國立臺灣大學管理學院財務金融研究所

碩士論文

Department of Finance

College of Management

National Taiwan University

Master Thesis

以混合雙指數 copula 模型定價合成型 CDO

Pricing synthetic CDO with mixed double exponential copula model

陳昱元

Yu-Yuan Chen

指導教授: 繆維中, 管中閔 博士

Advisor: Wei-Chung Miao, Chung-Ming Kuan Ph.D.

中華民國 113 年 7 月

July, 2024

國立臺灣大學碩士學位論文

口試委員會審定書

以混合雙指數 copula 模型定價合成型 CDO

Pricing synthetic CDO with mixed double exponential copula model

本論文係陳昱元君（R11723024）在國立臺灣大學財務金融研究所完成之碩士學位論文，於民國 113 年 7 月 11 日承下列考試委員審查通過及口試及格，特此證明

口試委員：______________________________

（指導教授）

所　　長：______________________________

致謝

謝謝

摘要

CDO（擔保債務憑證）的估值在近期疫情和經濟不確定性背景下，仍然是金融領域的一個重要課題。傳統的 Copula 模型，如單因子高斯 Copula，存在相關性偏斜和缺乏厚尾特徵等問題，導致校準和市場擬合上有顯著誤差。儘管已有各種具有厚尾特性的 Copula 模型被提出，如 Double t Copula 和 NIG Copula，但這些模型有些只能以近似方式求得分布，有些涉及過多的參數，導致計算時間增加和效率低下。各自都有部分缺陷。

本研究引入了高斯分佈和雙指數分佈混合（G-DE）Copula 模型，以解決過往 Copula 模型的局限性。G-DE Copula 模型保留了厚尾特徵，並具有卷積特性，消除了數值近似的需求，並減少了參數估計的複雜性，從而提高了計算效率。更甚是透過混合高斯分布，使其大幅提升市場報價擬合度。我們的研究將 G-DE 模型與傳統 Copula 模型進行比較，展示其在校準準確性、計算效率和整體性能方面的優勢。研究結果顯示，G-DE Copula 模型是 CDO 定價和風險管理的一個強有力的替代方案。

關鍵字：混合雙指數、Copula、因子模型、合成型 CDO

Abstract

The valuation of Collateralized Debt Obligations (CDOs) remains a crucial financial topic, especially in the context of the recent pandemic and economic uncertainties. Traditional copula models, such as the one-factor Gaussian copula, suffer from issues like correlation skew and lack of heavy-tailed characteristics, leading to biases in calibration and market fit. Although various heavy-tailed copula models have been proposed, including the double t copula and NIG copula, they either require complex numerical approximations or involve excessive parameters, resulting in increased computational times and inefficiencies.

This study introduces the Mixtures of Gaussian distribution and Double Exponential distribution (G-DE) copula model, addressing the limitations of previous copula models. The G-DE copula model retains heavy-tailed characteristics while possessing convolution property, eliminating the need for numerical approximations and reducing the complexity of parameter estimation. This results in improved computational efficiency and better market quote fitting. Our research compares the G-DE model with traditional copula models, demonstrating its advantages in calibration accuracy, computational efficiency, and overall performance. The findings highlight the G-DE copula model as a robust alternative for CDO pricing and risk management.

Keywords: Mixed Double Exponential, Copula, Factor model, Synthetic CDO

Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Motivation

After the subprime mortgage crisis of 2008 and the subsequent financial turmoil, issues related to the valuation of Collateralized Debt Obligations (CDOs) gradually diminished. However, with the recent outbreak of the pandemic, insurance companies have incurred significant losses due to unexpected claims on pandemic insurance policies. Moreover, the Federal Reserve's interest rate hikes have significantly impacted insurance companies. Thus, it is evident that statistical and mathematical issues regarding "default correlation" remain crucial.

CDOs, composed of a large number of homogeneous assets, are among the most popular financial assets studied for default correlation. The one-factor Gaussian copula model, initially proposed by Li (1999), was the market standard model due to its simplicity and ease of handling. However, there is a fundamental problem: if this model is used to calculate the implied correlation, we do not obtain the same correlation across the entire structure. This phenomenon is referred to as "correlation skew" in past literature. Furthermore, its lack of heavy-tailed characteristics makes it prone to more biases during calibration, resulting in failure to fit the market quote.

Past literature has proposed a series of copula models with different distributions incorporating tail dependence to address this issue. However, each model has its drawbacks. For instance, the double t copula model proposed by Hull et al. (2004) cannot be computed analytically; the NIG copula proposed by Kalemanova et al. (2007) requires too many parameters, resulting in longer calibration times.

Additionally, past literature has mainly focused on discussing the use of copulas with different distributions, with little discussion on the impact of adding factors. This is because these copula models are based on the mathematical assumption of the Large Homogeneous Portfolio (LHP) Approximation, which assumes that all underlying assets are homogeneous and follow the same correlation coefficient. Adding additional factors would affect the correlation coefficients of the assets by two common factors, allowing the underlying assets to have different correlation coefficients, thus violating the mathematical assumption of LHP. At that point, the computational efficiency advantages brought by LHP would no longer exist. However, increasing factors should also have their benefits, such as allowing the underlying assets to be grouped, with each group having certain correlations internally, while groups have another correlation with each other.

Therefore, this study will explore the effects of a two-factor model and compare various copula models proposed in past literature. Finally, this research will propose the Mixed Double Exponential Distribution Copula model and discuss the benefits it brings.

1.2 Literature review

Previous literature has proposed various distribution copula methods. For instance, the Student t copula in O' Kane and Schloegl (2003), the double t copula in Hull et al.

(2004), the Marshall-Olkin copula in Andersen and Sidenius (2004), and the NIG copula in Kalemanova et al. (2007). Although each of these copula models possesses heavy-tailed characteristics, allowing them to better fit the market quotes, they also have their respective shortcomings.

For example, the double t copula model is not stable under convolution, as noted in Kalemanova et al. (2007) and Choroś-Tomczyk et al. (2014). Therefore, the distribution of the double t-copula needs to be numerically approximated, as it cannot be computed analytically. The significant increase in computational time for such calculations led past literature to seek a different heavy-tailed distribution similar to the double t but stable under convolution. Kalemanova et al. (2007) thus proposed the NIG copula model, which is a generally flexible four-parameter distribution family capable of producing fat tails and skewness. This distribution is convolution stable under certain conditions, and its cumulative distribution function, density, and inverse distribution functions can still be computed sufficiently fast. However, despite the decrease in computational cost, the high number of parameters in the NIG copula model results in a significantly increased computational time required for calibration, and the absolute error among all tranches after calibration is not significantly reduced compared to the double t copula.

After the financial crisis, literature on CDO pricing models gradually diminished, but there are still a few studies focusing on improving their distribution. The relationship between the double t and NIG copula models, both possessing heavy-tailed characteristics, indeed enables a better fit to the market quotes. However, this also leads to certain tranches being overestimated. Xu (2006) first proposed the application of mixed distribution to copula, using the mixture copula model of multi-Gaussian distributions for CDO pricing. Wang et al. (2009) utilized a double mixture of t and Gaussian copula to price the

CDO tranches. Both of these studies found that these mixture copula models fitted better than the previous models. Subsequently, Yang et al. (2009) proposed a Mixtures of Gaussian distribution and NIG distribution (G-NIG) for pricing CDO, while Chen et al. (2014) proposed a Mixtures of Gaussian distribution and Variance Gamma distribution (G-VG). The heavy-tailed characteristics are captured by NIG and VG, while the rest are captured by Gaussian. Such properties resolve the issue of certain tranches being overestimated when using the double t and NIG models. Ultimately, these mixed distribution copula models outperform previous copula models.

In addition to the literature focusing on copula models based on various distributions, there have been successive studies targeting group-related dependencies. Choros et al. (2009) employed a multi-layer copula approach known as hierarchical Archimedean copulae (HAC) to construct a CDO pricing model with sectoral dependence. Okhrin and Xu (2017) utilized convex combination of copulae (cc-copula) to similarly develop a CDO pricing model incorporating sectoral dependence. However, these studies considering sectoral dependence require Monte Carlo methods for computation, lacking closed-form solutions for computational efficiency.

1.3 Contribution

This study will address two main issues. The first issue concerns the impact of adding factors on the model's calibration ability. As the number of factors increases, it becomes possible to construct closed-form solutions with sectoral dependence. This study uses a two-factor Gaussian model as an example, calibrating the model using simulated data and real market data, and then comparing and analyzing the calibration results.

The second issue pertains to the challenges encountered in previous literature when constructing copula models with heavy-tailed characteristics. The distribution of the double t Copula model can only be obtained through approximation, resulting in longer computation times. While the NIG distribution possesses convolution property, its numerous and complex parameters lead to longer calibration times and lack of intuitiveness. The G-NIG model resolves some issues related to overestimation of tranche quotes by using a mixture of Gaussian distributions, resulting in improved model performance. However, the increase in parameters prolongs the calibration process.

Therefore, this paper proposes the Mixtures of Gaussian distribution and Double Exponential distribution (G-DE) copula model. This distribution maintains heavy-tailed characteristics while also possessing convolution property, eliminating the need for approximate distribution acquisition and avoiding excessive parameters. This simplifies computation efficiency and enhances intuitiveness.

The remainder of the study proceeds as follows. In Chapter 2, we will introduce the concept of copulas, various copula models proposed in previous literature, CDO pricing methods, as well as the properties and derivation of the G-DE model we proposed. In Chapter 3, we will discuss the acquisition of CDO tranche data and calibration methods proposed in previous literature. In Chapter 4, we will analyze the calibration results of the two-factor Gaussian copula model using simulated market quotes. In Chapter 5, we will present and analyze the calibration results of various model using real market quotes. Finally, we will conclude the findings from the previous analyses in Chapter 6.

Chapter 2 Methodology

In this chapter, we will introduce the concept of copula, various copula models proposed in previous literature, CDO pricing methods, as well as the properties and derivation of the G-DE model we proposed.

2.1 Copula

Copula is a function which joints marginal distributions into a multivariate distribution. We present below the necessary definitions and useful properties.

2.1.1 Mathematical Definition of Copula

Theorem 2.1.1. *Sklar's Theorem c.f. Sklar (1959)*
For a random vector composed of a set of random variables $(X_1, X_2, \ldots, X_n)$*, there exists a unique copula function that transforms marginal distributions* $F_i(x_i) = P(X_i \leqq x_i)$ *into their joint cumulative distribution function* $H(x_1, x_2, \ldots x_n) = P(X_1 \leqq x_1, \ldots, X_n \leqq x_n)$.

Therefore, we have the following equation.

$$H(x_1, x_2, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) \tag{1}$$

If $F_i(\cdot)$ are continuous, then C is unique.

2.1.2 Types of Copula

Copulas can generally be divided into two categories: Elliptical copulas and Archimedean copulas. The cumulative distribution functions of elliptical copulas, such as Gaussian and Student-t distributions, form the class of elliptical copulas, which have been of high interest in credit risk modeling.

Here, we introduce the Gaussian copula. Let $X_1, \ldots, X_n$ be normally distributed random variables with means $\mu_1, \ldots, \mu_n$, standard deviations $\sigma_1, \ldots, \sigma_n$ and correlation matrix R. The distribution function $\boldsymbol{C_R}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ of the random variables $U_i := \Phi\left(\frac{X_i - \mu_i}{\sigma_i}\right), i \leq n$ is called the Gaussian copula to the correlation matrix **R**. With Sklar's Theorem (2.1.1), we can express the Gaussian copula as the following equation:

$$C_R(u_1, u_2, \ldots, u_n) = \Phi_R\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)\right), u_1, \ldots, u_n \in [0, 1] \tag{2}$$

Another important group are Archimedean copulae. The n-dimensional Archimedean copula function $C : [0, 1]^n \to [0, 1]$ is defined as:

$$C_\psi(u_1, \ldots, u_n) := \psi\left(\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_n)\right), u_1, \ldots, u_n \in [0, 1] \tag{3}$$

where $\psi : [0, \infty) \to [0, 1]$, is called the generator of the copula. A detailed review of the properties of Archimedean copulae can be found in McNeil and Nešlehová (2009)

2.2 CDO pricing

In this section, we present the modeling process of synthetic CDO pricing. We will show the general semi-analytic approach for synthetic CDO pricing following the steps in Kalemanova et al. (2007).

Essentially, pricing a synthetic CDO aims to ascertain the equitable value of every tranche within the structure. The protection buyer of a CDO tranche makes regular spread payments to the protection seller on predetermined dates, with the amount determined by the outstanding notional principal of each tranche. Should a default occur, compensation is provided to the protection buyer based on the loss incurred due to the default event.

Suppose a CDO with n underlying assets. The pricing of a synthetic CDO tranche that takes losses from K_A to K_D (with $0 \leq K_A < K_D \leq 1$) of the reference portfolio works in the same way as the pricing of a credit default swap. K_A and K_D define tranches, which are attachment and detachment points of a tranche expressed in percentage respectively.

Firstly, we denote the percentage portfolio loss by time t given j defaults as follows:

$$L(t, j) = \frac{(1 - \mathrm{R}) \cdot j}{n}, \tag{4}$$

where R denotes recovery rate.

Given the percentage portfolio loss $L(t, j)$ in Equation (4), the loss of the tranche

(K_A, K_D) can be written as:

$$L_{(K_A,K_D)}(t,j) = \frac{\min\left\{\max\left(0, L(t,j) - K_A\right), K_D - K_A\right\}}{K_D - K_A}$$
$$= \begin{cases} 0, & \text{if } L(t,j) < K_A \\ \dfrac{L(t,j) - K_A}{K_D - K_A}, & \text{if } K_A \le L(t,j) \le K_D \\ 1, & \text{if } L(t,j) > K_D \end{cases} \tag{5}$$

Then, the percentage expected loss of the trache (K_A, K_D) can be calculated as follows:

$$EL_{(K_A,K_D)}(t) = \sum_{j=1}^{n} L_{(K_A,K_D)}(t,j) \cdot P(N(t) = j) \tag{6}$$

where $N(t)$ denotes the number of defaults at time t.

We assume a known continuous portfolio loss distribution function, $F(t,x)$, which represents the cumulative density function of the percentage portfolio loss lower than x at time t for $x \in [0,1]$. Then, the percentage expected loss of the tranche (K_A, K_D) can be computed as follows:

$$\begin{aligned} EL_{(K_A,K_D)}(t) &= \frac{1}{K_D - K_A} \int_{K_A}^{1} \left(\min\left(x, K_D\right) - K_A\right) dF(t,x) \\ &= \frac{1}{K_D - K_A} \left(\int_{K_A}^{1} \left(x - K_A\right) dF(t,x) - \int_{K_D}^{1} \left(x - K_D\right) dF(t,x) \right) \end{aligned} \tag{7}$$

The derivation of the loss distribution $F(t,x)$ will be presented in the following section. Let us now focus on the pricing formula of the CDO tranche.

Assume that $t_0, t_1, \ldots, t_{m-1}, t_m = T$ denote the spread payment dates, where T represents the maturity of the synthetic CDO. With the assumption that interest rate, r, is constant, the discount factor is:

$$B\left(t_0, t_k\right) = e^{-r(t_k - t_0)}. \tag{8}$$

In valuing the CDO, it's enough to calculate the cumulative loss distribution at each payment date. This approach simplifies the process by bypassing the need to detail the entire trajectory of the loss process, though it requires specifying the timing of default occurrences between payment dates. While credit events can theoretically happen at any time, for the sake of obtaining a closed-form solution, we adopt a slightly simplified assumption that all defaults occur in the middle of a payment period.

The premium leg is therefore based on the expected average of the outstanding notionals between the two nearest payment dates:

$$\text{Premium Leg} = \sum_{k=1}^{T} B\left(t_0, t_k\right) \cdot \Delta t_k \cdot S^* \cdot \left[1 - \frac{EL_{(K_A,K_D)}\left(t_k\right) + EL_{(K_A,K_D)}\left(t_{k+1}\right)}{2}\right] \tag{9}$$

where $\Delta t_k = t_k - t_{k-1}$, and S^* denotes the breakeven tranche spread.

