A Review of the Use of Copulas in Credit Derivatives and the Development of Alternative Methodologies[1]

Slavica Lazic

Under the Supervision of
Prof Coenraad Labuschagne and Dr Diane Wilcox

Programme in Advanced Mathematics of Finance
School of Computational and Applied Mathematics
University of the Witwatersrand

Private Bag 3
P O WITS 2050
Johannesburg
South Africa

E-mail: `lazic.slavica@gmail.com`

August 16, 2010

[1]A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the Degree of Master of Science.

Abstract

Credit derivatives and their modelling have received a lot of attention in recent years. Dependence between assets is a crucial property where the contingent payment depends on a basket of underlying assets. Prior to the recent global economic crisis, copulas had earned the reputation of being key tools for capturing this dependence. However, their popularity has been subsequently lost. In this dissertation we will examine the theory surrounding copulas and their usefulness when applied to modelling credit derivatives. First, some general mathematical theory will be presented. Following this introduction, we will look at various copulas that have been suggested for the use in credit derivatives, such as the Gaussian copula, the t-copula and the Archimedean family of copulas. We will discuss the features of these copulas that may make them attractive for modelling credit derivatives. We will then turn our attention to the pitfalls of copulas that may have caused their recent lack of popularity. Finally, we will examine alternative models that have been put forward for capturing this dependence in credit derivatives.

Declaration

I declare that this thesis is my own, unaided work. It is being submitted for the Degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other university.

Slavica Lazic   Date

Acknowledgements

I would like to express my deep gratitude for Prof David Taylor for giving me the opportunity to do my studies, for giving me great insight into my area of research and for ensuring that this experience was fulfilling for me. I would also like to thank my supervisors Dr Nadia Uys, Prof Coenraad Labuschagne and Dr Diane Wilcox for providing me with guidance and constructive feedback and instilling in me confidence in my work. In addition, I am grateful to Stuart Cullender for rescuing the formatting of my dissertation. My parents have been crucial to my success and their commitment to supporting me made my studies possible. I would also like to thank Andrew Bizzell, fellow masters and doctorate students for their support, patience and encouragement.

Contents

Preface

A letter from the author to the reader.

Slavica Lazic
Johannesburg
December 2009

Chapter 1

Introduction

1.1 Introduction

Credit derivatives have become crucial financial instruments in recent times. Their development has been largely driven by the desire of financial institutions to hedge their credit exposures. According to [12] their popularity stems from the fact that they help financial firms to manage the credit risk on their books by dispersing parts of the risk through the wider financial sector, thereby reducing the concentration of risk. Major participants in the market for credit derivatives have been the banks, insurance companies and investment funds.

It is mentioned in [22] that according to a 2002 survey by the British Bankers' Association (BBA), the estimated size of the credit derivative market by the end of that year was already $2 trillion. The survey identified the so-called single-name credit default swaps (CDS) as accounting for almost half of the market volume at that time. These credit-derivatives have been traded over-the-counter.

The market for CDSs written on larger corporations is fairly liquid. They are a natural underlying security for many more complex credit derivatives, and models for pricing portfolio-related credit derivatives are usually calibrated to these quoted CDS spreads.

As financial innovation found its way into the market, credit derivatives became more complex and were increasingly based on entire loan portfolios. As a direct consequence of managing the risks of loan portfolios, a huge market for asset securitisation, such as various types of collateralised debt obligations (CDO), has emerged. With this development the mathematical modelling became increasingly complex. In addition, the purpose to which credit derivatives were put had broadened to include maximising expected returns through speculation - a risky strategy for any type of financial instrument.

According to [22] the roots of credit derivative modelling can be associated with the

Black and Scholes (1973) and Merton (1974) models. Credit risk has its foundations in the development of option-pricing techniques and the application to the study of corporate liabilities. The option-pricing literature, which views the bonds and stocks issued by a firm as contingent claims on the assets of the firm, is the first to provide a strong link between a statistical model describing default and an economic-pricing model

Due to its nature, credit risk is rich with institutional details of the underlying entity. However majority of the available models tend to deal with stylized versions of credit risk. In addition, the pricing of CDOs and the analysis of portfolios of loans or credit-risky securities lead to the question of modelling dependence between the defaults. Dependence modelling is necessary in trying to understand the risk of simultaneous defaults by entities on which the derivative is based. Such multiple defaults can affect the stability of the financial system with profound effects on the entire economy.

The market for credit derivatives had evolved so much that, according to [12], immediately prior to the credit crunch of 2007-2009 CDOs overtook the CDSs and became the most widely traded class of portfolio credit derivatives.

1.2 The Credit Crunch

The recent global financial crisis of 2007-2009 (also known as the credit crunch) has been extensively analysed and there is much literature to be found on it. Generally the literature is consistent in identifying the causes of the crisis. In a nutshell, defaults on home-loans around 2006-2007 led to either foreclosure or sale of houses to settle the mortgages. This sent the property prices declining, which in turn was a catalyst for further foreclosure and forced sales, and lead to the collapse of the United States' housing bubble, which peaked in approximately 2005-2006. The defaults upset the economic stability and damaged mortgage providers and other financial institutions globally.

The first trigger of the crisis was the rising of interest rates in the United States. Loan incentives based on easy initial terms and a long-term trend of rising housing prices had initially encouraged borrowers to assume difficult mortgages in the belief that they would be able to quickly refinance at more favourable terms. However, once interest rates began to rise and housing prices started to drop moderately in 2006-2007, refinancing became more difficult causing a large increase on mortgage defaults by the private sector. High default rates on "subprime" and adjustable rate mortgages (ARM), began to increase quickly thereafter.

Further, it was the decline in confidence, doubts regarding solvency of banks and mortgage institutions, and declines in credit availability that had an impact on global stock markets, which suffered large losses during 2008. Critics argued that credit rating agencies and investors failed to accurately price the risks involved with mortgage-related credit derivatives, and that governments did not adjust their regulatory practices to address the evolution in financial markets.

At the same time as the housing and credit booms, the amount of financial instruments such as the CDOs, which derived their value from mortgage payments and housing prices, greatly increased. As the demand for these instruments increased, more mortgages were issued and in turn more houses were sold. Often these instruments were many times removed from the underlying assets, increasing complexity reducing their transparency at the same time. But as housing prices declined, major global financial institutions that had borrowed and invested heavily in the subprime instruments reported significant losses. Defaults and losses on other loan types also increased significantly as the crisis expanded from the housing market to other parts of the economy. These losses impacted the ability of financial institutions to lend, slowing economic activity. The total losses are estimated in trillions of U.S. dollars globally.

1.3 The Role of Copulas

Modelling dependence between default events and between credit quality changes is, in practice, one of the biggest challenges of credit risk models. The most obvious reason for worrying about dependence is that it affects the distribution of loan portfolio losses and is therefore critical in pricing and valuing instruments such as the CDOs.

Prior to the financial crisis, copulas were one of the main tools for pricing and valuing credit derivatives. They allowed for a great level of flexibility in modelling the default behaviour of individual loans and still provided elegant solutions on a portfolio level.

Their theory has been around since the 1950s, though it is commonly cited that [23] was one of the first works applying copulas to the concept of defaults. In particular it is the Gaussian copula which became a market model for pricing and valuing instruments such as the CDOs.

With the financial crisis, fingers were pointed at the Gaussian copula for underestimating the risk of defaults in portfolio credit derivatives. Subsequently copulas altogether lost their popularity in the modelling of credit derivatives. We'll look at the question of whether this loss of popularity is justified, as well as what are the alternative methods of modelling credit derivatives and what these have to offer.

1.4 Research Objectives

There is a vast body of literature on the use of copulas in credit derivatives. As its first goal, this dissertation sets out to familiarise the reader with the copula theory.

It also undertakes to provide a detailed overview of the application of this theory to credit derivatives, using specific examples found in the literature.

The next task that this dissertation tackles is to provide a comprehensive answer to the question of why the sudden unpopularity of copulas altogether may have come about.

Finally, on criticising any methodology one should be ready to provide alternative solutions. With this in mind, the dissertation aims to provide a balanced review by discussing these alternative methods found in the literature. The author sets out to provide an insight into the evolution of credit derivative modelling and the direction in which it is going.

Since the dissertation reviews a diverse body of literature, practical issues such as simulation of results and computing times are not discussed.

1.5 Structure of the Dissertation

We first look at the various generic characteristics of credit derivatives in Chapter 2. The terminology commonly used in association with credit derivatives, as well as some basic mathematical concepts, is introduced. Some of the main single-name and basket credit derivatives are described. The relevant cashflows and the general approaches to pricing each of these derivatives is discussed.

Chapter 3 is dedicated to the theory of copulas. We follow the layout of [27] and start with some preliminary theory which will assist in defining the copula and its properties. After introducing the copulas some special cases are investigated. These will later assist us in recognising the features of various copulas. We compare various measures of dependence between variables and discuss desirable traits of these measures for our purposes. Lastly, we introduce a number of copula examples that have been used throughout literature for modelling credit derivatives.

Having established the theory of copulas, Chapter 4 sets out detailed examples of how copulas have been used in pricing and valuing credit derivatives on baskets of securities. We investigate the processes assumed for the underlying single security variables, such as time to default. Then we consider how these processes are brought together to a multivariate framework by a copula. The chapter is finished off with an examination of the drawbacks of using copulas for credit derivatives.

In order to provide a balanced review of the use copulas for valuing credit derivatives, we turn our attention in Chapter 5 to the alternate models. The chapter largely follows the chronological developments in credit derivative modelling and draws comparisons between different approaches as well as with the copulas.

Chapter 2

Credit Derivatives, Definitions, Prices

2.1 Introduction

One of the key risks taken on by investors in bonds or loans is *credit risk* - the risk that the bond or loan issuer will default on the debt. To meet the need of investors to hedge this risk, the market has developed instruments called *credit derivatives*. They were originally introduced to protect banks and other institutions against losses arising from default. As such they are instruments designed to off-load or take on credit risk. Credit derivatives have been used widely by banks, portfolio managers and corporate treasurers to enhance returns, trade credit for speculative purposes and as hedging instruments, although the recent credit crisis has created some reluctance in their use.

In simple words, credit risk or *default risk* (we will be using the terms interchangeably) is the risk that an obligor does not honour his/her payment obligations. Default risk has some important properties peculiar to it, which may make its quantitative modelling difficult, namely:

1. default events are rare and may occur unexpectedly; and
2. they may involve significant losses, but the size of these losses may not be known before default.

2.2 Basic Terminology

As suggested above, a credit derivative is a derivative security that is primarily used to transfer, hedge or manage credit risk, and hence whose payoff is materially affected by this risk. When looking at different credit derivatives it is important to consider the following items:

- *Reference credit* and the *reference credit asset*: Reference credit is the entity (or entities) whose defaults trigger the credit events. The set of assets issued by the reference credit is referred to as the reference credit asset. They are needed for the determination of the credit event and for the calculation of the recovery amount. Examples of reference credit assets include: residential property, commercial property; motor vehicles and student loans.

- Definition of *credit event*: A credit event can refer to a number of events, such as bankruptcy; failure to pay a certain obligation; repudiation; a rating downgrade below a certain threshold; or firm restructuring. The affected payments must exceed some pre-specified materiality threshold and the event must prevail over a certain grace period. It is evident that any probability of default on an instrument will depend on the exact definition of the credit event. For this reason it is important to define precisely what will trigger the credit event for the credit derivative in question.

- Assumption of *recovery*: Default may not necessarily result in the value of reference assets dropping to zero. Their recovery value (usually expressed in terms of a percentage of the notional asset value) is often used to determine the payment under the credit derivative.

- *Default payment*: This item refers to the payments made if the default event happens.

Of particular interest to our investigation are derivatives involving more than one security. On a portfolio level there is a risk of *clustering* of defaults, when more defaults occur jointly. We therefore need to refine the definition of credit or default event further in relation to a group of securities. For example the default can be specified to occur when the first security defaults (termed as first-to-default). Other possibilities for the definition of default could be second-to-default, or when the total loss on the basket of securities has exceeded a certain level.

2.3 Different Types of Credit Derivatives

Investment banks may offer their clients tailor-made solutions for their management of credit risk. This gives rise to innovation and creation of many different kinds of credit derivatives. However, the so-called vanilla instruments are still the most widely traded, especially since the recent credit crisis. In this section we will describe some of the common credit derivatives.

2.3.1 Credit Default Swaps

A *Credit Default Swap* (CDS) is an agreement between two counterparties to exchange sets of cashflows. The protection buyer will pay a fee for the protection, either at regular intervals or as a lump sum fee up front. In return, the protection seller will pay the default payment to the protection buyer if a default event happens during the term of the contract. This default payment is structured to replace the loss that the lender would incur upon a specified credit event. If a default does not occur up to the maturity of the default swap, the protection seller makes no payment.

The default payment could be specified in a number of ways, for example:

- Physical delivery of one or several of the reference assets against a repayment at par.
- Notional less post-default market value of the reference asset. This is referred to as cash settlement. The difficulty with this specification is that it may be hard to obtain a market value of an asset after it has defaulted.
- A pre-agreed fixed payoff, irrespective of the recovery rate (this is termed a default digital swap).

The price of the default protection that needs to be paid by the protection buyer to the protection seller is expressed as a rate called the *credit default swap spread*. The fee amount is then the CDS spread multiplied by the notional, often adjusted for the day count fraction, which takes into account conventions such as 360 day counts. On other words the reglar fee payable for the CDS is the notional amount multiplied by the spread. The first fee is usually payable at the end of the first period, and if a default happens between two fee payment dates, the accrued fee up to the time of the default must also be paid to the protection seller.

2.3.2 Collateralised Debt Obligation (CDO)

This is an instrument that securitises a basket of defaultable assets: loans, bonds, mortgages or even credit default swaps (in which case the CDO is referred to as *synthetic* CDO). Leading up to the recent credit crisis, synthetic CDOs became very popular because the issuer could create a portfolio with exposure to defaultable instruments without the requirement of owning them. An illustrative rationale for the development of CDOs can be found in [22].

The main aim of CDOs is for financial institutions to transfer some of their credit risk to investors and to free regulatory capital. The main advantage for investors is the ability to invest in products they would not have access to otherwise or not be permitted to invest in (which may enhance their portfolio diversification).

The assets which are to be securitised are sold to a company called a *Special Purpose Vehicle* (SPV), that is set up for this purpose. The sole purpose of the SPV is to hold the collateral assets and issue securities backed by this collateral into the capital markets. Investors are then offered the opportunity to invest in notes issued by this company.

However, before they are issued, the assets are re-packaged into tranches according to the level of perceived risk. Tranches are ranked by seniority and are defined by the lower and upper attachments points. Senior debt tranche is the least risky for investors. The mezzanine debt tranche has an intermediate risk level (there could be more than one mezzanine tranche level in a CDO). The equity tranche is the most risky for the investors. In return for its higher risk, the equity tranche provides the largest coupons.

Investors with increasing risk aversion select progressively more senior tranches, which provide correspondingly lower risk and returns. The order of payments resulting from the cashflows from the collateral assets is determined by the seniority of the tranches. The investors in the senior tranches have priority over the mezzanine tranche investors, who in turn have priority over the equity tranche holders.

In return for receiving periodic coupon payments (sometimes also referred to as premium payments), the investors will bear losses resulting from defaults of the collateralised assets. A default causes the amount of assets to decrease. Investors in a tranche with attachments level $a\%$ and detachment level $d\%$ will bear all losses in the portfolio in excess of $a\%$ and up to $d\%$ of the initial portfolio value. When tranches are issued they usually receive a credit rating by an independent credit rating agency (e.g. S&P or Moody's). Each tranche except the equity is rated. Prior to the recent credit crunch, the senior tranche used to have a Aaa\AAA rating.

2.3.3 Credit Linked Notes

A Credit Linked Note (CLN) is a security which has an interest payment and a fixed maturity structure similar to a vanilla bond. The performance of the note however, including the maturity value, is linked to the performance of a specified underlying asset or assets as well as that of the issuing entity.

A CLN is a combination of a credit derivative and a coupon bond that is sold as a fixed package. The coupon payments (and sometimes also the repayment of the principal) are reduced if a third party (the reference entity) experiences a default during the lifetime of the contract, so the buyer of a credit-lined note is providing the credit protection for the seller.

2.4 Credit Derivative Pricing

In this section we will be looking into the pricing of CDS and CDO contracts. In each case we will derive the cashflows for the protection buyer and protection seller's legs and then discuss how discounting and a suitable loss distribution would be applied to arrive at a fair price.

First, we will define the notation and discuss some concepts we will be using in this, as well as the following, chapters.

2.4.1 Notation and some Preliminary Concepts

Notation

The following notation will be used throughout this section:

- n is the number of assets included in the collateral portfolio backing a CDO, or the number of reference assets for a CDS.
- N_i is the notional amount of the i^{th} asset in the collateral portfolio.
- The total portfolio value is given by

$$N = \sum_{i=1}^{n} N_i.$$

- R_i is the assumed recovery rate of the i^{th} asset in the collateral portfolio, taking on values between 0 and 1. In the case where we are only dealing with one asset or we are only concerned with the recovery rate of one default event we denote its recovery rate simply by R. We will first assume that all recovery rates are deterministic. However, this assumption will later be relaxed.
- T is the maturity of the contract measured in years from the date of issue.
- $t_0 = 0$ is the contract initiation time.
- $0 < t_1, t_2, \ldots, t_z = T$ denotes the time points at which the CDS premiums are paid.
- τ_i is the default stopping time for the i^{th} asset with marginal default distribution

$$F_i(t) = P[\tau_i \leq t], \; t \geq 0.$$

  If we are only concerned with a single default event, we drop the subscript and denote the default time by τ.

- $S_i(t)$ is the corresponding survival function, i.e.

  $$S_i(t) = P[\tau_i > t] = 1 - F_i(t).$$

  In [23] it is pointed out that risky bond spread curves or asset swap spreads can be used to build so-called credit curves. A credit curve gives all marginal conditional default probabilities over a number of years.

- r is the continuously compounded risk free rate of interest. We are assuming it to be constant, though this assumption can be easily relaxed.

- $B(0,t)$ is the time $t = 0$ price for a zero-coupon risk-free bond with maturity value 1 at time t.

- $\tilde{p}_{(0,T)}$ is the fair spread for the CDS or the CDO tranche. A fair spread is that spread for which the value of the contract to both parties at inception is zero. For a CDO this is the rate applied to the outstanding tranche notional. The spread is dependent on the contract initiation time and its maturity.

- $1_{\tau_i<t}$ is the default indicator at time t for the i^{th} asset in the collateral portfolio. It is a counting process which jumps from 0 to 1 at the time of default of i^{th} asset. Another notation that will be used is X_i. So $X_i = 1_{\tau_i<t}$.

- The default event correlation between instruments i and j is generally given by

  $$\rho_{i,j}(t) = Corr\left[1_{\tau_i<t}, 1_{\tau_j<t}\right].$$

  Correlation between securities is one of key modelling aspects that we are addressing in this dissertation.

Hazard Rate

A frequent approach to defining the default distribution is to specify it via a hazard rate function, as done in [12], [22] and [31]. The hazard rate function gives the instantaneous default probability for a security that has survived until a time t. It is a conditional probability density function of the default time τ at exact time t, and is of the form

$$\lambda(t) = \frac{f(t)}{1 - F(t)} = -\frac{S'(t)}{S(t)},$$

where $S'(t)$ is the derivative of $S(t)$ with respect to t. Then the survival function can be expressed in terms of the hazard rate function

$$S(t) = e^{-\int_0^t \lambda(s)ds}.$$

The default arrival in this case is an inhomogeneous Poisson process, i.e. $\lambda(s)$ is a deterministic function. The density for τ is then

$$f(t) = S(t) \times \lambda(t).$$

A typical assumption is that the hazard rate is a constant, λ, over a certain period, such as a year. In this case the survival time follows an exponential distribution with parameter λ and the default arrival follows a homogeneous Poisson process. The survival probability over a time interval $[x, x+t]$ for $0 \le t \le 1$ can then be expressed as

$$_tp_x = e^{-\int_0^t h(s)ds} = e^{-\lambda t} = (p_x)^t, \tag{2.1}$$

where p_x denotes the survival probability over the time interval $[x, x+1]$.

In this setup modelling the default process is equivalent to modelling a hazard rate function. The advantage of this method is that the hazard rate function model can be easily adapted to more complicated situations, such as where there are several types of default or where we would like to consider stochastic default fluctuations. There are also a lot of similarities between the hazard rate function and the short rate, so many modelling techniques for the short rate processes can be borrowed to model the hazard rate function.

Probability Spaces and Filtrations

We will often consider a probability space $(\Omega, \mathcal{F}, P)$ and a random time to default τ defined on this space. In other words τ will be an $\mathcal{F}$-measurable random variable taking values in $[0, \infty]$.

A *filtration* $(\mathcal{F}_t)$ on $(\Omega, \mathcal{F})$ is an increasing family $\{\mathcal{F}_t : t \ge 0\}$ of sub-σ-algebras of $\mathcal{F} : \mathcal{F}_t \subset \mathcal{F}_s \subset \mathcal{F}$ for $0 \le t \le s < \infty$. For a generic filtration $(\mathcal{F}_t)$ we set $\mathcal{F}_\infty = \sigma\left(\bigcup_{t\ge0} \mathcal{F}_t\right)$. Filtrations are used to model the flow of information in a random system. So $\mathcal{F}_t$ represents the state of knowledge of an observer at time t, and $A \in \mathcal{F}_t$ is taken to mean that at time t the observer is able to determine if the event A has occurred.

Poisson Process

Poisson processes, touched upon in the previous section, are frequently used in modelling credit derivatives, most commonly in specifying the distribution of the number of defaults at a certain time point.

Given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, F)$, a Poisson process M with intensity $\lambda \ge 0$ is the unique process satisfying the following properties:

- $M_0 = 0$;
- M has independent increments; and
- $P\left(M_t - M_s = k\right) = \left(\left(\lambda\left(t-s\right)\right)^k / k!\right) \exp\left(-\lambda\left(t-s\right)\right)$ for $k \in M_0$ and $t > s \ge 0$.

It is interesting to note that the constant hazard rate λ is the parameter of the Poisson process. The construction of the Poisson process and its relationship with the exponentially distributed random waiting times can be found in [22].

Cox Process

Another name for the Cox process is the double stochastic Poisson process. The reason for this that the parameter λ in the Poisson process is now itself a stochastic process $\{\lambda_t : t \geq 0\}$.

Key Simplifying Assumptions

For simplicity, we will assume independence between default dates and interest rates. Empirical evidence shows that this assumption is not realistic, however our focus is on modelling the dependence between the default dates. Similarly, we assume that the recovery rates on the underlying assets are independent from default times and interest rates.

2.4.2 Credit Default Swaps

A CDS consists of two legs, one corresponding to the fee or premium payments and the other to the contingent default payment. The present value of a credit default swap can be viewed as the sum of the present values of its two legs. The par premium is the premium that makes the present value of the two legs equal.

Cashflows Under Each Leg

The CDS premium payments paid by the CDS buyer at some general time t_i, $1 \leq i \leq z$, can be represented by

$$N \times \tilde{p}_{(0,T)} \times (t_i - t_{i-1}).$$

However, if default occurs before time t_i only a fraction of the premium up to the time of default will be made by the CDS buyer. This premium payment at default can be expressed as

$$N \times \tilde{p}_{(0,T)} \times (\tau - t_{i-1}).$$

Looking now at the other leg, we can express the default payment made by the CDS seller by

$$N \times (1 - R).$$

Let us denote the value of the CDS contract at some future time $t_f \geq t_0$ by

$$\phi(t_0, t_f, T).$$

By construction the value of a fair CDS contract at initiation is equal to zero, i.e. $\phi(t_0, t_0, T) = 0$. However as the credit risk of the reference credit changes over time, the value of the CDS contract may increase or decrease.

Hedge-Based CDS Pricing (No-arbitrage pricing)

The principle of hedging or replication-based pricing is that if two portfolios have equivalent future payoffs, then the current value of both of these portfolios must be equal, else an arbitrage (risk-free profit) opportunity exists. This methodology is useful to spot mispricings in the market because it only relies on the payoff comparisons, and the results are robust as they are independent of any assumed pricing model. However, estimating future payoffs may not be a trivial exercise.

Firstly we need some simplifying assumptions about the default payment of a CDS:

- We assume that the default payment takes place at the time of default. The time delay through grace periods, dealer polls, etc. is ignored.
- We assume that the defaultable bond issued by the reference credit is the only deliverable asset of the CDS.
- We assume that the timing of the coupon payments and the CDS premium payments coincides. The bond and CDS market day count conventions are ignored.

The reasoning behind the use of a replicating strategy is that the purpose of the CDS is to hedge the credit risk of a defaultable bond issued by the reference credit. Thus a portfolio of a combined position in a defaultable bond and a CDS (written on the same reference credit that issued the defaultable bond), should trade close to the price of an equivalent default-free bond.

Replicating Strategy With Fixed-Coupon Bonds

This replicating strategy was presented in [33] and [10]. Consider two portfolios that are constructed at $t = 0$ and unwound at $t = T$ or at the time of default τ, whichever comes first.

Portfolio I The following is the composition of the portfolio:

- A long position in one defaultable coupon bond, which pays a coupon of $\tilde{c}$ at $0 < t_1 < t_2, \ldots, t_z = T$ and the principal N at maturity T.

- A long position in one CDS on this defaultable coupon bond, also maturing at T, with a premium $\tilde{p}_{(0,T)} \times N$.
- If the reference entity defaults before T, the portfolio is immediately unwound at the time of default.

Portfolio II The following is the composition of the portfolio:

- A long position in one default-free coupon bond, which pays a coupon of $c = \tilde{c} - \tilde{p}_{(0,T)} \times N$ at $0 < t_1 < t_2, \ldots, t_z = T$ and the principal at maturity T.
- If the reference entity defaults before T, the default-free bond is sold at time of default τ.

The cashflows of the two portfolios are the same if a default does not occur before T. If the payoffs at default are the same then, by the no-arbitrage argument, the initial values of the portfolios will be the same. By equating these two initial values we can calculate the par CDS spread $\tilde{p}_{(0,T)}$.

However the payoffs of the two portfolios at the time of default are not equivalent. In the case of *Portfolio I* we would get the recovery value NR from the defaultable bond and a payment of $N(1 - R)$ from the CDS, making the total payoff equal to the notional value. However, under *Portfolio II* the default-free bond will be sold at the time of default. There is no guarantee that its value will be equal to the notional (par) value at the time of default. The term structure of interest rates is dynamic, so even if the default-free bond was trading at par initially, there is no guarantee that it will trade near par at any other time in the future, except at maturity T. The value of the coupon-bearing bond will also vary in relation to the timing of the coupons. Its value will increase with accrued interest in between the coupon dates and then drop by the coupon amount c on the coupon payment date. For these reasons this replication strategy is an approximate one.

Replicating Strategy with Floating-Coupon Bonds

We will now replace the default-free fixed coupon bond in the previous replication strategy with a default-free floating coupon bond. A default-free floating coupon bond pays a coupon of $N \times L(t_{i-1}, t_i)$ at time t_i and the principal value N at maturity T. The rate $L(t_{i-1}, t_i)$ denotes the LIBOR (London Interbank Offered Rate) or a similar interest rate for the interval $[t_{i-1}, t_i]$. LIBOR represents the interest rate at which banks lend money to each other in Eurocurrency markets. It has been common practice to use it as a floating default free interest rate, since the default risk amongst banks was considered negligible, especially prior to the credit crisis, compared to companies in other sectors.

In order to achieve matching payoffs in survival we also need to use a defaultable bond that pays floating coupons. In order to match the initial value of the default-free bond, the defaultable bond needs to trade at par from the outset. Let us denote its value at time $t = 0$ by $\tilde{D}^{\tilde{c}'}(0,T)$ and the floating-rate coupon at time t_i by $\tilde{c}'_i = N(L(t_{i-1}, t_i) + p^{par})$, where the par spread p^{par} is chosen such that the value of the bond at issue is par.

Portfolio I The following is the composition of the portfolio:

- A long position in one defaultable floating-coupon bond, which pays a coupon of $\tilde{c}'_i$ at $0 < t_1 < t_2, \dots, t_z = T$ and the principal N at maturity T.
- A long position in one CDS on this defaultable coupon bond, also maturing at T, with a spread of $\tilde{p}_{(0,T)}$.
- If the reference credit defaults before T, the portfolio is unwound at time of default, $t = \tau$.

Portfolio II The following is the composition of the portfolio:

- A long position in one default-free floating-coupon bond, which pays a coupon of $N \times L(t_{i-1}, t_i)$ at $0 < t_1 < t_2, \dots, t_z = T$ and the principal N at maturity T.
- If the reference entity defaults before T, the default-free bond is sold at time of default $t = \tau$.

By design, the initial values of these two portfolios are identical. However, similarly to the replicating strategy using fixed-coupon bonds, the cashflows of the two portfolios differ at default by the amount of the accrued interest on the default-free bond, if default occurs between coupon payment dates. If default occurs in the interval $[t_{i-1}, t_i]$, the value of the default-free par bond is

$$D^{c'}(\tau) = N \times (1 + L(t_{i-1}, t_i)(\tau - t_{i-1})),$$

while *Portfolio I* pays out a total of the notional amount N at default, leaving $N \times L(t_{i-1}, t_i)(\tau - t_{i-1})$ as a difference. Adjustments can be made to the notional value of the CDS to compensate for this, though this difference is considered small.

Hence, the initial cashflows have been exactly matched and the default cashflows have been approximately matched. The survival payoffs differ by the difference between the CDS spread $\tilde{p}_{(0,T)}$ and the par spread p^{par}. Thus, in order to avoid arbitrage opportunities, these two payoffs must coincide, leading to

$$\tilde{p}_{(0,T)} = p^{par}.$$

Further details on this approach can be found in [33].

Problems with the Replicating Strategies

Apart from the approximations and differences already discussed, the problem with this methodology is that some of these replication instruments may not be available. Most reference credits only issue fixed-coupon bonds, if they issue any bonds at all. In some cases these bonds will contain call provisions or an equity convertibility option, which makes them unsuitable as CDS replication instruments.

Bond Price Based CDS Pricing

From the market prices of defaultable securities, one can extract the market's assessment of the issuer's default risk. By comparing the prices of an obligor's defaultable asset and a similar default-free asset, one can infer measurements (e.g. probability of default) of the obligor's credit risk. For further details, the reader is referred to [33].

Pricing the CDS from its Fundamental Cashflows

This approach has been discussed in [22]. We need to make some assumptions regarding the default intensity and the recovery rate on the reference asset. We will assume that we can model the default intensity of the underlying reference security as a Cox process with intensity process λ. The recovery rate R will be assumed constant.

In order to simplify the expressions, we will assume that the reference bond has a face value of 1. Its coupon dates are $0 < t_1 < t_2, \ldots, t_z = T$ and the maturity date is T.

The principle of this methodology is to obtain explicit expressions for the present value of the cashflows under each leg under the above assumptions. The par spread $\tilde{p}_{(0,T)}$, by its definition, is then obtained by equating the present values of each leg.

