Copulas in Equity and Credit Risk

Default-Dependent Intensity Models and Information-Based Setup

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20160506-115019-2
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20160506-115019-2]

Copulas in Equity and Credit Risk

Default-Dependent Intensity Models and Information-Based Setup

Dissertation

zur Erlangung des akademischen Grades eines

Doktor der Naturwissenschaften (Dr. rer. nat.)

der Fakultät für Mathematik und Naturwissenschaften
der Bergischen Universität Wuppertal vorgelegt von

M.Sc. Patrick Deuß

Dezember 2015

To My Wife Alexandra

∞

For Her Infinite Love

Without her endless patience,
extraordinary and great understanding,
and enduring encouragement
this thesis would never have been finished.

I love you.

Danksagung

Ich widme diese Arbeit meiner Frau Alexandra. Ohne ihre Geduld, ihre unentwegte Ermutigungen und ihr Verständnis hätte ich diese Arbeit wohl nicht fertiggestellt. Danke für Deine Liebe.

Der größte Dank gilt meinem Doktorvater Prof. Dr. Michael Günther, der - neben seinem fachlichen Input - ebenfalls eine außerordentliche Geduld und viel Verständnis über eine Zeit von ca. acht Jahren aufgebracht hat. Michael, ich danke Dir von ganzem Herzen, dass Du mir ein solches Projekt ermöglicht hast und wir es zusammen abgeschlossen haben.

Für die fachliche Unterstützung bedanke ich mich ganz besonders bei Dr. Jan-Philipp Hofmann (Deutsche Postbank AG), Prof. Dr. Barbara Rüdiger und Dr. Long Teng.

Philipp, vielen Dank für zahllose Gespräche zu möglichen Modellanwendungen, unterschiedlichen Wahrscheinlichkeitsmaßen, Spreads, etc., sowie für Deine stetige und schnelle Unterstützung, wenn ich sie brauchte.

Barbara, ich danke Dir für viele fachliche Diskussionen zur Stochastik, stochastischen Analysis und insbesondere zu Filtrationen, die die Arbeit bedeutend vorangetrieben haben.

Long, vielen Dank für den Review meiner Arbeit. Deine Anmerkungen waren besonders wertvoll.

Für unzählige fachliche Diskussionen, Hilfe und Unterstützung, Hinweise und Verbesserungen, neue Sichtweisen und Denkansätze sowie Marktdaten bedanke ich bei allen Kollegen der

- Deutsche Postbank AG - Dr. Jörg Kienitz, Daniel Wetterau, Dr. Manuel Wittke, Boris Tschierschke und Dr. Carsten Franz
- Barmenia Versicherungen - Tanja Dicke und Manuel Kasparek,
- Arbeitsgruppe AMNA - Dr. Kai Tappe, Dr. Cathrin van Emmerich und Dr. Christian Kahl.

Holger Rocke und Ulrich Andritzky danke ich für ihr Feedback zur Einleitung.

Eine berufsbegleitende Dissertation bringt besondere organisatorische Herausforderungen mit sich. Dr. Andreas Eurich (Barmenia Versicherungen) möchte an dieser Stelle einen besonderen Dank für seine Hilfe aussprechen.

Ich danke meinen Vorgesetzten beim BVR / bei der parcIT (Bernhard Krob und Stefan Schillmann), bei der Deutsche Postbank AG (Dr. Frank Beekmann, Christoph Blom, Susanne Röhrig und Martin Slowik) und bei den Barmenia Versicherungen (Frank Sievert) für ihr Verständnis und ihre Unterstützung.

Ein besonderer Dank gilt dem Team Kreditportfoliomodelle des BVR / der parcIT, das mir im letzten Jahr einen unschätzbaren Freiraum gegeben hat und somit erheblich zur Finalisierung der Arbeit beigetragen hat. Dies sind Dr. Jürgen Braun, Markus Hirsch und Jörg Lillpopp.

Zudem danke ich der Arbeitsgruppe AMNA der Universität Wuppertal. Hier möchte ich insbesondere Prof. Dr. Matthias Ehrhardt und Dr. Andreas Bartel erwähnen, die stets ein offenes Ohr für mich hatten. Darüber hinaus wäre ich ohne den technischen Support von Daniel Heubes aufgeworfen gewesen. Prof. Dr. Bruno Lang bin ich für seinen flexiblen Einsatz in der Prüfungskommission sehr dankbar.

Meiner Familie und meinen Freunden danke ich für die unermüdliche Unterstützung während der gesamten Arbeit. Insbesondere möchte ich meiner Schwester Carina, meinen Eltern Verena und Hans-Gerd sowie meinen Schwiegereltern Brigitte und Hans Claßen dafür danken, dass sie stets an mich geglaubt haben.

Table of Contents

List of Figures

Part I

Part II

Appendix

List of Tables

Part I

Part II

Appendix

List of Algorithms

Abbreviations and Notation

Abbreviations

1. cdf = cumulative distribution function - The expressions **distribution, distribution function** and **cumulative distribution function** are used synonymously. The plural form is denoted as cdfs.
2. i.i.d. = independent identical distributed
3. w.l.o.g. = without loss of generality
4. w.r.t. = with respect to
5. $\mathbb{P}-$a.s / $\mathbb{Q}-$a.s. = $\mathbb{P}$-almost surely / $\mathbb{Q}$-almost surely w.r.t. probability measures $\mathbb{P}$ / $\mathbb{Q}$
6. a.a. = almost all

Notation

Sets

1. $\mathbb{R} = (-\infty, \infty)$, $\bar{\mathbb{R}} = [-\infty, \infty]$
2. $\mathbb{R}^D = (-\infty, \infty)^D$, $\bar{\mathbb{R}}^D = [-\infty, \infty]^D$
3. $\mathbb{R}^+ = (0, \infty)$, $\mathbb{R}_{0,\infty} = [0, \infty]$
4. For an arbitrary set B, $\#B$ defines the number of elements in B.

Vectors

1. All vectors are denoted as column vectors.
2. If A is a matrix or a vector, we denote A' as its transposed.
3. $|A|$ is the determinant of matrix A.
4. $\mathbb{I}_D$ is the identity matrix of size $D \times D$.
5. $\mathbf{0}_D$ is the D-dimensional zero vector $(0, \ldots, 0)'$.
6. For $a, b \in \bar{\mathbb{R}}^D$, the expression $a \leq b$ states that $a_d \leq b_d$ for all $d = 1, \ldots, D$.
7. A random variable $X = (X_1, \ldots, X_D)'$ will be capitalized.
8. A realization of a random variable X is denoted with small letters $x = (x_1, \ldots, x_D)'$.
9. For a multidimensional stochastic process x_t the subscript for time-dependence is put to the lower right, the index for the dimension to the upper right of the vector, i.e.,

$$x_t = (x_t^1, \ldots, x_t^D)' \quad .$$

10. Let $\mathcal{D} = \{1, \ldots, D\}$ be an arbitrary set of indices. Let $\mathcal{E}$ be a subset of $\mathcal{D}$ with

$$\mathcal{E} = \{d_1, \ldots, d_E | d_e \in \mathcal{D}, e = 1, \ldots, E\} \quad .$$

Then, the vector $x(\mathcal{E})$ is defined as $x(\mathcal{E}) = (x_{d_1}, \ldots, x_{d_E})'$ with indices in $\mathcal{E}$.

Probability Theory

1. $\mathbb{P}$ and $\mathbb{Q}$ define probability measures. $\mathbb{P}_X$ denotes the probability measure of a random variable X, i.e., its distribution.
2. $\mathbb{E}[X]$ is the expectation of a random variable X.
3. $\mathbb{V}[X]$ is the variance of a random variable X.
4. $X \perp Y$ means that the random variables X and Y are independent.
5. $X \overset{d}{=} Y$ defines that random variables X and Y are equivalent in distribution.
6. $X_n \xrightarrow{d} X$ means that X_n converges in distribution to X for $N \to \infty$.
7. For random variables X and stochastic processes X_t we omit the dependence on $\omega \in \Omega$, when it is not explicitly necessary. We write $X(\omega) = X$ and $X_t(\omega) = X_t$.

Miscellaneous

1. $\mathbf{1}_A(x)$ is the indicator function:

$$\mathbf{1}_A(x) = \begin{cases} 0 & \text{for } x \notin A \\ 1 & \text{for } x \in A \end{cases}$$

2. The *sign-function* is defined as

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{for } x > 0 \\ 0 & \text{for } x = 0 \\ -1 & \text{for } x < 0 \end{cases} .$$

3. Let f and g be arbitrary functions and $a \in \bar{\mathbb{R}}$. The *Landau Symbols* are defined as

$$f(x) = \mathcal{O}(g(x)): \quad \lim_{x \to a} \frac{f(x)}{g(x)} = const \qquad \text{and}$$
$$f(x) = o(g(x)): \quad \lim_{x \to a} \frac{f(x)}{g(x)} = 0 \quad .$$

4. An empty sum is defined to be zero, i.e., $\sum_{i=n}^{m} x(i) = 0$ for $m < n$.

Univariate Distributions

General Distribution Function
Notation: $X \sim F$ - $F_X(x) = \mathbb{P}[X \leq x]$ for $x \in \bar{\mathbb{R}}$

Standard Uniform Distribution
Notation: $U \sim U(0, 1)$

Normal Distribution
Notation: $X \sim N(\mu, \sigma)$ - Abbreviation: $\Phi_{\mu,\sigma}(x)$

Standard Normal Distribution
Notation: $X \sim N(0, 1)$ - Abbreviation: $\Phi(x)$

***t*-distribution, ν degrees of freedom**
Notation: $X \sim t(\nu)$ - Abbreviation: $t_\nu(x)$

***χ^2*-distribution, ν degrees of freedom**
Notation: $X \sim \chi^2(\nu)$ - Abbreviation: $\chi^2_\nu(x)$

Multivariate Distributions

General Distribution Function
Notation: $X = (X_1, \ldots, X_D)' \sim F$ - defined as

$$F_X(x) = F_{X_1,\ldots,X_D}(x_1, \ldots, x_D) = \mathbb{P}[X_1 \leq x_1, \ldots, X_D \leq x_D] = \mathbb{P}[X \leq x]$$

for $x = (x_1, \ldots, x_D)' \in \bar{\mathbb{R}}^D$, $D \in \mathbb{N}$.

Uniform Distribution
Notation: $U \sim U(0,1)^D$

Normal Distribution
Notation: $X \sim N_D(\mu, \Sigma)$ - Abbreviation: $\Phi_{\mu,\Sigma}(x)$ defined as

$$\Phi_{\mu,\Sigma}(x) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma|^{\frac{1}{2}}} \int_{-\infty}^{x_D} \ldots \int_{-\infty}^{x_1} \exp\left(-(z-\mu)'\Sigma^{-1}(z-\mu)\right) dz_1 \ldots dz_D$$

with mean $\mu = (\mu_1, \ldots, \mu_D)' \in \mathbb{R}^D$ and covariance matrix $\Sigma \in \mathbb{R}^{D \times D}$.

Standard Normal Distribution
Notation: $X \sim N_D(0, P)$ - Abbreviation: $\Phi_P(x)$ with $\mu = \mathbf{0}_D$, correlation matrix $P \in \mathbb{R}^{D \times D}$.

Standard t-distribution
Notation: $X \sim t_D(\nu, P)$ - Abbreviation: $t_{\nu,P}(x)$ defined as

$$t_{\nu,P}(x) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_D} \frac{\Gamma\left(\frac{\nu+D}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)(\nu\pi)^{\frac{D}{2}}|P|^{\frac{1}{2}}} \left(1 + \frac{z'P^{-1}z}{\nu}\right)^{\frac{\nu+D}{2}} dz_1 \ldots dz_D \quad .$$

Γ denotes the gamma function, ν the degrees of freedom and $P \in \mathbb{R}^{D \times D}$ the correlation matrix.

Introduction

Terms, definitions and types of risks to which financial institutions are exposed are manifold. They are commonly differentiated by their source or their scope of application. Risk management and controlling departments mainly[1] distinguish between

- market risk - e.g., equity or interest rate risk,
- credit risk - e.g., default risk, and
- operational risk - e.g., individual mistakes of employees.

Identification of risk is sophisticated on account of the great variety of possible influencing factors. Due to the inherent uncertainty, risk is generally measured through probabilities. Thus, mathematical tools and schemes are frequently used for estimation and evaluation. Besides the resulting complexity, these models must be developed, calibrated to data and numerically implemented. For this process, conflicting standards and demands concerning

- internal and strategic objectives as well as
- regulatory requirements

must be considered. Risk aggregation and measuring dependencies between risk factors are additional issues.

In practice, these challenges are generally tackled by standard or simplified approaches. They are less complex, their handling is flexible and numerical implementation is fast. Therefore, their known drawbacks (e.g., under- and overestimation of risk) are accepted. In academia, many advanced models have been developed to eliminate these failures. As consequence, practical implementation is difficult and expensive.

In this thesis, we focus on modeling dependency structures by means of **copulas**[2]. Compared to standard frameworks, the copula approach is more elaborate but still provides a feasible implementation.

In a mathematical nutshell, the concept states that a multivariate distribution can be split into its one-dimensional marginal distributions and a coupling function denoted as copula. Transferred to an economic point of view, a multidimensional problem of risk aggregation can be separated into

- its single risk factors and
- its dependency structure - i.e., its copula.

The main advantage of this idea is that single risk factors are often easily measurable. Though, identification of dependencies is challenging. Copulas offer a large variety of dependency structures and hence a higher degree of freedom compared to standard approaches (multivariate normal distributions as a rule). However, the selection of the "right" copula is of central significance. In practice, copulas already cover various scopes of applications.

The first part of this thesis introduces to the mathematical background of copulas. As application, the value at risk of a stock portfolio is measured by the copula approach. Outcomes are compared to a standard multivariate normal benchmark.

The second part provides an introduction to credit risk and a detailed description of so-called intensity-based models. In this framework, we develop a new specification for modeling a copula- and default-dependent intensity. As conclusion, model construction and implementation are critically assessed.

[1] Further kinds (systemic, liquidity or model risk, for instance) exist and have ambiguous assignments to different areas.

[2] In literature, we casually find the plural form *copulae*.

Part I

Copula Functions and Applications in Equity Risk

1 Introductory Note to the First Part

Measuring equity risks of portfolios is an essential issue for financial institutions. It is often difficult to determine portfolio behavior at large. Single risk factors may easily be identifiable and transparent. However, dependencies between risk factors must be specified to assess their impact on the portfolio and to picture portfolio evolution.

The separation of risk drivers and dependency structure is met by the copula concept. From a mathematical point of view it is possible to split a multidimensional distribution (here the portfolio behavior) into its margins (here the single risk factors) and a copula (here the risk factors' dependency structure).

The first financial application of the copula approach was made by Li [41] in 2000. He uses the Gaussian copula to describe dependencies in credit default risk. In this first part we apply copulas in an equity risk framework. As we are free in selecting the copula, the main difficulty for these models is the choice of the **right** copula to identify the actual dependency structure. Our analysis is restricted to the Gaussian, t- and Clayton copula. A detailed extension to the great variety of copulas goes beyond the scope of this thesis.

We begin with an introduction to the concept, theory and simulation of copulas supported by examples in Chapter 2. Calibration procedures and evaluation of the goodness of fitting are presented in Chapter 3. In Chapter 4, a goodness-of-fit-test is applied to equity portfolios consisting of up to 25 companies of the German stock index DAX for different periods. Moreover, we simulate the value at risk for several portfolio sizes by above mentioned copula models. Outcomes are compared to a standard approach which is often applied in practice. Numerical results are presented and interpreted by a first conclusion. In Chapter 5, a recalibration procedure is developed for a second application, a backtesting of the value at risk. Again, simulation results of the three copulas approaches are compared to those of the standard model. A second conclusion is drawn, modeling assumptions and resulting problems are discussed. A summary of results and conclusions is given. Explanatory notes conclude part one.

2 Basic Copula Theory

In this chapter we give an introduction to copulas and describe their elementary features. Provided examples deliver insight into the topic. Two copula sampling approaches are presented in the last section.

Expressions marked by an asterisk (*) are explained in Appendix A.1. Appendices A.2 and A.3 yield a survey of the basic notation and characterization of multivariate distributions and conditional cdfs. For a small overview of random number generation, the reader is referred to A.4.

2.1 Definitions and Properties

Due to their substantial properties, it quickly becomes clear why copulas are a helpful tool for the analysis of multivariate dependence structures.

2.1 Definition (compare [15], [63])
A *D-dimensional copula* is a function C mapping from $[0,1]^D$ to $[0,1]$ such that

1. C is grounded*,
2. $C(1,\dots,1,u_d,1,\dots,1) = u_d$ for all $d \in \{1,\dots,D\}$, and
3. C is D-increasing*.

To get the idea about the definition, let $U = (U_1, \ldots, U_D)'$ be a D-dimensional random variable in which every U_d is standard uniformly distributed for each $d = 1, \ldots, D$. For every vector $u = (u_1, \ldots, u_D)' \in [0,1]^D$, we associate the copula C to the multivariate probability

$$C(u) = C(u_1, \ldots, u_D) = \mathbb{P}[U \leq u] = \mathbb{P}[U_1 \leq u_1, \ldots, U_D \leq u_D] \quad .$$

Furthermore, a copula fulfills the necessary and sufficient conditions to be a multivariate distribution function, see [37], Section 1.4.2 or [52], Definition 2.10.8. The second property of Definition 2.1 is required for uniform marginal distributions.

Thus, a copula function is a **multivariate distribution function with standard uniform margins**. We will always denote a copula by C instead of F_U.

It appears natural to question in how far multivariate cdfs, margins and copula functions can be be combined. The answer is given by Sklar's Theorem (1959), [61].

2.2 Theorem (compare [49], Theorem 5.3)
Let F be a multivariate distribution function with margins F_d, $d = 1, \ldots, D$.

1. Then, a copula $C : [0,1]^D \to [0,1]$ exists such that it holds
$$F(x_1, \ldots, x_D) = C(F_1(x_1), \ldots, F_D(x_D)) \tag{2.1}$$
for each $x = (x_1, \ldots, x_D)' \in \bar{\mathbb{R}}^D$. If the margins F_d, $d = 1, \ldots, D$, are continuous, C is unique. Otherwise, C is uniquely determined on Range $F_1 \times \ldots \times$ Range F_D.
2. If C is a copula and F_d, $d = 1, \ldots, D$, are univariate distributions, then F in (2.1) is a multivariate distribution function with margins F_d.

A proof can be found in [52], Section 2.3. ∎

A direct implication of the above theorem is

2.3 Corollary (compare [52], Corollary 2.10.10)
Let F be a multivariate cdf and C its copula in sense of Equation (2.1) with margins $F_1, \ldots, F_D$. Let $F_1^{-1}, \ldots, F_D^{-1}$ be their generalized inverse functions*. Then, for any $u = (u_1, \ldots, u_D)' \in [0,1]^D$ it holds
$$C(u_1, \ldots, u_D) = F(F_1^{-1}(u_1), \ldots, F_D^{-1}(u_D)) \quad . \tag{2.2}$$

Outline of proof:

Starting with equation (2.1), one exploits the quantile transformation property of the generalized inverse functions F_d^{-1}, $d = 1, \ldots, D$, to receive (2.2). ∎

Theorem 2.2 and Corollary 2.3 require some interpretation. First, Sklar's Theorem states that one can connect margins F_d, $d = 1, \ldots, D$, by a copula C to receive their multivariate distribution F. Second, Corollary 2.3 asserts that a copula C is exhibited by means of the multivariate distribution function F in combination with the generalized inverse functions F_d^{-1} of its margins F_d.

Hence, the copula C reveals the dependencies between the marginal distributions F_d of the multivariate cdf F. This separation of margins and multivariate distribution appears to be a powerful tool for applications in quantitative finance.

If F and F_d, $d = 1, \ldots, D$, are continuous and strictly increasing, we can compute the *copula density* c by applying the chain rule to (2.2) as

$$c(u_1, \ldots, u_D) = \frac{\partial C(u_1, \ldots, u_D)}{\partial u_1 \ldots \partial u_D} = \frac{f_X(F_1^{-1}(u_1), \ldots, F_D^{-1}(u_D))}{\prod_{d=1}^{D} f_d(F_d^{-1}(u_d))} \tag{2.3}$$

where f_d denotes the marginal density of X_d and f the multivariate density of X. Computation of these expressions are necessary for calibration issues and goodness-of-fit tests in Chapter 3.

2.4 Notation
The copula of a random variable $X \in \mathbb{R}^D$ with $X \sim F$ is denoted as C_X if indexing is necessary. The *copula of a distribution function* F is meant in sense of Equation (2.1), i.e., F can be represented in terms of a copula C and margins $F_1, \ldots, F_D$.

A further important property is that copulas are invariant under strictly increasing transformations - for the following statement compare [52], Theorem 2.4.3, or [49], Proposition 5.6.

2.5 Proposition (Invariance Property)
Let C_X be the copula of a random variable $X = (X_1, \ldots, X_D)' \in \mathbb{R}^D$ with continuous margins $F_1, \ldots, F_D$ and let $T_1, \ldots, T_D$ be strictly increasing transformations. Then, the transformed random variable

$$T(X) := (T(X_1), \ldots, T_D(X_D))' \in \mathbb{R}^D$$

has copula C_X as well, i.e., it holds

$$C_X = C_{T(X)} \quad .$$

A proof can be found in Appendix D.1.1. ∎

2.2 Examples

As pointed out, a copula represents the dependence structure of multivariate distributed random variables. Basically, we can divide copulas into three different categories, see [49], Section 5.1.2.

2.2.1 Fundamental Copulas

These kind of copulas define fundamental - in sense of elementary - dependency structures.

2.6 Definition (Fundamental Copulas)

1. The *independence* or *product copula* C_Π is defined as

$$C_\Pi(u_1, \ldots, u_D) := \prod_{d=1}^{D} u_d \quad .$$

2. The *comonotonicity copula* M is defined as

$$M(u_1, \ldots, u_D) := \min\{u_1, \ldots, u_D\}$$

and represents **perfect positive** dependence.

3. The *countermonotonicity function*[1] W is defined as

$$W(u_1, \ldots, u_D) := \max\left\{\sum_{d=1}^{D} u_d - D + 1, 0\right\}$$

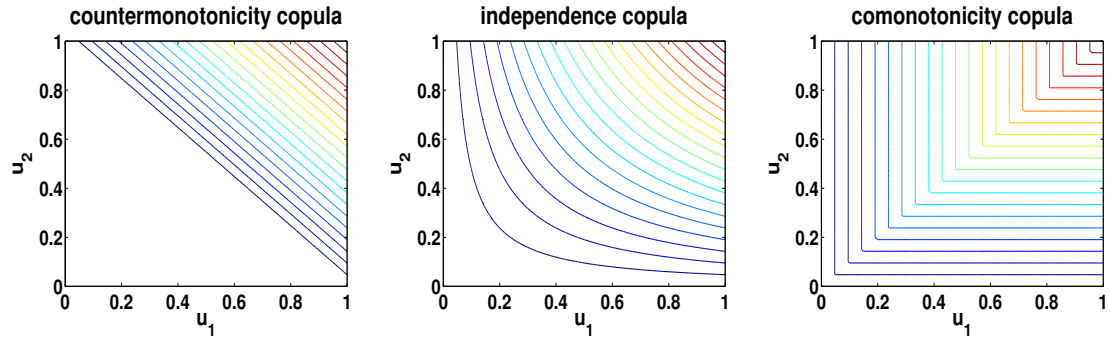
and represents **perfect negative** dependence.

One can show (see [49], Theorem 5.7 and proceeding proof) that every copula C is bounded by countermonotonicity W and comonotonicity M, i.e.,

$$W(u_1, \ldots, u_D) \leq C(u_1, \ldots, u_D) \leq M(u_1, \ldots, u_D) \quad .$$

W and M are called *Fréchet lower* and *upper bound*, respectively[2]. Figure F - 2.1 shows contour plots of the cumulative distribution functions of W, C_Π and M. It gives a graphical hint why countermonotonicity and comonotonicity can be interpreted as lower and upper bound for every copula.

[1] W is **not** a copula in general. Only in the case $D = 2$ $W(u_1, u_2)$ fulfills the requirements of Definition 2.1.
[2] Families of copulas that include the fundamental ones are said to be *comprehensive*. Their determining parameters allow a representation of W, C_Π and M.



F - 2.1 Figure: contour plots - fundamental copulas

2.2.2 Implicit Copulas

Implicit copulas are extracted from multivariate distributions and inherit their properties explaining the word **implicit**. We present two of the most commonly used in quantitative finance.

2.7 Definition

1. The *Gaussian copula* is defined as

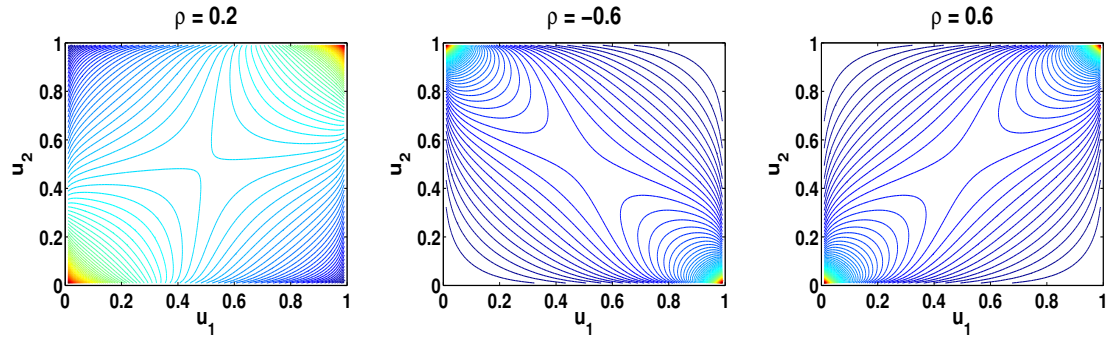
$$C_P^{Ga}(u_1,\dots,u_D,P) := \Phi_P(\Phi^{-1}(u_1),\dots,\Phi^{-1}(u_D))$$
$$= \int_{-\infty}^{\Phi^{-1}(u_D)} \dots \int_{-\infty}^{\Phi^{-1}(u_1)} \frac{1}{(2\pi)^{\frac{D}{2}}|P|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}z'P^{-1}z\right) dz_1 \dots dz_D$$

   in which Φ_P is the multivariate standard normal distribution with correlation matrix P and Φ^{-1} the inverse function of the univariate standard normal distribution.

2. The *t-copula* is defined as

$$C_{\nu,P}^{t}(u_1,\dots,u_D,\nu,P) := t_{\nu,P}\left(t_\nu^{-1}(u_1),\dots,t_\nu^{-1}(u_D)\right)$$
$$= \int_{-\infty}^{t_\nu^{-1}(u_D)} \dots \int_{-\infty}^{t_\nu^{-1}(u_1)} \frac{\Gamma\left(\frac{\nu+D}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)(\nu\pi)^{\frac{D}{2}}|P|^{\frac{1}{2}}} \left(1+\frac{z'P^{-1}z}{\nu}\right)^{\frac{\nu+D}{2}} dz_1 \dots dz_D$$

   in which $t_{\nu,P}$ is the multivariate t-distribution with ν degrees of freedom, correlation matrix P, and t_ν^{-1} the inverse function of the univariate t-distribution.



F - 2.2 Figure: contour plots - bivariate Gaussian copula density

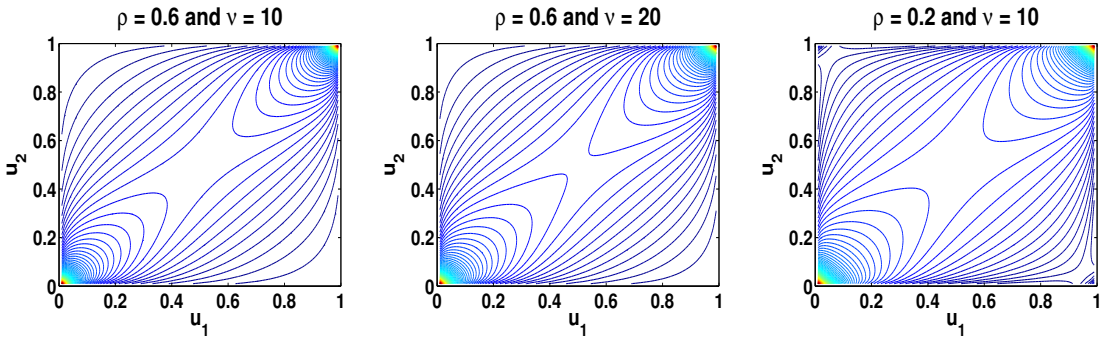
C_P^{Ga} and $C_{\nu,P}^t$ belong to the family of elliptical copula functions, a subclass of elliptical distributions, refer to [22], Chapter 5. Both copulas have closed form densities, for a derivation see [12], Sections 4.8.1 and 4.8.2. For the Gaussian Copula it holds

$$c(u_1,\ldots,u_D,P) = |P|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\xi'\left(P^{-1}-\mathbb{I}_D\right)\xi\right) \tag{2.4}$$

in which $\xi = \left(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_D)\right)'$. For $D=2$, contour plots for Equation (2.4) with different correlations are shown in Figure F - 2.2. For the t-copula we have

$$c(u_1,\ldots,u_D,\nu,P) = |P|^{-\frac{1}{2}} \frac{\Gamma\left(\frac{\nu+D}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}\right)^D \frac{\left(1+\frac{1}{\nu}\xi' P^{-1}\xi\right)^{-\frac{\nu+D}{2}}}{\prod_{d=1}^{D}\left(1+\frac{\xi_d^2}{\nu}\right)^{-\frac{\nu+1}{2}}} \tag{2.5}$$

in which $\xi = \left(t_\nu^{-1}(u_1),\ldots,t_\nu^{-1}(u_D)\right)'$. Again for $D=2$, Figure F - 2.3 illustrates contour plots for (2.5) with different degrees of freedom and correlations.



F - 2.3 Figure: contour plots - bivariate t-copula density

The driving parameter for the Gaussian copula is its correlation matrix P. Taking P as identity matrix $\mathbb{I}_D$, we receive the D-dimensional independence copula C_Π. Further, if all entries of P are one, we obtain the comonotonicity copula M. The same holds for the t-copula $C_{P,\nu}^t$. However, assuming $\nu < \infty$, $P = \mathbb{I}_D$ does **not** result in independence for the t-copula.

Due to the fact that copulas are invariant under strictly increasing transformations, a random variable $Y \sim \Phi_{\mu,\Sigma}$ with covariance matrix Σ has the same copula as $X \sim \Phi_{0,P} = \Phi_P$ with correlation matrix P. This is valid as the standardization from Y to X exclusively requires strictly increasing transformations, see [49], Proposition 5.6 and Subsection 5.1.2.

2.2.3 Archimedean Copulas

As seen in the previous subsection, implicit copulas are functions of given multivariate and marginal distributions. For **explicit** copulas this is not the case. They have an explicit, often a very simple closed form. Mostly, they belong to the family of **Archimedean** copulas.

Suppose that $\phi : [0,1] \to [0,\infty]$ is a continuous, strictly decreasing function with $\phi(1) = 0$ and $\phi(0) = \infty$ and let ϕ^{-1} be its inverse. Then, the function $C : [0,1]^D \to [0,1]$ given by

$$C(u_1,\ldots,u_D) = \phi^{-1}\left(\phi(u_1)+\ldots+\phi(u_D)\right) \tag{2.6}$$

is a copula if and only if ϕ^{-1} is completely monotonic* on $[0,\infty)$, e.g., see [52], Theorem 4.6.2. Families of form (2.6) are called *Archimedean copulas with generator* ϕ. One can easily prove that this class of copulas is symmetric and associative[3].

[3]For the bivariate case associativity means $C(C(u_1,u_2),u_3) = C(u_1,C(u_2,u_3))$.

2.8 Example (Clayton Copula)
Set $\phi_\theta(t) = \frac{1}{\theta}\left(t^{-\theta} - 1\right)$ with $\theta \geq -1$, $\theta \neq 0$. Its inverse $\phi^{-1}(s) = (\theta s + 1)^{-\frac{1}{\theta}}$ is completely monotonic for $\theta > 0$. For $D \geq 2$, the *Clayton* copula is given by

$$C_\theta^{Cl}(u_1, \ldots, u_D) = \left(\sum_{d=1}^{D}\left(u_d^{-\theta} - 1\right) + 1\right)^{-\frac{1}{\theta}} \quad . \tag{2.7}$$

For $\theta = -1$, we receive the Fréchet lower bound[4] W. Further, the Clayton copula family is comprehensive as $\lim_{\theta\to 0} C_\theta^{Cl} = C_\Pi$ and $\lim_{\theta\to\infty} C_\theta^{Cl} = M$, see [49], Table 5.4.

