about choice of the optimal smoothing parameter. The theory suggests that the bandwidth h should be proportional to $n^{-1/5}$, but the optimal constant of proportionality depends on the unknown density. A “rule of thumb” frequently used is the one devised in Silverman (1986):

$$\hat{h}_{opt} = 1.06 \min\left(\hat{\sigma}, \frac{IQR}{1.34}\right)n^{-1/5},$$

where $\hat{\sigma}$ and IQR are the sample standard deviation and the interquartile range (the difference between the 75th and the 25th percentile) of the considered random variable, respectively.

The in-built MATLAB function *ksdensity* employed for the numerical experiments allows to specify different parameters including the type of kernel smoother, the vector of values where the density estimate is to be evaluated, the function type to estimate (e.g. pdf, cdf) and the bandwidth. In particular, the default selected bandwidth is the optimal bandwidth for estimating normal densities (see Bowman and Azzalini (1997)):

$$\hat{h}_{\mathcal{N}} = \left(\frac{4}{3n} \right)^{1/5} \hat{\sigma},$$

where $\hat{\sigma}$ denotes the sample standard deviation of the considered random variable.

2.7 Summary

The first part of this chapter includes some methods to model four different scenarios for a financial portfolio, which are:

- Linear portfolio with normal returns;
- Linear portfolio with non-normal returns;
- Nonlinear portfolio with normal returns;
- Nonlinear portfolio with non-normal returns.

Our novel method, the Delta-Gamma-Q method, is devised to model a nonlinear portfolio with non-normal returns and provide its measure of risk. This method is described in subsection 2.6.2. The nonparametric density estimation is used to capture the features of the risk factors included in the portfolio.

All the included methods are generic and are also applicable to bond portfolios. In particular, the historical simulation method and the variance/covariance method will be employed later in chapter 6 for a comparison with our heuristic for measuring the risk of a fixed income portfolio.

Chapter 3

Mathematical tools for modeling risk in fixed income portfolios

3.1 Introduction

We limit ourselves to the risk of losses for fixed income portfolios which arises solely due to adverse movements of interest rates. Possible losses related to defaults are not considered, i.e. the portfolios are considered to be default-free. Note that the models and tools to address nonlinear portfolios of fixed income securities are completely different from those discussed in the previous chapter and are hence discussed separately in this chapter. Bond prices are function of the short rate $r(t)$, which is a latent variable. We use the Kalman Filter to derive the short rate using one linear model (the Vasicek model) and one nonlinear model (the CIR model). For both linear and nonlinear models there are one-factor and multiple-factors versions. One factor models prove inadequate to describe the term structure, that is why researchers often prefer to rely on three-factor models. Our choice was to opt for the more parsimonious two-factor models, based on some evidence that for one-step ahead prediction using two-factor models can be more accurate than one-step ahead prediction using three-factor models. A two factor model seems to offer a good compromise between the difficulty of calibrating a three factor model and poor accuracy of a one factor model. Hence a two factor model has been chosen for the current analysis. The chapter is organized as follows: section 3.2 provides preliminary definitions for interest rate models, while section 3.3 introduces the version of Kalman filter adopted. Sections 3.4 and 3.5 present the Vasicek

model and the CIR model, respectively, which will later be used in numerical experiments in chapter 6.

3.2 Preliminary definitions for interest rate models

For clarity, some essential definitions are henceforth briefly introduced. A zero-coupon bond is a contract that pays at its maturity, T , one unit of its currency. Its value at time t is denoted by $P(t, T)$, with $t < T$. Given the price of a pure discount bond having maturity T , the bond yield (or spot rate or zero-coupon yield) associated to a particular date t is given by:

$$y(t, T) = -\frac{\ln P(t, T)}{T - t}, \quad (3.1)$$

while the *forward rate* is defined¹ by:

$$f(t, T) = -\frac{\partial \ln P(t, T)}{\partial T}. \quad (3.2)$$

The instantaneous interest rate (or *short rate*) at time t is given by:

$$r(t) = f(t, t),$$

and the price of a zero-coupon bond with maturity T at time t is given by:

$$P(t, T) = \mathbb{E}_{\mathbb{Q}}[e^{-\int_t^T r(u)du} | \mathcal{F}_t],$$

where \mathcal{F}_t is the natural filtration for the process and the expectation is taken under a risk neutral measure \mathbb{Q} . Hence modeling variation of interest rate r over time affects the evolution of bond prices and other derivative prices, and ultimately bond yields. Bond yields are observable quantities, while $r(t)$ is a latent variable. We need to predict $r(t)$ if we want arbitrage-free forecasts on bond yields. In exponential affine models, yields depend affinely on the latent variable $r(t)$. In such cases, it is possible to estimate latent variables recursively in a computationally tractable fashion from observable bond yields using the Kalman filter, which is described in section 3.3.

¹Equations (3.1) and (3.2) can be equivalently rewritten as $P(t, T) = e^{-y(t, T)(T-t)}$ and $P(t, T) = e^{-\int_t^T f(t, u)du}$, respectively.

3.3 The Kalman filter

Bond prices are nonlinear functions of short rate, therefore representing their prices and predicting their future values require an adequate modeling and an accurate calibration of the parameters. Short rates are latent quantities that can be derived from observable bond prices $P(t_n, T_j)$ at discrete times t_n , $n = 1, 2, \dots, N$ and for discrete maturities $j = 1, 2, \dots, m$. As it will be clarified in later subsections 3.5.2 and 3.4.2 the short rate $r(t)$ is here modeled as sum of independent $r_1(t)$ and $r_2(t)$ both following a Vasicek and a CIR process, respectively. Accordingly, bond yields are modeled as affine functions of $r_1(t)$ and $r_2(t)$. The *Kalman filter* (KF), firstly proposed in Kalman (1960), is a recursive moment estimator often employed to model affine state space systems to retrieve hidden values (short rates) from measurable values (bond prices). We describe a general formulation of Kalman filter below.

Let us consider a discrete time, state space system:

$$\mathbf{r}_{n+1} = A\mathbf{r}_n + \mathbf{b} + \mathbf{w}_{n+1}, \quad (3.3)$$

$$\mathbf{y}_n = C\mathbf{r}_n + \mathbf{d} + \mathbf{z}_n, \quad (3.4)$$

where \mathbf{w}_n and \mathbf{z}_n are zero mean, Gaussian and uncorrelated random variables at each time t_n . A , \mathbf{b} , C , \mathbf{d} , $G = \mathbb{E}[\mathbf{w}_n \mathbf{w}_n^T]$ and $H = \mathbb{E}[\mathbf{z}_n \mathbf{z}_n^T]$ can form either:

- A linear system, if all of them are constants or are known functions of time;
- A nonlinear system, if one or more of them are a nonlinear function of r_n and t_n .

The variable \mathbf{y}_n is the only observable variable, while \mathbf{r}_n can not be directly observed and needs to be estimated. Each time-step $\Delta t = t_n - t_{n-1}$ is assumed to be constant. Equations (3.3) and (3.4) are referred to as the *transition equations* and the *measurement equations*, respectively. We have considered one linear model and one nonlinear model: the two-factor Vasicek model and the two-factor Cox-Ingersoll-Ross (CIR), respectively.

There are several different versions of the KF equations including the ones reported e.g. in Date and Ponomareva (2011) and in Date and Wang (2009). The KF consists in a set of

recursive equations; the one employed in this paper involves the following:

$$\mathbf{v}_n = \mathbf{y}_n - C\hat{\mathbf{r}}_{n|n-1} - \mathbf{d}, \quad (3.5)$$

$$\Sigma_n = CV_{n|n-1}C^T + H, \quad (3.6)$$

$$K_n = V_{n|n-1}C^T\Sigma_n^{-1}, \quad (3.7)$$

$$\hat{\mathbf{r}}_{n|n} = \hat{\mathbf{r}}_{n|n-1} + K_nv_n, \quad (3.8)$$

$$\hat{\mathbf{r}}_{n+1|n} = A\hat{\mathbf{r}}_{n|n} + \mathbf{b}, \quad (3.9)$$

$$V_{n+1|n} = AV_{n|n-1}A^T + G - AV_{n|n-1}C^T\Sigma_n^{-1}CV_{n|n-1}A^T. \quad (3.10)$$

The estimation of the variable of interest \mathbf{r}_n and the conditional variance of the estimate V_n based on information up to time t_{n-1} are respectively denoted by $\hat{\mathbf{r}}_{n|n-1}$ and $V_{n|n-1}$. Initial estimates $\hat{\mathbf{r}}_{0|0}$ and $V_{1|0}$ are assumed to be known or can themselves be parameterized. Innovations \mathbf{v}_n and their covariance matrix Σ_n are expressed by equations (3.5) and (3.6), while K_n , often referred to Kalman gain, is given by equation (3.7). Equation (3.8) represents the filtered state vector. The state vector and the covariance matrix predictions are respectively provided by equations (3.9) and (3.10). The calibration of the set of parameters which characterize matrices A , B , G , H and vectors \mathbf{b} , \mathbf{d} are obtained through maximum likelihood. Having the set of observations $Y = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\}$ and since $\mathbf{y}_{n+1|n} \sim \mathcal{N}(C\hat{\mathbf{r}}_{n+1|n} + \mathbf{d}, \Sigma_{n+1})$, we can express the loglikelihood function ignoring the constant terms:

$$\log L(Y) = -\frac{1}{2} \sum_{i=1}^N (\log |\Sigma_i| + \mathbf{v}_i^T \Sigma_i^{-1} \mathbf{v}_i). \quad (3.11)$$

The expression (3.11) can be maximized using standard solvers, such as `fminsearch` from MATLAB 7.9 used in this work. Once the optimal parameters are obtained, one can forecast successive values for the latent variable \mathbf{r} as long as new observations \mathbf{y} become available. This is implemented employing the recursive equations (3.5)-(3.10) with A , \mathbf{b} , C , \mathbf{d} , G and H expressed as function of the optimized parameters.

3.4 A linear system: The Vasicek model

This section presents how to model the short rate $r(t)$ initially using a single Gaussian source of uncertainty as in Vasicek (1977) and then extends it to two sources of uncertainty.

3.4.1 Single factor Vasicek model

Let us consider a short rate following an Ornstein-Uhlenbeck process under the risk neutral measure \mathbb{Q} :

$$dr(t) = k(\theta - r(t))dt + \sigma dW(t), \quad r(0) = r_0, \quad (3.12)$$

where r_0 , k , θ and σ are positive constants, and $W(t)$ is a \mathbb{Q} -Wiener process. The coefficients k , θ and σ are often referred to as “speed of reversion”, “long term mean level” and “instantaneous volatility”, respectively. Equation (3.12) captures the phenomenon of the mean reversion: interest rates move in a limited range, displaying a tendency to revert to the long term mean level θ , at a speed k . The short rate $r(t)$ conditional to \mathcal{F}_s is normally distributed with mean and variance (see, e.g. Brigo and Mercurio (2006)):

$$\mathbb{E}[r(t)|\mathcal{F}_s] = r(s)e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}), \quad (3.13)$$

$$\text{var}(r(t)|\mathcal{F}_s) = \frac{\sigma^2}{2k}(1 - e^{-2k(t-s)}). \quad (3.14)$$

Formulae (3.36) and (3.37) imply that, under the Vasicek dynamics, the interest rate $r(t)$ can assume negative values, although the probability of a negative value is extremely low for realistic values of parameters. The bond price function for the single factor Vasicek model has the following analytical form:

$$P(t, T, r(t)) = e^{E(t, T) - F(t, T)r(t)},$$

where

$$E(t, T) = \frac{(k^2(\theta - \frac{\sigma\lambda}{k}) - \frac{\sigma^2}{2})(F(t, T) - (T - t))}{k^2} - \frac{\sigma^2 F^2(t, T)}{4k}, \quad (3.15)$$

$$F(t, T) = \frac{1}{k}(1 - e^{-k(T-t)}), \quad (3.16)$$

and λ is the market price of risk. It can be noticed that the bond yield $y(t, T)$, as defined in (3.1), is an affine function of the short rate $r(t)$. The relationship between the interest rates and time to maturity is called term-structure of interest rates. Interest-rate models having bond yield which can be written as an affine function of the short rate are referred to as affine term-structure models.

3.4.2 Two-factor Vasicek model

The assumption of two-factor Vasicek model is that the short term interest rates are given by the sum of two state variables, each of them following an Ornstein-Uhlenbeck process. A single factor model has only one source of uncertainty, and hence accounts only for parallel shifts in yields. A two factor model can provide a better description of how term structure changes through time. Let us consider two independent state variables that follow linear, mean reverting Gaussian process under the risk neutral measure \mathbb{Q} :

$$r(t) = r_1(t) + r_2(t), \quad (3.17)$$

$$dr_i(t) = k_i(\theta_i - r_i(t))dt + \sigma_i dW_i(t), \quad r_i(0) = r_{i0}, \quad i = 1, 2, \quad (3.18)$$

where r_{i0} , k_i , θ_i and σ_i are positive constants, and $W_i(t)$ is a \mathbb{Q} -Wiener process. Each $r_i(t)$ conditional to \mathcal{F}_s is normally distributed with mean and variance:

$$\mathbb{E}[r_i(t)|\mathcal{F}_s] = r_i(s)e^{-k_i(t-s)} + \theta_i(1 - e^{-k_i(t-s)}), \quad i = 1, 2, \quad (3.19)$$

$$\text{var}(r_i(t)|\mathcal{F}_s) = \frac{\sigma_i^2}{2k_i}(1 - e^{-2k_i(t-s)}), \quad i = 1, 2. \quad (3.20)$$

We discretize equations (3.19) and (3.20) considering evenly spaced observation times $t_1 \leq t_2 \leq \dots \leq t_N$, with $t_{n+1} - t_n = \Delta t$, obtaining (see, e.g. Bolder (2001)) the following transition equations in the same form as expressed in (3.3):

$$\begin{bmatrix} r_{1\ n+1} \\ r_{2\ n+1} \end{bmatrix} = \begin{bmatrix} e^{-k_1\Delta t} & 0 \\ 0 & e^{-k_2\Delta t} \end{bmatrix} \begin{bmatrix} r_{1\ n} \\ r_{2\ n} \end{bmatrix} + \begin{bmatrix} \theta_1(1 - e^{-k_1\Delta t}) \\ \theta_2(1 - e^{-k_2\Delta t}) \end{bmatrix} + \begin{bmatrix} w_{1\ n+1} \\ w_{2\ n+1} \end{bmatrix}, \quad (3.21)$$

where $\mathbf{w}_{n+1|n} \sim \mathcal{N}(\mathbf{0}, G)$, with:

$$G = \begin{bmatrix} \frac{\sigma_1^2}{2k_1}(1 - e^{-2k_1\Delta t}) & 0 \\ 0 & \frac{\sigma_2^2}{2k_2}(1 - e^{-2k_2\Delta t}) \end{bmatrix}. \quad (3.22)$$

Let us assume that the each state variable that makes up the short interest rate follows a linear, mean reverting Gaussian process with the same volatility but a different drift function under a measure \mathbb{P} :

$$dr_i(t) = k_i(\tilde{\theta}_i - r_i(t))dt + \sigma_i d\tilde{W}_i(t), \quad r_i(0) = r_{i0}, \quad i = 1, 2, \quad (3.23)$$

where $\widetilde{W}_i(t)$ is a \mathbb{P} -Wiener process. Our assumption of an arbitrage-free market indirectly implies the existence of market price of risk processes $\lambda_{i,t}$ such that $\tilde{\theta}_i - \theta_i = \sigma_i \lambda_{i,t}$ for $i = 1, 2$ hold. It is common practice (see, e.g. De Rossi (2004) or Vasicek (1977)) to assume $\lambda_{i,t}$ to be constants, independent of t and $r_i(t)$.

The bond price function for the two-factor Vasicek model has the following analytical form:

$$P(t, T, r_1(t), r_2(t)) = e^{E(t, T) - F_1(t, T) r_1(t) - F_2(t, T) r_2(t)},$$

where

$$E(t, T) = \sum_{i=1}^2 \frac{(k_i^2(\theta_i - \frac{\sigma_i \lambda_i}{k_i}) - \frac{\sigma_i^2}{2})(F_i(t, T) - (T - t))}{k_i^2} - \frac{\sigma_i^2 F_i^2(t, T)}{4k_i}, \quad (3.24)$$

$$F_i(t, T) = \frac{1}{k_i}(1 - e^{-k_i(T-t)}), \quad i = 1, 2, \quad (3.25)$$

and λ_i is the market price of risk for the i^{th} factor. Even though $r_1(t)$ and $r_2(t)$ are uncorrelated, the bond prices (and the yields) are correlated via $r_1(t)$, $r_2(t)$. The measurement system we used involves the following relationship between zero-coupon yields and the price of zero-coupon bonds:

$$y(t, T) = -\frac{\ln P(t, T)}{T - t} = \frac{-E(t, T) + \sum_{i=1}^2 F_i(t, T) r_i(t)}{T - t}. \quad (3.26)$$

Using this equation at each t_n , for a set of m bonds with maturities T_1, \dots, T_m leads to the following vector valued equation:

$$\begin{bmatrix} y(t_n, T_1) \\ y(t_n, T_2) \\ \vdots \\ y(t_n, T_m) \end{bmatrix} = \begin{bmatrix} \frac{F_1(t_n, T_1)}{T_1 - t_n} & \frac{F_2(t_n, T_1)}{T_1 - t_n} \\ \frac{F_1(t_n, T_2)}{T_2 - t_n} & \frac{F_2(t_n, T_2)}{T_2 - t_n} \\ \vdots & \vdots \\ \frac{F_1(t_n, T_m)}{T_m - t_n} & \frac{F_2(t_n, T_m)}{T_m - t_n} \end{bmatrix} \begin{bmatrix} r_1(t_n) \\ r_2(t_n) \end{bmatrix} + \begin{bmatrix} -\frac{E(t_n, T_1)}{T_1 - t_n} \\ -\frac{E(t_n, T_2)}{T_2 - t_n} \\ \vdots \\ -\frac{E(t_n, T_m)}{T_m - t_n} \end{bmatrix} + \begin{bmatrix} z_1(t_n) \\ z_2(t_n) \\ \vdots \\ z_m(t_n) \end{bmatrix}, \quad (3.27)$$

where $z_i(t_n) \sim \mathcal{N}(0, H)$ are noise variables which reflect deviation of bond prices from the model price and $H = \text{diag}(h_1^2, h_2^2, \dots, h_m^2)$, where h_i are positive constants. Equations (3.19) and (3.20) provide the transition equation as in (3.3) and equations (3.27) supply measurement equation as in (3.4). Hence they form a linear state space system with bond

yields as observable variables, so that Kalman filtering equations described in the last section can be applied for model calibration and forecasting. These forecasts will then be used for predicting the tail losses of both simulated and real bond portfolios in sections 6.2 and 6.3, respectively.

3.5 A nonlinear system: The CIR model

This section, similarly to what is introduced in the previous section, presents how to model the short rate r initially using a single source of uncertainty as proposed in Cox et al. (1985) and then extends it to two sources of uncertainty.

