GAUSSIAN COPULA MODEL, CDOs AND THE CRISIS

Module 8 assignment

University of Oxford
Mathematical Institute

An assignment submitted in partial fulfillment of the MSc in Mathematical Finance

JUNE 5, 2016
1. Introduction

Since the late 1990s banks and other financial institutions were searching for new possibilities to increase returns while diversifying their own risks. The emerging new product class of (multi-name) credit derivatives seemed to be perfectly suitable to achieve this goal. Well-known and popular representatives of the multi-name credit derivatives were the so called **Collateralized Debt Obligations (CDOs)**, which are structured asset-backed securities. By CDOs, banks are allowed to form tradable securities (bonds) out of a pool of different types of mostly illiquid debt, e.g. loans or mortgages. The bank transfers the pool of debt assets to a special purpose vehicle (SPV), which issues new financial products by securitisation of the transferred assets. The new securities are divided in several risk categories, the so called **tranches**, which have different ratings and returns, depending on the risk they bear. The tranched securities pay interest to the investors out of the cashflows generated by the underlying debt assets.

However, there were already critical voices at the early 2000s by influential people of the financial sector, like Warren Buffet who termed credit derivatives as "financial weapons of mass destruction" in the year 2003. But the high profitability of these products for the issuers and their structure which also enabled the banks to outsource and reduce their own credit risk led to increased interest in those products. Though, one important question remained: Is there an easy and efficient way to price these credit derivatives? Central for pricing multi-name portfolio securities, like CDOs is the incorporation of the dependence between defaults of the assets in the underlying debt pool. From a mathematical perspective is the dependence structure between several random variables completely described by their joint distribution function. Therefore, the marginal default distributions of the single names in the portfolio and their default dependency have to be used for estimating the joint loss distribution of the underlying portfolio.

In the year 2000, David X. Li published a very tractable method (see Li [2000]), based on some relatively simple math, to tackle the problem of the dependence structure. He introduced a **copula function approach** to model default dependence by the usage of market prices of credit default swaps. These data can be used to estimate the correlation coefficients between the single titles which serve as input for his model. No historical default data was needed, which was a significant simplification for the market participants. The correlation coefficients and the marginal default distributions build the input for a copula function which delivers the multivariate distribution of all random variables as output. A significant weakness of the general model by Li is the assumption that the estimated correlation pairs remain constant over time and that all the input correlation information can easily be condensed into one final parameter which measures the overall default risk (cf. Brigo et al. [2010a] and Forslund and Johansson [2012]).

Due to the simple application in practice, Li’s model extended rapidly from bond investors and Wall Street banks to ratings agencies and regulators in subsequent years. This trend was compounded due to the economic situation in the United States. A longer phase of low interest rates ensured that a broad group of citizens in the USA, also low-income families, were offered mortgages by the banks (so called subprime-mortgages, if borrowers have a poor credit rating). The rather frequent lack of collateral for the mortgages was no big problem, because the house prices were rising and in case of a default of the borrower, the bank could sell the pledged house at a profit. Shortly a lot of banks possessed huge amounts of mortgages which have been securitised to highly profitable CDOs by SPVs and would then be transferred all over the world. Due to their good rating, especially of the CDO senior tranches, also conservative institutional investors had been very important customers. The development of the global issue volume of CDOs before and after the crisis is presented in Figure 1.

The problems began in the years 2005/2006, when the growth of the US economy weakened rapidly and the interest levels rose, so that many borrowers were no longer able to meet their interest and repayment obligations for their mortgages. They had to sold their houses. The result was an excess supply in the US housing market which leads to a collapse in prices. A lot of big banks had to write-off the remaining (nearly) unsecured mortgages as a loss. Thus the underlying of the CDOs became almost worthless overnight and therefore the CDOs, too.

At this stage it was too late to stop the financial crisis, which finally caused the downfall of some very renowned financial institutions.

The excessive use of securitization of high risk credits into CDOs was encouraged by the Gaussian copula
2. Collateralized Debt Obligations

A Collateralized Debt Obligation (CDO) is a fixed income security which belongs to the product class of the "Asset Backed Securities". This means, that the CDO is backed by a pool of reference assets, which are transformed into tradable securities by securitization. For the process of securitization the pool of reference assets is transferred to a special purpose vehicle (SPV), which is an off-balance sheet trust and bankruptcy-remote from the bank (the so called 'Originator') which originally owned the reference assets.

There is a variety of ways to classify the different kinds of CDOs. One possibility is to use the nature of the assets which build the underlying of the CDO. The following characteristics are popular:

1. Collateralized Loan Obligations (CLO): Loans are the building blocks for the CDO. The portfolio of loans and hence the credit risk is fully transmitted to the SPV.
2. Collateralized Bond Obligations (CBO): Bonds are the underlying assets. Equivalent as for CLOs.

3. Synthetic CDO: The credit risk is transferred via credit derivatives (e.g. credit default swaps, CDS) to the SPV. The underlying pool of assets remain physically at the originator.

This is just a small sample of different kinds of CDO types and possible underlyings (e.g. finer divisions as credit card receivables, auto loans, mortgages are used). In order to get a comprehensive overview of the different classifications, I refer to Bluhm and Overbeck [2006] and Engelmann [2010].