2.9 Example (Gumbel Copula)
Set $\phi_\theta(t) = (-\ln t)^\theta$ with $\theta > 1$. Its inverse $\phi^{-1}(s) = \exp\left(-s^{\frac{1}{\theta}}\right)$ is completely monotonic on $[0, \infty)$. For $D \geq 2$, the *Gumbel* copula is given by

$$C_\theta^{Gu}(u_1, \ldots, u_D) = \exp\left(\left(\sum_{d=1}^{D}\left(-\ln(u_d)\right)^{-\theta}\right)^{\frac{1}{\theta}}\right) \quad . \tag{2.8}$$

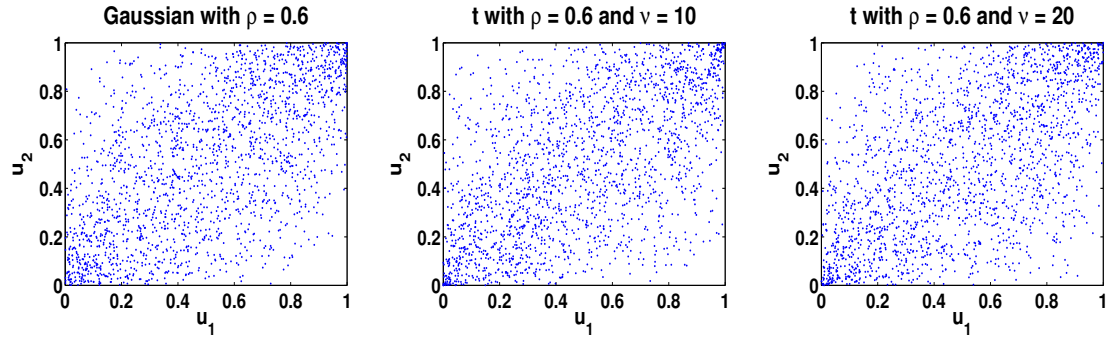
For $\theta = 1$ and $\theta \to \infty$, we receive independence C_Π and comonotonicity M, see [49], Table 5.4.

2.3 Simulation

For applications we need to generate random numbers distributed according to the copula in question. The majority of computing languages provides at least one pseudo-random generator for standard uniform distributed numbers. Moreover, many mathematical software programs supply an extensive quantity of random number generators of the most frequently used distributions for both the univariate and the multivariate case. The two main approaches, how to sample copula-distributed random variables, are presented in the preceding subsections 2.3.1 and 2.3.2. A substantial treatment of non-uniform random number generation is [18].

2.10 Notation

1. A realization $x = (x_1, \ldots, x_D)'$, drawn from a random variable $X = (X_1, \ldots, X_D)'$ distributed according to a distribution function F, is abbreviated by $x \sim F$ or $(x_1, \ldots, x_D)' \sim F$.
2. A realization $u = (u_1, \ldots, u_D)'$, drawn from a random variable $U = (U_1, \ldots, U_D)'$ distributed according to a copula function C, is abbreviated by $u \sim C$ or $(u_1, \ldots, u_D)' \sim C$.



F - 2.4 Figure: 2000 random samples - Gaussian and t-copula

[4] For $\theta = -1$, ϕ is not strict anymore.

2.3.1 Transformation Method

The transformation method is applicable for implicit copulas. This approach exploits the fact that a copula can be represented by Equation (2.2) as

$$C(u_1, \ldots, u_D) = F(F_1^{-1}(u_1), \ldots, F_D^{-1}(u_D)) \quad .$$

Let $X \in \bar{\mathbb{R}}^D$ be a random variable with $X \sim F$. If we apply the marginal distributions F_d, $d = 1, \ldots, D$, to each component X_d of X, we receive a new variable $U \in [0,1]^D$ which is distributed according to the copula C. Thus, this procedure allows to simulate random variables $U \sim C$ provided that the copula can be represented as in (2.2).

A - 2.1 Algorithm (Transformation Method)

Step 1

- draw $x = (x_1, \ldots, x_D)' \sim F$

Step 2 - for $d = 1, \ldots, D$:

- compute $u_d = F_d(x_d)$

Step 3

- receive $u = (u_1, \ldots, u_D)' \sim C$

In particular, this pseudo-algorithm can be applied to the Gaussian and the t-copula. We adopt the following algorithm from [49], Subsection 5.1.4. For the generation of multivariate normal and t-distributed variables, the reader is referred to Appendix A.4.

A - 2.2 Algorithm (Gaussian Copula Sampling)

Step 1

- draw $x = (x_1, \ldots, x_D)' \sim \Phi_P$

Step 2 - for $d = 1, \ldots, D$:

- compute $u_d = \Phi(x_d)$

Step 3

- receive $u = (u_1, \ldots, u_D)' \sim C_P^{Ga}$

A - 2.3 Algorithm (t-Copula Sampling)

Step 1

- draw $x = (x_1, \ldots, x_D)' \sim t_{\nu,P}$

Step 2 - for $d = 1, \ldots, D$:

- compute $u_d = t_\nu(x_d)$

Step 3

- receive $u = (u_1, \ldots, u_D)' \sim C_{\nu,P}^t$

Figure F - 2.4 shows random points sampled by Algorithms A - 2.2 and A - 2.3. The main advantage of the transformation method is a fast and simple implementation.

However, this method fails if distributions F and F_d, $d = 1, \ldots, D$, are not specified[5]. Thus, as general concept for copula sampling, the transformation approach is not suitable.

[5] For instance, this is the case for explicit copulas.

2.3.2 Conditional Sampling

The drawback of the previous section is solved by the so-called *conditional sampling* motivated in [12], Section 6.3, for instance[6]. Conditional simulation bases upon probability integral transformations (PIT's) by Rosenblatt [56][7]. As we will make use of PIT's in Section 3.3 as well, we go in for Rosenblatt's general result. A small reminder on conditional distributions is presented in Appendix A.3.

2.11 Notation
Let F be a multivariate cdf of $X \in \mathbb{R}^D$. Denote $F_{d|1,\dots,d-1}$ as the dth margin of X_d given $X_1, \dots, X_{d-1}$. This means it holds

$$F_{d|1,\dots,d-1}(x_d|x_1,\dots,x_{d-1}) = \frac{\dfrac{\partial^{d-1}F_{1,\dots,d}(x_1,\dots,x_d)}{\partial x_1,\dots\partial x_{d-1}}}{\dfrac{\partial^{d-1}F_{1,\dots,d-1}(x_1,\dots,x_{d-1})}{\partial x_1\dots\partial x_{d-1}}} = \frac{\dfrac{\partial^{d-1}F(x_1,\dots,x_d,\infty,\dots,\infty)}{\partial x_1,\dots\partial x_{d-1}}}{\dfrac{\partial^{d-1}F(x_1,\dots,x_{d-1},\infty,\dots,\infty)}{\partial x_1\dots\partial x_{d-1}}} \quad .$$

2.12 Theorem (compare [56])
Let $X = (X_1, \dots, X_D)' \in \mathbb{R}^D$ be a random variable with absolutely continuous cdf F and dth margins $F_{d|1,\dots,d-1}$.

1. Define D transformations $T_d : \mathbb{R} \to [0,1]$, $d = 1, \dots, D$, as

$$\begin{aligned} T_1(x_1) &= F_1(x_1) \quad , \\ T_2(x_2) &= F_{2|1}(x_2|x_1) \quad , \\ &\vdots \\ T_D(x_D) &= F_{D|1,\dots,D-1}(x_D|x_1,\dots,x_{D-1}) \quad . \end{aligned} \tag{2.9}$$

   Then, $v_d = T_d(x_d)$, $d = 1, \dots, D$, are independent and standard uniformly distributed.

2. Conversely, for an independent random variable $V \sim U(0,1)^D$, the inverse transformations $T_d^{-1}(V_d)$ of T_d, $d = 1, \dots, D$, generate a random variable $X = (X_1, \dots, X_D)'$ which is distributed according to F.

Rosenblatt proves the statement in his very article [56]. ∎

We can take advantage of the PIT by considering the multivariate cdf F in Theorem 2.12 as copula C. With $x = u \in [0,1]^D$ as a realization of $U \sim C$, the system (2.9) changes to

$$\begin{aligned} T_1(u_1) &= C_1(u_1) = C(u_1, 1, \dots, 1) = u_1 \quad , \\ T_2(u_2) &= C_{2|1}(u_2|u_1) = C(u_2, 1, \dots, 1|U_1 = u_1) \quad , \\ &\vdots \\ T_D(u_D) &= C_{D|1,\dots,D-1}(u_D|u_1,\dots,u_{D-1}) = \frac{\dfrac{\partial^{D-1}C(u_1,\dots,u_D)}{\partial u_1,\dots\partial u_{D-1}}}{\dfrac{\partial^{D-1}C(u_1,\dots,u_{D-1},1)}{\partial u_1\dots\partial u_{D-1}}} \quad . \end{aligned}$$

Using the second part of Theorem 2.12, one receives the recipe for generating copula-distributed random variables - Algorithm A - 2.4.

Conditional sampling turns out to be elegant in mathematical sense. However, this approach exhibits two major difficulties which are uniquely dependent on the structure of the concerned copula. First, it requires the computation of higher order conditional derivatives. Secondly, we have to invert this differentiation result.

[6] It is one of the mostly treated copula simulation methods, also compare [8], [22] or [63].
[7] See also [29].

The inversion of the derivative function cannot always be found analytically. In this case, the expression $u_d = C^{-1}_{d|1,\dots,d-1}(z_d|u_1,\dots,u_{d-1})$ can be received by numerical root finding of the equation $z_d = C_{d|1,\dots,d-1}(u_d|u_1,\dots,u_{d-1})$, see [12], Section 6.3. Simulation procedures are developed in [22], [26], [28] or [47] for instance[8].

A - 2.4 Algorithm (Conditional Sampling $U \sim C$)

Step 1

- draw $z_1,\dots,z_D$ independent and uniformly distributed random variables
- set $u_1 = z_1$

Step 2 - for $d = 2,\dots,D$:

- compute $u_d = C^{-1}_{d|1,\dots,d-1}(z_d|u_1,\dots,u_{d-1})$

Step 3

- receive $u = (u_1,\dots,u_D)' \sim C$

2.13 Proposition (Conditional Clayton Copula)

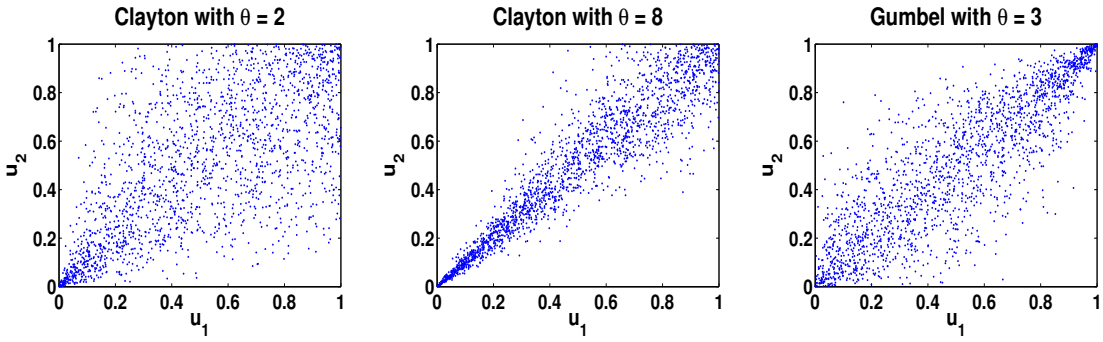
1. For $d = 2,\dots,D$, the conditional distribution functions $C^{Cl}_{d|1,\dots,d-1}$ for the Clayton copula (2.7) hold

$$C^{Cl}_{d|1,\dots,d-1}(u_d|u_1,\dots,u_{d-1};\theta) = \left(1 + \frac{u_d^{-\theta} - 1}{\sum_{i=1}^{d-1} u_i^{-\theta} - (d-1) + 1}\right)^{-\frac{1}{\theta}-(d-1)} . \tag{2.10}$$

2. For $d = 2,\dots,D$, let $z_d \sim U(0,1)$, $u_1,\dots,u_{d-1}$ be given. The inverse function of (2.10) is

$$\begin{aligned} u_d &= C^{-1}_{d|1,\dots,d-1}(z_d|u_1,\dots,u_{d-1};\theta) \\ &= \left(\left(z_d^{-\frac{\theta}{1+\theta(d-1)}} - 1\right)\cdot\left(\sum_{i=1}^{d-1} u_i^{-\theta} - (d-1) + 1\right) + 1\right)^{-\frac{1}{\theta}} . \end{aligned} \tag{2.11}$$

A proof can be found in Appendix D.1.3 and D.1.4. ∎



F - 2.5 Figure: 2000 random samples - Clayton and Gumbel copula

[8] For Archimedean copulas with an arbitrary generator ϕ, Cherubini et al. give a nice derivation for the expression $z_d = C_{d|1,\dots,d-1}(u_d|u_1,\dots,u_{d-1})$, compare [12], Theorem 6.1.

A - 2.5 Algorithm (Clayton Copula Sampling)

Given

- $\theta > 0$
- $C_\theta^{Cl}(u_1, \ldots, u_D) = \left(\sum_{d=1}^{D} \left(u_d^{-\theta} - 1 \right) + 1 \right)^{-\frac{1}{\theta}}$

Step 1

- draw i.i.d. $z_1, \ldots, z_D \sim U(0,1)$
- set $z_1 = u_1$

Step 2 - for $d = 2, \ldots, D$:

- compute $u_d = \left(\left(z_d^{-\frac{\theta}{1+\theta(d-1)}} - 1 \right) \cdot \left(\sum_{i=1}^{d-1} u_i^{-\theta} - (d-1) + 1 \right) + 1 \right)^{-\frac{1}{\theta}}$

Step 3

- receive $u = (u_1, \ldots, u_D)' \sim C_\theta^{Cl}$

2.14 Example (Gumbel Copula)
For the Gumbel copula C_θ^{Gu} (2.8), the conditional derivatives $C_{d|1,\ldots,d-1}(u_d|u_1, \ldots, u_{d-1})$ are **not** invertible. For a sampling algorithm, the reader is referred to [12], Subsection 6.3.2.

As one has to compute higher order conditional derivatives and an inversion of these differentiation results in each step, conditional sampling is computationally intensive. This is particularly relevant if the expressions $C^{-1}_{d|1,\ldots,d-1}$ do not have closed-form solutions and derivatives are unmanageable.

Several interesting ideas exists to overcome these difficulties. Whelan [64] presents an approach to simulate Archimedean copulas by partitioning and scaling one-dimensional draws to the right multivariate cdf. Further, the theory of pair copula functions avoids explicit multidimensional structures, e.g., compare [1].

2.3.3 Simulation of Conditional Copulas

Next, we combine the ideas of the previous Subsections 2.3.1 Transformation Method and 2.3.2 Conditional Sampling. The following approach allows to simulate conditional copulas (i.e., derivatives of copulas) for certain classes of multivariate distributions.

Suppose that the cdf F has a differentiable copula C. W.l.o.g. concerning index ordering, a derivative of a copula function can be specified as conditional distribution by

$$\begin{aligned}
\frac{\partial^d}{\partial u_1 \ldots \partial u_d} C(u_1, \ldots, u_D) &= C(u_{d+1}, \ldots, u_D | U_1 = u_1, \ldots, U_d = u_d) \\
&= \mathbb{P}\left[U_{d+1} \leq u_{d+1}, \ldots, U_D \leq u_D | U_1 = u_1, \ldots, U_d = u_d \right] \\
&= \mathbb{P}\Big[F_{d+1}^{-1}(U_{d+1}) \leq F_{d+1}^{-1}(u_{d+1}), \ldots, F_D^{-1}(U_D) \leq F_D^{-1}(u_D) | \\
&\qquad\qquad F_1^{-1}(U_1) = F_1^{-1}(u_1), \ldots, F_d^{-1}(U_d) = F_d^{-1}(u_d) \Big] \\
&= \mathbb{P}\left[X_{d+1} \leq x_{d+1}, \ldots, X_D \leq x_D | X_1 = x_1, \ldots, X_d = x_d \right] \\
&= F\left(x_{d+1}, \ldots, x_D | X_1 = x_1, \ldots, X_d = x_d \right) \quad .
\end{aligned} \tag{2.12}$$

Restricting this conditional multivariate cdf F to the classes of Gaussian and t-distributions, we take advantage of the following theorems. The subsequent results are collected from [23], [38], [63], and their respective references.

2.15 Theorem
Let $X \sim \Phi_{\mu,\Sigma}$ be a multivariate normal distribution with mean μ and covariance matrix Σ. We denote the *usual partition* of X, μ and Σ as

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \ , \ \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \ , \text{ and } \ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad \text{with}$$

- $X_1,\ \mu_1 \in \mathbb{R}^{d\times 1}$ and $\Sigma_{11} \in \mathbb{R}^{d\times d}$ for $0 < d < D$,
- $X_2,\ \mu_2 \in \mathbb{R}^{D-d\times 1}$ and $\Sigma_{22} \in \mathbb{R}^{D-d\times D-d}$,
- $\Sigma_{12} \in \mathbb{R}^{d\times D-d}$ and $\Sigma_{21} \in \mathbb{R}^{D-d\times d}$.

Then, the conditional random variable X_2 given X_1 has multivariate normal distribution $\Phi_{\mu_{2|1},\Sigma_{2|1}}$ with mean $\mu_{2|1}$ and covariance matrix $\Sigma_{2|1}$ computed as

$$\mu_{2|1} = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1) \quad \text{and} \quad \Sigma_{2|1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \quad .$$

The proof follows directly from Theorem 2.16 in [23], compare also Corollary 4.11 in [63]. ∎

2.16 Theorem
Let $X \sim t_{\nu,\mu,\Sigma}$ be a multivariate t-distribution with degrees of freedom ν, mean μ, and covariance matrix Σ. Let the usual partition of X, μ and Σ be defined as in Theorem 2.15.

Then, the conditional random variable X_2 given X_1 has multivariate t-distribution $t_{\nu_{2|1},\mu_{2|1},\Sigma_{2|1}}$ with degrees of freedom $\nu_{2|1}$, mean $\mu_{2|1}$, and covariance matrix $\Sigma_{2|1}$ computed as

$$\begin{aligned} \nu_{2|1} &= \frac{\nu}{\nu + d - 2} \quad , \\ \mu_{2|1} &= \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1) \quad \text{, and} \\ \Sigma_{2|1} &= \left(1 + (X_2 - \mu_2)' \left(\nu\Sigma_{22}\right)^{-1} (X_2 - \mu_2)\right) \left(\Sigma_{11} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right) \quad . \end{aligned}$$

Proof and derivation are described in [38], see particularly Section 3.2. ∎

Thus, if the unconditional distributions are given as multivariate Gaussian or t, their conditional distributions can be computed as shown in Theorems 2.15 and 2.16. Applying the general transformation method of Subsection 2.3.1 to these conditional cdfs provides the recipe to simulate conditional Gaussian or t-copulas due to Equation (2.12). We summarize this result in the following algorithms.

A - 2.6 Algorithm (Conditional Gaussian Copula Sampling)

Given

- $X \sim \Phi_P$, $U \sim C_P^{Ga}$
- $X_1 = x_1$ and its usual partition as in Theorem 2.15

Step 1

- draw $x_2 \sim \Phi_{\mu_{2|1},\Sigma_{2|1}}$

Step 2 - for relevant d of vector x_2:

- compute $u_d = \Phi(x_d)$

Step 3

- receive $u_2 \sim \frac{\partial^d}{\partial u_1 \ldots \partial u_d} C_P^{Ga}$

A - 2.7 Algorithm (Conditional t-Copula Sampling)

Given

- $X \sim t_{\nu,P}$, $C^t_{\nu,P}$
- $X_1 = x_1$ and its usual partition as in Theorem 2.16

Step 1

- draw $x_2 \sim t_{\nu_{2|1},\mu_{2|1},\Sigma_{2|1}}$

Step 2 - for relevant d of vector x_2:

- compute $u_d = t_\nu(x_d)$

Step 3

- receive $u_2 \sim \frac{\partial^d}{\partial u_1 \ldots \partial u_d} C^t_{\nu,P}$

2.17 Remark
The above given form for the usual partition is chosen due to readability and is interchangeable. The result holds in general, i.e., for all partitions.

Algorithms A - 2.6 and A - 2.7 start with the standard multivariate Gaussian and t-distribution, respectively. For both algorithms any general Gaussian or t-distribution is applicable.

3 Inference for Copulas

This chapter deals with the statistical inference for copulas, i.e., the estimation of parameters for copula models on the basis of historical data. We repeat miscellaneous approaches of likelihood calibration and present algorithms for the canonical maximum likelihood fitting of the Gaussian, t-, and Clayton copula. For these copulas, goodness-of-fit tests are illustrated.

3.1 Calibration

The preferred choice of copula parameter calibration to historical data is the maximum likelihood estimation (MLE)[1]. Besides, empirical estimation procedures and combinations of MLE and empirical approaches are applied in practice. Other theories of statistical inference fail to work, see [12], Section 5.1.

We restrict our analysis to parameter estimation and leave out the approximation of standard errors and confidence intervals. This is treated in [15], for instance. We mainly adopt our notation to [37], Chapter 10.

3.1 Assumption
Let $X = (X_1, \ldots, X_D)'$ be a $\mathbb{R}^D$-valued random variable. Suppose that X has multivariate cdf F and let $x_{\cdot,n} = (x_{1,n}, \ldots, x_{D,n})' \in \mathbb{R}^D$, $n = 1, \ldots, N$ with $N \in \mathbb{N}$, be **independent** realizations[2] of X. We denote

$$\Upsilon = \begin{pmatrix} x_{1,1} & \ldots & x_{1,N} \\ \vdots & \ddots & \vdots \\ x_{D,1} & \ldots & x_{D,N} \end{pmatrix}$$

as the *sample* or *observation matrix* of X.

[1] The idea of maximum likelihood estimation goes back to the explanations of Sir Ronald Aylmer Fisher at the beginning of the 1920s.
[2] That means $x_{\cdot,n}$ is independent of $x_{\cdot,m}$ for all $n \neq m$.

Equation (2.1) states that $F(x_1,\ldots,x_D) = C(F_1(x_1),\ldots,F_D(x_D))$ holds. As seen in Section 2.2, the copula C as well as the margins F_d depend on parameters defined as

- θ_c for the copula and
- θ_d for each F_d, $d = 1,\ldots,D$.

We set $\theta = (\theta_1,\ldots,\theta_D,\theta_c) \in \Theta$ in which $\Theta = (\Theta_1 \times \ldots \times \Theta_D \times \Theta_c)$ characterizes the *admissible parameter space*[3]. Thus, Equation (2.1) can be rewritten subject to the parameter θ as

$$F(x_1,\ldots,x_D,\theta) = C(F_1(x_1,\theta_1),\ldots,F_D(x_d,\theta_D),\theta_c) \quad . \tag{3.1}$$

We recall the canonical representation of the copula density function c of X, see Equation (2.3), and revise this expression in dependence of the parameters θ and $x = (x_1,\ldots,x_D)'$ instead of $u = (u_1,\ldots,u_D)'$ to

$$f(x_1,\ldots,x_D,\theta) = c\Big(F_1(x_1,\theta_1),\ldots,F_D(x_D,\theta_D),\theta_c\Big) \cdot \prod_{d=1}^{D} f_d(x_d,\theta_d) \quad . \tag{3.2}$$

3.1.1 Exact MLE

Roughly speaking, the principle of maximum likelihood estimation[4] relies on the maximization of the concerned multivariate density function with respect to the density's incorporated parameters given a (adequately large) number of realizations. In other words, MLE maximizes a multivariate density function f of Equation (3.2) with respect to its parameter θ provided a sample Υ is known.

Due to the independence assumption for the N realizations of X, see Assumptions 3.1, the likelihood function L of (3.2) is

$$L(\theta) = \prod_{n=1}^{N} f(x_{1,n},\ldots,x_{d,n},\theta) \quad . \tag{3.3}$$

The maximum likelihood estimator is now the value θ^* which maximizes (3.3). As it is strictly increasing, a logarithmic transformation of (3.3) will provide the same estimator θ^*. Thus, we define the log-likelihood function as

$$l(\theta) = \ln L(\theta) = \sum_{n=1}^{N} \ln f(x_{1,n},\ldots,x_{d,n},\theta) \tag{3.4}$$

and have simultaneously

$$\theta^* = \arg\max_{\theta\in\Theta} l(\theta) = \arg\max_{\theta\in\Theta} L(\theta)$$

for the required maximizer θ^*. This exact maximum (log-)likelihood estimator θ^* is obtained by solving the system[5]

$$\left(\frac{\partial l}{\partial\theta_1},\ldots,\frac{\partial l}{\partial\theta_D},\frac{\partial l}{\partial\theta_c}\right)' = \mathbf{0} \quad . \tag{3.5}$$

Due to relation (3.2), Equation (3.4) can be modified to

$$l(\theta) = \sum_{n=1}^{N} \ln\Big[c\Big(F_1(x_{1,n},\theta_1),\ldots,F_D(x_{D,n},\theta_D),\theta_c\Big)\Big] + \sum_{n=1}^{N}\sum_{d=1}^{D} \ln\left[f_d(x_{d,n},\theta_d)\right] \tag{3.6}$$

in which c is the copula density and f_d are the marginal densities, compare [12], Section 5.2. Generally, the optimization of (3.6), i.e., solving system (3.5), has to be carried out numerically.

[3] Parameters may be vectors. Dimensions of θ_c and θ_d, $d = 1,\ldots,D$, naturally determine the dimension of Θ.
[4] For an introduction to the MLE approach a variety of textbooks is available, see [21] for instance.
[5] Consider that the parameters θ_c and θ_d, $d = 1,\ldots,D$, may be vectors and the above derivatives change as they have to be taken with respect to each component of each parameter vector.

3.1.2 IFM - Inference Functions for Margins

The IFM are found in many compositions on copula calibration, e.g., compare [12], [15] or [37]. The motive is that computational optimization of (3.6) can be very intensive. The idea of the IFM method is to split maximization into two parts. This means we separately consider

$$l_d(\theta_d) = \sum_{n=1}^{N} \ln\left[f_d(x_{d,n}, \theta_d)\right] \quad \text{for } d = 1, \dots, D \text{ and} \tag{3.7}$$

$$l_c(\theta_c) = \sum_{n=1}^{N} \ln\Big[c\Big(F_1(x_{1,n}, \theta_1), \dots, F_D(x_{D,n}, \theta_D), \theta_c\Big)\Big] \quad . \tag{3.8}$$

In a first step, we independently optimize Equation (3.7) with respect to θ_d and receive

$$\hat{\theta}_d = \underset{\theta_d \in \Theta_d}{\arg\max}\ l_d(\theta_d)$$

for each $d = 1, \dots, D$. Next, values $(\hat{\theta}_1, \dots, \hat{\theta}_D)$ are used for maximizing (3.8), i.e., we search for

$$\hat{\theta}_c = \underset{\theta_c \in \Theta_c}{\arg\max} \sum_{n=1}^{N} \ln\Big[c\Big(F_1(x_{1,n}, \hat{\theta}_1), \dots, F_D(x_{D,n}, \hat{\theta}_D), \theta_c\Big)\Big] \quad .$$

Generally, the IFM estimator $\theta_{IFM} = (\hat{\theta}_1, \dots, \hat{\theta}_D, \hat{\theta}_c)$ is not equal to the MLE estimator θ^*. However, the IFM method is computationally less intensive than a direct optimization procedure for (3.6) and θ_{IFM} offers a sound initial value for the exact MLE.

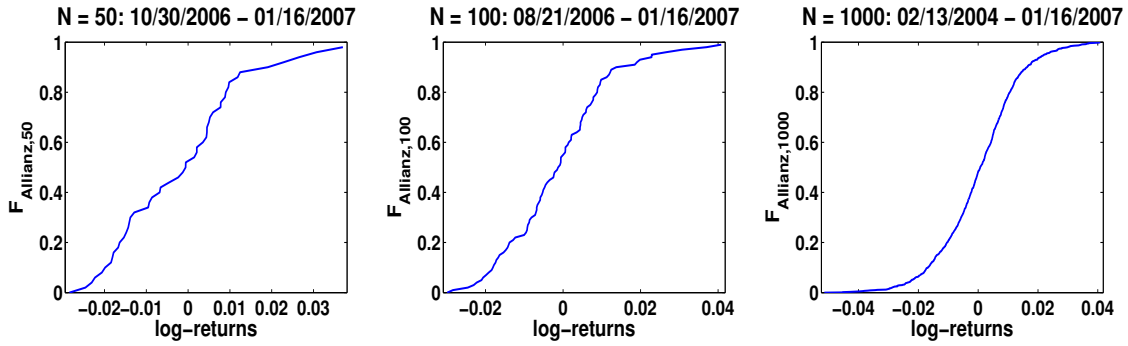
3.1.3 CML - Canonical Maximum Likelihood

The canonical maximum likelihood or pseudo-(log-)likelihood approach uses a combination of MLE and empirical estimation to calibrate copula parameters.

3.2 Definition (compare [15], Section 1.4)
Given an observation matrix Υ the *empirical cumulative distribution function (ecdf)*[6] for the dth margin is defined as

$$F_{dN}(x) = \frac{1}{N+1} \sum_{n=1}^{N} \mathbf{1}_{\{x_{d,n} \le x\}}(x) \quad . \tag{3.9}$$

When working with empirical distributions, a direct condition is that the number of realizations N must be large enough to allow for an adequate smoothness of the ecdf. Figure F - 3.1 shows the different behavior of empirical cdfs in dependence of the number of realizations N for Allianz AG. A sample size of $N = 1000$ guarantees a sufficient smoothness, for instance.



F - 3.1 Figure: empirical distribution functions for Allianz AG

[6] Sometimes one finds the expression *rescaled* ecdf.

Now, the CML employs the empirical distribution functions F_{dN} instead of the analytical margins F_d, $d = 1, \dots, D$. Thereby, it bypasses the parameter estimation for each margin as there exists none for ecdfs. The optimization of (3.6) is reduced to find the copula parameter θ_{CML} which is equal to step 2 of the IFM approach, i.e.,

$$\theta_{CML} = \underset{\theta_c \in \Theta_c}{\arg\max} \sum_{n=1}^{N} \ln \Big[c\Big(F_{1N}(x_{1,n}), \dots, F_{DN}(x_{D,n}), \theta_c \Big) \Big] \quad .$$

Thus, CML significantly diminishes computing time. Under Assumptions 3.1, θ_{CML} is a consistent estimator[7], see [27]. For an detailed analysis of the estimator θ_{CML}, the reader is referred to [15].

3.3 Notation

The observation matrix Υ transformed by the marginal ecdfs F_{dN} is denoted as *pseudo-observation matrix* Ψ and given as

$$\Psi = \begin{pmatrix} F_{1N}(x_{1,1}) & \dots & F_{1N}(x_{1,N}) \\ \vdots & \ddots & \vdots \\ F_{DN}(x_{D,1}) & \dots & F_{DN}(x_{D,N}) \end{pmatrix} \quad .$$

3.1.4 Non-Parametric Estimation

In literature, we also find non-parametric approaches to fit copulas to historical data. We touch on two methods which are known as *empirical* and *kernel copula approach*, respectively. In [16] and [17], Deheuvels introduces the concept of empirical copulas as a sort empirical multivariate distribution of the ranks of realizations. He shows that the empirical copula uniformly converges to the underlying copula. Another non-parametric estimation procedure is the kernel copula approach presented in [25]. The authors develop a smooth differentiable rebuilding of the copula via kernel functionals. Here, the advantage is that copula and margins and hence their parameters are not specified a priori.

3.2 Examples of CML Calibration

In practice, the pseudo-loglikelihood estimation in combination with the Gaussian or t-copula is frequently used for financial applications. Its costs for implementation, computing time as well as mathematical complexity for calibration are manageable. We assume that an observation matrix $\Upsilon \in \mathbb{R}^{D \times N}$, corresponding empirical cdfs F_{dN}, $d = 1, \dots, D$, and consequently a pseudo-observation matrix Ψ are given.

3.2.1 Fitting the Gaussian Copula

As argued in Subsection 3.1.3, for the CML approach the likelihood function for the Gaussian copula can be restricted to

$$\begin{aligned} l_{CML}^{Ga}(P) &= \sum_{n=1}^{N} \ln \left[c^{Ga}\left(F_{1N}(x_{1,n}), \dots, F_{DN}(x_{D,n}), P \right) \right] \\ &= \sum_{n=1}^{N} \ln \left[c^{Ga}\left(\Psi_{1,n}, \dots, \Psi_{D,n}, P \right) \right] \end{aligned} \tag{3.10}$$

in which c^{Ga} denotes the density of the Gaussian copula, see Equation (2.4).

[7] Let θ_N be the estimator for a sample of size N and θ^* the exact estimator. The estimator θ_N is *consistent* if it converges in sense of $\lim_{N \to \infty} \mathbb{P}[|\theta_N - \theta^*| \geq \epsilon] = 0 \quad \forall\, \epsilon \geq 0$.