The protection leg represents the difference between the residual principals at times t_k and t_{k-1}. In practice, compensation is provided immediately after a default occurs. However, for simplicity, we assume that compensation is paid only on the payment date. Consequently, the protection leg is:

$$\text{Protection Leg } = \sum_{k=1}^{T} B\left(t_0, t_k\right) \cdot \left(EL_{(K_A,K_D)}\left(t_k\right) - EL_{(K_A,K_D)}\left(t_{k-1}\right)\right) \tag{10}$$

To obtain the breakeven spread S^* of the CDO tranche, the premium leg (9) should equals the protection leg (10). Thus, this lead to the solution:

$$S^{*}=\frac{\sum_{k=1}^{T}\left[EL_{(K_A,K_D)}\left(t_k\right)-EL_{(K_A,K_D)}\left(t_{k-1}\right)\right]\cdot B\left(t_0,t_k\right)}{\sum_{k=1}^{T}\Delta t_k\left[1-\left(EL_{(K_A,K_D)}\left(t_k\right)+EL_{(K_A,K_D)}\left(t_{k+1}\right)\right)/2\right]\cdot B\left(t_0,t_k\right)} \tag{11}$$

However, certain tranches are priced differently from equation (11). Consider the equity tranche, for example, where the protection buyer pays an upfront fee once, at the inception of the trade and a fixed coupon of 500 basis points throughout the contract's duration. The upfront payment, denoted by $\alpha\left(t_0\right)$, is expressed as a percentage and is quoted in the market. The premium leg of the residual tranche is defined as follows:

$$\begin{aligned}\text{Premium Leg}^{*}&=\alpha\left(t_0\right)\left(K_D-K_A\right)M\\&+\sum_{k=1}^{T}B\left(t_0,t_k\right)\cdot\Delta t_k\cdot 500\cdot\left[1-\frac{EL_{(K_A,K_D)}\left(t_k\right)+EL_{(K_A,K_D)}\left(t_{k+1}\right)}{2}\right]\end{aligned} \tag{12}$$

where M denotes notional principal.

Again, according to the non-arbitrage assumption, the premium leg (12) should equals the protection leg (10). Thus, we can obtain the upfront payment $\alpha\left(t_0\right)$ as follows:

$$\alpha\left(t_0\right)=\frac{100}{K_D-K_A}\left[\text{Protection Leg}-500\cdot\text{Premium Leg}^{*}\right] \tag{13}$$

So far, we have introduced the pricing model for CDOs but have not yet detailed the derivation of the portfolio loss distribution mentioned in Equation (7). Thus, in the following sections, we introduce the factor copula model for correlated defaults and an analytical approximation method to calculate the expected loss of a tranche.

2.3 Gaussian copula model

Li (1999) introduced the One-Factor Gaussian Copula Model, which subsequently became the market standard. The following section demonstrates how to use this model to derive the loss distribution of the reference portfolio.

2.3.1 One-factor Gaussian copula model

We assume that the portfolio consists of n equally weighted underlying assets. The asset return of the i-th instrument, serving as the default indicator, can be expressed as follows:

$$x_i = a_i M + \sqrt{1 - a_i^2} Z_i \tag{14}$$

where $i = 1, ..., n$, M denotes the common factor, and Z_i denotes the idiosyncratic factor.

Both the common factor and idiosyncratic factor are independent random variables that follow Gaussian distributions. The correlation between x_i and x_j is represented by $a_i a_j$. Moreover, owing to the stability property of Gaussian distributions under convolution, the asset return x_i also follows a standard normal distribution.

In this study, we assume the marginal default distribution of the underlying assets follows a constant default intensity model. Consequently, the default probability of i-th asset before time t can be expressed as follows:

$$Q_i(t) = 1 - e^{-\lambda_i t} \tag{15}$$

where λ_i denotes the default intensity of i-th asset.

By Gaussian copula mapping[1], which is a percentile-to-percentile transformation, the default threshold can be efficiently derived as follows:

$$C_i(t) = \Phi^{-1}\left(Q_i(t)\right), \tag{16}$$

If $x_i(t) \leq C_i(t)$, it triggers a default event for this asset. Thus, the default probability of the i-th asset can be calculated as follows:

$$\begin{aligned} P(\tau_i \leq t \mid M) &= P\left[x_i \leq C_i(t) \mid M\right] \\ &= P\left[a_i M + \sqrt{1-a_i^2} Z_i \leq C_i(t) \mid M\right] \\ &= P\left[Z_i \leq \frac{C_i(t) - a_i M}{\sqrt{1-a_i^2}}\right] \\ &= \Phi\left(\frac{C_i(t) - a_i M}{\sqrt{1-a_i^2}}\right) \end{aligned} \tag{17}$$

where τ_i represents the default time of the i-th reference entity with marginal distribution $Q_i(t)$.

Assuming the portfolio is homogeneous, where $a_i = a$ and $C_i(t) = C(t)$ for all i, and the notional amounts and recovery rate R are uniform across all issuers, then the conditional default probability of all issuers in the portfolio given M is expressed as:

$$P(\tau \leq t \mid M) = \Phi\left(\frac{C(t) - aM}{\sqrt{1-a^2}}\right) \tag{18}$$

For simplicity, when the recovery rate $R = 0$, if j out of n underlying instruments default, the conditional default probability of experiencing a $\frac{j}{n}$ percentage loss of the port-

[1]The details of the transformation can be found in Hull et al. (2009)

folio is given by:

$$P\left[L(t) = \frac{j}{n}\middle| M\right] = \binom{n}{j}\Phi\left(\frac{C(t) - aM}{\sqrt{1-a^2}}\right)^j \left[1 - \Phi\left(\frac{C(t) - aM}{\sqrt{1-a^2}}\right)\right]^{n-j} \tag{19}$$

The unconditional loss distribution $P\left[L(t) = \frac{j}{n}\right]$ can be derived by integrating equation (19) with respect to the distribution of the common factor M as follows:

$$P\left[L(t) = \frac{j}{n}\right] = \int_{-\infty}^{\infty} \binom{n}{j}\Phi\left(\frac{C(t) - am}{\sqrt{1-a^2}}\right)^j \left(1 - \Phi\left(\frac{C(t) - am}{\sqrt{1-a^2}}\right)\right)^{n-j} \cdot \phi(m)dm \tag{20}$$

2.3.2 Gaussian Hermite Quadrature

The integration required for the loss distribution in equation (20) can be conducted using a technique known as Gaussian quadrature, as described by Hull et al. (2009). Gaussian Hermite Quadrature[2] approximating the integral by the formula:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-F^2/2} g(F)dy \approx \sum_{k=1}^{M} w_k g\left(F_k\right) \tag{21}$$

However, due to the computational complexity, particularly for large j in equation (20), it is advantageous to employ an approximation. The Large Homogeneous Portfolio approximation, as proposed by Vasicek (2002), presents a straightforward yet effective solution.

[2] The parameters w_k and F_k are determined based on the roots of Hermite polynomials. For further details on Gaussian quadrature, refer to Technical Note 21 described by Hull et al. (2009)

2.3.3 LHP Approximation Method

LHP Approximation assume an infinite number of homogeneous assets with identical pairwise default correlation in the reference portfolio. To derive the distribution of the percentage portfolio loss $F(t,x)$ as referenced in equation (7), we outline the methodology employed in Kalemanova et al. (2007).

We examine the cumulative probability of the percentage portfolio loss not exceeding x for $x \in [0,1]$, defined as:

$$F_n(t,x) = \sum_{j=0}^{[nx]} P\left[L(t) = \frac{j}{n}\right] \tag{22}$$

By substituting $s = \Phi\left(\frac{C(t)-am}{\sqrt{1-a^2}}\right)$ and incorporating equation (20) we derive the following expression for $F_n(t,x)$:

$$F_n(t,x) = -\int_0^1 \sum_{j=0}^{[nx]} \binom{n}{j} s^j (1-s)^{n-j} d\Phi\left(\frac{C(t) - \sqrt{1-a^2}\Phi^{-1}(s)}{a}\right) \tag{23}$$

Under LHP approximation, the number of underlying CDSs approaches infinity, the portfolio loss distribution can be expressed as follows:

$$F_\infty(t,x) = \lim_{n\to\infty}\left[-\int_0^1 \sum_{j=0}^{[nx]} \binom{n}{j} s^j (1-s)^{n-j} d\Phi\left(\frac{C(t) - \sqrt{1-a^2}\Phi^{-1}(s)}{a}\right)\right] \tag{24}$$

Since

$$\lim_{n\to\infty} \sum_{k=0}^{[nx]} \binom{n}{j} s^j (1-s)^{n-j} = \begin{cases} 0, \text{ if } x < s \\ 1, \text{ if } x > s \end{cases} \tag{25}$$

the portfolio loss distribution can finally be calculated as follows:

$$\begin{aligned} F_\infty(t,x) &= -\int_0^x d\Phi\left(\frac{C(t)-\sqrt{1-a^2}\Phi^{-1}(s)}{a}\right) \\ &= 1-\Phi\left(\frac{C(t)-\sqrt{1-a^2}\Phi^{-1}(x)}{a}\right) \\ &= \Phi\left(\frac{\sqrt{1-a^2}\Phi^{-1}(x)-C(t)}{a}\right) \end{aligned} \tag{26}$$

By differentiating equation (26), we may get the following formula:

$$f_\infty(t,x) = \frac{\sqrt{1-a^2}}{a}\frac{\phi\left(\frac{\sqrt{1-a^2}\Phi^{-1}(x)-C(t)}{a}\right)}{\phi\left(\Phi^{-1}(x)\right)}. \tag{27}$$

After the approximating distribution of portfolio loss is calculated, according to equation (7), it is possible to compute the expected tranche loss. To compute $EL_{(K_A,K_D)}(t)$, we first introduce a new variable to facilitate the change of variables. Let

$$y = \Phi^{-1}(x) \tag{28}$$

then

$$dy = \frac{dx}{\phi(\Phi^{-1}(x))} \tag{29}$$

Using this change of variables, we can rewrite the expected tranche loss as follows:

$$\begin{aligned} EL_{(K_A,K_D)}(t) &= \frac{1}{K_D-K_A}\left(\int_{K_A}^1 (x-K_A)\,dF_\infty(t,x) - \int_{K_D}^1 (x-K_D)\,dF_\infty(t,x)\right) \\ &= \frac{1}{K_D-K_A}\left\{\begin{array}{l} \int_{\Phi^{-1}(K_A)}^{\Phi^{-1}(1)} (\Phi(y)-K_A)\,\phi\left(\frac{\sqrt{1-a^2}y-C(t)}{a}\right)\frac{\sqrt{1-a^2}}{a}dy \\ -\int_{\Phi^{-1}(K_D)}^{\Phi^{-1}(1)} (\Phi(y)-K_D)\,\phi\left(\frac{\sqrt{1-a^2}y-C(t)}{a}\right)\frac{\sqrt{1-a^2}}{a}dy \end{array}\right\} \end{aligned} \tag{30}$$

Next, we we consider the scenario where the recovery rate R is non-zero. In this case, the complete loss of the tranche, representing K, occurs only when assets total amount of

$\frac{K}{1-R}$ have defaulted. Consequently, the expected loss of the tranche ranging from K to 1 is determined as follows:

$$EL_{(K,1)}(t) = \int_{\frac{K}{1-R}}^{1} (1-R)\left(x - \frac{K}{1-R}\right) dF_{\infty}(t,x) \tag{31}$$

From the derivation of equation (31), we can rewrite equation (30) as follows:

$$EL_{(K_A,K_D)}(t) = \frac{1}{K_D - K_A}\left\{\begin{array}{l} \int_{\Phi^{-1}\left(\frac{K_A}{1-R}\right)}^{\Phi^{-1}(1)} \left((1-R)\Phi(y) - K_A\right)\phi\left(\frac{\sqrt{1-a^2}y - C(t)}{a}\right)\frac{\sqrt{1-a^2}}{a}dy \\ -\int_{\Phi^{-1}\left(\frac{K_D}{1-R}\right)}^{\Phi^{-1}(1)} \left((1-R)\Phi(y) - K_D\right)\phi\left(\frac{\sqrt{1-a^2}y - C(t)}{a}\right)\frac{\sqrt{1-a^2}}{a}dy \end{array}\right\} \tag{32}$$

Finally, according to equation (11) and equation (13), the quote of CDO can be determined.

2.3.4 Two-factor Gaussian copula model

When utilizing the one-factor copula model, asset returns are influenced by only one common factor. Under this constraint, we are unable to construct more complex correlation matrix. Therefore, increasing the number of common factors allows for varying correlations among the underlying assets, facilitating the construction of more intricate dependency structures. For instance, the two-factor copula model can describe a correlation matrix divided into two groups.

As in equation (14) but with slight modifications, we assume that all underlying assets can be divided into two groups. The asset return of the issuer in i-th group in the portfolio can be expressed as follows:

$$x_i = a_i F_a + \beta_i F_b + \sqrt{1 - a_i^2 - \beta_i^2} Z_i \tag{33}$$

where $i = 1, 2$. F_a, F_b denotes the common factor, and Z_i denotes the idiosyncratic factor.

F_a, F_b, and Z_i are mutually independent random variables that follow Gaussian distributions. The correlation structure between x_i is dependent on two common factor F_a, F_b. To be specific, the correlation between x_i and x_j is $a_i a_j + \beta_i \beta_j$.

By doing so, we can describe the situation where the underlying assets have varying correlations with each other. Within each group (intra-class), the assets have a higher correlation, while between groups (inter-class), they have a lower correlation. If both assets come from the same group, i.e., x_1 and x_1 or x_2 and x_2, their correlation is $a_1 a_1 + \beta_1 \beta_1$ or $a_2 a_2 + \beta_2 \beta_2$. If the two assets come from different groups, i.e., x_1, x_2, their correlation is $a_1 a_2 + \beta_1 \beta_2$.

Similar to the derivation of the equation (17), we can then express conditional default probability of the issuer in i-th group in the portfolio given common factors F_a, F_b as follows:

$$Q_i = P(\tau_i \leq t \mid F_a, F_b) = \Phi\left(\frac{C_i(t) - a_i F_a - \beta_i F_b}{\sqrt{1 - a_i^2 - \beta_i^2}}\right) \tag{34}$$

Thus, the final unconditional loss distribution can be derived in a manner similar to equation (20), by integrating over the common factors F_a and F_b, yielding the following:

$$\begin{aligned} P\left[L(t) = \frac{j}{n}\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{j_1 + j_2 = j} \binom{n_1}{j_1} \binom{n_2}{j_2} \cdot Q_1^{j_1} \cdot (1 - Q_1)^{n_1 - j_1} \\ \cdot Q_2^{j_2} \cdot (1 - Q_2)^{n_2 - j_2} \cdot \phi\left(f_a, f_b\right) df_a df_b \end{aligned} \tag{35}$$

where $\phi\left(f_a, f_b\right)$ denotes bivariate probability density function of standard normal distribution with $\rho = 0$ due to independence of common factors.

However, equation (35) cannot be computed using the LHP approximation. This is because the LHP assumption requires all underlying assets to be homogeneous, leading to them being correlated with the same correlation coefficient. Therefore, we can only integrate this equation numerically, specifically using the Gaussian Quadrature approach. Additionally, since this equation involves a double integral, we need Two-dimensional Gaussian Hermite Quadrature.