Since the protection buyer pays a premium $\tilde{p}_{(0,T)}$ (the notional amount is 1) until maturity or earlier default, the value for this leg can be expressed as:

$$
\begin{aligned}
V^{pb} &= \mathbb{E}\left[\sum_{i=1}^{T} e^{-r\times i} 1_{\{\tau>i\}} \tilde{p}_{(0,T)}\right] \\
&= \tilde{p}_{(0,T)} \mathbb{E}\left[\sum_{i=1}^{T} \exp(-\int_0^i (r+\lambda_s)ds)\right] \\
&= \tilde{p}_{(0,T)} \sum_{i=1}^{T} \tilde{B}(0,i),
\end{aligned}
$$

where $\tilde{B}(0,i)$ denotes the time $t=0$ price of a risky zero-coupon bond, maturing at time t_i, with zero recovery.

The present value of the protection seller's leg can be expressed as:

$$\begin{aligned}
V^{ps} &= \mathbb{E}\left[e^{-r\tau}1_{\tau\leq T}\left(1-R\right)\right] \\
&= (1-R)\,\mathbb{E}\left[\int_0^T \lambda_t \exp\left(-\int_0^t \left(r+\lambda_s ds\right)\right) dt\right] \\
&= (1-R)\int_0^T e^{-rt}\times\mathbb{E}\left[\lambda_t \exp\left(-\int_0^t \lambda_s ds\right)\right] dt.
\end{aligned}$$

Since we are assuming independence between the default intensity and interest rates, the expression can be further simplified to

$$\begin{aligned}
V^{ps} &= (1-R)\int_0^T e^{-rt}\times\mathbb{E}\left[-\frac{\partial}{\partial t}\exp\left(-\int_0^t \lambda_s ds\right)\right] dt \\
&= \int_0^T B(0,t)\left(-\frac{\partial}{\partial t}S(t)\right) dt \\
&= (1-R)\int_0^T \hat{\lambda}(t)S(t)B(0,t)dt,
\end{aligned}$$

where $\hat{\lambda}$ is the hazard rate of the survival distribution, i.e.

$$S(t)\equiv\mathbb{E}\left[\exp\left(-\int_0^t \lambda_s ds\right)\right]\equiv\exp\left(-\int_0^t \hat{\lambda}_s ds\right).$$

By equating the values V^{pb} and V^{ps} we can solve for the par spread

$$\begin{aligned}
\tilde{p}_{(0,T)} &= \frac{((1-R)\int_0^T \hat{\lambda}(t)S(t)B(0,t)dt}{\sum_{i=1}^T \tilde{B}(0,i)} \\
&= \frac{(1-R)\int_0^T \hat{\lambda}(t)S(t)B(0,t)dt}{\sum_{i=1}^T B(0,i)S(i)}.
\end{aligned}$$

The integral appears since we are considering the settlement exactly at the default date. If instead we define the settlement as taking place on the same days as the swap payments and we let

$$\hat{Q}\left(\tau=i\right)=Q\left(\tau\in(i-1,i]\right)=S(i-1)-S(i),$$

then the previous expression for the CDS premium becomes

$$\tilde{p}_{(0,T)}=\frac{(1-R)\sum_{i=1}^T B(0,i)\hat{Q}\left(\tau=i\right)}{\sum_{i=1}^T B(0,i)S(i)}.$$

2.4.3 Pricing Credit Default Obligations

The following loss functions and the derivation for the par CDO spread can be found in [4].

Loss Functions

We define the default loss for instrument i as

$$L_i(t) = (1 - R_i)N_i 1_{\{\tau_i < t\}},$$

with R_i being the corresponding assumed recovery rate. This expression also represents the default payment made by the CDS seller.
Then, the cumulative loss on the collateral portfolio at time t is given by

$$L(t) = \sum_{i=1}^{n} L_i(t),$$

which is a pure jump process.

In the case of a CDO, we are interested in knowing how the total loss affects a tranche with attachment level $A = a\%N$ and detachment level $D = d\%N$. The seniority of the tranche is defined by the relative location of the tresholds A and D. For example, if $A = 0$, we are considering the equity tranche. Tranche $[A, D]$ suffers a loss at time t if and only if

$$A = a\%N < L(t) \leq d\%N = D,$$

hence the cumulative loss $L_{A,D}(t)$ on a given tranche is

$$L_{A,D}(t) = \begin{cases} 0, & \text{if } L(t) < A, \\ L(t) - A, & \text{if } A \leq L(t) \leq D, \\ D - A, & \text{if } L(t) \geq D. \end{cases}$$

Equivalently, this can be written as

$$L_{A,D}(t) = [L(t) - A]1_{\{L(t) \in [A,D]\}} + [D - A]1_{\{L(t) \in [D,N_T]\}}.$$

Note that the cumulative loss $L_{A,D}(t)$ is a pure jump process.

We will simplify the pricing problem by considering a homogeneous CDO, i.e. where we have the same notional N and recovery rate R for all assets in the collateral portfolio. Furthermore, we assume that the hazard rate λ is constant. We also use a constant correlation matrix where the correlation between all assets is equal, namely:

$$\rho_{i,j} = \begin{cases} \rho, & \text{if } i \neq j \text{ for } 0 < \rho < 1, \\ 1, & \text{if } i = j. \end{cases}$$

Using these assumptions, let us define

$$\mathcal{L}(t) = \frac{L(t)}{N}.$$

The tranche loss then becomes:

$$L_{A,D}(t) = N\left\{[\mathcal{L}(t) - a\%n]1_{\{\mathcal{L}(t) \in [a\%n, d\%n]\}} + [d\%n - a\%n]1_{\{\mathcal{L}(t) \in [d\%n, n]\}}\right\}.$$

CDO Spread

In order to price a CDO tranche, we first need to estimate the present value of the tranche losses triggered by credit events during the lifetime of the tranche. We express the default leg as the expected value of the default payment stream, discounted from the time of default:

$$DL = \mathbb{E}\left[\int_0^T B(0,t)dL_{A,D}(t)\right].$$

The default leg may be discretized as follows:

$$DL = \sum_{i=1}^{z} B(0,\tau_i)\left[L_{A,D}(\tau_i) - L_{A,D}(\tau_{i-1})\right] 1_{\{\tau_i \leq T\}},$$

where τ_i denote the sorted default times during the contract term and $L_{A,D}(\tau_0) = 0$.

The premium leg is the expected present value of the premium payments weighted by the outstanding capital (original tranche amount less accumulated losses at each payment date).

In the discrete case, the premium leg can be written as

$$PL = \mathbb{E}\left[\sum_{j=1}^{z} \tilde{p}_{(0,T)} B(0,t_j)\left[D - A\right] 1_{\{L(t)\in[0,A)\}} + \sum_{j=1}^{z} \tilde{p}_{(0,T)} B(0,t_j)[D - L(t)] 1_{\{L(t)\in[A,D]\}}\right],$$

where $D - A$ denotes the tranche size at inception and $D - L(t)$ is the outstanding tranche notional at time $t \in [0,T]$. Times $t_1,\ldots,t_z$ are premium payment dates.

The formula can also be written as:

$$PL = \mathbb{E}\left[N \sum_{j=1}^{z} \tilde{p}_{(0,T)} B(0,t_j) \min\{\max[d\%n - \mathcal{L}(t_j), 0], d\%n - a\%n\}\right].$$

Note that in the case of no defaults in the collateral pool (or up to a number of defaults such that the accumulated losses are less than A), the discounted premium is weighted by the total notional amount in the tranche. In the case of losses between A and D, the reference notional amount is reduced accordingly, until it is equal to 0 when the cumulative losses exceed the upper threshold D. When the tranche is wiped out, there are no more premium payments.

The par CDO spread is calculated by equating the present values of the two legs. Hence

$$\tilde{p}_{(0,T)} = \frac{\mathbb{E}\left[\int_0^T B(0,t)dL_{A,D}(t)\right]}{\mathbb{E}\left[N\sum_{j=1}^{z} B\left(0,t_i\right)\min\{\max[d\%n - L(t_j), 0], d\%n - a\%n\}\right]}.$$

Loss Distribution

So far we have not specified a loss distribution for different CDO tranches. This distribution needs to reflect the information on loss probability of the whole basket of securities as an entity. We have the same situation arising when valuing a CDS on a basket of securities. It is crucial to model the dependence structure between the securities correctly. This dependence will be the focus of the next two chapters in our review of copulas. Alternative approaches will be reviewed in Chapter 5.

Chapter 3

Properties of Different Copulas

3.1 Introduction

Copulas have been identified as an important tool for capturing dependence between random variables such as those representing underlying securities in a credit derivative payoff. The standard "operational" definition of a copula given in [13] is that a copula is a multivariate distribution defined on the unit cube $[0,1]^n$, where n denotes the number of variables or the dimension, with uniformly distributed marginals. It is a natural definition suggested by Sklar's theorem (which we will look at in this chapter), whereby a copula can be derived from a multivariate distribution function by transforming the univariate margins.

We will start this chapter by looking at the mathematical concepts that are needed in defining copulas. We will then look at the definition and features of copulas, in general. Special attention will be paid to various measures of dependence. Finally, we will discuss some examples of copulas which have been used in literature on credit derivatives, and discuss the features of these copulas which may make them attractive for this purpose.

3.2 Preliminary Theory

The following theory, which will enable us to better define copulas and their properties, is presented in [27].

Definition 3.2.1. *Let $S_1, S_2, \ldots, S_n$ be nonempty subsets of $\mathbb{R}$, and let H be an n-component real function such that $DomH = S_1 \times S_2 \times \cdots \times S_n$. Let $B = [\boldsymbol{a}, \boldsymbol{b}]$ be an n-box all of whose vertices are in $DomH$. Then the H-volume of B is given by*

$$V_H(B) = \sum sgn(\boldsymbol{c})H(\boldsymbol{c}),$$

where the sum is taken over all vertices $\mathbf{c}$ *of* B*, and by* $sgn(\boldsymbol{c})$ *we mean*

$$sgn(c) = \begin{cases} 1, & \text{if } c_k = a_k \text{ for an even number of } k\text{'s}, \\ -1, & \text{if } c_k = a_k \text{ for an odd number of } k\text{'s}. \end{cases}$$

The mechanics of this formula is easier to see in a 2-dimensional example. Let $B = [x_1, x_2] \times [y_1, y_2]$ be a rectangle whose vertices are in $DomH$. Then the H-volume of B is

$$V_H(B) = H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1).$$

The H-volume is equivalently the nth order difference of H on B. This, together with the following two definitions, will later enable us to specify probabilities of multivariate distributions.

Definition 3.2.2. *An* n*-place real function* H *is* n*-increasing if*

$$V_H(B) \geq 0$$

for all n*-boxes* B *whose vertices lie in* $DomH$*.*

Definition 3.2.3. *Let* $DomH = S_1 \times S_2 \times \cdots \times S_n$ *where each* S_k *has a smallest element* a_k*. Then* H *is said to be grounded if* $H(\boldsymbol{t}) = 0$ *for all* $\boldsymbol{t}$ *in* $DomH$ *such that* $t_k = a_k$ *for at least one* k*.*

Next, we will need to define the *margins* of a multivariate function in relation to H:

Definition 3.2.4. *If each* S_k *is nonempty and has a greatest element* b_k*, then the one-dimensional margins of* H *are the functions* H_k *given by* $DomH_k = S_k$ *and*

$$H_k(x) = H(b_1, \ldots, b_{k-1}, x, b_{k+1}, \ldots, b_n)$$

for all x *in* S_k*.*

In a similar manner, higher dimensional margins can be defined by fixing fewer places in H, i.e. by setting fewer function arguments to their largest elements.

Property 3.2.1. *Let* $S_1, S_2, \ldots, S_n$ *be nonempty subsets of* $\mathbb{R}$*, and let* H *be a grounded* n*-increasing function with domain* $S_1, S_2, \ldots, S_n$*. Then* H *is nondecreasing in each argument, i.e. if* $(t_1, \ldots, t_{k-1}, x, t_{k+1}, \ldots, t_n)$ *and* $(t_1, \ldots, t_{k-1}, y, t_{k+1}, \ldots, t_n)$ *are in* $DomH$ *and* $x < y$*, then*

$$H(t_1, \ldots, t_{k-1}, x, t_{k+1}, \ldots, t_n) \leq H(t_1, \ldots, t_{k-1}, y, t_{k+1}, \ldots, t_n).$$

Definition 3.2.5. *Let* F *be a distribution function. Then a quasi-inverse of* F *is any function* $F^{(-1)}$ *with domain* $\mathbb{I}$ *such that*

1. *if t is in RanF, then $F^{(-1)}(t)$ is any number x in $\mathbb{R}$ such that $F(x) = t$, i.e. for all t in RanF,*

$$F(F^{(-1)}(t)) = t;$$

2. *if t is not in RanF, then*

$$F^{(-1)}(t) = \inf\{x|F(x) \geq t\} = \sup\{x|F(x) \leq t\}.$$

If F is strictly increasing, then the above definition leads to the ordinary inverse function, which shall be denoted by F^{-1}.

3.3 Basic Theory of Copulas

We are now ready to define a copula and consider some of its basic properties. We will first define subcopulas as a certain class of grounded n-increasing functions with margins, and then define copulas as subcopulas with domain $\mathbb{I}^n$, in other words where the domain of each margin is $[0, 1]$.

Definition 3.3.1. *An n-dimensional subcopula (or n-subcopula) is a function C' with the following properties:*

1. *$DomC' = S_1 \times S_2 \times \cdots \times S_n$, where each S_k is a subset of $\mathbb{I}^n$ containing 0 and 1.*
2. *C' is grounded and n-increasing.*
3. *C' has (one-dimensional) margins $C'_k, k = 1, 2, \ldots, n$, which satisfy*

$$C'_k(u) = u \text{ for all } u \text{ in } S_k.$$

Definition 3.3.2. *An n-dimensional copula (or n-copula) is an n-subcopula C whose domain is $\mathbb{I}^n$.*

It is useful to note that for a given copula, the margins, as defined by Definition 3.3.2, are also themselves copulas.

We can rephrase the above Definitions 3.2.1 to 3.2.4 for the n-component real function with reference to copulas. An n-copula is a function C from $\mathbb{I}^n$ to $\mathbb{I}$ with the following properties:

1. For every $\mathbf{u}$ in $\mathbb{I}^n$, $C(\mathbf{u}) = 0$ if at least one coordinate of $\mathbf{u}$ is 0.
2. If all coordinates of $\mathbf{u}$ are 1 except u_k, then $C(\mathbf{u}) = u_k$.
3. For every $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{I}^n$ such that $\mathbf{a} \leq \mathbf{b}$,

$$V_C([\mathbf{a}, \mathbf{b}]) \geq 0.$$

We arrive at Sklar's theorem which is central to the theory of copulas and is the foundation of many of the applications, particularly in statistics. It highlights the role that copulas play in the relationship between multivariate distribution functions and their univariate margins.

Theorem 3.3.1. ***Sklar's Theorem.*** *Let H be an n-dimensional distribution function with margins $F_1, F_2, \ldots, F_n$. Then there exists an n-copula C such that for all $\mathbf{x}^n$ in $\mathbb{R}^n$,*

$$H(x_1, x_2, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)).$$

If $F_1, F_2, \ldots, F_n$ are all continuous, then C is unique; otherwise, C is uniquely determined on $RanF_1 \times RanF_2 \times \cdots \times RanF_n$. The converse also holds.

Lemma 3.3.1. *Let H be a joint distribution function with margins F and G. Then there exists a unique subcopula C' such that*

1. *$DomC' = RanF \times RanG$,*
2. *For all x, y in $\mathbb{R}$, $H(x, y) = C'(F(x), G(y))$.*

The proof of this lemma can be found in [27]. In the following corollary to Sklar's theorem, the link between multivariate and one-dimensional distribution functions via copulas becomes clear.

Corollary 3.3.1. *Let H, C, $F_1, F_2, \ldots, F_n$ be as defined in Sklar's Theorem above and let*
$F_1^{(-1)}, F_2^{(-1)}, \ldots, F_n^{(-1)}$ be quasi-inverses of $F_1, F_2, \ldots, F_n$, respectively. Then for any $\mathbf{u}$ in $\mathbb{I}^n$,

$$C(u_1, u_2, \ldots, u_n) = H\left(F_1^{(-1)}(u_1), F_2^{(-1)}(u_2), \ldots, F_n^{(-1)}(u_n)\right).$$

From Sklar's theorem and its corollary we see that, for continuous multivariate distribution functions, the univariate margins and the multivariate dependence structure can be separated, and this dependence structure can be specified by a copula.

Until now we have not specified the nature of the functions being discussed. For the purposes of credit derivatives the functions that we will need to deal with will be distribution functions of random variables (we will later discuss what these random variables usually represent). All of the above definitions and results, including Sklar's theorem, can be equally applied for random variables defined on a common probability space. We formalise this in the following theorem:

Theorem 3.3.2. *Let $X_1, X_2, \ldots, X_n$ be random variables with distribution functions $F_1, F_2, \ldots, F_n$, respectively, and joint distribution function H. Then there exists a copula C such that*

1. *$DomC = F_1 \times F_2 \times \cdots \times F_n$, where each F_k is a subset of $\mathbb{I}$ containing 0 and 1.*

2. *C is grounded and n-increasing.*

3. *C has margins C_k, $k = 1, 2, \ldots, n$ which satisfy*

$$C_k(u) = u, \text{ for all } u \text{ in } F_k.$$

We are now ready to look at some very specific copulas, whose properties will be significant in understanding the limiting behaviours of many other types of copulas. We begin with the *Fréchet-Hoeffding bounds*:

Definition 3.3.3. *Fréchet-Hoeffding Bounds*
If C' is any n-subcopula, then for every $\boldsymbol{u}$ in $DomC'$

$$W(\boldsymbol{u}) \leq C'(\boldsymbol{u}) \leq M(\boldsymbol{u}),$$

where

$$\begin{aligned} M(\boldsymbol{u}) &= \min(u_1, u_2, \ldots, u_n) \\ W(\boldsymbol{u}) &= \max(u_1 + u_2 + \cdots + u_n - n + 1, 0). \end{aligned}$$

Because every copula is a subcopula, the inequality above holds also for copulas. The function M is known as the Fréchet-Hoeffding upper bound and is always itself a copula. The function W is the Fréchet-Hoeffding lower bound. It is not a copula for $n > 2$.

Proof. Here we will present the proof for a two-dimensional case, as set out in [27].

Let (u, v) be an arbitrary point in $DomC'$. Using the marginal distributions and the fact that C' is 2-increasing, we can say $C'(u, v) \leq C'(u, 1) = u$ and $C'(u, v) \leq C'(u, 1) = u$, which combines to give $C'(u, v) \leq \min(u, v)$. Furthermore,

$$V_{C'}([u, 1], [v, 1]) \geq 0$$

implies $C'(u, v) \geq u + v - 1$, which when combined with the requirement that $C'(u, v) \geq 0$, yields $C'(u, v) \geq \max(u + v - 1, 0)$. □

Another type of copula that will prove very useful is the product copula:

$$\Pi^n(\mathbf{u}) = u_1 \times u_2 \times \cdots \times u_n.$$

We will later see that some copulas tend to Fréchet-Hoeffding bounds or the product copula in the limit as their parameters tend to certain values.

We come to the following theorem, which is stated without proof in [27].

Theorem 3.3.3. *For $n \geq 2$, let $X_1, X_2, \ldots, X_n$ be continuous random variables. Then*

1. *$X_1, X_2, \ldots, X_n$ are independent iff the n-copula of $X_1, X_2, \ldots, X_n$ is $\Pi^n(\mathbf{X})$, and*

2. *each of the random variables $X_1, X_2, \ldots, X_n$ is almost surely a strictly increasing function of any of the other random variables iff the n-copula of $X_1, X_2, \ldots, X_n$ is $M(\mathbf{X})$.*

It is worth noting that the first part follows from a well-known fact that $X_1, \ldots, X_n$ are independent if and only if $H(x_1, \ldots, x_n) = F_1(x_1) \times \cdots \times F_n(x_n)$ for all $x_1, \ldots, x_n$ in $\mathbb{R}$.

For $n = 2$, the lower bound W itself is a copula and it can be seen that it is a bivariate distribution function of the random vector $(U, 1-U)$, where U is uniformly distributed on $[0, 1]$. In this case we say that W describes perfect negative dependence. As suggested by the above theorem, M describes perfect positive dependence.

Theorem 3.3.4. *Let C' be a subcopula. Then for every $\boldsymbol{u}$ and $\boldsymbol{v}$ in $DomC'$,*

$$\left|C'(\boldsymbol{v}) - C'(\boldsymbol{u})\right| \leq \sum_{i=1}^{n} |v_i - u_i| .$$

Hence C' is uniformly continuous on its domain.

Next we will look at a concept called concordance ordering - a concept which will be needed later when we define measures of dependence such as Kendall's tau and Spearman's rho.

Definition 3.3.4. ***Concordance ordering***
If C_1 and C_2 are copulas, we say that C_1 is smaller than C_2 and write $C_1 \prec C_2$ if

$$C_1(u_1, u_2, \ldots, u_n) \leq C_2(u_1, u_2, \ldots, u_n)$$

for all $u_1, u_2, \ldots, u_n$ in $\mathbb{I}$.

Using this definition, together with Definition 3.3.3, we can say that W is smaller than every copula and M is larger than every copula.

The theorems and definitions above imply that copulas display invariance under strictly increasing transformations of individual components, keeping the dependence between these components constant. The following theorem formalises this very useful property.

Theorem 3.3.5. ***Strictly Increasing Transformations of Copulas.***
Let $X_1, X_2, \ldots, X_n$ be continuous random variables with copula $C_{X_1,X_2,\ldots,X_n}$. If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are strictly increasing on $RanX_1, RanX_2, \ldots, RanX_n$, respectively, then

$$C_{\alpha_1(X_1),\alpha_2(X_2),\ldots,\alpha_n(X_n)} = C_{X_1,X_2,\ldots,X_n}.$$

Here we will present the proof of this theorem for a 2-dimensional copula, as per [27]. The proof for higher dimensions follows in a straight-forward manner.

Proof. Let F_1 and F_2 denote the distribution functions of X_1 and X_2; and G_1 and G_2 denote the distribution functions of $\alpha_1(X_1)$ and $\alpha_2(X_2)$. Because $\alpha_1(X_1)$ and $\alpha_2(X_2)$ are strictly increasing,

$$G_k(x) = P[\alpha_k(X_k) \leq x] = P[X_k \leq \alpha_k^{-1}(x)] = F_k(\alpha_k^{-1}(x))$$

for any $k = 1, 2$. Thus, for any x_1, x_2 in $\mathbb{R}$

$$\begin{aligned}
& C_{\alpha_1(X_1),\alpha_2(X_2)}(G_1(x_1), G_2(x_2)) \\
= & \; P[\alpha_1(X_1) \leq x_1, \alpha_2(X_2) \leq x_2] \\
= & \; P[X_1 \leq \alpha_1^{-1}(x_1), X_2 \leq \alpha_2^{-1}(x_2)] \\
= & \; C_{X_1,X_2}(F_1(\alpha_1^{-1}(x_1), F_2(\alpha_2^{-1}(x_2)) \\
= & \; C_{X_1,X_2}(G_1(x_1), G_2(x_2)).
\end{aligned}$$

Because X_1 and X_2 are continuous, $RanF_1 = RanF_2$, from which it follows that

$$C_{\alpha_1(X_1),\alpha_2(X_2)} = C_{X_1,X_2}.$$

□

The following theorem, presented without proof, shows the simple transformations of copulas where the functions are not strictly increasing.

Theorem 3.3.6. *Let X and Y be continuous random variables with copula C_{XY}. Let α and β be strictly monotone on $RanX$ and $RanY$, respectively.*

1. *If α is strictly increasing and β is strictly decreasing, then*

$$C_{\alpha(X),\beta(Y)}(u, v) = u - C_{XY}(u, 1 - v).$$

2. *If α is strictly decreasing and β is strictly increasing, then*

$$C_{\alpha(X),\beta(Y)}(u, v) = v - C_{XY}(1 - u, v).$$

3. *If α and β are both strictly decreasing, then*

$$C_{\alpha(X),\beta(Y)}(u, v) = u + v - 1 + C_{XY}(1 - u, 1 - v).$$

Note that in each case the form of the copula is independent of the particular choices of α and β.

For some of the approaches to modelling credit derivatives the random variable of interest will represent the lifetime of an asset being modelled. We will want to look at the survival (as opposed to default) beyond a certain time. We therefore define the survival distribution and joint survival distributions:

$$\bar{F}(x) = P[X > x] = 1 - F(x)$$

and

$$\begin{aligned}\bar{H}(x,y) &= P[X > x, Y > y] \\ &= 1 - F(x) - G(y) + H(x,y) \\ &= \bar{F}(x) + \bar{G}(y) - 1 + C(F(x), G(y)) \\ &= \bar{F}(x) + \bar{G}(y) - 1 + C(1 - \bar{F}(x), 1 - \bar{G}(y)).\end{aligned}$$

From the above the natural definition of the survival copula $\hat{C}$ from $\mathbb{I}^2$ to $\mathbb{I}$ emerges as

$$\hat{C}(u,v) = u + v - 1 + C(1-u, 1-v).$$

This definition can be easily extended to higher dimensions. The survival copula couples the joint survival function to its univariate margins in the same way that a copula connects the joint distribution function to its margins.

3.4 Symmetry

In this section we discuss the symmetry properties of random variable distributions, particularly bivariate distributions, which will later be a tool in formulating some important tail properties of copulas.

If X is a random variable and a is a real number, we say that X is *symmetric* about a if for any x in $\mathbb{R}$, we have

$$P[X - a \leq x] = P[a - X \leq x].$$

In other words $X - a$ and $a - X$ have the same distribution.

When dealing with bivariate distributions there exist different kinds of symmetry, depending on how the two random variables relate to each other. Specifically, let X and Y be random variables and let (a,b) be a point in $\mathbb{R}^2$:

1. (X,Y) is *marginally symmetric* about (a,b) if X and Y are symmetric about a and b, respectively.
2. (X,Y) is *radially symmetric* about (a,b) if the joint distribution function of $X - a$ and $Y - b$ is the same as the joint distribution function of $a - X$ and $b - Y$.

3. (X, Y) is *jointly symmetric* about (a, b) if the following pairs of random variables have a common distribution function: $(X - a, Y - b)$, $(X - a, b - Y)$,$(a - X, Y - b)$ and $(a - X, b - Y)$.

The best known example of a *radially symmetric* distribution is the bivariate normal distribution.

Theorem 3.4.1. *Let X and Y be continuous random variables with joint distribution function H and margins F and G, respectively. Let (a, b) be a point in $\mathbb{R}^2$. Then (X, Y) is radially symmetric about (a, b) if and only if*

$$H(a + x, b + y) = \bar{H}(a - x, b - y) \text{ for all } (x, y) \text{ in } \mathbb{R}^2.$$

Theorem 3.4.2. *Let X and Y be continuous random variables with joint distribution function H, marginal distribution functions F and G, respectively, and copula C. Further, suppose that X and Y are symmetric about a and b, respectively. Then (X, Y) is radially symmetric about (a, b) if and only if $C = \hat{C}$, i.e. if*

$$C(u, v) = u + v - 1 + C(1 - u, 1 - v)$$

for all (u, v) in $\mathbb{I}^2$.

The proof can be found in [27].

A concept related to symmetry is *exchangeability* of random variables. Random variables X and Y are said to be *exchangeable* if the vectors (X, Y) and (Y, X) are identically distributed.

Theorem 3.4.3. *Let X and Y be continuous random variables with joint distribution function H, marginal distribution functions F and G, respectively, and copula C. Then X and Y are exchangeable if and only if $F = G$ and $C(u, v) = C(v, u)$ for all (u, v) in $\mathbb{I}^2$.*

3.5 Dependence

When looking at credit derivatives which are defined with respect to more than one reference credit asset, it is crucial to understand the dependence between these assets. Focal to our interests will be the extreme value or tail properties, as this is where we expect credit events to occur. There are various ways of defining measures of dependence, which will be the focus of this section.

We start with the well-known statistic - the *linear correlation coefficient*:

Definition 3.5.1. *Let X and Y be random variables with nonzero finite variances. The linear correlation coefficient is defined as*

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)} \times \sqrt{Var(Y)}}.$$

The linear correlation coefficient is also known as *Pearson's correlation* coefficient. It is invariant under strictly increasing *linear* transformations:

$$\rho(\alpha X+\beta,\gamma Y+\delta)=sign(\alpha\gamma)\rho(X,Y).$$

The popularity of the linear correlation coefficient stems from the ease with which they can be calculated, and the fact that they are the natural scalar measures of dependence in *elliptical* distributions (e.g. multivariate normal and multivariate t-distribution). However, most random variables are not jointly elliptically distributed. In modelling credit events we would choose to model a scenario using heavy-tailed distributions such as the t_2-distributions, in order to avoid understating the probability of a credit event occurring. In such cases the linear correlation coefficient is not even defined because of infinite second moments. Also, as its name suggests, the linear correlation coefficient can only capture linear dependence between variables.

We therefore turn our attention to two alternative measures of dependence discussed in [13], [23] and [27], which are also suitable for nonelliptical distributions.

Definition 3.5.2. ***Kendall's tau*** *for random variables X and Y is defined as*

$$\tau(X,Y)=P[(X_1-X_2)(Y_1-Y_2)>0]-P[(X_1-X_2)(Y_1-Y_2)<0],$$

where (X_2,Y_2) is an independent copy of (X_1,Y_1).

Kendall's tau is the probability of concordance less the probability of discordance. Note that the observations of random variables are concordant if $(x_i-x_j)(y_i-y_j)>0$ and discordant if $(x_i-x_j)(y_i-y_j)<0$. Informally, a pair of random variables is concordant if "large" values of one variable tend to be associated with "large" values of the other and "small" values of the one variable with "small" values of the other.

If X and Y are continuous random variables with copula C Kendall's tau is given by

$$\tau(X,Y)=4\int\int_{[0,1]^2}C(u,v)dC(u,v)-1.$$

Now, once again, we let X and Y be two random variables with joint distribution function H, margins F and G, and copula C. From these we take three independent sample pairs $(X_1,Y_1),(X_2,Y_2)$ and (X_3,Y_3). Spearman's rho is defined to be proportional to the probability of concordance minus the probability of discordance for the two vectors (X_1,Y_1) and (X_2,Y_3) - i.e. a pair of vectors with the same margins, but one vector has joint distribution function H, while the components of the other are independent.

Definition 3.5.3. ***Spearman's rho*** *for random variables X and Y is defined as*

$$\rho_S(X,Y)=3(P[(X_1-X_2)(Y_1-Y_3)>0]-P[(X_1-X_2)(Y_1-Y_3)<0]),$$

where $(X_1,Y_1),(X_2,Y_2)$ and (X_3,Y_3) are independent copies.

If X and Y are continuous random variables with copula C, Spearman's rho is given by

$$\rho_S(X,Y) = 12 \int\int_{[0,1]^2} C(u,v)dudv - 3.$$

Although both Kendall's tau and Spearman's rho measure the probability of concordance between random variables with a given copula, the values of ρ and τ are often quite different. The following theorem sets the bounds for the relationship between these two measures.