The "only" parameter to calibrate is the correlation matrix P. With (2.4) inserted in (3.10) and setting $\xi_n = (\Phi^{-1}(\Psi_{1,n}), \ldots, \Phi^{-1}(\Psi_{D,n}))'$, we receive

$$P_{CML}^{Ga} = \underset{P \in \Theta}{\arg\max} \; -\frac{N}{2} \ln |P| - \frac{1}{2} \sum_{n=1}^{N} \xi_n' (P^{-1} - \mathbb{I}_D) \xi_n \tag{3.11}$$

in which Θ is the set of all admissible[8] correlation matrices. The analytical solution of (3.11) is

$$P_{CML}^{Ga} = \frac{1}{N} \sum_{n=1}^{N} \xi_n' \xi_n \quad ,$$

see [46]. The effort for this optimization significantly grows with the number of dimensions. As the parameter space Θ is restricted to correlation matrices and the exact solution P_{CML}^{Ga} is generally[9] not element of Θ, maximization becomes even more complex.

To overcome this issue, Cherubini et al. [12], Subsection 5.4.1, propose to compute the correlation matrix $\tilde{P}_{CML}^{Ga}$ of the transformed observation matrix $\Phi^{-1}(\Psi)$[10]. A different approach is to calibrate the correlation matrix of the Gaussian copula by Spearman's rank correlation coefficient, refer to [49], Example 5.53. Here, we adopt the strategy of [12] to obtain the following calibration procedure.

A - 3.1 Algorithm (CML Calibration - Gaussian Copula)

Given

- observation matrix Υ

Step 1 - for $n = 1, \ldots, N$ and $d = 1, \ldots, D$:

- compute ecdf F_{dN} from Υ
- transform $F_{dN}(x_{d,n}) = \Psi_{d,n}$

Step 2 - for $n = 1, \ldots, N$:

- transform $\xi_{\cdot,n} = \left(\Phi^{-1}(\Psi_{1,n}), \ldots, \Phi^{-1}(\Psi_{D,n})\right)'$

Step 3 - for $i, j = 1, \ldots, D$:

- compute correlation $\rho_{i,j} = \rho\left(\xi_{i,\cdot}, \xi_{j,\cdot}\right)$

Step 4

- receive $\tilde{P}_{CML}^{Ga} = \begin{pmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,D} \\ \rho_{2,1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_{D-1,D} \\ \rho_{D,1} & \cdots & \rho_{D,D-1} & 1 \end{pmatrix}$

[8] I.e., positive definite and symmetric.

[9] Embrechts et al. [49] argue that P_{CML}^{Ga} is "in practice likely to be close to being a correlation matrix", p. 235.

[10] The idea is to transform the original data Υ by means of $\Phi^{-1}(F_{dN}(\Upsilon_{d,n}))$ for $d = 1, \ldots, D$ and $n = 1, \ldots, N$ to receive Gaussian data.

3.2.2 Fitting the t-Copula

Due to Equation (2.5) the loglikelihood function for the t-copula holds

$$\begin{aligned} l^t_{CML}(\nu, P) &= \sum_{n=1}^{N} \ln\left[c^t\left(\Psi_{1,n}, \ldots, \Psi_{D,n}, \nu, P\right)\right] \\ &= -\frac{N}{2}\ln|P| + N\left(\ln\left[\Gamma\left(\frac{\nu+D}{2}\right) - \Gamma\left(\frac{\nu}{2}\right)\right]\right) + N\cdot D\left(\ln\left[\Gamma\left(\frac{\nu}{2}\right) - \Gamma\left(\frac{\nu+1}{2}\right)\right]\right) \\ &\quad -\frac{\nu+D}{2}\sum_{n=1}^{N}\ln\left[1+\frac{1}{\nu}\xi' P^{-1}\xi\right] + \frac{\nu+1}{2}\sum_{n=1}^{N}\sum_{d=1}^{D}\ln\left[1+\frac{\xi^2_{d,n}}{\nu}\right] \end{aligned} \tag{3.12}$$

with pseudo-sample matrix Ψ, $\xi_n = \left(t_\nu^{-1}(\Psi_{1,n}), \ldots, t_\nu^{-1}(\Psi_{D,n})\right)'$ and $\xi_{d,n} = t_\nu^{-1}(\Psi_{d,n})$. We have to find the solution

$$(\nu^t_{CML}, P^t_{CML}) = \underset{(\nu,P)\in\Theta}{\arg\max}\; l^t_{CML}(\nu, P)$$

in which Θ denotes the admissible parameter space of degrees of freedom ν and correlation matrices P. Again, for low dimensions this optimization might be feasible. However, for higher dimensional orders a maximization of (3.12) is computationally very intensive. To deal with this matter, several authors split up the optimization procedure into two steps. The first step is to estimate the correlation matrix by use of rank correlation coefficients. In a second step, the remaining parameter ν is determined numerically. We outline this procedure below. For a detailed description the reader is referred to [22], [34], [42], or [43], for instance.

3.4 Definition (Kendall's τ - Kendall's rank correlation coefficient)

1. *Kendall's* τ for a multivariate random variable X is defined as

   $$\rho_\tau(X) = \operatorname{cov}(\operatorname{sign}(X - \tilde{X}))$$

   in which $\tilde{X}$ is an independent copy of X, see [49], p. 207.

2. The *standard estimator* of Kendall's τ for a sample matrix Υ is

   $$R_\tau(i,j) = \frac{2}{N(N-1)}\left(\sum_{n=1}^{N}\sum_{m>n}^{N}\operatorname{sign}(x_{i,n} - x_{i,m})(x_{j,n} - x_{j,m})\right) \quad \text{for } i,\ j = 1, \ldots, N \quad .$$

An overview on rank correlation is given in [49], Section 5.2.2. In particular, Kendall's τ

- is a measure of concordance[11] in sense of probability,
- is positive definite and symmetric (as covariance matrix),
- is invariant under strictly increasing transformations, and
- has values in $[-1, 1]$. X and $\tilde{X}$ are countermonotonic for the value $\rho_\tau = -1$ and comonotonic for the value $\rho_\tau = 1$.

3.5 Lemma

1. Let $U = (U_1, \ldots, U_D)' \in [0,1]^D$. Then, the correlation matrix $P_{i,j} = \rho_{i,j}$ of the t-copula $C^t_{\nu,P}$ can be determined by $\rho_\tau(U_i, U_j) = \frac{2}{\pi}\arcsin(\rho_{i,j})$ for $i,\ j = 1, \ldots, N$, as

   $$P_{i,j} = \rho_{i,j} = \sin\left(\frac{\pi}{2}\rho_\tau(U_i, U_j)\right) \quad .$$

2. For a pseudo-sample matrix Ψ the correlation matrix $\tilde{P}_{i,j} = r_{i,j}$ of the t-copula $C^t_{\nu,\tilde{P}}$ can be estimated by $R_\tau(U_i, U_j) = \frac{2}{\pi}\arcsin(r_{i,j})$ for $i,\ j = 1, \ldots, N$ as

   $$\tilde{P}_{i,j} = r_{i,j} = \sin\left(\frac{\pi}{2}R_\tau(U_i, U_j)\right) \quad .$$

[11] (x_1, y_1) and (x_2, y_2) are *concordant* in $\mathbb{R}^2$ if $(x_1 - y_1)(x_2 - y_2) > 0$ and *discordant* if $(x_1 - y_1)(x_2 - y_2) < 0$.

Outline of proof:

1. Proposition 5.37 in [49] shows a general result how to specify a correlation matrix for elliptical distributions. Apply this result to the t-copula. Then, the invariance property Kendall's τ implies the above given result, compare [49], Example 5.54.
2. Follows from 1) and the estimator R_τ for Kendall's τ. ∎

Lemma 3.5 gives the recipe to receive the correlation matrix $\tilde{P}$ for the t-copula. The estimation for the remaining parameter (degrees of freedom ν) has to be carried out numerically. CML optimization for the t-copula is summarized in following algorithm.

A - 3.2 Algorithm (CML Calibration - t-Copula)

Given

- observation matrix Υ

Step 1 - for $n = 1, \ldots, N$ and $d = 1, \ldots, D$:

- compute ecdf F_{dN} from Υ
- transform $\Psi_{d,n} = F_{dN}(x_{d,n})$

Step 2 - for $i,\, j = 1, \ldots, D$:

- compute $R_\tau(i,j) = \dfrac{2}{N(N-1)} \left(\sum\limits_{n=1}^{N} \sum\limits_{m>n}^{N} \operatorname{sign}(\Psi_{i,n} - \Psi_{i,m})(\Psi_{j,n} - \Psi_{j,m}) \right)$
- transform $\tilde{P}^t_{i,j} = \sin\left(\frac{\pi}{2} R_\tau(i,j)\right)$

Step 3

- receive correlation matrix $\tilde{P}^t_{CML}$
- compute $\tilde{\nu}_{CML} = \underset{\nu \in \Theta}{\arg\max} \sum\limits_{n=1}^{N} \ln \left[c^t\left(\Psi_{1,n}, \ldots, \Psi_{D,n}, \nu, \tilde{P} \right) \right]$
- receive CML estimates $(\tilde{\nu}_{CML}, \tilde{P}^t_{CML})$

3.6 Remark
The transformation in Step 2 does not always provide a positive definite matrix $\tilde{P}$. In this case, algorithms to generate nearest positive definite matrices can be applied such as an eigenvalue method for instance, see [49], Example 5.54 and Algorithm 5.55.

3.2.3 Fitting the Clayton Copula

The D-dimensional Clayton copula density is given by

$$c^{Cl}(u_1, \ldots, u_D, \theta) = \prod_{d=1}^{D} \left[(1 + (d-1)\theta) \cdot u_d^{-\theta-1} \right] \cdot \left(\sum_{d=1}^{D} u_d^{-\theta} - D + 1 \right)^{-\frac{1}{\theta} - D} . \tag{3.13}$$

A derivation is received by induction and can be found in Appendix D.1.2. For a pseudo-sample matrix Ψ, the loglikelihood function hence is

$$\begin{aligned} l^{Cl}_{CML}(\theta) &= \sum_{n=1}^{N} \ln c^{Cl}(\Psi_{1,n}, \ldots, \Psi_{D,n}, \theta) \qquad (3.14) \\ &= N \cdot \sum_{d=1}^{D} \ln\left[1 + (d-1)\theta\right] - (\theta+1) \sum_{n=1}^{N} \sum_{d=1}^{D} \ln\left[\Psi_{d,n}\right] - \left(\frac{1}{\theta} + D\right) \sum_{n=1}^{N} \ln\left[\sum_{d=1}^{D} \Psi_{d,n}^{-\theta} - D + 1 \right] . \end{aligned}$$

Optimization of (3.14) is not a CML estimation in the classical sense as there exists no implicit margins for the Clayton copula C^{Cl}, see Equation (2.7). Formula (3.14) appears to be as complex as the likelihood function for the t-copula. However, the one-dimensional search for θ eases maximization and is comparable to Step 3 of Algorithm A - 3.2. Again, numerical procedures have to be applied.

3.3 Goodness-of-Fit Test

When calibration has been performed, one needs to know whether the chosen copula fits to the given data. For this issue, the probability integral transformation (PIT), see [56] or Theorem 2.12, is a helpful tool. We recall that part one of this very theorem states that the transformations

$$\begin{aligned} T_1(u_1) &= C_1(u_1) \quad , \\ T_2(u_2) &= C_{2|1}(u_2|u_1) \quad , \\ &\vdots \\ T_D(u_D) &= C_{D|1,\ldots,D-1}(u_D|u_1,\ldots,u_{D-1}) \end{aligned}$$

produce a vector $(v_1,\ldots,v_D)' = (T_1(u_1),\ldots,T_D(u_D))'$ which is independent and standard uniformly distributed. Thus, $\Phi^{-1}(v_1),\ldots,\Phi^{-1}(v_D)$ are i.i.d. $N(0,1)$-distributed and

$$S = \sum_{d=1}^{D} (\Phi^{-1}(v_d))^2$$

is χ^2-distributed with D degrees of freedom. Consequently, we can apply a χ^2-test to S for the given data. We summarize this procedure in

A - 3.3 Algorithm (General Goodness-of-Fit Test)

Given

- pseudo-observation matrix Ψ
- copula parameters θ

Step 1 - for $n = 1,\ldots,N$:

- set $v_{1,n} = T_1(\Psi_{1,n}) = C_1(\Psi_{1,n}) = \Psi_{1,n}$
- for $d = 2,\ldots,D$:
  - compute $v_{d,n} = T_D(\Psi_{d,n}) = C_{d|1,\ldots,d-1}(\Psi_{d,n}|\Psi_{1,n},\ldots,\Psi_{d-1,n})$
- compute $S_n = \sum_{d=1}^{D} (\Phi^{-1}(v_{d,n}))^2$

Step 2

- perform χ^2-test to $S = (S_1,\ldots,S_N)'$

Computational efforts are highly dependent on the first step of the algorithm. We have to calculate conditional distributions, i.e., determining higher order derivatives, and evaluate these for each data sample n, $n = 1,\ldots,N$.

Tappe [63] shows how to perform PIT's for Gaussian and t-copula. As the derivation requires a lot of computation, we limit to the main result and borrow Algorithms 9 and 10 from [63].

Proposition 4.3 in [63] illustrates that the Gaussian copula with correlation matrix P has the dth conditional margin representation

$$C_{d|1,\ldots,d-1}(u_d|u_1,\ldots,u_{d-1}) = \Phi_{\mu^d,\sigma^d}\left(\Phi^{-1}(u_d)\right) \tag{3.15}$$

in which

- Φ_{μ^d,σ^d} is the univariate normal distribution with mean μ^d and variance σ^d,
- $\mu^d = P_{21}^d \cdot \left(P_{11}^d\right)^{-1} \cdot \left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_{d-1})\right)'$, and
- $\sigma^d = P_{22}^d - P_{21}^d \cdot \left(P_{11}^d\right)^{-1} \cdot P_{12}$

with the notation

$$P^d = \begin{pmatrix} P_{11}^d & P_{12}^d \\ P_{21}^d & P_{22}^d \end{pmatrix} \in \begin{pmatrix} \mathbb{R}^{(d-1)\times(d-1)} & \mathbb{R}^{(d-1)\times 1} \\ \mathbb{R}^{1\times(d-1)} & \mathbb{R} \end{pmatrix} .$$

P^d is the correlation for the d-margin copula $C_d^{Ga}(u_1, \ldots, u_d, 1, \ldots, 1)$. The above partition is called the d *by* $d-1$ *partition of* P^d.

For a detailed derivation, the reader is referred to [63] and the groundwork of Fang et al. [23]. Equation (3.15) results in the following[12]

A - 3.4 Algorithm (Goodness-of-Fit Test Gaussian Copula)

Step 1 - for $n = 1, \ldots, N$:

- set $v_{1,n} = \Psi_{1,n}$

Step 2 - for $n = 1, \ldots, N$ and $d = 2, \ldots, D$:

- compute $\mu^d = P_{21}^d \cdot \left(P_{11}^d\right)^{-1} \cdot \left(\Phi^{-1}(\Psi_{1,n}), \ldots, \Phi^{-1}(\Psi_{d-1,n})\right)'$
- compute $\sigma^d = P_{22}^d - P_{21}^d \cdot \left(P_{11}^d\right)^{-1} \cdot P_{12}$
- transform $v_{d,n} = \Phi_{\mu^d,\sigma^d}\left(\Phi^{-1}(\Psi_{d,n})\right)$

Step 3 - for $n = 1, \ldots, N$:

- compute $S_n = \sum_{d=1}^{D} (\Phi^{-1}(v_{d,n}))^2$

Step 4

- perform χ^2-test to $S = (S_1, \ldots, S_N)'$

Several tedious calculations, see [63], evidence that the dth margin $C_{d|1,\ldots,d-1}$ given $U_1, \ldots, U_{d-1}$ for the t-copula $C_{\nu,P}^t$ with ν degrees of freedom, correlation matrix P, and its d by $d-1$ partition P^d can be represented as

$$\begin{aligned} & C_{d|1,\ldots,d-1}^t(u_d|u_1, \ldots, u_{d-1}) \\ =& \frac{1}{2} + \frac{\Gamma\left(\frac{d+\nu}{2}\right)\left(1+\frac{s}{\nu}\right)^{-\frac{1}{2}}}{(\pi\nu)^{\frac{1}{2}}\,\Gamma\left(\frac{(d-1)+\nu}{2}\right)} \cdot \left(\frac{t-\mu^d}{\sigma^d}\right) \cdot h\left(\frac{1}{2}, \frac{d+\nu}{2}, \frac{3}{2}, -\frac{\left(\frac{t-\mu^d}{\sigma^d}\right)^2}{\nu+s}\right) \end{aligned} \tag{3.16}$$

in which

- $s = \sum_{i=1}^{d-1} t_\nu^{-1}(u_i)$, $t = t_\nu^{-1}(u_d)$, $\mu^d = P_{21}^d \cdot \left(P_{11}^d\right)^{-1} \cdot \left(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_{d-1})\right)'$,
- $\sigma^d = P_{22}^d - P_{21}^d \cdot \left(P_{11}^d\right)^{-1} \cdot P_{12}^d$, and
- Gaussian hypergeometric function[13] $h(a, b, c, z) = \sum_{k=1}^{\infty} \frac{\Gamma(k+a)}{\Gamma(a)} \frac{\Gamma(k+b)}{\Gamma(b)} \frac{\Gamma(c)}{\Gamma(k+c)} \frac{z^k}{k!}$.

[12] This is Algorithm 9 in [63].
[13] See [2] for specification.

For a detailed derivation the reader is referred to [63], pp. 38-41. Equation (3.16) provides the following result[14].

A - 3.5 Algorithm (Goodness-of-Fit Test t-Copula)

Step 1 - for $n = 1, \dots, N$:

- set $v_{1,n} = \Psi_{1,n}$

Step 2 - for $n = 1, \dots, N$ and $d = 2, \dots, D$:

- compute $\mu^d = P_{21}^d \cdot \left(P_{11}^d\right)^{-1} \cdot \left(t_\nu^{-1}(\Psi_{1,n}), \dots, t_\nu^{-1}(\Psi_{d-1,n})\right)'$
- compute $\sigma^d = P_{22}^d - P_{21}^d \cdot \left(P_{11}^d\right)^{-1} \cdot P_{12}^d$
- compute $s = \sum_{i=1}^{d-1} t_\nu^{-1}(\Psi_{i,n})$
- transform $t = t_\nu^{-1}(\Psi_{d,n})$
- compute $v_{d,n} = \frac{1}{2} + \frac{\Gamma\left(\frac{d+\nu}{2}\right)\left(1+\frac{s}{\nu}\right)^{-\frac{1}{2}}}{(\pi\nu)^{\frac{1}{2}}\,\Gamma\left(\frac{(d-1)+\nu}{2}\right)} \cdot \left(\frac{t-\mu^d}{\sigma^d}\right) \cdot h\left(\frac{1}{2}, \frac{d+\nu}{2}, \frac{3}{2}, -\frac{\left(\frac{t-\mu^d}{\sigma^d}\right)^2}{\nu+s}\right)$

Step 3 - for $n = 1, \dots, N$:

- compute $S_n = \sum_{d=1}^{D} (\Phi^{-1}(v_{d,n}))^2$

Step 4

- perform χ^2-test to $S = (S_1, \dots, S_N)'$

Due to Proposition 2.13 and Equation (2.10), a goodness-of-fit test for the Clayton copula is performed by

A - 3.6 Algorithm (Goodness-of-Fit Test Clayton Copula)

Step 1 - for $n = 1, \dots, N$:

- set $v_{1,n} = \Psi_{1,n}$

Step 2 - for $n = 1, \dots, N$ and $d = 2, \dots, D$:

- compute $v_{d,n} = \left(1 + \frac{\Psi_{d,n}^{-\theta} - 1}{\sum_{i=1}^{d-1} \Psi_{i,n}^{-\theta} - (d-1) + 1}\right)^{-\frac{1}{\theta} - (d-1)}$

Step 3 - for $n = 1, \dots, N$:

- compute $S_n = \sum_{d=1}^{D} (\Phi^{-1}(v_{d,n}))^2$

Step 4

- perform χ^2-test to $S = (S_1, \dots, S_N)'$

[14] This is Algorithm 10 in [63].

4 Application I - *Value at Risk of a Stock Portfolio*

In the introduction of this thesis we have described types of risk to which companies acting on financial markets are confronted. A key approach to assess market risk - to be more precise equity risk - is the concept of the *value at risk* (VaR). In our application we measure the VaR of a stock portfolio by copulas.

Our exemplary portfolios contain up to 25 companies of the German stock index DAX[1]. We calibrate the presented copulas to log-returns of the stock prices for different periods and test the goodness of their fit. Further, simulations allow to forecast the portfolio's VaR for the next period. Numerical result are interpreted and discussed in a conclusion.

In order to have a comparison to the outcomes of the copula approach, we evaluate the VaR assessed by the Black & Scholes stock price model as well.

4.1 Black & Scholes Stock Price Model

In many fields of quantitative finance researchers require stock price evolutions. Black and Scholes in [6] and Merton in [50] simultaneously set up a stock price model[2] commonly used for financial applications. Although it has several unrealistic assumptions[3], the model itself and modifications thereof are frequently applied as they

- provide a manageable implementation,
- have low computational costs, and
- return adequate as well as useful results.

The main assumption of the Black & Scholes model (BS) are log-normally distributed stock returns. Further, we assume that share prices evolve according to the stochastic differential equation

$$\frac{dS_t}{S_t} = \left(\mu + \frac{\sigma^2}{2}\right) dt + \sigma dW_t \tag{4.1}$$

in which σ is the *constant volatility*, $\mu + \frac{\sigma^2}{2}$ the *constant drift rate*, W_t a Brownian motion and some initial condition S_0.

If each price evolution is driven by **one** Brownian motion W_t^d, $d = 1, \ldots, D$, and their covariance matrix Σ can be decomposed by Cholesky factorization into $L \cdot L'$, a possible multivariate extension of Equation (4.1) is given by

$$\begin{pmatrix} \frac{dS_t^1}{S_t^1} \\ \vdots \\ \frac{dS_t^D}{S_t^D} \end{pmatrix} = \begin{pmatrix} \mu_1 + \frac{\sigma_1^2}{2} \\ \vdots \\ \mu_D + \frac{\sigma_D^2}{2} \end{pmatrix} dt + L \begin{pmatrix} dW_t^1 \\ \vdots \\ dW_t^D \end{pmatrix} \tag{4.2}$$

where $\sigma_d = \Sigma_{d,d}$. The VaR results computed by this model serve as comparison and benchmark for the outputs of the copula approach. One often calls this model *delta-normal approach* and connects it with Monte-Carlo simulations.

4.2 Value at Risk

Risk management departments are obliged to quantify the risk inherited in their stock portfolios. The high volatility of stock markets incapacitate long periods of forecasting. Results can be

[1] DAX composition of December 31, 2009, see Appendix A.5.
[2] The literal purpose of both papers is to price European options.
[3] Cont and Tankov give a survey on that topic in [14], Chapter 1.

imprecise or poorly conceived and interpretable. Most commonly, portfolio risk is measured on daily, weekly or monthly basis. Several key figures have been developed such as the mentioned value at risk and the *expected shortfall*[4] just to mention the well-known ones.

4.1 Definition (Value at Risk, compare [49], Definition 2.10)
Let $\alpha \in (0,1)$ be some confidence level and let pr denote the portfolio return. The *value at risk at confidence level* α is given by the smallest number l such that the probability that pr is below or equal to l is not larger than $1-\alpha$. Mathematically,

$$\mathrm{VaR}_\alpha = \inf\{l \in \mathbb{R} \,|\, \mathbb{P}[pr \leq l] \leq 1-\alpha\} \quad .$$

Thus, the VaR is the quantile function $q_{1-\alpha}$ of the portfolio return distribution F_{pr}, i.e.,

$$q_{1-\alpha} = F_{pr}^{-1}(1-\alpha)$$

with F_{pr}^{-1} as inverse of F_{pr}. Convenient confidence levels are $\alpha = 0.95$, $\alpha = 0.99$ and $\alpha = 0.995$.

Definition 4.1 states that with the probability of α the portfolio return pr exceeds l, i.e., the portfolio return pr is larger than the VaR_α - or in other words emphasizing the risk of the portfolio:

In $(1-\alpha)\times 100$ % of all cases the portfolio return is equal or below the VaR_α.

For a detailed insight to the world of quantitative risk management the reader is referred to [49].

4.2.1 VaR - Copula Approach

Suppose that our portfolio comprises D stocks and we are endowed with substantial data, i.e., N log-returns

$$\hat{s}_n(d) = \log\left(\frac{stock\ price_{n+1}(d)}{stock\ price_n(d)}\right) \qquad \text{for } n = 1,\ldots,N-1$$

and each asset $d = 1,\ldots,D$, such that the empirical distribution functions F_{dN}, $d = 1,\ldots,D$, are adequately smooth. Selected copulas (Gaussian, t and Clayton copula) are calibrated to the log-returns by the CML method presented in Subsection 3.1.3. Further, we run the goodness-of-fit test of Section 3.3 and simulate outcomes for the fitted copulas. Simulation results u_d are transformed by the generalized inverse function F_{dN}^{-1} of the ecdfs to obtain simulated log-returns

$$s_d = F_{dN}^{-1}(u_d)$$

for each asset $d \in \{1,\ldots,D\}$. Technically, this can be done by finding the optimal index

$$n^* = \min_{n\in\{1,\ldots,N\}} |u_d - F_{dN}(x_{d,n})| \tag{4.3}$$

and setting $s_d = x_{d,n^*}$ for each $d \in \{1,\ldots,D\}$. Note that the larger N the smaller are the errors of Equation (4.3). Thus, for a smooth transformation (copula simulations to log-returns) the number of data points should be extensive. Further, one replicates only those log-returns which have occurred in the past. Hence, the larger N the greater is the variety of simulated returns.

4.2 Assumption
We assume that our portfolio is equally weighted and that each asset has starting value $S_0^d = 1$ for $d = 1,\ldots,D$. Simulation or forecasting horizon is **one** period.

[4] The expected shortfall is a sort of expectation of the VaR, see Chapter 8, Notation 8.3.

Under the assumption of log-returns, the predicted portfolio return is computed as

$$pr = \sum_{d=1}^{D} \exp(s_d) \quad .$$

We perform Monte-Carlo simulation and produce $M \in \mathbb{N}$ portfolio returns

$$pr_m = \sum_{d=1}^{D} \exp(s_{d,m}) \quad \text{for } m = 1, \dots, M$$

which enables us to set up a distribution function of forecast portfolio returns as

$$\hat{F}_{pr}(x) = \frac{1}{M+1} \sum_{m=1}^{M} \mathbf{1}_{\{pr_m \leq x\}}(x) \quad . \tag{4.4}$$

We derive the VaR_α by means of the $(1-\alpha)$-quantile of Equation (4.4)

$$\text{VaR}_\alpha = \hat{F}_{pr}^{-1}(1-\alpha) \quad .$$

A - 4.1 Algorithm (VaR - Copula Approach)

Given

- observation matrix Υ
- empirical distribution functions F_{dN}, $d = 1, \dots, D$
- pseudo-observation matrix Ψ
- copula C
- number of Monte-Carlo simulations $M \in \mathbb{N}$ and confidence level $\alpha \in (0, 1)$

Step 1

- calibrate copula C to given data Υ and Ψ respectively
- perform goodness-of-fit test for C

Step 2 - Monte-Carlo simulation - for $m = 1, \dots, M$:

- draw $(u_{1,m}, \dots, u_{D,m})' \sim C$
- for $d = 1, \dots, D$:
  - compute $n_d^* = \min_{n \in \{1,\dots,N\}} |u_{d,m} - F_{dN}(x_{d,n})|$
  - set $s_{d,m} = x_{d,n_d^*}$
- compute portfolio return $pr_m = \sum_{d=1}^{D} \exp(s_{d,m})$

Step 3

- compute simulated distribution function

$$\hat{F}_{pr}(x) = \frac{1}{M+1} \sum_{m=1}^{M} \mathbf{1}_{\{pr_m \leq x\}}(x)$$

- compute $\text{VaR}_\alpha = \hat{F}_{pr}^{-1}(1-\alpha)$

4.2.2 VaR - BS Model

For calibration we recall that due to Itô's Lemma, e.g., see [60], the solution of Equation (4.2) is

$$\begin{pmatrix} S_t^1 \\ \vdots \\ S_t^D \end{pmatrix} = \begin{pmatrix} S_0^1 \\ \vdots \\ S_0^D \end{pmatrix} \cdot \exp\left(\mu t + L \cdot W_t\right) \tag{4.5}$$

in which $W_t = (W_t^1, \ldots, W_t^D)'$ is D-dimensional Brownian motion, $\mu = (\mu_1, \ldots, \mu_D)'$ is the *mean rate of return*, and $L \cdot L'$ is the Cholesky decomposition of Σ. Unbiased estimators for μ and Σ are

$$\hat{\mu}_d = \frac{1}{N}\sum_{n=1}^{N} x_{d,n} \quad \text{and} \quad \hat{\Sigma}_{i,j} = \frac{1}{N-1}\sum_{n=1}^{N}(x_{i,n} - \hat{\mu}_i)(x_{j,n} - \hat{\mu}_j) \quad \text{for } i,\, j = 1, \ldots, D \quad .$$

If we set $t = 1$ in Equation (4.5) and simulate $N(\hat{\mu}, \hat{\Sigma})$-distributed values[5] $(s_1^{BS}, \ldots, s_D^{BS})'$, we have forecast returns s_d^{BS}, $d = 1, \ldots, D$, for that period on which we calibrated the model. Thus, under Assumption 4.2, the portfolio return for the next period is obtained by

$$pr^{BS} = \sum_{d=1}^{D} \exp\left(s_d^{BS}\right) \quad .$$

The VaR_α^{BS} of the BS model for the confidence level α is gained by the $(1-\alpha)$-quantile of the predicted distribution function produced by a Monte-Carlo simulation.

A - 4.2 Algorithm (VaR - BS Model)

Given

- observation matrix Υ
- number of Monte-Carlo simulations $M \in \mathbb{N}$ and confidence level $\alpha \in (0, 1)$

Step 1 - for $d = 1, \ldots, D$:

- compute $\hat{\mu}_d = \frac{1}{N}\sum\limits_{n=1}^{N} x_{d,n}$
- set $\hat{\mu} = (\hat{\mu}_1, \ldots, \hat{\mu}_D)'$

Step 2 - for $i,\, j = 1, \ldots, D$:

- compute $\hat{\Sigma}_{i,j} = \frac{1}{N-1}\sum\limits_{n=1}^{N}(x_{i,n} - \hat{\mu}_i)(x_{j,n} - \hat{\mu}_j)$

Step 3 - Monte-Carlo simulation - for $m = 1, \ldots, M$:

- draw $(s_{1,m}^{BS}, \ldots, s_{D,m}^{BS})' \sim N(\hat{\mu}, \hat{\Sigma})$
- compute portfolio return $pr_m^{BS} = \sum\limits_{d=1}^{D} \exp\left(s_{d,m}^{BS}\right)$

Step 4

- compute simulated distribution function

$$\hat{F}_{pr}^{BS}(x) = \frac{1}{M+1}\sum_{m=1}^{M} \mathbf{1}_{\{pr_m^{BS} \le x\}}(x)$$

- compute $\text{VaR}_\alpha^{BS} = \left(\hat{F}_{pr}^{BS}\right)^{-1}(1-\alpha)$

[5] Explained in Appendix A.4.

4.3 Forecasting and Numerical Results

Forecasting results base on the sample matrix Υ. For instance, if calibration is carried out on daily log-returns, simulated outcomes are a prediction for the portfolio return of one day. This time horizon can be expanded or shortened to any arbitrary period with certain considerations on calibration and simulation specifications.

Two alternatives exist for this problem. The first is to calibrate the model to data which bases on the desired period. The advantage here is that a simulation of the model yields the prediction for the very time horizon. However, depending on the length of the period, calibration can be difficult due to scarce data[6].

The second is to fit the model on the basis of a shorter period and perform simulations for multiple periods. For example, one provides $K \in \mathbb{N}$ simulations of a daily graduated model and summarizes to receive the outcome for K days. The disadvantages for this approach is an increasing computing time due to **K** simulations instead of **one** simulation. Further, one predicts the performance by K daily returns instead of a K-days return which might cause inconsistent results.

4.3 Remark (Calibration of the BS Stock Price Model)

1. The BS stock price model is often calibrated to a yearly basis, but simulation of daily returns can be performed easily. One has to draw $N(\Delta t\mu, \Delta t\Sigma)$-distributed values with μ and Σ fitted on yearly basis and $\Delta t = \frac{1}{200}$[7].
2. In practical applications for longer periods, e.g., VaR for one year, the drift term of the price evolutions can be set to zero. For a non-zero drift, the trend superimposes the stochastic evolution of the share price. Risk - in terms of the volatility and the Brownian motion - becomes insignificant in this case.
3. For longer periods, the assumption of a constant volatility appears to be questionable.

The basis of the subsequent studies is an equally weighted portfolio with standardized initial prices and a forecasting horizon of one period - Assumption 4.2.

In our analysis we apply four different periods

- one day ($K = 1$),
- one week ($K = 5$),
- two weeks ($K = 10$), and
- one month ($K = 20$).