2.3.5 Two dimensional Gaussian Hermite Quadrature

Similar to the formula mentioned in equation (21), the Gaussian Hermite Quadrature in two-dimensional case can be expressed as follows:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)e^{-\left(x^2+y^2\right)}dxdy \approx \sum_{i=1}^{M}\sum_{j=1}^{M} g\left(x_i, y_j\right) w_i w_j \tag{36}$$

However, the computation described in equation (36) is time-consuming. Therefore, Jäckel (2005) proposed a pruning method to enhance the efficiency of multivariate Gaussian Hermite Quadrature without sacrificing precision:

$$\sum_{i=1,j=1}^{M,M} \mathbf{1}_{(w_i w_j > \theta_M)} \cdot w_i w_j f\left(a z_i + b z_j, b z_i + a z_j\right) \tag{37}$$

where

$$\theta_M := \frac{w_1 \cdot w_{\left[\frac{M+1}{2}\right]}}{M}, a := \frac{\sqrt{1+\rho}+\sqrt{1-\rho}}{2}, b := \frac{\sqrt{1+\rho}-\sqrt{1-\rho}}{2}.$$

By employing this method, dropping all points that would be weighted with a net weight below θ_M, it is possible to effectively reduce the computational complexity associated with double integration:

2.4 Double t copula model

The one-factor double t model, proposed by Hull et al. (2004), assumes Student t-distributions for both the common market factor M and the individual factors Z_i. The default times in this model are generated from the following equation:

$$x_i = a_i\sqrt{\frac{\nu_M - 2}{\nu_M}}M + \sqrt{1-a_i^2}\sqrt{\frac{\nu_Z - 2}{\nu_Z}}Z_i \tag{38}$$

where M and Z_i are t-distributed with ν_M and ν_Z degrees of freedom, respectively.

Since the Student t distribution is not stable under convolution, x_i are not t-distributed, and the copula is not a Student t copula. The default thresholds are then determined by:

$$C_i(t) = T^{*-1}(Q_i(t)) \tag{39}$$

where T^* denotes the distribution of x_i.

Consequently, the loss distribution $F_\infty(t,x)$ under LHP in equation (26) is given by:

$$F_\infty(t,x) = T\left(\frac{\sqrt{1-a^2}T^{-1}(x) - C(t)}{a}\right) \tag{40}$$

where T denotes the Student t-distribution function.

However, it is time-consuming to solve the integral in equation (7) analytically since the distribution T^* of x_i needs to be approximated numerically. Therefore, we need a distribution that exhibits heavy-tail properties and is also stable under convolution to address this issue.

2.5 NIG copula model

The one-factor NIG model, introduced by Kalemanova et al. (2007), utilizes NIG distributions for both the common market factor M and the individual factors Z_i. Unlike the double double t copula model proposed by Hull et al. (2004), the NIG distribution is stable under convolution. Before giving the one-factor NIG copula model, let us introduce the NIG distribution first.

The NIG (Normal Inverse Gaussian) distribution is a mixture of normal and inverse Gaussian distribution. A random variable X follows a NIG distribution with parameters α, β, μ and δ if its density function is of the following form:

$$f_{\mathcal{NIG}}(x; \alpha, \beta, \mu, \delta) = \frac{\delta\alpha \exp(\delta\gamma + \beta(x-\mu))}{\pi\sqrt{\delta^2 + (x-\mu)^2}} K_1\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right) \tag{41}$$

where $K_1(w) := \frac{1}{2}\int_0^{\infty} \exp\left(-\frac{1}{2}w\left(t + t^{-1}\right)\right) dt, 0 \leq |\beta| < \alpha$ and $\delta > 0$.

We denote the NIG distribution by $NIG(\alpha, \beta, \mu, \delta)$. If a random variable $X \sim NIG(\alpha, \beta, \mu, \delta)$, then the central moments of X can be calculated as follows:

$$\begin{aligned} &E[X] = \mu + \delta\frac{\beta}{\gamma}, \quad \text{Var}[X] = \delta\frac{\alpha^2}{\gamma^3} \\ &\text{Skew}[X] = 3\frac{\beta}{\alpha \cdot \sqrt{\delta\gamma}}, \quad \text{Kurt}[X] = 3 + 3\left(1 + 4\left(\frac{\beta}{\alpha}\right)^2\right)\frac{1}{\delta\gamma} \end{aligned} \tag{42}$$

where $\gamma := \sqrt{\alpha^2 - \beta^2}$

Next, we will introduce two important properties of the NIG distribution: the scaling property and the convolution property, the latter of which significantly enhances the computational efficiency compared to the double t copula model.

The scaling property is as follows:

$$X \sim NIG(\alpha, \beta, \mu, \delta) \Rightarrow cX \sim NIG\left(\frac{\alpha}{c}, \frac{\beta}{c}, c\mu, c\delta\right) \quad \text{for } c \in R, \tag{43}$$

The convolution property for independent random variables X_1 and X_2 is as follows:

$$\begin{aligned} & X_1 \sim NIG(\alpha, \beta, \mu_1, \delta_1) \quad \text{and} \quad X_2 \sim NIG(\alpha, \beta, \mu_2, \delta_2) \\ \Rightarrow \quad & X_1 + X_2 \sim NIG(\alpha, \beta, \mu_1 + \mu_2, \delta_1 + \delta_2) \end{aligned} \tag{44}$$

With these two properties considered, we can formulate the one-factor NIG copula model. The default times are then generated according to the following form:

$$x_i = a_i M + \sqrt{1 - a_i^2} Z_i \tag{45}$$

where M and Z_i are independent with random variables with

$$M \sim NIG\left(\alpha, \beta, -\frac{\beta\gamma^2}{\alpha^2}, \frac{\gamma^3}{\alpha^2}\right) \tag{46}$$

$$Z_i \sim NIG\left(\frac{\sqrt{1-a_i^2}}{a_i}\alpha, \frac{\sqrt{1-a_i^2}}{a_i}\beta, -\frac{\sqrt{1-a_i^2}}{a_i}\frac{\beta\gamma^2}{\alpha^2}, \frac{\sqrt{1-a_i^2}}{a_i}\frac{\gamma^3}{\alpha^2}\right) \tag{47}$$

Therefore, due to the properties outlined in quation (43) and equation (44), the asset return of the i-th instrument x_i follows a NIG distribution as follows:

$$x_i \sim NIG\left(\frac{\alpha}{a}, \frac{\beta}{a}, -\frac{1}{a}\frac{\beta\gamma^2}{\alpha^2}, \frac{1}{a}\frac{\gamma^3}{\alpha^2}\right) \tag{48}$$

For simplicity, we introduce the notation $F_{\mathcal{NIG}(s)}(x)$ to represent the distribution function $F_{\mathcal{NIG}}\left(x; s\alpha, s\beta, -s\frac{\beta\gamma^2}{\alpha^2}, s\frac{\gamma^3}{\alpha^2}\right)$. For instance, the distribution function of the common factor M is denoted as $F_{\mathcal{NIG}(1)}(x)$, and for the idiosyncratic factor Z_i, it is represented as $F_{\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}(x)$.

Then, following the same procedure as in equation (26), the loss distribution $F_\infty(t,x)$ can be expressed as follows:

$$F_\infty(t,x) = 1 - F_{\mathcal{NIG}(1)}\left(\frac{C(t) - \sqrt{1-a^2}F^{-1}_{\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}(x)}{a}\right) \tag{49}$$

where $C(t) = F^{-1}_{\mathcal{NIG}\left(\frac{1}{a}\right)}(Q(t))$, which denotes the default threshold based on NIG distribution.

By differentiating equation (49), we may get the following formula:

$$f_\infty(t,x) = \frac{\sqrt{1-a^2}}{a}\frac{f_{\mathcal{NIG}(1)}\left(\frac{C(t)-\sqrt{1-a^2}F^{-1}_{\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}(x)}{a}\right)}{f_{\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}\left(F^{-1}_{\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}(x)\right)}. \tag{50}$$

Again, following the same method used in equation (28) and equation (29), we define a new variable and differentiate it in order to facilitate the change of variables as follows:

$$\begin{aligned} y &= F^{-1}_{\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}(x), \\ dy &= \frac{dx}{f_{\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}\left(F^{-1}_{\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}(x)\right)} \end{aligned} \tag{51}$$

Therefore, we can calculate the expected tranche loss as follows:

$$
\begin{aligned}
&EL_{(K_A,K_D)}(t)\\
&= \frac{1}{K_D-K_A}\left\{\begin{array}{l}
\int_B^C \left((1-R)F_{\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}(y)-K_A\right) f_{\mathcal{NIG}(1)}\left(\frac{C(t)-\sqrt{1-a^2}y}{a}\right)\frac{\sqrt{1-a^2}}{a}dy\\
-\int_A^C \left((1-R)F_{\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}(y)-K_D\right) f_{\mathcal{NIG}(1)}\left(\frac{C(t)-\sqrt{1-a^2}y}{a}\right)\frac{\sqrt{1-a^2}}{a}dy
\end{array}\right\}
\end{aligned}
\tag{52}
$$

where $A = F^{-1}_{\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}\left(\frac{K_D}{1-R}\right)$, $\quad B = F^{-1}_{\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}\left(\frac{K_A}{1-R}\right)$, $\quad C = F^{-1}_{\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}(1)$.

2.6 G-NIG copula model

The one-factor G-NIG model, first proposed by Yang et al. (2009), utilizes a mixture of Gaussian distribution and NIG distribution for both the common market factor M and the individual factors Z_i.

Through mixing with a Gaussian distribution, the G-NIG copula model can improve the performance of previous models (e.g., double t, NIG) that exhibit deficiencies in certain scenarios due to their heavy-tailed characteristics. For instance, if the Gaussian distribution overestimates certain tranches while the NIG underestimates them, adjusting the proportion of the Gaussian distribution can accurately calibrate the CDO prices that were previously underestimated by the NIG model.

The G-NIG distribution is a mixture of a Gaussian distribution and a NIG distribution. A random variable $X \sim G-NIG(0,1;\alpha,\beta,\mu,\delta;p)$ is given by:

$$
X = \begin{cases} V, & \text{with probability } p \\ U, & \text{with probability } 1-p \end{cases}
\tag{53}
$$

where $V \sim N(0,1)$, $U \sim NIG(x;\alpha,\beta,\mu,\delta)$, and $p \in (0,1)$ represents the proportion of the Gaussian component in the mixture distribution. Denote the probability density function of X as $f_{G-\mathcal{NIG}}(x;0,1;\alpha,\beta,\mu,\delta;p)$, then

$$f_{G-\mathcal{NIG}}(x;0,1;\alpha,\beta,\mu,\delta;p) = \frac{p}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right) + (1-p) f_{\mathcal{NIG}}(x;\alpha,\beta,\mu,\delta) \tag{54}$$

Then, the one-factor G-NIG copula model can be formulated as follows:

$$x_i = a_i M + \sqrt{1-a_i^2} Z_i \tag{55}$$

where M and Z_i are independent with random variables with

$$M \sim G-NIG\left(0,1;\alpha,\beta,-\frac{\beta\gamma^2}{\alpha^2},\frac{\gamma^3}{\alpha^2};p\right) \tag{56}$$

$$Z_i \sim G-NIG\left(0,1;\frac{\sqrt{1-a_i^2}}{a_i}\alpha,\frac{\sqrt{1-a_i^2}}{a_i}\beta,-\frac{\sqrt{1-a_i^2}}{a_i}\frac{\beta\gamma^2}{\alpha^2},\frac{\sqrt{1-a_i^2}}{a_i}\frac{\gamma^3}{\alpha^2};p\right) \tag{57}$$

Again, for simplicity, we introduce the notation $F_{G-\mathcal{NIG}(s)}(x)$ to represent the distribution function $F_{G-\mathcal{NIG}}\left(x;0,1;s\alpha,s\beta,-s\frac{\beta\gamma^2}{\alpha^2},s\frac{\gamma^3}{\alpha^2};p\right)$. Then, following the same procedure as in equation (52), the expected tranche loss $EL_{(K_A,K_D)}(t)$ can be expressed as follows:

$$EL_{(K_A,K_D)}(t)$$
$$= \frac{1}{K_D-K_A}\left\{\begin{array}{l} \int_B^C \left((1-R)F_{G-\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}(y) - K_A\right) f_{G-\mathcal{NIG}(1)}\left(\frac{C(t)-\sqrt{1-a^2}y}{a}\right)\frac{\sqrt{1-a^2}}{a}dy \\ -\int_A^C \left((1-R)F_{G-\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}(y) - K_D\right) f_{G-\mathcal{NIG}(1)}\left(\frac{C(t)-\sqrt{1-a^2}y}{a}\right)\frac{\sqrt{1-a^2}}{a}dy \end{array}\right\} \tag{58}$$

where $A = F^{-1}_{G-\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}\left(\frac{K_D}{1-R}\right)$, $\quad B = F^{-1}_{G-\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}\left(\frac{K_A}{1-R}\right)$, $\quad C = F^{-1}_{G-\mathcal{NIG}\left(\frac{\sqrt{1-a^2}}{a}\right)}(1)$, and $C(t) = F^{-1}_{G-\mathcal{NIG}\left(\frac{1}{a}\right)}(Q(t))$.

2.7 G-DE copula model

Before we begin, we can summarize briefly. While the Gaussian copula serves as a standard model in the market due to its computational efficiency, it lacks heavy-tail characteristics, making it ineffective in fitting the market quotes accurately. Although the double t copula exhibits heavy-tail properties, its lack of convolution property leads to poor computational efficiency. On the other hand, the NIG copula possesses both heavy-tail characteristics and convolution property, requiring significantly less time for single computation compared to the double t copula. However, the introduction of additional parameters, such as α and β, in NIG leads to a significant increase in the time required for calibration and complicates the model. Finally, the G-NIG copula, by utilizing a mixture of Gaussian distributions, can fit the market quotes more precisely compared to NIG. However, similar to NIG, the G-NIG model suffers from complexity due to the excessive number of parameters, and the calibration time may even exceed that of NIG by several folds.

Therefore, in this section, we introduce the G-DE copula model proposed by us. It possesses heavy-tail characteristics and convolution property, resulting in excellent computational efficiency and a more accurate fit to market quotes. Furthermore, unlike NIG and G-NIG, it avoids excessive parameterization of the distribution while benefiting from the advantages of a mixed Gaussian distribution, enabling it to fit the market quotes more precisely. It is a model that offers various advantages in different aspects.

The one-factor G-DE model utilizes a mixture of Gaussian distribution and Double

Exponential distribution, also called Laplace distribution, for both both the common market factor M and the individual factors Z_i. Before giving the one-factor G-DE copula model, let us introduce the DE distribution first.

2.7.1 DE copula model

The DE distribution can be thought of as two exponential distributions spliced together along the abscissa. A random variable $X \sim DE(\mu, b)$ if its probability density function is:

$$f_{DE}(x;\mu,b) = \frac{1}{2b}\exp\left(-\frac{|x-\mu|}{b}\right) \tag{59}$$

where μ is a location parameter, and $b > 0$, which is sometimes referred to as the diversity, is a scale parameter.

Moreover, the double exponential distribution is straightforward to integrate because it utilizes the absolute value function. Its cumulative distribution function is as follows:

$$\begin{aligned} F_{DE}(x;\mu,b) &= \begin{cases} \frac{1}{2}\exp\left(\frac{x-\mu}{b}\right) & \text{if } x < \mu \\ 1 - \frac{1}{2}\exp\left(-\frac{x-\mu}{b}\right) & \text{if } x \geq \mu \end{cases} \\ &= \frac{1}{2} + \frac{1}{2}\operatorname{sgn}(x-\mu)\left(1 - \exp\left(-\frac{|x-\mu|}{b}\right)\right). \end{aligned} \tag{60}$$

Additionally, the central moments of X can be calculated as follows:

$$\begin{aligned} &E[X] = \mu, \quad \text{Var}[X] = 2b^2 \\ &\text{Skew}[X] = 0, \quad \text{Kurt}[X] = 6 \end{aligned} \tag{61}$$

Then, let us introduce the scaling property and the convolution property of the double exponential distribution. Although the sum of two double-exponentially distributed random variables does not follow a double exponential distribution, it is still stable under convolution.