3.5.1 A single factor CIR model

Let us consider a short rate following a CIR process under the risk neutral measure \mathbb{Q} :

$$dr(t) = k(\theta - r(t))dt + \sigma\sqrt{r(t)}dW(t), \quad r(0) = r_0, \quad (3.28)$$

where r_0 , k , θ and σ are positive constants, and $W(t)$ is a \mathbb{Q} -Wiener process. The CIR model, like the Vasicek model, is mean reverting and the short rate remains positive if the following condition is respected:

$$2k\theta > \sigma^2. \quad (3.29)$$

The short rate $r(t)$ conditional to \mathcal{F}_s has mean and variance:

$$\mathbb{E}[r(t)|\mathcal{F}_s] = r(s)e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}), \quad (3.30)$$

$$\text{var}(r(t)|\mathcal{F}_s) = r(s)\frac{\sigma^2}{k}(e^{-k(t-s)} - e^{-2k(t-s)}) + \theta\frac{\sigma^2}{2k}(1 - e^{-k(t-s)})^2. \quad (3.31)$$

The bond price function for the single factor CIR model has the following analytical form:

$$P(t, T, r(t)) = e^{E'(t, T) - F'(t, T) r(t)},$$

where

$$E'(t, T) = \ln \left(\frac{2\gamma e^{\frac{(\gamma+k+\lambda)(T-t)}{2}}}{(\gamma+k+\lambda)(e^{\gamma(T-t)} - 1) + 2\gamma} \right)^{\frac{2k\rho}{\sigma^2}}, \quad (3.32)$$

$$F'(t, T) = \frac{2(e^{\gamma(T-t)} - 1)}{(\gamma+k+\lambda)(e^{\gamma(T-t)} - 1) + 2\gamma}, \quad (3.33)$$

where

$$\gamma = \sqrt{(k+\lambda)^2 + 2\sigma^2},$$

and λ is the market price of risk. Proof of this result can be found, for example, in Bolder (2001).

3.5.2 Two-factor CIR model

The assumption of two-factor CIR model is that the short term interest rates are given by the sum of two state variables. Both the state variables include a “square-root” term in the diffusion coefficient of the instantaneous short-rate dynamics. Let us consider two independent state variables that follow this mean reverting process under the risk neutral measure \mathbb{Q} :

$$r(t) = r_1(t) + r_2(t), \quad (3.34)$$

$$dr_i(t) = k_i(\theta_i - r_i(t))dt + \sigma_i\sqrt{r_i(t)}dW_i(t), \quad r_i(0) = r_{i0}, \quad i = 1, 2, \quad (3.35)$$

where r_{i0} , k_i , θ_i and σ_i are positive constants, and $W_i(t)$ is a \mathbb{Q} -Wiener process. The mean and the variance of each $r_i(t)$ conditional to \mathcal{F}_s , for $i = 1, 2$, are given by:

$$\mathbb{E}[r_i(t)|\mathcal{F}_s] = r_i(s)e^{-k_i(t-s)} + \theta_i(1 - e^{-k_i(t-s)}), \quad (3.36)$$

$$var(r_i(t)|\mathcal{F}_s) = r_i(s)\frac{\sigma_i^2}{k_i}(e^{-k_i(t-s)} - e^{-2k_i(t-s)}) + \theta_i\frac{\sigma_i^2}{2k_i}(1 - e^{-k_i(t-s)})^2. \quad (3.37)$$

We discretize equations (3.36) and (3.37) considering evenly spaced observation times $t_1 \leq t_2 \leq \dots \leq t_N$, with $t_{n+1} - t_n = \Delta t$, obtaining the following transition equations in the same form as expressed in (3.3):

$$\begin{bmatrix} r_{1\ n+1} \\ r_{2\ n+1} \end{bmatrix} = \begin{bmatrix} e^{-k_1\Delta t} & 0 \\ 0 & e^{-k_2\Delta t} \end{bmatrix} \begin{bmatrix} r_{1\ n} \\ r_{2\ n} \end{bmatrix} + \begin{bmatrix} \theta_1(1 - e^{-k_1\Delta t}) \\ \theta_2(1 - e^{-k_2\Delta t}) \end{bmatrix} + \begin{bmatrix} w_{1\ n+1} \\ w_{2\ n+1} \end{bmatrix}, \quad (3.38)$$

where $\mathbf{w}_{n+1|n} \sim \mathcal{N}(\mathbf{0}, G')$, with:

$$G' = \begin{bmatrix} \zeta_1 & 0 \\ 0 & \zeta_2 \end{bmatrix}, \quad (3.39)$$

where

$$\zeta_i = \frac{\theta_i \sigma_i^2}{2k_i} (1 - e^{-k_i \Delta t})^2 + \frac{\sigma_i^2}{k_i} (e^{-k_i \Delta t} - e^{-2k_i \Delta t}) r_{in}, \quad i = 1, 2.$$

Let us assume that each state variable that makes up the short interest rate follows a CIR process with the same volatility but a different drift function under a measure \mathbb{P} :

$$dr_i(t) = k_i(\tilde{\theta}_i - r_i(t))dt + \sigma_i \sqrt{r_i(t)} d\tilde{W}_i(t), \quad r_i(0) = r_{i0}, \quad i = 1, 2, \quad (3.40)$$

where $\tilde{W}_i(t)$ is a \mathbb{P} -Wiener process. Our assumption of an arbitrage-free market indirectly implies the existence of processes $\lambda_{i,t}$ such that $\tilde{\theta}_i - \theta_i = \sigma_i \lambda_{i,t}$ for $i = 1, 2$ hold. It is common practice to assume $\lambda_{i,t}$ to be constants, independent of t and $r_i(t)$.

The bond price function for the two-factor CIR model has the following analytical form:

$$P(t, T, r_1(t), r_2(t)) = e^{\sum_{i=1}^2 (E'_i(t, T) - F'_i(t, T) r_i(t))},$$

where

$$E'_i(t, T) = \ln \left(\frac{2\gamma_i e^{\frac{(\gamma_i + k_i + \lambda_i)(T-t)}{2}}}{(\gamma_i + k_i + \lambda_i)(e^{\gamma_i(T-t)} - 1) + 2\gamma_i} \right)^{\frac{2k_i \theta_i}{\sigma_i^2}}, \quad i = 1, 2, \quad (3.41)$$

$$F'_i(t, T) = \frac{2(e^{\gamma_i(T-t)} - 1)}{(\gamma_i + k_i + \lambda_i)(e^{\gamma_i(T-t)} - 1) + 2\gamma_i}, \quad i = 1, 2, \quad (3.42)$$

where

$$\gamma_i = \sqrt{(k_i + \lambda_i)^2 + 2\sigma_i^2}, \quad i = 1, 2,$$

and λ_i is the market price of risk for the i^{th} factor. Even though $r_1(t)$ and $r_2(t)$ are independent, the bond prices (and the yields) are correlated via $r_1(t)$, $r_2(t)$. The measurement system we used involves the following relationship between zero-coupon yields and the price of zero-coupon bonds:

$$y(t, T) = -\frac{\ln P(t, T)}{T - t} = \frac{\sum_{i=1}^2 (F'_i(t, T) r_i(t) - E'_i(t, T))}{T - t}. \quad (3.43)$$

Using this equation at each t_n , for a set of m bonds with maturities T_1, \dots, T_m leads to the following vector valued equation:

$$\begin{bmatrix} y(t_n, T_1) \\ y(t_n, T_2) \\ \vdots \\ y(t_n, T_m) \end{bmatrix} = \begin{bmatrix} \frac{F'_1(t_n, T_1)}{T_1 - t_n} & \frac{F'_2(t_n, T_1)}{T_1 - t_n} \\ \frac{F'_1(t_n, T_2)}{T_2 - t_n} & \frac{F'_2(t_n, T_2)}{T_2 - t_n} \\ \vdots & \vdots \\ \frac{F'_1(t_n, T_m)}{T_m - t_n} & \frac{F'_2(t_n, T_m)}{T_m - t_n} \end{bmatrix} \begin{bmatrix} r_1(t_n) \\ r_2(t_n) \end{bmatrix} + \begin{bmatrix} -\frac{\sum_{i=1}^2 E'_i(t_n, T_1)}{T_1 - t_n} \\ -\frac{\sum_{i=1}^2 E'_i(t_n, T_2)}{T_2 - t_n} \\ \vdots \\ -\frac{\sum_{i=1}^2 E'_i(t_n, T_m)}{T_m - t_n} \end{bmatrix} + \begin{bmatrix} z_1(t_n) \\ z_2(t_n) \\ \vdots \\ z_m(t_n) \end{bmatrix}, \quad (3.44)$$

where $z_i(t_n) \sim \mathcal{N}(0, H')$ are noise variables which reflects deviation of bond prices from the model price and $H' = \text{diag}(h_1'^2, h_2'^2, \dots, h_m'^2)$, where $h_i'^2$ are positive constants. Equations (3.36) and (3.37) provide the transition equation as in (3.3) and equations (3.44) supply measurement equation as in (3.4). Hence they form a nonlinear state space system with bond yields as observable variables, so that approximate Kalman filtering can be applied for model calibration and forecasting. These forecasts will then be used for predicting the tail losses of both simulated and real bond portfolios in sections 6.2 and 6.3, respectively.

3.6 Summary

The first part of this chapter provides some preliminary definitions for interest rate models, followed by an introduction of Kalman filtering in the context of bond yields modeling. The single factor and two-factor representations for both Vasicek model and CIR model are then included along with some key features and mathematical relationships. As yield forecasting with one factor models tends to be far worse than that with two factor models (see, e.g. Date and Wang (2009)), only the two-factor versions of the presented models will be employed later in chapter 6 for the numerical experiments with simulated data and real data.

Chapter 4

Backtesting: Methodology

4.1 Introduction

Backtesting is a statistical tool to verify whether a model is adequate for its purpose. In the case of VaR models, it consists in checking that actual losses are in line with projected ones. It is crucial to check if predicted values of the measure of risk are reliable. If that is not the case one should reassess assumptions, include a different - and valid - set of parameters or provide an improved modeling methodology. In general, any VaR backtesting procedure follows the following steps:

1. The model is calibrated from one dataset, e.g. over time horizon ranging from t_1 to t_N . For example, this may involve finding the covariance matrix for variance/covariance method from historical data.
2. VaR is computed over another dataset, using the model parameters found in step 1. Typically, the second dataset will span the period $t_{N+1}, t_{N+2}, \dots, t_{N+m}$, where the entire period t_1 to t_{N+m} is in the past, i.e. data is already available. Since the actual losses during each period $t_{N+i} - t_{N+i-1}$ are known for $i = 1, 2, \dots, m$, one can compare the actual losses and the predicted losses and test statistically whether the model is acceptable.

We outline commonly used tests for VaR backtesting next.

4.2 The unconditional coverage test

Several authors recommend backtesting VaR models including Jorion (2007), Kupiec (1995) and Christoffersen (2003). The most common method to test a VaR model has been suggested in Kupiec (1995), where the author developed a 95% confidence region for the unconditional coverage test. The unconditional coverage test is the standard tool for backtesting models and is also recommended by Basel II. Hence this has been employed throughout the rest of the thesis. According to this procedure, a model is correctly calibrated when the number of exceptions (i.e. the portfolio losses exceeding VaR) is in line with the confidence level. If backtesting reveals too many exceptions then the risk is underestimated by the current model, hence one could reserve an insufficient required capital and suffer critical losses under extreme market movements. On the other hand too few exceptions signals an overestimated risk and that would lead to an inefficient allocation of capital, this situation is also not ideal for institutions that look for maximizing profits. Let's define \mathcal{I}_n as:

$$\mathcal{I}_n = \begin{cases} 0 & \text{if } \mathcal{L}_n \leq VaR_{\alpha,n|n-1} \\ 1 & \text{if } \mathcal{L}_n > VaR_{\alpha,n|n-1} \end{cases}$$

where $\mathcal{L}_n = -\Delta\Pi_n$ and $VaR_{\alpha,n|n-1}$ represent respectively the loss at time t_n and the α confidence level Value-at-Risk computed at time t_n given the information at time t_{n-1} . The number of exceptions is given by $\mathcal{X} = \sum_{n=1}^N \mathcal{I}_n$, where N is the total number of observations. Since each weekly outcome could lead to an exception or not, the random variable \mathcal{X} follows a binomial distribution:

$$\Pr[\mathcal{X} = x] = \binom{N}{x} p^x (1-p)^{N-x},$$

where $p = 1 - \alpha$, and α is the level for the selected VaR. Let us consider the number of exceptions in the sample, \tilde{x} , and define the failure rate as \tilde{x}/N . The null and the alternative hypothesis are in *Kupiec's test* (Kupiec (1995)) are as follows:

$$\begin{cases} H_0 : \alpha = 1 - \frac{\tilde{x}}{N} \\ H_1 : \alpha \neq 1 - \frac{\tilde{x}}{N} \end{cases}$$

so we test whether the observed failure rate differs significantly from the given confidence level p . The test statistic used is:

$$LR_{uc} = -2 \ln \left(\frac{(1-p)^{N-\tilde{x}} p^{\tilde{x}}}{\left[1 - \frac{\tilde{x}}{N}\right]^{N-\tilde{x}} \left(\frac{\tilde{x}}{N}\right)^{\tilde{x}}} \right) \sim \chi_1^2. \quad (4.1)$$

Using a 95% confidence interval this likelihood ratio test rejects the null hypothesis if $LR_{uc} > 3.841$. Table 4.1 displays 95% confidence regions of non rejection for the Kupiec's test:

α	N=250	N=500	N=1000
95%	$7 \leq \tilde{x} \leq 19$	$17 \leq \tilde{x} \leq 35$	$38 \leq \tilde{x} \leq 64$
99%	$1 \leq \tilde{x} \leq 6$	$2 \leq \tilde{x} \leq 9$	$5 \leq \tilde{x} \leq 16$

Table 4.1: Non rejection regions for Kupiec's test.

4.3 The independence test

The unconditional coverage test, on its own, focusses on the number of exceptions, but it does not consider whether they are clustered. Since large losses bunched in a small amount of time are more likely to cause disastrous events than single exceptions showing up occasionally (see Campbell (2005) for further details), it is crucial that the VaR model satisfies the independence property. The independence test, developed in Christoffersen (2003), is capable to reject a VaR with clustered exceptions. Let us define the indicator variable:

$$\mathcal{J}_n = \begin{cases} 1 & \text{if an exception occurs} \\ 0 & \text{otherwise} \end{cases}$$

and then define the transition probabilities $\pi_{ij} = \Pr[\mathcal{J}_n = i \text{ and } \mathcal{J}_{n+1} = j]$ so, as an example, π_{01} provides the probability of having an exception tomorrow given that today there were no exception. The first-order Markov sequence with transition probability matrix is given by:

$$\Pi = \begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix} = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix}.$$

If the exceptions sequence is independent over time then the probability of an exception tomorrow does not depend on today's outcome, i.e. $\pi_{01} = \pi_{11} = \pi$. In this case the null

and the alternative hypothesis are:

$$\begin{cases} H_0 : \pi_{01} = \pi_{11} \\ H_1 : \pi_{01} \neq \pi_{11} \end{cases}$$

To test it we use the following likelihood ratio test:

$$LR_{ind} = -2 \ln \left(\frac{(1 - \hat{\pi})^{N_{00} + N_{10}} \hat{\pi}^{N_{01} + N_{11}}}{(1 - \hat{\pi}_{01})^{N_{00}} \hat{\pi}_{01}^{N_{01}} (1 - \hat{\pi}_{11})^{N_{10}} \hat{\pi}_{11}^{N_{11}}} \right) \sim \chi_1^2, \quad (4.2)$$

where $\hat{\pi} = \frac{N_{01} + N_{11}}{N_{00} + N_{01} + N_{10} + N_{11}}$, $\hat{\pi}_{01} = \frac{N_{01}}{N_{00} + N_{01}}$ and $\hat{\pi}_{11} = \frac{N_{11}}{N_{10} + N_{11}}$. N_{ij} represents the number of days when state j follows state i , and i, j can only assume values 0 and 1.

4.4 The conditional coverage test

Since we are interested in understanding whether simultaneously the number of exceptions is correct and VaR exceptions are independent, we can test jointly this two features using the conditional coverage test:

$$LR_{cc} = LR_{uc} + LR_{ind} \sim \chi_2^2. \quad (4.3)$$

Using a 95% confidence interval this likelihood ratio test rejects the null hypothesis if $LR_{cc} > 5.991$. Hence, the 95% level critical values for LR_{uc} , LR_{ind} and LR_{cc} are 3.841, 3.841 and 5.991 respectively. Computation of statistics LR_{uc} and LR_{ind} as respectively specified in (4.1) and (4.2) provides the tool to accept or reject the model specification.

4.5 Summary

This chapter, after suggesting why backtesting is important and how a generic VaR backtesting procedure is performed, outlines the commonly used tests for backtesting VaR. The unconditional coverage test, the independence test and the conditional coverage test are explained and the critical values for the corresponding statistics are reported.

We will utilize these tests to verify the validity of VaR estimates in chapters 5 and 6.

Chapter 5

Nonlinear stock portfolios with fat-tailed risk factors: Numerical experiments

5.1 Introduction

In this chapter, we employ Delta-Gamma-Q method proposed in chapter 2 in simulation experiments as well in measuring the risk of a real financial portfolio. The efficacy of this method is tested against that of competing methods depending on normality. The chapter is structured as follows: section 5.2 provides the simulation experiments, while section 5.3 presents the test for a real financial portfolio.

5.2 Numerical experiments with simulated data

We first evaluate the Delta-Gamma-Q model for assets with non-normal returns considering a simple hypothetical portfolio π , in order to gain some insight in the performance of the proposed heuristic as compared to the performance of standard methods such as the Delta-Gamma Monte Carlo (also called partial Monte Carlo). The hypothetical portfolio is made up of one share each of m correlated fat-tailed stocks $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_m$ and m European call options $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_m$ having $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_m$ as underlying assets, respectively. Therefore

the portfolio value is $\Pi = \sum_{i=1}^m (\mathcal{S}_i + \mathcal{C}_i)$. Each European call option has payoff (or value of the call option at expiry) given by:

$$\max(\mathcal{S}_i(T) - K_i, 0), \quad i = 1, 2, \dots, m. \quad (5.1)$$

$\mathcal{S}_i(T)$ is the price of the i -th underlying stock at time T (maturity), and K_i is the i -th option strike price. Stocks \mathcal{S}_i have been simulated using:

$$\mathcal{S}_i(t) = \mathcal{S}_i(0)e^{(\mu_i - \sigma_i^2/2)t + \sqrt{t}v_i}, \quad i = 1, 2, \dots, m,$$

where $v = [v_1, v_2, \dots, v_m]^T$ is obtained by first generating $w \sim [w_1, w_2, \dots, w_m]^T$, then given the correlation matrix H , we set $v = C^T w$, where C is the Cholesky factor decomposition of H . w is chosen to have a fat-tailed distribution, the exact choice of which is discussed later in this section. Computation of each option value has been obtained using Black-Scholes formulae:

$$\mathcal{C}(\mathcal{S}, K, \sigma, r, T - t) = \mathcal{S}\Phi(d_1) - K\Phi(d_2)e^{-r(T-t)}, \quad (5.2)$$

where

$$d_1 = \frac{\ln(\mathcal{S}/K) + (r + \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}}, \quad d_2 = \frac{\ln(\mathcal{S}/K) + (r - \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}},$$

where T represents the maturity, $T - t$ is the time to maturity and r is the interest rate and

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}y^2} dy.$$

Note that these formulae will not be valid unless w above is normally distributed, i.e. unless the underlying continuous time process is not geometric Brownian motion. However, option prices and sensitivities are often computed under assumption that Black-Scholes formula holds true, irrespective of evidence to the contrary, see e.g. Shao et al. (2006). Sensitivities of interest for a European call option \mathcal{C} such as Delta, Gamma and Theta are given by:

$$\begin{aligned} \delta_{\mathcal{C}} &= \frac{\partial \mathcal{C}}{\partial \mathcal{S}} = \Phi(d_1), \quad \Gamma_{\mathcal{C}} = \frac{\partial^2 \mathcal{C}}{\partial \mathcal{S}^2} = \frac{\phi(d_1)}{\mathcal{S}\sigma\sqrt{T-t}}, \\ \Theta_{\mathcal{C}} &= \frac{\partial \mathcal{C}}{\partial t} = -\left[\frac{\mathcal{S}\sigma\phi(d_1)}{2\sqrt{T-t}} + rK\Phi(d_2)e^{-r(T-t)} \right], \end{aligned} \quad (5.3)$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$. In keeping with the standard market practice, we will use formulae (5.1)-(5.3) for prices and sensitivities in our computation. However, we assume that the *real* stock price dynamics are driven by fat tailed distributions rather than normal ones. The actual distributions of w_i and the parameters used are described later in this section.