Each observation matrix Υ_K hence contains log-returns for $K = 1$, 5, 10 and 20 days[8]. Calibration and simulation are performed on these periods for the Gaussian, t- and Clayton Copula.

We analyze portfolios which comprise up to $D = 25$ companies of the German stock index DAX[9]. Data was generated by Bloomberg, log-returns were computed from time-consistent closing prices including all corporate actions such as spin-offs, split-offs, etc., from January 1, 2000 till December 31, 2009.

Initially, a goodness-of-fit test is operated by Algorithms A - 3.4, A - 3.5 and A - 3.6 for the above mentioned copulas. Parameters D, **size of the portfolio**, and K, **the length of the period**, are altered. For χ^2-tests, we use the MATLAB routine `kstest2`[10] with a confidence level of 95%.

[6] Concerning minor data series, the assumption of empirical cdfs for the CML calibration can be dropped in favor of closed-form marginal distributions.

[7] For the trading year at stock exchanges one normally sets $\Delta t = \frac{1}{200}$ or $\Delta t = \frac{1}{250}$.

[8] Note that out of 2001 daily closing prices, we receive $N = 2000$ daily but only $N = 100$ monthly log-returns. The impact on the ecdfs is obvious and should be considered when discussing and evaluating results.

[9] Composition was assembled on December 31, 2009, see Appendix A.5. Not all companies were continuous members of the DAX.

[10] `kstest2` is a two sample Kolmogorov-Smirnov test, details are found in the MATLAB product help.

	Gaussian copula	***t*-copula**	**Clayton copula**
K = 1	*p*-value / hypothesis		
$D = 5$	10^{-20} / 1	0.0356 / 0.7716[12]	10^{-23} / 1
$D = 10$	10^{-62}/ 1	10^{-9}/ 1	10^{-91}/ 1
$D = 15$	10^{-91} / 1	10^{-17} / 1	10^{-116} / 1
$D = 20$	10^{-111} / 1	10^{-28} / 1	10^{-134} / 1
$D = 25$	10^{-126} / 1	10^{-39} / 1	10^{-168} / 1
K = 5	*p*-value / hypothesis		
$D = 5$	10^{-07} / 1	0.3525 / 0	10^{-08} / 1
$D = 10$	10^{-10}/ 1	0.0900 / 0.0745[12]	10^{-16}/ 1
$D = 15$	10^{-16} / 1	0.0058 / 1	10^{-22} / 1
$D = 20$	10^{-20} / 1	10^{-06} / 1	10^{-31} / 1
$D = 25$	10^{-24} / 1	10^{-08} / 1	10^{-40} / 1
K = 10	*p*-value / hypothesis		
$D = 5$	0.0027 / 1	0.7166 / 0	10^{-04} / 1
$D = 10$	10^{-05}/ 1	0.2700 / 0	10^{-09}/ 1
$D = 15$	10^{-07} / 1	0.0288 / 0.9468[12]	10^{-10} / 1
$D = 20$	10^{-09} / 1	10^{-04} / 1	10^{-15} / 1
$D = 25$	10^{-10} / 1	10^{-08} / 1	10^{-22} / 1
K = 20	*p*-value / hypothesis		
$D = 5$	0.1090 / 0.0022[12]	0.9804 / 0	0.0034 / 1
$D = 10$	0.0928 / 0.0021[12]	0.3866 / 0	10^{-05}/ 1
$D = 15$	0.0140 / 1	0.1854 / 0	10^{-05} / 1
$D = 20$	10^{-04} / 1	10^{-04} / 1	10^{-08} / 1
$D = 25$	10^{-05} / 1	10^{-05} / 1	10^{-08} / 1

T - 4.1 Table: χ^2-test results for different periods K and portfolio sizes D

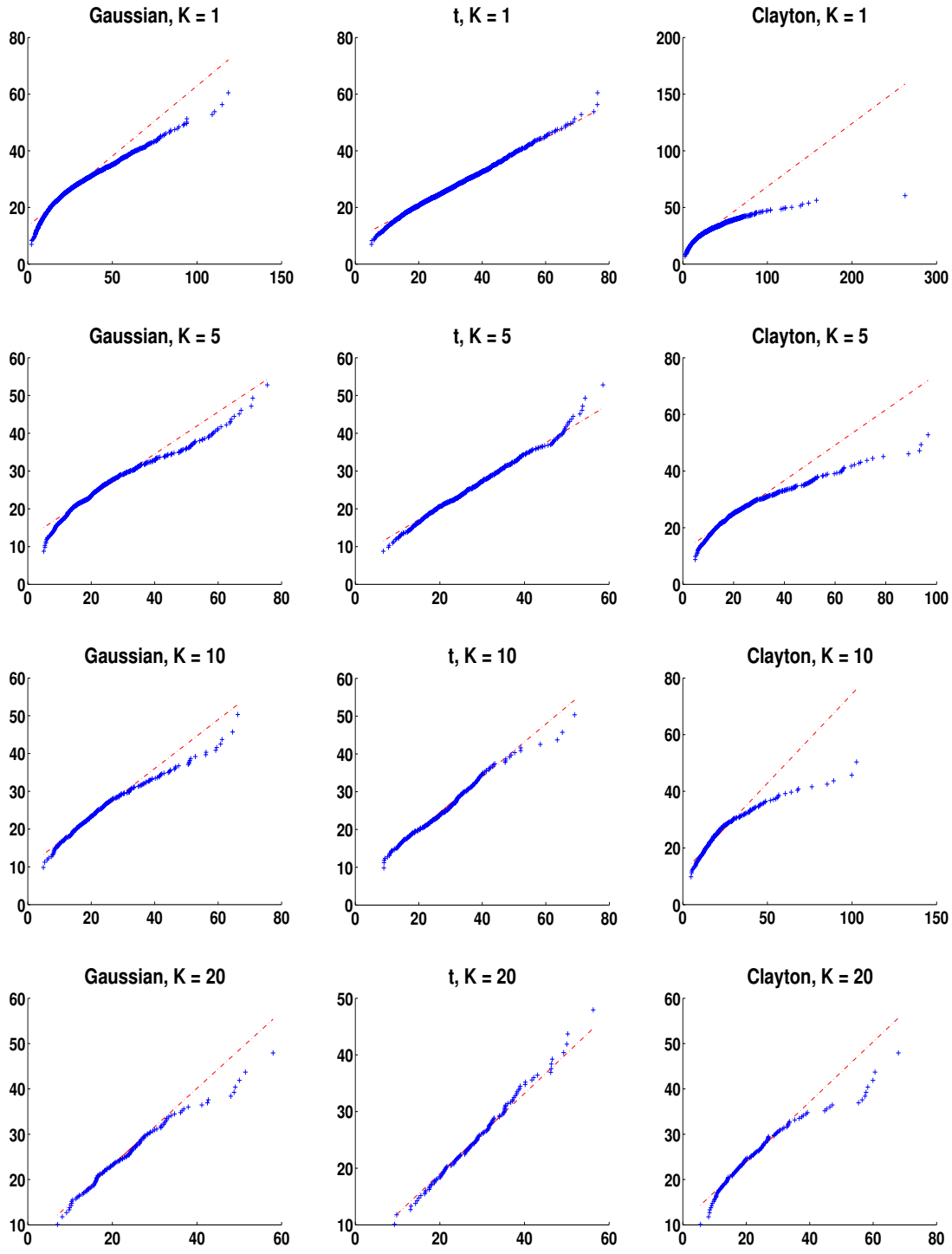
Moreover, we run a sort of Monte-Carlo simulation to average on χ^2-test results, i.e., the p-value and the hypothesis[11]. Additionally, quantile-quantile-plots are exemplified.

The goodness-of-fit test results are given in Table T - 4.1. A hypothesis of 1 rejects the assumption that the given data is distributed according to the different copulas. A hypothesis of 0 accepts this assumption. The p-value is the conventional p-value in statistics which indicates the probability of receiving a test statistic at least as extreme as the actually observed one. Figure F - 4.1 shows quantile-quantile-plots of the χ^2-test results compared to the sample data for a portfolio of $D = 25$ assets.

Further, we apply Algorithm A - 4.1 to the Gaussian, t- and Clayton copula. Results are compared to those of Algorithm A - 4.2. We vary parameters D and K. In addition, we test different confidence levels for the VaR, $\alpha = 0.95$, $\alpha = 0.99$ and $\alpha = 0.995$. The number of Monte-Carlo simulations is set to $M = 10.000$ for both algorithms. Computational results for a portfolio size of $D = 25$ assets are shown in Table T - 4.2. For $D = 5$, 10, 15 and 20, they can be found in Appendix A.6.

[11] To check fitting results on stability we repeat the χ^2-test with different χ^2-distributed random variables.

[12] As we average the hypothesis, i.e., running the `kstest` several times, these numbers indicate the percentaged values how often the hypothesis is accepted.



F - 4.1 Figure: χ^2-tests - quantile-quantile plots for different periods and $D = 25$ assets

	Gaussian copula	***t*-copula**	**Clayton copula**	**BS model**
D = 25	$\alpha = 0.95$			
$K = 1$	-0.023551	-0.023652	-0.021511	-0.024335
$K = 5$	-0.049081	-0.049826	-0.047979	-0.049615
$K = 10$	-0.073535	-0.074659	-0.074079	-0.074058
$K = 20$	-0.10551	-0.10567	-0.10872	-0.10669
D = 25	$\alpha = 0.99$			
$K = 1$	-0.038163	-0.040631	-0.03974	-0.035481
$K = 5$	-0.077904	-0.082861	-0.090017	-0.069702
$K = 10$	-0.12026	-0.13087	-0.14113	-0.10277
$K = 20$	-0.16502	-0.18577	-0.19622	-0.1459
D = 25	$\alpha = 0.995$			
$K = 1$	-0.047016	-0.048824	-0.048329	-0.038321
$K = 5$	-0.090212	-0.097282	-0.11583	-0.077151
$K = 10$	-0.13681	-0.15919	-0.17084	-0.11523
$K = 20$	-0.18544	-0.20148	-0.22479	-0.15849

T - 4.2 Table: VaR_α simulation results of A - 4.1 and A - 4.2 for $D = 25$

How to read Table T - 4.2

A $VaR_\alpha = -0.05$ tells us that in $\alpha \times 100\%$ of all cases the expected returns is larger than -0.05, i.e., larger than -5%. Equivalently, the expected loss is equal to or larger[13] than -5% with a probability of $1 - \alpha$.

4.4 A First Conclusion

4.4.1 Goodness-of-Fit Test Result

Goodness-of-fit tests mainly reject the assumption that data is distributed according to the different copulas - especially for high-dimensional portfolios and short periods. The hypothesis is supposed to be valid for the t-copula with parameter settings

- $K = 5$ and $D = 5$, 10[14],
- $K = 10$ and $D = 5$, 10,
- $K = 20$ and $D = 5$, 10, 20

and for the Gaussian copula with parameter settings

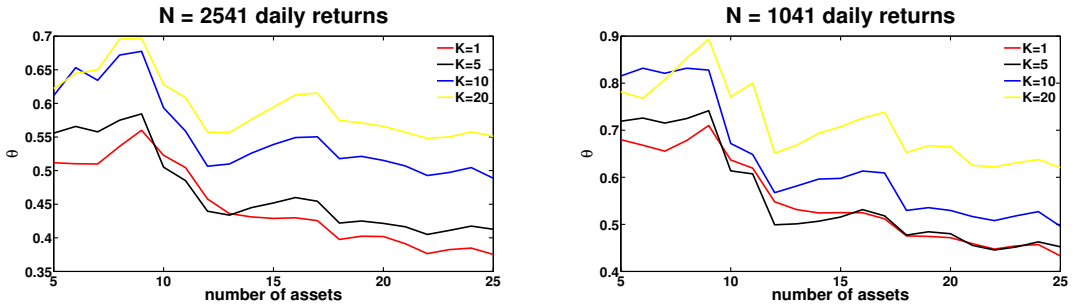
- $K = 20$ and $D = 5$[14], 10[14].

For longer periods and small portfolios, the t-copula gives the best fit results. From Table T - 4.1 we infer that

- the longer the period K and
- the smaller the portfolio size

the better are the p-values. It is important to mention that the portfolio set-up does not significantly influence fit test results. Outcomes are similar for different compositions. In contrast,

[13] In sense of more negative.
[14] Suitable with limited extent, compare to Footnote 12.



F - 4.2 Figure: Clayton copula parameter θ - dependence on sample sizes

portfolio size is a more dominating factor than portfolio configuration. The driving force is the dependency structure captured by the copula.

The t-copula provides the best fit followed by the Gaussian copula. Clayton copula gives the worst fit. This stands to reason as the Clayton copula parameter θ is calibrated to values in the interval [0.4, 0.7] for all combinations of K and D, see left hand side of Figure F - 4.2. This represents almost independence - a quite remote assumption. Note that θ is dependent on the size of the observation sample. For $N = 1041$ daily returns parameter are in the interval [0.5, 0.9], see right hand side of Figure F - 4.2. Moreover, the larger the portfolio the lower is θ. For different sample sizes, parameter evolutions in dependence of D are quite alike, Figure F - 4.2 vs. Figure F - A.1.

Graphically, the t-copula gives nice q-q-plots for all periods and a portfolio size of $D = 25$, see Figure F - 4.1. The Gaussian copula model misses the plots at the tails. This is a well known problem. The Clayton copula has the worst plots especially in the upper part of the distributions. Figures F - 4.1 and Table T - 4.1 indicate that the t-copula performs well for longer periods and small portfolios.

4.4 Summary (Fit Test Results)

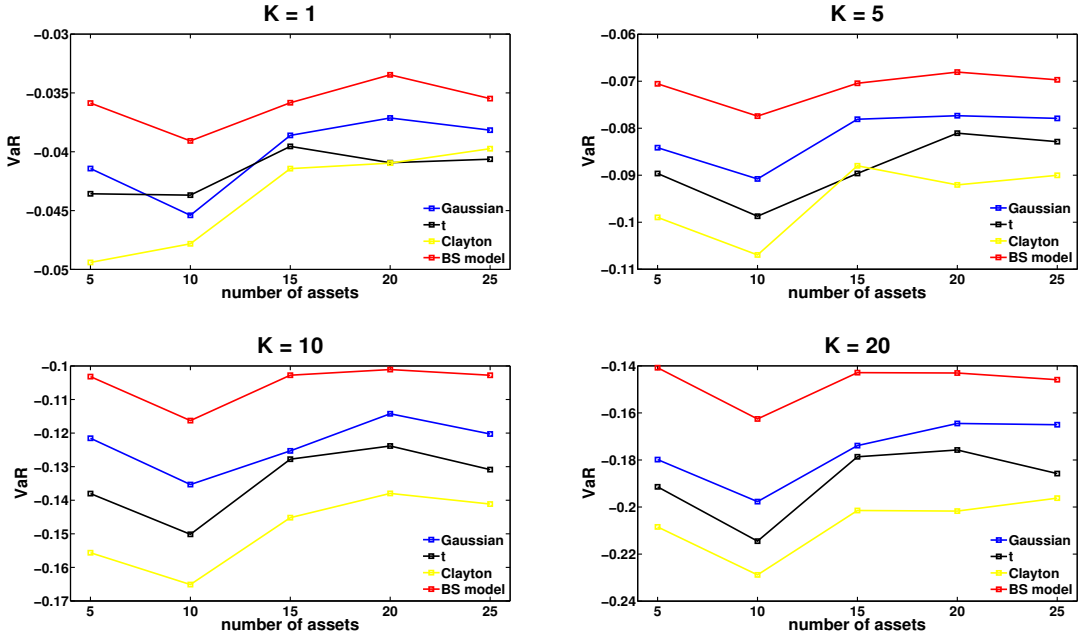
Under Assumption 4.2 we receive the following implications.

1. Goodness-of-fit test results are dependent on the dimensionality of the portfolio and the length of the period.
   - The smaller the portfolio and the longer the period, the better the fit.
   - The larger the portfolio and the shorter the period, the worse the fit.
2. For higher dimensional portfolios ($D > 15$), hypotheses that data is distributed according to the relevant copula are rejected.
3. The portfolio composition has negligible influence on fitting results.
4. Graphically, the t-copula evidences a good fit.
5. The Clayton copula parameter θ is dependent on the sample size N.

4.4.2 Simulation Results

We apply Algorithms A - 4.1 and A - 4.2 to the above given portfolio structures and parameter settings. For a confidence level of $\alpha = 0.95$, the three copula models and the BS model compute similar VaR's for all periods, see Table T - 4.2. This is essential as one can restrict to the standard black BS model if interested in a $VaR_{\alpha=0.95}$. The model is easier to implement and requires less computational time.

For the other confidence levels $\alpha = 0.99$ and $\alpha = 0.995$, the BS model simulates the smallest value at risk followed by the Gaussian and t-copula. The Clayton model estimates the largest VaR in



F - 4.3 Figure: $VaR_{\alpha=0.99}$ results - dependence on portfolio size D

sense of the largest potential loss. This observation holds for all four periods and is independent of the portfolio size, compare Table T - 4.2 with Table T - A.1.

The Clayton copula parameter is calibrated to approximately 0.6 for $D = 25$ which is close to the independence copula. Thus, it is not surprising that the Clayton model expects the highest risk. From an economic point of view, single returns are assumed to be almost mutually independent. This in turns means that only poor or even no diversification effects exits. It contradicts the economic perception that unsystematic risk is eliminated by diversification. Hence, the Clayton copula does not seem to be suitable for VaR forecasts.

Outcomes are consistent. We naturally expect the VaR for longer periods to be larger than the VaR for shorter ones. For all models this is true. Additionally, the differences for the VaR between each approach depend on the length of the period: the shorter the period the smaller the differences.

Figure F - 4.3 compares different value at risk simulations in dependence of portfolio size and and length of the period for a confidence level $\alpha = 0.99$. The largest value at risk is exhibited for $D = 10$ for all periods. It evidences that portfolio size effects VaR results. This outcome is consistent with the economic finding that unsystematic risk is reduced by diversification. In combination with goodness-of-fit test results, it suggests that the influence of dimensionality is not negligible.

4.5 Summary (Value at Risk Simulation)

1. All models (copula and benchmark) yield similar results for a confidence level of $\alpha = 0.95$.
2. For higher confidence levels, the Clayton copula has the largest risk expectation followed by the t- and Gaussian copula approach. The BS model simulates the smallest risk. This result holds for all periods and portfolio sizes.
3. The Clayton copula model assumes that single returns are almost independent for large samples. Consistent with goodness-of-fit results, it is not deemed to be suitable for VaR assessment on larger portfolios.
4. The value at risk is influenced by the portfolio size. For $D > 10$ diversification effects set in.
5. Value at risk simulation results are consistent with the goodness-of-fit test results, so far.

4.6 Remark
Performing goodness-of-fit tests means testing whether available data is distributed according to certain distributions. For VaR results we are only interested in a certain quantile of concerned distributions. Thus, simulation can provide adequate results, i.e., hitting the quantiles, although the hypothesis of the goodness-of-fit-test is rejected.

5 Application II - *Backtesting the Value at Risk*

Since the beginning of copula applications, researchers started to innovate approaches by launching time-dependent concepts into the copula theory. In 2001, Patton [54] introduced a first work on financial applications about time-varying exchange-rate dependence. He uses sub-σ-algebras to model conditional copulas and accordingly establishes a time-dependent component[1]. The main aspect is that pseudo-copulas[2] have the following representation

$$F(x_1, \ldots, x_D|\mathcal{F}) = C(F_1(x_1|\mathcal{G}_1), \ldots, F_D(x_D|\mathcal{G}_D)|(\mathcal{G}_1, \ldots, \mathcal{G}_D), \mathcal{F}) \quad .$$

in which $\mathcal{F}$ and $\mathcal{G}_1, \ldots, \mathcal{G}_D$ are σ-algebras, see [25]. The idea of filtration-conditioned copulas can be transferred to calibration issues in a simple way.

5.1 Recalibration - *Discrete Time-Dependent Copulas*

Suppose that we are endowed with a data series of length N, i.e., $\Upsilon \in \mathbb{R}^{D\times N}$. We calibrate the copula on basis of matrix Υ. As time passes, we receive new information. An additional data point $(x_{1,N+1}, \ldots, x_{D,N+1})'$ is revealed and Υ enlarges, i.e., $\Upsilon \in \mathbb{R}^{D\times N+1}$. As a consequence, the copula must be recalibrated resulting in new, possibly different parameters.

5.1 Notation
Let $\Upsilon^* \in \mathbb{R}^{D\times N^*}$ be a sample matrix and $t = 0, 1, \ldots, T$, $T \in \mathbb{N}$ and $T < \infty$, be a sequence of time points on which new relevant data is exposed. We denote a *discrete time-dependent copula parameter* as

$$\theta_t = (\theta_1^t, \ldots, \theta_D^t, \theta_c^t) \quad \text{for } t = 0, 1, \ldots, T$$

and define the *discrete time-dependent copula* as

$$C_t(u) = C(u_1, \ldots, u_D, \theta_1^t, \ldots, \theta_D^t, \theta_c^t) \quad \text{for } t = 0, 1, \ldots, T \quad .$$

If a recalibration is performed at every time point $t = 0, 1, \ldots, T$ (arrival of new information), we obtain a sequence of parameters θ_t which are constant on the interval $[t, t+1)$, $t = 0, \ldots, T$. This implies a series of discrete time-dependent copulas C_t.

5.2 Remark
1. The outcome for the time $t+1$ is simulated by the copula C_t, i.e., the copula calibrated at time t.
2. Especially high-dimensional recalibration is computationally expensive.
3. The length of the forecasting period K must be maintained for each recalibration step.

Recalibration can be performed in two ways. The first is to keep the size of the used matrix Υ constant. For every new data point $(x_{1,N+t}, \ldots, x_{D,N+t})$ one takes out the oldest entry $(x_{1,t}, \ldots, x_{D,t})$, $t = 1, \ldots, T$. Thus, Υ has a fixed size of $D\times N$. We denote this approach as *moving (re-)calibration*.

[1] The reader is also referred to the articles [8] and [25].
[2] These copulas do **not generally** satisfy the requirement of uniform margins, hence the name *pseudo-copulas*.

The second is to add the new information $(x_{1,N+t}, \ldots, x_{D,N+t})$ to the sample matrix. Hence, the used matrix Υ increases its size to $D \times N+t$ in every time step $t = 1, \ldots, T$. For CML recalibration, this is suggestive if data is scarce and N is consequently small, see Footnote 6 in Section 4.3. We call this procedure *added (re-)calibration*.

5.2 Backtesting

The idea of backtesting is to go **back** in time and **test** the model in question. Backtesting can be linked to recalibration. First, go back to a certain time point and restrict the available information and sample matrix to that very time. Then, simulate results for the next time step and compare these outputs with real outcomes. Now, recalibrate the model as new information is exposed. Again, simulate and compare to reality. This procedure is summarized in

A - 5.1 Algorithm (Backtesting VaR)

Given

- observation matrix $\Upsilon^* \in \mathbb{R}^{D \times N^*}$
- confidence level $\alpha \in (0, 1)$
- backtesting period T

Step 1 - for $t = 0, 1, \ldots, T-1$:

- set $N = N^* - T + t$
- set $\Upsilon \in \mathbb{R}^{D \times (1,\ldots,N)}$ for added calibration
- set $\Upsilon \in \mathbb{R}^{D \times (1+t,\ldots,N)}$ for moving calibration
- perform VaR-Algorithms A - 4.1 or A - 4.2, respectively
- receive VaR_α^t
- compare VaR_α^t with real-return[3] rr_t

  Under Assumption 4.2 the real-return rr_t of the period from t to $t+1$ is computed as

$$rr_t = \sum_{d=1}^{D} \exp\left(x_{d,N+t+1}\right) \quad .$$

Interpretation

For backtesting the value at risk of a portfolio, a precise interpretation can be given. We count the number of periods in which the VaR of the model is lower or equal to the real-return relative to the number of all periods, i.e.,

$$\eta := \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{1}_{\{VaR_\alpha^t \leq rr_t\}}(t) \quad .$$

If $\eta \approx 1 - \alpha$, the model passes the backtesting. The number of predicted outcomes, that are equal to the real-return, is approximately the expected number of shortfalls allowed by the predetermined confidence level α.

5.3 Example
For a backtesting of $T = 1000$ periods and a confidence level $\alpha = 0.99$, the number of simulation results not larger than the real-return is expected to be 10.

[3]For evaluation we have implicitly assumed that $S_t^1 = \ldots = S_t^D = 1$ holds for each period $t = 0, 1, \ldots, T-1$ consistent with Assumption 4.2.

confidence level $\alpha = 0.99$	**added calibration**	**moving calibration**
portfolio size	$D = 5$ / $D = 10$ / $D = 15$ / $D = 20$ / $D = 25$	

	$T = 1000$ and $K = 1$	
Gaussian copula	18 / 21 / 22 / 22 / 25	19 / 25 / 23 / 23 / 24
***t*-copula**	18 / 20 / 19 / 19 / 19	18 / 20 / 20 / 20 / 19
Clayton copula	11 / 15 / 16 / 15 / 17	11 / 14 / 15 / 13 / 14
BS model	21 / 30 / 25 / 24 / 26	23 / 30 / 26 / 25 / 30

	$T = 200$ and $K = 5$	
Gaussian copula	2 / 4 / 5 / 6 / 4	3 / 5 / 6 / 6 / 4
***t*-copula**	2 / 4 / 3 / 3 / 3	2 / 4 / 3 / 4 / 4
Clayton copula	2 / 3 / 2 / 1 / 2	2 / 2 / 2 / 2 / 1
BS model	7 / 10 / 8 / 7 / 7	10 / 11 / 9 / 9 / 11

	$T = 100$ and $K = 10$	
Gaussian copula	3 / 3 / 4 / 3 / 3	3 / 4 / 4 / 4 / 4
***t*-copula**	2 / 3 / 3 / 2 / 3	3 / 4 / 4 / 4 / 4
Clayton copula	2 / 2 / 2 / 2 / 2	2 / 2 / 2 / 2 / 2
BS model	3 / 4 / 4 / 3 / 4	5 / 4 / 4 / 4 / 4

	$T = 50$ and $K = 20$	
Gaussian copula	1 / 1 / 1 / 1 / 1	1 / 1 / 1 / 1 / 1
***t*-copula**	1 / 1 / 1 / 1 / 1	1 / 1 / 1 / 1 / 1
Clayton copula	1 / 1 / 1 / 1 / 1	1 / 1 / 1 / 1 / 1
BS model	1 / 1 / 1 / 1 / 1	1 / 3 / 1 / 1 / 2

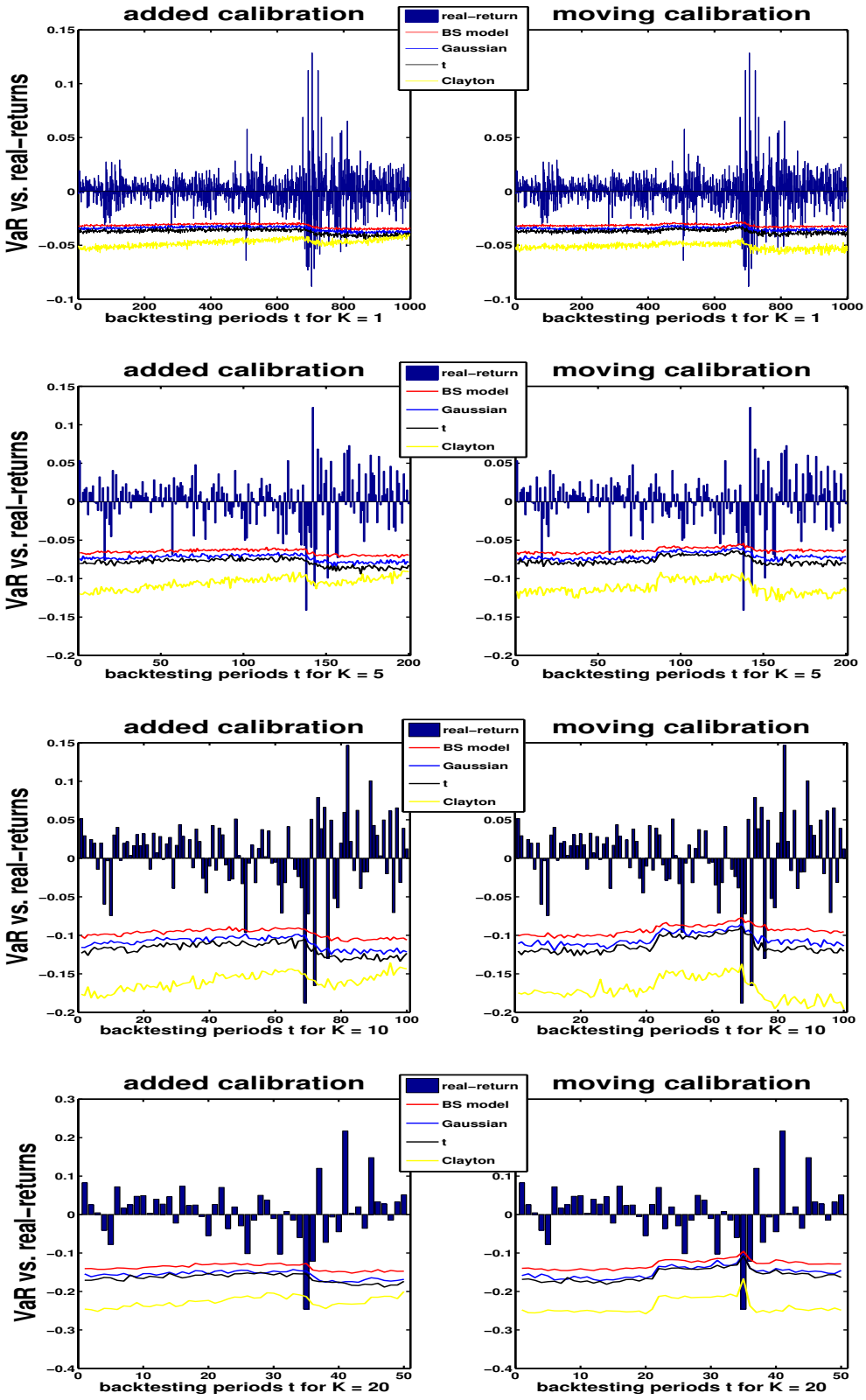
T - 5.1 Table: backtesting - number of shortfalls in tested periods

If the value of η significantly differs from $1 - \alpha$, the model has failed the backtesting. For the case $\eta < 1 - \alpha$, risk is overestimated. Too much capital is unnecessarily reserved. On the other hand, if $\eta > 1 - \alpha$, risk is underestimated. More capital is consumed than expected by the model.

5.3 Numerical Results

Backtesting is performed by Algorithm A - 5.1 for the three different copula models: Gaussian, *t*- and Clayton copula. The algorithm is also run for the standard BS model as a comparative benchmark. Numerical results of Section 4.3 show similar VaR results with confidence level of $\alpha = 0.95$ for all four approaches. Therefore, we restrict the analysis to $\alpha = 0.99$. As before, we perform a Monte-Carlo simulation of $M = 10000$ and test four different periods: $K = 1$, 5, 10 and 20 days. The backtesting period is set to $T = 1000$ days. Thus, we receive $T = 1000$ for $K = 1$, $T = 200$ for $K = 5$, $T = 100$ for $K = 10$ and $T = 50$ for $K = 20$. Further, we apply the moving and added recalibration.

Table T - 5.1 shows the number of shortfalls of the four approaches for different portfolio sizes $D = 5$, 10, 15, 20 and 25. It indicates how often the simulated VaR_α^t is smaller than the real-return rr_t. For a confidence level $\alpha = 0.99$, we expect this value approximately to be 1% of the length of the backtesting period T. In addition, Figure F - 5.1 gives a graphical comparison of the model outcomes with real-returns and portfolio size of $D = 25$. It displays the VaR evolution for each approach for the backtesting period. Real-returns breaking through solid VaR lines indicate shortfalls.



F - 5.1 Figure: backtesting for different periods and $D = 25$, $\alpha = 0.99$

As recalibration is performed in every backtesting time step, Algorithm A - 5.1 is computationally intensive especially for shorter periods. Fitting the models in two different ways - added and moving calibration - increases computational time.

5.4 A Second Conclusion

Figure F - 5.1 confirms VaR simulation results of Chapter 4 in the sense that the Clayton copula expects the largest risk followed by the t- and Gaussian copula. The BS stock price model has the lowest risk expectation. For a confidence level of $\alpha = 0.99$, this order is true for all periods and portfolio sizes validated by Table T - 5.1. The number of shortfalls in backtesting mainly maintains this observation. The most shortfalls occur for the BS model, the least for the Clayton copula.

As already evidenced by the goodness-of-fit test, all models fail to forecast daily periods. For weekly and two week periods approaches generally underestimate[4] risk. All the same, copula models for $K = 5$ and $K = 10$ and the BS model for $K = 10$ produce reasonable results. Best results are achieved for $K = 5$ days by the Clayton and t-copula. For $K = 20$ all models[5] provide adequate simulation results. Note that only $T = 50$ periods are back-tested. Thus, for monthly returns outcomes are acceptable with limited extent.