By securitization of the pool of underlying assets, their cashflows, as installment payments, are redirected to meet the coupon payments of the CDO investors. Normally a CDO is split in several risk categories, the **tranches**. A distinction is made in descending order between **Senior-, Mezzanine- and Equity-tranches**, which are divided commonly in several sub-tranches. Figure 2 shows a simple schematic sketch of the process.

![Figure 2: Securitization of an asset pool into tradeable securities: tranches of a CDO.](image)

The seniority of the tranches determines the order in which the periodic payment obligations (coupons) to the investors are done. The investors of the senior tranches receive their demands first, then the mezzanine investors and finally the equity investors. The level of returns of the different tranches declines with increasing seniority. Therefore, the senior tranches pay the lowest coupon and the equity tranche the highest. This is because of the increasing risk of the tranches with low seniority. Their payments are serviced not until all other higher tranche payment requirements are met, and additional they bear the losses triggered by credit events first. A credit event in the pool of reference assets leads to the default of the respective asset and hence to the loss of the underlying capital (up to a potential recovery value). The default of an asset reduces the value of the asset pool and therefore of the CDO which has the asset pool as underlying. Occurring losses are borne first by the investment products in the tranche with the lowest seniority, the equity tranche. The coupons which the equity investors receive are reduced to compensate the default. Each tranche has a **Attachment point** and a **Detachment point** which define the loss-making areas of the respective tranche. As stated before, the equity tranche bears the first losses (hence their attachment point = 0) until the portfolio losses exceed their detachment point, so that no notional remains in that tranche and no payment are made anymore. Further losses attack the notional of the next tranche, the mezzanine tranche, until their detachment point is reached and so on. This principle is illustrated in Figure 3.

The structure of a CDO ensures, that the senior tranche receive a better rating than the mezzanine and equity...
3. Introduction of the copula theory and the Gaussian copula model

A copula is a multivariate probability distribution function with uniform marginal distribution functions. Copulas allow to separate the problem of estimating a multidimensional distribution of several variables into the estimation of the individual marginal distributions and the joint dependency structure between these one-dimensional random variables. This is a major simplification for the pricing of multi-name credit derivatives, like CDOs, where one is mainly interested in the default dependence between different assets. The formal definition of a copula is the following one.

Definition 3.1. Let $U_1, \ldots, U_n$ uniformly distributed random variables in $[0, 1]$. A multivariate distribution function $C_\rho : [0, 1]^n \rightarrow [0, 1]$ with

$$C_\rho(u_1, u_2, \ldots, u_n) = \mathbb{P}[U_1 \leq u_1, U_2 \leq u_2, \ldots, U_n \leq u_n]$$

is called Copula-function, provided that the following conditions are met:
1. $C_\rho(u_1, u_2, \ldots, u_n) = 0$, if $u_i = 0$ for at least one $i \in \{1, 2, \ldots, n\}$.
2. $C_\rho(1, \ldots, 1, u_k, 1, \ldots, 1) = u_k$ for $k = 1, \ldots, n$.
3. $C_\rho$ is $n$-non-decreasing. This means, it is valid for each hyperrectangle $\prod_{i=1}^n [x_i, y_i] \subseteq [0, 1]^n$ that:

$$\sum_{(z_1, \ldots, z_n) \in \prod_{i=1}^n [x_i, y_i]} (-1)^{N(z_1, \ldots, z_n)} C_\rho(z_1, \ldots, z_n) \geq 0$$

with $N(z_1, \ldots, z_n) := |\{k \mid z_k = x_k\}|$.

Here $\rho$ is a general dependence parameter, which characterizes the correlation structure of the copula.

For a copula $C_\rho : [0, 1]^n \rightarrow [0, 1]$ and $n$ marginal distribution functions $F_1(x_1), \ldots, F_n(x_n)$ of random variables $X_1, \ldots, X_n$ it can be concluded that

$$C_\rho(F_1(x_1), \ldots, F_n(x_n)) = \mathbb{P}[U_1 \leq F_1(x_1), \ldots, U_n \leq F_n(x_n)]$$

$$= \mathbb{P}[F_1^{-1}(U_1) \leq x_1, \ldots, F_n^{-1}(U_n) \leq x_n]$$

$$= \mathbb{P}[X_1 \leq x_1, \ldots, X_n \leq x_n]$$

$$= F(x_1, \ldots, x_n)$$

Then $F(x_1, \ldots, x_n)$ is a copula-dependent multivariate distribution function with associated marginal distribution functions $F_1(x_1), \ldots, F_n(x_n)$.$^1$ For the marginal distribution functions of the $X_i$ you get:

$$C_\rho(F_1(\infty), \ldots, F_{i-1}(\infty), F_i(x_i), F_{i+1}(\infty), \ldots, F_n(\infty))$$

$$= \mathbb{P}[X_1 \leq \infty, \ldots, X_{i-1} \leq \infty, X_i \leq x_i, X_{i+1} \leq \infty, \ldots, X_n \leq \infty]$$

$$= \mathbb{P}[X_i \leq x_i]$$

$$= F_i(x_i)$$

The following theorem by Sklar is a very important result of the copula theory. It is the reversal of the just stated conclusion and enables the splitting of a multivariate distribution into their marginal distributions and an associated copula function.