The scaling property is as follows:

$$X \sim DE(\mu, b) \Rightarrow cX \sim DE\left(c\mu, cb\right) \text{ for } c \in R \tag{62}$$

Next, we consider the convolution property. When using a one-factor copula model, the common factor M and the idiosyncratic factor Z_i are typically assumed to follow distributions with zero mean and unit variance for the purpose of standardizing the distributions of both factors. For the double exponential (DE) distribution, the standardized DE has $\mu = 0$ and $b = \frac{1}{\sqrt{2}}$. Consequently, when calculating the probability density function of the sum of two DE-distributed variables, we only need to consider the case where $\mu = 0$.

Let X and Y both follow a DE distribution with zero mean. The pdf of the sum of these two random variables, denoted as ”2DE”, can be expressed as follows:

$$\begin{aligned}
&X \sim DE(0, a) = \frac{1}{2a} exp(-\frac{|x|}{a}) \\
&Y \sim DE(0, b) = \frac{1}{2b} exp(-\frac{|y|}{b}) \\
&f_{Z=X+Y}(z; a, b) = f_{2DE}(z; a, b) = \frac{1}{2} \frac{a \cdot exp(-\frac{|z|}{a}) - b \cdot exp(-\frac{|z|}{b})}{a^2 - b^2}
\end{aligned} \tag{63}$$

With equation (62) and equation (63) these two properties considered, we can formulate the one-factor DE copula model. Suppose that the i-th asset value X_i of the reference entities is given by

$$x_i = a_i M + \sqrt{1 - a_i^2} Z_i \tag{64}$$

where M and Z_i are standardized and independent with random variables with

$$
\begin{aligned}
M &\sim DE(0, \frac{1}{\sqrt{2}}) \\
Z_i &\sim DE(0, \frac{1}{\sqrt{2}})
\end{aligned}
\tag{65}
$$

The asset return of the i-th instrument, x_i, follows a 2DE distribution given by:

$$
x_i \sim 2DE\left(\frac{a_i}{\sqrt{2}}, \frac{\sqrt{1-a_i^2}}{\sqrt{2}}\right) \tag{66}
$$

with the corresponding pdf and cdf:

$$
\begin{aligned}
f_{2DE}(x; \frac{a_i}{\sqrt{2}}, \frac{\sqrt{1-a_i^2}}{\sqrt{2}}) &= \frac{1}{\sqrt{2}} \frac{a_i \cdot exp(-\frac{\sqrt{2}|x|}{a_i}) - \sqrt{1-a_i^2} \cdot exp(-\frac{\sqrt{2}|x|}{\sqrt{1-a_i^2}})}{2a_i^2 - 1} \\
F_{2DE}(x; \frac{a_i}{\sqrt{2}}, \frac{\sqrt{1-a_i^2}}{\sqrt{2}}) &= \begin{cases} \frac{1}{4a_i^2-2} \left\{ a_i^2 \cdot exp(-\frac{\sqrt{2}x}{a_i}) - (1-a_i^2) \cdot exp(-\frac{\sqrt{2}x}{\sqrt{1-a_i^2}}) \right\} \text{if } x < 0 \\ 1 - \frac{1}{4a_i^2-2} \left\{ a_i^2 \cdot exp(-\frac{\sqrt{2}x}{a_i}) - (1-a_i^2) \cdot exp(-\frac{\sqrt{2}x}{\sqrt{1-a_i^2}}) \right\} \text{if } x \geq 0 \end{cases}
\end{aligned}
\tag{67}
$$

The proof of this result can be found in Appendix A.1

Though the density function of the 2DE distribution is complicated, its moment generating function takes on a simpler form. Specifically, for the case of zero mean, the moment generating function of the DE distribution is expressed as:

$$
M_{DE}(t; 0, b) = \frac{1}{1 - b^2 t^2} \text{ for } |t| < \frac{1}{b} \tag{68}
$$

Additionally, the moment generating function of the sum of two random variables is obtained by multiplying the MGFs of the individual random variables. Thus, the moment generating function of the 2DE distribution, which is derived from the sum of a times the common factor and $\sqrt{1-a^2}$ times the idiosyncratic factor, can be expressed as follows:

$$M_{2DE}(t; \frac{a_i}{\sqrt{2}}, \frac{\sqrt{1-a_i^2}}{\sqrt{2}}) = \frac{1}{\left(1 - \frac{a^2t^2}{2}\right)} \cdot \frac{1}{\left(1 - \frac{(1-a^2)t^2}{2}\right)} \tag{69}$$

Considering a random variable $X \sim 2DE(\frac{a}{\sqrt{2}}, \frac{\sqrt{1-a^2}}{\sqrt{2}})$, through the derivation of the moment generating function in equation (69), we can obtain the central moments of the 2DE distribution as follows:

$$\begin{aligned} &E[X] = 0, \quad \text{Var}[X] = 1 \\ &\text{Skew}[X] = 0, \quad \text{Kurt}[X] = 6(a^4 - a^2 + 1) \end{aligned} \tag{70}$$

The proof of these central moments can be found in Appendix A.2

So far, we've determined that x_i follows a standardized 2DE distribution with zero mean and unit variance. Moreover, its kurtosis is $6(a^4 - a^2 + 1)$, indicating heavy-tail characteristic.

For simplicity, we introduce the notation $F_{DE(1)}(x)$ to represent the standard double exponential distribution function $F_{DE}\left(x; 0, \frac{1}{\sqrt{2}}\right)$, and $F_{2DE(a)}(x)$ to represent the distribution function $F_{2DE}\left(x; \frac{a}{\sqrt{2}}, \frac{\sqrt{1-a^2}}{\sqrt{2}}\right)$.

Following the same procedure as in equation (49), the loss distribution $F_\infty(t, x)$ can be expressed as follows:

$$F_\infty(t, x) = 1 - F_{DE(1)}\left(\frac{C(t) - \sqrt{1-a^2}F_{DE(1)}^{-1}(x)}{a}\right) \tag{71}$$

where $C(t) = F_{2DE(a)}^{-1}(Q(t))$, which denotes the default threshold based on 2DE distribution.

Then, following the same procedure as in equation (52), the expected tranche loss $EL_{(K_A,K_D)}(t)$ can be expressed as follows:

$$EL_{(K_A,K_D)}(t)$$
$$= \frac{1}{K_D - K_A} \left\{ \begin{array}{l} \int_B^C \left((1-R)F_{DE(1)}(y) - K_A\right) f_{DE(1)}\left(\frac{C(t)-\sqrt{1-a^2}y}{a}\right) \frac{\sqrt{1-a^2}}{a} dy \\ -\int_A^C \left((1-R)F_{DE(1)}(y) - K_D\right) f_{DE(1)}\left(\frac{C(t)-\sqrt{1-a^2}y}{a}\right) \frac{\sqrt{1-a^2}}{a} dy \end{array} \right\} \tag{72}$$

where $A = F^{-1}_{DE(1)}\left(\frac{K_D}{1-R}\right), \quad B = F^{-1}_{DE(1)}\left(\frac{K_A}{1-R}\right), \quad C = F^{-1}_{DE(1)}(1)$.

2.7.2 G-DE copula model

The G-DE distribution is a mixture of a Gaussian distribution and a DE distribution. A random variable $X \sim G - DE(0, 1; \mu, b; p)$ is given by:

$$X = \begin{cases} V, & \text{with probability } p \\ U, & \text{with probability } 1 - p \end{cases} \tag{73}$$

where $V \sim N(0, 1)$, $U \sim DE(x; \mu, b)$, and $p \in (0, 1)$ represents the proportion of the Gaussian component in the mixture distribution. Denote the probability density function of X as $f_{G-DE}(x; 0, 1; \mu, b; p)$, then

$$f_{G-DE}(x; 0, 1; \mu, b; p) = \frac{p}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) + (1 - p) \cdot f_{DE}(x; \mu, b) \tag{74}$$

With equation (62) and equation (63) these two properties considered, we can formulate the one-factor G-DE copula model. Suppose that the i-th asset value X_i of the reference entities is given by

$$x_i = a_i M + \sqrt{1 - a_i^2} Z_i \tag{75}$$

where M and Z_i are standardized and independent with random variables with

$$\begin{aligned} M &\sim G - DE(0, 1; 0, \frac{1}{\sqrt{2}}; p) \\ Z_i &\sim G - DE(0, 1; 0, \frac{1}{\sqrt{2}}; p) \end{aligned} \tag{76}$$

The asset return of the i-th instrument, x_i, follows a G-2DE distribution given by:

$$x_i \sim G - 2DE\left(0, 1; \frac{a_i}{\sqrt{2}}, \frac{\sqrt{1 - a_i^2}}{\sqrt{2}}; p\right) \tag{77}$$

with the corresponding pdf:

$$f_{G-2DE}(x; 0, 1; \frac{a_i}{\sqrt{2}}, \frac{\sqrt{1 - a_i^2}}{\sqrt{2}}; p) = \frac{p}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) + (1-p) \cdot f_{2DE}(x; \frac{a_i}{\sqrt{2}}, \frac{\sqrt{1 - a_i^2}}{\sqrt{2}}) \tag{78}$$

Next, we introduce the properties of the G-2DE. From equation (69), we know the moment generating function of the 2DE, and we have calculated its variance, kurtosis, and other central moments. With this information, we can easily compute the central moments of the mixed Gaussian and 2DE distribution. The calculation of the kurtosis of the mixed distribution involves the means and variances of each distribution. However, since both the Gaussian and 2DE distributions have zero mean and unit variance, these terms can be ignored. Consequently, a random variable $X \sim G - 2DE(0, 1; \frac{a}{\sqrt{2}}, \frac{\sqrt{1-a^2}}{\sqrt{2}}; p)$, the central moments of X can be computed as follows:

$$
\begin{aligned}
&E[X] = 0, \quad \text{Var}[X] = 1 \\
&\text{Skew}[X] = 0, \quad \text{Kurt}[X] = 3 \cdot p + (1-p) \cdot 6(a^4 - a^2 + 1)
\end{aligned}
\tag{79}
$$

Unlike the kurtosis of the 2DE in equation (70), which ranges between 4.5 and 6 (since $0 < a < 1$), the kurtosis of the G-2DE is more flexible. By adjusting the proportion p of the Gaussian distribution, the kurtosis can range between 3 and 6, providing more flexibility in its heavy-tailed properties.

Next, let's return to the derivation of the loss distribution and the expected tranche loss. Again, for simplicity, we introduce the notation $F_{G-DE(1)}(x)$ to represent the standard mixed Gaussian and double exponential distribution function $F_{G-DE}\left(x; 0, 1; 0, \frac{1}{\sqrt{2}}; p\right)$, and $F_{G-2DE(a)}(x)$ to represent the distribution function $F_{G-2DE}\left(x; 0, 1; \frac{a}{\sqrt{2}}, \frac{\sqrt{1-a^2}}{\sqrt{2}}; p\right)$.

Following the same procedure as in equation (49), the loss distribution $F_\infty(t, x)$ can be expressed as follows:

$$
F_\infty(t, x) = 1 - F_{G-DE(1)}\left(\frac{C(t) - \sqrt{1-a^2}F^{-1}_{G-DE(1)}(x)}{a}\right) \tag{80}
$$

where $C(t) = F^{-1}_{G-2DE(a)}(Q(t))$, which denotes the default threshold based on G-2DE distribution.

Finally, following the same procedure as in equation (52), the expected tranche loss

$EL_{(K_A,K_D)}(t)$ can be expressed as follows:

$$
\begin{aligned}
&EL_{(K_A,K_D)}(t) \\
&= \frac{1}{K_D - K_A} \left\{ \begin{array}{l} \int_B^C \left((1-R)F_{G-DE(1)}(y) - K_A\right) f_{G-DE(1)}\left(\frac{C(t)-\sqrt{1-a^2}y}{a}\right) \frac{\sqrt{1-a^2}}{a} dy \\ -\int_A^C \left((1-R)F_{G-DE(1)}(y) - K_D\right) f_{G-DE(1)}\left(\frac{C(t)-\sqrt{1-a^2}y}{a}\right) \frac{\sqrt{1-a^2}}{a} dy \end{array} \right\}
\end{aligned} \tag{81}
$$

where $A = F^{-1}_{G-DE(1)}\left(\frac{K_D}{1-R}\right), \quad B = F^{-1}_{G-DE(1)}\left(\frac{K_A}{1-R}\right), \quad C = F^{-1}_{G-DE(1)}(1).$

Up to this point, we have introduced the properties of various copula distributions and constructed their respective expected tranche loss calculation formulas. Combining equation (13) and equation (11), the CDO price can then be determined.

In the next chapter, we will introduce the data sources and calibration methods to evaluate the calibration ability, computational efficiency, and other aspects of each model.

Chapter 3 Data and Calibration

3.1 Synthetic CDO Data

Firstly, we begin by precisely defining what a synthetic CDO entails. According to Hull et al. (2009), if a CDO is constructed using a bond portfolio, it is termed a cash CDO. In practice, traditional CDOs typically invest in conventional debt instruments such as bonds, mortgages, and loans. In contrast, synthetic CDOs generate returns by investing in non-cash derivatives like credit default swaps (CDSs), options, and other contracts. In the context of this study, we specifically investigate a CDO structured through the sale of credit default swaps.

In order to compare the performance of the various copula models mentioned before, we will utilize real market quotes of synthetic CDOs. These copula models will be used to calibrate the market quotes, allowing us to compare their calibration results.

In this research, the financial instruments we use are CDS indices. There are currently two main families of corporate CDS indices: CDX NA IG and iTraxx Europe. CDX indices comprise companies from North America and Emerging Markets, managed by CDS Index Company and marketed by Markit Group Limited. Conversely, iTraxx indices include companies from the global market, administered by the International Index

Company, also owned by Markit. Both CDX and iTraxx indices consist of a reference portfolio of 125 credit default swaps. These portfolios are widely utilized in the market to define standard synthetic CDO tranches.

To validate the robustness and flexibility of various copula models, we will use iTraxx Europe and CDX NA IG data from different time periods to assess the calibration capabilities of each model. The market quote information for this study is sourced from Bloomberg. We will analyze three specific time periods:

1. 2005-2006: This period is extensively discussed in past literature and provides a baseline for comparison.

2. 2014-2015: This period allows us to verify if the model calibration results are consistent with previous findings after the financial crisis.

3. 2019-2020: This period covers the onset of the COVID-19 pandemic, enabling us to test whether the calibration results of various copula models under pandemic conditions align with previous conclusions.

We will perform model calibration using iTraxx Europe with 5-year maturity and CDX NA IG with 5-year maturity for these three periods. By examining these different periods, we aim to validate the robustness and flexibility of the copula models across varying market conditions.

In order to compare the calibration results of the various copula model we have mentioned before, we consider the example of the tranched Dow Jones iTraxx Europe with 5 years maturity. The reference portfolio consists of 125 credit default swap names. The standard tranches have attachment/detachment points at $3\%, 6\%, 9\%, 12\%$ and 22%. The investors in the tranches receive quarterly spread payments on the outstanding notional and compensate for losses when these hit the tranche they are invested in.

3.2 Data Quotation

Although iTraxx Europe and CDX NA IG share similarities, such as both having a reference portfolio consisting of 125 credit default swaps, they have several key differences. First, the standard tranches for iTraxx Europe have attachment/detachment points at 3%, 6%, 9%, 12%, and 22%. In contrast, the standard tranches for CDX NA IG have attachment/detachment points at 3%, 7%, 10%, 15%, and 30%.

Additionally, their pricing methods differ.

For iTraxx Europe in 2005-2006, the equity tranche (0-3%) is priced differently from the other tranches; it is quoted using an upfront rate. As recalled in equation (13), the protection buyer pays an upfront fee once at the inception of the trade and then pays a fixed coupon of 500 bps during the life of the contract.

In 2014-2015, iTraxx Europe had both the equity tranche (0-3%) and the 3-6% tranche quoted using the upfront rate. The remaining tranches continued to be quoted using the breakeven spread.

For the 2019-2020 period, iTraxx Europe's 0-3% and 3-6% tranches were also quoted using the upfront rate. However, unlike previous periods, the remaining tranches were limited to 6-12%, with no attachment/detachment points at 9% and 22%.