On the one hand, VaR results vary in dependence of portfolio size D indicating diversification effects. On the other hand, margins and copula parameters do influence outcomes as moving and added calibrations provide different VaR outcomes. Slightly more shortfalls for moving calibration in comparison to added fitting are observed.

Figure F - 5.1 clearly indicates what we have suspected in Chapter 4. The difference between VaR results of BS model and the elliptical copulas is apparent. The difference between the t- and Gaussian copula is small, whereas their difference to the Clayton copula approach is large. As expected, moving calibration is naturally more sensitive to data changes than the added one.

5.4 Summary (Backtesting Results)
Under Assumption 4.2 and the given backtesting periods, we receive the following implications.

1. Copula approaches outperform the BS model especially for shorter periods ($K = 1, 5, 10$).
2. Backtesting confirms goodness-of-fit tests and VaR simulation results of Chapter 4. Outcomes of Chapter 4 and Chapter 5 are consistent.
3. All models underestimate risk for the daily period.
4. Backtesting delivers acceptable results for monthly periods with the restriction that only $T = 50$ periods are tested.
5. For $K = 5$ and 10 outcomes are reasonable. However, models mainly underestimate risk.
6. Best results are generated by t-copula for weekly periods.

5.5 Remark
Although goodness-of-fit test results show that data is not distributed according to the given copulas (especially for higher dimensional portfolios), VaR backtesting yields quite acceptable results. This is due to the fact that for a confidence level of $\alpha = 0.99$ only 1% of the distribution is needed.

[4] For $K = 5$ and $T = 200$ the Clayton copula overestimates the risk for $D = 20$ within the added and for $D = 25$ within the moving calibration.
[5] Exception: BS model, moving calibration, $D = 10$ and $D = 25$.

6 Summary

Mathematical models simulate possible forecasts for the future. They do not guarantee to predict reality. Therefore, outcomes of models are used as decision support, serve as extensions for surveying market environment, and should be associated to the scope of application. Applying mathematical models means accepting several, sometimes unrealistic assumptions which are attended by difficulties.

6.1 Modeling Assumptions and Barriers

When calibrating models to historical data, one assumes that the past is rich enough to predict future. Of course, this assumption has oftentimes proven false. However, for many approaches in quantitative finance past is understood and used as best estimate for the future. Thus, models are calibrated to historical data[1].

Main assumption for our application is that multivariate distributions of returns are separable by

- marginal distributions of the single asset's return and
- a copula reflecting the dependence structure between margins.

Mathematically, this always holds. In practice, it bypasses the difficulty to fit multivariate distribution functions and provides more flexibility to model dependencies. However, it is quite challenging to estimate the right dependency structure, i.e., the right copula. Various copulas are available, but not all copulas are applicable or suited.

Further, copula calibration is not straightforward. For our application, we suppose that margins can be specified by their empirical distribution functions, and CML calibration can be applied. For this, a large observation matrix is required to receive adequately smooths ecdfs. Generally, this is possible for returns of small periods. For longer periods this may cause problems.

It has to be considered that backtesting approaches rely on assumptions as well. Interpreted as what-if-analysis, they allow to

- assess model performance,
- detect model failures and inconsistencies, and
- monitor model assumptions.

The BS stock price model serves as comparative benchmark as it is often used as standard VaR approach. Not dwelling the topic, the drawbacks and assumptions of this model are found in many textbooks on quantitative finance, e.g., see [49].

6.2 Final Conclusion

We merge conclusions of Chapters 4 and 5 and discuss whether and under which restrictions the presented copula models are applicable for equity risk assessment.

6.2.1 Assumptions for Application Results

We tested a sufficiently long period of $T = 1000$ days[2]. The backtesting period effects the available observation sample to which calibration is performed. For large samples, empirical distribution functions gain more accuracy and but increase computing time. For small samples, the opposite is true. To overcome this problem, one can substitute ecdfs by closed-form margins having the

[1] In contrast, quantitative models are calibrated to current market quotes for pricing issues.

[2] For practical implementation, normally more than one period is backtested to obtain better insight of suitability. As our period spreads over several years, its length may be assumed to be sufficient.

disadvantage that these have to be fitted as well. This has to be balanced. Moreover, using the example of the Clayton copula, the sample size also has impacts on the calibration of one parameter copula models. To some extent, relief for that problem can be achieved by time-varying parameters presented in Chapter 5.

6.2.2 Implementation

Computational time for backtesting is extensive, whereas it is neglectable for calibration and simulation issues. However, many copulas do not have closed-form solutions or cannot be inverted for conditional sampling. Numerical root finding increases computing time and might cause ill-posed or convergence problems, for instance. Thus, application of copula models and the choice of copulas must be weighed regarding their pros and cons with respect to practical implementation.

6.2.3 Numerical Results

Goodness-of-fit tests reject the hypotheses that portfolio returns are distributed according to the specified models for a portfolio size of larger than $D = 15$ and short periods. Concerning periods, these outcomes are reconfirmed by the VaR backtesting analysis.

VaR outcomes are affected by portfolio size. Concerning simulation, diversification effects are observable for $D > 10$. Concerning shortfalls in the backtesting analysis, the impact of diversification is less transparent. Besides, sample size influences calibration of copulas and margins. Main influencing parameter is the dependence structure, i.e., the copula.

Results are reasonable and consistent for all combinations of D and K. As assumed by the goodness-of-fit test, models underestimate portfolio risk in most cases. The shorter the period the larger is the deviation of the expected number of shortfalls. For a monthly period, all models correctly assess the risk for the stated backtest.

The failure of the BS stock price model is significant. The Clayton copula parameter θ is driven by the sample size N and calibrated to approximately 0.6 for large N representing almost independence. Thus, this calibration offers simulation results which are almost equal to independently adding up single risk factors. This fails the focus of risk evaluation. From this aspect, we conclude an additional problem: the model partly overestimates the risk.

6.1 Summary (Main Results)

1. For a confidence level of $\alpha = 0.95$, all models simulate equivalent outcomes. In this case, the BS model has to be preferred due to fast and easy simulation and calibration.
2. For a confidence level of $\alpha = 0.99$, copula approaches outperform the BS model. Crashes and high volatile bear markets are better captured by the copula models compared to BS one. However, they still underestimate risks.
3. The copula concept delivers reasonable and consistent results. Our studies reveal the t-copula to be the best estimate followed by the Gaussian copula. The Clayton copula partly overestimates risks.

6.2.4 Applicability

In practice, one searches for a balance of the right risk measure and manageable implementation without under- or overestimating risks to large extents. Models should be fast and provide consistent as well as acceptable results. t-copula and Gaussian copula are alternatives to the standard BS approach and can be solutions for that balance. Both are easy to simulate. The t-copula produces more accurate results whereas the Gaussian copula is easier to calibrate.

In general, our analysis underlines the known fact that the VaR is a slow responding risk measure. For all models, effects of large impacts are captured with delay.

Dependent on the specific scope of application, copula models generate more adequate risk assessments than the standard approach. Though, this is attended by higher complexity concerning mathematics, calibration and implementation issues.

6.3 Outlook and Literature

Quantitative applications in equity risk management are extensively treated. Available models produce reasonable as well as useable results. However, times of strengthening risk management regulation increase the demand for new and better models.

Concerning copula approaches, future research might deal with time-varying copula concepts allowing for more degrees of freedom. Especially one-parameter-driven copula models benefit from time-dependent parameters. Interconnection of approaches which barely fit together such as combining copulas with stochastic differential equations[3] might be researched in the future. Further, mutual effects of diversification (growing portfolio size) and calibration (influence of margins and copula parameters) as well as the impacts of margins and marginal calibration on copula fitting can be analyzed. Saving computational time by more efficient algorithms for calibration and simulation are always interesting issues for practical purposes. Generally, the analysis of large portfolios may be of interest for the future. By now, copula approaches are mainly applied for dependency modeling of few risk factors.

A mathematical treatment of copulas is Nelsen's monograph "An Introduction to Copulas" [52]. Especially two-dimensional copulas are discussed. For a general multidimensional extension, Joe's book on "Multivariate Models and Dependence Concepts" [37] is recommendable.

For practitioners, Cherubini et al. "Copulas in Finance" [12] give a comprehensive overview for standard applications of a variety of copulas. A detailed copula study on exchange rates is found in Dias dissertation "Copula Inference for Finance and Insurance" [15].

Chapter 5 of McNeil et al. "Quantitative Risk Management" [49] deals with copula applications. Furthermore, it is a good survey on risk management and its applications.

[3] Tankov developed the concept of Lévy copulas in [62].

Part II

Copula Functions in Intensity-Based Credit Risk Models

7 Introductory Note to the Second Part

In credit risk modeling, the specification of dependency structures for risk factors or counterparties might be even more essential as it is for other areas. Besides global market movements like crises, credit risk models are supposed to capture contagion effects such as defaults of other counterparties. A realistic description of these impacts is difficult. Dependent on the scope of application, the copula concept suits as feasible approach for dependence modeling in credit risk[1].

In Chapter 8, we provide an introduction to credit risk considering different types of risk and basic risk factors. Moreover, we classify quantitative models with respect to their purpose, setting, and perspective. Next, we introduce the mathematical background of so-called intensity-based models (Section 9.1) which is enlarged to a multidimensional setup (Section 9.2). Inspired by the approach of Schönbucher and Schubert (Section 9.3, [59]), we develop a new copula- and default-dependent intensity model (Chapter 10). In Chapter 11, model implementation and application are described, simulated outcomes are discussed. The part concludes with a summary in Chapter 12 explaining central methodological and numerical results.

8 Credit Risk

Concerning credit risk in institutions, risk management departments are mainly challenged by

- fulfilling capital requirements,
- controlling (internal) risk figures, and
- pricing issues

in - commonly - regulated frameworks (Basel or Solvency Accords, for instance). Subject of credit risk models are single debtors, portfolios, securities or transactions. Detailed descriptions and definitions of credit risk exposed contracts, products or portfolios are treated in [4], [7], [20], [48], [49] and [58], for instance. Here, we focus on an introduction to credit risk models and their driving risk factors. Following statements are basically collected from the above mentioned monographs.

8.1 Modeling Challenges and Types of Credit Risk

Obviousness of Risk. In direct contracts, credit risk is obviously revealed. If the obligor is not able to pay its obligations, the lender suffers losses. For some credit deals, the inherited risk is of indirect nature. For instance, a credit default swap (CDS) contract guards the protection buyer against failed payments of a reference company. However, the protection seller of this very CDS might be encountered to credit risk, too. In case of a possible default of the reference, he might not be able to absorb the remaining debt. For CDS evaluation, this risk may also be considered by the protection buyer.

Time-Dependence of Risk Factors and Assessable Credit Events. Risk factors depend on systematic and idiosyncratic developments. An incorporation of this time-dependence into a model can be challenging. Credit events such as rating transitions or defaults must be defined and measurable. In [58], these aspects are described as **arrival, timing**, and **recovery** risk.

Dependency Structures. Independence assumptions for risk factors are hardly maintainable for most applications. Impacts of credit events should be included. Schönbucher [58] characterizes these types as **default dependence** or **default correlation** risk. Moreover, **market** and **liquidity** risk factors are possibly linked with credit risk. Modeling these effects and dependencies is significant to the point of **systemic** risk.

[1] As mentioned, the first paper on copulas in credit risk was Li [41] in 2000.

Model Calibration and Validation. Model calibration is generally complex. In some areas, this becomes a core problem as market or internal default data is scarce, for example. Dependent on the scope of application, institutions are further obliged to validate their risk models under consideration of regulatory requirements[1].

8.2 Risk Factors

Credit risk is commonly explained by risk factors that influence the credit quality or creditworthiness of debtors and transactions[2]. The most general setup for measuring credit risk is to define the *loss variable* $L(d)$ for *single obligor* d as

$$L(d) = EAD_d \times LGD_d \times \mathbf{1}_{Default}(d) \quad \text{for } d = 1, \ldots, D \tag{8.1}$$

in which

- EAD_d is the *exposure at default*,
- LGD_d is the *loss given default*, and
- $\mathbf{1}_{Default}(d)$ is the *default indicator* with *default probability*[3]

$$\mathbb{E}[\mathbf{1}_{Default}(d)] = \mathbb{P}[Default(d)] = DP(d) \quad .$$

Aggregation on portfolio view is usually performed by summation of the loss variables in (8.1) and produces a *portfolio loss variable* L[4] as

$$L = \sum_{d=1}^{D} L(d) = \sum_{d=1}^{D} EAD_d \times LGD_d \times \mathbf{1}_{Default}(d) \quad . \tag{8.2}$$

Components can be defined as dependent stochastic processes or variables. They are integrated into credit models in miscellaneous ways and depend on the scope of application, setting, and perspective. Thus, the more complicated risk factors are designed and combined, the more difficult is the model in question.

8.2.1 EAD - The Exposure At Default

As the name indicates, the EAD specifies the exposure at the time the obligor defaults. It comprises *outstandings* and *commitments*. Roughly speaking, outstandings are actual obligations already drawn by the obligor whereas the commitments are the amount the lender has promised to lend, see [7], Subsection 1.1.2.

8.1 Example (EAD)
A private client has a bank account with a limit of 10.000 Euro. Say, he has drawn 7.500 Euro (commitments = 10.000 Euro, outstandings = 7.500 Euro). Now, the client defaults. In this case, the EAD is 7.500 Euro. However, the client could have drawn 10.000 Euro, which would have resulted in an EAD of 10.000 Euro.

The example shows a key problem of EAD estimation. An assessment of the expected outstandings related to the commitments is required **in advance**.

[1] For remarks on the internal validation process see for instance [10] or [55].
[2] Duffie and Singleton [20] describe credit risk as "unexpected changes in the credit quality of issuers or partners". Martin et al. [48] classify default risk and the risk of change in creditworthiness.
[3] The "right" probability measure depends on the problem we have to solve, see Section 8.4 for further explanation.
[4] Naturally, the portfolio loss variable can depend on the chosen segmentation and is applicable to sub-portfolios.

8.2.2 LGD - *The Loss Given Default*

The loss given default is defined as

$$LGD = 1 - R$$

in which R is the *recovery rate*. Thus, the LGD is that fraction of loss which the lender suffers if the borrower defaults. It is dependent on many different factors, e.g., the quality of collaterals[5] or the seniority of the borrowers debt claims. In many approaches, the recovery rate is assumed to be constant to reduce complexity[6].

8.2.3 DP - *The Default Probability*

In contrast to EAD and LGD as conditional variables (a default has occurred), the default probability DP is an unconditional variable. It assigns a probability to a possible default. In several models, it is connected to a *default time*.

Calibration for EAD, LGD, and DP is performed on market or internal data[7]. Depending on the purpose and scope of application, models focus on different risk factors with diverse complexities. Generally, each factor corresponds to a **consistent time horizon**.

8.3 Quantitative Credit Risk Models

Credit risk models are mainly differentiated by three categories:

model purpose	- credit risk management - regulatory capital requirements - credit risk management - bank-internal risk control - analysis of credit risky securities
model setting	- dynamic - static
model perspective	- single obligor - portfolio view

Model Purpose. Credit risk management models basically develop (book) values for balance sheets. In particular, *credit rating models* are used to meet capital requirements[8] predetermined by regulation authorities. Internal credit risk models such as *credit portfolio models* focus the assessment of the institution's risk bearing capacity[9]. They generate figures - e.g., the economic capital, risk limits or risk appetite - for internal risk control issues. To large extents, they are regulated by authorities. As further purpose, models are used to analyze and to price credit risky securities or derivatives.

Model Setting. Dynamic - in sense of a continuous-time framework - models are principally used for the assessment of credit risky securities and derivatives[10]. Static models are generally applied in credit risk management. Credit rating models are stationary for a fixed time horizon. Risk factors are typically estimated for one year. In addition, mixture models exist. Especially internal credit risk and portfolio models often incorporate both static and dynamic components.

[5] Collaterals are assets and / or securities which collateralize deals.
[6] Casually, one finds the notion of the *severity of loss*. The LGD is then defined as the expectation of a random variable describing the severity of the loss, i.e., $LGD = \mathbb{E}[SEVERITY]$.
[7] Internal data is mostly available for large portfolios with sufficient default occurrences, e.g., retail portfolios.
[8] Basel II / III Accords, Pillar I, Minimum Capital Requirement - In a nutshell, the regulatory capital is defined as reserves that have to be provided for each customer / portfolio open to credit risk.
[9] See Basel II / III Accords, Pillar II, Supervisory Review Process.
[10] In a dynamic setting, credit events are often designed as stopping times.

Model Perspective. Models for single debtors (e.g., rating models) aim to assess characteristic features of one obligor. Here, risk factors are commonly separated. A portfolio view (e.g., portfolio models) tends to capture interdependencies of sub-portfolios / portfolio members. It focuses on the aggregation of risk factors. Regarding pricing models, the perspective depends on the underlying product (e.g., single CDS or baskets).

8.3.1 Models for Loss Distributions

Models for loss distributions are usually applied for credit risk management purposes and analyze the portfolio loss variable L defined in Equation (8.2). Its corresponding cdf F_L is usually denoted as **portfolio loss distribution**. For modeling F_L, two approaches are popular, the *empirical* and the *analytical approximation*. For the empirical estimation, financial institutions use historical losses to generate F_L. Estimated losses are thus produced by simulations on the basis of former data. For the analytical approximation, one fits parameters of particular distributions to loss data and receives an analytical loss cdf. Mixture models exist here as well.

Regarding portfolio models in particular, the essential problem is to incorporate and to reproduce dependency structures. In practice, known credit risk portfolio models are CreditMetrics^{TM} by RiskMetrics / JP Morgan & Chase, CreditRisk$^+$ by Credit Suisse, and KMV maintained by Moody's[11]. Other approaches are rather treated in literature, see [7], Chapter 2.

8.3.2 Additional Risk Figures

The *expected loss* EL_d for a single obligor / EL for a portfolio is the expectation of the loss variable L_d / L of Equation (8.1) / Equation (8.2), i.e.,

$$\begin{aligned} EL_d &= \mathbb{E}\left[L_d\right] = \mathbb{E}\left[EAD_d \times LGD_d \times \mathbf{1}_{Default}(d)\right] \text{ and} \\ EL &= \mathbb{E}[L] = \mathbb{E}\left[\sum_{d=1}^{D} EAD_d \times LGD_d \times \mathbf{1}_{Default}(d)\right] \text{ respectively} \quad . \end{aligned}$$

EL_d and EL are used as regulatory and internal credit risk figures. Different definitions exist for the *unexpected loss* UL_d and UL. They depend on the relevant purposes for which the unexpected loss is determined. For the regulatory capital requirement[12], the UL_d is generally defined as a function of DP_d, LGD_d and EAD_d. Regarding internal management purposes, the UL is basically a predefined quantile of the loss distribution F_L[13] which leads to the concept of the *credit value at risk* ($CVaR$).

Hazardous risks for financial institutions are located in the exceedance of a certain quantile of the loss distribution. As capital is already reserved for expected losses, the computation of the $CVaR$ combines the respective quantile and the EL. Precisely, the $CVaR_\alpha(L)$ is generally defined as

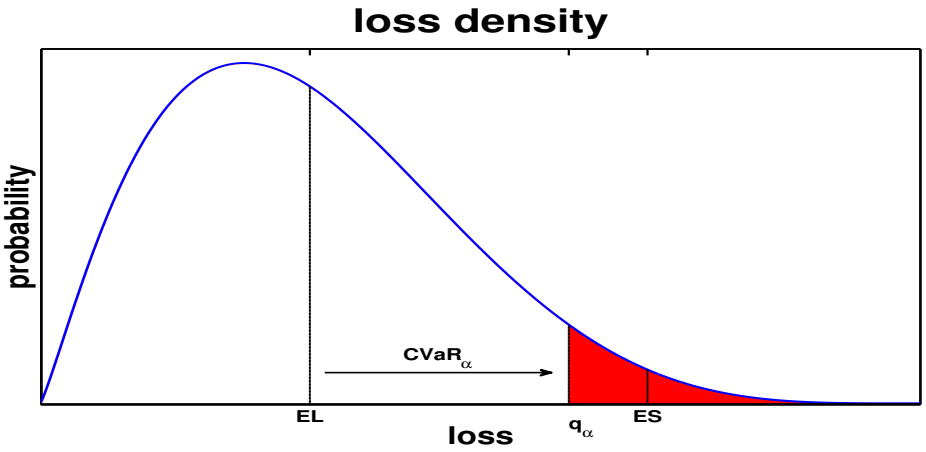
$$CVaR_\alpha(L) = q_\alpha(L) - EL$$

in which $\alpha \in (0,1)$ is the confidence level and $q_\alpha(L)$ is the quantile function of the loss L, i.e.,

$$q_\alpha(L) = \inf\left\{x \geq 0 | \mathbb{P}[L \geq x] \leq \alpha\right\} \quad . \tag{8.3}$$

8.2 Remark
In contrast to equity risk, credit risk losses are mainly defined as positive values. The credit value at risk $CVaR$ should not be mixed up with the equity value at risk VaR, Definition 4.1.

[11] For details, the reader is referred to the respective model documentation.
[12] In the Basel-II framework the UL_d for the individual obligor is estimated by the internal-rating-based formula, e.g., see [11] or [31].
[13] In some monographs, the unexpected loss is set as the standard deviation of the loss cdf which is commonly not used in practice.



F - 8.1 Figure: exemplary loss density with key risk figures

8.3 Notation (Loss Distribution, Economic Capital, and Expected Shortfall)

1. The *loss distribution function* F_L is the cdf of the loss variable L.
2. The credit value at risk $CVaR$ is often denoted as *unexpected loss*. The *economic capital* is the capital that covers the $CVaR$.
3. The *expected shortfall* or *expected tail loss*[14] of L with $\mathbb{E}[L] < \infty$ is given as

$$ES_\alpha = \frac{1}{1-\alpha} \int_\alpha^1 q_u(L)du = \mathbb{E}\left[L|L \geq q_\alpha(L)\right]$$

in which $\alpha \in (0,1)$ is a given confidence level. The last equivalence only holds for loss variable with continuous cdf F_L[15].

Figure F - 8.1 depicts a possible density function, the expected loss EL, the unexpected loss $CVaR = UL$, and the expected shortfall ES_α of an exemplary loss variable L.

8.3.3 Models for Default Probabilities

As central risk driver, various approaches of modeling default probability models have been worked out and studied. We picture a brief summary of existing DP models. Sound surveys are found in [20] or [49]. One categorizes

- *structural / firm-value*,
- *reduced-form* (also called *intensity-based* or *hazard rate*), and
- *mixture* models.

Structural or firm-value models "attempt to explain the mechanism by which default takes place", see [49], p. 331. Its foundation was laid by Merton [51] in 1974. He identifies a default if the company's liabilities are larger than the firm's assets at the end of a period[16]. Gradually, Merton's approach was extended - for instance - by:

- the model of Zhou

  a Merton model incorporating jumps

[14] *Conditional tail loss* or *conditional tail expectation* are conventional terms as well.
[15] The expected shortfall is sometimes defined as $ES_\alpha - \mathbb{E}[L]$.
[16] By means of these assumptions, Merton uses parts of the option pricing theory.

- first passage time or threshold models including copulas

  A default happens if a certain limit is exceeded or a modeling variable falls below a specified value.

- rating transitions or credit migration models

  Markov chain models indicate the issuer's creditworthiness or default by state variables.

Basically, structural approaches are static and applied for credit risk management purposes with some exceptions. Further, they provide the basis for many portfolio models.

In reduced-form models, an exogenous specified process - the default intensity - is used for default prediction. In contrast to structural models, it does not specify the mechanism of default, it is random. Credit migration models display the evolution of this process for the creditworthiness, whereas intensity-based ones only exhibit possible defaults. Commonly, reduced-form approaches are used for analysis and valuation of credit derivatives and thus belong to the class of dynamic schemes[17]. Further, one fundamental modeling aspect in this area is the underlying information.

As above, mixture models combine elements of different approaches, e.g., firm-value and reduced-form models or various approaches in the model itself. They can be both static or dynamic.

8.4 Risk-Neutral vs. Real World

Two general approaches exist to assess credit risk affected securities or products,

- the *risk-neutral* (also *financial* or *martingale*) pricing and
- the *real-world* (also *objective*, *historical* or *actuarial*) evaluation.

The significant difference between the risk-neutral and the real-world is the probability measure under which valuation is performed.

8.4.1 Risk-Neutral Probabilities

We denote a risk-neutral measure and default probabilities by $\mathbb{Q}$ and $DP_{\mathbb{Q}}$. In this world, prices of credit risky securities are expected discounted payoffs under $\mathbb{Q}$, i.e.,

$$\text{price} = \mathbb{E}_{\mathbb{Q}}[\text{discounted payoff}] \quad . \tag{8.4}$$

The measure $\mathbb{Q}$ causes any discounted payoff to be a $\mathbb{Q}$-martingale and produces *artificial* or *market implied* default probabilities $DP_{\mathbb{Q}}$. This means that investors are indifferent to risk. There is no need for risk compensation as the expected return on any payoff is the risk-free rate.

By the Fundamental Theorems of Asset Pricing, see Theorems B.11 and B.12 in Appendix B.3, we state that for a complete* model the measure $\mathbb{Q}$ is unique. For incomplete markets, this uniqueness is not given and it can be challenging to identify the right probability measure, see [49], Example 9.18. In practice, $DP_{\mathbb{Q}}$ can be linked to credit spreads, see [7], p. 187, or [49], p. 402. This relationship is useful for calibration.

8.4 Remark

1. Risk-neutral default probabilities are calibrated to current market quotes and spread data.
2. In credit risk, incomplete markets exists in a mathematical sense. This poses the problem of identifying the right risk-neutral probability measure.

[17] As classical paper the reader is referred to [19].

8.4.2 Real-World Probabilities

We denote the real-world measure and default probabilities as $\mathbb{P}$ and $DP_{\mathbb{P}}$. In this world, prices are computed as expected payoff under $\mathbb{P}$ plus a risk premium. This risk premium compensates the default risk of the issuer. Actuarial models basically price illiquid and non-traded credit securities, see [49], p. 401. They are also applied for credit risk management purposes.

Real-world default probabilities are calibrated to historical data. In contrast to risk-neutral probabilities, fitting is straightforward. The difference between these approaches is best explained by means of an example.

8.4.3 Different Default Probabilities - Example

Suppose that we have a defaultable bond with a coupon of 10% expiring in one year. Its actual nominal value is 100. Further, we assume that the risk-free rate is 5%, the historical default probability $DP_{\mathbb{P}}$ is 5%, and the recovery rate R in case of default is 40%.

At maturity, we receive 110 or 40 if the bond defaults. Discounting the payoff by the risk-free rate of 5% we get:

$$\frac{(1-0.05)\times 100\times 1.1+0.05\times 100\times 0.4}{1.05}=\frac{106.5}{1.05}=101.43 \quad .$$

However, the bond's actual price is 100. In our computation values are assumed to be certain. Thus, no risk premium is implied. We overstate the real price by 1.43, the risk premium. The solution $DP_{\mathbb{Q}}$ of

$$100=\frac{(1-DP_{\mathbb{Q}})\times 100\times 1.1+DP_{\mathbb{Q}}\times 100\times 0.4}{1.05}$$

gives a risk-neutral default probability of

$$DP_{\mathbb{Q}}\approx 0.0714>0.05=DP_{\mathbb{P}} \quad .$$

Interpretation

The market implied default probability is $DP_{\mathbb{Q}}=0.0714$. This is the same as a bond break even trade with the above given properties and a $DP_{\mathbb{P}}=DP_{\mathbb{Q}}=0.0714$. If all investors are indifferent to risk, they take over the higher default probability of 0.0714 and get the equivalent current price of 100 - the risk-neutral approach. If market participants assume $DP_{\mathbb{P}}$ to be 0.05, they demand for risk premium for default of 1.43 - the real-world approach.

9 Copulas in Dynamic Intensity-Based Models

In a nutshell, we derive the mathematical framework for dynamic intensity-based valuation of credit risky securities by means of Cox processes primarily introduced by Jarrow and Turnbull [36], Madan and Unal [45] and Lando [40]. This approach is often called *information-based modeling* as valuation largely depends on the underlying filtration.

In this dynamic setup, we describe the approach of Schönbucher and Schubert [59] who are the first presenting a continuous copula-dependent default model[1]. For a more detailed description of dynamic credit risk models, the reader is referred to [4], for instance.

[1] Li [41] introduced copula functions in a structural approach.

9.1 Mathematical Background

As basis for modeling, basic definitions and notations are indispensable. Valuation takes place on a fixed time horizon $[0,T]$ with $T<\infty$ and a filtered[2] probability space $(\Omega,\mathcal{H},(\mathcal{H}_t)_{t\in[0,T]},\mathbb{Q})$ with $\mathcal{H}=\mathcal{H}_T$ and a unique risk-neutral measure $\mathbb{Q}$. Due to the uniqueness of $\mathbb{Q}$, the market is supposed to be free of arbitrage*, see Theorems B.11 and B.12.

The mathematical construction of a filtration symbolizes the flow of market information[3]. The $\mathcal{H}_t$-measurability (or $\mathcal{H}_t$-adaptivity*) of a stochastic process X_t means that if the market participant is endowed with the information $\mathcal{H}_t$ at time $t\in[0,T]$, the value at t - and therefore the path from 0 to t - of the process X_t is known $\mathbb{Q}$-almost surely.

9.1 Notation

1. $\sigma\left(X_s:0\le s\le t\right)$ is the *σ-algebra* generated* by the stochastic process X_t. By construction, X_t is $\sigma\left(X_s:0\le s\le t\right)$-adapted.
2. All expectations are computed under the risk neutral measure $\mathbb{Q}$. For sake of readability, we omit the index $\mathbb{Q}$, i.e., $\mathbb{E}_{\mathbb{Q}}[X_t]=\mathbb{E}[X_t]$.
3. For a *default-free* (synonymously *risk-free* or *non-defaultable*) security or stochastic process, the risk of default is negligible. *Defaultable* securities or processes are exposed to default risk.

The intensity-based setup and its evaluation is highly dependent on the underlying information flow. In a first instance, we introduce the default-free filtration and pass on to the construction of default times. Next, we define the default information flow and describe its properties.

9.1.1 Default-Free Information

We suppose that market information of all default-free processes are united in one process Y_t. This process illustrates an incomplete state of economy as any default relevant information is not included.

9.2 Definition (Background Filtration)

1. The process Y_t, $t\in[0,T]$, is called *background process.*
2. The corresponding filtration $\mathcal{G}_t:=\sigma\left(Y_s:0\le s\le t\right)\subset\mathcal{H}_t$ is denoted as *background filtration.*
3. $\mathcal{G}=\mathcal{G}_T\subset\mathcal{H}$ is the *full background filtration.* It comprises all economic risk-free information of the whole interval $[0,T]$.

Consequently, all risk-free processes are $\mathcal{G}_t$- and hence $\mathcal{H}_t$-adapted. Further, discounting is performed by a risk-free short rate r_t defined as $\mathcal{G}_t$-adapted stochastic process[4].

9.3 Definition (Discount Factor)
Let r_t be the short rate process. The *stochastic discount factor* between two time instants t and T is defined as

$$D(t,T)=\exp\left(-\int_t^T r_s ds\right) \quad .$$

[2] All introduced filtrations are supposed to fulfill the usual conditions, see Definition B.1.
[3] *Information flow* and *information* are also used as expressions for the term filtration.
[4] For a survey on interest rate models, the reader is referred to [9].

9.4 Example
Due to Equation (8.4) and under information $\mathcal{H}_t$, the price for a **non-defaultable** zero-coupon bond $B(t,T)$ at time t with time to maturity T and payoff 1 is

$$B(t,T) = \mathbb{E}[D(t,T) \times 1 | \mathcal{H}_t] = \mathbb{E}\left[\exp\left(-\int_t^T r_s ds\right) | \mathcal{H}_t\right] = \mathbb{E}\left[\exp\left(-\int_t^T r_s ds\right) | \mathcal{G}_t\right] \quad .$$

The last identity holds as exclusively risk-free processes (namely r_t) are involved, i.e., filtration $\mathcal{H}_t$ can be replaced by $\mathcal{G}_t$.

9.1.2 Default Time Construction

Returning to Equation (8.1), we assume that there is no recovery ($R = 0 \Leftrightarrow LGD = 1$) and exposure is set to one ($EAD = 1$). Thus, the loss variable holds $L = \mathbf{1}_{Default}$ and only comprises the default probability. As we are interested whether default has already occurred at some time $t \in [0,T]$, the default indicator must contain both the default probability and the time of default.

9.5 Notation
1. We denote the *default time* by τ.
2. *Survival* and *default indicator* are the functions $\mathbf{1}_{\{\tau > t\}}$ and $\mathbf{1}_{\{\tau \leq t\}}$.

One crucial point for evaluation is the specification of the default time. Here, we construct default times via Poisson and Cox processes. The following derivation and framework can also be received by different approaches such as hazard rate functions or doubly stochastic random times, see [4] for instance. A comprehensive construction of Poisson processes by exponential distributed random variables is found in [14], Section 2.5. For a more intuitive generation the reader is referred to [58], Section 5.1 and 5.2.