Theorem 3.5.1. *Let X and Y be continuous random variables whose copula is C and let τ and ρ denote, respectively, Kendall's tau and Spearman's rho for this copula. Then*

$$\begin{aligned} \frac{3\tau - 1}{2} \leq \;&\rho\; \leq \frac{1 + 2\tau - \tau^2}{2}\;, \tau \geq 0, \;\textit{and} \\ \frac{\tau^2 + 2\tau - 1}{2} \leq \;&\rho\; \leq \frac{1 + 3\tau}{2}\;, \tau < 0. \end{aligned}$$

These bounds are illustrated and further discussed in [27].

We will now look at general features of a measure of concordance, as laid out in [27]. If X and Y are continuous and are joined by the copula C, then Kendall's tau and Spearman's rho satisfy the properties for a measure of concordance κ with the following properties:

- κ is defined for every pair (X,Y) of continuous random variables.
- $-1 \leq \kappa_{X,Y} \leq 1$, $\kappa_{X,X} = 1$ and $\kappa_{X,-X} = -1$.
- $\kappa_{X,Y} = \kappa_{Y,X}$.
- If C and $\tilde{C}$ are copulas such that $C \prec \tilde{C}$, then $\kappa_C \leq \kappa_{\tilde{C}}$.
- If (X_n, Y_n) is a sequence of continuous random variables with copulas C_n and if C_n converges pointwise to C, then $\lim_{n\to\infty} \kappa_{C_n} = \kappa_C$.

In addition, as for the linear correlation coefficient, κ is invariant under strictly increasing transformations of the random variables.

Copulas with perfect positive concordance measure (i.e. $\kappa_C = 1$) are often referred to as *comonotonic*, while those with perfect negative concordance measure - $\kappa_C = -1$ - are referred to as *countermonotonic*. We have already encountered an example of each: Fréchet-Hoeffding upper bound copula M is comonotonic; while Fréchet-Hoeffding lower bound W is countermonotonic.

We can also look at *tail dependence*, which relates to the amount of dependence in the upper-right-quadrant tail or lower-left-quadrant tail of a bivariate distribution. Tail dependence will be a useful measure when looking at suitability of a particular copula to use in credit derivatives, as credit events are expected to be rare. The amount of tail dependence is invariant under strictly increasing transformations of the random variables.

Definition 3.5.4. *Let X and Y be random continuous variables with marginal distribution functions F and G. The coefficient of upper tail dependence is*

$$d_U = \lim_{u \nearrow 1} P[Y > G^{-1}(u) | X > F^{-1}(u)],$$

provided that the limit $d_U \in [0,1]$ exists. Alternatively,

$$d_U = \lim_{u \nearrow 1} (1 - 2u + C(u,u))/(1-u). \tag{3.1}$$

In a similar manner, *lower tail dependence* can be defined as:

$$d_L = \lim_{u \searrow 0} C(u,u)/u.$$

3.6 Common Copulas in Credit Derivatives

We now turn our attention to specific copulas that have been covered in the literature addressing the modelling of dependence for credit derivatives. In this chapter our focus will not be on the underlying marginal distributions. However, it should be noted that the underlying variables generally represent the values of reference credit assets in relation to some threshold level or the times-to-default, which commonly have exponential distributions with the parameter being generated by a hazard rate function.

3.6.1 Gaussian Copula

Gaussian (or normal) copulas, from the family of elliptical copulas, are the most widely mentioned copulas in literature. In his paper, [23] points out that many default correlation models, such as the CreditMetrics model, implicitly use the Gaussian copula function

$$C_R^{Ga}(\mathbf{u}) = \Phi_n\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_n)\right),$$

where Φ_n has a correlation coefficient matrix $\mathbf{R}$.

In the bivariate case the Gaussian copula can be expressed as

$$C(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-R_{12}^2)^{1/2}} \times \exp\left(-\frac{s^2 - 2R_{12}st + t^2}{2(1-R_{12}^2)}\right) dsdt.$$

The most notable feature of the Gaussian copula is its lack of tail dependence. This is a major draw-back in modelling credit events, since our interest is in an event which we expect to be rare, i.e. in the tails of the distribution. The empirical evidence has shown that security values display lower tail dependence, so that dependencies increase in stressed market conditions.

3.6.2 t - Copulas

The following is a stochastic representation of an n-variate t-distribution with v degrees of freedom; mean μ and covariance matrix $v/(v-1)\Sigma$, can be found in [13]. We let

$$X =_d \mu + \frac{\sqrt{v}}{\sqrt{S}}\mathbf{Z},$$

where $\mu \in R^n$, $S \sim \chi_v^2$ and $\mathbf{Z} \sim N_n(\mathbf{0}, \Sigma)$ are independent.

The copula of $\mathbf{X}$ is

$$C_{v,R}^t(\mathbf{u}) = t_{v,R}^n(t_v^{-1}(u_1), t_v^{-1}(u_2), \ldots, t_v^{-1}(u_n)),$$

where $R_{ij} = \Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$ and where $t_{v,R}^n$ denotes the distribution function of $\sqrt{v}\mathbf{Y}/\sqrt{S}$, for which $S \sim \chi_v^2$ and $\mathbf{Y} \sim N_n(\mathbf{0}, R)$ are independent. The t-copulas have upper and lower tail dependence (the coefficient is the same since this is an elliptical copula), but the coefficient of tail dependence decreases as the degrees of freedom increase. This is because the the t-distribution starts behaving increasingly more like the Gaussian distribution as v is increased.

Both Gaussian and t-copulas are easily parametrised by the linear correlation matrix, but only t-copulas yield dependence structures with tail dependence.

3.6.3 Marshall-Olkin Copulas

Marshall-Olkin copulas are discussed in [13] and [27]. We will discuss the bivariate case in some detail.

We start with two components that are subject to shocks, which are fatal to one or both components. Let X_1 and X_2 denote the lifetimes of the two components. It is assumed that the shocks follow three independent Poisson processes with parameters $\lambda_1, \lambda_2, \lambda_{12} \geq 0$, depending on which component is affected. Then the times Z_1, Z_2 and Z_{12} of occurrence of these shocks are independent, exponential random variables with parameters λ_1, λ_2 and λ_{12}.

The joint survival probability is defined as:

$$\begin{aligned} S(x_1, x_2) &= P(X_1 > x_1, X_2 > x_2) \\ &= P(Z_1 > x_1)P(Z_2 > x_2)P(Z_{12} > \max(x_1, x_2)). \end{aligned}$$

After some algebraic simplification it can be shown that the survival copula for X_1, X_2 has the following form:

$$C_{\alpha_1,\alpha_2}(u_1, u_2) = \min(u_1^{1-\alpha_1} u_2, u_1 u_2^{1-\alpha_2}),$$

where $\alpha_1 = \lambda_{12}/(\lambda_1 + \lambda_{12})$ and $\alpha_2 = \lambda_{12}/(\lambda_2 + \lambda_{12})$.

In the credit derivative context, we can see that the Marshall-Olkin copula would be suitable for use in conjunction with the variables modelling time-to-default.

The various measures of dependence introduced earlier will give an indication of how useful the Marshall-Olkin copulas may be for modelling dependent credit events:

- Spearman's rho:

$$\rho_S(C_{\alpha_1,\alpha_2}) = \frac{3\alpha_1\alpha_2}{2\alpha_1 + 2\alpha_2 - \alpha_1\alpha_2}.$$

- Kendall's tau:

$$\tau(C_{\alpha_1,\alpha_2}) = \frac{\alpha_1\alpha_2}{\alpha_1 + \alpha_2 - \alpha_1\alpha_2}.$$

- Upper tail dependence:

$$d_U = \min(\alpha_1, \alpha_2).$$

It is evident that, by choosing appropriate values for α_1 and α_2, all the values in the interval $[0, 1]$ can be obtained for Spearman's rho and Kendall's tau, and also different levels of upper tail dependence achieved.

A complication for higher dimensional Marshall-Olkin copulas is that a large number of parameters is required. For an n-component system there are in total $2^n - 1$ shock processes, each in one-to-one correspondence with a nonempty subset of $\{1, \ldots, n\}$. It is pointed out in [13] that evaluating Kendall's tau or Spearman's rho rank correlation matrices is easily achieved since the bivariate margins of a Marshall-Olkin n-copula is a Marshall-Olkin 2-copula. However, given a (Kendall's tau or Spearman's rho) rank correlation matrix we cannot in general obtain a unique parametrisation of the copula. By setting the shock intensities for subgroups with more than two elements to zero we can obtain a natural parametrisation of the copula is this situation. However, this then means that the copula is restricted to bivariate dependence.

3.6.4 Archimedean Copulas

Archimedean copulas have been widely discussed in the literature. Unlike the elliptical copulas which are restricted to radial symmetry; these allow for a great variety of dependence structures and can be constructed with ease. In many financial applications (including credit derivative modelling) copulas are needed with a stronger dependence between big losses than between big gains. We will first define a general form of Archimedean copulas and then look at some examples. For a comprehensive list of Archimedean copulas and their features refer to [27].

Definition 3.6.1. *Let φ be a continuous, strictly decreasing function from $[0,1]$ to $[0,\infty]$ such that $\varphi(1) = 0$, and let $\varphi^{(-1)}$ be the pseudo-inverse of φ. Let C be the function from $[0,1]^2$ to $[0,1]$ given by*

$$C(u,v) = \varphi^{(-1)}(\varphi(u) + \varphi(v)). \tag{3.2}$$

Then C is a copula iff φ is convex.

The function φ is called the generator of the copula. If $\varphi(0) = \infty$, we say that φ is a *strict* generator. In this case $\varphi^{(-1)} = \varphi^{-1}$ and C is said to be a *strict* Archimedean copula.

The above definition is naturally extended to n-dimensions as follows:

$$C^n(\mathbf{u}) = \varphi^{(-1)}(\varphi(u_1) + \varphi(u_2) + \cdots + \varphi(u_n)),$$

where the superscript on C denotes dimension.

The following theorem is stated in [27] without proof. It gives elegant simplifications for constructions of Archimedean copulas.

Theorem 3.6.1. *Let C be an Archimedean copula with generator φ. Then:*

1. *C is symmetric; i.e. $C(u,v) = C(v,u)$ for all u, v in $\mathbb{I}$;*
2. *C is associative, i.e. $C(C(u,v),w) = C(u,C(v,w))$ for all u,v,w in $\mathbb{I}$;*
3. *If $c > 0$ is any constant, then $c\varphi$ is also a generator of C.*

The associative property of the Archimedean copulas suggests how higher-dimensional copulas can be constructed from the 2-dimensional ones. However, this method is not always successful. The following theorem gives necessary and sufficient conditions for the above function to be an n-copula.

Theorem 3.6.2. *Let φ be a continuous strictly decreasing function from $[0,1]$ to $[0,\infty]$ such that $\varphi(0) = \infty$ and $\varphi(1) = 0$, and let φ^{-1} denote the inverse of φ. If C^n is a function from $[0,1]^n$ to $[0,1]$ given by 3.2, then C^n is an n-copula for all $n \geq 2$ if and only if φ^{-1} is completely monotone on $[0,\infty)$.*

The following corollary shows that the generators suitable for extensions to arbitrary dimensions of Archimedean 2-copulas correspond to copulas which can model only positive dependence.

Corollary 3.6.1. *If the inverse φ^{-1} of a strict generator φ of an Archimedean copula C is completely monotone, then $C \succ \Pi$, i.e. $C(u,v) \geq uv$ for all u, v in* $[0,1]$.

For an Archimedean copula C with a given generator φ, the partial derivatives with respect to x_i and $x_j, i, j \leq n$, providing marginal distributions, are given by

$$C_{x_i}(x) = \frac{\varphi'(x_i)}{\varphi(C(x))}$$

$$C_{x_i,x_j}(x) = -\varphi(x_i)\varphi(x_j)\frac{\varphi''(C(x))}{(\varphi'(C(x)))^3}.$$

We will now look at some examples of 2-dimensional Archimedean copulas found in [13], which are frequently used in credit derivative modelling.

One-Parameter Gumbel Copula

The generator function is set to $\varphi(u) = (-\ln(u))^\theta$ for $\theta \in [1, \infty)$, which, using formula 3.2, leads to the following copula function:

$$C_\theta(u,v) = \exp\left(-\left[(-\ln(u))^\theta + (-\ln(v))^\theta\right]\right)^{1/\theta}.$$

Setting the parameter θ to 1 gives rise to the product copula with independent variables; while the copula approaches the Fréchet-Hoeffding upper bound as θ tends to infinity (i.e. it displays perfect positive dependence).

Let us look at the tail dependence of the Gumbel copula. Applying the definition of upper tail dependence per equation 3.1 we see that

$$\begin{aligned} d_U &= \lim_{u \nearrow 1}(1 - 2u + C(u,u))/(1-u) \\ &= 2 - \lim_{u \nearrow 1} 2^{1/\theta} u^{2^{1/\theta}-1} \\ &= 2 - 2^{1/\theta}. \end{aligned}$$

Thus for $\theta > 1$, C_θ has upper tail dependence. In a similar manner it can be shown that the Gumbel copula has no lower tail dependence.

It can be shown that Kendall's tau for this copula is

$$\tau(U,V) = 1 - 1/\theta. \tag{3.3}$$

Clayton Copula

The generator function for the Clayton copula is $\varphi(u) = (u^{-\theta} - 1)/\theta$, where $\theta \in [-1, \infty)\backslash \{0\}$. We thus obtain:

$$C_\theta(u, v) = \max\left(\left[u^{-\theta} + v^{-\theta} - 1\right]^{-1/\theta}, 0\right).$$

Setting $\theta = -1$ gives the Fréchet-Hoeffding lower bound (i.e. perfect negative dependence). The product copula is obtained in the limit as $\theta \to 0$, while in the limit $\theta \to \infty$ the copula becomes the Fréchet-Hoeffding upper bound.

Clayton copula has no upper tail dependence. Its coefficient of lower tail dependence is given by $d_L = 2^{-1/\theta}$ so the Clayton copula displays lower tail dependence for $\theta > 0$.

Kendall's tau for the Clayton copula is

$$\tau(U, V) = \theta/(\theta + 2). \tag{3.4}$$

The two-parameter extension of the Clayton copula is usually referred to as the *generalised Clayton copula*, and has the following form in the bivariate case:

$$C_{\theta,\delta}(u, v) = \left(\left((u^{-\theta} - 1)^\delta + (v^{-\theta} - 1)^\delta\right)^{1/\delta} + 1\right)^{-1/\theta}.$$

The generator function for this copula is $\varphi(u) = \theta^{-\delta}(u^{-\theta} - 1)^\delta$, with the requirement that $\theta \geq 0$ and $\delta \geq 1$. Its various dependence measures are given by the following formulae:

$$\tau(U, V) = \frac{(2 + \theta)\delta - 2}{(2 + \theta)\delta}$$

$$d_U = 2 - 2^{1/\delta}$$

$$d_L = 2^{-1/(\theta\delta)}.$$

Frank Family

Here $\varphi(u) = -\ln \frac{e^{-\theta u} - 1}{e^{-\theta} - 1}$ where $\theta \in \mathbb{R}\backslash \{0\}$. So we have:

$$C_\theta(u, v) = -\frac{1}{\theta} \ln\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right).$$

Letting the parameter $\theta \to -\infty$ produces the Fréchet-Hoeffding lower bound; the limit as $\theta \to 0$ gives the product copula and, in the limit $\theta \to \infty$, gives the Fréchet-Hoeffding upper bound. This is also the only family of Archimedean copulas with radial symmetry. It is important to note that the Frank family of copulas does not have tail dependence.

Concluding Remarks of Archimedean Copulas

It is evident that the Archimedean copulas form a very flexible family of copulas. Importantly, one can have only lower or upper tail dependence - a desirable feature since empirical evidence has shown that assets do not display the same level of dependence during market booms as they do during crashes. This family of copulas also has the useful property of being able to be expressed in closed form.

However, for multivariate extensions there is lack of parameter choice. They are also constrained in their dependence structure since all k-margins are identical - they are distribution functions of n exchangeable $U(0,1)$ random variables.

3.6.5 Nested Archimedean Copulas

As mentioned in the previous section, while Archimedean copulas are a very flexible family of copulas in the sense of various possible dependence structures, multivariate extensions of the exchangeable Archimedean copulas still have identical k-margins (where $2 \leq k \leq n$) and hence the dependence structure between these margins is the same. In the real world, securities will be correlated with each other to a different extent, depending on various factors such as geographical location or the industry sector. We would therefore like to be able to capture margins with different levels of dependence, still within the tractable Archimedean copula framework.

By nesting the generators within an Archimedean copula, as put forward by [24], it is possible to create non-exchangeable copulas that allow different degrees of dependence between different margins. As discussed in section introducing Archimedean copulas, and more specifically Corollary 3.6.1, the dependence within different bivariate margins will still be restricted to positive dependence.

A three-dimensional example of a nested Archimedean copula can be constructed using two Archimedean generators, φ_1 and φ_2, as follows

$$C(u_1, u_2, u_3) = \varphi_1\left(\varphi_1^{-1}(u_1) + \varphi_1^{-1} \circ \varphi_2\left(\varphi_2^{-1}(u_2) + \varphi_2^{-1}(u_3)\right)\right). \tag{3.5}$$

Conditions that ensure that this is a copula are that the inverses of generators φ_1 and φ_2 are completely monotonic decreasing functions, and the composition $\varphi_2 \circ \varphi_1^{-1} : [0,\infty] \to [0,\infty]$ is a completely monotonic increasing function. The nesting can be applied arbitrarily deeply. Generalising to a d-dimensional construction for $d \geq 2$, we have

$$C_d\left(u_1, \ldots, u_{d+1}; \varphi_1, \ldots, \varphi_d\right) = \varphi_1\left(\varphi_1^{-1}(u_1) + \varphi_1^{-1}(C_{d-1}\left(u_2, \ldots, u_{d+1}; \varphi_2, \ldots, \varphi_d\right))\right),$$

where $C_1(u_1, u_2; \varphi)$ refers to the bivariate copula defined in equation 3.2, $C_2(u_1, u_2, u_3; \varphi_1, \varphi_2)$ is defined by equation 3.5, and so on.

An additional issue, pointed out in [24], that we need to consider is what generators we can mix while ensuring that the conditions of our function being a copula are met. In the paper the generators discussed are of the same parametric family $\varphi(\cdot;\theta)$ for some parameter values θ_1 and θ_2. For many Archimedean generators, including the Gumbel and Clayton generators discussed in [24], it may be verified that the function $\varphi_1^{-1} \circ \varphi_2$ has completely monotonic derivative if and only if $\theta_2 \geq \theta_1$.

The above condition does have an implication for the kind of dependence structure that a nested Archimedean copula can capture. Suppose (U_1, U_2, U_3) is a random vector with multivariate distribution given by 3.5, where the generator is that for a one-parameter Gumbel or Clayton copula. Then the pairs (U_1, U_2) and (U_1, U_3) have bivariate marginal distributions that are Archimedean copulas with generator $\varphi(\cdot;\theta_1)$ and the pair (U_2, U_3) is distributed according to an Archimedean copula with generator $\varphi(\cdot;\theta_2)$. Referring to equations 3.3 and 3.4 for expressions of Kendall's tau, the requirement that $\theta_1 \leq \theta_2$ means that the pair (U_2, U_3) is more concordant than the pairs (U_1, U_2) and (U_1, U_3).

Chapter 4

Application of Copulas to Credit Derivatives

4.1 Introduction

When applied to credit derivatives such as credit default obligations or credit default swaps, copulas can provide a multivariate distribution of default times or some default trigger values for the portfolio of securities underlying that derivative. We can then use the expected value (conditional on some underlying information set) of such a default indicator, together with some discount mechanism, to either price or value the credit derivative.

Copulas can also be applied in the construction of risk measures by looking at the distribution of losses of a portfolio of securities and reading-off percentiles of interest. Furthermore, they can provide a useful tool for stress testing a variety of portfolios for extreme moves in correlation.

Two general types of default correlation models have been widely suggested in the literature: the reduced form models and the structural models. In brief, the reduced form models assume that the default intensities for different companies follow correlated stochastic processes. Structural models follow on from Merton's (1974) model and assume that a company defaults when the value of its assets falls below a certain trigger level. Default correlation is introduced into a structural model by assuming that the assets of different companies follow correlated stochastic processes. We will also be discussing these models in more detail in Chapter 5 when we look at alternative models for credit derivative, while useful additional literature can be found in [12].

However, both of the above models have been found to be computationally very time consuming for valuing the types of instruments we are considering. This has led market participants to model correlation using a factor copula model where the joint probability distribution for the times to default of many companies is constructed using a common underlying factor.

One of the key issues with the copula framework is the shift in focus from modelling dependency between default events up to a fixed time horizon (i.e. discrete variables) to the dependency between continuous random variables representing defaults, and this dependency not needing to be specified in respect of a specific time horizon.

In the sections that follow we will look at various approaches to modelling underlying single-obligor default processes and build-up on this setup with examinations into modelling dependencies between these obligors using a copula. Our focus will be on methodology, rather than the results obtained in the literature. In order to provide a balanced discussion, this will be followed by an analysis of possible drawbacks of such applications of copulas.

4.2 Some General Assumptions

While empirical evidence shows otherwise, for simplicity we will assume independence between default dates and interest rates. This is due to the fact that the most important task we address is the modelling of dependence between default events themselves. Similarly we generally assume that the recovery rates on the underlying assets are independent of default times and interest rates.

To simplify the expressions we will generally work with homogeneous portfolios, where all default probabilities are identical and where the asset correlation of any two counterparties is the same. It is worth reminding the reader that the copula theory, however, does not fall apart for non-homogeneous portfolios.

Finally, at time $t = 0$ all obligors are assumed to be in a non-default state.

4.3 From Individual to Correlated Default Probabilities - First Steps in Copula Applications

The individual distribution functions $F_i(t)$ of time until default $\tau_i(t)$ were introduced in Chapter 2. Having laid out the theory of copulas in Chapter 3, we are now ready to combine these functions into a single distribution function for defaults of more than one obligor in a portfolio.

The Gaussian copula, starting with its mentions in [23], is the most common example found in the literature, so it is fitting that we also take it as our starting point. As we have seen in Chapter 3, the Gaussian copula can be expressed as

$$F(t_1, t_2, \ldots, t_n) = \Phi_n\left(\Phi^{-1}(F_1(t_1)), \Phi^{-1}(F_2(t_2)), \ldots, \Phi^{-1}(F_n(t_n))\right),$$

where the distribution function of t_i is F_i; Φ_n is the n-dimensional normal cumulative distribution function (for the n obligors) with correlation coefficient matrix Σ.

In order to value credit derivatives written on a portfolio, Monte Carlo simulation can be used to obtain the distribution of default times. For each simulation performed we would have one scenario of the correlated default times $t_1, t_2, \ldots, t_n$, from which we can extract statistics of interest, for example the first-to-default time as $t = \min(t_1, t_2, \ldots, t_n)$.

As a simple illustration of how the Gaussian copula can be applied we follow the example provided in [23], where a two-year contract pays a nominal of 1 currency unit if the first default occurs during the first two years. The hazard rate was assumed to be a constant $\lambda = 0.1$ for $0 < t < \infty$, in other words the default arrival time followed a homogeneous Poisson process (see Chapter 2 for more detail on the Poisson process). The pairwise default correlation was specified between the asset values. If all the credits in the portfolio were independent, the hazard rate of the minimum survival time $T = \min(\tau_1, \tau_2, \ldots, \tau_n)$ would have been

$$\lambda_T = \lambda_1 + \lambda_2 + \cdots + \lambda_n = n\lambda.$$

The survival time of the first-to-default has a density function $f(t) = \lambda_T \times e^{-\lambda_T t}$, so the value of the contract is given by

$$V = \int_0^2 1 \times e^{-rt} f(t)dt = \int_0^2 e^{-rt}\lambda_T \times e^{-\lambda_t t}dt = \frac{\lambda_T}{r + \lambda_T}\left(1 - e^{-2\times(r+\lambda_T)}\right),$$

where r is the risk-free discount rate.

4.4 Modelling Multiple Survival Times

4.4.1 Extension of the Hazard Rate to a Cox Process

In Chapter 2 we touched upon the set-up where the hazard rate for default arrivals follows a non-negative, continuous, adapted stochastic process, i.e. a Cox process. Under a Cox process the default time τ can be expressed either as

$$\tau := \inf\left\{t : \int_0^t \lambda_s ds \geq \theta\right\},$$

where θ is an exponential random variable of parameter 1, independent of the intensity process, or as

$$\tau := \inf\left\{t \geq 0 : \check{N} > 0\right\},$$

where $\check{N}$ is the Cox process and τ is a stopping time with respect to some filtration generated by $\check{N}$.

Under both Poisson and Cox assumptions, modelling a default process is equivalent to modelling the intensity process.

In the next few sections we will be reviewing [21], where the assumption for this setup is that we are given a complete probability space $(\omega, \mathcal{B}, (\mathcal{F}_t), P)$.

4.4.2 Some Background Theory: From Poisson Towards Cox Processes

We draw a sequence (θ_n) of independent exponential random variables of parameter 1 and define T_n as the partial sum of the first n terms of the sequence

$$T_n = \sum_{i=1}^{n} \theta_i.$$

We also define the stochastic Poisson process of parameter 1:

$$M_t := \sum_{n=1}^{\infty} 1_{\{T_n \leq t\}}.$$

Then M is a process with independent and stationary increments and for all t, M_t follows a Poisson distribution of parameter t. The Poisson process M is a counting process, often referred to as a Lévy process, whose jumps are equal to 1.

We then pick a right-continuous with left limits (càdlag) non-decreasing function Λ such that $\Lambda_0 = 0$, $\Lambda_t < \infty$ for all t and $\Lambda_\infty = \infty$, and consider the time-changed Poisson process

$$\bar{M_t} = M_{\Lambda_t}.$$

This new process still has independent increments, but here $\bar{M_t} - \bar{M_s}$ has a Poisson distribution with parameter $\Lambda_t - \Lambda_s$, $(s \leq t)$. This is the *inhomogeneous Poisson process* and Λ is called the *intensity.* The most common assumption for Λ is $\Lambda_t = \int_0^t \lambda_s ds$, where λ_s is referred to as a density.

The Cox process is obtained by letting Λ be stochastic. Formally, conditionally on the knowledge of the intensity - that is the σ-field $\mathcal{F}_\infty^\Lambda = \sigma(\lambda_t, t \geq 0)$ - the Cox Process $\bar{M}$ is an inhomogeneous Poisson process of intensity Λ.

Conditioning on $\mathcal{F}_\infty \vee \mathcal{H}_t$, where $\mathcal{H}_t = \sigma(\tau > s, s \leq t)$ the filtration of the survival process $S_t = 1_{\{\tau > t\}}$, we still have

$$P\left(\tau > T | \mathcal{F}_\infty \vee \mathcal{H}_t\right) = \exp\left(-\int_t^T \lambda_s ds\right).$$

It can be seen that the default process $1_{\{\tau \leq t\}}$ is a Cox process stopped at τ.

4.4.3 Combining Multiple Survival Times Using Copulas

Using the above setting, the default time τ_i of each firm ($1 \leq i \leq n$, where n is the number of firms) can be expressed as

$$\tau_i := \inf \left\{ t : \int_0^t \lambda_s^i ds \geq \theta_i \right\},$$

where λ_s^i is the intensity of τ_i at time s, and θ_i is its threshold.

Multiple default times and the association between them can be introduced in three different ways:

1. The intensity processes λ_s^i of different firms can be correlated. However, it has been pointed out widely in the literature that correlating intensities does not permit us to obtain a high level of dependence between the default times. This is indeed a disadvantage since it has been observed in practice that default times become increasingly more correlated in stressed market conditions.

2. The survival approach: the joint survival function can be specified via a survival copula, as in [23]

   $$S(t_1, \ldots, t_n) = P(\tau_1 > t_1, \ldots, \tau_n > t_n) = \bar{C}_{\tau_1,\ldots,\tau_n}(S_1(t_1), \ldots, S_n(t_n)).$$

   The relevant filtration then is defined by:

   $$\mathcal{G}_t := \mathcal{F}_t \vee \sigma(\tau_i > s, s \leq t, i = 1, \ldots, n) = \mathcal{F}_t \vee \mathcal{H}_t.$$

3. The threshold approach: one can correlate the thresholds θ_i by assuming a specific copula for them, e.g. a survival copula:

   $$\begin{aligned} S(d_1, d_2, \ldots, d_n) &= Pr(\theta_1 > d_1, \theta_2 > d_2, \ldots, \theta_n > d_n) \\ &= \bar{C}_{\theta_1,\theta_2,\ldots,\theta_n}(S_1(d_1), S_2(d_2), \ldots, S_n(d_n)), \end{aligned}$$

   where d_i are the threshold levels (usually constant although they can themselves be stochastic) S_i is the survival function of θ_i and S is their joint survival function.

   Alternatively we can introduce a default countdown process $\gamma_i(t)$ as follows

   $$\gamma_i(t) := \exp\left(-\int_0^t \lambda_s^i ds\right), \tag{4.1}$$

   and define the time to default for firm i as

   $$\tau_i := \inf\left\{t \geq 0 : \gamma_t^i \leq U_i\right\},$$

   and since $\theta_i = -\ln U_i$, the copula of $\mathbf{U}^n = (U_1, \ldots, U_n)$ is the same as the survival copula of $\theta = (\theta_1, \ldots, \theta_n)$.

   The relevant filtration to be considered is again defined by:

   $$\mathcal{G}_t := \mathcal{F}_t \vee \sigma(\tau_i > s, s \leq t, i = 1, \ldots, n) = \mathcal{F}_t \vee \mathcal{H}_t.$$

The relationship between the (survival) copula of default times and that joining the thresholds, proved in [21], is:

$$\bar{C}_{\tau_1,\tau_2,\cdots,\tau_n}(S_1(t_1),\ldots,S_n(t_n)) = \mathbb{E}\left[\bar{C}_{\theta_1,\theta_2,\ldots,\theta_n}\left(\exp\left(-\int_0^{t_1}\lambda_s^1 ds\right),\ldots,\exp\left(-\int_0^{t_n}\lambda_s^n ds\right)\right)\right].$$

We will assume deterministic intensities, under which the survival and threshold approaches coincide:

$$\bar{C}_{\tau_1,\tau_2,\ldots,\tau_n} = \bar{C}_{\theta_1,\theta_2,\ldots,\theta_n}.$$

Let us denote the copula of the default times by C^τ and the copula of the thresholds by C^θ.

In the general case, if $C^\theta = C^\perp$, C^τ is the product copula if and only if the intensity processes are uncorrelated. One of the main differences between the two approaches is that there are two sources of correlation in the threshold approach: correlation between the intensity processes and correlation between the random thresholds.

4.4.4 Application of the Cox Process to Pricing and Risk Monitoring a CDO

In this section we will review the practical example considered in [4]. It involves a synthetic CDO called *EuroStoxx50*, which is composed of 50 single name credit default swaps on credits that belong to the DJ EuroStoxx50 equity index. Each reference credit has a notional equal to 10 million euros, making the total nominal amount of the collateral portfolio 500 million euros. This CDO comprises five tranches with the standard prioritised scheme: super senior tranche notes are paid before mezzanine and subsequent lower subordinations; after which any residual equity tranches are paid (refer to Chapter 2 for more details on CDO structuring).