Numerical experiments involved comparisons of five different methods:

- *Empirical method* is a Monte Carlo assessment of the chosen assets. Possible h day ahead portfolio values are obtained simulating M trajectories for the m stocks and adding their corresponding European call values, computed through (5.2). The simulation is performed with distribution of risk factors specified later in this section. The loss and profit distribution is achieved subtracting the stocks' initial values and the fair price of the call options from the simulated portfolio values;
- *Delta-Gamma-dP method*, also referred to as partial Monte Carlo earlier, is a computation according to formula (2.20). This involves simulating M trials for each \mathcal{S}_i , where $\Delta\mathcal{S}_i$ are obtained deducting $\mathcal{S}_i(0)$ from the corresponding stock price \mathcal{S}_i while δ and Γ are the ones provided in subsection 2.5.1;
- *Delta-Gamma-Q method* involves nonparametric estimation of pdf and cdf of m risk factors. Nonparametric estimation is able to capture the features of risk factors such as skewness and fat-tailedness. The probability conserving transformation (2.57) on each $\Delta\mathcal{S}_i$ provides cdf and pdf values to apply formula (2.58). Coefficients $[\delta_{\mathcal{Y}}]_i$ in (2.60) are computed approximately as $[\delta_{\mathcal{Y}}]_i \approx \mathbb{E}[d\mathcal{S}_i/d\mathcal{Y}_i][\delta]_i$, while elements $[\Gamma_{\mathcal{Y}}]_{ij}$ are computed approximately as

$$[\Gamma_{\mathcal{Y}}]_{ij} \approx \mathbb{E}[d\mathcal{S}_i/d\mathcal{Y}_i]\mathbb{E}[d\mathcal{S}_j/d\mathcal{Y}_j][\Gamma]_{ij},$$

where coefficients $d\mathcal{S}_k/d\mathcal{Y}_k$ are calculated using formula (2.56). The approximation $\mathbb{E}[d\mathcal{S}_i/d\mathcal{Y}_i]$ is introduced to decrease the computational burden of the algorithm and, simultaneously, to achieve VaR values close to those obtained through full Monte Carlo simulation. Being $\Delta\mathcal{Y}_i \sim \mathcal{N}(0, 1)$ we are in a framework similar to the one described by (2.21), it is therefore possible to find coefficients $a_{\mathcal{Y}}$, $b_{\mathcal{Y}}$ and $\lambda_{\mathcal{Y}}$'s. Using these coefficients one can invert the related characteristic function via FFT;

- *Delta-Gamma-dP Normal method* applies formula (2.20), i.e. assuming \mathcal{S}_i and \mathcal{C}_i to be

jointly normally distributed and ignoring the functional dependence between \mathcal{S}_i and \mathcal{C}_i ;

- *Delta-Gamma-dP inversion method* applies formula (2.24), i.e. uses a quadratic approximation assuming normal risk factors as in Delta-Gamma-dP above, but computes VaR using an inversion integral.

The last two models use normal factors with mean and variance that match sampling the mean and the variance of changes in the respective risk factor. This methodology allows a fair comparison between the first batch of three models considered and the last two. Making this assumption can be interpreted as follows: *provided that a portfolio evolves as described by the empirical model, what would be the VaR and CVaR that one would compute by employing the Delta-Gamma-Q model or a Delta-Gamma-Normal model?*

Probability distributions of risk factors for empirical and Delta-Gamma-dP have to be assessed. This can be done using either parametric (making use of a preselected model fitted on the already available dataset) or nonparametric estimation. Once the probability distribution is estimated, Monte Carlo simulation can be used to find VaR and CVaR for empirical as well as Delta-Gamma-dP methods. For Delta-Gamma-dP Normal and Delta-Gamma-dP inversion, VaR and CVaR are obtained by using formulae (1.1) and (1.3) for normal distribution and given confidence level.

Once VaR is computed using different methods, we wish to compare their accuracy. Confidence intervals and/or standard error estimates are usually used to perform this task. The author in Pritsker (1997) stated “*This is typically not done for Delta and Delta-Gamma based estimates of VaR since there is no natural method for computing a standard error or constructing a confidence interval.*” about this issue. One can use the empirical distribution from a Monte Carlo simulation to obtain confidence intervals for VaR estimates (95% confidence are typically calculated, but this can be easily generalized). The nonparametric confidence intervals, based on finite sample theory, are easy to compute and are valid for any continuous distribution of the random variable \mathcal{L} . Varying the sample size, M , of the Monte Carlo simulation changes the width of the confidence interval, according to the accuracy of VaR needed. Table 5.1 displays the index of ordered statistics to build 95% confidence intervals for 95% and 99% VaR for a different number of draws. Confidence intervals for

CVaR can also be obtained using formula (1.3) and table 5.1, but are not reported here.

Number of Draws	95% VaR		99% VaR	
	Lower Bound	Upper Bound	Lower Bound	Upper Bound
500	15	35	1	10
1,000	37	64	4	17
10,000	457	544	81	120
50,000	2,404	2,597	456	545
100,000	4,865	5,136	938	1,063

Table 5.1: 95% confidence intervals for Monte Carlo 95% and 99% VaR.

Our experiment considered a portfolio made up of fifty stocks and fifty European call options. Stock prices were generated using w_i distributed as Student's t with different degrees of freedom ν_i , ranging from 4 to 10. The values chosen for the simulation include: the time horizon is $h = 1$ day, the number of simulations required is $M = 10^4$, the interest rate $r = 0.05$, the maturities $T_i = 1$ year, for $i = 1, 2, \dots, 50$. The correlation matrices H are randomly generated. Strike prices K_i and initial values $\mathcal{S}_i(0)$ were chosen such that some of the options were in the money, some were at the money and some were out of the money. The parameters used in our simulation are in the range reported in table 5.2. The full table of values is omitted for brevity.

Value	Min	Max
K_i	2	99
μ_i	-4.3	4.1
$\mathcal{S}_i(0)$	12	123
ν_i	4	10

Table 5.2: Parameters range.

The proposed method provides a new covariance matrix which better reflects non-normality in risk factor distributions. Computations of 95% and 99% VaR and CVaR have been obtained keeping constant all the values except from the correlation matrix H which is allowed to vary and reported in tables 5.3–5.4:

Experiment	1	2	3	4
95% VaR				
Empirical	90.4 (87.9,93.3)	72.6 (71.1,74.7)	82.1 (79.4,84.5)	160.1 (151.8,167.7)
Δ - Γ -dP	92.3 (89.4,94.7)	74.1 (71.7,76.3)	82.5 (79.2,85.4)	162.6 (153.2,168.4)
Δ - Γ -Q	91.0	71.9	81.9	164.1
Δ - Γ -dP Normal	84.6 (81.6,87.5)	90.2 (87.4,92.8)	86.1 (83.8,89.9)	102.5 (96.4,107.8)
Δ - Γ -dP Inversion	84.0	91.0	85.4	104.4
99% VaR				
Empirical	131.5 (127.3,135.0)	108.0 (103.8,111.6)	117.8 (110.8,124.3)	246.9 (237.2,254.5)
Δ - Γ -dP	136.0 (131.9,140.9)	109.1 (106.6,112.6)	119.8 (115.2,124.6)	244.9 (238.1,253.2)
Δ - Γ -Q	135.4	109.4	118.2	250.0
Δ - Γ -dP Normal	127.4 (122.2,131.6)	135.6 (132.2,140.8)	126.8 (121.5,132.1)	152.0 (146.5,157.7)
Δ - Γ -dP Inversion	127.1	137.4	125.9	155.7

Table 5.3: Values of 95% and 99% VaR for the four experiments reported, with confidence intervals in brackets.

Experiment	1	2	3	4
95% CVaR				
Empirical	115.9	94.2	103.7	212.8
Δ - Γ -dP	118.4	95.5	105.5	212.7
Δ - Γ -Q	115.2	94.9	105.8	217.7
Δ - Γ -dP Normal	110.0	117.5	110.8	132.4
Δ - Γ -dP Inversion	110.8	119.1	111.2	134.8
99% CVaR				
Empirical	150.9	127.5	132.9	292.5
Δ - Γ -dP	155.6	125.3	141.0	288.5
Δ - Γ -Q	153.4	127.0	131.5	285.7
Δ - Γ -dP Normal	147.2	155.2	143.9	176.9
Δ - Γ -dP Inversion	145.8	157.9	142.5	180.4

Table 5.4: Values of 95% and 99% CVaR over four combination of parameters.

Confidence intervals for full and partial Monte Carlo methods are obtained using values in correspondence of 10,000 draws in table 5.1, while interval estimates are unavailable for $\Delta\text{-}\Gamma\text{-Q}$ and $\Delta\text{-}\Gamma\text{-dP}$ inversion since they are deterministic computations.

Percentage error is defined as $|E - X|/E \cdot 100$, where E is the empirical value for a measure of risk and represents the reference, while X is the alternative value for the same measure of risk. Table 5.5 gathers percentage errors relative to tables 5.3 and 5.4 for $\Delta\text{-}\Gamma\text{-dP}$, $\Delta\text{-}\Gamma\text{-Q}$, $\Delta\text{-}\Gamma\text{-dP}$ Normal and $\Delta\text{-}\Gamma\text{-dP}$ inversion:

Percentage Errors				
Experiment	1	2	3	4
95% VaR				
$\Delta\text{-}\Gamma\text{-dP}$	2.10	2.07	0.49	1.56
$\Delta\text{-}\Gamma\text{-Q}$	0.66	0.96	0.24	2.50
$\Delta\text{-}\Gamma\text{-dP}$ Normal	6.42	24.24	4.87	35.98
$\Delta\text{-}\Gamma\text{-dP}$ Inversion	7.08	25.34	4.02	34.79
99% VaR				
$\Delta\text{-}\Gamma\text{-dP}$	3.42	1.02	1.70	0.81
$\Delta\text{-}\Gamma\text{-Q}$	2.97	1.30	0.34	1.26
$\Delta\text{-}\Gamma\text{-dP}$ Normal	3.12	25.56	7.64	38.44
$\Delta\text{-}\Gamma\text{-dP}$ Inversion	3.35	27.22	6.88	36.94
95% CVaR				
$\Delta\text{-}\Gamma\text{-dP}$	2.16	1.38	1.74	0.05
$\Delta\text{-}\Gamma\text{-Q}$	0.60	0.74	2.03	2.30
$\Delta\text{-}\Gamma\text{-dP}$ Normal	5.09	24.73	6.85	37.78
$\Delta\text{-}\Gamma\text{-dP}$ Inversion	4.40	26.43	7.23	36.65
99% CVaR				
$\Delta\text{-}\Gamma\text{-dP}$	3.11	1.73	6.09	1.37
$\Delta\text{-}\Gamma\text{-Q}$	1.66	0.39	1.05	2.32
$\Delta\text{-}\Gamma\text{-dP}$ Normal	2.45	21.73	8.28	39.52
$\Delta\text{-}\Gamma\text{-dP}$ Inversion	3.38	23.84	7.22	38.32

Table 5.5: Percentage errors for the four experiments carried out.

The results indicate that, with varying covariance matrix H , while Delta-Gamma-dP and Delta-Gamma-Q provide a reasonably good approximation to the empirical value of VaR and CVaR, the two remaining models (based on normality assumption) do not provide a

consistent estimation of VaR and CVaR. In particular, using the empirical estimations as references, we can state that:

- in the first experiment 95% and 99% VaR are underestimated of about 7% and 3%, respectively and 95% and 99% CVaR are underestimated of about 5% and 3%, respectively;
- in the second experiment 95% and 99% VaR are overestimated of about 25% and 26%, respectively and 95% and 99% CVaR are overestimated of about 25% and 23%, respectively;
- in the third experiment 95% and 99% VaR are overestimated of about 4% and 7%, respectively and 95% and 99% VaR are both overestimated of about 7%;
- in the fourth experiment 95% and 99% VaR are underestimated of about 35% and 37%, respectively and 95% and 99% CVaR are underestimated of about 37% and 38%, respectively.

VaR and CVaR computed under the assumption of normal distributed risk factors are, according to expectations, unreliable. Results for varying parameters other than H are qualitatively similar. Portfolio composition changes such as the number of stocks included, the number and/or kind of (puts or calls) options included also lead to similar qualitative conclusions. Finally, the proposed methodology is quite general and works when some of the risk factors are normally distributed and/or present a certain degree of skewness. This was verified in simulation experiments with a few normal and centered skew- t distributions (see e.g. Azzalini and Capitanio (2003) for details about skew- t distributions) with a skewness absolute value greater than one components. Results in all these cases (varying other parameters, changing portfolio composition and changing distributional assumptions) do not add any additional information and are hence omitted for brevity.

We express hereafter some considerations about accuracy and computational time of the simulation for the three methods used for computation of VaR and CVaR: full Monte Carlo, partial Monte Carlo (or Delta-Gamma-dP) and Delta-Gamma-Q method. A fair comparison involves that all three methods require a nonparametric estimation. Let us recall that m is the number of risky factors and M is the number of simulations required. The empirical method or full Monte Carlo method is generally thought to give the most accurate estimates

of VaR for large sample sizes but it tends to be very time consuming, specially when the analytical solutions for some assets are not available. Furthermore, closed-form pricing formulae are not often available and options often need to be priced (and the sensitivities need to be computed) numerically, e.g. by solving a partial differential equation. In these cases, the time required will be several orders of magnitude higher, especially for Monte Carlo method. The Delta-Gamma-dP method or partial Monte Carlo produces estimates less accurate than full Monte Carlo but is less expensive in terms of time. As reported in Mina and Ulmer (1999), the partial Monte Carlo requires $\mathcal{O}(Mm^2)$ operations while the full Monte Carlo requires additional time to assess all of the positions M times. The Delta-Gamma-Q method demonstrates an accuracy which is comparable to the one of partial Monte Carlo and requires $\mathcal{O}(m^3)$ operations. Therefore, comparing computational costs it emerges that partial Monte Carlo is slower than the Delta-Gamma-Q method unless m is very large depending on the VaR confidence level α (e.g., about 1,000 for 95% VaR and about 5,000 for 99% VaR, as chosen in Mina and Ulmer (1999)). Speed performances of a typical portfolio made up of m stocks and m European call options with $m = 1, 2, 3, 4, 5, 10, 20, 30, 40, 50$ for the empirical, Delta-Gamma-dP and Delta-Gamma-Q has been measured using an Intel dual core i3 clocked at 2.66 GHz, with 3GB RAM and using MATLAB 7.9. Any computation includes estimation of 95% and 99% VaR and CVaR. The computational times (expressed in seconds) of the empirical, Delta-Gamma-dP and Delta-Gamma-Q methods are reported in table 5.6:

m	Empirical	Δ - Γ -dP	Δ - Γ -Q
1	0.090184	0.084911	0.025212
2	0.137924	0.103209	0.040598
3	0.158297	0.113802	0.043022
4	0.207286	0.137386	0.048445
5	0.233043	0.164483	0.051230
10	0.621080	0.206162	0.067723
20	0.832268	0.301063	0.163616
30	1.018157	0.373982	0.231480
40	1.288235	0.479643	0.288678
50	1.466571	0.563397	0.332650

Table 5.6: Net computational times for the computation of VaR and CVaR for the portfolio made up of m stocks and m European call options, with $m = 1, 2, 3, 4, 5, 10, 20, 30, 40, 50$.

The computational times in table 5.6 refer to computations with 10,000 simulations and do not include the time for nonparametric estimation of the risk factors. We set the full Monte Carlo method as the reference in terms of values and computational times. It can be noticed that both partial Monte Carlo and Delta-Gamma-Q are quantitatively comparable to the reference, as shown in tables 5.3-5.4. Furthermore, our method requires a lower amount of time than the other two methods considered for all the included values of m . As envisaged in the study of computation costs the Delta-Gamma-Q method for the computation of results is the quickest, since it exploits the computational speed of the FFT algorithm. Larger scale computational tests tend to be portfolio-specific. However, the qualitative conclusion that Delta-Gamma-Q method is comparable in its speed to Delta-Gamma-Normal method, but outperforms it in accuracy for non-normal risk factors, holds true.

5.3 Numerical experiments with real data

Having seen the performance of our method with a simulated portfolio with up to 50 assets, we now move to demonstrating it with a real financial portfolio and comparing it with Delta-Gamma-Normal method in computing VaR and CVaR. This section is divided into five subsections. Subsection 5.3.1 introduces the data employed and presents some of their descriptive statistics. The portfolio analyzed is constituted of four European options on FTSE100 index (which represents our single risk factor). A representative real portfolio with a relatively small number of assets is chosen for demonstration since it is easier to visualize data and report results with a small number of assets; it is also easy to reproduce our experiments based on the information provided here, if desired. Subsection 5.3.2 illustrates the choice and the features of the portfolio tested. Subsections 5.3.3 and 5.3.4 respectively report the results of backtesting using the Delta-Gamma-dP Normal approach and the Delta-Gamma-Q approach.

5.3.1 Data

For computing the VaR estimates and for backtesting, we use two sets of data:

1. A set of 501 daily closing prices for each of the four included European options on the underlying index FTSE100, according to the portfolio composition specified in tables 5.8–5.9;

2. A set of daily closing FTSE100 values which is used for estimating the parameters related to non-normality as described in section 2.6.1.

The portfolio analyzed includes observations from 10/07/2009 to 10/06/2011. The whole set of daily portfolio variations has been split in two 250-units subsets: an in-sample subset that covers observations from 13/07/2009 to 25/06/2010, and an out-of-sample subset that covers observations from 28/06/2010 to 10/06/2011. Summary statistics on FTSE100 daily returns R_t for the considered period are reported in table 5.7:

Statistic	Value
Mean	0.0722%
Standard Deviation	1.0349%
Minimum	-3.1815%
Maximum	5.1610%
Skewness	0.0072
Kurtosis	4.3989

Table 5.7: Summary statistics of the FTSE100 returns for the period 10/07/2009 – 10/06/2011.

The empirical distribution of daily returns is leptokurtic (i.e., its kurtosis exceeds the value 3), that indicates fat-tailedness, and slightly skewed. The Lilliefors test, used to test the null hypothesis that data come from a normal distribution when the sample is small (see e.g. Lilliefors (1967) for details), also rejects normality at both 5% and 1% significance level for the considered series. Figure 5.1 represents the returns histogram for FTSE100 for the considered period.