9.6 Definition (Poisson Process)
A non-decreasing integer-valued jump process (or counting process) $N_t : \Omega \times [0,T] \to \mathbb{N}$ with $N_0 = 0$ is called *Poisson process* with constant *intensity* $\lambda > 0$ if

1) it has independent increments and

2) the probability of n jumps, $n \in \mathbb{N}$, in any interval $[t,T]$ holds

$$\mathbb{Q}[N_T - N_t = n] = \frac{1}{n!}\lambda^n (T-t)^n \exp(-(T-t)\lambda) \quad . \tag{9.1}$$

The jump size of a Poisson process is one. Inter-arrival times - the time between two jumps - are exponentially distributed with parameter λ, see [58], Subsection 5.1.3. If the intensity is not constant but time-dependent, Equation (9.1) changes to

$$\mathbb{Q}[N_T - N_t = n] = \frac{1}{n!}\left(\int_t^T \lambda(s) ds\right)^n \exp\left(-\int_t^T \lambda(s) ds\right) \quad .$$

Then, N_t is an *inhomogeneous Poisson process*. Adding a random component $\omega \in \Omega$ to the intensity, $\lambda_t = \lambda_t(\omega)$ becomes a stochastic process - the *intensity process*. This leads to

9.7 Definition (Cox Process)
A *Cox process* (or *doubly stochastic process*) is a stochastic process N_t with stochastic intensity process $\lambda_t > 0$ if - conditioned on a realized path $\tilde{\lambda}_t$ of λ_t - N_t is an inhomogeneous Poisson process with intensity $\tilde{\lambda}_t$.

The definition seems clear at first sight. Knowing a realization of the intensity process $\tilde{\lambda}_t$, we are dealing with a inhomogeneous Poisson process. At a second glance, it is unusual as two sources of randomness are merged: the intensity process λ_t and the jump process N_t. Details on Cox processes are found in [32], for instance.

9.8 Remark

1. For intensity-based models and the later presented Schönbucher-Schubert framework (Section 9.3), the intensity process λ_t of N_t has to be non-negative and $\mathcal{G}_t$-adapted. Due to these properties, the theory of stochastic analysis shows that the process

$$M_t = N_t - \int_0^t \lambda_s ds$$

   is a martingale[5]. The integral is called *predictable compensator* of N_t. It can also be associated to a *compensator measure*. Vice versa, the process N_t can be generated by its compensator measure. A detailed and mathematically profound description for these aspects go beyond the scope of this thesis. For more information, the reader is referred to [4] or [58] for a more practical and to [35] and [44] for an in-depth mathematical treatment.

2. For our construction (Chapter 10), we demand the intensity λ_t to be positive due to economic as well as mathematical reasons. In intensity models, the process λ_t can be connected to credit spreads through default probabilities, e.g., see [9], Subsection 21.3.6. In an economic way of thinking, a zero (or negative) credit spread is generally associated to a riskless asset with no default probability and consequently zero (or negative) intensity process λ_t. From a mathematical point of view, model construction fails for a zero or negative intensity, i.e., the process N_t cannot jump.

Next, Cox and intensity processes are used for the construction of default times.

9.9 Assumption (Default Time)
The default time τ is defined as the first jump of a Cox process N_t with intensity process λ_t, mathematically

$$\tau := \inf\{t \in [0,T] : N_t > 0\} \quad . \tag{9.2}$$

The set $\{\tau > t\}$ reflects the path of N_t, i.e., we have the equivalence $\tau > t \Leftrightarrow N_t = 0$.

The setup for default times is resolved by Equation (9.2), but its generation or better its simulation remains to be clarified. Lando [40] observed that the time of the first jump of a Cox process with intensity λ_t can be constructed as

$$\tau = \inf\left\{ t \in [0,T] : \int_0^t \lambda_s ds \geq E \right\} \tag{9.3}$$

in which E is an exponentially distributed random variable. By $E \sim -\ln(U)$ with $U \sim U(0,1)$, we transform (9.3) to

$$\tau = \inf\left\{ t \in [0,T] : \gamma_t := \exp\left(-\int_0^t \lambda_s ds \right) \leq U \right\} \tag{9.4}$$

and denote the random variable U as *default trigger* and γ_t as *default countdown process.*

[5] This also holds for a general counting process with intensity λ_t.

9.1.3 Information-Based Modeling

As key element for a consistent model, the two different sources of randomness of the Cox process are separated and evaluated in dependence of corresponding filtrations. Therefore, the subsequent framework is often denoted as information-based modeling approach.

9.10 Definition (Default Filtration)
The *default filtration* $\mathcal{F}_t$ is generated by the Cox process N_t, mathematically

$$\mathcal{F}_t := \sigma\left(N_s : 0 \leq s \leq t\right) = \sigma\left(\{\tau \leq s\} : 0 \leq s \leq t\right) = \sigma\left(\{\tau > s\} : 0 \leq s \leq t\right) \subset \mathcal{H}_t \quad .$$

As a consequence, N_t is always $\mathcal{F}_t$- and hence $\mathcal{H}_t$-adapted. Respectively, the set $\{\tau > t\}$ is always $\mathcal{F}_t$- and $\mathcal{H}_t$-measurable. The filtration $\mathcal{F}_t$ comprises all information about defaults up to time $t \in [0, T]$. Similar to the background filtration $\mathcal{G}_t$, $\mathcal{F}_t$ can be considered as incomplete state of economy. In this case, risk-free information is not included.

9.11 Assumption (Information Modeling)

1. The *full filtration* $\mathcal{H}_t$ unites default and background filtration, i.e., $\mathcal{H}_t := \mathcal{F}_t \vee \mathcal{G}_t$.
2. The stochastic intensity λ_t of the Cox process N_t is supposed to be positive-valued and $\mathcal{G}_t$-adapted, see also Remark 9.8.
3. The default trigger U is independent of the full background filtration $\mathcal{G}$.

Assumption 9.11 is the foundation pillar of the information-based modeling in intensity models. The process $Z_t := (Y_t, N_t)$ reflects the complete state of economy symbolized by $\mathcal{H}_t$. The full information holds $\mathcal{H}_t = \sigma\left(Z_s : 0 \leq s \leq t\right)$.

As $\mathcal{G}_t$ does not incorporate default information, N_t cannot be $\mathcal{G}_t$-measurable. In contrast, Y_t cannot be $\mathcal{F}_t$-measurable because no default information from N_t is included. Note that - from economical point of view and in practice - the background filtration is influenced by default information and vice versa.

The principle of this setup aims for a conditional evaluation under the different filtrations. Suppose a realized path of λ_t on $[0, T]$ is given. According to Assumption 9.11, this information is incorporated in filtration $\mathcal{G}$. Thus, conditioning on $\mathcal{G}$, λ_t is deterministic.

The probability that the Cox process N_t with intensity process λ_t has n jumps in the interval $[t, T]$ is generally computed as

$$\mathbb{Q}[N_T - N_t = n] = \mathbb{E}\left[\mathbb{Q}[N_T - N_t = n|\mathcal{G}]\right] = \mathbb{E}\left[\frac{1}{n!}\left(\int_t^T \lambda_s ds\right)^n \exp\left(-\int_t^T \lambda_s ds\right)\right] \quad .$$

This equation evidences why λ_t **must** be $\mathcal{G}_t$-adapted. Conditioning on $\mathcal{G}$ is the recipe for solving problems in an intensity-based setup. It includes two steps:

1. Solve the problem on knowledge of $\mathcal{G}$. All expressions depending on default relevant processes are replaced by $\mathcal{G}$-adapted processes.
2. Take the expectation over all possible paths.

9.12 Lemma (Survival Probability)
Let $t \in [0, T)$ be given and $\tau > t$, i.e., the asset has not defaulted up to time t. Then, the survival probability from t to T is computed as

$$P(t, T) := \mathbb{E}\left[\mathbf{1}_{\{\tau > T\}}|\mathcal{H}_t\right] = \mathbb{E}\left[\exp\left(-\int_t^T \lambda_s ds\right)|\mathcal{G}_t\right] \quad .$$

A proof[6] can be found in Appendix D.2.1. ∎

[6] The proof is highly recommended to readers which are not familiar with the setup.

9.13 Example (Pricing a Defaultable Bond)
Let $\bar{B}(t,T)$ be the price of a zero-coupon, defaultable bond with maturity T, no recovery, and $EAD = 1$ at time t. Equation (8.4) and Notation 9.5 tell us that

$$\bar{B}(t,T) = \mathbb{E}\left[D(t,T) \cdot \mathbf{1}_{\{\tau>T\}} | \mathcal{H}_t\right].$$

If $\{\tau \leq t\}$, it is $\mathbf{1}_{\{\tau>T\}} = 0$ and $\mathbb{E}\left[D(t,T) \cdot \mathbf{1}_{\{\tau>T\}} | \mathcal{H}_t\right] = 0$. With $\{\tau > t\} \in \mathcal{H}_t$, we have

$$\mathbb{E}\left[D(t,T) \cdot \mathbf{1}_{\{\tau>T\}} | \mathcal{H}_t\right] = \mathbf{1}_{\{\tau>t\}} \mathbb{E}\left[D(t,T) \cdot \mathbf{1}_{\{\tau>T\}} | \mathcal{H}_t\right] \quad .$$

With Lemma 9.12 and $\mathcal{G}_t \subset \mathcal{H}$ it holds

$$\begin{aligned}\bar{B}(t,T) &= \mathbf{1}_{\{\tau>t\}} \mathbb{E}\left[D(t,T) \cdot \mathbf{1}_{\{\tau>T\}} | \mathcal{H}_t\right] = \mathbf{1}_{\{\tau>t\}} \mathbb{E}\left[\exp\left(-\int_t^T r_s + \lambda_s ds\right) | \mathcal{H}_t\right] \\ &= \mathbf{1}_{\{\tau>t\}} \mathbb{E}\left[\exp\left(-\int_t^T r_s + \lambda_s ds\right) | \mathcal{G}_t \vee \mathcal{F}_t\right] \quad .\end{aligned}$$

To get rid of the conditioning filtration $\mathcal{F}_t$, we exploit the proof of Lemma 9.12, see Appendix D.2.1. The default trigger U is independent of the filtration $\mathcal{G}$. This in turns means that U is independent of the filtration $\sigma\left(\exp\left(-\int_t^T r_s + \lambda_s ds\right)\right) \vee \mathcal{G}_t$. It can be shown that

$$\mathbb{E}\left[\exp\left(-\int_t^T r_s + \lambda_s ds\right) | \mathcal{G}_t \vee \sigma(U)\right] = \mathbb{E}\left[\exp\left(-\int_t^T r_s + \lambda_s ds\right) | \mathcal{G}_t\right] \quad ,$$

see [65]. Further, the inclusion $\mathcal{G}_t \subset \mathcal{G}_t \vee \mathcal{F}_t \subset \mathcal{G}_t \vee \sigma(U)$ leads to

$$\mathbb{E}\left[\exp\left(-\int_t^T r_s + \lambda_s ds\right) | \mathcal{G}_t \vee \mathcal{F}_t\right] = \mathbb{E}\left[\exp\left(-\int_t^T r_s + \lambda_s ds\right) | \mathcal{G}_t\right] \quad .$$

Thus, the price of the defaultable bond is

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

Comparing Example 9.4 with Example 9.13, we see that on the one hand $B(t,T)$ is risk-free and never defaults. Its return is simply the interest rate r_t. On the other hand, the defaultable bond is a risky asset. In financial markets, returns of a risky investment are generally larger than those of a risk-free strategy. This risk premium is compensated by the additional component $\lambda_t > 0$. Its return is $r_t + \lambda_t > r_t$. From the issuer's point of view, his default risk is punished by λ_t. Borrowing money gets more expensive.

9.14 Summary

1. The implementation of jump times of the Cox process N_t is realized by the connection of the default trigger U and the intensity process λ_t.
2. The separation of non-defaultable and defaultable information is the key element for evaluation in this setup.

We conclude with two algorithms for default time simulation on a fixed time grid, illustrated by [58], Section 7.7.1.1. The intensity process λ_t is assumed to be given as stochastic differential equation (SDE) discretized by some computation scheme[7] to $\lambda_{n\Delta t}$ for $n = 1, \ldots, N$ with step size $\Delta t = \frac{T}{N}$, $N \in \mathbb{N}$.

A - 9.1 Algorithm (Default Time Simulation)

Given

- initial value $\lambda_0 > 0$ and SDE parameters

Step 1 - for $n = 1, \ldots, N$

- draw $u_n \sim U(0,1)$[8]
- compute $\lambda_{n\Delta t}$ and $\exp(-\lambda_{n\Delta t}\Delta t)$
- compare $\exp(-\lambda_{n\Delta t}\Delta t) \leq u_n \quad \Rightarrow \quad$ default!
  - set $\lambda_{k\Delta t} = \lambda_{n\Delta t}$ for $k = n, \ldots, N$
  - set $\tau = n$
  - skip $n = N$
- compare $\exp(-\lambda_{n\Delta t}\Delta t) > U_n \quad \Rightarrow \quad$ no default!
  - set $n = n + 1$

A - 9.2 Algorithm (Direct Default Time Simulation)

Given

- initial value $\lambda_0 > 0$ and SDE parameters
- set $\gamma_0 = 1$

Step 1

- draw $u \sim U(0,1)$

Step 2 - for $n = 1, \ldots, N$

- compute $\lambda_{n\Delta t}$ and $\gamma_{n\Delta t} = \gamma_{(n-1)\Delta t} \exp(-\lambda_{n\Delta t}\Delta t)$
- compare $\gamma_{n\Delta t} \leq u \quad \Rightarrow \quad$ default!
  - set $\gamma_{k\Delta t} = \gamma_{(n-1)\Delta t}$ for $k = n, \ldots, N$
  - set $\tau = n$
  - skip $n = N$
- compare $\gamma_{n\Delta t} > u \quad \Rightarrow \quad$ no default!
  - set $n = n + 1$

Algorithm A - 9.1 is computationally more expensive as default triggers are simulated in each time step. However, it also seems to be more intuitive on a fixed time grid than Algorithm A - 9.2, see [58], pp. 215-216.

[7] Euler or Milstein scheme, for instance.
[8] Realizations u_n, $n = 1, \ldots, N$ are i.i.d.

9.2 The Multidimensional Model Setup

In the previous section, evaluation is performed for **one-dimensional** problems. Products such as baskets of defaultable assets or credit debt obligations (CDO's) are of **multidimensional** nature. Expansion of the model setup is straightforward.

For stochastic processes etc., the subscript for time-dependence is put to the lower right, the index for the dimension to the upper right of the vector, i.e.,

$$x_t = (x_t^1, \ldots, x_t^D)'.$$

Mathematical comparisons like $x_t > a$ are meant to be component-wise, i.e., $(x_t^1 > a, \ldots, x_t^D > a)'$.

9.15 Assumption (Multidimensional Model)
We have $D \in \mathbb{N}$ obligors or assets. We denote

1. the $\mathcal{G}_t$-adapted default intensity as $\lambda_t = (\lambda_t^1, \ldots, \lambda_t^D)' > 0$,
2. the default countdown process as
$$\gamma_t = (\gamma_t^1, \ldots, \gamma_t^D)' = \left(\exp\left(-\int_0^t \lambda_s^1 ds \right), \ldots, \exp\left(-\int_0^t \lambda_s^D ds \right) \right)' \quad ,$$
3. the default trigger as $U = (U^1, \ldots, U^D)' \perp \mathcal{G}$,
4. the default time as $\tau = (\tau^1, \ldots, \tau^D)'$ with $\tau^d := \inf\{t \in [0,T] : \gamma_t^d \leq U^d\}$, $d = 1, \ldots, D$,
5. the default filtration as $\mathcal{F}_t^d = \sigma\left(\{\tau^d > s\} : 0 \leq s \leq t\right)$,
6. the *full default filtration* as $\mathcal{F}_t = \bigvee_{d=1}^{D} \mathcal{F}_t^d$, and
7. the *partial* filtration as $\mathcal{H}_t^d = \mathcal{G}_t \vee \mathcal{F}_t^d$, $d = 1, \ldots, D$.

Remark that the full information still is $\mathcal{H}_t = \mathcal{G}_t \vee \mathcal{F}_t$. The $\mathcal{G}_t$-adaptivity for the process λ_t and the $\mathcal{G}$-independence of the default trigger U are just extended to a multivariate setting.

Several defaults can now occur on the valuation horizon $[0,T]$. In order to distinguish the defaulted obligors from the survivors, we introduce the following

9.16 Notation (Default Index Sets, Default Times and Defaulted Vectors)
Let $\mathcal{D} = \{1, \ldots, D\}$ be the index set of all positions / obligors.

1. The index set $\mathcal{E}_t$ denotes the *non-defaulted* obligors at time $t \in [0,T]$:
$$\mathcal{E}_t := \{d_1, \ldots, d_E | d_e \in \mathcal{D}, e = 1, \ldots, E\}$$
2. The index set $\mathcal{C}_t$ denotes the *defaulted* obligors at time $t \in [0,T]$[9]:
$$\mathcal{C}_t := \{\tilde{d}_1, \ldots, \tilde{d}_C | \tilde{d}_c \in \mathcal{D}, c = 1, \ldots, C\}$$
3. The vector of **potential default times** of the non-defaulted positions is defined as
$$\tau(\mathcal{E}_t) = \left(\tau^{d_1}, \ldots, \tau^{d_E}\right)'$$
with $\tau^{d_e} \in [t,T]$, $d_e \in \mathcal{E}_t$.

[9] It holds $\mathcal{E}_t \cup \mathcal{C}_t = \mathcal{D}$, $\mathcal{E}_t \cap \mathcal{C}_t = \emptyset$ for all $t \in [0,T]$ and $E + C = D$.

4. The vector of **realized default times**[10] is defined as

$$\tau(\mathcal{C}_t) = \left(\tau^{\tilde{d}_1}, \ldots, \tau^{\tilde{d}_C}\right)'$$

with $\tau^{\tilde{d}_c} \in [0,t]$, $\tilde{d}_c \in \mathcal{C}_t$.

5. A **non-defaulted vector** is denoted as

$$x_s(\mathcal{E}_t) = (x_s^{d_1}, \ldots, x_s^{d_E})$$

at time $s \in [t,T]$ with indices in $\mathcal{E}_t$.

6. A **defaulted vector** is denoted as

$$x_\tau(\mathcal{C}_t) = \left(x_{\tau^{\tilde{d}_1}}^{\tilde{d}_1}, \ldots, x_{\tau^{\tilde{d}_1}}^{\tilde{d}_C}\right)'$$

with indices in $\mathcal{C}_t$ and default times $\tau(\mathcal{C}_t)$ at time $t \in [0,T]$.

The challenge for the multivariate setting is to incorporate dependencies. Financial crises evidence that products or assets do mutually influence each other, particularly in case of defaults of other companies, countries or securities. Capturing these dependencies and impacts are often described as *default contagion*.

Many approaches have been developed to model these contagion effects. The first who use copula functions in a dynamic intensity-based configuration are Schönbucher and Schubert [59] in 2001.

9.3 The Schönbucher - Schubert Framework

Schönbucher and Schubert's main idea is to introduce a dependency structure to default times via copulas. In the multidimensional setup, each countdown process is compared to a default trigger which is uniformly distributed, i.e., $\gamma_t^d \leq U^d$, $d = 1, \ldots, D$. The Schönbucher-Schubert approach assumes that these default triggers to be distributed according to a copula, i.e.,

$$U = (U^1, \ldots, U^D)' \sim C \quad . \tag{9.5}$$

If a default occurs, the model locks the countdown process γ_t^d for $t \in [\tau^d, T]$. This implies that the default trigger U^d is set to γ_t^d and a conditional multivariate probability (i.e., a conditional copula) must be evaluated. Hence, the copula C is assumed to be continuously differentiable. Due to Assumption 9.15, the variable $U \sim C$ is independent of the background filtration $\mathcal{G}$. Naturally, Equation (9.5) reflects default contagion underlined by the following result.

9.17 Proposition (compare Propositions 4.3 and 4.6, Lemma 4.4 in [59])

1. Assume that up to time $t \in [0,T)$ no default has occurred, i.e., $\mathcal{E}_t = \mathcal{D}$ and $\tau(\mathcal{E}_t) > t$.

   The survival probability from t to T for **a single obligor** $d \in \{1, \ldots, D\}$ is given by

$$Q_d(t,T) := \mathbb{E}\left[\mathbf{1}_{\{\tau^d > T\}} | \mathcal{H}_t\right] = \mathbb{E}\left[\frac{C\left(\gamma_t^1, \ldots, \gamma_t^{d-1}, \gamma_T^d, \gamma_t^{d+1}, \ldots, \gamma_t^D\right)}{C(\gamma_t)} | \mathcal{H}_t\right] \quad .$$

   The survival probability from t to T for **all obligors** is given by

$$Q(t,T) := \mathbb{E}\left[\mathbf{1}_{\{\tau > T\}} | \mathcal{H}_t\right] = \mathbb{E}\left[\frac{C(\gamma_T)}{C(\gamma_t)} | \mathcal{H}_t\right] \quad .$$

[10] Each default time is uniquely identifiable by its value (time of default) and its index.

2. Assume that the set of defaulted obligors up to time $t \in [0,T)$ is $\mathcal{C}_t$, i.e., $\tau(\mathcal{C}_t) \leq t$.

   The survival probability from t to T of **a single obligor** $d \in \mathcal{E}_t$ is given by

$$R_d(t,T) := \mathbb{E}\left[\mathbf{1}_{\{\tau^d \in \mathcal{E}_t > T\}} | \mathcal{H}_t\right] = \mathbb{E}\left[\frac{\frac{\partial^C}{\partial \lambda_t(\mathcal{C}_t)} C\left(\gamma_t(\mathcal{E}_t \setminus \{d\}), \gamma_T^d; \gamma_\tau(\mathcal{C}_t)\right)}{\frac{\partial^C}{\partial \lambda_t(\mathcal{C}_t)} C(\gamma_t(\mathcal{E}_t); \gamma_\tau(\mathcal{C}_t))} | \mathcal{H}_t\right] \quad .$$

   The survival probability from t to T of **all non-defaulted obligors** is given by

$$R(t,T) := \mathbb{E}\left[\mathbf{1}_{\{\tau(\mathcal{E}_t) > T\}} | \mathcal{H}_t\right] = \mathbb{E}\left[\frac{\frac{\partial^C}{\partial \lambda_t(\mathcal{C}_t)} C\left(\gamma_T(\mathcal{E}_t); \gamma_\tau(\mathcal{C}_t)\right)}{\frac{\partial^C}{\partial \lambda_t(\mathcal{C}_t)} C(\gamma_t(\mathcal{E}_t); \gamma_\tau(\mathcal{C}_t))} | \mathcal{H}_t\right] \quad .$$

Concerning the notation in the second part it holds

$$\frac{\partial^C}{\partial \lambda_t(\mathcal{C}_t)} C\left(\gamma_t(\mathcal{E}_t \setminus \{d\}), \gamma_T^d; \gamma_\tau(\mathcal{C}_t)\right) := \mathbb{Q}\left[U(\mathcal{E}_t \setminus \{d\}) \leq \gamma_t(\mathcal{E}_t \setminus \{d\}), U^d \leq \gamma_T^d | U(\mathcal{C}_t) = \gamma_\tau(\mathcal{C}_t)\right]$$

and

$$\frac{\partial^C}{\partial \lambda_t(\mathcal{C}_t)} C\left(\gamma_T(\mathcal{E}_t); \gamma_\tau(\mathcal{C}_t)\right) := \mathbb{Q}\left[U(\mathcal{E}_t) \leq \gamma_T(\mathcal{E}_t) | U(\mathcal{C}_t) = \gamma_\tau(\mathcal{C}_t)\right] \quad .$$

A proof is given in Appendix D.2.2. ∎

Interpretation

Proposition 9.17 shows that the survival probabilities are dependent on

- the copula function and its derivatives
- the intensities and default countdown processes of both, the concerned obligor and the other positions, and
- the occurred defaults.

Thus, the model includes default contagion. The countdown process is fixed at the revealed default time. We have to evaluate a multivariate distribution function C at certain values $U(\mathcal{C}_t) = \tau(\mathcal{C}_t) = \gamma_\tau(\mathcal{C}_t)$ which is naturally done by taking the derivative of C with respect to the variables with index $d \in \mathcal{C}_t$.

Analyzing the **real** default intensities of the Schönbucher-Schubert model, some drawbacks surface. Under certain regularity conditions[11], the intensity of obligor d is received by

$$h_t^d = -\frac{\partial}{\partial T} Q_d(t,T)\Big|_{T=t} = \lambda_t^d \gamma_t^d \frac{\frac{\partial}{\partial \lambda_t^d} C(\gamma_t)}{C(\gamma_t)} \tag{9.6}$$

if no default has occurred up to time $t \in [0,T)$. If $\mathcal{C} \neq \emptyset$ for $t \in [0,T)$, it holds

$$h_t^d = -\frac{\partial}{\partial T} R_d(t,T)\Big|_{T=t} = \lambda_t^d \gamma_t^d \frac{\frac{\partial^C}{\partial \lambda_t(\mathcal{C}_t) \partial \lambda_t^d} C\left(\gamma_t(\mathcal{E}_t); \gamma_\tau(\mathcal{C}_t)\right)}{\frac{\partial^C}{\partial \lambda_t(\mathcal{C}_t)} C\left(\gamma_t(\mathcal{E}_t); \gamma_\tau(\mathcal{C}_t)\right)} \quad . \tag{9.7}$$

[11] See [59], Section 2.3 and references therein.

Equations (9.6) and (9.7) can be derived by straightforward calculation, see also [59][12]. Under information $\mathcal{H}_t$ (i.e., knowing the evolution of the countdown process γ_t and defaults up to time t), the real intensities h_t^d, $d = 1, \ldots, D$, differ from model intensities λ_t^d. In the multidimensional setup filtration $\mathcal{H}_t$ contains more information than $\mathcal{H}_t^d$. Therefore, λ_t^d is often called *pseudo-intensity*. Here, the main obstacle of Schönbucher-Schubert approach is unveiled. The real default intensities h_t^d are observable in the market. For modeling purposes, pseudo-intensities λ_t^d need to be determined.

When it comes to calibration issues, two problems become apparent. They are

1. the choice of the "right" copula[13] and
2. the extraction of the model intensities λ_t^d out of the observable intensities h_t^d for $d = 1, \ldots, D$ in combination with the copula function. This can result in an inaccessible inverse problem.

If λ_t follows a **continuous** stochastic differential equation, the process h_t also satisfies an SDE which jumps in case of defaults, see Section 4.4 in [59]. The structure and the complexity of this SDE is dependent on

- defaults,
- copula C and its derivatives,
- intensities λ_t^d and corresponding countdown processes γ_t^d.

As for default probabilities, default contagion is directly visible by the SDE for h_t as it jumps if assets default.

Concerning consistency of the Schönbucher-Schubert model, it holds $\mathbb{E}\left[h_t^d | \mathcal{H}_t^d\right] = \lambda_t^d$. So, if we restrict our information to the one-obligor case, we will indeed recover the intensity. Moreover, the same result holds if the used copula function is chosen as independent copula Π. In this way, the approach is consistent.

9.18 Summary (Schönbucher-Schubert Framework)

1. Schönbucher and Schubert develop a copula-dependent model in an intensity-based credit risk framework which embraces default contagion.
2. In case of continuous-time pseudo-intensities, the setup might be calibrated. As intensities h_t^d, $d = 1, \ldots, D$, are observable in the market, the pseudo-intensities λ_t^d, $d = 1, \ldots, D$, are reproducible via Equations (9.6) and (9.7). However, this inversion problem is possibly unsolvable.
3. Due to the previous aspect, practical implementation is seldom.

10 A Copula- and Default-Dependent Intensity Model

In this chapter, we develop a reduced-form credit risk model which connects ideas and elements of the Cox process framework, the Schönbucher-Schubert approach and the concept of copulas. The common intensity-based design is modified as evaluations use altered conditioning filtrations.

At this point, we focus on a generalized mathematical construction, i.e., introduced processes, distributions and time grid are universally valid. Model intentions and specifications as well as application issues are tackled later. Requirements on the design are that

- defaults interfere with model components (i.e., the model reproduces default contagion) and
- intensities are stochastic and interdependent.

[12] As auxiliary comment, note that the expectations of $Q_d(t,T)$ and $R_d(t,T)$ vanish as the derivatives are evaluated at time $T = t$.

[13] See also discussion in Chapter 6.

We start with the familiar multivariate setting of Chapter 9. Let N_t, $t \in [0,T]$, be a D-dimensional jump process with stochastic intensity process

$$\lambda_t = \left(\lambda_t^1, \ldots, \lambda_t^D\right)'$$

which is supposed to be càdlàg and non-negative[1]. Default times τ^d, $d = 1, \ldots, D$, are modeled as first jump of the concerned component of N_t, i.e.,

$$\tau^d := \inf\left\{t \in [0,T] : N_t^d > 0\right\} \quad .$$

In this framework - initially ignoring dependence structures and default contagion effects - survival probabilities[2] from 0 up to time $t \in [0,T]$ are given as

$$\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t\}}\right] = \mathbb{P}\left[\tau^d > t\right] = \mathbb{P}\left[N_t^d = 0\right] = \mathbb{E}\left[\exp\left(-\int_0^t \lambda_s^d ds\right)\right]$$

for each position $d \in \{1, \ldots, D\}$. Under a realized path $\tilde{\lambda}_t^d$ of λ_t^d, we receive

$$\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t\}} | \tilde{\lambda}_t^d\right] = \mathbb{P}\left[\tau^d > t | \tilde{\lambda}_t^d\right] = \mathbb{P}\left[N_t^d = 0 | \tilde{\lambda}_t^d\right] = \exp\left(-\int_0^t \tilde{\lambda}_s^d ds\right) \quad ,$$

see Lemma 9.12, respectively [40] or [58]. We borrow the default mechanism of [59] and introduce the default trigger level as

$$U = (U^1, \ldots, U^D)' \sim U(0,1)^D$$

which is independent of all other modeling variables. In contrast to [59], we assume that U^d, $d = 1, \ldots, D$, are mutually independent[3]. As observed, default times are constructable as

$$\tau^d := \inf\left\{t \in [0,T] : \exp\left(-\int_0^t \lambda_s^d ds\right) \leq U^d\right\} \tag{10.1}$$

for $d = 1, \ldots, D$. As in the Schönbucher-Schubert framework, the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ supports all subsequent processes and variables. All introduced filtrations are subsets of $\mathcal{A}$ and are enlarged by zero-sets $\mathcal{N}$ of $\mathcal{A}$, see [59], pp. 12–15. Further, we denote the σ-algebras

1. $\mathcal{G}_t := \sigma\left(\lambda_s : 0 \leq s \leq t\right) \vee \mathcal{N}$ for $t \in [0,T]$ as *intensity filtration* and
2. $\mathcal{F}_t := \sigma\left(N_s : 0 \leq s \leq t\right) \vee \mathcal{N}$ for $t \in [0,T]$ as *default filtration* and

Each component λ_t^d, $d = 1, \ldots, D$, of the intensity process λ_t is $\mathcal{G}_t$-adapted by construction.

10.1 Remark
In contrast, the common reduced-form approach and the Schönbucher-Schubert model operate on a background filtration to which the intensity process is adapted. It contains **all** market and non-default information. The default filtration is identical to the one in [59].

For the multidimensional setup we take over Notation 9.16 concerning

- the index sets for non-defaulted and defaulted positions ($\mathcal{E}_t$ and $\mathcal{C}_t$),
- non-defaulted and defaulted vectors ($x_s(\mathcal{E}_t)$ and $x_s(\mathcal{C}_t)$) for $s,\ t \in [0,T]$, and
- potential and realized default times ($\tau(\mathcal{E}_t)$ and $\tau(\mathcal{C}_t)$).

[1] Compare to Assumption 1 in [59].
[2] The probability measure (real-world / risk-neutral) depends on the relevant application of the model. As basic approach we use the notation $\mathbb{P}$.
[3] In [59], the premise is $U \sim$ copula C.

Demands on the construction of the intensity process are that λ_t

- is stochastic, càdlàg and non-negative,
- covers multivariate dependence, and
- interferes with defaults.