**Theorem 3.1** (Sklar). Let $F$ be a multivariate $n$-dimensional distribution function and $F_1, \ldots, F_n$ the related marginal distribution functions. Then there exists a $n$-dimensional copula $C_\rho$, so that the following equation holds for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$:

$$F(x_1, \ldots, x_n) = C_\rho(F_1(x_1), \ldots, F_n(x_n)).$$

If all marginal distribution functions $F_1, \ldots, F_n$ are continuous, then the copula function $C_\rho$ is unique, as stated e.g. in O’Kane [2008] or Bluhm and Overbeck [2006]. For a proof of the theorem I refer to Nelsen [2007].

It should be emphasized, that Sklar’s theorem "only" ensures the existence of a copula, but the appropriate choice of a copula function which matches the specific model requirements is fraught with risk. There are a variety of copula functions used in practice which can model different dependence structures. At times of the financial crisis the **Gaussian copula function** was widely used in pricing multi-name credit derivatives due to Li’s model approach, which became market standard quickly in those days.

**Definition 3.2.** Let $\Phi$ be the distribution function of the one-dimensional standard normal distribution and $\Phi_\Sigma$ the distribution function of the $n$-dimensional standard normal distribution with positive definite correlation matrix $\Sigma$. Then the $n$-dimensional **Gaussian copula** $C_\Sigma^\Phi$ is defined as follows:

$$C_\Sigma^\Phi(u_1, \ldots, u_n) = \Phi_\Sigma(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n))$$

$^1$Cf. Li [2000]
for all \((u_1, \ldots, u_n) \in [0,1]^n\). For \(n = 2\) you get:

\[
C_{\rho_{12}}^\Phi(u_1, u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi(1 - \rho_{12}^2)^{1/2}} \exp \left(-\frac{s^2 - 2\rho_{12} \cdot s \cdot t + t^2}{2(1 - \rho_{12}^2)}\right) ds dt,
\]

with \((u_1, u_2) \in [0,1]^2\) and \(\rho_{12}\) states for the correlation coefficient of the bivariate standard normal distribution.

The Gaussian copula approach for CDO tranche modeling uses the copula function, introduced in Definition 3.2, to incorporate default correlation in the underlying asset pool.

Let \(\tau_i\) be a random variable, describing the default time of the \(i\)-th asset in the underlying pool and \(F(t_i) = \mathbb{P}[\tau_i \leq t_i]\) the marginal default time distribution function for \(i = 1, \ldots, M\) and \(M\) is the number of assets in the pool. To model the joint default times of all assets in the underlying pool \(F(t_1, \ldots, t_M) = \mathbb{P}[\tau_1 \leq t_1, \ldots, \tau_M \leq t_M]\) for all \((t_1, \ldots, t_M) \in \mathbb{R}_+^M\) you can make use of Sklar’s theorem, which ensures the existence of a copula \(C : [0,1]^M \rightarrow [0,1]\), such that

\[
F(t_1, \ldots, t_M) = C(F_1(t_1), \ldots, F_M(t_M)).
\]

Li used in his publication a Gaussian copula \(C =: C_{\Sigma}^\Phi\). The most common used version of Li’s model is the One-factor Gaussian copula model, which offers a high analytical tractability by assuming that the portfolio of underlying assets is homogeneous (therefore the model is named “Homogeneous Large Pool Gaussian Copula model”). Each asset in the underlying portfolio belongs to a company \(i\) with asset value \(X_i\) for \(i = 1, \ldots, M\). In the one-factor framework, the value of company \(i\) is modeled as

\[
X_i = \sqrt{\rho}Y + \sqrt{1-\rho}Z_i, \quad \text{for } i = 1, \ldots, M,
\]

where \(Y, Z_1, \ldots, Z_M\) are i.i.d. \(N(0,1)\) and \(\rho \in (0,1)\). \(Y\) is a systematic risk factor, which describes a kind of market risk, common to all companies and the \(\{Z_i\}_{i=1}^M\) present idiosyncratic risk factors, which are specific for each company \(i\). One of the core assumptions of this model is the flat correlation between each pair of companies due to the homogeneity of the asset pool. This leads to one value \(\rho\) for the correlation of every pair of assets.

Put it all together, the transpose of the value vector \((X_1, \ldots, X_M)\) has a multivariate normal distribution:

\[
\begin{pmatrix}
X_1 \\
\vdots \\
X_M
\end{pmatrix} \sim N^M \left[
\begin{pmatrix}
0 \\
\vdots \\
0
\end{pmatrix},
\begin{pmatrix}
1 & \rho & \rho & \ldots & \rho \\
\rho & 1 & \rho & \ldots & \rho \\
\rho & \rho & 1 & \ldots & \rho \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \rho & \rho & \ldots & 1
\end{pmatrix}
\right].
\]

A company \(i\) defaults if their asset value \(X_i\) falls below a threshold. For linking the single default time \(\tau_i\) to the one-factor model, the following relation is used

\[
X_i = \Phi^{-1}(F_i(\tau_i)) \quad (\Leftrightarrow \tau_i = F_i^{-1}(\Phi(X_i))).
\]

After the marginal distribution functions \(\{F_i\}_{i=1}^M\) for the \(\{\tau_i\}_{i=1}^M\) are determined, you can use equation (1) with a Gaussian copula \(C_{\Sigma}^\Phi\) to estimate the joint default distribution of the asset pool and after that, Li’s one-factor model is fully specified.