5.3.2 Test portfolio

The portfolio tested includes four European options on the same underlying: the index FTSE100. Its composition is changed periodically to mirror the variations occurring in a typical, actively traded options portfolio. To be specific, the portfolio is made up of two pairs of options, each pair consisting of a call option and one put option having common strike price K and maturity T . The daily portfolio π_t is made up of $\mathcal{C}_t^{(1)}$, $\mathcal{P}_t^{(1)}$, $\mathcal{C}_t^{(2)}$ and $\mathcal{P}_t^{(2)}$, with $t = 1, 2, \dots, 501$. Therefore the daily portfolio values Π_t are computed as $\Pi_t = \mathcal{C}_t^{(1)} + \mathcal{P}_t^{(1)} + \mathcal{C}_t^{(2)} + \mathcal{P}_t^{(2)}$ and the daily changes in portfolio values are computed as

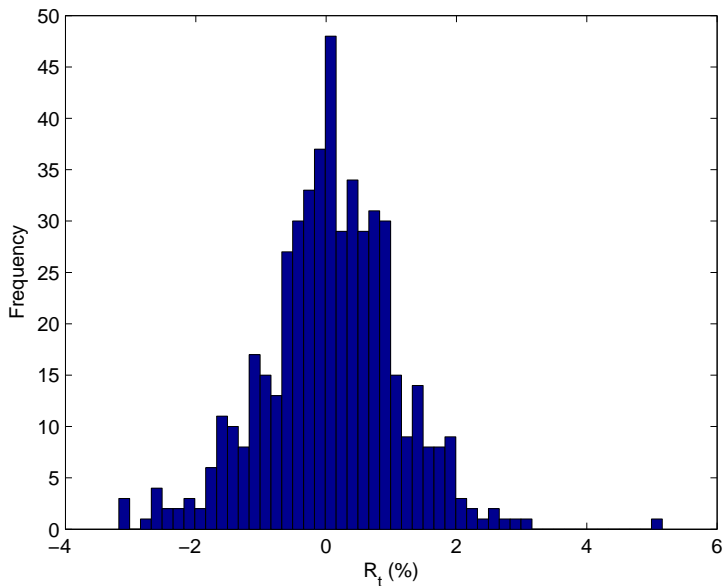


Figure 5.1: FTSE100 returns histogram for the period 10/07/2009 – 10/06/2011.

$\Delta\Pi_{t-1} = \Pi_t - \Pi_{t-1}$ for $t = 2, 3, \dots, 501$. We decided to change the portfolio composition over time, including options having different characteristics (strike prices and maturities), to highlight that the method does not depend on specific features of the portfolio. As mentioned earlier, nonlinear portfolios are rarely static over a long period of time. To generate a large enough data sample, with the same underlying risk factor for backtesting, it makes sense to use a portfolio of options which evolves over time. The whole length of time under test is split in six different length intervals. Intervals with different lengths provide again a more general framework. At every change of interval one of the two call-put pairs is dropped and is alternately replaced by another call-put pair of options, with both the call and the put having the same strike price and maturity. Table 5.8 displays the portfolio composition chosen, whereas figure 5.2 provides a graphical representation of the portfolio composition as a function of time.

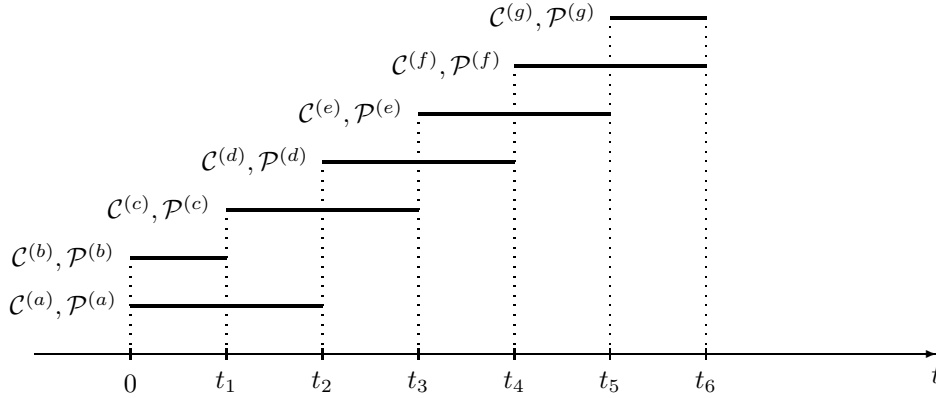


Figure 5.2: Graphical representation of portfolio composition in time. Options are included in the portfolio in correspondence of the relative thick segment.

Note that most options have a life which is shorter than what one needs for a reasonably large backtesting data sample, which makes the proposed changes in portfolio over time (keeping the risk factor the same) a sensible alternative for backtesting VaR methodologies.

Interval	from	to	Length	$\mathcal{C}_t^{(1)}$	$\mathcal{P}_t^{(1)}$	$\mathcal{C}_t^{(2)}$	$\mathcal{P}_t^{(2)}$
1	10/07/2009	10/11/2009	88	$\mathcal{C}^{(a)}$	$\mathcal{P}^{(a)}$	$\mathcal{C}^{(b)}$	$\mathcal{P}^{(b)}$
2	11/11/2009	26/02/2010	78	$\mathcal{C}^{(a)}$	$\mathcal{P}^{(a)}$	$\mathcal{C}^{(c)}$	$\mathcal{P}^{(c)}$
3	01/03/2010	30/07/2010	110	$\mathcal{C}^{(d)}$	$\mathcal{P}^{(d)}$	$\mathcal{C}^{(c)}$	$\mathcal{P}^{(c)}$
4	02/08/2010	30/11/2010	87	$\mathcal{C}^{(d)}$	$\mathcal{P}^{(d)}$	$\mathcal{C}^{(e)}$	$\mathcal{P}^{(e)}$
5	01/12/2010	15/03/2011	75	$\mathcal{C}^{(f)}$	$\mathcal{P}^{(f)}$	$\mathcal{C}^{(e)}$	$\mathcal{P}^{(e)}$
6	16/03/2011	10/06/2011	63	$\mathcal{C}^{(f)}$	$\mathcal{P}^{(f)}$	$\mathcal{C}^{(g)}$	$\mathcal{P}^{(g)}$

Table 5.8: Portfolio composition in the different intervals.

Strike prices and maturities of each pair of options $\mathcal{C}^{(l)}, \mathcal{P}^{(l)}$, $l \in \{a, b, c, d, e, f, g\}$ are reported in table 5.9:

Asset Label	a	b	c	d	e	f	g
Strike Price	3,800	2,000	3,000	4,800	3,700	5,400	2,600
Maturity	18/06/10	18/06/10	17/09/10	17/12/10	18/03/11	17/06/11	17/06/11

Table 5.9: Considered strike prices and maturities.

5.3.3 Applying the Delta-Gamma-dP Normal method: Results

The first experiment consists in computing 95% and 99% VaR of the considered portfolio using Delta-Gamma-dP Normal model and assessing its reliability through unconditional and conditional tests. The descriptive statistics of $\Delta\mathcal{S}$ are gathered in table 5.10.

Statistic	In-sample	Out-of-sample
Mean	3.6772	2.8773
Standard Deviation	57.6257	53.3055
Minimum	-170.8800	-157.4600
Maximum	264.4000	141.4700
Skewness	-0.1112	-0.0386
Kurtosis	4.6705	3.3710

Table 5.10: Descriptive statistics of the $\Delta\mathcal{S}$.

A glance at the statistics reported in table 5.10 suggests that $\Delta\mathcal{S}$ has a different type of distribution in the two different subsets: the in-sample subset appears not normally distributed and its Lilliefors test rejects at levels 5% and 1% that $\Delta\mathcal{S}$ comes from a distribution in the normal family, while for the out-of-sample subset, the Lilliefors test does not reject at levels 5% and 1% that $\Delta\mathcal{S}$ comes from a distribution in the normal family.

Sensitivities $\Theta_t^{(i)}$, $\delta_t^{(i)}$ and $\Gamma_t^{(i)}$, with $i \in \{\mathcal{C}^{(1)}, \mathcal{P}^{(1)}, \mathcal{C}^{(2)}, \mathcal{P}^{(2)}\}$, according to the portfolio composition reported in table 5.8 are computed for both the in-sample data and out-of-sample data using steps (1) and (2) from the procedure described in subsection 5.3.4. Θ_t , δ_t and Γ_t are derived summing up the correspondent sensitivities of the options included at time t . The Delta-Gamma-dP Normal approximation becomes therefore:

$$\Delta\Pi_t = \Theta_t\Delta t + \delta_t\Delta\mathcal{S} + \frac{1}{2}\Gamma_t\Delta\mathcal{S}^2. \quad (5.4)$$

Assuming that $\Delta\mathcal{S} \sim \mathcal{N}(0, \sigma^2)$, it is possible to use the transformation $\Delta\mathcal{S} = \sigma\mathcal{Z}$, where $\mathcal{Z} \sim \mathcal{N}(0, 1)$. The equation (5.4) can be rewritten as:

$$\Delta\Pi_t = \Theta_t\Delta t + \delta_t\sigma\mathcal{Z} + \frac{1}{2}\Gamma_t\sigma^2\mathcal{Z}^2. \quad (5.5)$$

For each of the subsets one can utilize the standard deviation of $\Delta\mathcal{S}$ and compute the Delta-Gamma-dP Normal VaR. The time horizon h is set to 1 day. Unconditional and conditional tests are reported in table 5.11:

Subset	α	\hat{x}	N_{00}	N_{01}	N_{10}	N_{11}	LR_{uc}	LR_{ind}	LR_{cc}	A/R
In-sample	95%	14	222	13	13	2	0.1827	1.1758	1.3385	A
	99%	10	230	10	10	0	12.9555	0.8336	13.7891	R
Out-of-sample	95%	15	222	13	14	1	0.4961	0.0326	0.5286	A
	99%	4	242	4	4	0	0.7691	0.1301	0.8992	A

Table 5.11: Summary of test results for the Delta-Gamma-dP Normal model.

The values in bold represents either values outside the non-rejection confidence intervals or values above the corresponding critical value. In particular the last column of each table indicates whether the VaR, at the specified level of confidence, estimated for the in-sample subset and the out-of-sample subset is accepted (A) or rejected (R). Statistics for the in-sample subset show that the 95% VaR estimate is acceptable being all the values LR_{uc} , LR_{ind} and LR_{cc} below the respective critical values. Instead, the large number of failures and relative statistics suggest that the estimation using the Delta-Gamma-dP Normal method for the 99% VaR is not appropriate. This result confirms the insight that just using the standard deviation of a risk factor, neglecting its fat-tailed behavior could lead to an acceptable VaR for relative low levels of α (up to 95%), but could fail to provide suitable VaR for higher values of α , underestimating its actual value. This might explain a fraction of what has happened during the recent financial crisis: managers, having relied on Gaussian-based models and ignoring the non-normality of risk factors, obtained VaR values that were wrong (even catastrophically, in some cases). The statistics LR_{uc} , LR_{ind} and LR_{cc} for the out-of-sample subset are all below the respective critical values for both the α levels considered. This is not surprising since the out-of-sample dataset is close to being normal; also see table 5.10.

5.3.4 Applying the Delta-Gamma-Q method: Results

The second experiment involves two phases:

- Calibrate the parameters of the Delta-Gamma-Q model for a non-normal factor using in-sample data;

- Assessing the model validity through conditional and unconditional tests using out-of-sample data.

Daily 1-day horizon $VaR_{t|t-1}$ estimates are obtained using the appropriate option prices occurred on day $t-1$. The first 250-unit subset of losses, covering the period from 13/07/2009 to 25/06/2010, is used for parameter calibration while the second 250-unit subset of losses, covering the period from 28/06/2010 to 10/06/2011 is used for model validation using the parameter $D = \mathbb{E}[d\mathcal{S}/d\mathcal{Y}]$ estimated in the in-sample subset. The following part describes how the estimates $VaR_{t|t-1}$ are computed. The payoff for an European call option is given by (5.1), while the payoff for an European put option is given by:

$$\max(K - \mathcal{S}(T), 0). \quad (5.6)$$

The calibration step itself requires the following steps for each time interval considered:

1. finding the implied risk-free rates which are needed in finding δ and Γ ;
2. determining sensitivities δ, Γ and Θ ;
3. working out coefficients that capture the index FTSE100 non-normality;
4. calculating coefficients $\delta_{\mathcal{Y}}, \Gamma_{\mathcal{Y}}$ and $\Theta_{\mathcal{Y}}$;
5. deriving coefficients $a_{\mathcal{Y}}, b_{\mathcal{Y}}$ and $\lambda_{\mathcal{Y}}$.

Given that daily implied volatilities of the call and the put option in each pair, $\sigma_{\mathcal{C}}$ and $\sigma_{\mathcal{P}}$, are also available data, one can estimate the risk-free rate for each pair of options by minimizing the quantity:

$$\left[(\mathcal{C}_i - \mathcal{C}(\mathcal{S}, K, \sigma_{\mathcal{C}}, r, T - t))^2 + (\mathcal{P}_i - \mathcal{P}(\mathcal{S}, K, \sigma_{\mathcal{P}}, r, T - t))^2 \right], \quad i = 1, 2,$$

where r is the only unknown, since option values \mathcal{C}_i and \mathcal{P}_i , index value \mathcal{S} , strike price K and time to maturity $T - t$ are all known. The function $\mathcal{C}(\cdot)$ refers to the Black-Scholes formulae to compute European call option prices and is provided by (5.2), while the function $\mathcal{P}(\cdot)$ refers to the Black-Scholes formulae to compute European put option and is given by:

$$\mathcal{P}(\mathcal{S}, K, \sigma, r, T - t) = Ke^{-r(T-t)}\Phi(-d_2) - \mathcal{S}\Phi(-d_1), \quad (5.7)$$

where d_1 and d_2 are the ones expressed in (5.2). Functional dependence of variables on t is suppressed for notational brevity. Therefore two risk-free rates r_1 and r_2 are obtained for the two call-put option pairs. This minimization was carried out in MATLAB 7.9 using an inbuilt routine `fminbnd`, which uses golden section search and parabolic interpolation. Sensitivities Delta, Gamma and Theta for European call options can be calculated using (5.3), while corresponding sensitivities for European put options are given by:

$$\begin{aligned}\delta_{\mathcal{P}} &= \frac{\partial \mathcal{P}}{\partial \mathcal{S}} = \Phi(d_1) - 1, \quad \Gamma_{\mathcal{P}} = \frac{\partial^2 \mathcal{P}}{\partial \mathcal{S}^2} = \frac{\phi(d_1)}{\mathcal{S}\sigma\sqrt{T-t}}, \\ \Theta_{\mathcal{P}} &= \frac{\partial \mathcal{P}}{\partial t} = -\frac{\mathcal{S}\sigma\phi(d_1)}{2\sqrt{T-t}} + rK\Phi(-d_2)e^{-r(T-t)}.\end{aligned}\tag{5.8}$$

Having two pairs of options in our test portfolio as described in subsection 5.3.2 we obtain two sets of sensitivities applying the appropriate risk-free rate to the corresponding pair of options.

A crucial role is played by the coefficient $D = \mathbb{E}[d\mathcal{S}/d\mathcal{Y}]$ which is able to encapsulate the possible non-normality of the risk factor \mathcal{S} . Its value for the computation of VaR_t would be computed using a sample average over a “window” of FTSE100 data of length n . Using trial and error in the in-sample subset we found that a suitable width for this risk factor window \bar{n} is 150. We used three 150 wide rolling windows for the in-sample subset: the first from 10/07/2009 to 04/02/2010, the second from 18/09/2009 to 15/04/2010, the third from 27/11/2009 to 25/06/2010. The value D has been computed for each rolling window using formula (2.55) and their average value was used to verify whether the model is valid using the conditional and unconditional coverage tests for the in-sample subset. This simple moving average heuristic with overlapping windows provides a smoothing effect. For assessment of out-of-sample data, D is computed using a single window of width 150 from 27/11/2009 to 25/06/2010 (i.e., the in-sample data immediately prior to the start of out-of-sample data set). The assessment of modeling for out-of-sample data is done using a window of width 150 from 27/11/2009 to 25/06/2010, for the computation of the value D needed for the conditional and unconditional coverage tests.

Coefficients $a_{\mathcal{Y}}$, $b_{\mathcal{Y}}$ and $\lambda_{\mathcal{Y}}$ are calculated using formulae as reported in Section 2.6.2. Results for the in-sample Delta-Gamma-Q and for the out-of-sample Delta-Gamma-Q are shown in table 5.12.

Subset	α	\tilde{x}	N_{00}	N_{01}	N_{10}	N_{11}	LR_{uc}	LR_{ind}	LR_{cc}	A/R
In-sample	95%	14	223	13	13	1	0.1827	0.0620	0.2447	A
	99%	5	240	5	5	0	1.9568	0.2041	2.1609	A
Out-of-sample	95%	14	223	14	13	0	0.1827	1.5400	1.7226	A
	99%	2	246	2	2	0	0.1084	0.0323	0.1407	A

Table 5.12: Summary of test results for the Delta-Gamma-Q modeling of real data.

Results, computed using formulae mentioned earlier in this section, show that all tests for both 95% VaR and 99% VaR are below the respective critical values, therefore cannot be rejected. The proposed model hence seems to provide acceptable VaR estimates in both the sub-samples considered and for both the confidence intervals, in contrast with Delta-Gamma-Normal method outlined in subsection 5.3.3 earlier. The results using 5-day time horizon were found to be consistent with these findings and are omitted for brevity.

5.4 Summary

In this chapter we carry out numerical experiments with both simulated data and real data to test the behavior of the proposed new heuristic. In section 5.2 we compute measures of risk of four simulated portfolios each made up of fifty fat-tailed stocks and fifty European call options using five different methods: the empirical method, the Delta-Gamma-dP method, the Delta-Gamma-Q method, the Delta-Gamma-dP Normal method and the Delta-Gamma-dP inversion method. We compare the performance of the Delta-Gamma-Q method in terms of VaR and CVaR percentage errors with the other methods, using the empirical method as the reference. Furthermore, we compare the computational times of portfolios including m stocks and m European call options for $m = 1, 2, 3, 4, 5, 10, 20, 30, 40, 50$ for computing VaR using the empirical method, Delta-Gamma-dP and Delta-Gamma-Q.

In section 5.3 we measure the risk for a portfolio that includes four European options, all having the same underlying: the FTSE100. We compute 95% and 99% VaR for an in-sample subset and an out-of-sample subset using two models: the Delta-Gamma-dP Normal method and the Delta-Gamma-Q method. The results indicate that relying on the assumption of Gaussianity of ΔS can lead to wrong VaR estimates, while our method, which is able to include effects such as fat-tailedness and skewness, can provide acceptable VaR estimates.

Chapter 6

Nonlinear fixed income portfolios: Numerical experiments

6.1 Introduction

In this chapter, we use the mathematical tools described in chapters 3 and 4 to provide backtesting results for simulated as well as real financial data for fixed income portfolios of sovereign bonds. Results are presented for both normal (with Vasicek short rate model) and non-normal (with CIR model) risk factors, while the portfolio value is nonlinear in the risk factors in both the cases. The chapter is organized as follows: section 6.2 presents experiments using simulated data, while section 6.3 provides results using real financial data.

6.2 Numerical experiments with simulated data

In this subsection we assess and compare Value-at-Risk for bond portfolios computed through Monte Carlo (MC) simulation, historical simulation (HS) method and variance/covariance (VC) method using simulated data. In subsection 6.2.1 the MC simulation is performed using the two-factor Vasicek model and in subsection 6.2.2 a sensitivity analysis for the same type of process is reported. In subsection 6.2.3 the MC simulation is performed using

the two-factor CIR model and in subsection 6.2.4 a sensitivity analysis for the same type of process is reported. This will help to gain some insight about the performances of both MC simulations compared to the remaining considered methods. Subsection 6.2.5 comments on the simulation experiments reported in the previous subsections.