10.1 Generating an Intensity Process with Conditional Margins

Let λ_t be defined as intensity of N_t above. First, we discretize the interval into N equidistant[4] parts with

$$t_0 = 0,\, t_N = T \quad \text{and} \quad t_n - t_{n-1} = \Delta t \quad \text{for } n = 1, \ldots, N.$$

The intensity is assumed to be constant on sub-intervals $[t_n, t_{n+1})$ and to be driven by jumps at time steps t_n[5] denoted as

$$\Delta\lambda_n = (\Delta\lambda_n^1, \ldots, \Delta\lambda_n^D)' \quad \text{for } n = 1, \ldots, N \quad .$$

Thus, the change of the process from t_{n-1} to t_n is given by

$$\tilde{\lambda}_{t_n} := \tilde{\lambda}_{t_{n-1}} + \Delta\lambda_n$$

or written as vector

$$\begin{pmatrix} \tilde{\lambda}_{t_n}^1 \\ \vdots \\ \tilde{\lambda}_{t_n}^D \end{pmatrix} = \begin{pmatrix} \tilde{\lambda}_{t_{n-1}}^1 + \Delta\lambda_n^1 \\ \vdots \\ \tilde{\lambda}_{t_{n-1}}^D + \Delta\lambda_n^D \end{pmatrix} \quad .$$

We suppose that the intensity λ_t is constructed by a positive initial value $\lambda_0 = (\lambda_0^1, \ldots, \lambda_0^D)'$ and the sum of its changes up to time t. Thus, paths are representable as

$$\tilde{\lambda}_t^d = \tilde{\lambda}_{t_n}^d = \lambda_0^d + \sum_{j=1}^{n} \Delta\lambda_j^d \quad \text{with } t \in [t_n, t_{n+1}) \tag{10.2}$$

for $n = 0, \ldots, N$ and each $d = 1, \ldots, D$. In vectorial form we have

$$\tilde{\lambda}_t = \begin{pmatrix} \tilde{\lambda}_t^1 \\ \vdots \\ \tilde{\lambda}_t^D \end{pmatrix} = \begin{pmatrix} \tilde{\lambda}_{t_n}^1 \\ \vdots \\ \tilde{\lambda}_{t_n}^D \end{pmatrix} = \begin{pmatrix} \lambda_0^1 + \sum\limits_{j=1}^{n} \Delta\lambda_j^1 \\ \vdots \\ \lambda_0^D + \sum\limits_{j=1}^{n} \Delta\lambda_j^D \end{pmatrix} \quad \text{with } t \in [t_n, t_{n+1}) \quad . \tag{10.3}$$

As intensity process, λ_t must remain non-negative. As stronger restriction, we demand λ_t to be positive. To ensure this positivity we extend the above construction. We impose that

$$\tilde{\lambda}_{t_n}^d = \tilde{\lambda}_{t_{n-1}}^d + \Delta\lambda_n^d > 0$$

or equivalently

$$\Delta\lambda_n^d > -\tilde{\lambda}_{t_{n-1}}^d = -\lambda_0^d - \sum_{j=1}^{n-1} \Delta\lambda_j^d \tag{10.4}$$

holds for each time step t_n, $n = 1, \ldots, N$, and each asset $d = 1, \ldots, D$. The change $\Delta\lambda_n^d$ is therefore dependent on λ_0^d and $\Delta\lambda_1^d, \ldots, \Delta\lambda_{n-1}^d$.

[4] For an arbitrary time grid, model calibration would have to be performed for each single time step which would not be target-oriented for implementation.

[5] We allow a final jump at $t_N = T$.

Suppose now that $\Delta\lambda = (\Delta\lambda^1, \dots, \Delta\lambda^D)'$ is a realization of a D-dimensional random variable

$$\Lambda = (\Lambda^1, \dots, \Lambda^D)' \sim F$$

with $\Lambda_d \sim F^d =: F^d_{\Delta t}$ for $d = 1, \dots, D$. The index Δt indicates the dependence of the margins on the chosen time grid. Initially ignoring dependence structures, we restrict to the one-dimensional case and generalize Equation (10.2) for all paths as

$$\lambda^d_t = \lambda^d_{t_n} = \lambda^d_0 + \sum_{j=1}^{n} \Lambda^d_j \quad \text{for } t \in [t_n, t_{n+1}),\ n = 0, \dots, N \tag{10.5}$$

in which Λ^d_n denotes the nth draw of the marginal random variable Λ^d. To ensure inequality (10.4), Λ^d_n has to be larger than $-\lambda^d_{t_{n-1}}$. If the range of F^d includes $-\lambda_{t_{n-1}}$, Λ^d_n cannot be distributed according to F^d. However, the distribution $F^d_n(\Delta t) := F^d_n$ of Λ^d_n can be obtained by a **conditional** cdf of F^d and the $\mathcal{G}_t$-adaptivity of λ_t as

$$\begin{aligned}
& F^d_n(\Delta\lambda^d_n, \Delta t, \mathcal{G}_{t_{n-1}}) = F^d_n(\Delta\lambda^d_n) = \mathbb{P}\left[\Lambda^d_n \le \Delta\lambda^d_n \middle| \mathcal{G}_{t_{n-1}}\right] \\
:= \ & F^d\left(\Delta\lambda^d_n \middle| \left\{\Lambda^d > -\lambda^d_{t_{n-1}}\right\} \vee \mathcal{G}_{t_{n-1}}\right) = \mathbb{P}\left[\Lambda^d \le \Delta\lambda^d_n \middle| \left\{\Lambda^d > -\lambda^d_{t_{n-1}}\right\} \vee \mathcal{G}_{t_{n-1}}\right] \\
= \ & \frac{\mathbb{P}\left[-\lambda^d_{t_{n-1}} < \Lambda^d \le \Delta\lambda^d_n | \mathcal{G}_{t_{n-1}}\right]}{\mathbb{P}\left[\Lambda^d > -\lambda^d_{t_{n-1}} | \mathcal{G}_{t_{n-1}}\right]} = \frac{\mathbb{P}[\Lambda^d \le \Delta\lambda^d_n] - \mathbb{P}\left[\Lambda^d \le -\lambda^d_{t_{n-1}} | \mathcal{G}_{t_{n-1}}\right]}{1 - \mathbb{P}\left[\Lambda^d \le -\lambda^d_{t_{n-1}} | \mathcal{G}_{t_{n-1}}\right]} \\
= \ & \frac{F^d(\Delta\lambda^d_n) - F^d\left(-\lambda^d_{t_{n-1}} | \mathcal{G}_{t_{n-1}}\right)}{1 - F^d\left(-\lambda^d_{t_{n-1}} | \mathcal{G}_{t_{n-1}}\right)}
\end{aligned} \tag{10.6}$$

with $\Delta\lambda^d_n \in (-\lambda_{t_{n-1}}, \infty)$ and for $n = 1, \dots, N$. The construction is based on the fact that $\lambda^d_{t_{n-1}}$ is known. This is in turn controlled by the realizations of Λ^d_j, $j = 1, \dots, n-1$ which are revealed due to conditioning on the filtration $\mathcal{G}_{t_{n-1}}$. Thus, changes $\Delta\lambda^d_n$, $n = 1, \dots, N$, are path-dependent.

10.2 Remark (Time-Grid Dependence of Distributions)
The random variables Λ^d_n must be related to the partition of the interval $[0, T]$. This means, distributions $F^d_{\Delta t} = F^d$ and $F^d_n(\Delta t) = F^d_n$ have to be adjusted to the given time grid.

10.3 Definition (Admissible Distribution)
We say that a marginal distribution F^d, $d = 1, \dots, D$, is *admissible on the time grid* Δt if it is valid for construction (10.6), i.e., it holds

$$F^d(-\lambda_{t_n}) < 1 \text{ for all } n = 0, \dots, N-1.$$

In particular, all distributions with range $(\infty, -\lambda^d_{t_n}]$, $n = 0, \dots, N-1$, are not admissible. For sake of readability we abbreviate $a^d_n := \lambda^d_{t_{n-1}}$ and $\mathcal{G}_n := \mathcal{G}_{t_{n-1}}$.

10.4 Remark

1. The characteristics of F^d_n can considerably differ from the original distribution F^d.
2. If intended, the margins F^d can be cut off at both ends instead of one. The above construction of the conditional distribution has to be modified to

$$\begin{aligned}
G^d_n(\Delta\lambda^d_n, \Delta t, \mathcal{G}_n) =& G^d_n(\Delta\lambda^d_n) = \mathbb{P}\left[\Lambda^d_n \le \Delta\lambda^d_n | \mathcal{G}_n\right] := F^d\left(\Delta\lambda^d_n \middle| \left\{-a^d_n < \Lambda^d \le a^d_n\right\} \vee \mathcal{G}_n\right) \\
=& \frac{F^d(\Delta\lambda^d_n) - F^d\left(-a^d_n | \mathcal{G}_n\right)}{F^d\left(a^d_n | \mathcal{G}_n\right) - F^d\left(-a^d_n | \mathcal{G}_n\right)}
\end{aligned} \tag{10.7}$$

with $\Delta\lambda^d_n \in \left(-a^d_n, a^d_n\right]$.

3. If F^d is symmetrically distributed around zero[6], specification (10.7) yields that expectation and symmetry are maintained, i.e., $\mathbb{E}[\Lambda_n^d] = \mathbb{E}[\Lambda^d]$ and $G_n^d(x) = 1 - G_n^d(-x)$.

In order to consider both approaches - Equations (10.6) and (10.7) - we redefine

$$F_n^d(\Delta\lambda_n^d) := F^d\left(\Delta\lambda_n^d \middle| \left\{-a_n^d < \Lambda^d \leq \kappa_n^d\right\} \vee \mathcal{G}_n\right) = \frac{F^d(\Delta\lambda_n^d) - F^d\left(-a_n^d|\mathcal{G}_n\right)}{\kappa_n^d - F^d\left(-a_n^d|\mathcal{G}_n\right)} \tag{10.8}$$

with

$$\kappa_n^d := \begin{cases} 1 & \text{for } F_n^d \text{ as in (10.6)} \\ F^d\left(a_n^d|\mathcal{G}_n\right) & \text{for } G_n^d \text{ as in (10.7).} \end{cases} \tag{10.9}$$

If the context is clear, we omit the conditioning term for distribution functions, i.e.,

$$F^d\left(-\lambda_{t_{n-1}}^d\right) = F^d\left(-\lambda_{t_{n-1}}^d|\mathcal{G}_{t_{n-1}}\right) \quad \text{respectively} \quad F^d\left(-a_n^d\right) = F^d\left(-a_n^d|\mathcal{G}_n\right) \quad .$$

Turning back to the analysis, we require

10.5 Assumption (Discrete-Time Setup)

1. The time steps $t_n - t_{n-1} = \Delta t$ are fixed for $n = 1, \ldots, N$ with $N \in \mathbb{N}$.
2. The random variable Λ_n^d is distributed according to F_n^d as defined in Equation (10.8) with κ_n^d as in Equation (10.9) for $n = 1, \ldots, N$ and $d = 1, \ldots, D$.
3. All distributions are continuous, strictly increasing and admissible in sense of Definition 10.3.

10.6 Lemma

Under Assumption 10.5 each marginal intensity process λ_t^d is $\mathbb{P}$-almost surely positive.

A proof can be found in Appendix D.2.3. ∎

Generation of the intensity processes (10.5) via $\Lambda_n^d \sim F_n^d$ of Equations (10.8) and (10.9) assures positivity. It leaves to be shown how random variables $\Lambda_n^d \sim F_n^d$ are generated. The analytical and numerical creation of Λ_n^d is not straightforward because F_n^d is a **conditional** distribution. We bypass the problem as follows. For the conditional cdf F_n^d have in mind that

$$F_n^d : \left(-a_n^d, \kappa_n^d\right) \to (0, 1)$$

holds. Therefore, it is

$$v_n^d := F_n^d\left(\Delta\lambda_n^d\right) = \frac{F^d(\Delta\lambda_n^d) - F^d\left(-a_n^d\right)}{\kappa_n^d - F^d\left(-a_n^d\right)} \in (0, 1) \quad . \tag{10.10}$$

As F^d is a continuous and strictly increasing cdf, it is uniquely invertible. Inversion yields

$$\Delta\lambda_n^d = \left(F_n^d\right)^{-1}\left[v_n^d\right] \tag{10.11}$$

or equivalently

$$\Delta\lambda_n^d = \left(F^d\right)^{-1}\left[v_n^d\left(\kappa_n^d - F^d\left(-a_n^d\right)\right) + F^d\left(-a_n^d\right)\right]$$

with

$$\lim_{v_n^d \to 0}\left(F^d\right)^{-1}\left[v_n^d\left(\kappa_n^d - F^d\left(-a_n^d\right)\right) + F^d\left(-a_n^d\right)\right] = -a_n^d \quad ,$$

$$\lim_{v_n^d \to 1}\left(F^d\right)^{-1}\left[v_n^d\left(\kappa_n^d - F^d\left(-a_n^d\right)\right) + F^d\left(-a_n^d\right)\right] = \begin{cases} \infty & \text{for } \kappa_n^d = 1 \\ a_n^d & \text{for } \kappa_n^d = F^d\left(a_n^d\right) \end{cases} \quad .$$

[6] For symmetric distributions with non-zero means, results are similar. Random variables just need to be shifted.

Thus, the random variables Λ_n^d with range $\left(-a_n^d, \kappa_n^d\right)$ can be derived by

$$\Lambda_n^d = \left(F^d\right)^{-1}\left[V_n^d\left(\kappa_n^d - F^d\left(-a_n^d\right)\right) + F^d\left(-a_n^d\right)\right] \sim F_n^d \tag{10.12}$$

in which V_n^d is the nth draw of a standard uniformly distributed random variable V^d. Note that the time index n is not yet necessary, but it will be needed as time indication in subsequent sections. Moreover, Lemma 10.6 still holds for the above construction (10.12)[7]. For the one-dimensional case, Equation (10.12) gives us the recipe for the generation of random variables $\Lambda_n^d \sim F_n^d$:

- draw v_n^d from $V_n^d \sim U(0,1)$, i.i.d. for $n = 1, \ldots, N$
- perform transformation (10.12)
- receive $\Delta\lambda_n^d$ from $\Lambda_n^d \sim F_n^d$

10.7 Assumption
For all $d \in \{1, \ldots, D\}$ and $m, n = 1, \ldots, N$ and $m \neq n$, V_m^d is independent of V_n^d.

This assumptions reveals a sort of "pseudo" time-independence as well as the path-dependence of the random variables Λ_n^d, $n = 1, \ldots, N$.

10.8 Lemma (Path-Dependent Increments)
Under Assumptions 10.5, 10.7, and relevant filtrations $\mathcal{G}_m$ and $\mathcal{G}_n$, the increments Λ_m^d and Λ_n^d derived as in Equation (10.12) are *pseudo-independent* for all $m < n$, i.e., it holds

$$\mathbb{P}\left[\Lambda_m^d \leq x_1, \Lambda_n^d \leq x_2 | \mathcal{G}_m\right] = \mathbb{P}\left[\Lambda_m^d \leq x_1 | \mathcal{G}_m\right] \cdot \mathbb{E}\left[\mathbb{P}\left[\Lambda_n^d \leq x_2 | \mathcal{G}_n\right] | \mathcal{G}_m\right] \quad .$$

A proof can be found in Appendix D.2.5. ∎

Summarized, the representation of the intensity process (10.5) through (10.12) is given as

$$\begin{aligned} \lambda_t^d = \lambda_{t_n}^d &= \lambda_0^d + \sum_{j=1}^{n} \Lambda_j^d \\ &= \lambda_0^d + \sum_{j=1}^{n} \left(F^d\right)^{-1}\left[V_j^d\left(\kappa_j^d - F^d\left(-a_j^d\right)\right) + F^d\left(-a_j^d\right)\right] \end{aligned} \tag{10.13}$$

for $d = 1, \ldots, D$ and $n = 0, \ldots, N$.

Equation (10.13) generates a strictly positive path of the intensity process λ_t^d driven by path-dependent increments. It would have been sufficient if we had conditioned concerned equations on the filtration generated by the one-dimensional intensity $\sigma(\lambda_{t_n}^d) \subset \mathcal{G}_{t_n}$, $n = 0, \ldots, N$. However, when introducing dependencies between intensities in the next section, it is inevitable to condition on the whole intensity filtration. For the multidimensional process $\lambda_t = (\lambda_t^1, \ldots, \lambda_t^D)'$, the dependency structure of its driving random variables

$$\Lambda_n = (\Lambda_n^1, \ldots, \Lambda_n^D)' \sim F_n \quad \text{for } n = 1, \ldots, N$$

is not yet specified. Due to Equation (10.12), one can already guess that via the uniformly distributed random variables

$$V_n = (V_n^1, \ldots, V_n^D)' \quad \text{for } n = 1, \ldots, N$$

the dependence structure will be introduced by copula functions.

10.9 Remark
Increments are pseudo-independent and not identically distributed for each iteration t_{n-1} to t_n, $n = 1, \ldots, N$. Therefore, the setup (10.13) does **not** generate Lévy processes.

[7] A proof can be found in Appendix D.2.4.

10.2 Incorporating Copula Functions

Suppose that the general random variable $\Lambda = (\Lambda^1, \ldots, \Lambda^D)' \sim F$ has copula function K, i.e.,

$$F(\Delta\lambda^1, \ldots, \Delta\lambda^D) = K(F^1(\Delta\lambda^1), \ldots, F^D(\Delta\lambda^D)) \tag{10.14}$$

for given $\Delta\lambda^d$ of Λ^d with marginal distribution functions F^d, $d = 1, \ldots, D$. One implication of this theorem, compare Corollary 2.3, is that for given v^d of $V^d \sim U(0,1)$, $d = 1, \ldots, D$, we have

$$F\left((F^1)^{-1}[v^1], \ldots, (F^D)^{-1}[v^D]\right) = K(v^1, \ldots, v^D) \quad .$$

and therefore

$$V = \left(V^1, \ldots, V^D\right) \sim K \quad .$$

10.10 Assumption (Dependence Structure of Random Changes for Intensities)
The standard uniform random variables V_n^d, $d = 1, \ldots, D$ and $n = 1, \ldots, N$, of Equation (10.10) are assumed to be distributed to copula K of the general random variable $\Lambda \sim F$, i.e.,

$$V_n = \left(V_n^1, \ldots, V_n^D\right) \sim K \quad \text{for } n = 1, \ldots, N \quad .$$

This premise implies the following important result for the dependence structure of the random variable Λ_n.

10.11 Proposition
Under Assumptions 10.5 and 10.10, the random variables $\Lambda \sim F$ and $\Lambda_n \sim F_n$ have the identical and unique copula function K. In mathematical terms that means

$$F_n(\Delta\lambda_n^1, \ldots, \Delta\lambda_n^D) = K(F_n^1(\Delta\lambda_n^1), \ldots, F_n^D(\Delta\lambda_n^D)) \quad \text{for } n = 1, \ldots, N \quad .$$

Proof

The general random variables $\Lambda \sim F$ is assumed to have copula K, compare Equation (10.14). Further, as Λ is independent of the filtration $\mathcal{G}_n$ for $n = 1, \ldots, N$, it is

$$\begin{aligned} F(\Delta\lambda_n^1, \ldots, \Delta\lambda_n^D) &= \mathbb{P}\left[\Lambda^1 \le \Delta\lambda_n^1, \ldots, \Lambda^D \le \Delta_n^D\right] \\ &= \mathbb{P}\left[\Lambda^1 \le \Delta\lambda_n^1, \ldots, \Lambda^D \le \Delta_n^D | \mathcal{G}_n\right] \\ &= K(F^1(\Delta\lambda_n^1), \ldots, F^D(\Delta\lambda_n^D)) \quad . \end{aligned}$$

For $\Lambda_n \sim F_n$ given filtration $\mathcal{G}_n$, it holds

$$\begin{aligned} F_n(\Delta\lambda_n^1, \ldots, \Delta\lambda_n^D) &= \mathbb{P}\left[\Lambda_n^1 \le \Delta\lambda_n^1, \ldots, \Lambda_n^D \le \Delta_n^D | \mathcal{G}_n\right] \\ &= \mathbb{P}\Bigg[\left(F^1\right)^{-1}\left[V_n^1\left(\kappa_n^1 - F^1\left(-a_n^1\right)\right) + F^1\left(-a_n^1\right)\right] \le \Delta\lambda_n^1, \ldots, \\ &\qquad \left(F^D\right)^{-1}\left[V_n^D\left(\kappa_n^D - F^D\left(-a_n^D\right)\right) + F^D\left(-a_n^D\right)\right] \le \Delta_n^D | \mathcal{G}_n\Bigg] \\ &= \mathbb{P}\left[V_n^1 \le \frac{F^1(\Delta\lambda_n^1) - F^1\left(-a_n^1\right)}{\kappa_n^1 - F^1\left(-a_n^1\right)}, \ldots, V_n^D \le \frac{F^D(\Delta\lambda_n^D) - F^D\left(-a_n^D\right)}{\kappa_n^D - F^D\left(-a_n^D\right)} | \mathcal{G}_n\right] \end{aligned}$$

As the expressions $F^d\left(-a_n^d\right)$, $d = 1, \ldots, D$, are known under $\mathcal{G}_n$, we get

$$\begin{aligned} F_n(\Delta\lambda_n^1, \ldots, \Delta\lambda_n^D) &= \mathbb{P}\left[V_n^1 \le F_n^1(\Delta\lambda_n^1), \ldots, V_n^D \le F_n^D(\Delta\lambda_n^D) | \mathcal{G}_n\right] \\ &= K\left(F_n^1(\Delta\lambda_n^1), \ldots, F_n^D(\Delta\lambda_n^D)\right) \end{aligned}$$

for each $n = 1, \ldots, N$.

We have used the definition of Λ_n^d - Equation (10.12), the intensity filtration $\mathcal{G}_n$ for the evaluation of $F_n^d(-a_n^d)$ and κ_n^d in each component, the definition of F_n^d - Equation (10.8), and Assumption 10.10 with $V_n \sim K$. In addition to the equivalence of the copula, we remember that F and F_n^d are assumed to be continuous. The copula K is therefore unique in both cases due to Sklar's Theorem 2.1. ∎

This result is important for the application of the model. Copula K and marginal distributions Λ^d, $d = 1, \dots, D$, can directly be calibrated to market data. A fitting of the conditional margins Λ_n^d is redundant.

10.12 Remark
The multivariate cdf of F and F_n are naturally different in each time step.

$$F(\Delta\lambda_n^1, \dots, \Delta\lambda_n^D) = K(F^1(\Delta\lambda_n^1), \dots, F^D(\Delta\lambda_n^D))$$

is generally not equal to

$$\begin{aligned} F_n(\Delta\lambda_n^1, \dots, \Delta\lambda_n^D) &= K(F_n^1(\Delta\lambda_n^1), \dots, F_n^D(\Delta\lambda_n^D)) \\ &= K\Bigg(F^1\left(\Delta\lambda_n^1 \middle| \left\{-a_n^1 < \Lambda^1 \leq \left(F^1\right)^{-1}\left[\kappa_n^1\right]\right\} \vee \mathcal{G}_n\right), \dots, \\ &\qquad F^D\left(\Delta\lambda_n^D \middle| \left\{-a_n^D < \Lambda^D \leq \left(F^D\right)^{-1}\left[\kappa_n^D\right]\right\} \vee \mathcal{G}_n\right)\Bigg) \end{aligned}$$

due to the different margins F^d and F_n^d, $d = 1, \dots, D$ and $n = 1, \dots, N$.

10.13 Summary
The dependency structure of the increments Λ_n - copula function K - remains steady whereas their multivariate distribution - cdf F_n - varies in each time step t_n, $n = 1, \dots, N$.

Under these considerations, we extend (10.13) to multidimensional, copula-dependent intensity model as

$$\begin{aligned} \lambda_t^d = \lambda_{t_n}^d &= \lambda_0^d + \sum_{j=1}^{n} \Lambda_j^d \\ &= \lambda_0^d + \sum_{j=1}^{n} \left(F^d\right)^{-1}\left[V_j^d\left(\kappa_j^d - F^d\left(-a_j^d\right)\right) + F^d\left(-a_j^d\right)\right] \end{aligned} \tag{10.15}$$

for $d = 1, \dots, D$, $n = 0, \dots, N$, and $t \in [t_n, t_{n+1})$ with

$$V_j = (V_j^1, \dots, V_j^D)' \sim K$$

such that

$$\Lambda_j = (\Lambda_j^1, \dots, \Lambda_j^D)' \sim F_j$$

for $j = 1, \dots, n$. The design for processes and random variables are preserved, a dependence structure is introduced.

By now, we have not yet reasoned the multi-time step construction. The above framework could be implemented as one period and repeated N times. Results will be equivalent to an N time construction if time grids are identical for both setups. Incorporating defaults to the model makes a multi-time step basis indispensable.

10.14 Remark
The key to launch defaults into our approach is to exploit the information generated by the intensity λ_t and the Cox process N_t simultaneously. In other words, information of a càdlàg process on a discrete-time grid (λ_t) is combined with the filtration of a continuous-time framework (N_t).

10.3 Including a First Default

For an adaption on default events, we initially suppose that at time $\tau^{\tilde{d}_1} = (0, T]$ the **first** default occurs for position $\tilde{d}_1 \in \mathcal{D}$. The first default time has been determined, i.e.,

$$\tau^{\tilde{d}_1} = \inf\left\{t \in [0,T] : N_t^{\tilde{d}_1} > 0\right\}$$

or equivalently - see Equation (10.1)

$$\exp\left(-\int_0^{\tau^{\tilde{d}_1}} \lambda_s^{\tilde{d}_1} ds\right) \leq U^{\tilde{d}_1}$$

with index sets $\mathcal{C}_{\tau^{\tilde{d}_1}} = \{\tilde{d}_1\}$ and $\mathcal{E}_{\tau^{\tilde{d}_1}} = \mathcal{D} \setminus \{\tilde{d}_1\}$ for the defaulted and non-defaulted indices at time $\tau^{\tilde{d}_1}$, see Notation 9.16. Model expansion is driven by two fundamental assumptions.

10.3.1 Assumption - Locked Processes and Variables for the Defaulted Position

Similar - but not identical - to the Schönbucher-Schubert approach we assume that no further assessment for the defaulted obligor is needed. Processes are locked in and remain constant for the rest of the period. In particular, it holds

$$\lambda_t^{\tilde{d}_1} = \lambda_{\tau^{\tilde{d}_1}}^{\tilde{d}_1} \quad \text{for } t \in [\tau^{\tilde{d}_1}, T] \quad .$$

for the intensity process. As they are independent of all other modeling variables, the default trigger variables U^d, $d = 1, \ldots, D$, are not affected. Assume that $\tau^{\tilde{d}_1} \in [t_{n-1}, t_n)$. Due to the càdlàg property, it holds

$$\exp\left(-\sum_{j=0}^{n-1}\left(\lambda_{t_j}^{\tilde{d}_1} \cdot \Delta t\right) - \lambda_{t_{n-1}}^{\tilde{d}_1} \cdot \left(\tau^{\tilde{d}_1} - t_{n-1}\right)\right) \leq U^{\tilde{d}_1} \quad \text{for } t \in [\tau^{\tilde{d}_1}, T]$$

and for the intensity process

$$\lambda_t^{\tilde{d}_1} = \lambda_{t_{n-1}}^{\tilde{d}_1} \quad \text{for } t \in [\tau^{\tilde{d}_1}, T] \quad .$$

10.3.2 Assumption - Default Impact on Non-Defaulted Positions

As further assumption we expect that the default **instantaneously** affects the evolution of the intensity process λ_t, i.e., an additional random change

$$\Delta\phi_{\tau^{\tilde{d}_1}} = \left(\Delta\phi_{\tau^{\tilde{d}_1}}^1, \ldots, \Delta\phi_{\tau^{\tilde{d}_1}}^D\right)'$$

is triggered at default time $\tau^{\tilde{d}_1}$. As representation for a path $\tilde{\lambda}_t$ of λ_t, we receive

$$\tilde{\lambda}_t = \lambda_0 + \sum_{j=1}^{n-1} \Delta\lambda_j + \Delta\phi_{\tau^{\tilde{d}_1}} \quad \text{for } t \in [\tau^{\tilde{d}_1}, t_n) \quad .$$

Analogous to Section 10.1, we suppose that the *generalized default change*

$$\Delta\phi = \left(\Delta\phi^1, \ldots, \Delta\phi^D\right)'$$

is a realization of a D-dimensional random variable

$$\Phi = \left(\Phi^1, \ldots, \Phi^D\right)'$$

with multivariate cdf H, differentiable copula L and continuous margins $H^1, \ldots, H^D$, i.e.,

$$H\left(\Delta\phi^1, \ldots, \Delta\phi^D\right) = L\left(H^1\left(\Delta\phi^1\right), \ldots, H^d\left(\Delta\phi^D\right)\right)' \quad .$$

Just as for the realizations $\Delta\lambda_n$ of Λ_n, $n = 1, \ldots, N$, the default change $\Delta\phi_{\tau^{\tilde{d}_1}}$ has to preserve positivity of the intensity process Λ_t. We assume that $\Delta\phi^d_{\tau^{\tilde{d}_1}}$ is drawn from the conditional random variable $\Phi^d_{\tau^{\tilde{d}_1}}$ distributed according to

$$\begin{aligned} & H^d_{\tau^{\tilde{d}_1}}\left(\Delta\phi^d_{\tau^{\tilde{d}_1}}, \mathcal{G}_{t_{n-1}}\right) = H^d_{\tau^{\tilde{d}_1}}\left(\Delta\phi^d_{\tau^{\tilde{d}_1}}\right) \\ := \; & H^d\left(\Delta\phi^d_{\tau^{\tilde{d}_1}} | \left\{-\lambda^d_{t_{n-1}} < \Phi^d \leq \mu^d_{\tau^{\tilde{d}_1}}\right\} \vee \mathcal{G}_{t_{n-1}}\right) \\ = \; & \frac{H^d\left(\Delta\phi^d_{\tau^{\tilde{d}_1}}\right) - H^d\left(-\lambda^d_{t_{n-1}} | \mathcal{G}_{t_{n-1}}\right)}{\mu^d_{\tau^{\tilde{d}_1}} - H^d\left(-\lambda^d_{t_{n-1}} | \mathcal{G}_{t_{n-1}}\right)} \quad \text{for } d \in \mathcal{E}_{\tau^{\tilde{d}_1}} \end{aligned} \tag{10.16}$$

with admissible, continuous margins H^d and possible choices

$$\mu^d_{\tau^{\tilde{d}_1}} := \begin{cases} 1 \\ H^d\left(-\lambda^d_{t_{n-1}} | \mathcal{G}_{t_{n-1}}\right) \end{cases} , \tag{10.17}$$

compare to Equations (10.8) and (10.9) in Section 10.1. Again - see (10.10) and (10.11) - it holds

$$w^d_{\tau^{\tilde{d}_1}} := H^d_{\tau^{\tilde{d}_1}}\left(\Delta\phi^d_{\tau^{\tilde{d}_1}}\right) = \frac{H^d(\Delta\phi^d_{\tau^{\tilde{d}_1}}) - H^d\left(-\lambda^d_{t_{n-1}}\right)}{\mu^d_{\tau^{\tilde{d}_1}} - H^d\left(-\lambda^d_{t_{n-1}}\right)} \in (0,1)$$

and by inversion we receive

$$\Delta\phi^d_{\tau^{\tilde{d}_1}} = \left(H^d_{\tau^{\tilde{d}_1}}\right)^{-1}\left[w^d_{\tau^{\tilde{d}_1}}\right] = \left(H^d\right)^{-1}\left[w^d_{\tau^{\tilde{d}_1}}\left(\mu^d_{\tau^{\tilde{d}_1}} - H^d\left(-\lambda^d_{t_{n-1}}\right)\right) + H^d\left(-\lambda^d_{t_{n-1}}\right)\right] \quad .$$

Thus, with $W^d_{\tau^{\tilde{d}_1}} \sim U(0,1)$ the construction of the random default change is definable as

$$\Phi^d_{\tau^{\tilde{d}_1}} = \left(H^d\right)^{-1}\left[W^d_{\tau^{\tilde{d}_1}}\left(\mu^d_{\tau^{\tilde{d}_1}} - H^d\left(-\lambda^d_{t_{n-1}}\right)\right) + H^d\left(-\lambda^d_{t_{n-1}}\right)\right] \sim H^d_{\tau^{\tilde{d}_1}} \tag{10.18}$$

which maintains positivity of the intensity for $d \in \mathcal{E}_{\tau^{\tilde{d}_1}}$. For $t \in [\tau^{\tilde{d}_1}, t_n)$, the process is now representable as

$$\begin{aligned} \lambda^d_t = \; & \lambda^d_0 + \sum_{j=1}^{n-1} \Lambda^d_j + \Phi^d_{\tau_{\tilde{d}_1}} \\ = \; & \lambda^d_0 + \sum_{j=1}^{n-1} \left(F^d\right)^{-1}\left[V^d_j\left(\kappa^d_j - F^d\left(-\lambda^d_{t_{j-1}}\right)\right) + F^d\left(-\lambda^d_{t_{j-1}}\right)\right] \\ & + \left(H^d\right)^{-1}\left[W^d_{\tau_{\tilde{d}_1}}\left(\mu^d_{\tau_{\tilde{d}_1}} - H^d\left(-\lambda^d_{t_{j-1}}\right)\right) + H^d\left(-\lambda^d_{t_{j-1}}\right)\right] \quad . \end{aligned} \tag{10.19}$$

for $d \in \mathcal{E}_{\tau^{\tilde{d}_1}}$. Just as in Section 10.2 we make the following

10.15 Assumption (Dependence Structure of Default Change)
The random variables $W_{\tau^{\tilde{d}_1}}$, $d = 1, \ldots, D$, of Equation (10.19) are assumed to be distributed to the copula L of the generalized default change variable $\Phi \sim H$, i.e.,

$$W_{\tau^{\tilde{d}_1}} = \left(W^1_{\tau^{\tilde{d}_1}}, \ldots, W^D_{\tau^{\tilde{d}_1}}\right)' \sim L \quad .$$

10.16 Proposition
Under filtration $\mathcal{G}_{t_{n-1}}$ the random variables $\Phi \sim H$ and $\Phi_{\tau^{\tilde{d}_1}} \sim H_{\tau^{\tilde{d}_1}}$ have the identical and unique copula function L, i.e.,

$$H_{\tau^{\tilde{d}_1}}\left(\Delta\phi^1_{\tau^{\tilde{d}_1}},\ldots,\Delta\phi^D_{\tau^{\tilde{d}_1}}\right) = L\left(H^1_{\tau^{\tilde{d}_1}}\left(\Delta\phi^1_{\tau^{\tilde{d}_1}}\right),\ldots,H^D_{\tau^{\tilde{d}_1}}\left(\Delta\phi^D_{\tau^{\tilde{d}_1}}\right)\right) \quad .$$

The proof is identical - except for the different notation - to the one of Proposition 10.11. ∎

We recall that the multivariate cdfs H and $H_{\tau^{\tilde{d}_1}}$ generally differ.