CDX NA IG presents a more complex case. In 2005-2006, the CDX NA IG equity tranche (0-3%) was also quoted using the upfront rate, with other tranches quoted using the breakeven spread.

However, for the 2014-2015 period, CDX NA IG' s attachment/detachment points were reduced to 3%, 7%, and 15%. Additionally, all tranches were quoted using the upfront rate, and the fixed coupon paid each period varied. For instance, the 0-3% tranche required a spread payment of 500 bps each period, while the 3-7% and 7-15% tranches required a spread payment of 100 bps each period, significantly differing from iTraxx.

Finally, in the 2019-2020 period, CDX NA IG maintained its attachment/detachment points at 3%, 7%, and 15%, with all tranches quoted using the upfront rate. Unlike previous periods, all tranches during this time required a fixed coupon payment of 100 bps per period, rather than the earlier requirement of 500 bps for the 0-3% tranche.

Due to the complexity of the pricing method for CDX NA IG, this study will primarily focus on the market quotes of iTraxx Europe. CDX NA IG will only be used with data from 2005-2006, which is commonly used in previous literature, for calibration purposes.

3.3 Calibration Method

The calibration method used in this study follows the approach proposed by Kalemanova et al. (2007) and applies the same method across all copula models to ensure fairness in comparing the calibration results. According to Kalemanova et al. (2007), by calibration, we mean finding optimal parameters for all copula models to fit the equity tranche quote and minimize the absolute error across all tranches. In other words, we will adjust the model parameters so that the calculated upfront rate of the equity tranche matches the market quote, and then minimize the calculation errors for the remaining tranches.

Additionally, Bloomberg provides not only the market quotes for each tranche but also the tranche base correlation. In this study, we assume the marginal default distribution of the underlying assets follows a constant default intensity model. We will first estimate the default intensity λ of the one-factor Gaussian copula model with LHP using the 0-3% base correlation and the market quote for the 0-3% tranche provided by Bloomberg. This estimated default intensity will then be used for model calibration. The constant recovery rate R and risk-free interest rate r are assumed to be 40% and 5%, respectively.

Through this calibration method, we can evaluate the accuracy of each model by comparing the expected valuation calculated by the models with the market quotes. The smaller the absolute error, the more accurate the model is in fitting the market quote.

Chapter 4 Simulation Results

In this chapter, we aim to verify the effectiveness of the two-factor copula model. Due to the correlation in a two-factor model being determined by two common factors F_a and F_b, it can capture a more complex correlation matrix. However, because of this, the correlation between underlying assets will no longer follow a single correlation coefficient, thus violating the homogeneous assumption of the LHP.

Therefore, before proceeding to the empirical analysis, we want to examine some simulation results to see if sacrificing the computational efficiency advantages of the LHP, the multi-factor copula model can bring significant improvements. Suppose the underlying assets have a group structure, consisting of many groups with high intra-group correlation and low inter-group correlation. In this chapter, we will evaluate whether the two-factor Gaussian copula model can capture the characteristics of grouped-correlation and make its calibration results significantly superior to the one-factor model.

4.1 Simulation Method

Recall equation (16), where we discussed the concept of using Gaussian copula mapping. By employing the percentile-to-percentile transformation, we can obtain the default threshold. Similarly, in the Monte Carlo simulation method, we can use the copula mapping technique to construct correlated default times. The following outlines the method for simulating the correlated time to default.

The Gaussian copula mapping in equation (16) can be rewritten as $Q(t) = \Phi(u)$. Here, t is the correlated default time we aim to sample. Therefore, by combining equation (15), we can express the correlated default times τ_i as follows:

$$\tau_i = Q_i^{-1}(\Phi(u)) = \frac{-\ln(1-\Phi(u))}{\lambda}, \quad i = 1, \ldots, n \tag{82}$$

The following algorithm generate random variates $(u_1, u_2, \ldots, u_n)$ which are determination of correlated uniform variates on [0,1] from the Gaussian copula with the correlation matrix R. We can construct $\Phi(u)$ as follows:

1. Find the Cholesky decomposition A of the correlation matrix R, such that R $= A \cdot A^T$;
2. Simulate n independent standard normal random variates $Z = (z_1, z_2, \ldots, z_n)^T$
3. Set u = A · **Z**
4. Computing $\Phi(u)$. The vector $\Phi(u)$ is a random variate from the n-dimensional Gaussian copula C_R.

Recall that the i th obligor is deemed to default before $t \in [0, T]$ if $\tau_i \leq t$. Then the loss variable is defined as follows:

$$\Gamma_i(t) = \mathbf{1}\left(\tau_i \leq t\right) \tag{83}$$

where **1** denotes the indicator function.

Similar to equation (4), the portfolio loss process via simulation, which averages the losses of all obligors, can be represented by the following equation:

$$L(t) = \frac{1}{n}\sum_{i=1}^{n}(1-R)\,\Gamma_i(t), \quad t \in [0,T] \tag{84}$$

With the portfolio loss process defined, we can utilize equation (11) and equation (13) to compute the breakeven upfront rate and spread quote. Subsequently, we can simulate market quotes based on the complex correlation matrix with grouped correlations, and validate the calibration capability of the two-factor Gaussian copula model against these quotes.

Next, we introduce the configuration of the correlation matrix R. We will simulate market quotes under two scenarios: division of underlying assets into two groups, and division into six groups.

Firstly, under the LHP assumption, all underlying assets are homogeneous, implying that all underlying assets are correlated the same correlation coefficient. The correlation matrix under LHP assumption can then be expressed as follows:

$$R = \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{bmatrix} \tag{85}$$

where ρ denotes the correlation coefficient.

If the underlying assets can be divided into several groups, the correlation matrix will no longer be as depicted in equation (85). Taking the example of dividing the underlying assets into two groups, we assume that all 125 companies are split into 2 groups, with 60 and 65 companies respectively. Intra-group exhibits higher correlation, while inter-group shows lower correlation. Under these assumptions, the correlation matrix can be written as follows:

$$
R = \begin{bmatrix}
1 & \rho_g & \cdots & \rho_g & \rho & \rho & \cdots & \rho \\
\rho_g & 1 & \cdots & \rho_g & \rho & \rho & \cdots & \rho \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\rho_g & \rho_g & \cdots & 1 & \rho & \rho & \cdots & \rho \\
\rho & \rho & \cdots & \rho & 1 & \rho_g & \cdots & \rho_g \\
\rho & \rho & \cdots & \rho & \rho_g & 1 & \cdots & \rho_g \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\rho & \rho & \cdots & \rho & \rho_g & \rho_g & \cdots & 1
\end{bmatrix} \tag{86}
$$

where ρ_g represents the intra-group correlation and ρ denotes the inter-group correlation coefficient.

Finally, assuming underlying assets are divided into six groups. Similar to the assumption of two groups, we assume higher correlation within each group and lower correlation between groups. All 125 companies are split into 6 groups, comprising 20, 20, 20, 20, 20, and 25 companies respectively. Under these assumptions, the correlation matrix can be represented as follows:

$$
R = \begin{bmatrix}
1 & \dots & \rho_g & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho \\
\vdots & \ddots & \vdots & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho \\
\rho_g & \dots & 1 & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho \\
\rho & \rho & \rho & 1 & \dots & \rho_g & \rho & \rho & \rho & \rho & \rho & \rho & \rho \\
\rho & \rho & \rho & \vdots & \ddots & \vdots & \rho & \rho & \rho & \rho & \rho & \rho & \rho \\
\rho & \rho & \rho & \rho_g & \dots & 1 & \rho & \rho & \rho & \rho & \rho & \rho & \rho \\
\vdots & & & & & & \ddots & & & & & & \vdots \\
\rho & \rho & \rho & \rho & \rho & \rho & \rho & 1 & \dots & \rho_g & \rho & \rho & \rho \\
\rho & \rho & \rho & \rho & \rho & \rho & \rho & \vdots & \ddots & \vdots & \rho & \rho & \rho \\
\rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho_g & \dots & 1 & \rho & \rho & \rho \\
\rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & 1 & \dots & \rho_g \\
\rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \vdots & \ddots & \vdots \\
\rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho_g & \dots & 1
\end{bmatrix} \tag{87}
$$

where ρ_g represents the intra-group correlation and ρ denotes the inter-group correlation coefficient.

4.2 Simulation Results

Firtly, we examine the scenario where all underlying assets can be divided into two groups, as represented by the correlation matrix in equation (86). In this scenario, we assume high intra-group correlation $\rho_g = 0.7$ and lower inter-group correlation $\rho = 0.3$. We then simulate market quotes using a Gaussian copula with this correlation matrix. In addition, when using the two-factor Gaussian models as represented by equation (35), it is necessary to specify the total number of companies in Group 1 and Group 2, denoted

	Simulated Quotes	One-factor Gaussian	Two-factor Gaussian	Simulated Quotes	One-factor Gaussian	Two-factor Gaussian	Simulated Quotes	One-factor Gaussian	Two-factor Gaussian
λ	0.01			0.15			0.5		
0-3%	1241.88	1241.89	1241.88	8604.58	8614.94	8604.54	9561.26	9587.9	9561.17
3-6%	407.75	387.45	408.16	5469.23	5283.77	5470.51	17496.94	17777.47	17498.24
6-9%	267.03	246.61	267.06	4314.44	4100.23	4316.12	14100.28	13944.24	14097.41
9-12%	188.63	179.34	188.19	3604.12	3360.81	3603.24	12001.69	11639.08	12003.43
12-22%	98.59	95.48	98.22	2624.38	2447.95	2622.54	9123.42	8732.91	9123.4
absolute error		53.13	1.25		819.41	5.69		1189.69	5.92
ρ	0.3	0.57	0.31	0.3	0.55	0.3	0.3	0.53	0.3
ρ_g	0.7	-	0.7	0.7	-	0.7	0.7	-	0.7
calibration time		0.38	392.22		0.52	329.67		0.38	315.55

Table 4.1: Calibration results using quotes simulated by two group Gaussian copula with different default intensity λ

as n_1 and n_2. We assume that n_1 and n_2 are known and are 60 and 65, respectively. With this setup, we will compare the calibration performance of the one-factor Gaussian and two-factor Gaussian models. The results are shown in Table 4.1.

We compare the cases with different default intensities λ. First, considering the Simulated Quotes, the 0-3% tranche is the equity tranche and is thus quoted as a breakeven upfront rate. The other tranches are quoted as breakeven spreads. As λ increases, the quotes also increase. This makes sense because as the default probability increases, the protection buyer needs to pay a higher upfront fee/spread to protect defaultable assets.

Next, we look at the calibration results for the one-factor and two-factor models. Based on our previously mentioned calibration method, we first fit the 0-3% equity tranche upfront rate and then minimize the absolute errors of the other tranches. As a result, we see that the calibrated equity tranche upfront fees for both models are close to the simulated quotes, and the remaining tranches' errors are minimized as much as possible. Therefore, we can evaluate the model calibration quality by summing the absolute errors across all tranches.

	Simulated Quotes	One-factor Gaussian	Two-factor Gaussian	Simulated Quotes	One-factor Gaussian	Two-factor Gaussian	Simulated Quotes	One-factor Gaussian	Two-factor Gaussian
λ	0.01			0.15			0.5		
0-3%	2224.08	2224.1	2224.1	9122.58	9148.76	9126.99	9675.02	9693.27	9682.92
3-6%	476.98	440.24	464.8	7371	7124.4	7278.94	23749.38	24008.73	23776.22
6-9%	262.48	245.2	262.02	5437.55	5174.13	5409.09	18010.63	17769.73	17925.24
9-12%	152.23	151.53	161.94	4293.57	4090.27	4328.8	14532	14265.64	14527.18
12-22%	57.58	65.72	63.52	2849.56	2750.53	2927.79	10158.99	10000.05	10251.88
absolute error		62.85	28.3		812.34	233.99		925.56	209.95
ρ	0.3	0.42	0.04	0.3	0.4	0.12	0.3	0.38	0.19
ρ_g	0.7	-	0.59	0.7	-	0.58	0.7	-	0.53
calibration time		0.43	559.52		0.42	321.39		0.36	299.73

Table 4.2: Calibration results using quotes simulated by six group Gaussian copula with different default intensity λ

We can see that the absolute errors for the two-factor Gaussian model are significantly lower than those for the one-factor Gaussian model. Additionally, the calibrated ρ and ρ_g values are close to the correlation matrix used in our simulated quotes. The calibration performance of the two-factor Gaussian model significantly outperforms that of the one-factor Gaussian model.

Next, we examine the scenario where all underlying assets can be divided into six groups, as represented by the correlation matrix in equation (87). Similarly, in this scenario, we assume high intra-group correlation $\rho_g = 0.7$ and lower inter-group correlation $\rho = 0.3$. We then simulate market quotes using a Gaussian copula with this correlation matrix. In addition, when using the two-factor Gaussian models, we also assume n_1 and n_2 are known and are 60 and 65, respectively. The calibration results are shown in Table 4.2.

In the scenario where the underlying assets are divided into six groups, the absolute error of the two-factor Gaussian model starts to increase, and ρ, ρ_g no longer approximate the correlation matrix used in the simulated quotes. However, its absolute error remains significantly lower than that of the one-factor Gaussian model. This indicates that the calibration performance of the two-factor Gaussian model still significantly outperforms that of the one-factor Gaussian model.

Thus, we can draw a preliminary conclusion that whether the underlying assets are divided into two groups or six groups, the two-factor Gaussian model performs better in calibration compared to the one-factor Gaussian model, and is more capable of capturing the grouped-correlation.

However, there is also an issue. The calibration time for the two-factor Gaussian model is significantly longer than that for the one-factor Gaussian model. This is because the one-factor Gaussian model, under the LHP approximation, has very efficient closed-form mathematical solutions. In contrast, the two-factor Gaussian model, without the LHP assumption, requires double integration over two common factors, resulting in lower computational efficiency.

Next, we examine the scenario where all underlying assets can be divided into two groups. The only difference is that in this case, simulated market quotes are generated using an NIG copula instead of Gaussian. We aim to evaluate whether the two-factor Gaussian model still outperforms the one-factor Gaussian model when market quotes are simulated using different copula distributions. For convenience, we assume parameters $\alpha = 1$ and $\beta = 0$ for the NIG distribution. The calibration results are presented in Table 4.3.

We observe that the absolute error of the two-factor Gaussian model is no longer significantly lower than that of the one-factor Gaussian model. Furthermore, the calibrated parameters ρ and ρ_g from the two-factor Gaussian model are identical to ρ from the one-factor Gaussian model, indicating that the addition of an extra factor has no impact on the

	Simulated Quotes	One-factor Gaussian	Two-factor Gaussian	Simulated Quotes	One-factor Gaussian	Two-factor Gaussian	Simulated Quotes	One-factor Gaussian	Two-factor Gaussian
λ	0.005			0.15			0.5		
0-3%	-274.16	-274.17	-274.17	8810.26	8809.81	8808.83	9609.16	9584.86	9584.81
3-6%	125.63	185.32	183.81	5740.73	5792.69	5808.91	17681.07	17654.62	17658.74
6-9%	81.14	113.43	116.13	4377.14	4400.33	4410.7	13699.13	13868.71	13847.61
9-12%	61.12	85.08	80.22	3635.28	3576.77	3595.85	11550.54	11583.58	11599.7
12-22%	38.95	43.24	43.62	2686.48	2534.69	2545.1	8850.23	8705.66	8706.67
absolute error		120.24	116.95		285.46	282.56		373.64	363.53
ρ	0.3	0.63	0.63	0.3	0.5	0.47	0.3	0.53	0.53
ρ_g	0.7	-	0.63	0.7	-	0.54	0.7	-	0.53
calibration time		0.24	539.07		0.28	329		0.37	338.73

Table 4.3: Calibration results using quotes simulated by two group NIG copula with different default intensity λ

calibration results.