In subsections 6.2.1 a short term interest rate path is generated using a two-factor Vasicek model as specified by formulae (3.17) and (3.18), while in subsection 6.2.3 a short term interest rate is generated using a two-factor CIR model as specified by formulae (3.34) and (3.35). As mentioned earlier in section 3.3, at any time t_n the simulated bond yields \mathbf{y}_n are given by:

$$\mathbf{y}_n = C\mathbf{r}_n + \mathbf{d} + \mathbf{z}_n,$$

where C and \mathbf{d} are explicitly expanded in equation (3.27) for the Vasicek model and in equation (3.44) for the CIR model. The measurement data are given by simulating a set of m bonds having different maturities whose prices are driven by the corresponding generated interest rate. The KF recursive equations (3.5) to (3.10) are allowed to change their values according to the measurement data provided. In particular the state vector and the covariance matrix predictions $\hat{\mathbf{r}}_{n+1|n}$ and $V_{n+1|n}$ computed at any time t_n are useful to obtain the VaR at time t_{n+1} for a specified bond portfolio through full Monte Carlo simulation.

The bond portfolios under study consist of m bonds having each maturity T_j and price $P_{n,j}$ at time t_n for $j = 1, 2, \dots, m$. Each portfolio composition is not changing during the period considered. At time $n + 1$ the value of each bond portfolio is given by:

$$\Pi_{n+1} = \sum_{j=1}^m \omega_j P_{n+1,j}, \quad (6.1)$$

where $P_{n+1,j}$ and Π_{n+1} are, respectively, the price of the j^{th} bond and the portfolio net worth at time t_{n+1} and ω_j is the quantity of the j^{th} bond held. We are interested in characterizing the distribution of Π_n .

Each $P_{n+1,j}$ using the Vasicek model is given by:

$$P_{n+1,j} = e^{E_{n+1,j} - F_{1\ n+1,j} r_{1\ n+1} - F_{2\ n+1,j} r_{2\ n+1}}, \quad (6.2)$$

where r_1 and r_2 are the factors, and E , F_1 and F_2 are the known functions depending on $T_j - t_{n+1}$ as specified in (3.24) and (3.25).

Each $P_{n+1,j}$ using the CIR model is given by:

$$P_{n+1,j} = e^{E'_{1n+1,j} + E'_{2n+1,j} - F'_{1n+1,j} r_{1n+1} - F'_{2n+1,j} r_{2n+1}}, \quad (6.3)$$

where r_1 and r_2 are the factors, and E'_1 , E'_2 , F'_1 and F'_2 are the known functions depending on $T_j - t_{n+1}$ as specified in (3.41) and (3.42).

Since for the Vasicek model $\mathbf{r}_{n+1|n} \sim \mathcal{N}(\hat{\mathbf{r}}_{n+1|n}, V_{n+1|n})$, one can perform a full Monte Carlo simulation to obtain an approximation of the distribution of Π_{n+1} . At any time t_n , Π_n is exactly known, based on actual bond prices. We simulate the distribution of Π_{n+1} , find the distribution of predicted loss $\Pi_{n+1} - \Pi_n$ and obtain the Value-at-Risk at a specific confidence level. For the CIR model, one can approximate the transitional distribution of $r_{n+1|n}$ by a normal distribution, with a correct conditional mean and variance as given in equations (3.38)-(3.39). We are going to apply the same methodology to compute VaR using both Vasicek model and CIR model. The Value-at-Risk obtained from the distribution of predicted losses will be compared with the actual losses to perform the backtesting. Once computed 95% and 99% VaR it is also possible to obtain their nonparametric confidence intervals which are based on finite sample theory (see Pritsker (1997) for more details on finite sample theory). The number of Monte Carlo simulations determines the degree of accuracy of VaR, since it affects the width of the confidence interval.

We consider an interest rate path for both Vasicek model and CIR model, compute bond portfolio values and assess the reliability of 95% and 99% VaR estimates for MC method, HS method and VC method using the backtesting procedure reported in chapter 4. Each simulation has been run choosing ‘typical’ values (see Castellanos Pinzon (2008)) and it involves computing a weekly interest rate using the Euler discretization of stochastic differential equations (3.17) and (3.18) for the Vasicek modeling and equations (3.34) and (3.35) for the CIR modeling. The procedure for computing VaR using Monte Carlo simulation consists of the following four steps:

- Considering a suitable number of yields obtained by simulating batches of bonds each

- having different maturities;
- Estimating KF parameters using an in-sample subset implementing equations (3.5)-(3.10) and maximum likelihood as described earlier;
 - Selecting the set of estimated parameters such that they suitably fit the in-sample subset;
 - Computing 95% and 99% VaR using Monte Carlo simulation as previously described.

The choice of the dataset is related to the standard of its desired backtesting. To achieve an adequate level of reliability one requires to consider a sufficiently large number of values. We opted for 250 values. The procedure adopted considers an in-sample subset consisting of 200 yields for each of the four batches of simulated bonds to estimate the vector of parameters and then use the estimated values to calculate 50 one-step ahead yield predictions, that will be compared with the corresponding out-of-sample actual values. The choice of employing 50 one-step ahead yield predictions comes from empirical evidence suggesting that out-of-sample fitting using real data is rather good for a number of time steps in a range of 50–75 data, while outside this range results of fitting are poorer. Repeating five times this procedure, shifting both in-sample and the out-of-sample of 50 values, as explained next, allow us to compute the required 250 non-overlapping values:

In-sample	Out-of-sample
1 – 200	201 – 250
51 – 250	251 – 300
101 – 300	301 – 350
151 – 350	351 – 400
201 – 400	401 – 450

The time-step $\Delta t = \frac{1}{52}$ (i.e. weekly data is used), while the whole interest rate path generated consists of 450 simulations.

The procedure for the HS method and the VC method involves using the in-sample subsets to obtain the quantiles $q_{1-\alpha}$ and $q_{1-\alpha}^N$, respectively, that will be employed to compute 95% and 99% VaR in the out-of-sample subset. Bond prices included in the considered portfolio to compute the relevant statistics $q_{1-\alpha}$ and $q_{1-\alpha}^N$ are computed using the formula:

$$P_{n,j} = e^{-y_{n,j}(T_j - t_n)}, \quad j = 1, 2, \dots, m,$$

where $y_{n,j}$ represents the realization of the j^{th} bond yield simulated at time t_n .

6.2.1 Simulation using the two-factor Vasicek model

The values to simulate the two-factor Vasicek model described by equations (3.17) and (3.18) are reported in the table 6.1:

i	x_{i0}	k_i	θ_i	σ_i	λ_i
1	0.015	0.375	0.044	0.015	-0.18
2	0.025	0.02	0.014	0.01	-0.0001

Table 6.1: Coefficients used for the simulation of the two-factor Vasicek model.

The starting values set for the initialization of the KF algorithm are:

$$\hat{\mathbf{r}}_{0|0} = \begin{bmatrix} 0.02 \\ 0.02 \end{bmatrix} \quad \text{and} \quad V_{1|0} = 3 \cdot 10^{-5} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The observable measurements are supplied by five bonds whose features are reported in table 6.2:

j	$T_j(\text{years})$	h_j^2
1	0.5	0.0009 ²
2	1	0.0011 ²
3	1.5	0.0010 ²
4	2	0.0012 ²
5	5	0.0006 ²

Table 6.2: Features of the bonds providing the measurement values.

T_j and h_j^2 represent the maturity and the variance of the zero mean noise term that perturbs the measured bond yield, respectively. The goodness of fit reached for both the in-sample subset and the out-of-sample subset was assessed considering the relative absolute error (RAE), defined as:

$$\text{RAE} = \frac{|\text{simulated rate} - \text{predicted rate}|}{\text{simulated rate}}.$$

The features of the bonds included in the portfolio under study are reported in table 6.3:

j	$T_j(\text{years})$	ω_j
1	1	15,000
2	2	35,000
3	5	30,000

Table 6.3: Features of the bonds included in the simulated portfolio.

where T_j are the maturities and ω_j is the initial number of the j^{th} bond held. Table 6.4 reports the parameter estimation for the five in-sample subset considered, and their corresponding mean of the relative absolute error for both the in-sample (indicated as $MRAE$) and the out-of-sample (indicated as $MRAE^*$):

Subset	1	2	3	4	5
k_1	0.3127	0.2987	0.2942	0.3025	0.3211
θ_1	0.0551	0.0335	0.0005	0.0229	0.0536
σ_1	0.0161	0.0176	0.0177	0.0166	0.0145
λ_1	-0.2821	-0.5838	-0.9375	-0.6707	-0.0944
k_2	0.0586	0.0452	0.0222	0.0105	0.0733
θ_2	0.0131	0.0005	0.0083	0.0062	0.0076
σ_2	0.0109	0.0100	0.0093	0.0082	0.0110
λ_2	-0.0002	-0.0014	-0.0014	-0.0001	-0.0014
h_1	0.0006	0.0006	0.0007	0.0007	0.0007
h_2	0.0011	0.0011	0.0010	0.0010	0.0011
h_3	0.0009	0.0009	0.0009	0.0008	0.0008
h_4	0.0011	0.0011	0.0011	0.0011	0.0011
h_5	0.0009	0.0009	0.0010	0.0010	0.0010
$MRAE(\%)$	0.89	0.91	1.13	1.36	1.40
$MRAE^*(\%)$	1.28	1.46	1.15	1.09	0.83

Table 6.4: Estimated parameters for the subset considered, and their corresponding $MRAE$ and $MRAE^*$.

Local modeling of interest rate allows to obtain an overall good fitting as highlighted by low $MRAE$ and $MRAE^*$ values reported in table 6.4. Figure 6.1 displays an instance of the simulation of the considered interest rate r and its prediction using Kalman filtering. We used the estimated parameters to perform the Monte Carlo simulation as described earlier in section 6.2 to obtain estimations of one-step ahead 95% and 99% VaR. We also computed the 95% and 99% VaR using the HS method and VC method previously described. Unconditional and conditional tests using the considered simulation for Monte Carlo simulation, historical simulation method and variance/covariance method are reported in table 6.5.

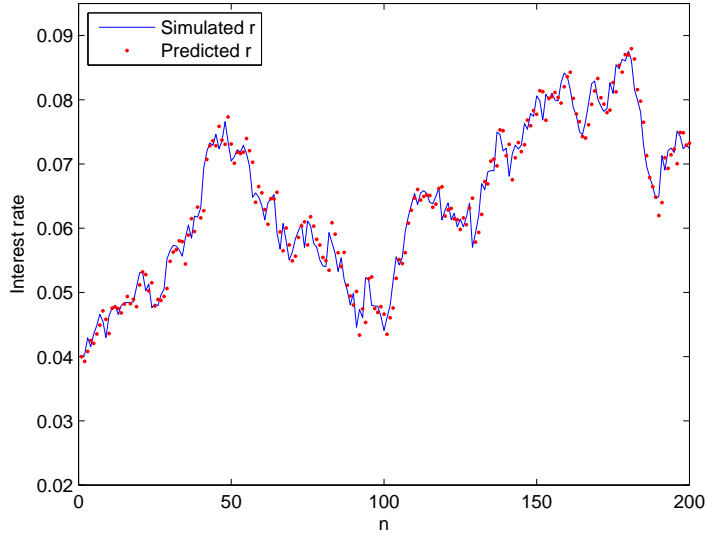


Figure 6.1: An instance of the simulated interest rate r considered and its prediction using Kalman filtering.

Method	α	\hat{x}	N_{00}	N_{01}	N_{10}	N_{11}	LR_{uc}	LR_{ind}	LR_{cc}	A/R
MC	95%	12	228	10	10	2	0.0213	2.5109	2.5322	A
	99%	4	242	4	4	0	0.7691	0.1301	0.8992	A
HS	95%	13	225	12	12	1	0.0208	0.1528	0.1736	A
	99%	7	236	7	7	0	5.4970	0.4033	5.9003	R
VC	95%	15	221	14	14	1	0.4961	0.0122	0.5082	A
	99%	9	232	9	9	0	10.2290	0.6724	10.9014	R

Table 6.5: Summary of test results for the considered simulation.

Table 6.5 reports the 95% and 99% VaR backtesting outcomes for an instance of a process described by equations (3.17) and (3.18) having values reported in table 6.1. The values in bold represents either values outside the non-rejection confidence intervals or values above the corresponding critical value. In particular the last column of each table indicates whether the VaR, at the specified level of confidence, estimated using either MC simulation, HS method or VC method is accepted (A) or rejected (R). In this instance while 95% and 99% VaR estimates obtained using MC are both accepted, 95% VaR estimates using HS and VC are accepted but the 99% VaR estimates using HS and VC estimates are both rejected. We

will comment on these results later together with the results on CIR model, in subsection 6.2.5.

6.2.2 Sensitivity study for the two-factor Vasicek model

In this subsection we perform some sensitivity analysis on the Vasicek model adopted in subsection 6.2.1. We consider the parameters included in table 6.1 and apply small changes to a couple of them. Then we simulate a process using the new set of parameter values and carry out the analysis reported in subsection 6.2.1 to understand whether little deviations from the initial considered model lead to similar conclusions.

We chose to change $\theta_1 = 0.46$ and $\sigma_2 = 0.012$. Therefore the new values to simulate the two-factor Vasicek model described by equations (3.17) and (3.18) are reported in the table 6.6:

i	x_{i0}	k_i	θ_i	σ_i	λ_i
1	0.015	0.375	0.046	0.015	-0.18
2	0.025	0.02	0.014	0.012	-0.0001

Table 6.6: New set of coefficients used for the simulation of the two-factor Vasicek model.

The starting values set for the initialization of the KF algorithm, the features of the five bonds for the observable measurements and the features of the bonds included in the simulated portfolio are kept constant. The goodness of fit is assessed using the RAE. Table 6.7 reports the new parameter estimation for the five in-sample subset considered, and their corresponding mean of the relative absolute error for both the in-sample (indicated as $MRAE$) and the out-of-sample (indicated as $MRAE^*$):

Subset	1	2	3	4	5
k_1	0.3204	0.3530	0.3798	0.3491	0.3928
θ_1	0.0534	0.0370	0.0468	0.0517	0.0338
σ_1	0.0156	0.0133	0.0171	0.0144	0.0161
λ_1	-0.3012	-0.2611	-0.4316	-0.2686	-0.6370
k_2	0.0285	0.0594	0.0168	0.0393	0.0485
θ_2	0.0111	0.0090	0.0158	0.0018	0.0101
σ_2	0.0121	0.0122	0.0112	0.0105	0.0126
λ_2	-0.0001	-0.0005	-0.0003	-0.0002	-0.0001
h_1	0.0008	0.0014	0.0011	0.0009	0.0007
h_2	0.0011	0.0010	0.0011	0.0014	0.0012
h_3	0.0014	0.0011	0.0013	0.0009	0.0010
h_4	0.0012	0.0010	0.0013	0.0011	0.0010
h_5	0.0005	0.0005	0.0003	0.0008	0.0006
$MRAE(\%)$	1.21	0.79	0.90	1.14	1.52
$MRAE^*(\%)$	1.05	1.17	1.31	1.22	1.01

Table 6.7: Estimated parameters for the new subset considered, and their corresponding $MRAE$ and $MRAE^*$.

Local modeling of interest rate using the new set of values allows to obtain an overall good fitting as highlighted by low $MRAE$ and $MRAE^*$ values reported in table 6.7. We used the estimated parameters to perform the Monte Carlo simulation as described earlier in subsection 6.2.1 to obtain estimations of one-step ahead 95% and 99% VaR. We also computed the 95% and 99% VaR using the HS method and VC method previously described. Unconditional and conditional tests using the considered simulation for Monte Carlo simulation, historical simulation method and variance/covariance method are reported in table 6.8.

Method	α	\tilde{x}	N_{00}	N_{01}	N_{10}	N_{11}	LR_{uc}	LR_{ind}	LR_{cc}	A/R
MC	95%	13	225	12	12	1	0.0208	0.1528	0.1736	A
	99%	3	244	3	3	0	0.0949	0.0729	0.1678	A
HS	95%	16	218	16	16	0	0.9514	2.1897	3.1411	A
	99%	7	237	6	6	1	5.4970	1.8520	7.3490	R
VC	95%	17	217	16	16	1	1.5403	0.0252	1.5655	A
	99%	9	233	8	8	1	10.2290	1.0122	11.2412	R

Table 6.8: Summary of test results for the considered new simulation.

Table 6.8 reports the 95% and 99% VaR backtesting outcomes for an instance of a process described by equations (3.17) and (3.18) having values reported in table 6.6. The values in

bold represents either values outside the non-rejection confidence intervals or values above the corresponding critical value. Again, the last column of each table indicates whether the VaR, at the specified level of confidence, estimated using either MC simulation, HS method or VC method is accepted (A) or rejected (R). In this instance of the new process while 95% and 99% VaR estimates obtained using MC are both accepted, 95% VaR estimates using HS and VC are accepted but the 99% VaR estimates using HS and VC estimates are both rejected. These results are compatible with the ones obtained using the original set of parameters.

6.2.3 Simulation using the two-factor CIR model

The values to simulate the two-factor CIR model described by equations (3.34) and (3.35) are reported in the table 6.9:

i	x_{i0}	k_i	θ_i	σ_i	λ_i
1	0.015	0.6	0.025	0.075	-0.22
2	0.025	0.01	0.013	0.015	-0.05

Table 6.9: Coefficients used for the simulation of the two-factor CIR model.

The starting values set for the initialization of the KF algorithm are:

$$\hat{\mathbf{r}}_{0|0} = \begin{bmatrix} 0.02 \\ 0.02 \end{bmatrix} \quad \text{and} \quad V_{1|0} = 3 \cdot 10^{-5} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The observable measurements are supplied by four bonds whose features are reported in table 6.10:

j	$T_j(\text{years})$	$h_j'^2$
1	0.5	0.0009 ²
2	1	0.0011 ²
3	2	0.0012 ²
4	5	0.0006 ²

Table 6.10: Features of the bonds providing the measurement values.

T_j and $h_j'^2$ represent the maturity and the variance of the zero mean noise term that perturbs the measured bond yield, respectively. The features of the bonds included in the portfolio

under study in this subsection are the same used in subsection 6.2.1, reported in table 6.3. Similarly, the goodness of fit reached for both the in-sample subset and the out-of-sample subset was assessed considering the RAE. Table 6.11 reports the parameter estimation for the five in-sample subsets considered, and their corresponding mean of the relative absolute error for both the in-sample (indicated as $MRAE$) and the out-of-sample (indicated as $MRAE^*$):

Subset	1	2	3	4	5
k_1	0.6206	0.5998	0.6507	0.6368	0.6306
θ_1	0.0343	0.0289	0.0361	0.0229	0.0203
σ_1	0.0696	0.0634	0.0566	0.0730	0.0800
λ_1	-0.2751	-0.1754	-0.2235	-0.2514	-0.2354
k_2	0.0109	0.0121	0.0135	0.0140	0.0098
θ_2	0.0063	0.0094	0.0156	0.0145	0.0130
σ_2	0.0102	0.0112	0.0125	0.0136	0.0153
λ_2	-0.0490	-0.0883	-0.0332	-0.0753	-0.0415
h'_1	0.0001	0.0006	0.0007	0.0006	0.0006
h'_2	0.0013	0.0011	0.0012	0.0010	0.0011
h'_3	0.0013	0.0009	0.0009	0.0009	0.0008
h'_4	0.0013	0.0011	0.0012	0.0009	0.0011
$MRAE(\%)$	1.38	1.20	1.15	1.12	1.09
$MRAE^*(\%)$	0.98	1.18	1.29	0.90	0.81

Table 6.11: Estimated parameters for the subset considered, and their corresponding $MRAE$ and $MRAE^*$.