Application to Pricing

Since one cannot observe a time to default series, one is forced to use alternative proxies to get the desired parameters both for the marginals and the copula itself. Spreads of the single-name credit defaults swaps (CDSs), which can be observed in the market, for each obligor were considered to be best proxies for such a task.

For the joint default behaviour the authors considered the Gaussian, Student t, Clayton and Frank copulas. Although the suitability of the Gaussian copula to default modelling has been questionable and recently proved inappropriate, is a common starting point for copula analysis, beginning with [23]. Student t copula is a natural next step due to its tail dependence and asymptotical behaviour which tends towards the Gaussian copula. Clayton and Frank copulas are from the Archimedean family of copulas, which have recently become popular. For more properties of each of these copulas refer to Chapter 3.

In order to get the simulations for each obligor's time to default, it is assumed that marginal densities follow an exponential distribution for which the hazard term has been derived from the CDS curve at a particular date for each obligor. Each hazard term is assumed as follows:

$$\lambda_t = \frac{\tilde{p}(0,t)}{1-R},$$

where the recovery rate R is fixed to 30%, as is common practice in the market, and $\tilde{p}(0,t)$ represents the CDS spread at a given time t. The intensity process obtained in this case is under the risk neutral measure, while using historical default rates would lead to the process under a historical measure.

As mentioned in [4], the above relationship can be derived by considering the cash-flows exchanged during a CDS and the fact that a fair valuation requires the total amount of payment received by the protection seller to equal the expected loss the seller would pay the buyer if a credit event occurs.

As discussed in Chapter 2, the relationship between the hazard rate and the cumulative default probability is given by

$$F(t) = 1 - \exp\left[-\int_0^t \lambda(s)ds\right]. \tag{4.2}$$

In the real world we cannot deduce the hazard rate for all periods of time, but only for a finite set of times for which we have information on the CDS spread. In the example considered there are three observed points implied by the single name CDS premium at terms 1 year, 3 years and 5 years. In general, we can have N points in time $t_1, t_2, \ldots, t_N$. Hence, the hazard rate function we obtain is a stepwise constant function of time based on the observable values of λ.

Our cumulative distribution function for each obligor's time to default becomes:

$$F(t) = 1 - \exp\left(-\sum_{j=1}^{k} \lambda_j \delta_j\right),$$

where $h_j = h(t_j)$, $\delta_j = t_j - t_{j-1}$ and

$$k = \begin{cases} 1, & \text{if } t \leq t_1 \\ 2, & \text{if } t_1 < t \leq t_2 \\ \vdots & \\ N, & \text{if } t > t_{N-1} \end{cases}$$

This methodology has been used for each reference credit in the portfolio in order to obtain the draws for each time to default, after having generated the draws from the chosen copula.

Application to Risk Monitoring

For risk monitoring, the recovery rate across the portfolio was also changed in [4] into a random variable. It was generated from a Beta distribution, independent from the times to default and interest rates. The Beta distribution has a range $0 \leq x \leq 1$ and shape parameters $v > 0$, $w > 0$. Its density is given by:

$$\frac{x^{v-1}(1-x)^{w-1}}{B(v,w)},$$

where $B(v,w) = \int_0^1 u^{v-1}(1-u)^{w-1}du$. It is well-known that its mean is equal to $v/(v+w)$ and variance equal to $vw/[v+w)^2(v+w+1)]$. The parameters v, w were chosen by fixing the mean equal to 50% and variance equal to 30% (this data was reported in a Moody's study on recoveries from defaulted corporate bonds).

Various levels of VaR (value at risk) were computed for each tranche's cumulative loss and the results from different copulas were compared. The paper found that more losses were simulated from the Student t copula than the Gaussian or the two Archimedean copulas. The result is not surprising, especially when one notes that the chosen degrees of freedom in [4] was 8 for the Student t copula. The lower the degrees of freedom, the higher the tail dependence, increasing the probability of multiple defaults. The Gaussian and Frank copulas do not have tail dependence. It would be of interest to extend this investigation to the conditional tail expectations, loss amounts or number of defaults.

4.4.5 Application of Modelling Survival Times to a Simple First-to-Default Product

In this section we review the example considered by [21], namely a product that pays one unit of currency is case of a default, where default is defined as first-to-default of a portfolio of n assets. The stopping time is thus $\tau = \bigwedge_{i=1}^{n} \tau_i$.

The default intensities were simulated with $\lambda_t^i := (W_t^i)^2$ where $W^n = (W_1, \cdots, W_n)$ is a vector of n correlated $(\mathcal{F}_t)$ Brownian motions with a correlation matrix denoted by ρ^W.

In the case where the payoff happens at maturity T, we can write the expression for the price at time t, where $t < T$ and $t < \tau$, as:

$$\mathbb{E}\left[\exp\left(-\int_t^T r_s ds\right) 1_{\{\tau \leq T\}} | \mathcal{G}_t\right] = B(t,T) - \mathbb{E}\left[\exp\left(-\int_t^T r_s ds\right) 1_{\{\tau > T\}} | \mathcal{G}_t\right],$$

where $B(t,T)$ is the default-free zero-coupon price of maturity T at time t so that $B(t,T) = \exp\left(-\int_t^T r_s ds\right)$. In the case of the survival approach it is not possible to get simpler expressions. However, for the threshold approach one can obtain the following expression for the above expectation term on the right:

$$\begin{aligned} &\mathbb{E} \left[\exp\left(-\int_t^T r_s ds \right) 1_{\{\tau > T\}} | \mathcal{G}_t \right] \\ &= \mathbb{E} \left[\exp\left(-\int_t^T r_s ds \right) \frac{\hat{C}^\theta \left(\exp(-\int_0^T \lambda_s^1 ds), \dots, \exp(-\int_0^T \lambda_s^n ds) \right)}{\hat{C}^\theta \left(\exp(-\int_0^t \lambda_s^1 ds), \dots, \exp(-\int_0^t \lambda_s^n ds) \right)} | \mathcal{F}_t \right]. \end{aligned}$$

Here the copula considered by [21] is the Gaussian copula.

In case of a payoff at default, the threshold model requires an expression for conditional density of τ given the filtration $\mathcal{F}_\infty \vee \mathcal{H}_t$. First lets define

$$\xi_t(dv) = -\frac{\partial}{\partial v} P\left(\tau > v | \mathcal{F}_\infty \vee \mathcal{H}_t\right) dv = -\frac{\partial}{\partial v} \frac{\hat{C}^\theta \left(\exp(-\int_0^v \lambda_s^1 ds), \cdots, \exp(-\int_0^v \lambda_s^n ds) \right)}{\hat{C}^\theta \left(\exp(-\int_0^t \lambda_s^1 ds), \cdots, \exp(-\int_0^t \lambda_s^n ds) \right)} dv.$$

By substituting the above expression into the expectations above we can see that, on $\{\tau > t\}$

$$\begin{aligned} \mathbb{E} \left[\exp\left(-\int_t^\tau r_s ds \right) 1_{\{\tau \le T\}} | \mathcal{G}_t \right] &= \mathbb{E} \left[\mathbb{E} \left[\exp\left(-\int_t^\tau r_s ds \right) 1_{\{\tau \le T\}} | \mathcal{F}_\infty \vee \mathcal{H}_t \right] | \mathcal{G}_t \right] \\ &= \mathbb{E} \left[\int_t^T \xi_t(dv) \exp\left(-\int_t^v r_s ds \right) | \mathcal{F}_t \right]. \end{aligned}$$

4.5 One Factor Models

One factor models are very widely used in the literature. Examples can be found in [12], [23] and [31] among others. In these models the specification of direct relationships between each pair of securities can be avoided. This makes their approach simple for dealing with a large number of names and leads to very tractable pricing results.

We define a variable x_j, $1 \le j \le n$, by

$$x_j = a_j E + \sqrt{1 - a_j^2} Z_j, \tag{4.3}$$

where $a_j^2 \le 1$, and E and Z_j's have independent probability distributions with mean 0 and variance 1. The variable x_j can be thought of as a default indicator variable for the j^{th} obligor. The dependence between obligors is captured through the common factor E, which is the same for all j, while Z_j is an idiosyncratic component affecting only x_j. For simplicity, it is usually assumed that all Z_j's have the same probability distribution. We can see that the correlation between x_i and x_j is $a_i a_j$. From this relationship it follows that conditional on knowledge of E, the x_j's are independent.

We shall denote the cumulative probability distribution of x_j by F_j, which is often assumed to be normal.

Now, let us suppose that t_j is the random variable representing time to default of the j^{th} obligor and that its cumulative probability distribution is Q_j. It is generally assumed that this distribution function is known. The copula model maps x_j to t_j on a "percentile to percentile" basis. The point $t_j = t$ is mapped to $x_j = x$, where

$$x = F_j^{-1}\left[Q_j(t)\right].$$

What the copula model here achieves is to map the variables of interest (i.e. time to default) into other more manageable variables and then defines a correlation structure between those variables.

From equation 4.3 we have

$$Prob\left(x_j < x|E\right) = G_j\left[\frac{x - a_j E}{\sqrt{1 - a_j}}\right],$$

where G_j is the cumulative probability distribution of Z_j. It follows that

$$Q_j\left(t|E\right) = Prob\left(t_j < t|E\right) = G_j\left\{\frac{F_j^{-1}\left[Q_j(t)\right] - a_j E}{\sqrt{1 - a_j^2}}\right\} \tag{4.4}$$

and the conditional probability that the obligor will survive beyond time t is

$$S_j\left(t|E\right) = 1 - G_j\left\{\frac{F_j^{-1}\left[Q_j(t)\right] - a_j E}{\sqrt{1 - a_j^2}}\right\}.$$

The variable E contains the information about the default environment for the time frame of the model. The realisation of E happens at time zero and governs the default outcomes for the whole duration until expiry of the contract. This makes the default environment constant for the whole life of the model - it does not evolve with realisation of new information. This implies that there is no stochastic evolution for hazard rates or CDS spreads in the model. Once E has been determined the cumulative probability of default Q_j is a known function of time.

When this model is used to value securities such as a CDO tranche we set up a procedure to calculate expected cash flows on the tranche conditional on E and then integrate over E to obtain the unconditional expected cashflows.

4.5.1 Gaussian One Factor Models

If E and Z_j have standard normal distributions then the resulting copula is the Gaussian copula, as set out in [23]. In this case x_j also has a standard normal distribution so that $G_j = F_j = \Phi$ for all j, where Φ is the cumulative standard normal distribution function.

The standard market model until recently was the one-factor Gaussian copula model with constant pairwise correlations, constant CDO spreads, and constant default intensities for all companies in the reference portfolio. A single recovery rate of 40% was usually assumed. This made the calculations straightforward as the probability of k or more defaults by time T conditional on the value of the factor E could be calculated form a binomial distribution.

4.5.2 Alternative One Factor Models

While the Gaussian model is most commonly discussed in literature, empirical evidence has proved that this is not the most suitable model to use for credit derivative modelling. We will now look at how the one factor framework can be applied to other copulas which may be more fitting for our purposes. We review the alternatives presented in [3].

Stochastic Correlation

We use the following specification of the latent variables as a simple way of introducing stochastic correlation:

$$x_i = B_i\left(\rho E + \sqrt{1-\rho^2} Z_i\right) + (1 - B_i)\left(\beta E + \sqrt{1-\beta^2} Z_i\right),$$

where B_i are Bernoulli random variables which assign probabilities to the two possible states for each latent variable - one corresponding to low correlation and the other to high correlation. Parameters ρ and β are the correlation parameters, with $0 \leq \beta \leq \rho \leq 1$. We have factor exposure ρ with probability p and β with probability $1 - p$. Again, E and Z_i are independent Gaussian random variables.

As before, we define the default times as $\tau_i = Q_i^{-1}(\Phi(x_i))$ for $i = 1, \ldots, n$. The default times are independent conditionally on E and can be written as

$$Prob(t_i < t|E) = p\Phi\left(\frac{-\rho E + \Phi^{-1}(Q_i(t))}{\sqrt{1-\rho^2}}\right) + (1-p)\Phi\left(\frac{-\beta E + \Phi^{-1}(Q_i(t))}{\sqrt{1-\beta^2}}\right).$$

As one may suspect, the two state model can be generalised to allow for fully stochastic correlations as follows:

$$x_i = \tilde{\rho}_i E + \sqrt{1-\tilde{\rho}_i^2} Z_i,$$

where $\tilde{p}_1, \ldots, \tilde{p}_n$ are independent stochastic correlations.

Student t Copula

In the Student t approach, our vector of interest $(x_1, \ldots, x_n)$ follows a Student t distribution with v degrees of freedom, where $x_i = \sqrt{W} X_i$, and $X_i = \rho E + \sqrt{1-\rho^2} Z_i$ is the already-studied one factor Gaussian copula model. W is independent from $(X_1, \ldots, X_n)$ and follows an inverse Gamma distribution with parameters equal to $\frac{v}{2}$ (or equivalently $\frac{v}{W}$ follows a χ_v^2 distribution). The degrees of freedom v is a parameter that would be determined using some goodness-of-fit test (see [4] for an example).

If we denote the distribution function of the standard univariate Student t by t_v, the time until default for obligor i can be expressed as $\tau_i = Q_i^{-1}(t_v(x_i))$. Conditionally on (E, W) default times are independent and

$$Prob(t_i < t | E, W) = \Phi\left(\frac{-\rho E + W^{-1/2} t_v^{-1}(Q_i(t))}{\sqrt{1-\rho^2}}\right).$$

What we have is a two factor model.

Double t Copula

This model is based on the same idea as the Gaussian one-factor copula model, except that the common and idiosyncratic factors have t distributions. This captures the property of heavier tails of distributions of asset values. It is not necessary for the two t distributions to have the same degrees of freedom.

The latent variables have the following form

$$x_i = \rho \left(\frac{v-2}{v}\right)^{1/2} E + \sqrt{1-\rho^2} \left(\frac{\bar{v}-2}{\bar{v}}\right)^{1/2} Z_i,$$

where E and Z_i are independent random variables following Student t distributions with v and $\bar{v}$ degrees of freedom respectively, and $0 \leq \rho \leq 1$. As the Student distribution is not stable under convolution, the x_i's do not follow Student distributions - the corresponding copula is not a Student copula.

The default times are then given by: $\tau_i = Q_i^{-1}(F_i(x_i))$ for $i = 1, \ldots, n$ where F_i is the distribution function of x_i. Then we have

$$Prob(t_i < t | E) = t_v\left(\left(\frac{\bar{v}}{\bar{v}-2}\right)^{1/2} \frac{F_i^{-1}(Q_i(t)) - \rho \left(\frac{v-2}{v}\right)^{1/2} E}{\sqrt{1-\rho^2}}\right).$$

Clayton Copula

In this model, the random variable denoting the common factor, E, follows a standard Gamma distribution with shape parameter $1/\theta$ where $\theta > 0$. Laplace transform, Ψ, is used to transform its probability density function. As the result of this transform we obtain $\psi(s) = \int_0^\infty f(x)e^{-sx}dx = (1+s)^{-1/\theta}$.

The latent variables x_i's are defined as:

$$x_i = \psi\left(-\frac{\ln U_i}{E}\right), \tag{4.5}$$

where $U_1, \ldots, U_n$ are independent uniform random variables, which are also independent of E. As in the previous examples, the relationship between the default times and corresponding latent variables is

$$\tau_i = Q_i^{-1}(x_i). \tag{4.6}$$

We have obtained a one factor representation with E as the common factor. Applying equations 4.5 and 4.6 together with the result of the Laplace transform we can see that the conditional default probabilities can be expressed as:

$$Prob(t_i < t|E) = \exp\Big(E \times (1 - Q_i(t)^{-\theta})\Big).$$

Low levels of the latent variable are associated with shorter survival times. For this reason, E is also referred to as a "frailty" factor.

In [3] it is highlighted that the joint distribution function of the x_i's is indeed the Clayton copula with generator $\varphi(t) = t^{-\theta} - 1$ and copula form

$$C_\theta(u_1, \ldots, u_n) = \varphi^{-1}(\varphi(u_1) + \cdots + \varphi(u_n))$$

for any set of values $(u_1, \ldots, u_n) \in [0,1]^n$.

4.5.3 Extension to Many Factors

Although it is not often done in practice, the one factor model can be extended to incorporate many common factors. The equation 4.3 becomes

$$x_i = a_{i1}E_1 + a_{i2}E_2 + \cdots + a_{im}E_m + Z_i\sqrt{1 - a_{i1}^2 - a_{i2}^2 - \cdots - a_{im}^2},$$

where $a_{i1}^2 + a_{i2}^2 + \cdots + a_{im}^2 < 1$ and the E_j, $1 \le j \le m$, have independent distributions with zero mean and unit variance. The correlation between x_i and x_j is then $a_{i1}a_{j1} + a_{i2}a_{j2} + \cdots + a_{im}a_{jm}$. Equation 4.4 similarly becomes

$$Q_i(t|E_1, E_2, \cdots, E_m) = G_i\left\{\frac{F_i^{-1}[Q_i(t)] - a_{i1}E_1 - a_{i2}E_2 - \cdots - a_{im}E_m}{\sqrt{1 - a_{i1}^2 - a_{i2}^2 - \cdots - a_{im}^2}}\right\}.$$

4.5.4 Comparison of the One Factor Model Results to Market Data of a k^{th} to Default CDS

We now look at a specific application found in [19] which considers a 5-year k^{th} to default CDS on a basket of $n = 10$ reference entities in the situation where the expected recovery rate, R, is 40%. The default probabilities are generated by Poisson processes with constant default intensities, λ_i $(1 \leq i \leq 10)$, so that the survival probability for each entity is as per equation 2.1 of Chapter 2. In the base case $\lambda_i = 0.01$ for all i and the correlation between all pairs of entities is 0.3.

The results of the various copulas were compared to the market prices as at 4 August 2004 of the CDS indices listed on the Dow Jones. In the base case a Gaussian copula was used (i.e. E and Z_i having standard normal distributions). The results were compared against the following variations:

1. E has a t-distribution with 5 degrees of freedom and the Z_i's are normal;
2. E is normally distributed and the Z_i's have a t-distribution with 5 degrees of freedom;
3. both E and the Z_i's have t-distributions with 5 degrees of freedom.

The last case of the double t copula was found to fit the market prices reasonably well. However, one must bear in mind that the market conditions prevalent at the time of the study were very different to the current market conditions or the conditions at the time of the recent financial crisis. It would be interesting to see what kind of results would have been obtained under these more recent conditions.

4.5.5 Comparison of the CDS and CDO Pricing Using One Factor Models

In this section we review the examples found in [3]. The examples were artificial in the sense that all the reference credits in the portfolio had the same single-name CDS premiums and the same deterministic recovery rate. We do not focus on the actual premium rates obtained, but rather the findings in the paper.

At this point it is worth noting that, as in most of the literature, a restriction is imposed to the cases where a copula is a symmetric function with respect to its coordinates (i.e. correlations between any two reference assets are the same).

Investigation of CDS Premiums

The first example in [3] compared the first-to-default and k^{th}-to-default CDS premiums for Gaussian, Student t and Clayton copulas, as well as the Marshall-Olkin copula. In the first part of the exercise the dependence parameters were set such to obtain the same premium across models for $n = 25$ names; while in the second case they were calibrated to ensure that it is the first-to-default premiums which are equal for each model.

It was found that in the premiums obtained in both exercised are quite similar for the Gaussian, Student t and Clayton copulas, while the Marshall-Olkin copula produced materially different premiums. It would be interesting to investigate the extent to which changing the degrees of freedom for the t copula would have affected these results. In addition, given the departure from rest of the results, some investigation may be warranted into whether it is the Marshall-Olkin copula which produced more reliable results. Comparing these results against actual market data may have been able to answer these questions. Intrestingly, the distribution of the default times was not discussed in the paper.

Investigation of CDO Premiums

In this part of the investigation by [3], in order to compare the different pricing models the dependence parameters were set such to obtain the same equity tranche premiums. Once the equity tranches were matched, the premiums for the mezzanine tranches were computed and compared across models. This procedure effectively compares the tail behaviour of different copulas.

The Clayton and Student t copulas provided results that were in line with that of the Gaussian model. As for the CDS premiums, the premiums computed under the Marshall-Olkin copula were found to be fairly different, except for the extreme cases of independence and comonotonicity. The double t model lies between these two extremes, i.e. the Gaussian and Marshall-Olkin copulas.

The paper discusses the various investigations done to understand the underlying driver for these differences. The two measures of dependence - Kendall's tau and tail dependence coefficients - were not able to provide an explanation for the similarities and differences. In fact, for the tail dependence coefficients it was discovered that the probability of default over the CDO's lifetime did not actually lie in the tail, making tail dependence irrelevant. However, it was found that the bivariate default probabilities, $Q(\tau_i \leq T, \tau_j \leq T)$ for $i \neq j$, are very close for the Gaussian, Clayton and Student t copulas. They were stronger for the double t models and even stronger for the Marshall-Olkin copula.

The different models were also tested for their ability to produce a smile on the pricing tranches on the iTraxx CDO index, as can be observed in the market. The parameters for each model were calibrated on the market quote for the $[0 - 3\%]$ equity tranche.

The Gaussian, Clayton and Student t copulas could not create the correlation smiles implied by the index tranches. Marshall-Olkin copula created a smile, but the prices for the mezzanine tranches were underestimated, while those for the senior tranches were overestimated. The stochastic correlation model provided a reasonable fit, though it overestimated the mezzanine tranche premiums.

Concluding Remarks of the Pricing Exercise

As a conclusion, [3] found that any two models associated with the same distributions of conditional joint default probabilities would lead to the same joint distribution of default indicators and eventually to the same CDO premiums. The closeness of default probabilities for the Gaussian, Clayton and Student t copulas would imply that Clayton and Student t copulas would not provide a material improvement over the Gaussian copula. This is a startling statement given the common argument that the main failure of the Gaussian copula is the lack of tail dependence.

4.5.6 Implied Copula Approach for One Factor Models

It is possible to formulate the one factor model in terms of the conditional hazard rates, instead of the conditional times to default, as set out in [18]. For this approach we use the relationship between hazard rates and default times as specified in the Cox process.

Per [18], we define $h_j(t|E)$ to be the hazard rate at time t for company j conditional on the common background factor E. From equation 4.2 the relationship between $h_j(t|E)$ and $Q_j(t|E)$ is

$$Q_j(t|E) = 1 - \exp\left[-\int_0^t h_j(\tau|E)d\tau\right],$$

or equivalently

$$h_j(t|E) = \frac{dQ_j(t|E)/dt}{1 - Q_j(t|E)}.$$

This equation can be applied together with equation 4.4 to calculate the hazard rate as a function of time for alternative values of E. This conditional function of time is generally referred to as the hazard rate path.

Hence, the one-factor copula model can imply a set of hazard rate paths. The probability of each path occurring is determined by the probability distribution of E. A problem found with the Gaussian copula is that it models the uncertainty about the future hazard rate decreasing with time.

Each hazard rate path represents a future credit environment (FCE), and the set of all possible hazard rate paths then forms the distribution of future credit environments (DFCE).

As already discussed, a default model is often specified starting with the hazard rate and its distribution. In [18] this specification of DFCE is taken as the starting point. It is argued that the copula does not even need to be specified - it is implied by the specification of the DFCE.

However, [18] acknowledges that there are potentially many different distributions of future credit events that are consistent with the observed market data. Therefore calibrating the model to the market data would lead to a degree of uncertainty about the value of any non-standard contract, since it would be unclear which of the possible implied copula models is the "correct" one.

4.5.7 Modelling Correlation Between the Default Rate and Recovery Rate Using One Factor Models

Empirical evidence has shown that there is a strong negative correlation between the default rate and the recovery at default. However, for simplicity this feature has not generally been captured in modelling and, instead, the recovery rate is usually set to a pre-specified constant (as already discussed, market participants tend to use 30% to 40%).

However, [19] had applied the one factor model to incorporate the probability distribution of the recovery rate. This recovery rate is automatically correlated with the default rates through the factor model. Again, the common background factor is E, for which the lower values indicate earlier defaults. The recovery rate R is set to be positively dependent on E. In other words, higher values of E, which indicate longer pre-default times, lead to higher recovery rates and so lower values of loss-given-default.

A random variable x_R is defined by:

$$x_R = a_R E + \sqrt{1 - a_R^2} Z_R,$$

where $a_R^2 < 1$ and Z_R has a zero-mean, unit variance distribution that is independent of E. The copula maps x_R to the probability distribution of the recovery rate on a percentile-to-percentile basis. If G_R is the probability distribution for Z_R, F_R is the unconditional probability distribution for x_R, and Q_R is the unconditional probability for R, then by 4.4

$$Prob\left(R < R^* | M\right) = G_R \left\{ \frac{F_R^{-1}\left[Q_R\left(R^*\right)\right] - a_R E}{\sqrt{1 - a_R^2}} \right\}$$

for some $R^* \in [0\%, 100\%]$.It was pointed out in [19] that the impact of the correlation between the background factor and the recovery rate is significant, particularly for senior tranches. Without the correlation these tranches are relatively "safe". With the correlation they are vulnerable to a bad year where probabilities of default are high and recovery rates are low. This is a significant feature of the model, backed by empirical evidence.

When recovery rates are correlated with the probability of loss the expected loss is increased if the default intensity is the same as in the uncorrelated case. As a result the break-even rate for every tranche is increased with senior tranches more seriously affected.

4.5.8 One Factor Model for the Grouped t Copula

We consider a portfolio of n obligors, where the systematic risk of each obligor is explained by a set of risk factors. It is assumed that the marginals are normally distributed, which is argued in [7] to be reasonable for a short time horizon T (in the particular example outlined in [7] T is one month).

The asset value monthly log returns Y_i for obligor each i are modelled according to the following factor model:

$$Y_i = a_i \mathbf{c}_i' \mathbf{E} + \sqrt{1 - a_i^2} s_i Z_i,$$

where $\mathbf{E}$ is the vector of monthly risk factor log returns (i.e. some macro-economic variables) with a grouped t copula and normally distributed marginals. Also $\mathbb{E}[\mathbf{E}] = 0$ and Z_i has a standard normal distribution independent of $\mathbf{E}$. The parameters $\mathbf{c}_i$ are in the range $[0, 1]$. Analogously to the other one factor models, a_i is the coefficient of determination of systematic risk and $s_i = Var(Y_i) = \mathbf{c}_i' Cov(\mathbf{E})\, \mathbf{c}_i$, $1 \leq i \leq n$. The distribution function of Y_i is $F_i(x) = Prob\,(Y_i \leq x)$.

In this set-up a default event for obligor i is defined to occur when the value of the random variable Y_i drops below a pre-specified default threshold for the obligor d_i, i.e. when $Y_i \leq d_i$. Let $\tilde{p}_i$ be the (unconditional) probability of default for obligor i, so that $\tilde{p}_i = F_i(d_i)$. It is common practice to obtain the estimates for $\tilde{p}_i$ from some internal or external rating system. The conditional probability of default for counterparty i given the risk factors $\mathbf{E}$ can then be written as

$$Q_i(\mathbf{E}) = P[Y_i \leq d_i | \mathbf{E}] = \Phi\left(\frac{F_i^{-1}(\tilde{p}_i) - a_i \mathbf{c}_i' \mathbf{E}}{\sqrt{1 - a_i^2}\sqrt{\mathbf{c}_i' Cov(\mathbf{E}) \mathbf{c}_i}}\right),$$

where Φ denotes the standard normal distribution function. The distribution function of Y_i is unknown. A way around this is to work with the estimated conditional probability of default $\hat{Q}_i(\mathbf{E})$ obtained by replacing $F_i^{-1}(\tilde{p}_i)$ in the equation above by the empirical quantile $\hat{F}_i^{-1}(\tilde{p}_i)$.

To value the estimated future losses on a portfolio we need a default indicator. Therefore, we introduce a state variable X_i for counterparty i at time T which takes the values in $\{0, 1\}$: the value 1 represents the default state and value 0 the non-default state. So $X_i = 1$ is and only if $Y_i \leq d_i$.

The default model described above is applied to each single obligor in the credit portfolio. Conditional on the realisation of a scenario $\mathbf{E}$ counterparty defaults are simulated from independent Bernoulli distributions with parameters $\hat{Q}_i(\mathbf{E})$. Let N_i be the exposure of obligor i, $1 \leq i \leq n$, and R_i its recovery rate. The assumption made in [7] is that $1 - R_i$ is uniformly distributed on $(0, 1)$, and hence effectively that R_i is uniformly distributed.

The total credit loss under the realisation of scenario $\mathbf{E}$ can be expressed as

$$\tilde{L}\left(\mathbf{E}\right) = \sum_{i=1}^{n} X_i\left(\mathbf{E}\right)\left(1 - R_i\right) N_i,$$

which is an extension of the application in [3] covered in the earlier sections.

In practice the credit loss distribution can be obtained by a three stage procedure:

1. Simulation of the monthly risk factor log returns $\mathbf{E}$ from a grouped t-copula with normal marginals;
2. For each obligor i, simulation of the conditional default indicator $X_i(\mathbf{E}) \in \{0, 1\}$ from a Bernoulli distribution with parameter $\hat{Q}_i(\mathbf{E})$;
3. Estimation of the credit loss distribution over a large set of scenarios for $\mathbf{E}$, by integrating exposures and loss given default in the loss function $\tilde{L}(\mathbf{E})$.

4.6 Modelling Shocks with Marshall-Olkin Copula

Marshall-Olkin copulas are associated with shock models, where shocks can be fatal or non-fatal. The fatal shock model is simpler and is more often used in literature. We will focus on its application as outlined in [13].

Consider n obligors, where each nonempty subset of obligors is assigned a shock which is fatal to all components of that subset (i.e. it leads to default of all obligors in the subset). Let O denote the set of all nonempty such subsets of $\{1, \ldots, n\}$. There are in total $2^n - 1$ shock processes, each in one-to-one correspondence with a nonempty subset of $\{1, \ldots, n\}$. Let $X_1, \ldots, X_n$ denote the time-until-default of each obligor, and assume that shocks assigned to different subsets o, $o \in O$ follow independent Poisson processes with intensities λ_o.

Let Z_o, $o \in O$, denote the time of first occurrence of a shock event for o. Then the occurrence times Z_os are independent exponential random variables with parameters λ_o, and the time until default $X_j = \min_{\{o: j \in o\}} Z_o$ for $j = 1, \ldots, n$.

The random vector of the times to default $(X_1, \ldots, X_n)^T$ has and n-dimensional Marshall-Olkin distribution whose survival copula is a Marshall-Olkin n-copula.

Alternatively we can apply the factor model to express each latent variable in terms of a single common and idiosyncratic shocks. We thus have $X_i = \min(E, Z_i)$, $i = 1, \ldots, n$ where E, Z_i are independent exponentially distributed random variables with parameters α, $1-\alpha$, $\alpha \in (0,1)$. We have the survival copula

$$\hat{C}\left(u_1, \ldots, u_n\right) = \min\left(u_1^{\alpha}, \ldots, u_n^{\alpha}\right) \prod_{i=1}^{n} u_i^{1-\alpha}.$$

The default times are then defined as:

$$\tau_i = S_i^{-1}\left(\exp\left(-\min\left(E, Z_i\right)\right)\right),$$

where S_i, $1 \le i \le n$ are the survival distributions of each obligor.