In practice the marginal defaults are often modeled as the first jump time of a Poisson process with an intensity rate \(\lambda_i(t)\), which describes the instantaneous default probability of the company \(i\) at time \(t\), and with exponentially distributed jumps

\[
F_i(t) = \mathbb{P}[\tau_i \leq t] = 1 - \exp \left(-\int_{t_0}^t \lambda_i(u)du\right), \quad t \geq t_0.
\]

Often it is assumed that every individual company has a time constant intensity function \(\lambda_i(t) = \lambda_i\), which leads to the default distribution \(F_i(t) = 1 - \exp \left(-\lambda_i \cdot (t - t_0)\right)\). These exponential marginals distribution functions, together with the market observed prices for the CDO tranches (≡ market spreads) can be used to
4. Limitations and drawbacks of the Gaussian copula in the context of the Financial Crisis

As already indicated previously, the Gaussian copula model suffers from two main problems, firstly the inconsistencies in implied CDO tranche correlation estimation and secondly the failure in modeling extremal events due to the missing tail dependence. Obviously, there are still other shortcomings (e.g. see Donnelly and Embrechts [2010] or Brigo et al. [2010a]), but the two mentioned problems are the most striking ones.

Inconsistencies in implied correlations estimation:
The implied correlations for each CDO tranche are calculated in such a way that their observed market spread agree with the used pricing model, in this case with one-factor Gaussian copula model. There are two approaches which are widely used in the practice, the compound correlation and the base correlation method. The compound correlation approach is similar to the concept of implied volatility. Each tranche is considered separately and is priced using a single correlation parameter as input. The parameter is calculated out of a market spread by inverting the pricing model. By an iteration process you receive the compound correlations of all tranches, which produces a spread that matches the market quote. The presented one-factor Gaussian model uses only one correlation $\rho$ for all tranches to specify the loss distribution and the price, so you should get equal correlation values for each tranche by the compound correlation approach. That is, however, precisely the problem, because this method does not produce a unique solution for all tranches, since the compound correlation is a function of both, the tranche attachment and detachment point (cf. Lehnert et al. [2005] and Torresetti et al. [2006]). An example of the typical resulting correlation "smile" is illustrated in Figure 4 on the left hand side.

![Figure 4](image.png)

**Figure 4**: Left graphic: Compound correlations of DJ iTraxx Europe 5-year tranches as of 2004/02/19. Right graphic: DJ iTraxx Europe 5-year 3-6% Tranche spread in bp (y axis), Input Correlation (x axis), as of 2004/02/16. It can be seen that for the traded spread of 227bp, there are two possible correlations. Source for both: JPMorgan.

Additional, many empirical evidence shows that the mezzanine tranches are not monotonic in correlation. In consequence, the estimation of the correlation parameter can provide two different results that lead to the same market spread. Moreover, sometimes the compound correlation parameter can’t be calibrated to the market data, in these cases it does simply not exist as presented in Figure 4 on the right hand side.

The main idea behind the concept of the base correlation is that every tranche can be decomposed into two first loss tranches, which means that you only have to value tranches with lower attachment point zero. For
example, to replicate a long position in a tranche with the attachment points \( a_i \) and detachment point \( d_i \) you can take simultaneously a long position in a first-loss-tranche with detachment point \( d_i \) and a short position in a first-loss-tranche with attachment point \( a_i \). By using a bootstrapping technique, you can estimate all base correlations successively. The concept of base correlations overcomes some limitations of the compound correlations, e.g. it provides a unique implied correlation for fixed attachment points and a wider range of tranche spreads can be inverted into a base correlation value. Additional the approach makes it possible to price tranches with arbitrary attachment and detachment points which are not traded actively by linearly interpolation of the traded tranches.

But base correlations also have some drawbacks. The problem of different correlation parameters for each tranche remains as well as there could be still situations where you can’t calibrate the correlation to the market spread. Furthermore an new problem arises, the expected tranche loss might become negative under some market conditions, which violates the basic no-arbitrage constraints. Figure 5 illustrates the appealed drawbacks.

![Figure 5: Left graphic: Base correlations of DJ iTraxx Europe 5-year tranches as of 2004/02/19. Source: JPMorgan. Right graphic: Expected tranche loss as a function of time derived from the base correlations calibrated to the DJ iTraxx Series 5 10-year tranche spreads as of 2005/08/03. Source Brigo et al. [2010b].](image_url)

The described inconsistencies of both correlation approaches lead to big problems in the modeling of fair tranche prices and can cause high losses if you apply delta-hedging strategies based on the tranche valuations. Examples on that and further surveys can be found in Torresetti et al. [2006], Brigo et al. [2010a] and Lehnert et al. [2005].

**The lack of tail dependence:**

The Gaussian copula approach can’t model tail dependence. Therefore, the occurrence of extreme events in the upper and lower tails of the joint distribution of several random variables is dramatically underestimated. This has a big impact on the valuation of CDOs, because the simultaneous default of a large number of assets in the underlying portfolio can’t be considered in the model. Especially in times of crisis, this might cause dangerous inaccurate appraisals of structured products, like CDO tranches.

In the following I focus for illustration purposes on the bivariate case and start with a definition.

**Definition 4.1.** Let \((X,Y)\) be a bivariate random vector and \(X\) and \(Y\) two continuous random variables with associated distribution functions \(F_X\) and \(F_Y\).