We used the estimated parameters to perform the MC simulation as described earlier in section 6.2 to obtain estimations of one-step ahead 95% and 99% VaR. We also computed the 95% and 99% VaR using the HS method and VC method previously described. Unconditional and conditional tests using the considered simulation for Monte Carlo simulation, historical simulation method and variance/covariance method are reported in table 6.12.

Method	α	\tilde{x}	N_{00}	N_{01}	N_{10}	N_{11}	LR_{uc}	LR_{ind}	LR_{cc}	A/R
MC	95%	19	213	18	18	1	3.0905	0.1764	3.2669	A
	99%	6	238	6	6	0	3.5554	0.2951	3.8505	A
HS	95%	24	204	22	22	2	8.8777	0.0509	8.9286	R
	99%	5	240	5	5	0	1.9568	0.2041	2.1609	A
VC	95%	22	208	20	20	2	6.2590	0.0025	6.2615	R
	99%	5	240	5	5	0	1.9568	0.2041	2.1609	A

Table 6.12: Summary of test results for the considered simulation.

Table 6.12 reports the 95% and 99% VaR backtesting outcomes for an instance of a process described by equations (3.34) and (3.35) having values reported in table 6.9. Again, the values in bold represents either values outside the non-rejection confidence intervals or values above the corresponding critical value. In particular the last column of each table indicates whether the VaR, at the specified level of confidence, estimated using either MC simulation, HS method or VC method is accepted (A) or rejected (R). In this instance while 95% and 99% VaR estimates obtained using MC are both accepted, 95% VaR estimates using HS and VC are both rejected but the 99% VaR estimates using HS and VC estimates are accepted.

6.2.4 Sensitivity study for the two-factor CIR model

In this subsection we perform some sensitivity analysis on the CIR model adopted in subsection 6.2.3. We consider the parameters included in table 6.9 and apply small changes to a couple of them. Then we simulate a process using the new set of parameter values and carry out the analysis reported in subsection 6.2.3 to understand whether little deviations from the initial considered model lead to similar conclusions.

We chose to change $\theta_1 = 0.023$ and $\sigma_2 = 0.0148$. Therefore the new values to simulate the two-factor CIR model described by equations (3.34) and (3.35) are reported in the table 6.13:

i	x_{i0}	k_i	θ_i	σ_i	λ_i
1	0.015	0.6	0.023	0.075	-0.22
2	0.025	0.01	0.013	0.0148	-0.05

Table 6.13: New set of coefficients used for the simulation of the two-factor CIR model.

The starting values set for the initialization of the KF algorithm, the features of the four bonds for the observable measurements and the features of the bonds included in the simulated portfolio are kept constant. The goodness of fit is assessed using the RAE. Table 6.14 reports the new parameter estimation for the five in-sample subset considered, and their corresponding mean of the relative absolute error for both the in-sample (indicated as $MRAE$) and the out-of-sample (indicated as $MRAE^*$):

Subset	1	2	3	4	5
k_1	0.5967	0.5901	0.6086	0.6433	0.5681
θ_1	0.0253	0.0263	0.0244	0.0251	0.0233
σ_1	0.0792	0.0774	0.0819	0.0668	0.0685
λ_1	-0.1786	-0.1521	-0.2539	-0.2472	-0.1994
k_2	0.0110	0.0080	0.0094	0.0116	0.0099
θ_2	0.0093	0.0160	0.0136	0.0129	0.0128
σ_2	0.0153	0.0131	0.0178	0.0125	0.0156
λ_2	-0.0704	-0.0678	-0.0502	-0.0401	-0.0422
h'_1	0.0005	0.0012	0.0008	0.0005	0.0013
h'_2	0.0009	0.0012	0.0009	0.0009	0.0007
h'_3	0.0009	0.0015	0.0008	0.0014	0.0012
h'_4	0.0008	0.0007	0.0005	0.0007	0.0003
$MRAE(\%)$	0.99	0.84	0.93	1.19	0.98
$MRAE^*(\%)$	1.31	1.44	1.08	1.26	0.95

Table 6.14: Estimated parameters for the new subset considered, and their corresponding $MRAE$ and $MRAE^*$.

Local modeling of interest rate using the new set of values allows to obtain an overall good fitting as highlighted by low $MRAE$ and $MRAE^*$ values reported in table 6.14. We used the estimated parameters to perform the Monte Carlo simulation as described earlier in subsection 6.2.3 to obtain estimations of one-step ahead 95% and 99% VaR. We also computed the 95% and 99% VaR using the HS method and VC method previously described. Unconditional and conditional tests using the considered simulation for Monte Carlo simulation, historical simulation method and variance/covariance method are reported in table 6.15.

Method	α	\tilde{x}	N_{00}	N_{01}	N_{10}	N_{11}	LR_{uc}	LR_{ind}	LR_{cc}	A/R
MC	95%	17	216	17	17	0	1.5403	2.4829	4.0232	A
	99%	5	240	5	5	0	1.9568	0.02041	2.1609	A
HS	95%	20	212	18	18	2	4.0395	0.1108	4.1503	R
	99%	4	242	4	4	0	0.7691	0.1301	0.8992	A
VC	95%	20	210	20	20	0	4.0395	3.4827	7.5222	R
	99%	3	244	3	3	0	0.0949	0.0729	0.1678	A

Table 6.15: Summary of test results for the considered new simulation.

Table 6.15 reports the 95% and 99% VaR backtesting outcomes for an instance of a process described by equations (3.34) and (3.35) having values reported in table 6.13. The values in bold represents either values outside the non-rejection confidence intervals or values above

the corresponding critical value. Again, the last column of each table indicates whether the VaR, at the specified level of confidence, estimated using either MC simulation, HS method or VC method is accepted (A) or rejected (R). In this instance of the new process while 95% and 99% VaR estimates obtained using MC are both accepted, 95% VaR estimates using HS and VC are accepted but the 99% VaR estimates using HS and VC estimates are both rejected. These results are compatible with the ones obtained using the original set of parameters.

6.2.5 Comments on the simulation experiments

One can notice that the estimation for the two-factor Vasicek model uses five yields, while the estimation for the two-factor CIR model only requires four yields. A higher number of yields has been required by the Vasicek algorithm to achieve a goodness of fit comparable to the one achieved by the CIR algorithm. Specifically, using only four yields we managed to obtain a *MRAE* between 2% and 2.6%, which could result on its own acceptable unless one considers the maximum *RAE* that could reach values in a range between 16% and 24%. Employing these estimates could lead to a high number of biased VaR estimates, hence the choice of a different number of yields. For both the estimations the *MRAE* is just above 1% and the maximum *RAE* ranges between 2% and 6%.

It might be worth noticing that even under substantial changes in the initialization values this KF algorithm setting seems to produce similar results.

Other instances (not reported here) of the considered interest rate simulation lead to the conclusion that MC and VC methods, at least in this context, are not reliable for the estimation of 95% and 99% VaR. As reported in the previous two subsections, their backtesting fail to be accepted. The assumption that the past provides a fair representation of the immediate future seems to be not always true. The ability to estimate VaR using MC and VC methods seems to depend on the specific path simulated. This might also depend on the (implicit) assumption of portfolio returns being independent and identically distributed, which appears to be unrealistic. Furthermore all the rates of returns are given equal weights, where it might be more appropriate to assign different weights according to the fact that data further away from the present have a lower predicting influence compared to the closer ones. Monte Carlo simulation, provided that the fitting is sufficiently good, leads to VaR

estimates that are not rejected by the unconditional and conditional tests.

Varying the coefficients for the simulation of the interest rate as reported in subsections 6.2.2 and 6.2.4 leads to similar conclusions. Similarly, varying the features of the bonds included in the considered portfolio leads to similar conclusions and are not reported here for brevity.

Having seen the performance of our method with a simulated portfolio, we carry out the same tests using real bond prices.

6.3 Numerical experiments with real data

The aim of this section is to compute 95% and 99% VaR using Vasicek and CIR Monte Carlo simulations for an actual portfolio of bonds and to compare their performances against the results obtained using the HS method and the VC method, as described in the subsection 6.2. This is achieved in three steps:

1. Calibrating the two-factor Vasicek term structure model and the two-factor CIR term structure model using Kalman filter for the chosen dataset;
2. Computing the VaR at the required confidence level, as described in subsection 6.2;
3. Backtesting the one-step ahead forecasting, as described in section 4.1.

Unlike in the simulation experiments, the interest rate values are unknown. Hence we need to compare bond yields. The goodness of fitness can be assessed considering the RAE, which is defined for the experiments using real data as:

$$\text{RAE} = \frac{|\text{observed yield} - \text{predicted yield}|}{\text{observed yield}}.$$

Subsection 6.3.1 introduces the data employed and explains how it is used to calibrate the parameters, while subsection 6.3.2 reports the results of backtesting the estimates obtained with the proposed method.

6.3.1 Data

The dataset employed for this experiment consists of 450 weekly yields computed on three groups of US government bonds from 28/06/2001 to 11/10/2012: three batches of very short term bonds (5 to 7 months maturity), two batches of short term bonds (4.5 to 5.5 years maturity) and one batch of medium term bonds (10 years maturity). All the data was retrieved from Datastream. This data was split into five in-sample and five out-of-sample data sets using exactly the same procedure as used in subsection 6.2 for simulated data set of same size. We calibrated five two-factor Vasicek models and five two-factor CIR models on the six different in-sample yields data-sets mentioned above, each with 14 parameters. Each in-sample estimation is used to produce one-step ahead forecasting of yields. The procedure adopted for the experiments using real data is similar to the one chosen for the experiment using simulated data. We used five in-sample subsets and follow the four steps described in subsection 6.2. Computation of statistics $q_{1-\alpha}$ and $q_{1-\alpha}^N$ to obtain VaR through HS and VC methods are calculated using actual bond values.

6.3.2 Results

Tables 6.16 and 6.17 display the estimated parameters for the two-factor Vasicek model and the two-factor CIR model obtained using the mentioned five in-sample data subsets.

Subset	1	2	3	4	5
k_1	0.7030	0.7118	0.7095	0.07023	0.06891
θ_1	0.0056	0.0047	0.0049	0.0048	0.0045
σ_1	0.0321	0.0314	0.0332	0.0327	0.0288
λ_1	-0.4591	-0.4606	-0.4842	-0.4751	-0.4553
k_2	0.0255	0.0275	0.0231	0.0241	0.0219
θ_2	0.0035	0.0035	0.0028	0.0030	0.0029
σ_2	0.0142	0.0154	0.0138	0.0144	0.0127
λ_2	-0.2652	-0.2639	-0.2528	-0.2509	-0.2629
h_1	0.0009	0.0011	0.0010	0.0009	0.0013
h_2	0.0012	0.0006	0.0011	0.0011	0.0011
h_3	0.0013	0.0010	0.0006	0.0009	0.0009
h_4	0.0007	0.0010	0.0007	0.0011	0.0007
h_5	0.0009	0.0004	0.0012	0.0011	0.0009
h_6	0.0010	0.0008	0.0009	0.0008	0.0009
$MRAE(\%)$	1.22	1.34	1.18	1.36	1.15
$MRAE^*(\%)$	1.16	1.09	1.29	1.12	1.20

Table 6.16: Estimated parameters for two-factor Vasicek model.

Subset	1	2	3	4	5
k_1	0.6102	0.6089	0.6095	0.6064	0.5984
θ_1	0.0201	0.0221	0.0223	0.0217	0.0206
σ_1	0.0424	0.0411	0.0402	0.0408	0.0394
λ_1	-0.3861	-0.3775	-0.3747	-0.3721	-0.3788
k_2	0.0304	0.0315	0.321	0.0313	0.0308
θ_2	0.0125	0.0131	0.0128	0.0127	0.0125
σ_2	0.0134	0.0142	0.0138	0.0136	0.0133
λ_2	-0.0192	-0.0208	-0.0205	-0.0201	-0.0189
h'_1	0.0011	0.0012	0.0007	0.0009	0.0009
h'_2	0.0009	0.0010	0.0011	0.0011	0.0012
h'_3	0.0013	0.0010	0.0009	0.0008	0.0008
h'_4	0.0008	0.0010	0.0011	0.0010	0.0008
h'_5	0.0008	0.0009	0.0012	0.0011	0.0009
h'_6	0.0001	0.0008	0.0010	0.0007	0.0010
$MRAE(\%)$	1.19	1.32	1.24	1.25	1.21
$MRAE^*(\%)$	1.06	1.22	1.31	1.23	1.17

Table 6.17: Estimated parameters for two-factor CIR model.

The estimated values were used to calculate 50 one-step ahead interest rate prediction, and used to carry out Monte Carlo simulations as described in subsection 6.2. Each Monte Carlo simulation consists of 10,000 trials. The portfolio considered includes an initial number of 50,000 of each of the six bonds. Table 6.18 reports conditional and unconditional tests using the considered simulation for Monte Carlo simulation, historical simulation method and variance/covariance method using real data.

Method	α	\tilde{x}	N_{00}	N_{01}	N_{10}	N_{11}	LR_{uc}	LR_{ind}	LR_{cc}	A/R
MC Vasicek	95%	16	219	15	15	1	0.9514	0.0006	0.9520	A
	99%	5	240	5	5	0	1.9568	0.2041	2.1609	A
MC CIR	95%	15	220	15	15	0	0.4961	1.9162	2.4123	A
	99%	4	242	4	4	0	0.7691	0.1301	0.8992	A
HS	95%	17	218	15	15	2	1.5403	0.5996	2.1399	A
	99%	7	237	6	6	1	5.4970	1.8520	7.3490	R
VC	95%	16	220	14	14	2	0.9514	1.6762	2.6276	A
	99%	6	238	6	6	0	3.5554	0.2952	3.8506	A

Table 6.18: Summary of test results for the experiments with real data.

Both Vasicek and CIR Monte Carlo simulations and variance/covariance method provide acceptable estimations of both 95% and 99% VaR, while historical simulation method just bears an acceptable 95% VaR but fail to provide a valid 99% VaR. As highlighted in the previous section some assumptions such that the past provides a fair representation of the immediate future and that rates of return are normally distributed might be not realistic. In this instance, historical simulation method produce too many exceptions in estimates of 99% VaR, therefore the evaluation of 99% VaR is rejected. The calibrated model seems accurate enough for the purpose of estimating quantiles, since the conditional and unconditional tests are non-rejected for both 95% and 99% VaR estimates.

In general, both Vasicek and CIR models are employed to fit the term structure using real financial data. Empirical evidence suggests that the Vasicek model performs better than the CIR model for pricing bonds in terms of MRAE and MRAE*, see Wang (2008). One can also observe that having constraints on coefficients of CIR model for interest rate positivity plus constraints on allowed correlation coefficients of risk factors may lead to poor fits to bond prices, see e.g. Dai and Singleton (2000). For measuring the risk of portfolios they seem to perform equally well.

Both the models employed in the numerical experiments reported in this thesis seem to provide an accurate goodness of fit. Furthermore, the summary test results confirm that 95% and 99% VaR estimates using the two-factor Vasicek model and the two-factor CIR model are equally adequate.

6.4 Summary

In this chapter we carry out numerical experiments with both simulated data and real data to test the behavior of the proposed methods. In subsections 6.2.1 and 6.2.3 we simulate a path using a Vasicek model and a CIR model, respectively. In each of these subsections we estimate the parameters of the respective simulated path and then perform Monte Carlo simulation to obtain one-step ahead 95% and 99% VaR estimates for a generic portfolio. We also computed the 95% and 99% VaR using the historical simulation method and the variance/covariance method. Next, we backtested the results obtained using Monte Carlo simulation, historical simulation method and variance covariance method and compare their performances.

In section 6.3 we measure the risk for a portfolio that includes an initial number of 50,000 of six batches of bonds. The employed bond batches consist of three batches of very short term bonds, two batches of short term bonds and one batch of medium term bonds. We estimate the parameters for both the Vasicek model and the CIR model, and then estimate 95% and 99% VaR. We also compute 95% and 99% VaR using historical simulation method and variance/covariance method, backtest the results and compare their performances.

The results indicate that historical simulation method and variance/covariance method, relying on the assumption that the past provides a fair representation of the immediate future, could provide biased 95% and 99% VaR estimations. This fact suggests that the assumption, at least in this context, is not realistic. Using Monte Carlo simulation employing either the two-factor Vasicek model or the two-factor CIR model seem to provide acceptable 95% and 99% VaR estimations, as long as the relative parameters display a good degree of fitness.

Chapter 7

Summary of contributions and future research

In this chapter the main contributions are summarized and some suggestions for further research are reported. The main focus of the thesis was to investigate computationally efficient ways of measuring the risk of portfolios whose values are nonlinear functions of possibly non-Gaussian risk factors. The research presented draws on mathematical tools from a variety of fields such as applied statistics (probability conserving transformation), signal processing (Fast Fourier Transforms) and systems theory (Kalman filtering). Specifically, two methods for computation of Value-at-Risk were investigated:

1. Portfolios where both the risk factors and the nonlinearities have no specific structure. In this case, we developed a method, called Delta-Gamma-Q, which allows us to compute VaR and CVaR through a combination of Delta-Gamma-Normal model in Glasserman (2003) and probability conserving transformation in Sornette et al. (2000). In this method, the marginal distributions of risk factors are mapped through nonlinear changes of variables onto Gaussian distributions. A new covariance matrix can therefore be computed and it redefines the dependence among transformed risk factors. Delta-Gamma coefficients obtained under normal conditions are multiplied by factors that take into account the shape of risk factors and then fast Fourier transform allows us to perform a quick computation of VaR and CVaR values. This method was described in chapter 2 and was demonstrated through comprehensive experiments

on simulated as well as real data in chapter 5. These experiments show that using Delta-Gamma-dP Normal method for non-normal risk factors leads to misleading results while our method corrects the bias to a significant extent. The numerical results are shown to be consistent across a range of parameter values, across two different confidence levels and across two different time horizons. The method presented here emphasized the fact that using methods that do not recognize the lack of normality can lead to rather biased estimates of both VaR and CVaR, especially in nonlinear portfolios. The method presented features a good degree of flexibility since the use of nonparametric estimation can capture the distribution characteristics of risk factors to be analyzed and the computational effort is lower than partial Monte Carlo simulation.

This study focusses only on dealing with possible skewness and fat tails of the portfolio distribution. As mentioned in subsection 2.6.2, it would be interesting to see if the proposed framework can be extended to prediction of VaR in a dynamic setting, perhaps including phenomena such as autoregressive behavior and clustering of volatility. Further, the current work is limited to analysis of a given (or fixed) portfolio. The use of probability conserving transformation and the subsequent transformed correlations in selecting a portfolio which optimizes risk/return characteristics also indicates a potentially interesting direction for future research.

2. The second method uses a specific nonlinearity (bond prices) and specific structure for the risk factors (either Vasicek or CIR model for the short rate). This generates scenarios which are arbitrage-free and provides a parsimonious as well as realistic means of generating scenarios for Monte Carlo based evaluation of VaR for fixed income portfolios in a liquid (and hence arbitrage-free) market. The KF-based method requires simulating only a vector of two random variables for one step ahead forecasts and is hence computationally cheaper as compared to principal components analysis using more than two principal components. Numerical experiments with simulated data as well as real treasuries data confirm the utility of our method in measuring tail risk.