Again, default times are conditionally independent on E and the conditional survival probabilities are given by

$$Prob(t_i > \tau | E) = 1_{E > \ln S_i(t)} \times S_i(t)^{1-\alpha}.$$

This model allows for simultaneous defaults with positive probability.

The fatal shock models can be extended to non-fatal shock models by introducing Bernoulli random variables for each shock to represent the probability of the shock event leading to a loss.

4.7 Dynamic Modelling Through Copulas

So far all the copula models we examined were static, in the sense that the default environment (represented by the common factor E in the factor approach) was defined at inception for the entire time horizon. In [32] an attempt is made to specify default probabilities which evolve over time, within the copula models.

If the behaviour of the securities is correlated, then it is expected that upon default of one security the probability of other securities defaulting increases immediately. This new information will cause a jump in the default probabilities of the remaining securities at the time of default of the first security (i.e. conditional probabilities of survival are affected).

The approach proposed in [32] is to extend the intensity-based approach to incorporate default correlations through a copula, keeping the individual intensity dynamics and model calibration unaffected. We will examine the theory put forward in the following sections.

4.7.1 Preliminaries

The basic probability space for the model is $\left(\Omega, \tilde{\mathcal{F}}, \mathcal{Q}\right)$. The sample space Ω is assumed to be large enough to support all processes that are introduced. All subsequently introduced filtrations are subsets of $\tilde{\mathcal{F}}$ and augmented by the zero-sets of $\tilde{\mathcal{F}}$. The probability measure $\mathcal{Q}$ can - but need not - be a martingale measure for the specific filtrations considered.

The background filtration $(\mathcal{G}_t)_{t\in[0,T]}$ represents the information about the development of general market variables such as share prices, default-free interest rates or exchange rates, and all default-relevant information except explicit information on the occurrence or non-occurrence of defaults. Thus, $(\mathcal{G}_t)_{t\in[0,T]}$ can also contain information on credit spread movements and rating transitions (except for transitions to "default"). It is analogous to the common background factor E of the factor model, except for the notable difference that $(\mathcal{G}_t)_{t\in[0,T]}$ is not constant over $[0,T]$.

We assume that the time of default of obligor i, τ_i $i = 1,\ldots,n$, is the first time when the default countdown process $\gamma_i(t)$ reaches the level of the trigger variable U_i:

$$\tau_i := \inf\left\{t : \gamma_i(t) \leq U_i\right\},$$

where:

1. The default trigger variables U_i, $i = 1,\ldots,n$ are random variables on the unit interval $[0,1]$. Also $\sigma(U_i) =: \mathcal{U}_i$ is the information generated by knowledge of U_i.

2. The default countdown process $\gamma_i(t)$ is defined as per equation 4.1:

$$\gamma_i(t) := \exp(-\int_0^t \lambda_i(u)du).$$

3. The pseudo-default intensity $\lambda_i(t)$ is a nonnegative cádlág (continues á droite, limites á gauche) stochastic process which is adapted to the filtration $(\mathcal{G}_t)_{t\in[0,T]}$ of the background process. We denote $\Lambda_i(t) := \int_0^t \lambda_i(s)ds$ as the integral of the intensity.

Furthermore we define the default and survival indicator processes $X_i(t) := \mathbf{1}_{\{\tau\leq t\}}$ and $I_i(t) := \mathbf{1}_{\{\tau_i>t\}}$ respectively. Filtration $(\mathcal{F}^i_t)_{t\in[0,T]}$ is the augmented filtration that is generated by $X_i(t)$.

The process $\lambda_i(t)$ is called *pseudo* default intensity because it coincides with the default intensity of obligor i only in the "independence" case, or if information is restricted to information about obligor i alone. In general, it will *not* be the default intensity.

A crucial element of this set-up is the careful specification of the available information, since different information will result in different default probabilities. In [32] the following filtrations are further introduced:

Definition 4.7.1.

1. *Filtration* $(\tilde{\mathcal{H}}_t^i)_{t\in[0,T]}$ *contains information about the default or survival of obligor* i *up to time* t*, and complete information about the background process:*

$$\tilde{\mathcal{H}}_t^i := \sigma(\mathcal{F}_t^i \cup \mathcal{G}).$$

2. *Filtration* $(\mathcal{H}_t^i)_{t\in[0,T]}$ *contains information about the default or survival of obligor* i *up to time* t*, and partial information about the background process up to time* t*:*

$$\mathcal{H}_t^i := \sigma(\mathcal{F}_t^i \cup \mathcal{G}_t).$$

3. *Filtration* $(\tilde{\mathcal{H}}_t)_{t\in[0,T]}$ *reflects information about the defaults of all obligors until* t*, and complete information about the background process:*

$$\tilde{\mathcal{H}}_t = \sigma\left(\bigcup_{i=1}^{n} \tilde{\mathcal{H}}_t^i\right).$$

4. *Filtration* $(\mathcal{H}_t)_{t\in[0,T]}$ *is the equivalent of* $(\tilde{\mathcal{H}}_t)_{t\in[0,T]}$*, but with the information on the background process restricted to* $[0,t]$ *:*

$$\mathcal{H}_t = \sigma\left(\bigcup_{i=1}^{n} \mathcal{H}_t^i\right).$$

5. *Filtration* $(\mathcal{F}_t)_{t\in[0,T]}$ *contains only default information of all* n *obligors up to time* t *:*

$$(\mathcal{F}_t)_{t\in[0,T]} = \sigma\left(\bigcup_{i=1}^{n} \mathcal{F}_t^i\right).$$

These filtrations enable us to model the intensity of the default process of one obligor independent of the information about the default behaviour of the remaining $n-1$ obligors. It can be seen that the default process X_i would have a different intensity under the filtration $(\mathcal{H}_t^i)_{t\in[0,T]}$ than under $(\mathcal{H}_t)_{t\in[0,T]}$. Accordingly, we need to define different survival probabilities.

Definition 4.7.2. *For each obligor* i, $i = 1, \dots, n$ *define at time* t, $0 \le t \le T$*:*

1. *the survival probability until* T *given the information* $(\mathcal{H}_t^i)_{t\in[0,T]}$*:*

$$S_i'(t,T) := \mathbb{E}^Q[I_i(T)|\mathcal{H}_t^i],$$

2. *the survival probability until* T *given the information* $(\tilde{\mathcal{H}}_t^i)_{t\in[0,T]}$*:*

$$\tilde{S}_i'(t,T) := \mathbb{E}^Q[I_i(T)|\tilde{\mathcal{H}}_t^i].$$

An assumption is also made that the default threshold U_i, $i = 1, \dots, n$, is uniformly distributed on $[0, 1]$ under $(Q, \mathcal{H}_0^i)$, and $\mathcal{U}_i$ is independent from $\mathcal{G}_\infty$ under Q. Under this assumption only the *marginal* distribution of the U_i is prescribed as the filtration $(\mathcal{H}_t^i)_{t\in[0,T]}$ does not contain information about the other U_j, $j \neq i$.

Under the above assumption, and given that $\tau_i > t$, the survival probabilities are:

$$\tilde{S}_i'(t,T) = \frac{\gamma_i(T)}{\gamma_i(t)} = e^{-\int_t^T \lambda_i(s)ds},$$

$$S_i'(t,T) = \mathbb{E}^Q\left[\frac{\gamma_i(T)}{\gamma_i(t)}|\mathcal{H}_t^i\right] = \mathbb{E}^Q[e^{-\int_t^T \lambda_i(s)ds}|\mathcal{H}_t^i], \tag{4.7}$$

which is a proposition proved in [32]. So we can express and define the intensity of the default process through the survival probabilities.

The following proposition shows that λ_i is indeed the default intensity of obligor i if the only information at hand is that concerning the general economic environment and the particular obligor.

Proposition 4.7.1. *The intensity of the default indicator process $X_i(t)$ under the filtration $(\mathcal{H}_t^i)_{t\in[0,T]}$ is*

$$-\frac{\partial}{\partial T} S_i'(t,T)|_{T=t} = 1_{\{\tau^i > t\}}\lambda_i(t).$$

This proposition can be proved by the differentiation of 4.7. We have essentially reduced the model to a standard default risk intensity-type model if the information set is reduced to $(\mathcal{H}_t^i)_{t\in[0,T]}$.

4.7.2 Modelling Dependent Defaults

We now look at an application of dynamic intensity modelling discussed in 4.7. The copula of joint defaults is still specified for a given time horizon, in other words as a static framework. Dependence between the defaults of the n obligors is introduced through the specification of the joint distribution of the random variables $U_1, \dots, U_n$. We will see how the default intensity of an obligor is affected by the knowledge of the default of another obligor.

The two copulas which are considered in [32] are the Gumbel and Clayton copulas. They are commonly used examples of the Archimedean family of copulas. The exact form of the Archimedean copula is specified through its generator function (refer to Chapter 3 for more details). As we have already seen earlier, Gumbel copula has upper tail dependence, while the Clayton copula has lower tail dependence.

For the Gumbel copula the default intensity $\tilde{\lambda}_i$ is:

$$\tilde{\lambda}_i(t) = \frac{\phi'(\gamma_i)}{C(\gamma)\phi'(C(\gamma))}\gamma_i\lambda_i \qquad = \left(\frac{\Lambda_i}{\|\Lambda\|_\theta}\right)^{\theta-1}\lambda_i,$$

where C denotes the (Gumbel) copula. In [32] it has been shown that the jump in the default intensity of obligor i on the default of obligor j can be expressed as:

$$\tilde{\lambda}_i^{-j}(t) = \left(-\frac{C(\gamma)\phi''(C(\gamma))}{\phi'(C(\gamma))}\right)\tilde{\lambda}_i \qquad = \left(1 + \frac{(\theta-1)}{\|\Lambda\|_\theta}\right)\tilde{\lambda}_i,$$

where we denote $\|x\|_\theta := \left(\sum_{i=1}^n |x_i|^\theta\right)^{1/\theta}$ for the θ-norm in $\mathbb{R}^n$ and $\Lambda_i(t) := \int_0^t \lambda_i(s)ds$.

Similarly, for the Clayton copula the default intensity is

$$\tilde{\lambda}_i(t) = \frac{\phi'(\gamma_i)}{C(\gamma)\phi'(C(\gamma))}\gamma_i\lambda_i = \left(\frac{C(\gamma)}{\gamma_i}\right)^\alpha \lambda_i$$

and the jump in the intensity on obligor j's default:

$$\tilde{\lambda}_i^{-j}(t) = \left(-\frac{C(\gamma)\phi''(C(\gamma))}{\phi'(C(\gamma))}\right)h_i \qquad = (1+\alpha)h_i.$$

With the Clayton copula there is a constant jump in the default intensity (and therefore a constant increase to the credit spread) by a factor $(1+\alpha)$ at default of another obligor j. While no defaults happen there is a drift correction to the default intensity which has an effect of reducing these default intensities. One can also directly obtain the distribution of default times.

For the Gumbel copula, the intensity parameter $h_i(t)$ depends on the factor $\frac{\Lambda_i(t)}{\|\Lambda(t)\|_\theta}$, which represents the dependence structure of the default times. This is the i-th component of the cumulative intensity vector Λ, normalised to one in the θ-norm. For constant pseudo-intensities λ_i, this factor will be constant, thus making the model time-invariant before defaults. In contrast to this, the jump factor when a default happens is not constant but approaches 1 as time proceeds, thus preventing the default intensities from increasing without bound.

4.8 Discussion on the Drawbacks of Copulas

We have seen that copulas provide an elegant solution to a complex problem of modelling the behaviour of a basket of securities or obligors. The methodology of finding a solution is broken down into two steps: the first step deals with modelling each obligor's univariate marginal distribution; the second step consists of specifying a copula, which summarizes all the dependencies between margins. We are provided with the flexibility of being able to specify different marginal distributions for each obligor.

However, no study of a model would be complete without highlighting its possible pitfalls. Therefore the natural progression of our study of copulas is to examine the aspects of which one needs to be aware if copulas are used for modelling dependencies in credit derivatives.

4.8.1 Estimation of Copulas

A natural question to pose when a copula model is being set-up is whether one should first fully specify the univariate marginals or whether the whole model should be specified is the same step. Apart from deciding on the form of univariate marginals and the form of the copula (i.e. whether it should be for example Gaussian, Clayton, or say Marshall-Olkin), the parameters for both need to be estimated. This question is a central theme of [14], and we will look into its arguments in this section.

Let us assume that the true copula for our model belongs to a parametric family $\mathcal{C} = \{C_\theta : \theta \in \Theta\}$, where Θ is the parameter space. We can use the maximum likelihood methods (ML) to estimate consistent and asymptotically normally distributed estimates of the parameter θ, which belong to the parameter space Θ. The test can be specified in two different set-ups. The first method assumes our knowledge of the parametric forms of the univariate marginal distributions. Each margin is plugged into the likelihood function, which is then maximised with respect to parameter θ. Alternatively we can make no assumptions about the parametric form of the marginals, and instead plug the univariate empirical cumulative distribution functions into the likelihood function for a so-called semi-parametric method.

The first method depends on the correct specification of all the marginal distributions. This constraint may be too onerous, and could lessen the interest of working with copulas. The semi-parametric estimation procedure, where margins are left unspecified, does not suffer from this inconvenience but is less efficient to apply. In addition there is no guarantee that the specified copula would be the most suitable one for our modelling. Also the standard inference for independent and identically distributed (i.i.d) data does not hold with time-dependent data. This means that test statistics delivered by standard ML method under the usual assumption of i.i.d. data should be handled with care.

The choice among possible copula specifications can also be done rigorously via so-called goodness-of-fit (GOF) tests. A GOF test for multivariate distributions may be set-out in two different ways:

- $\mathcal{H}_0 : F = F_0$, against $\mathcal{H}_a : F \neq F_0$, when the null hypothesis is simple, or
- $\mathcal{H}_0 : F \in \mathcal{F}$, against $\mathcal{H}_a : F \notin \mathcal{F}$, when the null hypothesis is composite.

Here F_0 denotes some known cumulative distribution function, and $\mathcal{F} = \{F_\theta : \theta \in \Theta\}$ is some known parametric family of multivariate distribution functions.

In a one-dimensional framework the problem is relatively simple. It is pointed out in [14] that by considering the transformation of a random variable X_i by its distribution function F_i, the corresponding empirical process tends weakly to a uniform Brownian bridge under the null hypothesis. For this a lot of well-known distribution-free GOF statistics are available. In a multidimensional framework, it is more difficult to build distribution-free GOF tests because the law of transformed variable is no longer distribution-free.

In a similar manner, the GOF problem for copulas can be set out based on either of the two assumptions:

- $\mathcal{H}_0 : C = C_0$, against $\mathcal{H}_a : C \neq C_0$, when the null hypothesis is simple, or
- $\mathcal{H}_0 : C \in \mathcal{C}$, against $\mathcal{H}_a : C \notin \mathcal{C}$, when the null hypothesis is composite.

Here, C_0 denotes some known copula and $\mathcal{C} = \{C_\theta : \theta \in \Theta\}$ is some known parametric family of copulas. The difficulty of applying the same method as for multivariate distributions is that the univariate cumulative distribution functions F_js are now unknown.

In general, the GOF test for copulas has not yet been tackled rigorously to provide reliable asymptotic estimates.

A possible approach is to perform a test for each marginal distribution separately as a first step. If each marginal model is accepted, then a test of the whole multidimensional distribution can be implemented. However, such a procedure is heavy, and it is still necessary to deal with a multidimensional GOF test.

4.8.2 Too Much Choice Can Be a Bad Thing

A common argument in literature against the copula approach is that there are almost too many copulas to choose from. This is the case even if we specify desirable attributes of the copula for modelling dependence of credit derivatives, such as tail dependence.

In fact there is an infinite number of one-factor copulas that can exactly fit observed market prices. It is not clear how one would choose amongst these copulas to select the "correct" model. However, the implications of choosing a particular copula come to light when our chosen model is then used to model some non-standard contract, or when we want to apply our model to extreme events such as probability of default.

Copulas, by their nature, are specified independently of the marginal distributions. Therefore the knowledge of the parametric form of the marginals does not aid in the specification of the copula. This approach may be considered as against the intuitive understanding of the system to be modelled. Indeed, in [30] the author puts forward a question "if one marginal is Gaussian and the other Student, should we use a Gaussian copula or a t-copula?". The fact that the answer to this question does not come to mind immediately highlights the arbitrariness in the copula philosophy.

4.8.3 Modelling Dependence

It is pointed out in [30] that the copula approach does not normally correspond to any natural multivariate structure arising from some underlying dynamic. The exception is the Marshall-Olkin copula which is associated with a system containing components which are subject to certain shocks leading to failure of either or both components. Some Archimedean copulas do arise naturally in shared frailty models for dependent lifetimes, this being the origin of the Clayton copula in particular.

As already suggested in sections above, copulas do not enable us to model the source of dependence or the common factor itself. Copulas give us the tools to model dependence as a separate entity, but the set-up is not conducive to obtaining insight into the cause of this dependence. In such a setting we are less likely to understand dependence we are modelling and be aware of how it may evolve over time.

4.8.4 Other Considerations

A commonly recognised main "unit" in modelling defaults is the hazard rate. While copulas can also be specified with the hazard rate as the underlying default mechanism, the set-up prevents a dynamic evolution of the hazard rates or, for that matter, credit spreads. An attempt at the dynamic modelling of hazard rates has been made in [32], though the solution is cumbersome, and the model would need to be manually adjusted on each obligor's default. The copula model is inappropriate for valuing contracts such as a one-year option on a five-year CDO because this instrument depends on hazard rates between years one and five conditional on what we observe happening during the first year.

Some attempts have been made to incorporate a dependence between the default time and the recovery rate, as seen in [19]. However, while it is intuitive for dependence to also exist between default probabilities and interest rates, models to date have not tackled the inclusion of this dependence. Interest rates have been assumed to be either constant or to be a deterministic function of time.

One of the main problems that must be kept in mind is that the copula and its parameters are fitted to "normal" market conditions, and not the conditions of interest for credit derivative modelling. We can only guess how our copula will perform in default environments. However, as this is an issue of availability of data, other approaches to modelling credit derivatives and dependencies between defaults are also not immune to this shortcoming.

Chapter 5

Alternative Methods for Capturing Dependence

5.1 Introduction

We have had an in-depth investigation into the theory of copulas and their applications to modelling dependent defaults, this being a crucial element of baskets of credit-risky securities. Having identified in the last chapter the areas where copulas are lacking, we now investigate other options available for modelling dependent defaults. This chapter presents some of these alternative modelling techniques covered in the literature.

5.2 Some General Modelling Concepts

It has frequently been recognised, and specifically noted in [22], that, in general, the mechanisms for obtaining dependence are all versions of three underlying themes:

- Default probabilities are influenced by common background variables which are observable. As in all factor models, we then need to specify the joint movements of the factors and how default probabilities depend on these factors.
- Default probabilities depend on unobserved background variables, and the occurrence of an event causes updating of these latent variables, which in turn causes a reassessment of the default probability of the remaining assets.
- There is a direct contagion in which a default event of one firm directly causes a default of another firm or, in the least, a deterioration of credit quality of this second firm.

It is likely that we would need to calculate the distribution of the number of defaults among a large collection of issuers, which is very cumbersome unless some sort of homogeneity assumption is made. This naturally leads to the use of binomial distributions. In fact, much of correlation modelling can be seen as variations of the binomial distribution using mixture distributions.

The binomial distribution here provides a framework for studying the distribution of the number of defaults over a pre-specified time horizon. In addition we introduce the mixture distribution which randomizes the default probability within the binomial, effectively creating dependence between the securities modelled. It mimics a situation where a common background variable affects a collection of firms, as in the factor model application to copulas (see Chapter 4 for details). It is important to note here that the default probability will depend on the chosen time horizon. We will discuss this model in more detail in the next section.

Throughout this chapter n represents the number of firms in our collection. The default indicator of firm i is denoted by X_i and is equal to 1 if firm i defaults and 0 otherwise. The number of defaults is denoted by M_n so that

$$M_n = X_1 + \cdots + X_n.$$

The default probability $\tilde{p}$ is itself a random variable, taking values in the interval $[0, 1]$, which is independent of all the X_i. For simplicity it is the same variable for all X_i. Conditional on $\tilde{p}$ the random variables $X_1, X_2, \ldots, X_n$ are independent. For large portfolios it is the distribution of $\tilde{p}$ which determines the loss distribution. The more variability that there is in the mixture distribution, the more correlation there is between default events, and hence the more weight there is in the tails of the loss distribution.

5.3 The Mixture Model Approach

As mentioned above, mixture models introduce correlation between variables (which in our case represent default indicators) through the stochastic modelling of the default probability. Given a specific realization of this default probability, the defaults of individual firms are independent. We will discuss some examples of the mixture models covered in [12].

Definition 5.3.1. ***Bernoulli mixture model*** *Given a random variable Ψ, the random vector $\mathbf{X}^n = (X_1, \ldots, X_n)'$ follows a Bernoulli mixture model with random variable Ψ if there are functions $p_i : \mathbb{R} \to [0, 1]$, $1 \leq i \leq n$, such that conditional on Ψ the components of $\mathbf{X}^n$ are independent Bernoulli random variables satisfying $P(X_i = 1|\Psi = \psi) = p_i(\psi)$.*

Then for a specific default scenario $\mathbf{x}^n = (x_1, \ldots, x_n)'$ in $\{0, 1\}^n$ we have that

$$P(\mathbf{X} = \mathbf{x}|\Psi = \psi) = \prod_{i=1}^{n} p_i(\psi)^{x_i} (1 - p_i(\psi))^{1-x_i}.$$

The unconditional distribution of the default indicator vector $\mathbf{X}^n$ is obtained by integrating over the distribution of the random variable Ψ. In particular, the default probability of company i is given by $\bar{p}_i = P\left(X_i = 1\right) = \mathbb{E}\left(p_i\left(\Psi\right)\right)$.

Since default is expected to be a rare event, it is possible to approximate Bernoulli random variables with Poisson random variables in Poisson mixture models. With this model a company may potentially default more than once in the period we are analysing, however with a very low probability, especially if we keep the Poisson parameters λ_i fairly small. We will use the notation $\tilde{X}_i \in \{0, 1, 2, \ldots\}$ for the counting random variable giving the number of defaults of company i. It is then the random variables $\tilde{X}_i$, $1 \le i \le n$, that have Poisson distributions. The formal definition parallels the definition of a Bernoulli mixture model.

Definition 5.3.2. ***Poisson mixture model*** *Given p and Ψ as in Definition 5.3.1, the random vector $\tilde{\mathbf{X}}^n = \left(\tilde{X}_1, \ldots, \tilde{X}_n\right)'$ follows a Poisson mixture model with random variable Ψ if there are functions $\lambda_i : \mathbb{R} \to (0, \infty)$, $1 \le i \le n$, such that conditional on $\Psi = \psi$ the random vector $\tilde{\mathbf{X}}^n$ is a vector of independent Poisson distributed random variables with rate parameter $\lambda_i\left(\psi\right)$.*

We define the random variable $\tilde{M} = \sum_{i=1}^{n} \tilde{X}_i$ which, for small Poisson parameters λ_i, approximately represents the number of defaulting companies. Given the realisation of the random variable Ψ, $\tilde{M}$ it is the sum of conditionally independent Poisson variables and therefore its distribution satisfies

$$P\left(\tilde{M} = k|\Psi = \psi\right) = \exp\left(-\sum_{i=1}^{n} \lambda_i\left(\psi\right)\right) \frac{\left(\sum_{i=1}^{n} \lambda_i\left(\psi\right)\right)^k}{k!}.$$

If $\tilde{\mathbf{X}}^n$ follows a Poisson mixture model and we re-define our default indicators as $X_i = I_{\{\tilde{X} \ge 1\}}$ then $\mathbf{X}^n$ follows a Bernoulli mixture model and the mixing variables are related by $p_i\left(\cdot\right) = 1 - \exp\left(-\lambda_i\left(\cdot\right)\right)$.

Often a simplification is introduced to this set-up by making the probability functions p_i all identical. In this case the Bernoulli mixture model is said to be *exchangeable*, since the random vector $\mathbf{X}^n$ is exchangeable. The number of defaults then has a Binomial distribution with parameter $Q := p_1\left(\psi\right)$.

The following mixing distributions of parameter Q are frequently used in Bernoulli mixture models:

- **Beta mixing distribution**: here we assume that $Q \sim \text{Beta}\left(a, b\right)$ for some parameters $a > 0$ and $b > 0$.
- **Probit-normal mixing distribution**: we set $Q = \Phi\left(\mu + \sigma\Psi\right)$ for $\Psi \sim N\left(0, 1\right)$, $\mu \in R$ and $\sigma > 0$, where Φ is the standard normal distribution function. It turns out that this model can be viewed as a one-factor version of the CreditMetrics model.

- **Logit-normal mixing distribution**: we set $Q = F(\mu + \sigma\Psi)$ for $\Psi \sim N(0,1)$, $\mu \in R$ and $\sigma > 0$, where $F(x) = (1+\exp(-x))^{-1}$ is the distribution function of a so-called logistic distribution.

The tail of $\tilde{M}$ is essentially determined by the tail of the mixing variable Q. However, once the default probability and default correlation have been chosen, the parametric form of the mixing distribution is of minor importance.

The majority of useful threshold models can be represented as Bernoulli mixture models. The following Lemma, which is proved in [12] shows the role of mixture models within the threshold model framework.

Lemma 5.3.1. *Let (V, D) be a threshold model for an n-dimensional random vector $\mathbf{V}^n$. If $\mathbf{V}^n$ has a p-dimensional conditional independence structure with conditioning variable Ψ, then the default indicators $X_i = I_{\{V_i \leq d_{i1}\}}$ follow a Bernoulli mixture model with factor Ψ, where the conditional default probabilities are given by $p_i(\psi) = P(V_i \leq d_{i1} | \Psi = \psi)$.*

CreditRisk+ has been an industry model for credit risk which was proposed by Credit Suisse Financial Products in 1997. The model has the structure of a Poisson mixture model, where the factor Ψ consists of s independent, gamma-distributed random variables. The (stochastic) parameter $\lambda_i(\Psi)$ of the conditional Poisson distribution for firm i is given by $\lambda_i(\Psi) = k_i w_i' \Psi$ for a constant $k_i > 0$. The non-negative factor weights $w_i = (w_{i1}, \ldots, w_{is})'$ satisfy the criteria $\sum_j w_{ij} = 1$, and s is the number of independent $Ga(\alpha_j, \beta_j)$-distributed factors $\Psi_1, \ldots, \Psi_s$ with parameters set to be $\alpha_j = \beta_j = \sigma_j^{-2}$ for some $\sigma_j > 0$. This parametrization of the gamma variables ensured that we have $\mathbb{E}(\Psi_j) = 1$, $var(\Psi_j) = \sigma_j^2$ and $\mathbb{E}(\lambda_i(\Psi)) = k_i \mathbb{E}(w_i'\Psi) = k_i$. Observe that in this model the default probability is given by $P(X_i = 1) = P\left(\tilde{X}_i > 0\right) = \mathbb{E}\left(P\left(\tilde{X} > 0 | \Psi\right)\right)$. Since $\tilde{X}_i$ is Poisson given Ψ, we have that

$$\mathbb{E}\left(P\left(\tilde{X} > 0|\Psi\right)\right) = \mathbb{E}\left(1 - \exp\left(-k_i w_i'\Psi\right)\right) \approx k_i \mathbb{E}\left(w_i'\Psi\right) = k_i,$$

where the approximation holds because k_i is typically small. Hence k_i is approximately equal to the default probability for firm i.

Then the distribution of $\tilde{M} = \sum_{i=1}^n \tilde{X}_i$ is conditionally Poisson and satisfies

$$\tilde{M}|\Psi = \psi \sim Poi\left(\sum_{i=1}^n k_i w_i' \psi\right).$$

While this model is simple to understand, one of the shortcomings of the mixture model is that the assumption of homogeneity places some serious restrictions on the portfolio. It is expected that companies in different industrial sectors, as well as companies in different economies, have different default probabilities. The homogeneity assumption effectively reduces the scope of portfolios for which defaults can be accurately modelled.

5.4 A Contagion Model

Contagion models have become widely used in default modelling. The concept of contagion means that once a firm defaults, it may bring down other firms with it. We look into the model which incorporates contagion in a binomial-type model, as described in [22] and [8].

In the general mixture model it is the common dependence on the background variable $\tilde{p}$ that induces the correlation in the default events. While it is possible in this model to obtain all correlations between 0 and 1, we need to assume large fluctuations in $\tilde{p}$ to obtain a significant correlation. A more direct way that does not push the marginal probabilities up as sharply, while still inducing correlation, is to have direct contagion.

The model is constructed with us distinguishing between direct defaults and defaults triggered through a contagion event. We thus introduce Y_{ij} - an "infection" variable - which, when equal to 1, implies that default of firm i immediately triggers default of firm j. Now let us assume that all X_i, Y_{ij} are independent Bernoulli variables with $P(X_i = 1) = p$ and $P(Y_{ij} = 1) = q$. Let the new combined default indicator of firm i be given by

$$Z_i = X_i + (1 - X_i)\left(1 - \prod_{j \neq i}(1 - X_j Y_{ji})\right).$$

This expression is equal to 1 either if there is a direct default of firm i (in other words if $X_i = 1$) or if there is no direct default *and* the entire second term of the sum is 0. This second term is 0 precisely when at least one of the factors X_j, Y_{ji} is 1, which happens when firm j defaults and infects firm i. Now the number of defaults is

$$M_n = Z_1 + \cdots + Z_n.$$

One of the desirable features of this method is that we have constructed the model for dependence using independent variables, which are easier to handle than dependent variables. We shall see later in this chapter another model based on this approach.

5.5 Continuous Time Analogue of the Binomial Model

We can extend the stochastic one-period binomial model with identical default intensities of firms into a continuous time model, as described in [22] . This model is considered to be dynamic in a sense since the default intensity of the next firm to default depends on the number of firms which have already defaulted.

We consider a state space $0, 1, \ldots, n$, where each state represents the number of non-defaulted firms (n being the total number of firms). Assuming that only one firm can default at a time and that every firm has a default intensity λ, the intensity of the first default among n firms is $n\lambda$. Therefore, we define the intensity of the so-called pure death process going from state i to state $i-1$ as $\lambda_{i,i-1} = i\lambda$ for $i = 1, \ldots, n$, and we let state 0 be the absorbing state. It is then easy to compute the distribution of the number of firms alive at some future date t. If the starting population is n then, since we can view each firm as having an exponential lifetime with intensity parameter λ, the probability of there being k firms left at time t is

$$\binom{n}{k} \exp\left(-k\lambda t\right) \left(1 - \exp\left(-\lambda t\right)\right)^{n-k}.$$

This expression corresponds to all the ways in which k firms survive and $n-k$ firms default within the given time period. Note the difference between the probability of this process having k defaults and the probability that n Poisson processes with intensity λ experience k defaults. The latter probability is larger since no firms leave the population after a default.