The random variables \(X\) and \(Y\) are **upper tail dependent**, if

\[
\lambda_U := \lim_{u \to 1} \mathbb{P}[Y > F_Y^{-1}(u) \mid X > F_X^{-1}(u)]
\]

exists and \(\lambda_U \in (0,1]\). \(X\) and \(Y\) are independent in the upper tail if \(\lambda_U = 0\).

The random variables \(X\) and \(Y\) are **lower tail dependent**, if

\[
\lambda_L := \lim_{u \to 0} \mathbb{P}[Y \leq F_Y^{-1}(u) \mid X \leq F_X^{-1}(u)]
\]
The theorem can be applied: The family of Gaussian copulas is symmetrical in its arguments. Therefore, the second statement from

\[ \rho \]

with a Gaussian copula

\[ C \]

let

\[ \text{Theorem 4.1.} \]

The next theorem is stated in Embrechts et al. [2003] and connects tail dependence with the copula approach.

\[ \lambda_U = \lim_{u \searrow 1} \frac{1 - 2u + C(u,u)}{1 - u} \]

and

\[ \lambda_L = \lim_{u \searrow 0} \frac{C(u,u)}{u} \]

The copula \( C \) models upper resp. lower tail dependence, if \( \lambda_U \in (0,1] \) resp. \( \lambda_L \in (0,1] \). \( C \) has no upper resp. lower tail dependence, if \( \lambda_U = 0 \) resp. \( \lambda_L = 0 \).

2. In addition, if we assume that \( C \) has symmetrical arguments, this means \( C(u,v) = C(v,u) \forall u,v \in [0,1] \), then we get:

\[ \lambda_U = \lim_{u \searrow 1} 2 \cdot \mathbb{P}[V > u \mid U = u] \]

and

\[ \lambda_L = \lim_{u \searrow 0} 2 \cdot \mathbb{P}[V < u \mid U = u] \]

A proof of Theorem 4.1 can be found in the appendix A.1.

Theorem 4.1 can be used to show that the Gaussian copula approach fails to model tail dependence. Again I focus on the bivariate case. Let \( \tau_1 \) and \( \tau_2 \) be the default time variables of two companies with associated one-dimensional distribution functions \( F_1 \) and \( F_2 \). The multivariate distribution function of \( (\tau_1, \tau_2) \) is modeled with a Gaussian copula \( C_{\rho_{12}} \), with estimated correlation coefficient \( \rho_{12} \). Hence:

\[ \rho = \rho_{12} \]

The family of Gaussian copulas is symmetrical in its arguments. Therefore, the second statement from Theorem 4.1 can be applied:

\[ \lambda_U = \lim_{u \searrow 1} \mathbb{P}[\tau_2 > F_2^{-1}(u) \mid \tau_1 > F_1^{-1}(u)] \]

\[ = 2 \cdot \lim_{u \searrow 1} \mathbb{P}[F_2(\tau_2) > u \mid F_1(\tau_1) = u] \]

The distribution function \( \Phi \) of the standard-normal distribution does not have a boundary point in the positive range \( (\mathbb{R}^+) \), so you can express the limit value as follows:

\[ 2 \cdot \lim_{u \searrow 1} \mathbb{P}[F_2(\tau_2) > u \mid F_1(\tau_1) = u] = 2 \cdot \lim_{x \to \infty} \mathbb{P}[-\frac{1}{2}(F_2^{-1}(\tau_2)) > x \mid \Phi^{-1}(F_1(\tau_1)) = x] \]

\[ = 2 \cdot \lim_{x \to \infty} \mathbb{P}[X > x \mid Y = x] \]

(2)

with \( X = \Phi^{-1}[F_2(\tau_2)], Y = \Phi^{-1}[F_1(\tau_1)] \sim \mathcal{N}(0,1) \) and correlation parameter \( \rho \) of \( (X,Y) \).

In the bivariate standard-normal case, you get for the conditional distribution:

\[ X|Y = x \sim \mathcal{N}(\rho x, 1 - \rho^2) \]

\[ \text{Cf. Nelsen [2007]} \]
This result and the transformation of a normally distributed random variables to a standard-normal random variable\(^3\) applied to equation (2) leads to:

\[
2 \cdot \lim_{x \to \infty} \Pr[X > x \mid Y = x] = 2 \cdot \lim_{x \to \infty} \left[ 1 - \Phi\left( \frac{x - \rho x}{\sqrt{1 - \rho^2}} \right) \right] = 2 \cdot \lim_{x \to \infty} \left[ 1 - \Phi\left( x \cdot \sqrt{\frac{(1 - \rho)^2}{(1 - \rho)(1 + \rho)}} \right) \right] = 2 \cdot \lim_{x \to \infty} \left[ 1 - \Phi\left( x \cdot \sqrt{\frac{1 - \rho}{1 + \rho}} \right) \right] = 0 = \lambda_U \quad \text{for } \rho \in (-1, 1)
\]

Equivalent calculation steps result in \(\lambda_L = 0\) for \(\rho \in (-1, 1)\). Alternative, you can use the fact, that the Gaussian copula is radially symmetrical.\(^4\) From this we have \(\lambda_U = \lambda_L\).

Thus, the Gaussian copula is asymptotically independent in both tails for \(\rho \in (-1, 1)\).