The analysis included here is limited to compute the Value-at-Risk of a bond portfolio. A possible future research direction could be to extend portfolios composition including other nonlinear interest instruments such as interest rate caps and floors.

Appendix A

Function Documentation

A substantial amount of MATLAB-based re-useable code was produced during this project. This Appendix provides a description of the code developed and used to compute the results.

`Like.m`

The function `Like` is in the form: `[LRuc,LRind,LRcc]=Like(N,p,n,n00,n01,n10,n11)`.

The function takes as input arguments:

1. N is the total number of observations,
2. p is the VaR confidence level,
3. n is the number of exceptions in the sample,
4. $n00$ is the number of days where a state 0 is followed by a state 0,
5. $n01$ is the number of days where a state 0 is followed by a state 1,
6. $n10$ is the number of days where a state 1 is followed by a state 0,
7. $n11$ is the number of days where a state 1 is followed by a state 1.

The output arguments are:

1. $LRuc$ is the unconditional coverage test,
2. $LRind$ is the independence test,
3. $LRcc$ is the conditional coverage test.

`Empirical50.m`

The script used to simulate a portfolio made up of 50 correlated stocks and 50 options employing the empirical method with parameters as specified in section 5.2. It returns 95% and 99% VaR & CVaR and confidence intervals for 95% and 99% VaR, as reported in tables 5.3–5.4.

`RemainingSimulation50.m`

The script used to simulate a portfolio made up of 50 correlated stocks and 50 options employing the Delta-Gamma-dP method, the Delta-Gamma-Q method, the Delta-Gamma-dP Normal method and the Delta-Gamma-dP inversion method with parameters as specified in section 5.2. It returns 95% and 99% VaR & CVaR and confidence intervals for 95% and 99% VaR (where relevant), as reported in tables 5.3–5.4.

`BacktestingReal.m`

The script used to compute the measures of risk for the real portfolio considered in section 5.3. It returns 95% and 99% VaR & CVaR using the Delta-Gamma-dP Normal method and 95% and 99% VaR & CVaR using the Delta-Gamma-Q method, employing the algorithms mentioned in subsections 5.3.3 and 5.3.4, respectively.

`SimV.m`

The script used to simulate a sample of an interest rate path using the Vasicek process specified in subsection 6.2.1. For each maturity of the bond included it computes and visualizes the corresponding yield.

`SimCIR.m`

The script used to simulate a sample of an interest rate path using the Cox-Ingersoll-Ross process specified in subsection 6.2.3. For each maturity of the bond included it computes and visualizes the corresponding yield.

`EstimationAlgo.m`

The script used to compute parameter estimations and *MRAE* for both Vasicek model and CIR model, reported in tables 6.4 and 6.11, respectively. It requires simulations of yields and their associated times to maturity generated using `SimV` and `SimCIR`, respectively, as an

input. It returns the estimated parameters and MRAE of the corresponding model for the subset considered. This script calls the functions `KFV` and `KFCIR`.

`KFV.m`

The function `KFV` is in the form:

`[f , xP1 , xP2 , PP11 , PP12 , PP21 , PP22] =KFV(x , y1 , t1 , y2 , t2 , y3 , t3 , y4 , t4 , y5 , t5)`.

The function takes as input arguments:

1. x is the vector of parameters for the Vasicek Kalman filter,
2. $y1$ is the vector with the in-sample yields of the first simulated bond,
3. $t1$ is the vector with the in-sample times to maturity for the first simulated bond,
4. $y2$ is the vector with the in-sample yields of the second simulated bond,
5. $t2$ is the vector with the in-sample times to maturity for the second simulated bond,
6. $y3$ is the vector with the in-sample yields of the third simulated bond,
7. $t3$ is the vector with the in-sample times to maturity for the third simulated bond,
8. $y4$ is the vector with the in-sample yields of the fourth simulated bond,
9. $t4$ is the vector with the in-sample times to maturity for the fourth simulated bond.
10. $y5$ is the vector with the in-sample yields of the fourth simulated bond,
11. $t5$ is the vector with the in-sample times to maturity for the fourth simulated bond.

The output arguments are:

1. f is the value of the likelihood function,
2. $xP1$ is the value of the estimation $\hat{r}_{1|n}$,
3. $xP2$ is the value of the estimation $\hat{r}_{2|n}$,
4. $PP11$ is the value in the matrix $V_{n+1|n}$ indexed as 11,
5. $PP12$ is the value in the matrix $V_{n+1|n}$ indexed as 12,
6. $PP21$ is the value in the matrix $V_{n+1|n}$ indexed as 21,

7. $PP22$ is the value in the matrix $V_{n+1|n}$ indexed as 22.

KFCIR.m

The function KFCIR is in the form:

$[f, xP1, xP2, PP11, PP12, PP21, PP22] = \text{KFCIR}(x, y1, t1, y2, t2, y3, t3, y4, t4)$.

The function takes as input arguments:

1. x is the vector of parameters for the Cox-Ingersoll-Ross Kalman filter,
2. $y1$ is the vector with the in-sample yields of the first simulated bond,
3. $t1$ is the vector with the in-sample times to maturity for the first simulated bond,
4. $y2$ is the vector with the in-sample yields of the second simulated bond,
5. $t2$ is the vector with the in-sample times to maturity for the second simulated bond,
6. $y3$ is the vector with the in-sample yields of the third simulated bond,
7. $t3$ is the vector with the in-sample times to maturity for the third simulated bond,
8. $y4$ is the vector with the in-sample yields of the fourth simulated bond,
9. $t4$ is the vector with the in-sample times to maturity for the fourth simulated bond.

The output arguments are:

1. f is the value of the likelihood function,
2. $xP1$ is the value of the estimation $\hat{r}_{1|n}$,
3. $xP2$ is the value of the estimation $\hat{r}_{2|n}$,
4. $PP11$ is the value in the matrix $V_{n+1|n}$ indexed as 11,
5. $PP12$ is the value in the matrix $V_{n+1|n}$ indexed as 12,
6. $PP21$ is the value in the matrix $V_{n+1|n}$ indexed as 21,
7. $PP22$ is the value in the matrix $V_{n+1|n}$ indexed as 22.

PredictV.m

The script used to compute $MRAE^*$ for the Vasicek simulation as described in subsection 6.2.1 and reported in table 6.4. It requires simulations of yields and their associated times to maturity generated using SimV. It returns the $MRAE^*$ for the subset considered. This script calls the functions KFV and ForeV.

ForeV.m

The function ForeV is in the form:

```
[X1,X2,PP11,PP12,PP21,PP22]=ForeV(x,xPe1,xPe2,Pe11,Pe12,Pe21,Pe22,  
y1,t1,y2,t2,y3,t3,y4,t4,y5,t5).
```

The function takes as input arguments:

1. x is vector of the estimated parameters for the subset considered,
2. $xPe1$ is the initialization value $\hat{r}_{1|0}$,
3. $xPe2$ is the initialization value $\hat{r}_{2|0}$,
4. $Pe11$ is the value in the initialization matrix $V_{1|0}$ indexed as 11,
5. $Pe12$ is the value in the initialization matrix $V_{1|0}$ indexed as 12,
6. $Pe21$ is the value in the initialization matrix $V_{1|0}$ indexed as 21,
7. $Pe22$ is the value in the initialization matrix $V_{1|0}$ indexed as 22,
8. $y1$ is the vector with the out-of-sample yields of the first simulated bond,
9. $t1$ is the vector with the out-of-sample times to maturity for the first simulated bond,
10. $y2$ is the vector with the out-of-sample yields of the second simulated bond,
11. $t2$ is the vector with the out-of-sample times to maturity for the second simulated bond,
12. $y3$ is the vector with the out-of-sample yields of the third simulated bond,
13. $t3$ is the vector with the out-of-sample times to maturity for the third simulated bond,
14. $y4$ is the vector with the out-of-sample yields of the fourth simulated bond,

15. $t4$ is the vector with the out-of-sample times to maturity for the fourth simulated bond,
16. $y5$ is the vector with the out-of-sample yields of the fifth simulated bond,
17. $t5$ is the vector with the out-of-sample times to maturity for the fifth simulated bond.

The output arguments are:

1. $X1$ is the out-of-sample estimation $\hat{r}_{1n|n}$,
2. $X2$ is the out-of-sample estimation $\hat{r}_{2n|n}$,
3. $PP11$ is the value in the matrix $V_{n+1|n}$ indexed as 11,
4. $PP12$ is the value in the matrix $V_{n+1|n}$ indexed as 12,
5. $PP21$ is the value in the matrix $V_{n+1|n}$ indexed as 21,
6. $PP22$ is the value in the matrix $V_{n+1|n}$ indexed as 22.

PredictCIR.m

The script used to compute $MRAE^*$ for the CIR simulation as described in 6.2.3 and reported in table 6.11. It requires simulations of yields and their associated times to maturity generated using **SimCIR**. It returns the $MRAE^*$ for the subset considered. This script calls the functions **KFCIR** and **ForeCIR**.

ForeCIR.m

The function **ForeCIR** is in the form:

```
[X1,X2,PP11,PP12,PP21,PP22]=ForeCIR(x,xPe1,xPe2,Pe11,Pe12,Pe21,Pe22,
y1,t1,y2,t2,y3,t3,y4,t4).
```

The function takes as input arguments:

1. x is vector of the estimated parameters for the subset considered,
2. $xPe1$ is the initialization value $\hat{r}_{10|0}$,
3. $xPe2$ is the initialization value $\hat{r}_{20|0}$,
4. $Pe11$ is the value in the initialization matrix $V_{1|0}$ indexed as 11,

5. *Pe12* is the value in the initialization matrix $V_{1|0}$ indexed as 12,
6. *Pe21* is the value in the initialization matrix $V_{1|0}$ indexed as 21,
7. *Pe22* is the value in the initialization matrix $V_{1|0}$ indexed as 22,
8. *y1* is the vector with the out-of-sample yields of the first simulated bond,
9. *t1* is the vector with the out-of-sample times to maturity for the first simulated bond,
10. *y2* is the vector with the out-of-sample yields of the second simulated bond,
11. *t2* is the vector with the out-of-sample times to maturity for the second simulated bond,
12. *y3* is the vector with the out-of-sample yields of the third simulated bond,
13. *t3* is the vector with the out-of-sample times to maturity for the third simulated bond,
14. *y4* is the vector with the out-of-sample yields of the fourth simulated bond,
15. *t4* is the vector with the out-of-sample times to maturity for the fourth simulated bond.

The output arguments are:

1. *X1* is the out-of-sample estimation $\hat{r}_{1\ n|n}$,
2. *X2* is the out-of-sample estimation $\hat{r}_{2\ n|n}$,
3. *PP11* is the value in the matrix $V_{n+1|n}$ indexed as 11,
4. *PP12* is the value in the matrix $V_{n+1|n}$ indexed as 12,
5. *PP21* is the value in the matrix $V_{n+1|n}$ indexed as 21,
6. *PP22* is the value in the matrix $V_{n+1|n}$ indexed as 22.

HS & VC - V.m

The script used to simulate an interest rate using the two-factor Vasicek model using the parameters specified in subsection 6.2.1 and to compute VaR employing historical simulation method and variance/covariance method as described in section 2.3. It returns the values \tilde{x} ,

N_{00} , N_{01} , N_{10} and N_{11} , which will provide the statistics for backtesting the VaR estimates using the function `Like`.

`HS & VC - CIR.m`

The script used to simulate an interest rate using the two-factor CIR model using the parameters specified in subsection 6.2.3 and to compute VaR employing historical simulation method and variance/covariance method as described in section 2.3. It returns the values \tilde{x} , N_{00} , N_{01} , N_{10} and N_{11} , which will provide the statistics for backtesting the VaR estimates using the function `Like`.

`CompuVaRV.m`

The script used to compute the 95% and 99% VaR using the Vasicek simulation explained in section 6.2. It requires bond yields simulations generated using `SimV` and calls the functions `KFV` and `ForeV`.

`CompuVarCIR.m`

The script used to compute the 95% and 99% VaR using the CIR simulation in section 6.2. It requires bond yields simulations generated using `SimCIR` and calls the functions `KFCIR` and `ForeCIR`.

`EstimationAlgo-r.m`

The script used to compute parameter estimations and *MRAE* for both Vasicek model and CIR model, reported in tables 6.16 and 6.17, respectively. It requires real bond yields and their associated times to maturity as an input. It returns the estimated parameters and *MRAE* of the corresponding model for the subset considered. This script calls the functions `KFVr` and `KFCIRr`.

`KFVr.m`

The function `KFVr` is in the form:

`[f, xP1, xP2, PP11, PP12, PP21, PP22]=KFVr(x, y1, t1, y2, t2, y3, t3, y4, t4, y5, t5, y6, t6).`

The function takes as input arguments:

1. x is the vector of parameters for the Vasicek Kalman filter,

2. $y1$ is the vector with the in-sample yields of the first bond,
3. $t1$ is the vector with the in-sample times to maturity for the first bond,
4. $y2$ is the vector with the in-sample yields of the second bond,
5. $t2$ is the vector with the in-sample times to maturity for the second bond,
6. $y3$ is the vector with the in-sample yields of the third bond,
7. $t3$ is the vector with the in-sample times to maturity for the third bond,
8. $y4$ is the vector with the in-sample yields of the fourth bond,
9. $t4$ is the vector with the in-sample times to maturity for the fourth bond,
10. $y5$ is the vector with the in-sample yields of the fifth bond,
11. $t5$ is the vector with the in-sample times to maturity for the fifth bond,
12. $y6$ is the vector with the in-sample yields of the sixth bond,
13. $t6$ is the vector with the in-sample times to maturity for the sixth bond.

The output arguments are:

1. f is the value of the likelihood function,
2. $xP1$ is the value of the estimation $\hat{r}_{1n|n}$,
3. $xP2$ is the value of the estimation $\hat{r}_{2n|n}$,
4. $PP11$ is the value in the matrix $V_{n+1|n}$ indexed as 11,
5. $PP12$ is the value in the matrix $V_{n+1|n}$ indexed as 12,
6. $PP21$ is the value in the matrix $V_{n+1|n}$ indexed as 21,
7. $PP22$ is the value in the matrix $V_{n+1|n}$ indexed as 22.

`KFCIRr.m`

The function `KFCIRr` is in the form:

`[f, xP1, xP2, PP11, PP12, PP21, PP22]=KFCIRr(x, y1, t1, y2, t2, y3, t3, y4, t4, y5, t5, y6, t6).`

The function takes as input arguments:

1. x is the vector of parameters for the Cox-Ingersoll-Ross Kalman filter,
2. $y1$ is the vector with the in-sample yields of the first bond,
3. $t1$ is the vector with the in-sample times to maturity for the first bond,
4. $y2$ is the vector with the in-sample yields of the second bond,
5. $t2$ is the vector with the in-sample times to maturity for the second bond,
6. $y3$ is the vector with the in-sample yields of the third bond,
7. $t3$ is the vector with the in-sample times to maturity for the third bond,
8. $y4$ is the vector with the in-sample yields of the fourth bond,
9. $t4$ is the vector with the in-sample times to maturity for the fourth bond,
10. $y5$ is the vector with the in-sample yields of the fifth bond,
11. $t5$ is the vector with the in-sample times to maturity for the fifth bond,
12. $y6$ is the vector with the in-sample yields of the sixth bond,
13. $t6$ is the vector with the in-sample times to maturity for the sixth bond.

The output arguments are:

1. f is the value of the likelihood function,
2. $xP1$ is the value of the estimation $\hat{r}_{1n|n}$,
3. $xP2$ is the value of the estimation $\hat{r}_{2n|n}$,
4. $PP11$ is the value in the matrix $V_{n+1|n}$ indexed as 11,
5. $PP12$ is the value in the matrix $V_{n+1|n}$ indexed as 12,
6. $PP21$ is the value in the matrix $V_{n+1|n}$ indexed as 21,
7. $PP22$ is the value in the matrix $V_{n+1|n}$ indexed as 22.

PredictVr.m

The script used to compute $MRAE^*$ for the Vasicek estimation using real data as described in 6.3 and reported in table 6.16. It requires real bond yields and their associated times to maturity as an input. It returns the $MRAE^*$ for the subset considered. This script calls the functions `KFVr` and `ForeVr`.

ForeVr.m

The function `ForeVr` is in the form:

```
[yf1a,yf1b,yf1c,yf1d,yf1e,yf1f,xP1,xP2,PP11,PP12,PP21,PP22]=ForeVr(x,x1,x2,  
Pe11,Pe12,Pe21,Pe22,y1,t1,y2,t2,y3,t3,y4,t4,y5,t5,y6,t6).
```

The function takes as input arguments:

1. x is vector of the estimated parameters for the subset considered,
2. $x1$ is the initialization value $\hat{r}_{1|0}$,
3. $x2$ is the initialization value $\hat{r}_{2|0}$,
4. $Pe11$ is the value in the initialization matrix $V_{1|0}$ indexed as 11,
5. $Pe12$ is the value in the initialization matrix $V_{1|0}$ indexed as 12,
6. $Pe21$ is the value in the initialization matrix $V_{1|0}$ indexed as 21,
7. $Pe22$ is the value in the initialization matrix $V_{1|0}$ indexed as 22,
8. $y1$ is the vector with the out-of-sample yields of the first real bond,
9. $t1$ is the vector with the out-of-sample times to maturity for the first real bond,
10. $y2$ is the vector with the out-of-sample yields of the second real bond,
11. $t2$ is the vector with the out-of-sample times to maturity for the second real bond,
12. $y3$ is the vector with the out-of-sample yields of the third real bond,
13. $t3$ is the vector with the out-of-sample times to maturity for the third real bond,
14. $y4$ is the vector with the out-of-sample yields of the fourth real bond,
15. $t4$ is the vector with the out-of-sample times to maturity for the fourth real bond,

16. *y5* is the vector with the out-of-sample yields of the fifth real bond,
17. *t5* is the vector with the out-of-sample times to maturity for the fifth real bond,
18. *y6* is the vector with the out-of-sample yields of the sixth real bond,
19. *t6* is the vector with the out-of-sample times to maturity for the sixth real bond.

The output arguments are:

1. *yfla* is the out-of-sample forecast of the first bond yield,
2. *yflb* is the out-of-sample forecast of the second bond yield,
3. *yflc* is the out-of-sample forecast of the third bond yield,
4. *yfld* is the out-of-sample forecast of the fourth bond yield,
5. *yfle* is the out-of-sample forecast of the fifth bond yield,
6. *yflf* is the out-of-sample forecast of the sixth bond yield,
7. *xP1* is the value of the estimation $\hat{r}_{1|n}$,
8. *xP2* is the value of the estimation $\hat{r}_{2|n}$,
9. *PP11* is the value in the matrix $V_{n+1|n}$ indexed as 11,
10. *PP12* is the value in the matrix $V_{n+1|n}$ indexed as 12,
11. *PP21* is the value in the matrix $V_{n+1|n}$ indexed as 21,
12. *PP22* is the value in the matrix $V_{n+1|n}$ indexed as 22.

PredictCIRr.m

The script used to compute $MRAE^*$ for the CIR estimation using real data as described in 6.3 and reported in table 6.17. It requires real bond yields and their associated times to maturity as an input. It returns the $MRAE^*$ for the subset considered. This script calls the functions **KFCIRr** and **ForeCIRr**.

ForeCIRr.m

The function ForeCIRr is in the form:

$[yf1a,yf1b,yf1c,yf1d,yf1e,yf1f,xP1,xP2,PP11,PP12,PP21,PP22]=\text{ForeCIRr}(x,x1,x2,Pe11,Pe12,Pe21,Pe22,y1,t1,y2,t2,y3,t3,y4,t4,y5,t5,y6,t6)$.