It is also straightforward to include a common background variable variable just as in the Cox setup: given an intensity process λ and assuming that, conditional on the sample path of this process, we have

$$P_{k,\lambda}(t) = \binom{n}{k} \exp\left(-k \int_0^t \lambda_s ds\right) \left(1 - \exp\left(- \int_0^t \lambda_s ds\right)\right)^{n-k}.$$

5.6 Intensity Correlation Through Factor Specification

One of the most common approaches to handling correlation between a large number of issuers is to impose a factor structure on the default intensities. This approach has also been the building block of copula work in [23] and we have looked at this setup in Chapter 4.

The goal with the factor approach is to reduce the amount of parameter specification by quantifying the part of the marginal intensity which comes from a set of common factors when specifying the marginal distributions. A good introductory example can be seen in [19], where the factor structure is introduced to model the value of assets in a portfolio. When we impose a factor structure to default intensities, we divide the structure of the intensity of an individual firm i into two independent components, one coming from a common factor and one being idiosyncratic. In [22] the following simple structure is set out:

$$\lambda^i(t) = v^c(t) + v^i(t).$$

The common factor v^c can be split further into industry sector components and general economy components, but each additional split introduces its own estimation problems.

If we apply this split in a Cox process setup we get the survival probability of an issuer to be

$$\begin{aligned} P\left(\tau^i > T\right) &= \mathbb{E}\left[\exp\left(\int_0^T \lambda^i(s)ds\right)\right] \\ &= \mathbb{E}\left[\exp\left(\int_0^T v^i(s)ds\right)\right]\mathbb{E}\left[\exp\left(\int_0^T v^c(s)ds\right)\right], \end{aligned}$$

where τ^i is the default time of the ith obligor. The correlation between firms arises because of the common intensity component. To avoid the problems with the lack of homogeneity, we will need to consider pools of issuers whose marginal intensities and whose split between global and idiosyncratic risk are similar enough to be grouped. In this case we are back to a mixed binomial model in which, after conditioning on the common factor, defaults are independent.

An example specification of systematic risk found in [22] is

$$dv_t^c = \kappa^c\left(\theta^c - v_t^c\right)dt + \sigma^c\sqrt{v_t^c}dW_t^c + dJ_t^c,$$

where

$$J_t^c = \sum_{i=1}^{N_t^c} \epsilon_i^c,$$

and N_t^c is a Poisson process with intensity $l^c \geq 0$ while $(\epsilon_i)_{i=1}^{\infty}$ is a sequence of independent and identically distributed exponential random variables with mean μ^c. This specification represents a mean-reverting Brownian motion process with jumps of random sizes (given by a Poisson process) at random times (given by the exponential random variable).

It is then common to assume that the idiosyncratic risk is specified in exactly the same form, replacing all superscripts "c" by "i". In particular, the driving Brownian motions and the jump processes are independent. In this specification jumps in the total intensity occur at a rate of

$$l = l^c + l^i.$$

We let

$$\rho = l^c/l$$

denote the fraction of jumps in the individual firm's intensity that is due to common jumps. It turns out that if one assumes that

$$\kappa := \kappa^c = \kappa^i,\ \sigma := \sigma^c + \sigma^i,\ \text{and}\ \mu := \mu^c + \mu^i,$$

then the sum of two affine processes v^i and v^c is again an affine process with parameters κ, σ, μ as defined above and with jump rate and mean reversion level given as

$$l = l^i + l^c \text{ and } \theta = \theta^i + \theta^c.$$

This means that as long as we keep these two sums constant, we can vary the contribution from the systematic and idiosyncratic risk without changing the marginal rate. With this we are effectively controlling the correlation of default rates without having to resort to large variations of the background variable.

5.7 Reduced-Form Credit Risk Models

Structural models generally attempt to account at some level of detail for the events leading to a default. A problem of the structural approach is that it is difficult in such a model to deal systematically with the multiplicity of situations that can lead to default. That is why structural models are sometimes viewed as unsatisfactory as a basis for a practical modelling framework, particularly when multi-name products such as nth to default swaps and collateralised debt obligations are involved.

Reduced-form models are more commonly used in practice on account of their tractability and fewer assumptions being required regarding the nature of the debt obligations involved and circumstances leading to default. Most reduced-form models are based on a random time of default, modelled as the time at which the integral of a random intensity process first hits a certain critical level, this level itself being an independent random variable.

Before we begin our review of various reduced-form models we introduce some necessary mathematical concepts.

5.7.1 Mathematical tools

We consider a probability space $(\Omega, \mathcal{F}, P)$ and a random time to default τ defined on this space. As usual, we denote $F(t) = P(\tau \leq t)$ and $S(t) := 1 - F(t)$ as the corresponding survival probability function. Again we define the jump-to-default indicator process (X_t) associated with τ by $X_t = I_{\{\tau \leq t\}}$ for $t \geq 0$.

We assume that the only observable quantity is the random time τ or, equivalently, the associated jump indicator process (X_t). The appropriate filtration is therefore given by $(\mathcal{H}_t)$ with

$$\mathcal{H}_t = \sigma\left(\{X_u : u \leq t\}\right).$$

By definition, τ is an $(\mathcal{H}_t)$-stopping time, as $\{\tau \leq t\} = \{X_t = 1\} \in \mathcal{H}_t$ for all $t \geq 0$; moreover $(\mathcal{H}_t)$ is the smallest filtration with this property.

The following Lemma, stated without proof, lays out a useful way of expressing expectations conditional on the filtration $(\mathcal{H}_t)$.

Lemma 5.7.1. *Let τ be a random time with jump indicator process $X_t = I_{\{\tau \leq t\}}$ and natural filtration $(\mathcal{H}_t)$. Then, for any integrable random variable V and any $t \geq 0$,* *we have*

$$\mathbb{E}\left(I_{\{\tau>t\}}V|\mathcal{H}_t\right) = I_{\{\tau>t\}}\frac{\mathbb{E}\left(V;\tau>t\right)}{P\left(\tau>t\right)}.$$

Proposition 5.7.1. *Let τ be a random time with absolutely continuous distribution function $F(t)$ and hazard-rate function $\lambda(t)$. Then $\tilde{M}_t := X_t - \int_0^{t\wedge\tau}\lambda(s)\,ds$, $t \geq 0$, is an (H_t)-martingale, where H is a $\mathcal{H}_t$-measurable random variable.*

The proofs of Lemma 5.7.1 and Proposition 5.7.1 can be found in [12].

Additional information is typically generated by background processes, often modelled as diffusions or continuous-time Markov chains, representing, for instance, economic activity in a country or in an industry sector, risk-free interest rates or rating transitions between the non-default states. Formally, we represent this additional information by some filtration $(\mathcal{F}_t)$ on $(\Omega, \mathcal{F}, P)$.

This leads us to introduce a new filtration $(\mathcal{G}_t)$ by

$$\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t,\ t \geq 0,$$

meaning that $\mathcal{G}_t$ is the smallest σ-algebra that contains $\mathcal{F}_t$ and $\mathcal{H}_t$. Obviously τ is an $(\mathcal{H}_t)$ stopping time and hence also a $(\mathcal{G}_t)$-stopping time. In the context of credit risk models the filtration $(\mathcal{G}_t)$ contains information about the background processes and the occurrence or non-occurrence of default up to time t, and thus typically corresponds to the information available to investors.

The following Lemma is stated without proof in [12].

Lemma 5.7.2. *For every $\mathcal{G}_t$-measurable random variable V there is some $\mathcal{F}_t$-measurable random variable $\tilde{V}$ such that $VI_{\{\tau>t\}} = \tilde{V}I_{\{\tau>t\}}$.*

In economic terms this Lemma tells us that before a default occurs all information is generated by the background filtration $(\mathcal{F}_t)$. Now we turn to conditional expectations with respect to $\mathcal{G}_t$.

Lemma 5.7.3. *For every integrable random variable V we have*

$$\mathbb{E}\left(I_{\{\tau>t\}}V|\mathcal{G}_t\right) = I_{\{\tau>t\}}\frac{\mathbb{E}\left(I_{\{\tau>t\}}V|\mathcal{F}_t\right)}{P\left(\tau>t|\mathcal{F}_t\right)}.$$

Note that Lemma 5.7.3 allows us to replace certain conditional expectations with respect to $\mathcal{G}_t$ by conditional expectations with respect to background information $\mathcal{F}_t$.

Proof. $\mathbb{E}\left(I_{\{\tau>t\}}V|\mathcal{G}_t\right)$ is $\mathcal{G}_t$-measurable and zero on $\{\tau \leq t\}$. By Lemma 5.7.2 there is therefore an $\mathcal{F}_t$-measurable random variable $\tilde{Z}$ such that $\mathbb{E}\left(I_{\{\tau>t\}}V|\mathcal{G}_t\right) = I_{\{\tau>t\}}\tilde{Z}$. Since $\mathcal{F}_t \subset \mathcal{G}_t$, taking conditional expectations with respect to $\mathcal{F}_t$ yields

$$\mathbb{E}\left(I_{\{\tau>t\}}V|\mathcal{F}_t\right) = P\left(\tau>t|\mathcal{F}_t\right)\tilde{Z}.$$

Hence $\tilde{Z} = \mathbb{E}\left(I_{\{\tau>t\}}V|\mathcal{F}_t\right)/P\left(\tau > t|\mathcal{F}_t\right)$, which proves the lemma. □

We can extend the reduced-form models by replacing the usually deterministic hazard rates with stochastically-modelled hazard rates. These are referred to as models with doubly stochastic random times - also called *conditional Poisson* or *Cox* random times in the literature. However, before we formally define doubly stochastic random times, let us first introduce the cumulative hazard rate.

Definition 5.7.1. ***Cumulative hazard function***. *The function $\Lambda(t) := -\ln(S(t))$ is called the cumulative hazard function of the random time τ. If F is absolutely continuous with density f, then the corresponding hazard rate function is $\lambda(t) := f(t)/(1-F(t)) = f(t)/S(t)$.*

Definition 5.7.2. ***Doubly stochastic random times*** *A random time τ is called doubly stochastic with respect to the background filtration $(\mathcal{F}_t)$ if it admits the $(\mathcal{F}_t)$-conditional hazard-rate process (λ_t), if $\Lambda(t)$ is strictly increasing, and if, for all $t > 0$,*

$$P\left(\tau \le t|\mathcal{F}_\infty\right) = P\left(\tau \le t|\mathcal{F}_t\right).$$

The above conditioning means that, given the past values of the information process, the future values do not contain any extra information for predicting the probability that the default time τ occurs before some future time t.

We have seen in proposition 5.7.1 that the jump indicator process (X_t) can be turned into an $(\mathcal{H}_t)$-martingale if we subtract the process $\int_0^{t\wedge\tau}\lambda(s)\,ds$. Here we generalize this result to doubly stochastic random times.

Proposition 5.7.2. *Let τ be a doubly stochastic random time with $(\mathcal{F}_t)$-conditional hazard-rate process (λ_t). Then $\tilde{M}_t := X_t - \int_0^{t\wedge\tau}\lambda_s ds$ is a $(\mathcal{G}_t)$-martingale.*

Given the set-up in this section, a non-negative $(\mathcal{G}_t)$-adapted process (λ_t) is called a $(\mathcal{G}_t)$-*martingale intensity* process of the random time τ if $\tilde{M}_t := X_t - \int_0^{t\wedge\tau}\lambda_s ds$ is a $(\mathcal{G}_t)$-martingale.

In reduced-form credit risk models, (λ_t) is usually called the *default intensity* of the default time τ. This martingale intensity is uniquely defined on $\{t < \tau\}$.

We can also model doubly stochastic random times through a factor model with hazard rate $\lambda_t = h(\mathbf{\Psi}_t)$ (under a risk-neutral measure $\mathbb{Q}$). Here $(\mathbf{\Psi})$ is some d-dimensional process representing economic factors, which is adapted to the background filtration $(\mathcal{F}_t)$; h is a function from $\mathbb{R}^d$ to $\mathbb{R}_+$.

The risk-neutral default probability of a corporation can be estimated from credit-spread data for bonds issued by that corporation. Market quotes for CDS spreads can also be used to infer risk-neutral default probabilities.

Martingale Modelling

In a complete market, the only thing that matters for the pricing of derivative securities is the $\mathbb{Q}$-dynamics of the traded underlying assets. When building a model for pricing derivatives it is therefore a natural shortcut to model the objects of interest - such as interest rates, default times and the price processes of traded bonds - directly under some exogenously specified measure $\mathbb{Q}$. This approach is called *martingale modelling.*

Denoting by $B(t) > 0$ the default-free savings account and by $\mathcal{G}_t$ the information available to investors at time t, we have the following formula for the price at time $t \leq T$ of a security whose value at T is given by the $\mathcal{F}_t$-measurable random variable $H \geq 0$:

$$H_t = B(t)\,\mathbb{E}^{\mathbb{Q}}\left(B(T)^{-1} H|\mathcal{G}_t\right).$$

Martingale modelling ensures that the resulting model is arbitrage free. However, as pointed out in [12], it has two drawbacks. First, historical information is, to a large extent, useless in estimating model parameters, as these may change in the transition from real-world measure to equivalent martingale measure. Second, realistic models for pricing credit derivatives are typically incomplete, so that one cannot eliminate all risk by dynamic hedging. In those situations one is interested in the distribution of the remaining risk under the actual risk measure $\mathbb{P}$, so martingale modelling alone is not sufficient. In summary, the martingale-modelling approach is most suitable in situations where the market for underlying securities is reasonably liquid.

5.7.2 Pricing with Doubly Stochastic Default Times

We will follow the application set out in [12] and consider a firm whose default time is given by a doubly stochastic random time as per Definition 5.7.2. The economic background filtration represents the information generated by an arbitrage-free and complete model for non-defaultable security prices. More precisely, let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{Q})$ denote a filtered probability space, where $\mathbb{Q}$ is the equivalent martingale measure. Prices of default-free securities such as default-free bonds are $(\mathcal{F}_t)$-adapted processes. By (r_t) we denote the default-free rate of interest and $\{B_{tT}\}_{0\leq t<T<\infty} = \exp\left(-\int_t^T r_s ds\right)$ the default-free discount factor.

Let τ be the default time of some company under consideration and let $X_t = I_{\{\tau\leq t\}}$ be the associated default indicator process. We set $\mathcal{H}_t = \sigma(\{X_s : s \leq t\})$ and $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$; we assume that default is observable and that investors have access to the information contained in the background process $(\mathcal{F}_t)$, so that the information available to investors at time t is given by $\mathcal{G}_t$.

We consider a market for credit products which is liquid enough that we may use the martingale-modelling approach, and we use $\mathbb{Q}$ as the pricing measure for defaultable securities. Finally, we assume that, under $\mathbb{Q}$, the default time τ is a doubly stochastic random time with background filtration $(\mathcal{F}_t)$ and hazard-rate process (λ_t).

In the following theorem, whose proof is presented in [12], we show that valuation of payments in the event of default can be reduced to a pricing problem in a default-free security market with adjusted default-free interest rate.

Theorem 5.7.1. *Suppose that, under $\mathbb{Q}$, the default time τ is doubly stochastic with background filtration $(\mathcal{F}_t)$ and hazard-rate process (λ_t). Define $R_s := r_s + \lambda_s$. Let H be an $\mathcal{F}_T$-measurable promised payment which is made at time T if there is no default. Let Z_τ denote a recovery payment at the time of default τ, and $Z = (Z_t)_{t\geq 0}$ be an $(\mathcal{F}_t)$-adapted stochastic process. Assume that the random variables $\exp\left(-\int_t^t r_s ds\right) \times |H|$ and $\int_t^T |Z_s\lambda_s| \exp\left(-\int_t^T R_u du\right)$ are all integrable with respect to $\mathbb{Q}$. Then the following identities hold:*

$$\mathbb{E}^{\mathbb{Q}}\left(\exp\left(-\int_t^T r_s ds\right) I_{\{\tau>T\}} H|\mathcal{G}_t\right) = I_{\{\tau>t\}}\mathbb{E}^{\mathbb{Q}}\left(\exp\left(-\int_t^T R_s ds\right) H|\mathcal{F}_t\right), \quad (5.1)$$

$$\mathbb{E}^{\mathbb{Q}}\left(I_{\{\tau>t\}}\exp\left(-\int_t^\tau r_s ds\right) Z_\tau I_{\{\tau\leq T\}}|\mathcal{G}_t\right) = I_{\{\tau>t\}}\mathbb{E}^{\mathbb{Q}}\left(\int_t^T Z_s\lambda_s \exp\left(-\int_t^s R_u du\right) ds|\mathcal{F}_t\right). \quad (5.2)$$

Proof. The integrability conditions ensure that all conditional expectations are well defined. We start with the pricing formula 5.1 for the vulnerable claim. Define the $\mathcal{F}_T$-measurable random variable $\tilde{H} := \exp\left(-\int_t^T r_s ds\right) H$. We obtain, using a result in [12] on modelling under conditional expectations of additional information, that

$$\mathbb{E}^{\mathbb{Q}}\left(\tilde{H} I_{\{\tau>T\}}|\mathcal{G}_t\right) = I_{\{\tau>t\}}\mathbb{E}^{\mathbb{Q}}\left(\exp\left(-\left(\Lambda_T - \Lambda_t\right)\right)\tilde{H}|\mathcal{F}_t\right).$$

Using the relation $\Lambda_T - \Lambda_t = \int_t^T \lambda_s ds$ and the definition of $\tilde{H}$, we immediately obtain that the right-hand side equals $I_{\{\tau>t\}}\mathbb{E}^{\mathbb{Q}}\left(\exp\left(-\int_t^T R_s ds\right) H|\mathcal{F}_t\right)$.

Next we turn to 5.2. We obtain from Lemma 5.7.3 that

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} & \left(I_{\{\tau>t\}}\exp\left(-\int_t^\tau r_s ds\right) Z_\tau I_{\{\tau\leq T\}}|\mathcal{G}_t\right) \\ = & \; I_{\{\tau>t\}}\frac{\mathbb{E}^{\mathbb{Q}}\left(I_{\{\tau>t\}}\exp\left(-\int_t^\tau r_s ds\right) Z_\tau I_{\{\tau\leq T\}}|\mathcal{F}_t\right)}{P\left(\tau>t|\mathcal{F}_t\right)}. \end{aligned} \quad (5.3)$$

Now note that

$$P\left(\tau\leq t|\mathcal{F}_T\right) = 1 - \exp\left(-\int_0^t \lambda_s ds\right),$$

so the conditional density of τ given $\mathcal{F}_T$ equals $f_{\tau|\mathcal{F}_T}(t) = \lambda_t \exp\left(-\int_0^t \gamma_s ds\right)$. Hence

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}} \left(I_{\{\tau>t\}} \exp\left(-\int_t^\tau r_s ds\right) Z_\tau I_{\{\tau\leq T\}} | \mathcal{F}_T \right) \\
= &\int_t^T \exp\left(-\int_t^s r_u du\right) Z_s \lambda_s \exp\left(-\int_0^s \lambda_u du\right) ds \\
= &\exp\left(-\int_0^t \lambda_u du\right) \int_t^T Z_s \lambda_s \exp\left(-\int_t^s R_u du\right) ds.
\end{aligned}$$

Therefore we obtain, using iterated conditional expectations, that

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}} \left(I_{\{\tau>t\}} \exp\left(-\int_t^\tau r_s ds\right) Z_\tau I_{\{\tau\leq T\}} | \mathcal{F}_t \right) \\
= &\exp\left(-\int_0^t \lambda_u du\right) \mathbb{E}^{\mathbb{Q}} \left(\int_t^T Z_s \lambda_s \exp\left(-\int_t^s R_u du\right) ds | \mathcal{F}_t \right)
\end{aligned}$$

the identity 5.2 follows because of 5.3.
□

5.7.3 Affine Models

We now turn our focus to ways of evaluating the conditional expectations of equations 5.1 and 5.2. In most models where default is modelled by a doubly stochastic random time, (r_t) and (λ_t) are modelled as functions of some p-dimensional Markovian state variable process $(\mathbf{\Psi}_t)$ with state price given by the domain $D \subset \mathbb{R}^p$. In this set-up $R_t := r_t + \lambda_t$ is of the form $R_t = R(\mathbf{\Psi}_t)$ for some function $R : D \subseteq \mathbb{R}^p \to \mathbb{R}_+$, and the natural background filtration is given by $(\mathcal{F}_t) = \sigma(\{\mathbf{\Psi}_s : s \leq t\})$. Hence we have to compute conditional expectations of the form $\mathbb{E}\left(\exp\left(-\int_t^T R(\mathbf{\Psi}_s) ds\right) g(\mathbf{\Psi}_T) | \mathcal{F}_t\right)$ for a generic $g : D \subset \mathbb{R}^p \to \mathbb{R}_+$. Since $(\mathbf{\Psi}_t)$ is a Markov process, this conditional expectation is given by some function $f(t, \mathbf{\Psi}_t)$ of time and current value $\mathbf{\Psi}_t$ of the state variable process.

Defining the economic factor process $(\mathbf{\Psi}_t)$ is thus a crucial part of the model, and one given a fair amount of attention in the literature. Where it is assumed that (Ψ_t) is given by a one-dimensional diffusion process, it is the unique solution of the SDE

$$d\Psi_t = \mu(\Psi_t) dt + \sigma(\Psi_t) dW_t, \;\; \Psi_0 = \psi \in D, \tag{5.4}$$

with the state space given by the domain $D \subset \mathbb{R}$. Here (W_t) is a standard, one-dimensional Brownian motion on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t))$, and μ and σ are continuous functions from D to $\mathbb{R}$, respectively $\mathbb{R}_+$.

Another very popular affine model is the square-root diffusion model proposed by [6] as a model of the short rate of interest. In this model (Ψ_t) is given by the solution of the SDE

$$d\Psi_t = \kappa\left(\bar{\theta} - \Psi_t\right)dt + \sigma\sqrt{\Psi_t}dW_t,\ \ \Psi_0 = \psi > 0$$

for parameters κ, $\bar{\theta}$, $\sigma > 0$ and state space $D = [0, \infty)$. Here (Ψ_t) is a mean reverting process: if Ψ_t deviates from the mean-reversion level $\bar{\theta}$, the process is pulled back towards $\bar{\theta}$. Moreover, if the mean reversion is sufficiently strong relative to the volatility, the trajectories never reach zero.

We also briefly discuss an extension of the basic model 5.4, where the economic factor process (Ψ_t) follows a diffusion with jumps. Adding jumps to the dynamics of (Ψ_t) is considered consistent with phenomena observed in practice, in addition to which it provides more flexibility for modelling default correlations in models with conditionally independent defaults. Here we assume that (Ψ_t) is the unique solution of the SDE

$$d\Psi_t = \mu\left(\Psi_t\right)dt + \sigma\left(\Psi_t\right)dW_t + dZ_t,\ \ \Psi_0 = \psi \in D. \tag{5.5}$$

The process (Z_t) is a pure jump process whose jump intensity at time t is equal to $\lambda^Z\left(\Psi_t\right)$ for some function $\lambda^Z : D \to \mathbb{R}_+$ and whose jump-size distribution has degrees of freedom v on $\mathbb{R}$. Intuitively this means that given the trajectory $\left(\Psi_t\left(\omega\right)\right)_{t\geq 0}$ of the factor process, (Z_t) jumps at the jump times of an inhomogeneous Poisson process with time-varying intensity $\lambda^Z\left(t, \Psi_t\right)$; the size of the jumps has degrees of freedom v.

We look at the example of a jump-diffusion model for CDO pricing given in [9]. There the dynamics of (Ψ_t) are given by

$$d\Psi_t = \kappa\left(\bar{\theta} - \Psi_t\right)dt + \sigma\sqrt{\Psi_t}DW_t + dZ_t, \tag{5.6}$$

for parameters κ, $\bar{\theta}$, $\sigma > 0$. The jump process (Z_t) has a constant jump intensity greater than zero and exponentially distributed jump sizes with parameter $1/\mu$. Following [9], this model is sometimes called a *basic affine jump diffusion* model. Note that these assumptions imply that the mean v is equal to μ and that v has support $[0, \infty)$, so that (Ψ_t) has only upward jumps. Hence the existence of a solution to 5.6 follows from the existence of solutions in the pure diffusion case.

5.7.4 Conditionally Independent Defaults

Introduction

The simplest reduced-form models for portfolio credit risk are models with conditionally independent defaults. In this class default times are independent given the realization of some observable economic background process, much like the factor models discussed earlier in this chapter as well as in Chapter 4. These models are also essentially an extension of the static Bernoulli mixture models introduced earlier in this chapter.

Throughout our analysis we restrict ourselves to models without simultaneous defaults. As we are working with continuous-time models, this assumption is realistic.

As in the previous sections $(\mathcal{F}_t)$ represents our background filtration, typically generated by some observable process $(\mathbf{\Psi}_t)$ representing economic factors. Moreover, we introduce the filtrations $\{\mathcal{H}_t^i\}$, $1 \le i \le n$, $(\mathcal{H}_t)$ and $(\mathcal{G}_t)$ by

$$\mathcal{H}_t^i = \sigma\left(\{X_{s,i} : s \le t\}\right), \mathcal{H}_t = \mathcal{H}_t^i \vee \cdots \vee \mathcal{H}_t^n \text{ and } \mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t.$$

Here $\{\mathcal{H}_t^i\}$ is the filtration generated by default observation for obligor i alone; $(\mathcal{H}_t)$ is the filtration generated by default observation for all the obligors; $(\mathcal{G}_t)$ contains default information for all the obligors and observable background information and thus represents the information available to investors at time t. Often $(\mathcal{H}_t)$ is called the *internal filtration* generated by the default times τ_i, $1 \le i \le n$.

Conditionally Independent Default Times

Although we have already been using the term, let us now formalize the definition of conditionally independent default times, as laid out in [12].

Definition 5.7.3. *Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with background filtration $(\mathcal{F}_t)$ and random default times $\tau_1, \ldots, \tau_n$, the τ_i are conditionally independent doubly stochastic random times if*

- *each of the τ_i is a doubly stochastic random time with background filtration $(\mathcal{F}_t)$ and $(\mathcal{F}_t)$-conditional hazard-rate process $(\lambda_{t,i})$; and*
- *the random variables $\tau_1, \ldots, \tau_n$ are conditionally independent given $\mathcal{F}_\infty$, i.e. we have, for all $t_1, \ldots, t_n > 0$*

$$P\left(\tau_1 \le t_1, \ldots, \tau_n \le t_n | \mathcal{F}_\infty\right) = \prod_{i=1}^{n} P\left(\tau_i \le t_i | \mathcal{F}_\infty\right).$$

The following lemma, stated without proof in [12], gives properties of the first default time T_1.

Lemma 5.7.4. *Let $\tau_1, \ldots, \tau_n$ be conditionally independent doubly stochastic random times with hazard-rate processes $(\lambda_{t,1}), \ldots, (\lambda_{t,n})$. Then T_1 is a doubly stochastic random time with $(\mathcal{F}_t)$-conditional hazard-rate process $\bar{\lambda}_t := \sum_{i=1}^{n} \lambda_{t,1}$, $t \ge 0$.*

Note that we have already seen this concept while we were studying the pure death process as the continuous analogue of the mixture binomial model. However, here we have more flexibility due to the fact that hazard rates are not required to be identical.

The following proposition shows that, for conditionally independent defaults, martingale intensities and hazard rates coincide.

Proposition 5.7.3. *Let $\tau_1, \dots \tau_n$ be conditionally independent doubly stochastic random times with hazard rate processes $(\lambda_{t,1}), \dots, (\lambda_{t,n})$. Then the process $\tilde{M}_{t,i} := X_{t,i} - \int_0^{t\wedge\tau_i} \lambda_{s,i} ds$ is a $(\mathcal{G}_t)$-martingale, with $(\mathcal{G}_t)$ representing the information available to investors at time t.*

Suppose that $\tau_1, \dots, \tau_n$ are conditionally independent doubly stochastic random times. Consider a single-name credit product with maturity T whose pay-off H depends only on the default history of firm i and on the evolution of default-free security prices. This product is thus $\mathcal{G}_T^i$-measurable. A typical example is a vulnerable claim of the form $H = I_{\{\tau_i > T\}} V$ for an $\mathcal{F}_T$-measurable random variable V. The conditional independence of τ_i and τ_j for $i \neq j$ means that default information of obligor $j \neq i$ is of no use in predicting the default of obligor i. This enable us to say that

$$\mathbb{E}^{\mathbb{Q}}\left(\exp\left(-\int_t^T r_s ds\right) H|\mathcal{G}_t^i\right) = \mathbb{E}^{\mathbb{Q}}\left(\exp\left(-\int_t^T r_s ds\right) H|\mathcal{G}_t\right),\ t \leq T,$$

where (r_t) is the $\mathcal{F}_t$-adapted default-free short rate. Hence pricing formulas for single-name credit products obtained in a single-firm model with doubly stochastic default time remain valid in a portfolio model with conditionally independent default times. If we go beyond conditional independence this is no longer true.

In most models with conditionally independent defaults, hazard rates are modelled as linear combinations of independent affine diffusions, possibly with jumps. A typical model is as follows:

$$\lambda_{t,i} = \lambda_{i0} + \sum_{j=1}^{p} \lambda_{ij} \Psi_{t,j}^{syst} + \Psi_{t,i}^{id},\ 1 \leq i \leq n. \tag{5.7}$$

Here $\left(\Psi_{t,j}^{syst}\right)$, $1 \leq j \leq p$, and $\left(\Psi_{t,i}^{id}\right)$, $1 \leq i \leq n$, are basic affine jump diffusions as in 5.5. The factor weights λ_{ij}, $0 \leq j \leq p$, are non-negative constants. Obviously, $\left(\mathbf{\Psi}_t^{syst}\right)$ represents the common or systematic factors, whereas $\left(\Psi_{t,i}^{id}\right)$ is an idiosyncratic factor process affecting only the hazard rate of obligor i. Note that the weight of the idiosyncratic factor can be incorporated into the parameters of the dynamics of $\left(\Psi_t^{id}\right)$, so we do not need an extra factor weight. Throughout this section we assume that the background filtration is generated by $\left(\mathbf{\Psi}_t^{syst}\right)$ and $\left(\Psi_{t,i}^{id}\right)$, $1 \leq i \leq n$. In practical applications of the model, the current value of these processes is derived from observed prices of defaultable bonds.

However, it is often claimed that the default correlation values (i.e. Pearson correlation coefficient) which can be attained in models with conditionally independent defaults are too low compared with empirical default correlations.

An Application of Conditionally Independent Default Times to First-to-Default Swaps

Finally, as an application we review an example in [12] which involves the pricing of first-to-default swaps in models with conditionally independent defaults. We consider a portfolio of n firms. Premium payments on the swap are due at N points in time $0 < t_1 < \ldots < t_N =: T$. Provided that the time of first default $T_1 > t_i$, the premium at time t_i is of the form $p\,(t_i - t_{i-1})$ for some spread p; at T_1 premium payments stop. For simplicity we neglect accrued premium payments at time of default. The default payment occurs at time T_1 provided $T_1 < T$.

We assume that the payment depends on the identity ξ_1 of the first defaulting firm, where $\xi_m \in \{1, \ldots, n\}$ denotes the identity of the m-th firm to default, but is otherwise deterministic. In other words, there are constants $l_1, \ldots, l_n$ such that the default payment is equal to l_i if $T_1 < T$ and $\xi_1 = i$. As usual, the fair spread $\tilde{p}$ of the swap is the value of p such that at $t = 0$ the default payment leg and the premium payment leg have the same value.