\section*{5. Alternative to the Gaussian copula - mixture of two copulas}

As presented in the previous section 4, the Gaussian copula can’t be used to model tail dependence in joint distribution functions, which is a major weakness (amongst other flaws) of this approach as it turned out during the Financial Crisis. In order to remedy this problem, at least in part, the choice of another copula to model the joint defaults of assets in an underlying pool for a CDO, may be more suitable. In financial markets it could be expected that large joint losses of several companies, resp. assets occur more often than large joint gains, at least if you rely on more conservative valuation approaches. Especially in times of crisis, this is no bad assumption. Therefore a way to model asymmetric tail dependence is required, which leads to the class of Archimedean copulas. Two representatives of this class are the Gumbel and the Clayton copula, which are both asymmetric copulas, that are able to model upper and lower tail dependence, but each only in one direction. However, it is possible to create a mixture of both, which is still a copula function and has the property of asymmetric tail dependence. Before this point is considered in more detail I refer to the book of Nelsen [2007] to get an introduction of the theory of Archimedean copulas.

\textbf{Theorem 5.1.} Let \(C_1^{\rho_1}\) and \(C_2^{\rho_2}\) two copula functions and \(\theta \in [0, 1]\). A linear convex combination \(C_\rho\) of \(C_1^{\rho_1}\) and \(C_2^{\rho_2}\) is still a copula function:

\[
C_{\rho_1,\rho_2}^{12}(u_1, u_2, \ldots, u_n; \theta) = \theta C_1^{\rho_1}(u_1, u_2, \ldots, u_n) + (1 - \theta)C_2^{\rho_2}(u_1, u_2, \ldots, u_n).
\]

A proof of this can be found in Nelsen [2007]. For a mixture of a Clayton \(C_{\rho_C}\) and a Gumbel \(C_{\rho_G}\) copula in the bivariate case you get

\[
C_{\rho_C,\rho_G}^{CG}(u_1, u_2; \theta) = \theta C_{\rho_C}^C(u_1, u_2) + (1 - \theta)C_{\rho_G}^G(u_1, u_2)
\]

with

\[
C_{\rho_C}^C(u_1, u_2) = (u_1^{-\rho_C} + u_2^{-\rho_C} - 1)^{-\frac{1}{\rho_C}}, \quad \forall \rho_C > 0
\]

and

\[
C_{\rho_G}^G(u_1, u_2) = \exp\left( - \left[ (\log u_1)^{\rho_G} + (\log u_2)^{\rho_G} \right] \frac{1}{\rho_G} \right), \quad \forall \rho_G \in [-1, \infty).
\]

If viewed separately, the following values for the tail dependence parameters can be estimated for the Clayton and Gumbel copula, as stated in the appendix A.2:

\begin{enumerate}
  \item \textbf{Clayton copula}: \(\lambda_U = 0\) and \(\lambda_L = 2 - \frac{1}{\rho_C}\)
\end{enumerate}

\(^3\)It holds that: \(W \sim N(\mu, \sigma^2) \implies \frac{W - \mu}{\sigma} \sim N(0, 1)\)

\(^4\)Cf. McNeil et al. [2005], ‘radial symmetric’ means: Let \((U_1, \ldots, U_n) \sim C\) with appropriate copula \(C\). \(C\) is called radial symmetric if: \((U_1 - 0.5, \ldots, U_n - 0.5) \overset{d}{=} (0.5 - U_1, \ldots, 0.5 - U_n) \iff U \overset{d}{=} 1 - U\) with \(1, U \in \mathbb{R}^{1 \times n}\)
2. **Gumbel copula**: \( \lambda_U = 2 - 2^{\frac{1}{\rho_G}} \) and \( \lambda_L = 0 \).

This means that a Clayton copula can be used to model lower tail dependence and a Gumbel copula is able to model upper tail dependence, by appropriate selection of \( \rho_C \) and \( \rho_G \). Therefore, a mixture of both can be used to model asymmetric tail dependence for the upper and lower tail. The parameters \( \lambda_{12}^U \) and \( \lambda_{12}^L \) for a combination of two copulas can easily be estimated, if you know the tail dependence parameters of every single copula \( \lambda_1^U \), \( \lambda_2^U \) and \( \lambda_1^L \), \( \lambda_2^L \) and if you have defined the weight coefficient \( \theta \in [0, 1] \):

\[
\begin{align*}
\lambda_{12}^U &= \theta \lambda_1^U + (1 - \theta) \lambda_2^U \\
\lambda_{12}^L &= \theta \lambda_1^L + (1 - \theta) \lambda_2^L.
\end{align*}
\]

The parameter \( \theta \) should be estimated by market calibration under consideration of extreme events, such as the simultaneous default of many companies as observed during the Financial Crisis. By the mixture of two copulas you gain much flexibility in modeling dependence structures compared to the Gaussian copula approach and this might be one possibility to overcome one of the most serious flaws of this approach.

### 6. Conclusion

The exploding popularity, beginning in the early 2000’s, of the securitization of loan receivables by financial institutions and especially the pricing of the resulting multi-name credit derivatives by **copula approaches** can be illustrated with a simple example: While a Google search for the word "copula" yields about 10,000 hits in the year 2003 as stated in Mikosch [2006], the same search results in approximately 1,3 million hits in the year 2007 as written by the Financial Times.

The Gaussian copula approach presented in Li [2000], provides a fast and very practicable opportunity to price the highly complex multi-name credit derivatives, mostly CDOs, using only a few market observable parameters. In addition to the simplicity of the valuation approach, the financial institutions obtained very lucrative profits by issuing CDOs and as a side effect, they got rid of huge parts of their credit risks. Consequently it was not surprising that the market for these products grew ever more rapidly and that early warning signals were ignored.