The function takes as input arguments:

1. x is vector of the estimated parameters for the subset considered,
2. $x1$ is the initialization value $\hat{r}_{1|0}$,
3. $x2$ is the initialization value $\hat{r}_{2|0}$,
4. $Pe11$ is the value in the initialization matrix $V_{1|0}$ indexed as 11,
5. $Pe12$ is the value in the initialization matrix $V_{1|0}$ indexed as 12,
6. $Pe21$ is the value in the initialization matrix $V_{1|0}$ indexed as 21,
7. $Pe22$ is the value in the initialization matrix $V_{1|0}$ indexed as 22,
8. $y1$ is the vector with the out-of-sample yields of the first real bond,
9. $t1$ is the vector with the out-of-sample times to maturity for the first real bond,
10. $y2$ is the vector with the out-of-sample yields of the second real bond,
11. $t2$ is the vector with the out-of-sample times to maturity for the second real bond,
12. $y3$ is the vector with the out-of-sample yields of the third real bond,
13. $t3$ is the vector with the out-of-sample times to maturity for the third real bond,
14. $y4$ is the vector with the out-of-sample yields of the fourth real bond,
15. $t4$ is the vector with the out-of-sample times to maturity for the fourth real bond,
16. $y5$ is the vector with the out-of-sample yields of the fifth real bond,
17. $t5$ is the vector with the out-of-sample times to maturity for the fifth real bond,
18. $y6$ is the vector with the out-of-sample yields of the sixth real bond,
19. $t6$ is the vector with the out-of-sample times to maturity for the sixth real bond.

The output arguments are:

1. *yfla* is the out-of-sample forecast of the first bond yield,
2. *yflb* is the out-of-sample forecast of the second bond yield,
3. *yflc* is the out-of-sample forecast of the third bond yield,
4. *yfl d* is the out-of-sample forecast of the fourth bond yield,
5. *yfle* is the out-of-sample forecast of the fifth bond yield,
6. *yflf* is the out-of-sample forecast of the sixth bond yield,
7. *xP1* is the value of the estimation $\hat{r}_{1 n|n}$,
8. *xP2* is the value of the estimation $\hat{r}_{2 n|n}$,
9. *PP11* is the value in the matrix $V_{n+1|n}$ indexed as 11,
10. *PP12* is the value in the matrix $V_{n+1|n}$ indexed as 12,
11. *PP21* is the value in the matrix $V_{n+1|n}$ indexed as 21,
12. *PP22* is the value in the matrix $V_{n+1|n}$ indexed as 22.

HS & VC - `r.m`

The script used to compute 95% and 99% VaR employing historical simulation method and variance/covariance method as described in section 6.3. It returns the values \tilde{x} , N_{00} , N_{01} , N_{10} and N_{11} , which will provide the statistics for backtesting the VaR estimates using the function `Like`.

`CompuVaRVr.m`

The script used to compute the 95% and 99% VaR using the Vasicek simulation explained in section 6.3. It requires real bond yields and calls the function `ForeVr`.

`CompuVaRCIRr.m`

The script used to compute the 95% and 99% VaR using the CIR simulation explained in section 6.3. It requires real bond yields and calls the function `ForeCIRr`.

Bibliography

- C. Acerbi. Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26(7):1505–1518, 2002.
- C. Acerbi and D. Tasche. Expected shortfall: a natural coherent alternative to Value at Risk. *Economic Notes*, 31:379–388, 2002.
- K. J. Adams and D.R. Van Deventer. Fitting yield curves and forward rate curves with maximum smoothness. *The Journal of Fixed Income*, 4:52–62, 1994.
- C. Albanese, K. Jackson, and P. Wiberg. A new Fourier transform algorithm for Value-at-Risk. *Quantitative Finance*, 4(3):328–338, 2004.
- C. Alexander. *Market risk analysis. Vol. 4, Value-at-Risk models*. John Wiley & Sons, Ltd., 2008.
- P. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.
- A. Azzalini and A. Capitanio. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew- t distribution. *Journal of the Royal Statistical Society Series B*, 65(2):367–389, 2003.
- S. H. Babbs and K. B. Nowman. Kalman filtering of generalized Vasicek term structure models. *Journal of Financial and Quantitative Analysis*, 34:115–130, 1999.
- A. A. Balkema and L. De Haan. Residual lifetime at great age. *Annals of Probability*, 2:792–804, 1974.
- Basel Committee. Basel Committee on banking supervision international convergence of capital measurement and capital standards, 2006.

- D. J. Bolder. Affine term-structure models: theory and implementation. *Bank of Canada, Working Paper No. 2001-15*, 2001.
- A. W. Bowman and A. Azzalini. *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*. Oxford Science Publications, 1997.
- D. Brigo and F. Mercurio. *Interest Rate Models – Theory and Practice*. Springer Finance. Springer, Berlin, 2006.
- M. Britten-Jones and S. M. Schaefer. Non-linear Value-at-Risk. *European Finance Review*, 2:161–187, 1999.
- S. D. Campbell. A review of backtesting and backtesting procedures. *Working paper in Finance and Economic Discussion series, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington D.C., USA*, 2005.
- P. Carr, D. B. Madan, and R. H. Smith. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2:61–73, 1999.
- A. Carvalhal and S. P. Daumas. A minimizing risk strategy for Brazilian fixed-income portfolios. *Latin American Business Review*, 11(1):1–18, 2010.
- G. Castellacci and M. J. Siclari. The practice of Delta-Gamma VaR: Implementing the quadratic portfolio model. *European Journal of Operational Research*, 150(3):529–545, 2003.
- J. B. Castellanos Pinzon. Estimating multifactor uncorrelated Vasicek and CIR model for the UK term structure. *Centre for Computational Finance and Economic Agents, University of Essex. Working Paper Series, WP024-08*, 2008.
- A. Cerný. Introduction to fast Fourier transform in finance. *The Journal of Derivatives*, 12:73–88, 2004.
- G. B. Chaplin and K. P. Sharp. Analytic solutions for bonds and bond options under n correlated stochastic processes. *Institute of Insurance and Pension Research Working Paper, University of Waterloo*, 1993.
- R. R. Chen. A two-factor preference-free model for interest rate sensitive claims equilibrium model of the term structure of interest rates. *Journal of Futures Markets*, 15(3):345–372, 1995.

- Y. Chen, W. Härdle, and V. Spokoiny. Portfolio value at risk based on independent component analysis. *Journal of Computational and Applied Mathematics*, 205(1):594–607, 2007.
- P. F. Christoffersen. *Elements of Financial Risk Management*. Academic Press, 2003.
- T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 1991.
- J. C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.
- Q. Dai and K. J. Singleton. Specification analysis of affine term structure models. *The Journal of Finance*, 55(5):1943–1978, 2000.
- A. Damodaran. Value at Risk. Research and Papers - NYU Stern School of Business. <http://pages.stern.nyu.edu/~adamodar/pdfiles/papers/VAR.pdf>, 2007. [Online; accessed 12-June-2013].
- G. Darbha. Value-at-Risk for fixed income portfolios. A comparison of alternative models. *National Stock Exchange, Mumbai, India*, 2001.
- P. Date and K. Ponomareva. Linear and non-linear filtering in mathematical finance: a review. *IMA Journal of Management Mathematics*, 22(3):195–211, 2011.
- P. Date and C. Wang. Linear Gaussian affine term structure models with unobservable factors: Calibration and yield forecasting. *European Journal of Operational Research*, 195(1):156–166, 2009.
- G. De Rossi. Kalman filtering of consistent forward rate curves: a tool to estimate and model dynamically the term structure. *Journal of Empirical Finance*, 11(2):277–308, 2004.
- D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6(4):379–406, 1996.
- D. Duffie and J. Pan. An overview of Value-at-Risk. *Journal of Derivatives*, 4(3):7–49, 1997.
- L. El-Jahel, W. Perraudin, and P. Sellin. Value-at-Risk for derivatives. *The Journal of Derivatives*, 6(3):7–26, 1999.
- P. Embrechts, S. I. Resnick, and G. Samorodnitsky. Extreme value theory as a risk management tool. *North American Actuarial Journal*, 3(2):30–41, 1999.

- J. Fan and I. Gijbels. *Local Polynomial Modelling and Its Applications*. Chapman & Hall / CRC, 1996.
- V. Fernández. Extreme value theory: Value at Risk and returns dependence around the world. Documentos de Trabajo 161, Centro de Economía Aplicada, Universidad de Chile, 2003.
- A. Feuerverger and A. C. M. Wong. Computation of Value-at-Risk for nonlinear portfolios. *The Journal of Risk*, 3(1):37–55, 2000.
- R. Fiori and S. Iannotti. Scenario based principal component Value-at-Risk: an application to Italian banks' interest rate risk exposure. *Journal of Risk*, 9(3):63–99, 2007.
- M. Fisher, D. Nychka, and D. Zervos. Fitting the term structure of interest rates with smoothing splines. *Board of Governors of the Federal Reserve System (U.S.), Finance and Economics Discussion Series*, 95(1):1–30, 1995.
- R. A. Fisher and L. H. C. Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of The Cambridge Philosophical Society*, 24:180–190, 1928.
- I. Fraga Alves and C. Neves. *International Encyclopedia of Statistical Science - Extreme value distributions (Chapter)*. Springer Berlin Heidelberg, 2011.
- M. S. Gibson and M. Pritsker. Improving grid-based methods for estimating Value at Risk of fixed-income portfolios. *Board of Governors of the Federal Reserve System (U.S.). Finance and Economics Discussion Series*, 25:1–29, 2000.
- J. Gil-Pelaez. Note on the inversion theorem. *Biometrika*, 38(3-4):481–482, 1951.
- M. Gilli and E. Kéllezi. An application of extreme value theory for measuring financial risk. *Computational Economics*, 27(2):207–228, 2006.
- P. Glasserman. *Monte Carlo Methods in Financial Engineering (Stochastic Modelling and Applied Probability)*. Springer, New York, NY, 2003.
- P. Glasserman, P. Heidelberger, and P. Shahabuddin. Portfolio Value-at-Risk with heavy-tailed risk factors. *Mathematical Finance*, 12:239–269, 2000.

- S. Goedecker. Fast radix 2, 3, 4, and 5 kernels for fast Fourier transformations on computers with overlapping multiply-add instructions. *SIAM Journal on Scientific Computing*, 18(6):1605–1611, 1997.
- G. A. Holton. History of Value-at-Risk: 1922-1998. Method and hist of econ thought, EconWPA, 2002.
- J. Hull and A. White. Value-at-Risk when daily changes in market variables are not normally distributed. *Journal of Derivatives*, 5:9–19, 1998.
- J. C. Hull. *Options, futures and other derivatives*. Prentice Hall, 2009.
- F. Jamshidian and Y. Zhu. Scenario simulation: theory and methodology. *Finance and Stochastics*, 1(1):43–67, 1997.
- S. R. Jaschke. The Cornish-Fisher-expansion in the context of Delta-Gamma-normal approximations. *Journal of Risk*, 4(4):1–14, 2001.
- N. Jegadeesh and G. G. Pennacchi. The behavior of interest rates implied by the term structure of Eurodollar futures. *Journal of Money, Credit and Banking*, 28(3):426–446, 1996.
- A. F. Jenkinson. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348):158–171, 1955.
- N. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions*, volume 2. Wiley, 1994.
- P. Jorion. *Value at Risk: the new benchmark for managing financial risk*. McGraw-Hill, 2007.
- R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82:35–45, 1960.
- J. S. Kamdem. Value-at-Risk and Expected Shortfall for quadratic portfolio of securities with mixture of elliptic distributed risk factors. Technical report, Laboratoire de Mathématiques - Université de Reims, 2003.

- A. A. A. Kanto. A note on calculation of CVaR for student's t distribution. Helsinki School of Economics, Working Papers, W-369, May 2004.
- S. Kotz and S. Nadarajah. *Extreme Value Distributions: Theory and Applications*. Imperial College Press, 2000.
- P. H. Kupiec. Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, 3(2):73–84, 1995.
- P. K. Kytte and M. R. Schäferkötter. *Handbook of Computational Methods for Integration*. Chapman & Hall/CRC, 2004.
- P. Lévy. *Calcul de probabilités*. Gauthier-Villars, Paris, 1924.
- H. W. Lilliefors. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62:399–402, 1967.
- R. Litterman and J. Scheinkman. Common factors affecting bond returns. *The Journal of Fixed Income*, 1:54–61, 1991.
- P. A. Lynn and W. Fuerst. *Introductory Digital Signal Processing with Computer Applications, Second Edition*. Wiley, New York, 2000.
- R. G. Lyons. *Understanding Digital Signal Processing - Second Edition*. Prentice Hall, New Jersey, 2004.
- J. H. McCulloch. The tax-adjusted yield curve. *Journal of Finance*, 30(3):811–830, 1975.
- A. J. McNeil. Calculating quantile risk measures for financial return series using extreme value theory. <http://e-collection.ethbib.ethz.ch/eserv/eth:25077/eth-25077-01.pdf>, 1998. [Online; accessed 12-June-2013].
- C. Menn and S. T. Rachev. Calibrated FFT-based density approximations for α -stable distributions. *Computational Statistics & Data Analysis*, 50(8):1891–1904, 2006.
- T. Mikosch. Copulas: tales and facts. *Extremes*, 9:3–20, 2006.
- J. Mina and A. Ulmer. Delta-Gamma four ways. <http://www.texnology.com/risk.pdf>, 1999. [Online; accessed 12-June-2013].

- M. Montero and A. Kohatsu-Higa. Malliavin calculus applied to finance. *Physica A: Statistical Mechanics and its Applications*, 320:548–570, 2003.
- R. B. Nelsen. *An Introduction to Copulas (Springer Series in Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006. ISBN 0387286594.
- E. Ngoussou. Value at Risk using skew elliptical distributions. http://www.affi.asso.fr/uploads/Externe/1f/CTR_FICHER_38_1226070144.pdf, 2004. [Online; accessed 12-June-2013].
- M. Nozari, S. M. Raei, P. Jahangiri, and M. Bahramgiri. A comparison of heavy-tailed VaR estimates and filtered historical simulation: Evidence from emerging markets. *International review of Business Research Papers*, 6(4):347–359, 2010.
- M. Paolella. *Intermediate Probability: A Computational Approach*. Wiley-Interscience, New York, NY, USA, 2007.
- G. M. Phillips and P. J. Taylor. *Theory and applications of numerical analysis*. Academic Press, 1996.
- S. Pichler and K. Selitsch. A comparison of analytical VaR methodologies for portfolios that include options. <http://symposium.fbv.uni-karlsruhe.de/8th/papers/pichler.pdf>, 1999. [Online; accessed 12-June-2013].
- J. Pickands. Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1):119–131, 1975.
- M. Pritsker. Evaluating Value-at-Risk methodologies: accuracy versus computational time. *Journal of Financial Services Research*, 12(2-3):201–242, 1997.
- C. Qian and R. H. Gerlach. The two-sided Weibull distribution and forecasting of financial risk. *OME Working Paper*, pages 1–36, 2011.
- M. Rahman, Z. Jian-Ping, and S. Yi-Hang. An application of copula theory in the estimation of Value at Risk in case of Bangladeshi stock markets. *International Research Journal of Finance & Economics*, 71:118–131, 2011.
- C. R. Rao. *Linear Statistical Inference and its Applications (2nd Edition)*. John Wiley & Sons, 2001.

- R.T. Rockafellar and S. Uryasev. Optimization of conditional Value-at-Risk. *Journal of Risk*, 2:21–41, 2000.
- C. Rouvinez. Going Greek with VaR. *Risk*, 10(2):57–65, 1997.
- A. Rozga and J. Arnerić. Dependence between volatility persistence, kurtosis and degrees of freedom. *Investigacion Operacional*, 30(1):32–39, 2009.
- H. Sak and C. Haksöz. A copula-based simulation model for supply portfolio risk. *The Journal of Operational Risk*, 6(3):15–38, 2011.
- Q. Shao, H. Wang, and H. Yu. A calibrated scenario generation model for heavy-tailed risk factors. *IMA Journal of Management Mathematics*, 17:289–303, 2006.
- A. Z. Sheikh and H. Qiao. Non-normality of market returns: A framework for asset allocation decision-making. *The Journal of Alternative Investments*, 12(3):8–35, 2010.
- B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, 1986.
- A. Sklar. Random variables, joint distribution functions, and copulas. *Kybernetika*, 9(6):449–460, 1973.
- D. J. Smith. A primer on bond portfolio value at risk. *Advances in Financial Education : Journal of the Financial Education Association*, 7(1-2):1–14, 2009.
- D. Sornette, P. Simonetti, and J. V. Andersen. ϕ^q -field theory for portfolio optimization: “fat tails” and nonlinear correlations. *Physics Reports*, 335:19–92, 2000.
- S. V. Stoyanov, S. T. Rachev, B. Racheva-Iotova, and F. J. Fabozzi. Fat-tailed models for risk estimation. *Department of Economics and Business Engineering, Karlsruhe Institute of Technology (KIT). Working Paper Series in Economics, WP024-08*, 30:1–22, 2011.
- G. Szegö. Measures of risk. *Journal of Banking & Finance*, 26(7):1253–1272, 2002.
- D. Takahashi and Y. Kanada. High-performance radix-2, 3 and 5 parallel 1-d complex FFT algorithms for distributed-memory parallel computers. *Journal of Supercomputing*, 15:207–228, 2000.
- N. Taleb. *Dynamic Hedging*. Wiley, New York, 1997.

- C. Temperton. A generalized prime factor FFT algorithm for any $n = 2^p 3^q 5^r$. *SIAM Journal on Scientific and Statistical Computing*, 13(3):676–686, 1992.
- O. Vasicek. An equilibrium characterisation of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.
- O. Vasicek and G. Fong. Term structure modeling using exponential splines. *Journal of Finance*, 37(2):339–348, 1982.
- S. Venkataraman. Value at Risk for a mixture of normal distributions: the use of quasi-Bayesian estimation techniques. *Economic Perspectives*, pages 2–13, 1997.
- R. von Mises. La distribution de la plus grande de n valeurs. *Selected Papers - American Mathematical Society, Providence, RI*, 2:271–294, 1954.
- I. C. Wang. *Dynamic Interest Rate Models and Their Applications*. PhD thesis, Brunel University, 2008.
- J. Wang. Generating daily changes in market variables using a multivariate mixture of normal distributions. <http://www.informs-sim.org/wsc01papers/035.PDF>, 2001. [Online; accessed 12-June-2013].
- J. G. Wendel. The non-absolute convergence of Gil-Pelaez’ inversion integral. *The Annals of Mathematical Statistics*, 32(1):338–339, 1961.
- P. Wiberg. Computation of Value-at-Risk: The fast convolution method, dimension reduction and perturbation theory. <http://www.cs.toronto.edu/pub/reports/na/wiberg.phd.ps.gz>, 2002. [Online; accessed 12-June-2013].
- T. C. Wilson. Infinite wisdom. *Risk*, 6:37–45, 1993.
- T. C. Wilson. Value at risk. *Risk Management and Analysis*, 1:61–124, 1998.
- P. Zangari. An improved methodology for measuring VaR. *RiskMetrics Monitor*, second quarter 1996a.
- P. Zangari. How accurate is the Delta-Gamma methodology? *RiskMetrics Monitor*, third quarter 1996b.

P. Zangari. A VaR methodology for portfolios that include options. RiskMetrics Monitor, first quarter 1996c.