Since, in practice, first-to-default swaps are always priced relative to traded single-name CDSs, it is natural to adopt the martingale-modelling approach. We assume that under the equivalent martingale measure $\mathbb{Q}$ the default times τ_i are conditionally independent doubly stochastic random times with hazard rates of the form 5.7. Further, the risk-free short rate (r_t) is also assumed to be of the form 5.7. In this set-up, for generic swap spread p the value of the premium payment equals

$$V^{prem} = \sum_{i=1}^{N} \mathbb{E}^{\mathbb{Q}} \left(\exp\left(-\int_0^{t_i} r_s ds \right) I_{\{T_1 > t_i\}} \right) p\,(t_i - t_{i-1}) \,.$$

Using Theorem 5.7.1 and Lemma 5.7.4 we get

$$\mathbb{E}^{\mathbb{Q}} \left(\exp\left(-\int_0^{t_n} r_s ds \right) I_{\{T_1 > t_n\}} \right) = \mathbb{E}^{\mathbb{Q}} \left(\exp\left(-\int_0^{t_n} \left(r_s + \sum_{i=1}^{n} \lambda_{s,i} \right) ds \right) \right).$$

For hazard rates and a risk-free short rate of the form 5.7 this can be expressed as a product of expectations of the form $\mathbb{E}^{\mathbb{Q}} \left(\exp\left(-c \int_0^{t_i} \Psi_s ds \right) \right)$ for some constant c and a one-dimensional affine jump diffusion (Ψ_t). So the premium payments can be computed using the methods developed for the affine models.

Next we turn to the default payments. We have

$$V^{def} = \sum_{i=1}^{n} l_i \mathbb{E}^{\mathbb{Q}} \left(\exp\left(-\int_0^{T_1} r_s ds \right) I_{\{T_1 \leq T\}} I_{\{\xi_1 = i\}} \right).$$

We begin by computing

$$\mathbb{E}^{\mathbb{Q}} \left(\exp\left(-\int_0^{T_1} r_s ds \right) I_{\{T_1 \leq T\}} I_{\{\xi_1 = i\}} | \mathcal{F}_\infty \right).$$

Conditioning on T_1 we obtain that this equals

$$\int_0^T \exp\left(-\int_0^t r(s)\,ds\right) Q\left(\xi_1 = i | T_1 = t, \mathcal{F}_\infty\right) f^{\mathbb{Q}}_{T_1|\mathcal{F}_\infty}(t)\,dt, \tag{5.8}$$

where $f^{\mathbb{Q}}_{T_1|\mathcal{F}_\infty}(t)$ is the $\mathbb{Q}$-density of T_1 given $\mathcal{F}_\infty$. By Lemma 5.7.4 we know that

$$f^{\mathbb{Q}}_{T_1|\mathcal{F}_\infty}(t) = \bar{\lambda}(t)\exp\left(-\int_0^t \bar{\lambda}(t)\,ds\right).$$

Moreover, it can be shown that

$$Q\left(\xi_1 = i | T_1 = t, \mathcal{F}_\infty\right) = \lambda_i(t)/\bar{\lambda}(t).$$

Hence 5.8 equals $\int_0^T \lambda_i(t)\exp\left(-\int_0^t \left(r(s)+\bar{\lambda}(s)\right)ds\right)dt$. To compute the value of V^{def} we thus have to compute $\mathbb{E}^{\mathbb{Q}}\left(\int_0^T \lambda_{t,i}\exp\left(-\int_0^t r_s + \bar{\lambda}_s ds\right)dt\right)$. If the default payments l_i are all identical, the first-to-default swap can be priced like a single-name CDS, with the hazard rate of the default time given by $\left(\bar{\lambda}_t\right)$; this follows immediately from Lemma 5.7.4.

In certain special cases higher-order default swaps can be evaluated analytically. However, in most cases that are practically relevant one has to use Monte Carlo simulation.

5.7.5 Default Contagion in Reduced-Form Models

Default Contagion and Default Dependence

A more sophisticated model for dependent defaults includes models with interacting intensities. A common feature of these models is the presence of a default contagion, where the conditional default probability of a non-defaulted firms jump (usually upwards) given the additional information that some other firm has defaulted. This would happen through the jumps in their default intensities at the default times of other firms in the portfolio. Modelling default contagion might help to explain the clustering of defaults around economic recessions.

Let us denote the *ordered default times* by $T_0 < T_1 < \cdots < T_n$, where $T_0 = 0$ and $T_m = \min\{\tau_i : \tau_i > T_{m-1}, 1 \le i \le n\}$ for $1 \le m \le n$. The identity of the firm defaulting at time is denoted by T_m, where $\xi_m \in \{1,\ldots,n\}$. We can then set

$$A_m = \{1 \le i \le n : X_i(T_m) = 0\} = \{1,\ldots,n\}\backslash\{\xi_1,\ldots,\xi_m\}$$

to represent the set of non-defaulted firms immediately after time T_m.

Martingale intensities. We start with a general result which characterizes the martingale default intensities of dependent default times. In specifying the filtration we use, we assume that investors only have access to the default history of firms in the portfolio under consideration, i.e. we are interested in martingale properties with respect to the internal filtration $(\mathcal{H}_t)$. Note that $\mathcal{H}_t$ can be described as $\mathcal{H}_t = \sigma\left(\{(T_N, \xi_N) : 1 \le j \le N\}\right)$, with ξ_i, $1 \le i \le n$ being the identity of the i-th to default firm. This coincides with the general abstract definition of the σ-algebra of events observable up to some stopping time.

The following theorem on the martingale default intensities is given without proof in [12]:

Theorem 5.7.2. *Consider default times $\tau_1, \ldots, \tau_n$ and denote by $(\mathcal{H}_t)$ the corresponding internal filtration. Suppose that for every $0 \le m \le n-1$ and every $i \in \{1, \ldots, n\}$ there is a random mapping $g_i^{(m)} : \Omega \times \mathbb{R}_+ \to \mathbb{R}_+$ measurable with respect to the σ-algebra $\mathcal{H}_{T_m} \otimes \mathcal{B}(\mathbb{R}_+)$, such that*

$$P\left(T_{m+1} - T_m \le s, \xi_{m+1} = i | \mathcal{H}_{T_m}\right)(\omega) = \int_0^s g_i^{(m)}(\omega, u)\, du,\ 1 \le i \le n.$$

Then the martingale default intensity of default indicator $(X_{t,i})$ with respect to $(\mathcal{H}_t)$ is given by

$$\lambda_{t,i}(\omega) = \frac{g_i^{(m)}(\omega, t - T_m)}{P\left(T_{m+1} > t | \mathcal{H}_{t_m}\right)(\omega)},\ t_m < t \le T_{m+1}. \tag{5.9}$$

The form 5.9 for the martingale intensity is quite natural. If investors observe only past and present defaults, they obtain significant new information only at time points $T_1(\omega), \ldots, T_m(\omega)$. Hence we expect the martingale default intensity $(\lambda_{t,i})$ of some firm $i \in A_m$ (a surviving firm) to evolve in a deterministic fashion for $t \in (T_m, T_{m+1}]$ and to change with the random arrival of new information at T_{m+1}.

We need to note that in general the marginal hazard rates $\lambda_i(t) \ne \lambda_{t,i}$, $t > 0$, where

$$\lambda_i(t) = \lim_{h \to 0} \frac{1}{h} P\left(\tau_i \le t + h | \tau_i > t\right).$$

Hence $\lambda_t(t)$ gives the instantaneous default probability of firm i given $\tau_i > t$, whereas $\lambda_{t,i}$ gives the instantaneous default probability given $\tau_i > t$ *and* the default history of all other firms in the portfolio. With (conditionally) independent defaults on the other hand, the additional information about the default history of firms $j \ne i$ in the portfolio is of no use in predicting the default time of firm i, and we have $\lambda_{t,i} = \lambda_i(t)$.

Conditional survival functions. Let $T_m \le t < T_{m+1}$ for some $0 \le m \le n-1$. We would like to compute the conditional survival function $S_{\tau_i|\mathcal{H}_t}$ for some firm $i \in A_m$. To simplify the notation we assume from now on that the indices have been permuted in such a way that $A_m^c = \{1, \ldots, m\}$ and $A_m = \{m+1, \ldots, n\}$, i.e. the defaulted firms correspond to the first m firms in our index set. Put $\tilde{\tau}_1 = (\tau_1, \ldots, \tau_m)'$ and $\tilde{\tau}_2 = (\tau_{m+1}, \ldots, \tau_n)'$. As an intermediate step we consider $S_{\tilde{\tau}_2|\tilde{\tau}_1}(t_1, \ldots, t_{n-m}|\tilde{\tau}_1)$, the conditional survival function of the last $n-m$ firms given the vector of the default times of the first m firms.

We have the following lemma, whose proof is given in [12].

Lemma 5.7.5. *Assume that the vector* $(\tau_1, \ldots, \tau_n)$ *has a density. Then*

$$S_{\tilde{\tau}_2|\tilde{\tau}_1}(t_1, \ldots, t_{n-m}|\tau_1, \ldots, \tau_m) = \frac{\frac{\partial^m}{\partial t_1 \ldots \partial t_m} S(\tau_1, \ldots, \tau_m, t_1, \ldots, t_{n-m})}{\frac{\partial^m}{\partial t_1 \ldots \partial t_m} S(\tau_1, \ldots, \tau_m, 0, \ldots, 0)}.$$

Finally, we turn to the conditional survival function $S_{\tau_i|\mathcal{H}_t}$. At time t default information consists of the vector $\tilde{\tau}_1$ of the default times of the firms from A_m^c and of the atom $B := \{\tilde{\tau}_2 > t\}$. Hence we have, for $i \in \{m+1, \ldots, n\}$ and $T \geq t$, using Lemma 5.7.5, that

$$\begin{aligned} S_{\tau_i|\mathcal{H}_t}(T) &= P(\tau_i > T|B, \tilde{\tau}_1) \\ &= \frac{P(\tau_i > T, \tilde{\tau}_2 > t|\tilde{\tau}_1)}{P(\tilde{\tau}_2 > t|\tilde{\tau}_1)} \\ &= \frac{\frac{\partial^m}{\partial t_1 \ldots \partial t_m} S(\tau_1, \ldots, \tau_m, t, \ldots, T, \ldots, t)}{\frac{\partial^m}{\partial t_1 \ldots \partial t_m} S(\tau_1, \ldots, \tau_m, t, \cdots, t \ldots, t)}. \end{aligned}$$

Application of Default Contagion to Pricing a First-to-Default Swap

Our example will follow the one laid out in [12]. In line with the common approach we assume that the risk-free short rate $r(t) \geq 0$ is deterministic. Again $B(t) = \exp\left(\int_0^t r(s)\, ds\right)$ denotes the default-free savings account. Premiums are due at times $0 < t_1 < \ldots < t_N = T$, provided that no default has yet occurred. If $T_1 < T$ and, specifically if $\xi_1 = i$, there is a default payment equal to the constant l_i. In this set-up the value at time $t = 0$ of the default payment leg, paid by the CDS seller, equals

$$V^{def} = \sum_{i=1}^{n} l_i \mathbb{E}^{\mathbb{Q}} \left(B(\tau_i)^{-1} I_{\{\tau_i = T_1\}} I_{\{\tau_i \leq T\}} \right).$$

If we condition on the time to i-th default τ_i, we get, for each term of this sum,

$$\mathbb{E}^{\mathbb{Q}} \left(B(\tau_i)^{-1} I_{\{T_1 = \tau_i\}} I_{\{\tau_i \leq T\}} \right) = \int_0^T B(t)^{-1} Q(\tau_i = T_1|\tau_i = t) f_i(t)\, dt,$$

where $f_i(t)$ is the marginal density of τ_i. Now Lemma 5.7.5 yields

$$Q(\tau_i = T_1|\tau_i = t) = Q(\tau_j > t \text{ for all } j \neq i|\tau_i = t) = -\frac{1}{f_i(t)} \frac{\partial S}{\partial t_i}(t, \ldots, t),$$

from which we obtain

$$V^{def} = -\sum_{i=1}^{n} l_i \int_0^T B(t)^{-1} \frac{\partial S}{\partial t_i}(t, \ldots, t)\, dt.$$

Note that, by definition, $Q(T_1 > t) = S(t_i, \ldots, t_i)$; hence the value at $t = 0$ of the premium payments (assuming a generic swap spread p) is given by

$$V^{prem} = p \sum_{i=1}^{N} B(t_i)^{-1}(t_i - t_{i-1}) S(t_i, \ldots t_i).$$

Default Contagion and Interacting Intensities

In copula models the dependence structure of the default times is exogenously specified; the form of the resulting default contagion can then be deduced from the model. In models with interacting intensities, on the other hand, the impact of defaults on the default intensities of surviving firms is exogenously specified; the joint distribution of the default times is then endogenously derived. The main drawback of models with interacting intensities is the fact that the marginal distribution of individual default times is typically not available in closed form.

In models with interacting intensities the martingale default intensity of firm i belonging to a given portfolio is given by an exogenously specified function $\lambda_i(t, \mathbf{Y}_t)$ of time and the current state $\mathbf{Y}_t$ of the portfolio. The dependence on the current state of the portfolio is the major innovation of the model; in this way the counterparty credit risk can be modelled explicitly. It is straightforward to extend the model to stochastic default intensities of the form $\lambda_i(\mathbf{\Psi_t}, \mathbf{Y}_t)$ for some observable background process $(\mathbf{\Psi}_t)$.

It is convenient to model the default indicator process $(\mathbf{X}_t)$ in a model with interacting intensities as a time-inhomogeneous continuous-time Markov chain.

Continuous-time Markov chains. A time-inhomogeneous continuous-time Markov chain $(\mathbf{Y}_t)$ on a finite space S is characterized by non-negative and bounded *transition rate functions* $\lambda(t, \mathbf{y}, \mathbf{x})$, $\mathbf{x}$, $\mathbf{y} \in S$, $\mathbf{x} \neq \mathbf{y}$, $t \geq 0$, with the following interpretation. Fix $t \geq 0$ and let $T := \inf\{s \geq t : \mathbf{Y}_s \neq \mathbf{Y}_t\}$, i.e. T gives the time of the first jump of the chain after time t. Define, for $\mathbf{y} \in S$,

$$\lambda(t, \mathbf{y}, \mathbf{y}) := -\sum_{\mathbf{x} \in S, \mathbf{x} \neq \mathbf{y}} \lambda(t, \mathbf{y}, \mathbf{x}), \ t \geq 0,$$

and denote by $\mathcal{H}_t := \sigma(\{\mathbf{Y}_s : s \leq t\})$ the internal filtration of the chain. Then

$$P(T > s | \mathcal{H}_t) = P(T > s | \mathbf{Y}_t) = \exp\left(\int_t^s \lambda(u, \mathbf{Y}_t, \mathbf{Y}_t)\, du\right), \ s \geq t.$$

In the special case of a time-homogeneous Markov chain where the transition rate functions are independent of time, given $\mathcal{H}_t$, the random variable $(T - t)$ (the waiting time for the next jump after time t) is thus $Exp(-\lambda(\mathbf{Y}_t, \mathbf{Y}_t))$ distributed.

Construction of interacting intensities via Markov chains. Now we turn to the formal construction of models with interacting intensities. Set $S := \{0,1\}^n$ and define, for $\mathbf{x} \in S$ and $i \in \{1, \ldots, n\}$, the state $\mathbf{x}^i$ by $x_j^i = x_j$ for $j \in \{1, \ldots, n\} \backslash i$ and $x_i^i = 1 - x_i$, i.e. $\mathbf{x}^i$ is constructed from $\mathbf{x}$ by flipping the ith coordinate. Given non-negative and bounded functions $\lambda_i : [0, \infty) \times S \to \mathbb{R}_+$, for $1 \leq i \leq n$, we define the default indicator process $(\mathbf{X}_t)$ as a time-inhomogeneous continuous-time Markov chain with state space S and transition rates

$$\lambda(t, \mathbf{x}, \mathbf{y}) = \begin{cases} I_{\{x_i=0\}} \lambda_i(t, \mathbf{x}), & \text{if } \mathbf{y} = \mathbf{x}^i \text{ for some } i \in \{1, \ldots, n\}, \\ 0, & \text{otherwise.} \end{cases}$$

The above relation implies that the chain can jump only to those neighbouring states $\mathbf{X}_t^i$ that differ from the current state $\mathbf{X}_t$ by exactly one default. There are no joint defaults. If $X_t^i = 0$, the probability that firm i defaults in the small time interval $[t, t+h)$, i.e. the probability of jumping to the neighbouring state $\mathbf{X}_t^i$ in $[t, t+h)$, is approximately equal to $h\lambda_i(t, \mathbf{X}_t)$.

The definition of $(\mathbf{X}_t)$ suggests that $(\lambda_i(t, \mathbf{X}_t))$ is the martingale default intensity of firm i. The transition probabilities of the chain $(\mathbf{X}_t)$ are given by

$$p(t, s, \mathbf{x}, \mathbf{y}) := P(\mathbf{X}_s = \mathbf{y} | \mathbf{X}_t = \mathbf{x}), \mathbf{x}, \mathbf{y} \in S, \ 0 \leq t \leq s < \infty.$$

Models for the default intensities. The functions $\lambda_i(t, \mathbf{x})$ are an essential ingredient in any model with interacting intensities. We look at several specifications proposed in the literature, such as [20], that study a model with stochastic background process (Ψ_t), but with a restriction to special form of interacting intensities called the *primary-secondary framework*. This concept is along similar principles to those in the contagion models that we have examined earlier in the chapter.

In this framework firms are divided into two classes: primary and secondary. The default intensity of primary firms depends only on (Ψ_t); the default intensity of secondary firms depends on (Ψ_t) and on the default state of the primary firms. This simplifying assumption facilitates the mathematical analysis of the model.

Here is a specific example from the paper with $n = 2$. The background process (Ψ_t) is the short rate of interest (r_t). The default intensities are then given by

$$\lambda_1(r_t, \mathbf{X}_t) = a_{10} + a_{11} r_t \text{ and } \lambda_2(r_t, \mathbf{X}_t) = a_{20} + a_{21} r_t + a_{22} I_{\{X_{t,1}=1\}},$$

so company one is a primary firm and company two is a secondary firm. However, under the primary-secondary framework, cyclical default dependence, such as a situation where the default intensity of firm i is affected by the default of firm $j \neq i$, and vice versa, cannot be modeled.

Similarly, [34] sets out a model where the whole portfolio enters an "enhanced risk state" after the first default. Default intensities of the form

$$\lambda_i(t, \mathbf{X}_t) = a_0 + a_1 I_{\{\mathbf{X} \neq \mathbf{0}\}}, \ i \in \{1, \ldots, n\}, \ a_0, a_1 > 0$$

are used. Hence, at the first default time T_1, the default intensities of the surviving firms jump from a_0 to $a_0 + a_1$. The assumption of identical default intensities for all firms implies that the portfolio is homogeneous, i.e. that the default times $(\tau_1, \ldots, \tau_n)$ are exchangeable.

5.8 Information-Based Approaches Using Brownian Bridges

5.8.1 Some More Mathematical Concepts

In the above sections we have already seen the background information processes being explicitly specified. In [2] the credit risk model is build around the market filtration modelling. This is the paper that we will review in this section.

The background information process is assumed to be generated by one or more independent market information processes. Each such information process carries partial information about the values of the market factors that determine future cashflows.

This framework satisfies an overall dynamic consistency condition that makes it suitable as a basis for practical modelling situations where frequent recalibration may be necessary.

Here, the reduced-form credit risk models were taken as a starting point. However, in [2] it was recognised that crucial information is lost with the use of these models. Namely, they do not adequately take into account the fact that defaults are typically associated directly with a failure in the delivery of a contractually agreed cash flow. Reduced-form models also do not allow for the modelling of the rise and fall of credit spreads.

The key assumption made about cashflows of the underlying instruments is that partial information about each such cashflow is available to market participants at earlier times. However, a Gaussian noise process, conditioned to vanish at the time of the required cashflow, is disguising this information.

We specify the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ with filtration $\{\mathcal{F}_t\}_{0\leq t<\infty}$. The probability measure $\mathbb{Q}$ is the risk-neutral measure, and $\{\mathcal{F}_t\}$ is market filtration. All asset-price processes and other information-providing processes accessible to market participants are adapted to $\{\mathcal{F}_t\}$.

Absence of arbitrage is assumed as well as the existence of a pricing kernel, which ensure the existence of a unique risk-neutral measure, even though the market may be incomplete. Again we use the default-free discount bond system in its general form $\{B_{tT}\}_{0\leq t<T<\infty} = \exp\left(-\int_t^T r_s ds\right)$. It follows that if the integrable random variable H_T represents a cash flow occurring at T, then its value H_t at any earlier time t is given by

$$H_t = B_{tT}\mathbb{E}\left[H_T|\mathcal{F}_t\right].$$

Of course, being a random variable, the true value of H_T is not known until time T; that is, H_T is $\mathcal{F}_T$-measurable, but not $\mathcal{F}_t$-measurable for $t < T$. However we assume that partial information regarding the value of H_T is available at earlier times. This information will in general be imperfect. The paper assumes the following form for the $\{\mathcal{F}_t\}$-adapted process:

$$\xi_t = \sigma H_T t + \beta_{tT}.$$

The process $\{\xi_t\}$ is referred to as a *market information process*. The process $\{\beta_{tT}\}_{0\le t\le T}$ is a *standard Brownian bridge* on the time interval $[0, T]$. This means that it is a Gaussian process satisfying $\beta_{0T} = 0$ and $\beta_{TT} = 0$. In addition $\mathbb{E}[\beta_{tT}] = 0$ and $\mathbb{E}[\beta_{sT}\beta_{tT}] = s(T-t)/T$ for all s, t satisfying $0 \le s \le t \le T$. It is also assumed that $\{\beta_{tT}\}$ is independent of H_T, and thus represents pure noise. Market participants do not have direct access to $\{\beta_{tT}\}$; that is to say $\{\beta_{tT}\}$ is not assumed to be adapted to $\{\mathcal{F}_t\}$.

The motivation for the use of a bridge process to represent the noise is laid out as follows. It is assumed that initially all available market information is taken into account in the determination of the price; which includes *a priori* probabilities of any credit events. As time passes the variance of the Brownian bridge steadily increases for the first half of its trajectory, reflecting uncertainty increasing about the final outcome. Eventually, however, the variance falls to zero at the maturity of the underlying, as its true value is revealed.

The parameter σ in this model represents the rate at which the true value of H_T is revealed. Thus, if σ is low the market information process comprises a lot of noise until close to the maturity of the asset; on the other hand, if σ is high true information about H_T is revealed quickly. This enables the use of inaccessible stopping times to be avoided.

The paper assumes that the only market information available about H_T at times earlier than T comes from observations of $\{\xi_t\}$. Specifically, if we denote by $\mathcal{F}_t^\xi$ the subalgebra of $\mathcal{F}_t$ generated by $\{\xi_s\}_{0\le s\le t}$, then the simplifying assumption made is that $\{\mathcal{F}_t\} = \{\mathcal{F}_t^\xi\}$.

Our aim is to determine the price-process $\{S_{tT}\}_{0\le t\le T}$ for a credit-risky asset with payout H_T, i.e.

$$S_{tT} = B_{tT} H_{tT},$$

where H_{tT} denotes the conditional expectation of the asset payout:

$$H_{tT} = \mathbb{E}\left[H_T|\mathcal{F}_t^\xi\right].$$

5.8.2 Baskets of Credit-Risky Bonds

The first application we will examine is that of a basket of correlated bonds with various maturities. For simplification a set of digital bonds is considered and the notation of the bonds will be labelled in chronological order with respect to maturity. So H_{T_1} denotes the payoff of the bond that expires first; H_{T_2} $(T_2 \ge T_1)$ the payoff of the bond that matures after T_1; and so on.

The paper proposes to model this set of dependent random variables in terms of an underlying set of independent market factors. To achieve this let X denote the random variable associated with the payoff of the first bond: $H_{T_1} = X$. The random variable X takes on the values $\{1, 0\}$ with *a priori* probabilities $\{p, 1-p\}$. The payoff of the second bond H_{T_2} can then be represented in terms of three independent random variables: $H_{T_2} = XX_1 + (1 - X)X_0$. Here X_0 takes the values $\{1, 0\}$ with the probabilities $\{p_0, 1 - p_0\}$, and X_1 takes the values $\{1, 0\}$ with the probabilities $\{p_1, 1 - p_1\}$.

Since these random variables are independent, the *a priori* probability that the second bond does not default is $p_0 + p\,(p_1 - p_0)$. To represent the payoff of the third bond four additional independent random variables need to be introduced:

$$H_{T_3} = XX_1X_{11} + X\,(1 - X_1)\,X_{10} + (1 - X)\,X_0X_{01} + (1 - X)\,(1 - X_0)\,X_{00}.$$

The market factors $\{X_{ij}\}_{i,j=0,1}$ here take the values $\{0, 1\}$ with probabilities $\{p_{ij}, 1-p_{ij}\}$.

Extending the above representation, the payoff expression for a generic bond in the basket is:

$$H_{T_{n+1}} = \sum_{\{k_j\}=1,0} X^{\omega(k_1)} X_{k_1}^{\omega(k_2)} X_{k_1k_2}^{\omega(k_3)} \cdots X_{k_1k_2\ldots k_{n-1}}^{\omega(k_n)} X_{k_1k_2\ldots k_{n-1}k_n}.$$

Here, for any random variable X we define $X^{\omega(0)} = 1 - X$ and $X^{\omega(1)} = X$.

In general, if we have a basket of n digital bonds with arbitrary *a priori* default probabilities and correlations, we can introduce $2^n - 1$ independent digital random variables to represent the n correlated random variables associated with bond payoff.

One advantage of the decomposition into independent market factors is that the analytical tractability for pricing of the basket is retained. We have already seen a variation of this type of modelling in the contagion models with infection variables, earlier in the chapter. The idea is also not dissimilar to the concept of shocks used in the Marshall-Olkin copulas.

We also have $2^n - 1$ independent Brownian bridges to represent the noise associated with the independent market factors:

$$\xi_t^{k_1k_2\ldots k_n} = \sigma_{k_1k_2\ldots k_n} X_{k_1k_2\ldots k_n} t + \beta_{tT_{n+1}}^{k_1k_2\ldots k_n}.$$

The number of independent factors grows rapidly with the number of bonds in the portfolio. As a consequence, a market that consists of correlated bonds is in general highly incomplete. This fact provides an economic justification for the creation of products such as CDSs and CDOs that enhance the "hedgeability" of such portfolios.

A possible approach to contain the number of independent factors required in this model is to apply some level of nesting, where bonds with similar features are grouped together. The model can then be applied at multiple levels, along the same principles as the nested Archimedean copulas we investigated in an earlier chapter.

5.8.3 Homogeneous Baskets

The number of independent factors can be reduced for homogeneous baskets of digital bonds, where each bond matures at time T. Our goal is to model default correlations in the portfolio, and in particular to model the flow of market information concerning default correlation. Let us write H_T for the payoff at time T of this homogeneous portfolio, and set

$$H_T = n - X_1 - X_1X_2 - X_1X_2X_3 - \cdots - X_1X_2 \dots X_n,$$

where the random variables $\{X_j\}_{j=1,2,\dots,n}$, each taking the values $\{0,1\}$, are assumed to be independent. Thus if $X_1 = 0$, then $H_T = n$; if $X_1 = 1$ and $X_2 = 0$, then $H_T = n-1$; if $X_1 = 1$, $X_2 = 1$, and $X_3 = 0$, then $H_T = n-2$; and so on. If we use the notation $p_j = Q(X_j = 1)$ and $q_j = Q(X_j = 0)$ for $j = 1, 2, \dots, n$, then we can see that $Q(H_T = n) = q_1$, $Q(H_T = n-1) = p_1q_2$, $Q(H_T = n-2) = p_1p_2q_3$ and so on.

Resulting from this setting, we can see that if $p_1 << 1$ but $p_2, p_3, \dots, p_k$ are large, then our environment is that of low default probability and high default correlation; in other words the probability of a default occurring in the portfolio is small, but conditional on at least one default occurring, the probability of several defaults is high.

We again introduce a set of independent information processes $\{\eta_t^j\}$ defined by

$$\eta_t^j = \sigma_j X_j t + \beta_{tT}^j,$$

where $\{\sigma_j\}_{j=1,2,\dots,n}$ are parameters, and $\{\beta_{tT}^j\}_{j=1,2,\dots,n}$ are independent Brownian bridges. The market filtration $\{\mathcal{F}_t\}$ is generated collectively by $\{\eta_j\}_{j=1,2,\dots,n}$, and for the portfolio value $H_T = B_{tT}\mathbb{E}\left[H_T|\mathcal{F}_t\right]$ we have

$$H_{tT} = B_{tT}\left[n - \mathbb{E}_t\left[X_1\right] - \mathbb{E}_t\left[X_1\right]\mathbb{E}_t\left[X_n\right] - \cdots - \mathbb{E}_t\left[X_1\right]\mathbb{E}_t\left[X_2\right]\dots\mathbb{E}_t\left[X_n\right]\right]. \quad (5.10)$$

Each of the conditional expectations appearing here for X_j, $j = 1, \dots, n$, taking the values $\{0,1\}$, can be calculated as follows:

$$Q\left(X_j = x_i|\eta_t^j\right) = \frac{p_i\rho\left(\eta_t|X_j = x_i\right)}{\sum_i p_i\rho\left(\eta_t|X_j = x_i\right)},$$

where the conditional density function $\rho\left(\eta|X_j = x_i\right)$, $\eta \in R$, for the random variable η_t is defined by the relation

$$Q\left(\eta_t \leq x|X_j = x_i\right) = \int_{-\infty}^{x} \rho\left(\eta|X_j = x_i\right)d\eta.$$

We then use the fact that conditional on $X_j = x_i$ the random variable η_t is normally distributed with mean $\sigma x_i t$ and variance $t\left(T-t\right)/T$ to obtain

$$\rho\left(\eta|X_j = x_i\right) = \frac{1}{\sqrt{2\pi t\left(T-t\right)/T}}\exp\left(-\frac{1}{2}\frac{\left(\eta - \sigma x_i t\right)^2}{t\left(T-t\right)/T}\right).$$

This then leads to

$$Q\left(X_j = x_i | \eta_t^j\right) = \frac{p_i \exp\left[\frac{T}{T-t}\left(\sigma x_i \eta_t - \frac{1}{2}\sigma^2 x_i^2 t\right)\right]}{\sum_i p_i \exp\left[\frac{T}{T-t}\left(\sigma x_i \eta_t - \frac{1}{2}\sigma^2 x_i^2 t\right)\right]},$$

which can then be directly used in the expectations of equation 5.10.

The resulting dynamics for $\{H_t\}$ can then be used to describe the evolution of correlations in the portfolio. For example, if $\mathbb{E}_t[X_1]$ is low and $\mathbb{E}_t[X_2]$ is high, then the conditional probability at time t of a default at time T is small; whereas if $\mathbb{E}_t[X_1]$ were to increase suddenly, then the conditional probability of two or more defaults at time T would rise as a consequence.