The result from Table 4.3 contrasts sharply with the results from Table 4.1 and Table 4.2, where the two-factor Gaussian model exhibited clear superiority. In this case, when the simulated quotes use an NIG copula distribution and the calibration model uses a Gaussian copula distribution, the two-factor Gaussian model does not demonstrate an advantage. This may suggest that when the copula distribution used for calibration is incorrect, the improvement in model calibration accuracy due to adding an additional factor and its associated changes in correlation structure is not significant.

Finally, we again consider the scenario where underlying assets are divided into six groups and simulate market quotes using an NIG copula. We assume the parameters of the NIG distribution as $\alpha = 1$ and $\beta = 0$. The calibration results are presented in Table 4.4.

We observe that the absolute error of the two-factor Gaussian model remains similar to that of the one-factor Gaussian model, with no significant reduction. This trend is particularly evident at default intensities $\lambda = 0.005$ and $\lambda = 0.5$, where the two-factor Gaussian model shows no clear advantage. Only at $\lambda = 0.15$, the absolute error of the

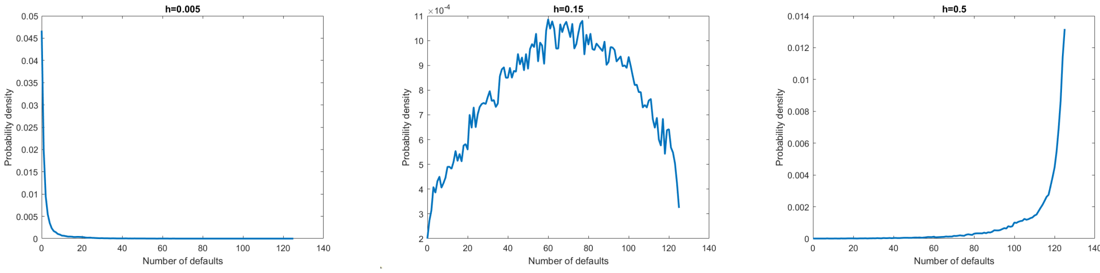
	Simulated Quotes	One-factor Gaussian	Two-factor Gaussian	Simulated Quotes	One-factor Gaussian	Two-factor Gaussian	Simulated Quotes	One-factor Gaussian	Two-factor Gaussian
λ		0.005			0.15			0.5	
0-3%	261.42	261.43	261.43	9209.88	9243.11	9214.91	9702.76	9703.38	9700.87
3-6%	161.81	197.89	197.61	7911.45	7694.9	7833.22	24825.89	24938.47	24886.66
6-9%	88.84	108.32	108.51	5791.65	5498.11	5756.04	18500.41	18328.82	18368.35
9-12%	45.8	67.03	66.89	4541.98	4295.81	4568.64	14730.38	14629.48	14686.41
12-22%	14.48	29.21	29.39	2958.25	2836.9	3043.52	10012.84	10170.88	10220.11
absolute error		91.53	91.47		877.6	225.78		543.1	444.06
ρ	0.3	0.5	0.5	0.3	0.36	0.05	0.3	0.36	0.29
ρ_g	0.7	-	0.5	0.7	-	0.57	0.7	-	0.43
calibration time		0.28	448.06		0.55	299.71		0.4	311.63

Table 4.4: Calibration results using quotes simulated by six group NIG copula with different default intensity λ

two-factor Gaussian model is noticeably lower than that of the one-factor Gaussian model.

The calibration of the two-factor model performs better when default intensity $\lambda = 0.15$. Possible explanations include the following: Figure 4.1 illustrates the simulated distribution of number of defaults from the NIG copula for different default intensities λ. It is observed that at $\lambda = 0.005$ and $\lambda = 0.5$, there is a tendency towards either all defaults or no defaults. Due to high intra-group correlation, the default/non-default status of one company tends to influence the entire group to default or not. This diminishes the significance of grouped-correlation. In other words, This suggests that grouped-correlation becomes significant only when the default intensity is neither too high nor too low. Therefore, at $\lambda = 0.15$, the two-factor Gaussian model is able to capture grouped-correlation, resulting in better calibration compared to the one-factor Gaussian model.

Finally, let's provide a brief summary. First, when the copula distribution used to depict the correlation is incorrect, the improvement in model accuracy from adding factors is not significant. As shown in Table 4.1 and Table 4.2, the two-factor Gaussian copula model effectively fits quotes simulated by a Gaussian copula with grouped correlation.



$\lambda = 0.005$ $\lambda = 0.15$ $\lambda = 0.5$

Figure 4.1: Distribution of number of defaults from NIG copula simulation

However, as shown in Table 4.3 and Table 4.4, it fails to accurately capture the effect of grouped correlation simulated by an NIG copula.

Second, the effect of grouped correlation varies with default intensity: as seen in the Figure 4.1, the impact of grouped correlation is noticeable only when default intensity λ is neither extremely high nor low. When default intensity is extreme, high intra-group correlation causes simultaneous defaults or non-defaults within the group, skewing the probability distribution towards either 0 or all 125 defaults. This explains why Table 4.4 shows that the two-factor Gaussian outperforms the one-factor Gaussian significantly only at $\lambda = 0.15$.

Therefore, we can infer that the effect of use of different copula distribution is more significant than that of adding factors. Furthermore, empirical data shows a relatively low default intensity, further diminishing the effect of grouped correlation. Based on these points, we believe that focusing on the analysis of different copula distributions on empirical data will provide more significant insights.

Chapter 5 Empirical Results

In the previous chapters, we have observed that even when the underlying assets exhibit complex grouped-correlation, increasing the number of factors does not improve the calibration of the model if the copula distribution used to describe the correlation is incorrect. Additionally, the empirical data shows very low default intensity, which makes the grouped-correlation characteristics less apparent. Therefore, in the following sections, when comparing the calibration results of the empirical models, we will focus on examining how using different distributions impacts the calibration effectiveness of the models.

5.1 Calibration Results

We first examine the 5-year maturity CDX.NA.IG, with results presented in Table 5.1. Market quotes of CDX.NA.IG tranches as of September 7, 2005, are listed in the first column. The 0-3% tranche, referred to as the equity tranche, is quoted using an upfront rate, whereas the other tranches are quoted using spreads.

First, we look at the 1FG and 2FG columns, representing the one-factor and two-factor Gaussian copula models, respectively. It can be observed that the calibration performance of the 2FG model is similar to that of the 1FG model. This reinforce our conclusion from the previous chapter that adding factors does not improve model calibration

	Market	1FG	2FG	Double t	NIG	G-NIG	DE	G-DE
$0-3\%$	870.00	870.03	870.03	869.90	870.00	870.00	870.00	870.01
$3-7\%$	132.00	174.53	174.57	90.11	96.98	126.41	99.57	126.77
$7-10\%$	36.00	69.40	69.38	48.91	50.76	47.97	50.57	57.32
$10-15\%$	21.50	30.58	30.55	33.60	33.46	21.33	32.48	31.74
$15-30\%$	10.30	6.03	6.03	19.18	17.53	6.02	17.01	12.52
absolute error		89.27	89.26	75.78	68.97	22.02	64.69	39.01
ρ		0.36	0.35	0.43	0.41	0.36	0.43	0.41
ρ_g		-	0.36	-	-	-	-	-
α		-	-	-	0.90	0.17	-	-
p		-	-	-	-	0.82	-	0.47
calibration time		0.24	2158.73	316.68	371.19	15957.37	5.31	17.52

Table 5.1: Pricing CDX.NA.IG(2005/09/07) tranches with different copula model

performance when the copula distribution used to describe correlation is incorrect.

Next, we examine the Double t, NIG, and DE columns, where DE refers to the double-exponential copula model discussed in section 2.7.1. It can be seen that the calibration performances of these heavy-tailed copula models are similar. However, the DE model significantly outperforms the others in terms of calibration time. The reason is that the Double t copula model lacks the convolution property, and its distribution can only be approximated without an analytical solution, leading to a long time required to compute market quotes. The NIG copula model has the convolution property, reducing the computation time significantly. However, despite having an analytical solution, the cumulative distribution function (CDF) of the NIG distribution still requires integration for calculation. Additionally, the NIG model has numerous parameters, such as α and β, which provide flexibility in producing skewness and kurtosis but also prolong the calibration process to find the optimal parameters. In this study, we set $\beta = 0$ to simplify the calibration process. However, even with $\beta = 0$ to reduce calibration time, it is still longer than that of the Double t model. In contrast, the DE copula model, with its convolution property and heavy-tailed feature, is less complex than the NIG distribution, and its CDF, as shown in equation (67), does not require integration. This gives it substantial advantages in terms of model complexity, computational efficiency, calibration time, and calibration performance.

Subsequently, we consider the G-NIG and G-DE columns. By mixing a portion of the Gaussian distribution, these copula models address the issue of underestimation in the second tranche (3-7%) common to traditional heavy-tailed copula models. In Figure 5.2, we can see that the relevant loss distribution functions for all calibrated copula models are drawn. This figure demonstrates that G-NIG and G-DE can achieve what double t, NIG, and DE cannot by adjusting the proportion of the Gaussian distribution. Both G-NIG and G-DE show significant improvements in calibration performance compared to the Double t, NIG, and DE models.

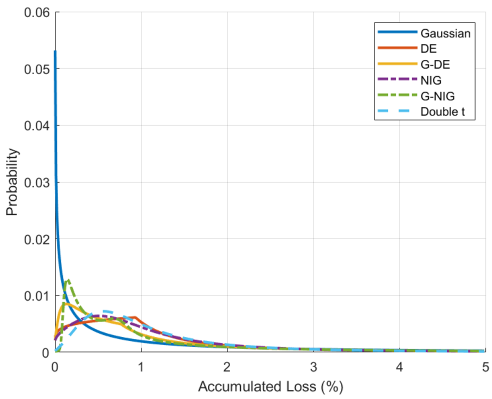


Figure 5.1: CDX(2005/09/07)
Density function

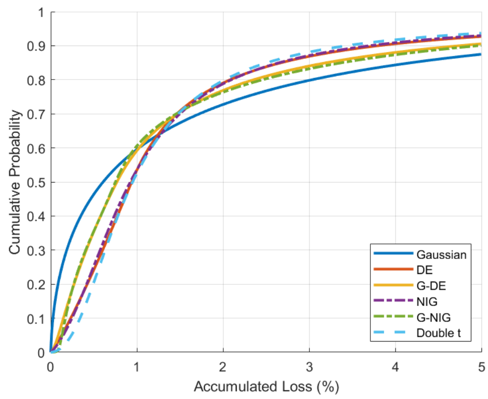


Figure 5.2: CDX(2005/09/07)
Cumulative distribution function

Finally, let's compare G-NIG and G-DE. The introduction of the α parameter in the G-NIG model provides additional flexibility in shape, leading to better calibration results and closer fitting to market quotes. However, similar to the NIG model, the numerous parameters in the G-NIG model result in low calibration efficiency. In this table, the

	Market	1FG	2FG	Double t	NIG	G-NIG	DE	G-DE
$0-3\%$	1180.00	1180.06	1180.06	1180.50	1180.00	1180.00	1180.01	1180.01
$3-6\%$	81.00	147.92	150.53	66.00	77.84	80.55	82.87	82.87
$6-9\%$	25.80	45.94	45.98	31.79	36.92	27.27	37.32	37.32
$9-12\%$	14.30	16.70	15.98	20.96	23.28	13.72	22.54	22.54
$12-22\%$	8.20	3.15	2.65	12.36	12.46	6.89	11.63	11.63
absolute error		94.52	96.94	31.80	27.52	3.81	25.06	25.06
ρ		0.23	0.10	0.29	0.26	0.25	0.32	0.32
ρ_g		-	0.33	-	-	-	-	-
α		-	-	-	0.77	0.24	-	-
p		-	-	-	-	0.67	-	0
calibration time		0.23	2215.27	284.93	491.44	5329.41	3.84	22.25

Table 5.2: Pricing ITRX(2005/09/05) tranches with different copula model

calibration time for the G-NIG model is up to 900 times longer than that of the G-DE model.

This highlights the major advantage of our proposed G-DE model. The G-DE copula model possesses the convolution property, heavy-tailed characteristics, simplicity in model structure, and a very straightforward cumulative distribution function (CDF) that requires no integration. Furthermore, the use of a mixture of Gaussian distributions significantly enhances its calibration performance. In conclusion, we consider G-DE to be an ideal model for evaluating synthetic CDOs.

Next, we examine the 5-year iTraxx Europe for the same period. The results are presented in Table 5.2.

Similar to CDX.NA.IG, the 0-3% tranche represents the equity tranche priced using upfront rates, while other tranches are priced using spreads. The conclusions are largely consistent with our previous findings: the calibration performance of the 2FG model is similar to 1FG, while Double t, NIG, and DE outperform Gaussian models. Double t's lack of convolution property and the inherent complexity of the NIG model result in significantly longer calibration times compared to DE. Unlike the previous conclusion, here the G-DE model produces results identical to the DE model. The use of mixture of Gaussian distribution show no impact on calibrating the market quote. In contrast, G-NIG exhibits vastly superior calibration performance compared to all models, almost perfectly fitting the market quote. In fact, Figure 5.4 shows that, except for Gaussian, the CDFs of other models are quite similar. However, the complexity of the G-NIG model allows its CDF to better match the loss distribution of the real-world market, as evident in Figure 5.3, where the density function of G-NIG can depict more complex shapes to fit the market quote more accurately.

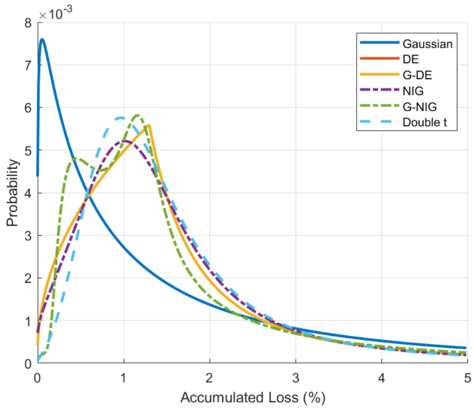


Figure 5.3: ITRX(2005/09/05)
Density function

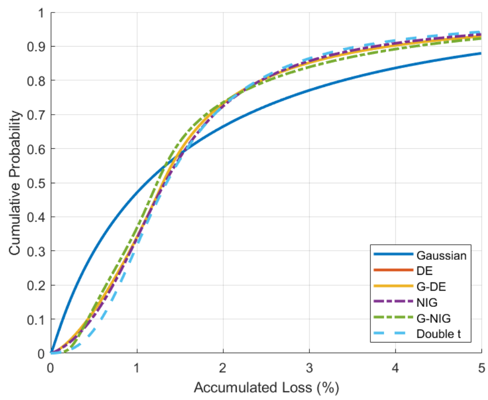


Figure 5.4: ITRX(2005/09/05)
Cumulative distribution function

Nevertheless, G-NIG faces the ongoing issue of significantly longer calibration times, approximately 200 times more than G-DE. This reinforces our emphasis on the advantages of the G-DE copula model. Although G-DE copula model sacrifices some precision in calibration results, it achieves unparalleled calibration efficiency compared to all the other copula models.