The limitations of the Gaussian copula approach, discussed in Section 4 and further drawbacks were already well known and published before the Financial Crisis. However, there were a large number of more or less populist articles, like "The formula that felled Wall St" (Financial Times), "Wall Street’s Math Wizards forgot a few variables" (New York Times) or "Recipe for disaster: the Formula that killed Wall Street" (Wired magazine) which blamed the academics and mathematicians to be responsible for the Financial Crisis. In their readable work Brigo et al. [2010a] rejected this critique and clarified that in particular academics had warned before the crisis to rely blindly on the Gaussian copula approach.

In order to make it perfectly clear, a mathematical model alone, cannot be responsible for the Financial Crisis and its consequences. All involved parties, namely risk managers and financial institutions must be aware of the flaws and risks of the models they apply. They are responsible to choose appropriate models for the applications under consideration and they need to know what are the structures and properties of these models. The question remains if there is a "right" copula, which should have been used to prevent the crisis? The mixture of two copulas, e.g. a Clayton and Gumbel copula, as introduced in Section 5 can overcome some of the drawbacks of the Gaussian copula approach, but others still remain and maybe new problems arise. Hence, as stated before, it should be borne in mind that you are always dealing 'only' with a mathematical model and a model is not able to fully capture all factors having an effect in reality - therefore do not blindly trust to it. Only if this is clear to the people and they generally act accordingly, the next Financial Crisis can be prevented.

### References

A. Mathematical appendix

The proof is based on the sketch presented in Embrechts et al. [2003].

**Proof.** The proof is only executed for $\lambda_U$. However, the assertions for $\lambda_L$ follow equivalently.

1. The first assertion: For $\lambda_U$ and a bivariate copula $C$ with $(X,Y) \sim C(F_X(X),F_Y(Y))$ the following is valid:

$$
\lambda_U = \lim_{u \to 1} P[Y > F_Y^{-1}(u) \mid X > F_X^{-1}(u)] = \lim_{u \to 1} P[F_Y(Y) > u \mid F_X(X) > u]
$$
A. Mathematical appendix

= \lim_{u \to 1} \frac{\mathbb{P}[F_Y(Y) > u, F_X(X) > u]}{\mathbb{P}[F_X(X) > u]}
= \lim_{u \to 1} \frac{1 - (\mathbb{P}[F_X(X) \leq u] + \mathbb{P}[F_Y(Y) \leq u] - \mathbb{P}[F_X(X) \leq u, F_Y(Y) \leq u])}{1 - \mathbb{P}[F_X(X) \leq u]}
= \lim_{u \to 1} \frac{1 - 2u + C(u, u)}{1 - u}
\text{ , because of } F_X(X), F_Y(Y) \sim U[0, 1]

(3)

2. The second assertion: The numerator and denominator of equation (3) converge towards 0 for \( u \to 1 \), so you can use L'Hôpital’s rule:

\[ \lim_{u \to 1} 1 - 2u + C(u, u) = 1 - 2 + 1 = 0 \quad \text{and} \quad \lim_{u \to 1} 1 - u = 1 - 1 = 0 \]

It follows

\[ \lim_{u \to 1} \frac{1 - 2u + C(u, u)}{1 - u} = \lim_{u \to 1} \left[ -2 + \frac{\partial C(u, u)}{\partial u} \right] \]
\[ = - \lim_{u \to 1} \left[ -2 + \frac{\partial C(s, t)}{\partial s} \Bigg|_{s=t=u} + \frac{\partial C(s, t)}{\partial t} \Bigg|_{s=t=u} \right] \]

(4)

In the next step I introduce an interim result which is needed in the following:

\[ \mathbb{P}[V \leq v \mid U = u] = \lim_{\Delta u \to 0} \mathbb{P}[V \leq v \mid u \leq U \leq u + \Delta u] \]
\[ = \lim_{\Delta u \to 0} \frac{\mathbb{P}[V \leq v, u \leq U \leq u + \Delta u]}{\mathbb{P}[u \leq U \leq u + \Delta u]} \]
\[ = \lim_{\Delta u \to 0} \frac{\mathbb{P}[V \leq v, U \leq u + \Delta u] - \mathbb{P}[V \leq v, U \leq u]}{\mathbb{P}[U \leq u + \Delta u] - \mathbb{P}[U \leq u]} \]
\[ = \lim_{\Delta u \to 0} \frac{C(u + \Delta u, v) - C(u, v)}{\Delta u} \]
\[ = \frac{\partial}{\partial u} C(u, v) \]

Hence, you get

\[ \mathbb{P}[V > v \mid U = u] = 1 - \frac{\partial}{\partial u} C(u, v) \]

and

\[ \mathbb{P}[U > u \mid V = v] = 1 - \frac{\partial}{\partial v} C(u, v) \]

From this result and equation (4), you can deduce the second assertion:

\[ \lambda_U = - \lim_{u \to 1} \left[ -2 + \frac{\partial}{\partial s} C(s, t) \Bigg|_{s=t=u} + \frac{\partial}{\partial t} C(s, t) \Bigg|_{s=t=u} \right] \]
\[ = \lim_{u \to 1} \left[ 1 - \frac{\partial}{\partial s} C(s, t) \Bigg|_{s=t=u} + 1 - \frac{\partial}{\partial t} C(s, t) \Bigg|_{s=t=u} \right] \]
\[ = \lim_{u \to 1} \mathbb{P}[V > u \mid U = u] + \mathbb{P}[U > u \mid V = u] \]
\[ = 2 \cdot \lim_{u \to 1} \mathbb{P}[V > u \mid U = u] \]