5.9 Gamma Information Processes and Modelling Cumulative Losses

5.9.1 Introduction

We now look at a variation of the above concept of modelling losses in the framework of information processes. We will examine the model in [1], where the accumulation processes are modelled using Gamma bridges.

As usual, the time period $[0, T]$ is fixed. At time T a contract pays a random cashflow H_T, which is assumed to be positive and represents the terminal value of some accumulation process. The value process for H_T at $0 \leq t \leq T$ is given by $\{S_t\}$, and the filtration representing the flow of information available to market participants is represented by $\{\mathcal{F}_t\}$. The pricing measure is $\mathbb{Q}$. Then the expression for the value process is

$$S_t = B_{tT}\mathbb{E}\left[H_T | \mathcal{F}_t\right].$$

It is assumed that $\{\mathcal{F}_t\}$ is generated by an aggregate claims process $\{\xi_t\}$, where for each t the variable ξ_t represents the totality of claims known at t to be payable at T. The paper assumes that $\{\xi_t\}$ takes the form

$$\xi_t = H_T \gamma_{tT}, \tag{5.11}$$

where $\{\gamma_{tT}\}$ is a gamma bridge over $[0, T]$, independent of H_T. The claims process can thus be decomposed into the "signal" H_T and an independent "noise" $\{\gamma_{tT}\}$. This product representation of the gamma information process can be considered natural, since many properties of the Brownian bridge that hold additively have multiplicative analogues for gamma bridges. We will formally introduce the gamma bridge process in the next two sections.

5.9.2 Gamma Processes and Associated Martingales

We fix a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, where $\mathbb{Q}$ is the risk-neutral measure. By a *standard gamma process* $\{\gamma_t\}_{0\leq t<\infty}$ on $(\Omega, \mathcal{F}, \mathbb{Q})$ with growth rate m (which has units of inverse time), we mean a process with independent increments such that $\gamma_0 = 0$, and such that the random variable γ_t has a gamma distribution with mean and variance mt. In other words, writing $\mathbb{Q}[\gamma_t \in dx] = g(x)dx$, the density of γ_t is

$$g(x) = 1_{\{x>0\}} \frac{x^{mt-1}e^{-x}}{\Gamma\left[mt\right]}.$$

Here $\Gamma\left[a\right] = \int_0^\infty x^{a-1}e^{-x}dx$ is the gamma function. Using the well-known identity $\Gamma\left[a+1\right] = a\Gamma\left[a\right]$, we obtain $\mathbb{E}\left[\gamma_t\right] = mt$, justifying the interpretation of m as the *mean growth rate*.

It can be also shown that $var\left[\gamma_t\right] = mt$, and from the independent increments property it follows that $cov\left[\gamma_t, \gamma_u\right] = mt$ for $u \geq t$. Also, $\gamma_u - \gamma_t$ is gamma distributed with parameter $m\left(u-t\right)$. This implies that the increments of $\{\gamma_t\}$ have a stationary probability law in the sense that $\gamma_{u+h} - \gamma_{t+h}$ has the same distribution as $\gamma_u - \gamma_t$. In addition, $\{\gamma_t - mt\}$ and $\{\left(\gamma_t - mt\right)^2 - mt\}$ are martingales.

To add flexibility to the gamma processes, the paper introduces the "scaled" gamma processes with a separate parameter for growth rate μ and spread σ. This is a process $\{\Gamma_t\}_{0\leq t<\infty}$ with independent increments such that $\Gamma_0 = 0$ and such that Γ_t has a gamma distribution with mean μt and variance $\sigma^2 t$. Defining $m = \mu^2/\sigma^2$ and $\kappa = \sigma^2/\mu$, we have $\mu = \kappa m$ and $\sigma^2 = \kappa^2 m$. One can think of m as a "standardized growth rate", and κ as a scale. The density of Γ_t is given by

$$\mathbb{Q}\left[\Gamma_t \in dx\right] = 1_{\{x>0\}} \frac{\kappa^{-mt}x^{mt-1}e^{-x/\kappa}}{\Gamma\left[mt\right]}dx. \tag{5.12}$$

If $\{\Gamma_t\}$ is a scaled gamma process with standardized growth rate m and scale κ, then $\{\kappa^{-1}\Gamma_t\}$ is a gamma process with growth rate m.

5.9.3 Properties of a Gamma Bridge Process

Let $\{\gamma_t\}_{0\leq t<\infty}$ be a standard gamma process with growth rate m, and for fixed time horizon T define the process $\{\gamma_{tT}\}_{0\leq t<T}$ by setting $\gamma_{tT} = \gamma_t/\gamma_T$. Obviously $\gamma_{0T} = 0$ and $\gamma_{TT} = 1$. The process $\{\gamma_{tT}\}$ is referred to as the *standard gamma bridge* over $[0, T]$ associated with $\{\gamma_t\}$. It can be shown that the random variable γ_{tT} has a beta distribution, i.e. if $\mathbb{Q}\left[\gamma_{tT} \in dy\right] = f(y)dy$ then

$$f(y) = 1_{\{0<y<1\}} \frac{y^{mt-1}\left(1-y\right)^{m(T-t)-1}}{B\left[mt, m\left(T-t\right)\right]},$$

where $B\left[a,b\right] = \Gamma\left[a\right]\Gamma\left[b\right]/\Gamma\left[a+b\right] = \int_a^b y^{a-1}\left(1-y\right)^{b-1}dy$ is known as the beta function. One can deduce that $\mathbb{E}\left[\gamma_{tT}\right] = t/T$ and that $var\left[\gamma_{tT}\right] = t\left(T-t\right)/T^2\left(1+mT\right)$.

So the expectation of γ_{tT} does not depend on the growth rate m, and the variance of γ_{tT} decreases with increasing m.

A useful property proved in [1] is that if $\{\gamma_t\}_{0\leq t<\infty}$ is a standard gamma process, then the random variables γ_t/γ_T and γ_T are independent for $0 \leq t \leq T$.

The gamma process $\{\gamma_t\}$ has the Markov property, which means that for $a > 0$

$$\mathbb{Q}\left[\gamma_t < a|\gamma_s, \gamma_{s_1}, \gamma_{s_2}, \ldots, \gamma_{s_n}\right] = \mathbb{Q}\left[\gamma_t < a|\gamma_s\right],$$

for all $t \geq s \geq s_1 \geq s_2 \geq \cdots \geq s_n \geq 0$ and for all $n \geq 1$. By a similar argument it can be seen that the gamma bridge is Markovian, so that

$$\mathbb{Q}\left[\frac{\gamma_t}{\gamma_T} < a|\frac{\gamma_s}{\gamma_T}, \frac{\gamma_{s_1}}{\gamma_T}, \frac{\gamma_{s_2}}{\gamma_T}, \cdots\right] = \mathbb{Q}\left[\frac{\gamma_t}{\gamma_T} < a|\frac{\gamma_s}{\gamma_T}\right].$$

5.9.4 Valuation of Aggregate Claims

The objective is to calculate the value at t of a contract that pays H_T at T. It is assumed that H_T is strictly positive and integrable continuous random variable. The value S_t of the contract at $t \leq T$ is given by $S_t = B_{tT}\mathbb{E}\left[H_T|\mathcal{F}_t\right]$, where $\mathcal{F}_t = \sigma\left(\{\xi_s\}_{0\leq s\leq t}\right)$.

The paper shows that the aggregate claims process $\{\xi_t\}_{0\leq t\leq T}$ has the Markov property, which means that

$$\mathbb{Q}\left[\xi_t < a|\mathcal{F}_s\right] = \mathbb{Q}\left[\xi_t < a|\xi_s\right]$$

for all s, t such that $0 \leq s \leq t \leq T$. Given this and the fact that H_T is $\mathcal{F}_T$-measurable, the expression for S_t can be simplified to take the form

$$S_t = B_{tT}\mathbb{E}\left[H_T|\xi_t\right].$$

Applying the the distribution function of the gamma bridge, it can be seen that the above formula amounts to

$$S_t = B_{tT}\frac{\int_{\xi_t}^{\infty} p(x)x^{2-mT}\left(x-\xi_t\right)^{m(T-t)-1}dx}{\int_{\xi_t}^{\infty} p(x)x^{1-mT}\left(x-\xi_t\right)^{m(T-t)-1}dx},$$

where $p(x)$ is the density of H_T. The result follows by the use of the conditional density process for H_T, $\{\pi_t(x)\}$, and the Bayes formula:

$$\pi_t(x) = \frac{p(x)\,\rho\left(\xi_t|H_T = x\right)}{\int_0^{\infty} p(x)\,\rho\left(\xi_t|H_T = x\right)dx},$$

where $\rho\left(\xi_t|H_T = x\right)$ is the conditional density for ξ_t.

We are in a position to price a simple credit contract C_{tT} that pays at T an amount equal to the total loss incurred by the portfolio, in excess of some threshold K. The expression is given by

$$C_{tT} = B_{tT} \int_0^\infty (x-K)^+ \pi_t(x)\, dx = B_{tT} \frac{\int_{\xi_t}^\infty (x-K)^+ p(x)\, x^{1-mT} (x-\xi_t)^{m(T-t)-1} dx}{\int_{\xi_t}^\infty p(x)\, x^{1-mT} (x-\xi_t)^{m(T-t)-1} dx}$$

for $t < T$ and by $C_{TT} = (X_T - K)^+$. This means that once a time t has been reached such that $\xi_t \geq K$, then $C_{uT} = S_u - B_{uT}K$ for $t \leq u \leq T$; in other words when a sufficient number of credit events has occurred, the option is sure to expire in-the-money. It is straightforward to see that the pay-off function can be adjusted to accommodate a segregation of the loss allocation into tranches, hence enabling us to price CDOs.

5.9.5 Gamma-Distributed Terminal Gains

When the terminal aggregate loss H_T is gamma distributed with mean κmT and variance $\kappa^2 mT$ for some κ, [1] shows that the value process $\{S_t\}$ has a particularly simple structure. Let $\{\gamma_t\}$ be a standard gamma process with rate m and let $\{\gamma_{tT}\}$ be the associated bridge. Then H_T and $\kappa\gamma_T$ have the same distribution; but since γ_T and $\{\gamma_{tT}\}$ are independent, $\{H_T\gamma_{tT}\}$ and $\{\kappa\gamma_T\gamma_{tT}\}$ have the same probability law. Therefore $\{\xi_t\}$ has the same law as $\{\kappa\gamma_t\}$ and hence is a $\mathbb{Q}$-gamma process with scale κ and standard growth rate m.

Although the $\mathbb{Q}$-gamma process has independent increments, the cumulative gains process 5.11 has dependent increments. In particular, for the covariance of ξ_s and $\xi_t - \xi_s$ in the general case we have

$$cov\left[\xi_s, \xi_t - \xi_s\right] = \frac{ms(t-s)}{T(mT+1)} \mathbb{E}\left[H_T^2\right] - \frac{s(t-s)}{T^2} \left(\mathbb{E}\left[H_T\right]\right)^2.$$

Hence a necessary condition for independent increments is given by $\left(\mathbb{E}\left[H_T\right]\right)^2 = mTvar\left[H_T\right]$.

The value process for various claims in the $\mathbb{Q}$-gamma model can be worked out explicitly. Using the density of H_T given by 5.12 and carrying out the integration, it can be shown that the expression for the value process is:

$$S_t = B_{tT}\left(\xi_t + \kappa m (T-t)\right).$$

Next we move onto working out the value of a call-type option or a CDO tranche with a threshold K. In the $\mathbb{Q}$-gamma model we have $C_{tT} = B_{tT}\mathbb{E}\left[(\xi_T - K)^+ | \xi_t\right]$, and hence by use of the independent increments property we deduce that

$$\begin{aligned} C_{tT} &= B_{tT} \int_{(K-\xi_t)/\kappa}^\infty (\kappa z + \xi_t - K) \frac{z^{m(T-t)-1} e^{-z}}{\Gamma\left[m(T-t)\right]} dz \\ &= B_{tT} \left[\kappa \frac{\Gamma\left[m(T-t)+1, (K-\xi_t)/\kappa\right]}{\Gamma\left[m(T-t)\right]} - (K-\xi_t) \frac{\Gamma\left[m(T-t), (K-\xi_t)/\kappa\right]}{\Gamma\left[m(T-t)\right]}\right], \end{aligned}$$

where $\Gamma[a, z] = \int_z^\infty x^{a-1}e^{-x}dx$ denotes the incomplete gamma integral.

The paper points out that the theory put forward here can be developed further to take into account multiple cash flows, families of interdependent assets, and filtrations of greater complexity. Each asset is defined by a series of one or more cash flows, each such cash flow being dependent on a set of one or more independent market factors (X-factors). Each X-factor has an associated information process, which may be of the Brownian bridge type or the gamma bridge type, or some generalization thereof; and the market filtration is taken to be generated collectively by this set of information processes.

5.9.6 Factor Model Framework

Another recent application of the gamma processes which is gaining popularity is in the modelling of an underlying asset value in a structural model. Here the value S_t at time t is assumed to follow

$$S_t = S_0 \exp\left(-\Gamma(t; m, \kappa) + \mu t\right),$$

where $\Gamma(t; m, \kappa)$ denotes the scaled gamma process and $\mu = m \ln\left(1 + \kappa^{-1}\right)$.

The correlation between different assets in the basket is then introduced through a factor model, where the gamma process for each asset is modelled as a combination of a common global gamma process and an independent idiosyncratic process. The total intensity rate m is divided into the global intensity αm and the idiosyncratic intensity $(1 - \alpha) m$, where α captures the dependence between securities. The gamma process of the ith entity is then represented by

$$\Gamma_i(t) = \Gamma_g(t; \alpha m, \kappa) + \Gamma_i(t; (1 - \alpha) m, \kappa,),$$

where $\Gamma_g(t; \alpha m, \kappa)$ is the common global gamma process and $\Gamma_i(t; (1 - \alpha) m, \kappa)$ is the idiosyncratic gamma process.

A common simplification is to use the standard gamma process, instead of the scaled one, in other words to set $\kappa = 1$.

Apart from desirable qualities, such as the Markovian property of the gamma process, the source and extent of correlation is more transparent than in many other models. We are also in the familiar, well-researched, framework of factor models.

5.10 Coupled Stochastic Differential Equations

We now revisit the multivariate distribution theory with another approach that allows for heterogeneous marginals, as presented in [30]. This approach is to consider multivariate distributions as arising naturally from coupled stochastic differential equations (SDEs), rather than being artificially imposed by a copula. The paper considers a first step of the approach based on the equilibrium situation. The marginals here are one of the classic Pearson family of distributions. Fokker-Planck equations are used to associate the distributions to the corresponding SDEs.

Fokker-Planck equation describes the time evolution of a probability density function and is also known as the Kolmogorov forward equation. In one spatial dimension x, the Fokker-Planck equation for a process with drift $\mu(x_t, t)$ and diffusion $\Sigma(x_t, t)$ is defined as

$$\frac{\partial}{\partial t} f(x,t) = -\frac{\partial}{\partial x}\left(\mu(x_t,t) f(x,t)\right) + \frac{\partial^2}{\partial x^2}\left(\Sigma(x_t,t) f(x,t)\right).$$

The Fokker-Planck equation can be used for computing the probability densities of stochastic differential equations. Consider the Itô stochastic differential equation

$$dX_t = \mu(X_t,t)\;dt + \sigma(X_t,t)\;dW_t,$$

where $X_t \in \mathbb{R}^n$ is the state and $W_t \in \mathbb{R}^n$ is a standard n-dimensional Wiener process. If the initial distribution is X_0 distributed $f(x,0)$, then the probability density $f(x,t)$ of the state X_t is given by the Fokker-Planck equation with the drift $\mu_i(x,t)$ and diffusion terms

$$\Sigma_{ij}(x,t) = \frac{1}{2}\sum_k \sigma_{ik}(x,t)\,\sigma^T_{kj}(x,t).$$

In general, it is recognised that considering an approach to multivariate distributions based on coupled SDEs is a complicated time-dependent problem.

5.10.1 The Quantilized Fokker-Planck Equation

The starting point of the approach is the univariate SDE

$$dx_t = \mu(x_t,t)\,dt + \Sigma(x_t,t)\,dW_t, \tag{5.13}$$

where W_t is a standard Brownian motion. Let $f(x,t)$ denote the associated time-dependent probability density function. The quantile function $Q(u,t)$ associated with this density function is defined by the condition

$$\int_{-\infty}^{Q(u,t)} f(x,t)\,dx = u$$

and satisfies various conditions associated with density functions and the Fokker-Planck equation. The required results are the non-linear ordinary differential equation (ODE)

$$\frac{\partial^2 Q(u,t)}{\partial u^2} = -\frac{\partial \log\left(f\left(Q(u,t),t\right)\right)}{\partial Q}\left(\frac{\partial Q(u,t)}{\partial u}\right)^2$$

and the quantilized Fokker-Planck equation (QFPE) in the form

$$\frac{\partial Q}{\partial t} = \mu(Q,t) - \frac{1}{2}\frac{\partial \Sigma^2}{\partial Q} + \frac{1}{2}\Sigma^2(Q,t)\left(\frac{\partial Q(u,t)}{\partial u}\right)^{-2}\frac{\partial^2 Q(u,t)}{\partial u^2}. \tag{5.14}$$

5.10.2 Stochastic Equilibrium and the Pearson Family

The system defined in equation 5.13 is said to be in stochastic equilibrium if the density is time-independent. Then the quantile function is also time-independent and satisfies the non-linear ODE arising from equation 5.14, with the coefficients of the SDE also assumed to be time-independent. With some computation we can obtain the following equation

$$\Sigma^{-2}(Q)\left(\frac{d\Sigma^2}{dQ} - 2\mu(Q)\right) = -\frac{\partial \log(f(Q))}{\partial Q}.$$

Now, Pearson's distributions are linked to choices of density function for which, for constants a, b, c, m we have

$$-\frac{\partial \log(f(Q))}{\partial Q} = \frac{Q-m}{a+bQ+cQ^2}.$$

Combining the above two conditions, we can see that we are interested in that collection of SDEs for which

$$\Sigma^{-2}(Q)\left(\frac{d\Sigma^2}{dQ} - 2\mu(Q)\right) = \frac{Q-m}{a+bQ+cQ^2}. \tag{5.15}$$

5.10.3 Reconstructing Pearson Distributions from SDE's

The above equation is a differential constraint on allowable drift and volatility functions. It is necessary to find solutions of the constraint that will generate solutions of the equilibrium equations that are stable and, for practical purposes, arise naturally from a variety of initial conditions.

A few examples of Pearson's types are given in [30], including the Gaussian and the Student t. We will use the one-sided exponential case as an example, as presented in the paper. Here $m = c = a = 0$, making the right-hand side of equation 5.15 a constant. One can come up with diverse plausible choices for the drift and volatility, though not all of it will generate a stable equilibrium. A natural choice is the SDE for the CIR interest rate model:

$$dx_t = \frac{1}{2}\Sigma_0^2(1-\lambda x_t)\,dt + \Sigma_0\sqrt{x_t}dW_t,$$

which has the remaining Pearson parameter $b = \lambda^{-1}$, which is also the mean of the equilibrium exponential distribution, with density

$$f(x) = \lambda e^{-\lambda x}.$$

5.10.4 The Bivariate Case

A natural means of construction of a bivariate or multivariate system with any of the Pearson marginals is to write down a coupled set of SDEs where the dependency arises from the correlation in the underlying Brownian motions. So in the bivariate case we should consider the pair of SDEs

$$dx_{1t} = \mu_1 (x_{1t}, t)\, dt + \Sigma_1 (x_{1t}, t)\, dW_t^1$$

$$dx_{2t} = \mu_2 (x_{2t}, t)\, dt + \Sigma_2 (x_{2t}, t)\, dW_t^2,$$

where

$$\mathbb{E}\left[dW_t^1 dW_t^2\right] = \hat{\rho} dt,$$

and $\hat{\rho}$ denotes the underlying correlation of the Brownian motions. The Fokker-Planck equation for the joint density function can be specified from the above equations. For the equilibrium case we have the following two-dimensional PDE:

$$\frac{\partial}{\partial x_1}\left[-\mu_1 f + \frac{1}{2}\frac{\partial}{\partial x_1}\left(\Sigma_1^2 f\right)\right] + \frac{\partial}{\partial x_2}\left[-\mu_2 f + \frac{1}{2}\frac{\partial}{\partial x_2}\left(\Sigma_2^2 f\right)\right] + \frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2}\left(\hat{\rho}\Sigma_1\Sigma_2 f\right) = 0.$$

5.10.5 Application with a Fat-Tailed Distribution

If one marginal is Gaussian and the other Student, taking the first variable to be Student with degrees of freedom v, the equilibrium PDE comes out to be

$$\frac{\partial}{\partial x_1}\left[\left(1 - \frac{1}{v}\right) x_1 f + \frac{\partial}{\partial x_1}\left(\left(1 + \frac{x_1^2}{v}\right) f\right)\right] + \frac{\partial}{\partial x_2}\left[x_2 f + \frac{\partial f}{\partial x_2}\right] + 2\hat{\rho}\frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2}\left(f\sqrt{1 + \frac{x_1^2}{v}}\right) = 0,$$

and the associated SDE for simulation of the correlated system is

$$dx_{1t} = -\frac{1}{2}\Sigma_0^2\left(1 - \frac{1}{v}\right) x_{1t} dt + \Sigma_0\sqrt{1 + \frac{x_{1t}^2}{v}} dW_t^1,$$

$$dx_{2t} = -\frac{1}{2}\Sigma_0^2 x_{2t} dt + \Sigma_0 dW_t^2.$$

We have already recognised the appropriateness of fat-tailed distributions in modelling default events.

Chapter 6

Conclusion

6.1 What Happened to the Copulas?

Having looked at the theory of copulas, their applications to credit derivatives, some drawbacks, as well as alternative methods of modelling credit derivatives of asset portfolios, one may be excused for still wanting to ask "what exactly happened to the copulas?".

The recent global financial crisis brought to light the lack in the understanding and the shortcomings of existing credit risk modelling, especially that of dependent defaults. The Gaussian copula was to blame for underestimating default probabilities through the lack of extreme event dependence. And the empirical evidence showed that this finger-pointing was justified.

However, as [25] points out, even such an apparently inadequate model as the Gaussian copula does assign some probability to the extreme events, whether it is of the order of 0.1% or 5%. For sound financial institutions that aim to maintain a AA or A rating these probabilities are large enough to take note of the risks. It it argued in the paper that it was a matter of management decision whether to treat those probabilies as large enough to warrant more stringent risk management and capital provision, or to willingly take the risks and ride the wave.

We have also seen in this dissertation the variety of copulas available for modelling credit derivatives. Their features are also equally varied, including the strength of dependence of defaults in extreme circumstances. So while the Gaussian copula may be written off as a feasible model for credit derivatives, there seems to be no reason to put for example the Clayton copula in the same corner.

A simple answer to the above question may therefore be that copulas simply lost their popularity and their trustworthiness, following an inappropriate choice for the credit derivatives.

6.2 What Conclusions Can We Draw?

We have identified the main drawbacks of copulas in this dissertation. However, it was also noted that other models come with their warnings. It is a natural feature of any model to have its simplifications and pitfalls, and is by no means an indication that the model is not appropriate.

Perhaps one of the most useful lessons from this dissertation is the identification of the pitfalls of each model. This creates an awareness of the possible limitations of the results obtained and the necessary caution with which these results should be interpreted.

Essentially we understand that we cannot have a perfect model for portfolio credit derivatives. But we need to be aware what are the key features that the model should have, and then what are the desirable features of lesser importance. Tail dependency, as we have learnt, is a key feature.

A difficult task is also created by the scarcity of data. Even with the recent credit crisis most of the data available is in respect of the normal market conditions. No models are immune to this problem. In addition, there is the question of how does one account for the fact that credit events have recently happened - does one adjust the default probabilities or is this a validation of the a priori probabilities?

It has become common knowledge that complex derivatives are more difficult to price, no matter what model is used. They are more difficult to understand and are often many times removed (e.g. CDO squared) from the underlying assets. In addition, the models for dealing with such derivatives are usually still calibrated using prices of the more common credit derivatives, such as single name CDSs.

Following the credit crunch one may be warranted in pondering whether credit derivatives have a future. Nevertheless, the need for transferring credit risk is still present. So credit derivatives are here to stay. The question is whether their use will still extend to speculative purposes. The purpose to which these instruments are put will drive the demand for the specific type of credit derivatives in the market, which will in turn dictate the importance of modelling dependence or some other characteristics.

In addition the need to understand the credit risk, with its dependencies, is greater than ever. This is the case regardless of whether there is a high demand for portfolio-based credit derivatives. The credit risk models of dependent defaults contribute to improved understanding of this area.

6.3 Further Research Directions

We have learnt some significant lessons on credit derivative modelling since the copula days, in particular the fact that the models such as the Gaussian copula tend to understate the probability of multiple defaults. We are well aware that a credit derivative model needs to have heavy tails. We can also probably agree that a desirable model would take into account some information process and model the evolution of the default probabilities. And, more importantly, the chosen model should support our understanding of the source of dependence, by allowing some modelling of this common source itself.

However we are still largely back at the drawing board. A new, robust model is urgently needed. It will be of great interest to hear the reasons for a particular new market model arising and how its pricing for more complex derivatives compares to the models used to date.

One of the elements that should also be tackled is the dependence between default rates and inflation or interest rates. All the models that we looked at so far have assumed independence between these variables, although it is intuitively obvious that the probabilities of institutional defaults are linked to these economic indicators.

There is also much to be seen with regards to the future use and nature of credit derivatives themselves and the instruments that will become commonly traded in the market.

Lastly, there is an increasing emphasis on risk management and suitable capital allocation. In the past this has been mainly driven by the regulatory authorities and the development of the Basel II accord, issued by the Basel Committee on Banking Supervision. However, since the financial crisis there is an increased buy-in from the financial institutions themselves. The methodology applied to modelling dependence between credit derivatives translates very well for applications to risk management and capital allocation.

Bibliography

[1] BRODY D.C., HUGHSTON L.P. *and* MACRINA A. (2008). Dam rain and cumulative gain. Proceedings of the Royal Society London A, Vol. 464, No. 2095, 1801-1822.

[2] BRODY D.C., HUGHSTON L.P. *and* MACRINA A. (2007). Beyond hazard rates: a new framework for credit-risk modelling.Advances in Mathematical Finance, Festschrift volume in honour of Dilip Madan, edited by R. Elliott, M. Fu, R. Jarrow & J. Y. Yen. Birkhuser and Springer.

[3] BURTSCHELL X., GREGORY J., *and* LAURENT J.P. (2008). A comparative analysis of CDO pricing models. Working paper. BNP-Paribas.

[4] CHERUBINI U., LUCIANO E. *and* VECCHIATO W. (2004). *Copula methods in finance*. Wiley Finance.

[5] CHOUNDHRY M. (2004). *Structured Credit Products: Credit Derviatives and Synthetic Securitisation*. Wiley Finance.

[6] COX J.C., INGERSOLL J.E. *and* ROSS S.A. (1985). A theory of term structure of interest rates. *Econometrica*, 53(2):385-407.

[7] DAUL S., DE GIORGI E., LINDSKOG F. *and* McNEIL A.J. (2003). Using the grouped t-copula. RISK, November 2003: 73-76.

[8] DAVIS M. *and* LO V. (2001). Infectious defaults. *Quantitative Finance* 1:382-387.

[9] DUFFIE D. *and* GÂRLEANU (2001). Risk and valuation of collateralized debt obligation. *Financial Analysts Journal* 57(1):41-59.

[10] DUFFIE D. (1999) Credit swap valuation. *Financial Analysts Journal* 55(1):73-87.

[11] EMBRECHTS P. (2009). Copulas: a personal view. *Journal of Risk and Insurance* 76: 639-650.

[12] EMBRECHTS P., FREY R. *and* McNEIL A.J. (2005). *Quantitative risk management*. Princeton University Press.

[13] EMBRECHTS P., LINDSKOG F. *and* McNEIL A.J. (2003). Modelling dependence with copulas and applications to risk management. *Handbook of heavy-tailed distributions in finance* edited by Rachev S.T. Elsevier/North-Holland.

[14] FERMANIAN J.D. *and* SCAILLET O. (2004). Some statistical pitfalls in copula modelling for financial applications. FAME Research paper No. 108.

[15] FREY R. *and* McNEIL A.J. (2003). Dependent defaults in models of portfolio credit risk. *Journal of Risk*, 6(1).

[16] FREY R., McNEIL A.J. *and* NYFELER M.A. (2001). Copulas and credit models. *RISK*, October 2001: pages 111-114.

[17] HULL J., PREDESCU M. *and* WHITE A. (2006). The valuation on correlation-dependent credit derivatives using a structural model. Working paper. University of Toronto.

[18] HULL J. *and* WHITE A. (2006). The perfect Copula. Working paper. University of Toronto.

[19] HULL J. *and* WHITE A. (2004). Valuation of a CDO and n^{th} to default CDS without Monte Carlo simulation. *Journal of Derivatives* 12(2).

[20] JARROW R.A. *and* YU F. (2001). Counterparty risk and pricing of defaultable securities. *Journal of Finance*. 50(1):53-85.

[21] JOUANIN J.F., RAPUCH G., RIBOULET G. *and* RONCALLI T. (2001). Modelling dependence for credit derivatives with copulas. Working paper. Groupe de Recherche Opérationnelle. Crédit Lyonnais.

[22] LANDO D. (2004). *Credit Risk Modeling: Theory and Applications*. Princeton University Press.

[23] LI D. (2000). On default correlation: A copula function approach. Working paper 99-07. Risk Metrics Group.

[24] McNEIL A.J. (2007). Sampling Nested Archimedean Copulas. Journal of Statistical Computation and Simulation, 78(6): 567-581.

[25] LÖFFLER G. (2009). The US Housing Crash: Model or Management Failure?. GARP Risk Professional. December 2009: 56-57.

[26] MIKOSCH T. (2006). Copulas: Tales and Facts. *Extremes*, 9: pages 3-20. Rejoinder: pages 55-62.

[27] NELSEN R.B. (2005). *An Introduction to Copulas*. Second Edition. Springer.

[28] PARBHOO P. (2006). *Copula Methods for Pricing Collateralised Debt Obligations*. School of Computational and Applied Mathematics. University of the Witwatersrand.

[29] SCHMIDT T. (2006). Coping with copulas. In Rank, J. (Ed.). Copulas: From theory to application in finance. Risk Books.

[30] SHAW W.T. *and* MUNIR A. (2009). Dependency without copulas or ellipticity. European Journal of Finance, doi: 10.1080/13518470802697402

[31] SCHÖNBUCHER P.J. (2003). *Credit Derivatives Pricing Models: Models, Pricing, Implementation*. Wiley Finance.

[32] SCHÖNBUCHER P.J. *and* SCHUBERT D. (2001). Copula-dependent default risk in intensity models. Working paper. Bonn University.

[33] TEIXEIRA P. (2007) *Hedging Credit Risk Using Equity Derivatives*. School of Computational and Applied Mathematics. University of the Witwatersrand.

[34] YU F. (2004). Correlated defaults and the valuation of defaultable securities. Working Paper. University of California, Irvine.