Next, let's consider the 5-year maturity iTraxx Europe as of 16-Apr-2006. The results

	Market	1FG	2FG	Double t	NIG	G-NIG	DE	G-DE
$0-3\%$	2400.00	2400.05	2400.05	2400.20	2400.00	2400.00	2400.00	2400.00
$3-6\%$	63.00	99.56	99.53	56.08	62.97	62.89	64.16	64.16
$6-9\%$	18.00	9.34	9.41	19.46	19.34	18.29	18.78	18.78
$9-12\%$	9.00	0.93	0.92	11.07	9.21	8.66	8.69	8.69
$12-22\%$	4.00	0.03	0.03	5.63	3.35	3.31	3.25	3.25
absolute error		57.27	57.32	12.08	2.23	1.43	3.00	3.00
ρ		0.08	0.08	0.15	0.10	0.11	0.17	0.17
ρ_g		-	0.08	-	-	-	-	-
α		-	-	-	0.76	0.59	-	-
p		-	-	-	-	0.13	-	0
calibration time		0.41	2037.4	258.04	307.73	1174.91	10.58	25.63

Table 5.3: Pricing ITRX(2006/04/13) tranches with different copula model

are presented in Table 5.3. The findings are generally consistent with previous results: the performance of the 2FG model compared to 1FG has not improved, and copula models with heavy-tail characteristics outperform Gaussian models. However, in this instance, except for Double t, all other copula models exhibit good calibration. From Figure 5.6, it can be observed that Double t deviates slightly from other models in Accumulated loss 0-2%, resulting in less effective calibration compared to other copula models. Additionally, DE itself nearly fits the market quote, allowing G-DE to accurately fit the market quote without needing to adjust the proportion of Gaussian distribution. Finally, DE and G-DE maintain excellent calibration efficiency, with calibration times significantly lower than other copula models with heavy-tail characteristics.

So far, the previous results demonstrate the computational and calibration efficiency advantages of G-DE. To validate the robustness of fitness and flexibility of our G-DE copula model, we conducted further calibration using market quotes of 5-year iTraxx Europe from various periods. First, we analyze the results for 07-Aug-2014 iTraxx Europe, pre-

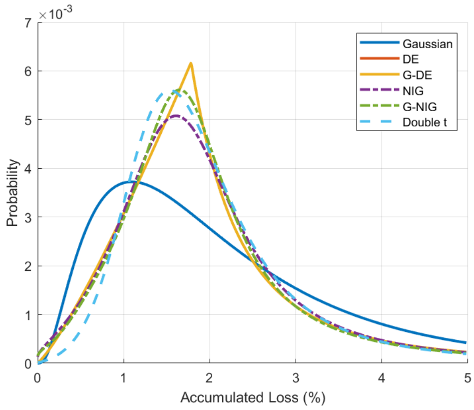


Figure 5.5: ITRX(2006/04/13)
Density function

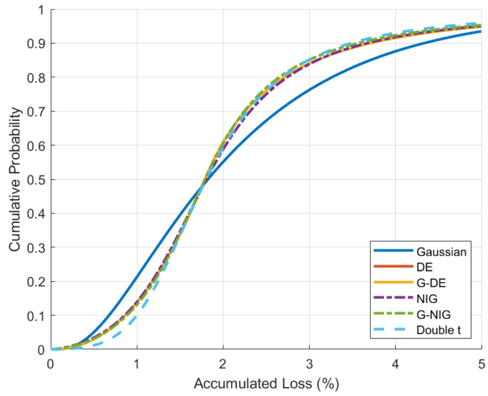


Figure 5.6: ITRX(2006/04/13)
Cumulative distribution function

	Market	1FG	2FG	Double t	NIG	G-NIG	DE	G-DE
$0-3\%$	987.50	986.76	988.09	988.63	987.51	987.51	987.63	987.51
$3-6\%$	-1087.50	-728.23	-722.15	-1311.36	-1092.81	-1091.00	-1275.34	-1092.27
$6-9\%$	128.00	200.36	197.63	124.32	149.77	140.52	125.71	148.42
$9-12\%$	77.00	129.58	133.13	95.70	109.24	92.07	94.60	106.07
$12-22\%$	45.50	67.76	67.66	66.96	68.62	46.41	64.70	64.53
absolute error		506.48	513.29	267.70	82.47	32.01	226.92	73.31
ρ		0.54	0.52	0.62	0.61	0.53	0.60	0.59
ρ_g		-	0.56	-	-	-	-	-
α		-	-	-	1.63	0.10	-	-
p		-	-	-	-	0.73	-	0.39
calibration time		0.45	2324.20	304.79	87.45	8570.47	4.05	27.34

Table 5.4: Pricing ITRX(2014/08/07) tranches with different copula model

sented in Table 5.4. The market quotes here, apart from the equity tranche (0-3%) priced with upfront rate, also include the 3-6% tranche similarly quoted with upfront rate. Thus, negative market quotes indicate that at the inception of the contract, the protection seller pays the protection buyer an upfront rate, and subsequently, the protection buyer pays a 500 basis point spread per period.

Similar to previous findings, increasing the factors in copula models did not enhance calibration performance, while copula models with heavy-tail properties outperformed the Gaussian model. The calibration results for Double t and DE were similar, but they were inferior to those of NIG, G-NIG, and G-DE. This is primarily attributed to the 3-6% tranche. As previously concluded in Table 5.1, traditional heavy-tailed copula models tend to underestimate significantly in the second tranche. Here, both Double t and DE also undervalue the market quotes in the second tranche(3-6%). Therefore, the use of a mixture of Gaussians leads to a significant improvement in G-DE compared to DE.

Another noteworthy point is that in this calibration, NIG also performs well. Figure 5.8 shows that the cumulative loss distributions of NIG and G-DE are very similar. The calibration performance of NIG here is favorable due to parameter α, which allows it to flexibly capture the shape of the distribution and thus avoid underestimating the second tranche quotes excessively. Finally, G-NIG remains the model that best fits the market quote, with only 32 basis points absolute error. In Figure 5.7, it can also be observed that G-NIG deviates slightly from NIG and G-DE, resulting in a better fit to real-world market loss distribution. However, it is emphasized that the main drawback of NIG and G-NIG lies in their model complexity, leading to considerable calibration time. In this regard, G-DE once again proves its computational efficiency while maintaining good calibration performance.

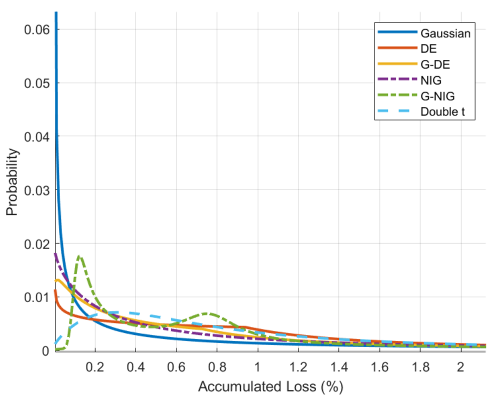


Figure 5.7: ITRX(2014/08/07)
Density function

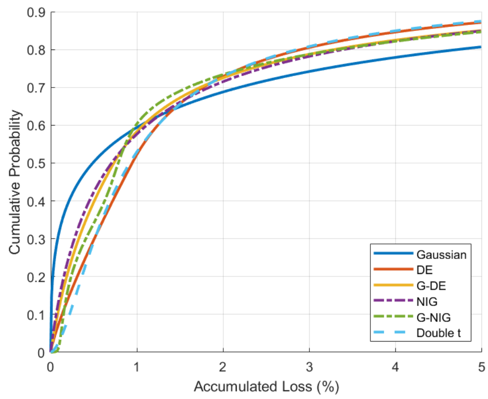


Figure 5.8: ITRX(2014/08/07)
Cumulative distribution function

	Market	1FG	2FG	Double t	NIG	G-NIG	DE	G-DE
0 − 3%	637.50	637.45	637.44	637.67	637.54	637.54	637.55	637.54
3 − 6%	-1165.00	-1065.08	-1066.19	-1575.76	-1175.24	-1181.16	-1521.37	-1176.86
6 − 9%	54.50	135.42	136.62	83.18	126.11	121.77	86.74	125.11
9 − 12%	30.25	87.28	85.61	62.80	82.40	76.18	63.20	80.72
12 − 22%	24.50	38.30	38.47	43.01	41.09	34.47	41.42	39.28
absolute error		251.68	250.27	490.50	150.61	139.34	438.49	147.73
ρ		0.49	0.48	0.56	0.52	0.50	0.55	0.52
ρ_g		-	0.50	-	-	-	-	-
α		-	-	-	3.46	0.10	-	-
p		-	-	-	-	0.93	-	0.85
calibration time		0.28	1045.62	223.32	427.43	11873.32	4.96	13.35

Table 5.5: Pricing ITRX(2015/08/18) tranches with different copula model

Next, we consider the iTraxx Europe as of 18-Aug-2015. The results are presented in Table 5.5. Comparing 2FG to 1FG, there is still no improvement noted. Additionally, similar to previous results, both Double t and DE fail to outperform the Gaussian copula, and even perform significantly worse than Gaussian. In contrast, NIG, G-NIG, and G-DE exhibit superior performance over Gaussian, Double t, and DE. Figure 5.10 again illustrates that the cumulative loss distributions of NIG, G-NIG, and G-DE exhibit striking similarity.

However, the calibration results of G-DE still outperform NIG and are only slightly inferior to G-NIG. Furthermore, its calibration efficiency remains significantly better than both models. This once again demonstrates that G-DE possesses robust calibration capabilities while also offering substantial computational efficiency advantages.

Subsequently, we examine the iTraxx Europe as of 01-Apr-2019. The results are presented in Table 5.6. Here, there are only three tranches remaining, with the first two tranches priced using upfront rates. Similar to previous findings, comparing 2FG to 1FG

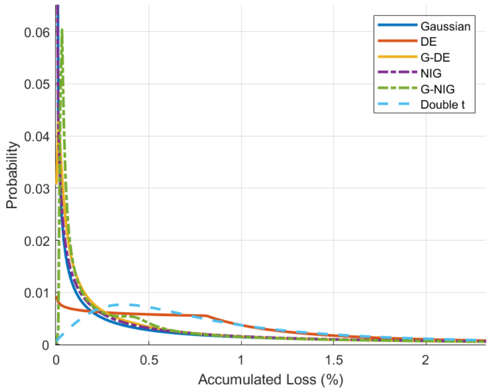


Figure 5.9: ITRX(2015/08/18)
Density function

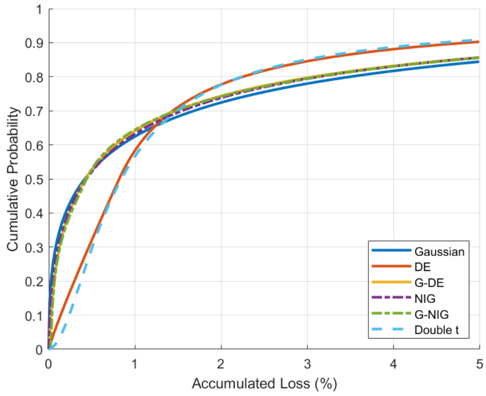


Figure 5.10: ITRX(2015/08/18)
Cumulative distribution function

	Market	1FG	2FG	Double t	NIG	G-NIG	DE	G-DE
$0-3\%$	3886.20	3885.25	3885.32	3886.39	3886.33	3886.33	3888.25	3886.33
$3-6\%$	865.40	1115.92	1115.56	157.11	828.03	827.74	37.27	828.90
$6-12\%$	110.32	402.57	403.00	279.93	366.88	366.26	263.43	368.01
absolute error		542.78	542.84	877.91	293.93	293.61	981.25	294.20
ρ		0.42	0.40	0.55	0.47	0.47	0.56	0.46
ρ_g		-	0.43	-	-	-	-	-
α		-	-	-	2.51	1.08	-	-
p		-	-	-	-	0.72	-	0.81
calibraion time		0.36	2467.65	175.32	117.39	863.19	3.01	14.79

Table 5.6: Pricing ITRX(2019/04/01) tranches with different copula model

shows no improvement, and both Double t and DE fail to outperform the Gaussian copula. Conversely, NIG, G-NIG, and G-DE outperform Gaussian. we can observe that the valuation of Double t and DE in the second tranche reveals significant underestimation, leading to poor calibration results. In contrast, the estimation provided by NIG, G-NIG, and G-DE closely match the market quotes in the second tranche. Figure 5.12 also illustrates that the distribution functions of these three models are very similar and effectively fit the market quotes. However, G-DE continues to exhibit significantly lower calibration times compared to NIG and G-NIG models, reaffirming its advantageous properties.

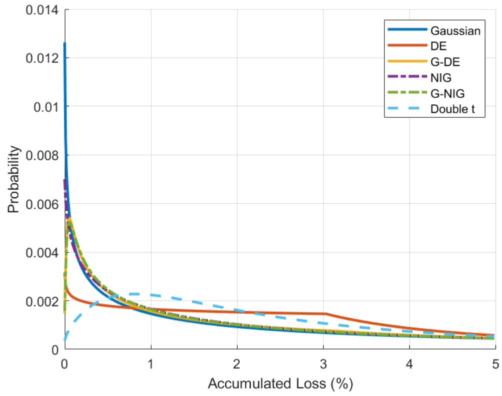


Figure 5.11: ITRX(2019/04/01)
Density function

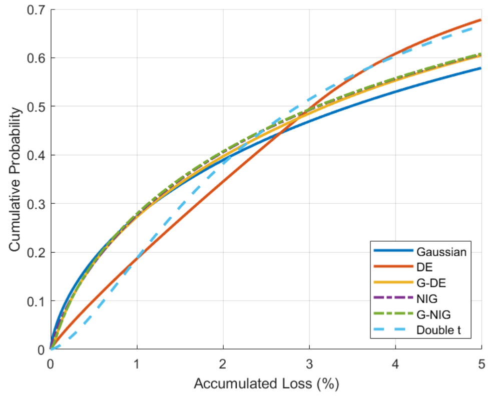


Figure 5.12: ITRX(2019/04/01)
Cumulative distribution function

	Market	1FG	2FG	Double t	NIG	G-NIG	DE	G-DE
$0-3\%$	4235.00	4229.27	4229.53	4234.63	4235.18	4235.20	4235.58	4235.20
$3-6\%$	1270.50	2150.82	2155.59	1286.89	1207.16	1206.71	1118.90	1206.54
$6-12\%$	194.56	728.58	724.88	543.45	523.18	508.93	497.89	515.14
absolute error		1414.35	1415.41	365.28	392.14	378.17	454.92	384.53
ρ		0.59	0.59	0.72	0.72	0.72	0.74	0.72
ρ_g		-	0.59	-	-	-	-	-
α		-	-	-	1.23	0.43	-	-
p		-	-	-	-	0.50	-	0.10
calibration time		0.30	2356.72	183.77	84.25	2846.45	5.02	18.88

Table 5.7: Pricing ITRX(2020/04/01) tranches with different copula model

Finally, we examine the iTraxx Europe as of 01-Apr-2020. The results are presented in Table 5.7. The calibration results for 2FG remain similar to those of 1FG, and other heavy-tailed copula models outperform Gaussian significantly.

Additionally, the calibration results for Double t, NIG, G-NIG, and G-DE are all similar, with only DE slightly falling behind the aforementioned models. However, there is no significant difference among them. Figure 5.14 also shows that except for Gaussian, the cumulative loss distribution functions of all models are quite similar. G-DE continues to demonstrate good computational efficiency and solid calibration performance.

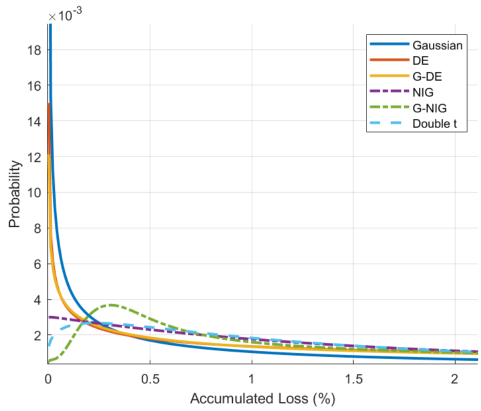


Figure 5.13: ITRX(2020/04/01)
Density function

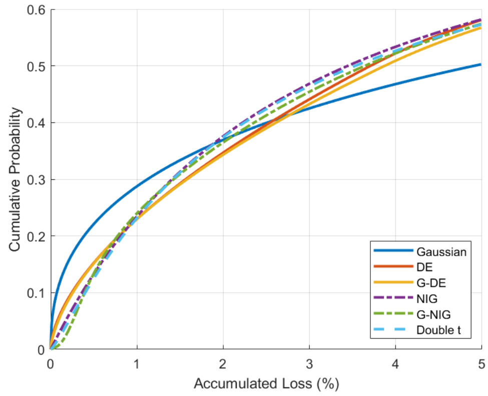


Figure 5.14: ITRX(2020/04/01)
Cumulative distribution function

5.2 Calibration Summary

Finally, based on these calibration results, we will summarize the characteristics of each model in Table 5.8 and assign scores based on empirical results. Firstly, the one-factor Gaussian, serving as the market standard pricing model, is characterized by its simplicity, which results in very high computational efficiency, making it the most efficient model in calibration among all. However, its lack of heavy-tailed characteristics results in us assigning it five-stars for calibration efficiency and one-star for calibration capability.