(5)
A. Mathematical appendix

A.2. Tail dependence of bivariate Gumbel and Clayton copula’s

For a bivariate Gumbel copula it holds:

\[ C^G_{\rho_G}(u, u) = \exp \left( - \left[ (-\log u)^{\rho_G} + (-\log u)^{\rho_G} \right] \frac{1}{\rho_G} \right) \]
\[ = \exp \left( - \left[ -2(\log u)^{\rho_G} \right] \frac{1}{\rho_G} \right) \]
\[ = \exp \left( 2 \frac{1}{\rho_G} \log u \right) \]
\[ = u^{2 \frac{1}{\rho_G}}. \]

The use of Theorem 4.1 finally leads to the parameter for the upper and lower tail dependence coefficient.

**Upper tail dependence coefficient for a bivariate Gumbel copula:**

\[ \lambda_U = \lim_{u \to 1} \frac{1 - 2u + C^G_{\rho_G}(u, u)}{1 - u} \]
\[ = \lim_{u \to 1} \frac{1 - 2u + u^{2 \frac{1}{\rho_G}}}{1 - u} \]
\[ = 2 - 2 \frac{1}{\rho_G}. \]

The numerator and denominator of equation (6) converge towards 0 for \( u \to 1 \), so you can use L’Hôpital’s rule:

\[ = \lim_{u \to 1} \frac{-2 + \frac{1}{\rho_G} u^{2 \frac{1}{\rho_G}} - 1}{-1} \]
\[ = 2 - 2 \frac{1}{\rho_G}. \]

**Lower tail dependence coefficient for a bivariate Gumbel copula:**

\[ \lambda_L = \lim_{u \to 0} \frac{C^G_{\rho_G}(u, u)}{u} \]
\[ = \lim_{u \to 0} \frac{u^{2 \frac{1}{\rho_G}}}{u} \]

Again with L’Hôpital’s rule it follows:

\[ = \lim_{u \to 0} \frac{2 \frac{1}{\rho_G} u^{2 \frac{1}{\rho_G}} - 1}{1} \]
\[ = 0 \]

Equivalently, you get for a bivariate Clayton copula:

\[ C^C_{\rho_C}(u, u) = \left( u^{-\rho_C} + u^{-\rho_C} - 1 \right)^{-\rho_C^{-1}} \]
\[ = (2u^{-\rho_C} - 1)^{-\rho_C^{-1}}. \]

Consequently you can deduce the upper and lower tail dependence coefficient by usage of Theorem 4.1 and L’Hôpital’s rule.
### A. Mathematical appendix

**Upper tail dependence coefficient for a bivariate Clayton copula:**

\[
\lambda_U = \lim_{u \uparrow 1} \frac{1 - 2u + C_{\rho_C}^C(u, u)}{1 - u} = \lim_{u \uparrow 1} \frac{1 - 2u + (2u^{-\rho_C} - 1)^{-\rho_C^{-1}}}{1 - u} = \lim_{u \uparrow 1} \frac{-2 + 2u^{-\rho_C} - (2u^{-\rho_C} - 1)^{-\rho_C^{-1}}}{-1} = 0
\]

**Lower tail dependence coefficient for a bivariate Clayton copula:**

\[
\lambda_L = \lim_{u \downarrow 0} \frac{C_{\rho_C}^C(u, u)}{u} = \lim_{u \downarrow 0} \frac{(2u^{-\rho_C} - 1)^{-\rho_C^{-1}}}{u} = \lim_{u \downarrow 0} \frac{(2u^{-\rho_C} - 1)^{-\rho_C^{-1}}}{(u^{-\rho_C})^{-\rho_C}} = \lim_{u \downarrow 0} (2 - u^{-\rho_C})^{-\rho_C^{-1}} = 2^{-\rho_C}
\]

### A.3. Tail dependence and mixtures of two copula’s

For the upper tail dependence of a mixture copula \(C^{12}\), consisting of a convex combination of two copula’s \(C^1\) and \(C^2\), it can be deduced

\[
\lambda_{U}^{12} = \lim_{u \uparrow 1} \frac{1 - 2u + C^{12}(u, u)}{1 - u} = \lim_{u \uparrow 1} \frac{1 - 2u + \theta C^1(u, u) + (1 - \theta)C^2(u, u)}{1 - u} = \theta \lim_{u \uparrow 1} \frac{1 - 2u + C^1(u, u)}{1 - u} + (1 - \theta) \lim_{u \uparrow 1} \frac{1 - 2u + C^2(u, u)}{1 - u} = \theta \lambda_U^1 + (1 - \theta) \lambda_U^2.
\]

Analogously you can estimate the lower tail dependence of \(C^{12}:

\[
\lambda_{L}^{12} = \lim_{u \downarrow 0} \frac{C^{12}(u, u)}{u} = \lim_{u \downarrow 0} \frac{\theta C^1(u, u) + (1 - \theta)C^2(u, u)}{u} = \theta \lambda_L^1 + (1 - \theta) \lambda_L^2.
\]