Model	Characteristic		Performance	
	Convolution Property	Mixture of Gaussian	Calibration Efficiency	Accuracy
1FG	✓		★★★★★	★
2FG	✓		★★	★
Double t			★★★	★★
NIG	✓		★★★	★★★★
G-NIG	✓	✓	★	★★★★★
DE	✓		★★★★★	★★
G-DE	✓	✓	★★★★	★★★★

Table 5.8: Comparison of various copula models

As for the two-factor Gaussian, consistent with the findings in Chapter 4, empirical data does not show significant evidence of grouped correlation due to lower default intensities. Additionally, the choice of copula distribution has a more pronounced effect than adding factors. The calibration results also show that 2FG does not demonstrate superiority over 1FG in any of the empirical results. Therefore, we also assign it a one-star rating for its calibration capability. Additionally, 2FG cannot use LHP approximation and can only use double integration to obtain the loss distribution. Thus, we assign it a calibration efficiency rating of two-stars.

Moreover, traditional heavy-tailed copula models do not exhibit robust performance in calibration; Double t and DE even underperform Gaussian in certain empirical scenarios. Therefore, we assign these two models two-stars for their calibration capability. In contrast, the flexibility provided by the parameter α in NIG prevents excessive underestimation in estimating the second tranche. The use of a mixture of Gaussian distributions in G-DE also enhances its ability to fit the market quote in the second tranche. Therefore, we assign these two models four-stars for their calibration capability.

Overall, G-NIG demonstrates the best overall calibration capability. The parameter α and the use of a mixture of Gaussian distributions provide flexibility to the model. Therefore, we assign G-NIG five-stars for its calibration capability. However, as emphasized throughout this study, excessive parameters lead to model complexity and prolonged calibration times. For NIG and G-NIG, this is a significant drawback. Therefore, we assign them three-stars and one-star for calibration efficiency, respectively. As for Double t, although the lack of convolution property results in very low computational efficiency, it does not have an excessive number of parameters to calibrate. Thus, its calibration efficiency is similar to that of NIG, and we also assign it three-stars.

On top of that, DE and G-DE perform exceptionally well in terms of calibration efficiency: their convolution property and the simple computation of cumulative distribution functions (CDFs) result in very high computational efficiency. Additionally, their few parameters and simple model structure contribute to efficient calibration. Therefore, we assign DE and G-DE five-stars and four-stars for calibration efficiency, respectively.

Finally, the heavy-tailed characteristics and use of mixture of Gaussian distributions in G-DE enhance its calibration capability. While its calibration capability is second only to G-NIG, its calibration efficiency surpasses that of G-NIG and other copula models. Combining these empirical findings, we conclude that the G-DE copula model outperforms in many aspects, making it an ideal synthetic CDO pricing model.

Chapter 6 Conclusions

This study first emphasizes the importance of copula distribution, which has a greater impact than adding factors. Therefore, this research focuses on the improvement of copula distributions. We introduce various copula models with different distributions proposed in the literature and compare them with our proposed mixed Gaussian and double exponential distribution (G-DE) copula model. Finally, this study summarizes the various advantages of G-DE compared to previously proposed copula models.

Compared to double t, G-DE has the convolution property, allowing its distribution to be obtained without approximation. Compared to NIG, its CDF is simpler and does not require integration, significantly enhancing computational efficiency. Compared to G-NIG, the fewer parameters and simpler model structure of G-DE result in much higher calibration efficiency, without a significant disadvantage in calibration capability.

Therefore, this study concludes that the G-DE copula model outperforms in many aspects and is an ideal model for synthetic CDO pricing.

References

Andersen, L. and Sidenius, J. (2004). Extensions to the gaussian copula: Random recovery and random factor loadings. Journal of Credit Risk Volume, 1(1):05.

Chen, J., Liu, Z., and Li, S. (2014). Mixed copula model with stochastic correlation for cdo pricing. Economic Modelling, 40:167–174.

Choros, B., Härdle, W. K., and Okhrin, O. (2009). Cdo pricing with multifactor and copulae models. Technical report, Citeseer.

Choroś-Tomczyk, B., Härdle, W. K., and Overbeck, L. (2014). Copula dynamics in cdos. Quantitative Finance, 14(9):1573–1585.

Hull, J. et al. (2009). Options, futures and other derivatives/John C. Hull. Upper Saddle River, NJ: Prentice Hall,.

Hull, J., White, A., et al. (2004). Valuation of a cdo and an nth to default cds without monte carlo simulation. Journal of Derivatives, 12(2):8–23.

Jäckel, P. (2005). A note on multivariate gauss-hermite quadrature. London: ABN-Amro. Re.

Kalemanova, A., Schmid, B., Werner, R., et al. (2007). The normal inverse gaussian distribution for synthetic cdo pricing. Journal of derivatives, 14(3):80.

Li, D. X. (1999). On default correlation: A copula function approach. Available at SSRN 187289.

McNeil, A. J. and Nešlehová, J. (2009). Multivariate archimedean copulas, d-monotone functions and ℓ 1-norm symmetric distributions.

Okhrin, O. and Xu, Y. F. (2017). A comparison study of pricing credit default swap index tranches with convex combination of copulae. The North American Journal of Economics and Finance, 42:193–217.

O' Kane, D. and Schloegl, L. (2003). An analytical portfolio credit model with tail dependence. Quantitative Credit Research, Lehman Brothers.

Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. In Annales de l'ISUP, volume 8, pages 229–231.

Vasicek, O. (2002). The distribution of loan portfolio value. Risk, 15(12):160–162.

Wang, D., Rachev, S. T., and Fabozzi, F. J. (2009). Pricing tranches of a cdo and a cds index: Recent advances and future research. Risk Assessment: Decisions in Banking and Finance, pages 263–286.

Xu, G. (2006). Extending gaussian copula with jumps to match correlation smile. Wachovia Securities, online at http://www. defaultrisk. com.

Yang, R., Qin, X., and Chen, T. (2009). Cdo pricing using single factor mg-nig copula model with stochastic correlation and random factor loading. Journal of mathematical analysis and applications, 350(1):73–80.

Appendix A

A.1 Derivation of cdf of 2DE Distribution

In this appendix, we provide the detailed proof of the 2DE distribution cumulative density function (67) introduced in Section 2.7.1. We will derive the density function for the case when $x \geq \mu$, which implies $x \geq 0$. As for the case when $x < 0$, due to its symmetric nature, a similar equation can be applied.

$$
\begin{aligned}
F_{2DE}(x) &= \int_{-\infty}^{x} f(u; \frac{a}{\sqrt{2}}, \frac{\sqrt{1-a^2}}{\sqrt{2}}) du \\
&= \int_{-\infty}^{x} \frac{1}{\sqrt{2}} \frac{a \cdot exp(-\frac{\sqrt{2}|u|}{a}) - \sqrt{1-a^2} \cdot exp(-\frac{\sqrt{2}|u|}{\sqrt{1-a^2}})}{2a^2-1} du \\
&= \frac{1}{\sqrt{2}} \frac{1}{2a^2-1} \left\{ \int_{-\infty}^{0} + \int_{0}^{x} \right\} a \cdot exp(-\frac{\sqrt{2}|u|}{a}) - \sqrt{1-a^2} \cdot exp(-\frac{\sqrt{2}|u|}{\sqrt{1-a^2}}) du \\
&= \frac{1}{2} + \frac{1}{\sqrt{2}} \frac{1}{2a^2-1} \int_{0}^{x} a \cdot exp(-\frac{\sqrt{2}u}{a}) - \sqrt{1-a^2} \cdot exp(-\frac{\sqrt{2}u}{\sqrt{1-a^2}}) du \\
&= \frac{1}{2} + \frac{1}{\sqrt{2}} \frac{1}{2a^2-1} \left[-\frac{1}{\sqrt{2}} a^2 \cdot exp(-\frac{\sqrt{2}u}{a}) + \frac{1}{\sqrt{2}} (1-a^2) \cdot exp(-\frac{\sqrt{2}u}{\sqrt{1-a^2}}) \right]_{0}^{x} \\
&= \frac{1}{2} + \frac{1}{\sqrt{2}} \frac{1}{2a^2-1} \left\{ \frac{1}{\sqrt{2}} (-a^2 \cdot exp(-\frac{\sqrt{2}x}{a}) + (1-a^2) \cdot exp(-\frac{\sqrt{2}x}{\sqrt{1-a^2}})) - \frac{1}{\sqrt{2}} (1-2a^2) \right\} \\
&= \frac{1}{2} + \frac{1}{2} \frac{2a^2-1}{2a^2-1} + \frac{1}{2} \frac{1}{2a^2-1} \left\{ -a^2 \cdot exp(-\frac{\sqrt{2}x}{a}) + (1-a^2) \cdot exp(-\frac{\sqrt{2}x}{\sqrt{1-a^2}}) \right\} \\
&= 1 - \frac{1}{4a^2-2} \left\{ a^2 \cdot exp(-\frac{\sqrt{2}x}{a}) - (1-a^2) \cdot exp(-\frac{\sqrt{2}x}{\sqrt{1-a^2}}) \right\} \text{(for } x \geq 0\text{)}
\end{aligned}
$$

A.2 Proof of central moments of 2DE Distribution

Before we calculate the central moments, we need to obtain derivatives of all orders of its moment generating function. The first to fourth derivatives of the moment generating function with respect to t are as follows:

$$M_{2DE}^{(1)}(t) = \frac{\partial}{\partial t}\left(\frac{1}{\left(1-\frac{a^2t^2}{2}\right)\left(1-\frac{(1-a^2)t^2}{2}\right)}\right) =$$

$$\frac{a^2t}{\left(1-\frac{a^2t^2}{2}\right)^2\left(1-\frac{(1-a^2)t^2}{2}\right)} + \frac{\left(1-a^2\right)t}{\left(1-\frac{a^2t^2}{2}\right)\left(1-\frac{(1-a^2)t^2}{2}\right)^2}$$

$$M_{2DE}^{(2)}(t) = \frac{\partial^2}{\partial t^2}\left(\frac{1}{\left(1-\frac{a^2t^2}{2}\right)\left(1-\frac{(1-a^2)t^2}{2}\right)}\right) = \frac{2a^2\left(1-a^2\right)t^2}{\left(1-\frac{a^2t^2}{2}\right)^2\left(1-\frac{(1-a^2)t^2}{2}\right)^2}+$$

$$\frac{\frac{2\left(1-a^2\right)^2t^2}{\left(1-\frac{(1-a^2)t^2}{2}\right)^3}+\frac{1-a^2}{\left(1-\frac{(1-a^2)t^2}{2}\right)^2}}{1-\frac{a^2t^2}{2}} + \frac{\frac{a^2}{\left(1-\frac{a^2t^2}{2}\right)^2}+\frac{2a^4t^2}{\left(1-\frac{a^2t^2}{2}\right)^3}}{1-\frac{(1-a^2)t^2}{2}}$$

$$M_{2DE}^{(3)}(t) = \frac{\partial^3}{\partial t^3}\left(\frac{1}{\left(1-\frac{a^2t^2}{2}\right)\left(1-\frac{(1-a^2)t^2}{2}\right)}\right) =$$

$$\frac{3a^2t\left(\frac{2\left(1-a^2\right)^2t^2}{\left(1-\frac{(1-a^2)t^2}{2}\right)^3}+\frac{1-a^2}{\left(1-\frac{(1-a^2)t^2}{2}\right)^2}\right)}{\left(1-\frac{a^2t^2}{2}\right)^2} + \frac{\frac{6\left(1-a^2\right)^2t}{\left(1-\frac{(1-a^2)t^2}{2}\right)^3}+\frac{6\left(1-a^2\right)^3t^3}{\left(1-\frac{(1-a^2)t^2}{2}\right)^4}}{1-\frac{a^2t^2}{2}}+$$

$$\frac{3\left(1-a^2\right)t\left(\frac{a^2}{\left(1-\frac{a^2t^2}{2}\right)^2}+\frac{2a^4t^2}{\left(1-\frac{a^2t^2}{2}\right)^3}\right)}{\left(1-\frac{(1-a^2)t^2}{2}\right)^2} + \frac{\frac{6a^6t^3}{\left(1-\frac{a^2t^2}{2}\right)^4}+\frac{6a^4t}{\left(1-\frac{a^2t^2}{2}\right)^3}}{1-\frac{(1-a^2)t^2}{2}}$$

$$M_{2DE}^{(4)}(t) = \frac{\partial^4}{\partial t^4}\left(\frac{1}{\left(1-\frac{a^2t^2}{2}\right)\left(1-\frac{(1-a^2)t^2}{2}\right)}\right) =$$

$$\frac{\frac{36\left(1-a^2\right)^3 t^2}{\left(1-\frac{(1-a^2)t^2}{2}\right)^4} + \frac{6\left(1-a^2\right)^2}{\left(1-\frac{(1-a^2)t^2}{2}\right)^3} + \frac{24\left(1-a^2\right)^4 t^4}{\left(1-\frac{(1-a^2)t^2}{2}\right)^5}}{1-\frac{a^2t^2}{2}} +$$

$$\frac{4a^2t\left(\frac{6\left(1-a^2\right)^2 t}{\left(1-\frac{(1-a^2)t^2}{2}\right)^3} + \frac{6\left(1-a^2\right)^3 t^3}{\left(1-\frac{(1-a^2)t^2}{2}\right)^4}\right)}{\left(1-\frac{a^2t^2}{2}\right)^2} +$$

$$6\left(\frac{a^2}{\left(1-\frac{a^2t^2}{2}\right)^2} + \frac{2a^4t^2}{\left(1-\frac{a^2t^2}{2}\right)^3}\right)\left(\frac{2\left(1-a^2\right)^2 t^2}{\left(1-\frac{(1-a^2)t^2}{2}\right)^3} + \frac{1-a^2}{\left(1-\frac{(1-a^2)t^2}{2}\right)^2}\right) +$$

$$\frac{4\left(1-a^2\right)t\left(\frac{6a^6t^3}{\left(1-\frac{a^2t^2}{2}\right)^4} + \frac{6a^4t}{\left(1-\frac{a^2t^2}{2}\right)^3}\right)}{\left(1-\frac{(1-a^2)t^2}{2}\right)^2} + \frac{\frac{24a^8t^4}{\left(1-\frac{a^2t^2}{2}\right)^5} + \frac{36a^6t^2}{\left(1-\frac{a^2t^2}{2}\right)^4} + \frac{6a^4}{\left(1-\frac{a^2t^2}{2}\right)^3}}{1-\frac{(1-a^2)t^2}{2}}$$

Next, we can calculate the central moments by setting $t = 0$. Assuming $X \sim 2DE\left(\frac{a}{\sqrt{2}}, \frac{\sqrt{1-a^2}}{\sqrt{2}}\right)$, the first to fourth central moments are as follows:

$$E[X] = 0$$

$$E\left[X^2\right] = 1$$

$$E\left[X^3\right] = 0$$

$$E\left[X^4\right] = 6(a^4 - a^2 + 1)$$

Finally, we can calculate Kurtosis as follows:

$$\text{Kurt}[X] = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] = \frac{E\left(X^4\right) - 4E\left(X^3\right)\mu + 6E\left(X^2\right)\mu^2 - 3\mu^4}{\left(\sigma^2\right)^2}$$

$$= 6(a^4 - a^2 + 1)$$