10.3.3 Resulting Impacts

As a consequence of the postulation

$$\lambda^{\tilde{d}_1}_t = \lambda^{\tilde{d}_1}_{\tau^{\tilde{d}_1}} \quad \text{for } t \in [\tau^{\tilde{d}_1}, T]$$

for the **defaulted** position $\tilde{d}_1$, the random variable $W^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}$ generating the default change $\Phi^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}$ is already fixed. Thus, a realization of $\Phi^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}$ must hold

$$\Delta\phi^{\tilde{d}_1}_{\tau^{\tilde{d}_1}} = \left(H^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}\right)^{-1}\left[\tilde{w}^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}\right] = \left(H^{\tilde{d}_1}\right)^{-1}\left[\tilde{w}^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}\left(\mu^{\tilde{d}_1}_{\tau^{\tilde{d}_1}} - H^{\tilde{d}_1}\left(-\lambda^{\tilde{d}_1}_{t_{n-1}}\right)\right) + H^{\tilde{d}_1}\left(-\lambda^{\tilde{d}_1}_{t_{n-1}}\right)\right] = 0 \ .$$

Via inversion, we receive

$$\tilde{w}^{\tilde{d}_1}_{\tau^{\tilde{d}_1}} = \frac{H^{\tilde{d}_1}(0) - H^{\tilde{d}_1}\left(-\lambda^{\tilde{d}_1}_{t_{n-1}}\right)}{\mu^{\tilde{d}_1}_{\tau^{\tilde{d}_1}} - H^{\tilde{d}_1}\left(-\lambda^{\tilde{d}_1}_{t_{n-1}}\right)} \quad . \tag{10.20}$$

Hence, the copula-distributed realizations

$$w_{\tau^{\tilde{d}_1}} = \left(w^1_{\tau^{\tilde{d}_1}},\ldots,w^D_{\tau^{\tilde{d}_1}}\right)' \in (0,1)^D$$

are now computed on the condition $W^{\tilde{d}_1}_{\tau^{\tilde{d}_1}} = \tilde{w}^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}$. This means that the generating random variable

$$W_{\tau^{\tilde{d}_1}} = \left(W^1_{\tau^{\tilde{d}_1}},\ldots,w^{\tilde{d}_1-1}_{\tau^{\tilde{d}_1}},W^{\tilde{d}_1}_{\tau^{\tilde{d}_1}} = \tilde{w}^{\tilde{d}_1}_{\tau^{\tilde{d}_1}},W^{\tilde{d}_1+1}_{\tau^{\tilde{d}_1}},\ldots,W^D_{\tau^{\tilde{d}_1}}\right)' \in (0,1)^D$$

is distributed according to the **conditional** copula[8]

$$\begin{aligned}
& L\left(w^1_{\tau^{\tilde{d}_1}},\ldots,w^{\tilde{d}_1-1}_{\tau^{\tilde{d}_1}},w^{\tilde{d}_1}_{\tau^{\tilde{d}_1}} = \tilde{w}^{\tilde{d}_1}_{\tau^{\tilde{d}_1}},w^{\tilde{d}_1+1}_{\tau^{\tilde{d}_1}},\ldots,w^D_{\tau^{\tilde{d}_1}}\right) \\
= \ & \mathbb{P}\left[W^1_{\tau^{\tilde{d}_1}} \le w^1_{\tau^{\tilde{d}_1}},\ldots,W^{\tilde{d}_1-1}_{\tau^{\tilde{d}_1}} \le w^{\tilde{d}_1-1}_{\tau^{\tilde{d}_1}},W^{\tilde{d}_1+1}_{\tau^{\tilde{d}_1}} \le w^{\tilde{d}_1+1}_{\tau^{\tilde{d}_1}},\ldots,W^D_{\tau^{\tilde{d}_1}} \le w^D_{\tau^{\tilde{d}_1}} \,|\, \left\{W^{\tilde{d}_1}_{\tau^{\tilde{d}_1}} = \tilde{w}^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}\right\} \vee \mathcal{G}_{t_{n-1}}\right] \\
= \ & \frac{\partial}{\partial w^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}} L\left(w^1_{\tau^{\tilde{d}_1}},\ldots,w^D_{\tau^{\tilde{d}_1}}\right)\Big|_{w^{\tilde{d}_1}_{\tau^{\tilde{d}_1}} = \tilde{w}^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}}
\end{aligned}$$

with $\tilde{w}^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}$ given in Equation (10.20). For the evaluation of $\tilde{w}^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}$, filtration $\mathcal{G}_{t_{n-1}}$ is inevitably required as the value $\lambda^{\tilde{d}_1}_{t_{n-1}}$ must be known. Conditioning on $\left\{W^{\tilde{d}_1}_{\tau^{\tilde{d}_1}} = \tilde{w}^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}\right\}$ implicitly assumes that we know the defaulted index at time $\tau^{\tilde{d}_1}$. This information is incorporated in the set $\mathcal{C}_{\tau^{\tilde{d}_1}}$. Consequently, evaluation of conditional margins and copulas as well as measurability of random changes must be reconsidered. We will tackle this problem later on.

[8] A derivative of a copula fulfills properties of Definition 2.1 under certain conditions. Anyway, it has range $[0,1]$ and is non-decreasing and is therefore interpretable as distribution function. For further details refer to [52] and [12], Section 4.1.

If no further default happens in $[\tau^{\tilde{d}_1}, t_n]$, the assumption $\lambda_t^{\tilde{d}_1} = \lambda_{\tau^{\tilde{d}_1}}^{\tilde{d}_1} = \lambda_{t_{n-1}}^{\tilde{d}_1}$ includes the restriction $\Delta\lambda_n^{\tilde{d}_1} = 0$ for the next random change at time t_n - equivalently to $\Delta\phi_{\tau^{\tilde{d}_1}}^{\tilde{d}_1} = 0$. Thus, it holds

$$\Delta\lambda_n^{\tilde{d}_1} = \left(F_n^{\tilde{d}_1}\right)^{-1}\left[\tilde{v}_n^{\tilde{d}_1}\right] = \left(F^{\tilde{d}_1}\right)^{-1}\left[\tilde{v}_n^{\tilde{d}_1}\left(\kappa_n^{\tilde{d}_1} - F^{\tilde{d}_1}\left(-\lambda_{\tau^{\tilde{d}_1}}^{\tilde{d}_1}\right)\right) + F^{\tilde{d}_1}\left(-\lambda_{\tau^{\tilde{d}_1}}^{\tilde{d}_1}\right)\right] = 0 \quad .$$

for the **defaulted** counterparty $\tilde{d}_1$. Via inversion, we receive

$$\tilde{v}_n^{\tilde{d}_1} = \frac{F^{\tilde{d}_1}(0) - F^{\tilde{d}_1}\left(\lambda_{\tau^{\tilde{d}_1}}^{\tilde{d}_1}\right)}{\kappa_n^{\tilde{d}_1} - F^{\tilde{d}_1}\left(\lambda_{\tau^{\tilde{d}_1}}^{\tilde{d}_1}\right)} = F_n^{\tilde{d}_1}(0) \quad . \tag{10.21}$$

As above, the generating random variable

$$V_n = \left(V_n^1, \ldots, V_n^{\tilde{d}_1 - 1}, V_n^{\tilde{d}_1} = \tilde{v}_n^{\tilde{d}_1}, V_n^{\tilde{d}_1 + 1}, \ldots, V_n^D\right)' \in (0,1)^D$$

is distributed according

$$\begin{aligned} & K\left(v_n^1, \ldots, v_n^{\tilde{d}_1 - 1}, v_n^{\tilde{d}_1} = \tilde{v}_n^{\tilde{d}_1}, v_n^{\tilde{d}_1 + 1}, \ldots, v_n^D\right) \\ = \; & \mathbb{P}\left[V_n^1 \le v_n^1, \ldots, V_n^{\tilde{d}_1 - 1} \le v_n^{\tilde{d}_1 - 1}, V_n^{\tilde{d}_1 + 1} \le v_n^{\tilde{d}_1 + 1}, \ldots, V_n^D \le v_n^D \middle| \left\{V_n^{\tilde{d}_1} = \tilde{v}_n^{\tilde{d}_1}\right\} \vee \mathcal{G}_{\tau^{\tilde{d}_1}}\right] \\ = \; & \frac{\partial}{\partial v_n^{\tilde{d}_1}} K\left(v_n^1, \ldots, v_n^D\right)\Big|_{v_n^{\tilde{d}_1} = \tilde{v}_n^{\tilde{d}_1}} \end{aligned}$$

with $\tilde{v}_n^{\tilde{d}_1}$ as in Equation (10.21). Conditioning on $\mathcal{G}_{t_{n-1}}$ is needed as the evaluation of $\tilde{v}_m^d$ for $m = n, \ldots, N$ involves $\lambda_{\tau^{\tilde{d}_1}}^{\tilde{d}_1} = \lambda_{t_{n-1}}^{\tilde{d}_1}$, conditioning on $\left\{V_n^{\tilde{d}_1} = \tilde{v}_n^{\tilde{d}_1}\right\}$ provides information about the defaulted index to which the derivative is taken. This is contained in filtration $\mathcal{F}_{t_n}$.

Note that for the construction of V_n only the defaulted positions must be known. For the random change Λ_n this is different, see Equation (10.12), as the value of $\lambda_{\tau^{\tilde{d}_1}}$ must be revealed for conditional margins. Thus, the information for the generating random variable V_n and for the random change Λ_n has to cover

- the defaulted index at time t_n - i.e., $\mathcal{F}_{t_n}$,
- the information about the value of the intensity process before t_n - i.e., $\mathcal{G}_{\tau^{\tilde{d}_1}}$, and
- the information about the conditional copula at t_n - i.e., $\left\{V_n^{\tilde{d}_1} = \tilde{v}_n^{\tilde{d}_1}\right\} \subset \mathcal{F}_{t_n} \vee \mathcal{G}_{\tau^{\tilde{d}_1}}$

which leads to a different demand on the measurability for the random variable Λ_n. If we define

$$\mathcal{H}_{t_n} = \mathcal{F}_{t_n} \vee \mathcal{G}_{\tau^{\tilde{d}_1}}$$

and assume that Λ_n is $\mathcal{H}_{t_n}$-measurable, the available information about default, intensity, and copula will be sufficient to generate Λ_n and therefore λ_{t_n}.

Above considerations have to be generalized and extended to an arbitrary number of defaults. Besides the mentioned impacts on

- random changes (assumption of enlarged measurability for Λ and Φ),
- generating random variables of K and L (V and W),
- copulas (partial derivatives), and
- the intensity process,

it also leaves to be shown to what extent dependencies of Λ and Φ change.

10.4 Including Defaults

The extension to the general case of several defaults has to be handled attentively. First, we recall the definition of realized default times

$$\tau(\mathcal{C}_t) = \left(\tau^{\tilde{d}_1}, \ldots, \tau^{\tilde{d}_C}\right)' \leq t \in [0,T] \quad .$$

of Notation 9.16. We suppose that default times are stated in chronological order, i.e.,

$$\tau^{\tilde{d}_1} \leq \tau^{\tilde{d}_2} \leq \ldots \leq \tau^{\tilde{d}_{C-1}} \leq \tau^{\tilde{d}_C} \quad .$$

The upper right index indicates the original position $\tilde{d}_1 \in \mathcal{D}$ which has defaulted first. Further, we denote λ_{t^-} as value of the intensity process directly before $t \in [0,T]$ and sharpen

10.17 Definition (Random Default Change)
The random variable $\Phi_\tau = \left(\Phi_\tau^1, \ldots, \Phi_\tau^D\right)'$ is called *random default change* of the intensity process λ_t and is **instantaneously** triggered at all realized default times. The time τ indicates the next occurring default[9].

Additionally, we specify the measurability of the random changes and revise their conditional evaluations.

10.4.1 Measurability of Conditional Random Changes Λ_n and Φ_τ

Let $s \in (t_{n-1}, t_n)$ be the first default time in this very interval. A random default change Φ_s is triggered. For the generation of Φ_s, we need the information about defaulted indices up to and **inclusively** time s which is contained in $\mathcal{F}_s$[10]. Further, the level of the intensity process before default time s (i.e., $\lambda_{s^-} = \lambda_{t_{n-1}}$) for each position (default and non-defaulted) is required. This information is provided by filtration $\mathcal{G}_{t_{n-1}}$[11].

Suppose, the next default happens at $t \in (s, t_n)$. Again, filtration $\mathcal{F}_t$ is used for Φ_t. However, the information about the level of the intensity before t is now comprised in filtration $\mathcal{G}_s$ due to the default change Φ_s at time s (shift from $\lambda_{t_{n-1}}$ to λ_s).

The distance between two defaults (here s and t) can be infinitesimal small in the continuous-time setup. Therefore, we denote

$$\mathcal{G}_{t^-} := \sigma\left(\lambda_s : 0 \leq s < t\right) \quad .$$

Filtration $\mathcal{G}_{t^-}$ comprehends the level of the intensity process directly before t and is applied for

10.18 Definition (Filtration for Random Changes)
Let $t \in [0,T]$ and $\mathcal{G}_t$, $\mathcal{G}_{t^-}$ and $\mathcal{F}_t$ be defined as above. The information flow

$$\mathcal{H}_t := \mathcal{F}_t \vee \mathcal{G}_{t^-}$$

is denoted as *measurability filtration*. Random changes Λ and Φ are assumed to be $\mathcal{H}_t$-measurable.

Due to the altered measurability and the extension to several defaults, the design of the random variables Λ_n and Φ_τ is affected. On the one hand the marginal constructions (via the generating variables V_n and W_τ) are concerned, on the other hand their dependence structure (the multivariate cdfs and copulas K and L) is influenced.

[9] For $\mathcal{C}_t = \left\{\tilde{d}_1, \ldots, \tilde{d}_C\right\}$ at time $t \in [0,T]$, this implies $\tau = \tau^{\tilde{d}_{C+1}}$ with $t < \tau$, for instance.
[10] Note that $\mathcal{F}_{t_{n-1}} \neq \mathcal{F}_s$.
[11] In (t_{n-1}, s) no further change is accrued. In addition, it holds $\mathcal{G}_s \neq \mathcal{G}_{t_{n-1}}$ due to Φ_s.

10.4.2 Impact on Margins of Random Changes

For the non-defaulted positions, the random change Λ_n^d, Equation (10.12), and the random default change Φ_τ^d, Equation (10.18), must be revised to

$$\Lambda_n^d = \left(F^d\right)^{-1}\left[V_n^d\left(\kappa_n^d - F^d\left(-\lambda_{t_n^-}^d|\mathcal{H}_{t_n}\right)\right) + F^d\left(-\lambda_{t_n^-}^d|\mathcal{H}_{t_n}\right)\right] \tag{10.22}$$

for $d \in \mathcal{E}_{t_n}$, $n = 1, \ldots, N$, with $V_n^d \sim U(0,1)$ and possible choices $\kappa_n^d = \begin{cases} 1 \\ F^d\left(-\lambda_{t_n^-}^d|\mathcal{H}_{t_n}\right) \end{cases}$ and

$$\Phi_\tau^d = \left(H^d\right)^{-1}\left[W_\tau^d\left(\mu_\tau^d - H^d\left(-\lambda_{\tau^-}^d|\mathcal{H}_\tau\right)\right) + H^d\left(-\lambda_{\tau^-}^d|\mathcal{H}_\tau\right)\right] \tag{10.23}$$

for $d \in \mathcal{E}_\tau$, $\tau \in (0,T]$, with $W_\tau^d \sim U(0,1)$ and possible choices $\mu_\tau^d = \begin{cases} 1 \\ H^d\left(-\lambda_{\tau^-}^d|\mathcal{H}_\tau\right) \end{cases}$.

Their cdfs are now given as

$$\begin{aligned} F_n^d(\Delta\lambda_n^d(t_n)) :=& \quad F^d\left(\Delta\lambda_n^d|\left\{-\lambda_{t_n^-}^d < \Lambda^d \le \kappa_n^d\right\} \vee \mathcal{H}_{t_n}\right) \\ =& \quad \frac{\mathbb{P}[\Lambda^d \le \Delta\lambda_n^d] - \mathbb{P}\left[\Lambda^d \le -\lambda_{t_n^-}^d|\mathcal{H}_{t_n}\right]}{\mathbb{P}\left[\Lambda^d \le \left(F^d\right)^{-1}[\kappa_n^d]\right] - \mathbb{P}\left[\Lambda^d \le -\lambda_{t_n^-}^d|\mathcal{H}_{t_n}\right]} \\ =& \quad \frac{F^d(\Delta\lambda_n^d) - F^d\left(-\lambda_{t_n^-}^d|\mathcal{H}_{t_n}\right)}{\kappa_n^d - F^d\left(-\lambda_{t_n^-}^d|\mathcal{H}_{t_n}\right)} \quad \text{for } d \in \mathcal{E}_{t_n} \text{ and } n = 1, \ldots, N \end{aligned} \tag{10.24}$$

and

$$\begin{aligned} H_\tau^d(\Delta\phi_\tau^d(\tau)) :=& \quad H^d\left(\Delta\phi_\tau^d|\left\{-\lambda_{\tau^-}^d < \Phi^d \le \mu_\tau^d\right\} \vee \mathcal{H}_\tau\right) \\ =& \quad \frac{\mathbb{P}[\Phi^d \le \Delta\phi_\tau^d] - \mathbb{P}\left[\Phi^d \le -\lambda_{\tau^-}^d|\mathcal{H}_\tau\right]}{\mathbb{P}\left[\Phi^d \le \left(H^d\right)^{-1}[\mu_\tau^d]\right] - \mathbb{P}\left[\Phi^d \le -\lambda_{\tau^-}^d|\mathcal{H}_\tau\right]} \\ =& \quad \frac{H^d(\Delta\phi_\tau^d) - H^d\left(-\lambda_{\tau^-}^d|\mathcal{H}_\tau\right)}{\mu_\tau^d - H^d\left(-\lambda_{\tau^-}^d|\mathcal{H}_\tau\right)} \quad \text{for } d \in \mathcal{E}_\tau \text{ and } \tau \in (0,T] \quad . \end{aligned} \tag{10.25}$$

For the defaulted positions, the realizations of the random changes

Λ_n^d for $d \in \mathcal{C}_{t_m}$, $m = n, \ldots, N$ and $\tau^d \le t_m$ and

Φ_τ^d, for $d \in \mathcal{C}_\tau$, $\tau \in (0,T]$ and $\tau^d \le \tau$

have to satisfy

$$\begin{aligned} \Delta\lambda_n^d &= \left(F_n^d\right)^{-1}[\tilde{v}_n^d] = \left(F^d\right)^{-1}\left[\tilde{v}_n^d\left(\kappa_n^d - F^d\left(-\lambda_{t_n^-}^d|\mathcal{H}_{t_n}\right)\right) + F^d\left(-\lambda_{t_n^-}^d|\mathcal{H}_{t_n}\right)\right] = 0 \quad , \\ \Delta\phi_\tau^d &= \left(H_\tau^d\right)^{-1}[\tilde{w}_\tau^d] = \quad \left(H^d\right)^{-1}\left[\tilde{w}_\tau^d\left(\mu_\tau^d - H^d\left(-\lambda_{\tau^-}^d|\mathcal{H}_\tau\right)\right) + H^d\left(-\lambda_{\tau^-}^d|\mathcal{H}_\tau\right)\right] = 0 \quad . \end{aligned}$$

These equations are equivalent to

$$\begin{aligned} \Delta\lambda_n^d &= \left(F^d\right)^{-1}\left[\tilde{v}_n^d\left(\kappa_{\tau^d}^d - F^d\left(-\lambda_{\tau^d-}^d|\mathcal{H}_{\tau^d}\right)\right) + F^d\left(-\lambda_{\tau^d-}^d|\mathcal{H}_{\tau^d}\right)\right] \quad = 0 \quad , \\ \Delta\phi_\tau^d &= \left(H^d\right)^{-1}\left[\tilde{w}_\tau^d\left(\mu_{\tau^d}^d - H^d\left(-\lambda_{\tau^d-}^d|\mathcal{H}_{\tau^d}\right)\right) + H^d\left(-\lambda_{\tau^d-}^d|\mathcal{H}_{\tau^d}\right)\right] \quad = 0 \end{aligned}$$

as filtrations $\mathcal{H}_{t_n}$ and $\mathcal{H}_\tau$ contain the same information as $\mathcal{H}_{\tau^d}$ if restricted to the relevant position for the one-dimensional case. Further, it holds $\lambda_t^d = \lambda_{\tau^d}^d$ for $d \in \mathcal{C}_t$ and $t \ge \tau^d$. Via inversion we receive

$$\tilde{v}_n^d = F_n^d(0) = \frac{F^d(0) - F^d\left(-\lambda^d_{\tau^d-}|\mathcal{H}_{\tau^d}\right)}{\kappa^d_{\tau^d} - F^d\left(-\lambda^d_{\tau^d-}|\mathcal{H}_{\tau^d}\right)} =: F^d_{\tau^d}(0,t_n) \tag{10.26}$$

for $\tau^d \leq t_n,\ n = 1,\ldots,N$, and $d \in \mathcal{C}_{t_n}$ as well as

$$\tilde{w}_{\tau^d}^d = H_\tau^d\left(\Delta\phi^d_{\tau^d}\right) = \frac{H^d(0) - H^d\left(-\lambda^d_{\tau^d-}|\mathcal{H}_{\tau^d}\right)}{\mu^d_{\tau^d} - H^d\left(-\lambda^d_{\tau^d-}|\mathcal{H}_{\tau^d}\right)} =: H^d_{\tau^d}(0,\tau) \tag{10.27}$$

for $\tau^d \leq \tau$ and $d \in \mathcal{C}_\tau$.

Locking the process in case of defaults implies that no further information is needed for the defaulted position. Therefore, filtration $\mathcal{H}_{\tau^d}$ would be sufficient for one dimension. However, fixing generating variables V_n^d to $\tilde{v}_n^d$ and W_τ^d to $\tilde{w}_\tau^d$ has consequences on the multivariate dependence.

10.4.3 Impacts on Copulas of Random Changes

For the multivariate extension, the generating random variables are denoted as

$$\begin{aligned} V_n(\mathcal{E}_t) &= \left(V_n^{d_1}(t),\ldots,V_n^{d_E}(t)\right)' \in (0,1)^E \\ \text{and } W_\tau(\mathcal{E}_t) &= \left(W_\tau^{d_1}(t),\ldots,W_\tau^{d_E}(t)\right)' \in (0,1)^E \end{aligned}$$

for the non-defaulted positions $d_e \in \mathcal{E}_t$, $e = 1,\ldots,E$ (as upper right index) with

- the random change Λ_n^d at step $t_n \in (t,T]$ (as lower right index) and
- the random default change Φ_τ^d at the next default time $\tau \in (t,T]$ (as lower right index)

under information $\mathcal{H}_t$ (defined as argument t) for Equations (10.22) and (10.23).

Note that $V_n(\mathcal{E}_s)$ and $V_n(\mathcal{E}_t)$ - $W_\tau(\mathcal{E}_s)$ and $W_\tau(\mathcal{E}_t)$ respectively - with $s < t$ can be different due to the altered information from $\mathcal{H}_s$ to $\mathcal{H}_t$. Thus, generating variables depend on the available information at the given time which reasons the additional argument.

Further, the realizations $v_n(\mathcal{C}_t)$ and $w_\tau(\mathcal{C}_t)$ of the random variables

$$\begin{aligned} V_n(\mathcal{C}_t) &= \left(V_n^{\tilde{d}_1}(t),\ldots,V_n^{\tilde{d}_C}(t)\right)' \in (0,1)^C \\ \text{and } W_\tau(\mathcal{C}_t) &= \left(W_\tau^{\tilde{d}_1}(t),\ldots,W_\tau^{\tilde{d}_C}(t)\right)' \in (0,1)^C \end{aligned}$$

for defaulted positions $\tilde{d}_c \in \mathcal{C}_t$, $c = 1,\ldots,C$, (as upper right index) are fixed as

$$\begin{aligned} v_n(\mathcal{C}_t) &= \tilde{v}_n(\mathcal{C}_t) \quad \text{for } \tau(\mathcal{C}_t) \leq t \leq t_n \\ \text{and } w_\tau(\mathcal{C}_t) &= \tilde{w}_\tau(\mathcal{C}_t) \quad \text{for } \tau(\mathcal{C}_t) \leq t \leq \tau \end{aligned}$$

under information $\mathcal{H}_t$. Due to Equations (10.26) and (10.27) it holds

$$\begin{aligned} \tilde{v}_n(\mathcal{C}_t) &= \left(\tilde{v}_n^{\tilde{d}_1}(t),\ldots,\tilde{v}_n^{\tilde{d}_C}(t)\right)' = \left(F^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}(0,t),\ldots,F^{\tilde{d}_C}_{\tau^{\tilde{d}_C}}(0,t)\right)' \text{ with } \tau(\mathcal{C}_t) \leq t \leq t_n \text{ under } \mathcal{H}_t \\ \text{and } \tilde{w}_\tau(\mathcal{C}_t) &= \left(\tilde{w}_\tau^{\tilde{d}_1}(t),\ldots,\tilde{w}_\tau^{\tilde{d}_C}(t)\right)' = \left(H^{\tilde{d}_1}_{\tau^{\tilde{d}_1}}(0,t),\ldots,H^{\tilde{d}_C}_{\tau^{\tilde{d}_C}}(0,t)\right)' \text{ with } \tau(\mathcal{C}_t) \leq t \leq \tau \text{ under } \mathcal{H}_t. \end{aligned}$$

Again, n and τ indicate the next random and default change. The information at hand is subscripted by argument t. It marks the defaulted indices and brings along all defaults up to time t, i.e., $\tau(\mathcal{C}_t) \leq t$.

Note that the evaluation of the intensity process λ_t for the next change at t_n or τ is now an **expected value** if $t < t_n$ or $t < \tau$. Only if $t = t_n$ or $t = \tau$, the value of λ_t is known under $\mathcal{H}_t$.

For sake of readability, we abbreviate

$$V_n(t) = (V_n(\mathcal{E}_t); V_n(\mathcal{C}_t) = \tilde{v}_n(\mathcal{C}_t)) \quad \text{and} \quad W_\tau(t) = (W_\tau(\mathcal{E}_t); W_\tau(\mathcal{C}_t) = \tilde{w}_\tau(\mathcal{C}_t))$$

with realizations $v_n(t)$ and $w_\tau(t)$ and fixed $\tilde{v}_n(\mathcal{C}_t)$ and $\tilde{w}_\tau(\mathcal{C}_t)$ as above. Notations $V_n(t)$ and $W_\tau(t)$ do not intend to permute the original order of V_n and W_τ.

As observed in Subsection 10.3.3, the locked random variables also influence the copula functions. Therefore, we make the following

10.19 Assumption (Differentiability of Copulas K and L)
Copulas K and L of the random changes for the intensity process are supposed to be sufficiently differentiable with respect to all defaultable indices.

We further assume that the generating random variables $V_n(t)$ and $W_\tau(t)$ are distributed according to the conditional copulas

$$\begin{aligned} K(t) = K(v_n(t)) :&= K(v_n(\mathcal{E}_t); v_n(\mathcal{C}_t) = \tilde{v}_n(\mathcal{C}_t)) \\ &= \frac{\partial^C}{\partial v_n(\mathcal{C}_t)} K(v_n(\mathcal{E}_t); v_n(\mathcal{C}_t)) \Big|_{v_n(\mathcal{C}_t)=\tilde{v}_n(\mathcal{C}_t)} \\ &= \mathbb{P}[V_n(\mathcal{E}_t) \le v_n(\mathcal{E}_t) | \mathcal{H}_t] \quad \text{for } t \le t_n \text{ and } n = 1, \dots, N \end{aligned} \tag{10.28}$$

and

$$\begin{aligned} L(t) = L(w_\tau(t)) :&= L(w_\tau(\mathcal{E}_t); w_\tau(\mathcal{C}_t) = \tilde{w}_\tau(\mathcal{C}_t)) \\ &= \frac{\partial^C}{\partial w_\tau(\mathcal{C}_t)} L(w_\tau(\mathcal{E}_t); w_\tau(\mathcal{C}_t)) \Big|_{w_\tau(\mathcal{C}_t)=\tilde{w}_\tau(\mathcal{C}_t)} \\ &= \mathbb{P}[W_\tau(\mathcal{E}_t) \le w_\tau(\mathcal{E}_t) | \mathcal{H}_t] \quad \text{for } t \le \tau \text{ and } \tau \in (0, T] \quad . \end{aligned} \tag{10.29}$$

Filtration $\mathcal{H}_t$ contains

- the information about the value of the intensity process directly before t (i.e., λ_{t^-}),
- the index of the defaulted obligors at time t, and therefore
- the evaluation of the generating random variable of all defaulted positions $\{V(\mathcal{C}_t) = \tilde{v}(\mathcal{C}_t)\}$.

The arguments of K and L ($v_n(t)$ and $w_\tau(t)$) indicate the next step or default (n and τ) and the available information (t for $\mathcal{H}_t$)[12]. The revised dependence structures of the new random variables

- $\Lambda_n(t) := (\Lambda_n(\mathcal{E}_t); \Lambda_n(\mathcal{C}_t))$ through $V_n(t)$ with
  - $\Lambda_n(\mathcal{E}_t) = \left(\Lambda_n^{d_1}(t), \dots, \Lambda_n^{d_E}(t)\right)'$ for $t \le t_n$ and $d_e \in \mathcal{E}_t$, $e = 1, \dots, E$
  - $\Lambda_n(\mathcal{C}_t) = \left(\Lambda_n^{\tilde{d}_1}(t), \dots, \Lambda_n^{\tilde{d}_C}(t)\right)' = 0$ for $\tau(\mathcal{C}_t) \le t \le t_n$ and $d_c \in \mathcal{C}_t$, $c = 1, \dots, C$
  - $\Lambda_n(t) \sim F_n(t)$ and copula $K(t)$ of Equation (10.28).
- $\Phi_\tau(t) := (\Phi_\tau(\mathcal{E}_t); \Phi_\tau(\mathcal{C}_t))$ for $t \le \tau$ through $W_\tau(t)$ with
  - $\Phi_\tau(\mathcal{E}_t) = \left(\Phi_\tau^{d_1}(t), \dots, \Phi_\tau^{d_E}(t)\right)'$ for $t \le \tau$ and $d_e \in \mathcal{E}_t$, $e = 1, \dots, E$
  - $\Phi_\tau(\mathcal{C}_t) = \left(\Phi_\tau^{\tilde{d}_1}(t), \dots, \Phi_\tau^{\tilde{d}_C}(t)\right)' = 0$ for $\tau(\mathcal{C}_t) \le t \le \tau$ and $d_c \in \mathcal{C}_t$, $c = 1, \dots, C$
  - $\Phi_\tau(t) \sim H_\tau(t)$ and copula $L(t)$ of Equation (10.29).

are comparable to the original random variables Λ and Φ.

[12] Filtration $\mathcal{H}_t$ is needed for the evaluation of $\tilde{v}(\mathcal{C}_t)$. If $\tilde{v}(\mathcal{C}_t)$ did not comprise former values of λ_t, it would be sufficient to use $\mathcal{C}_t$ as condition.

10.20 Proposition
Let $t \in (0,T]$ and sets $\mathcal{E}_t$, $\mathcal{C}_t$ be given. The original random variables[13]

$$\begin{aligned} \Lambda = \Lambda(t) \sim F \text{ with } \Lambda(t) = (\Lambda(\mathcal{E}_t), \Lambda(\mathcal{C}_t)) \text{ given } \Lambda(\mathcal{C}_t) = 0 \\ \text{and} \quad \Phi = \Phi(t) \sim H \text{ with } \Phi(t) = (\Phi(\mathcal{E}_t), \Phi(\mathcal{C}_t)) \text{ given } \Phi(\mathcal{C}_t) = 0 \end{aligned}$$

and the random change variables

$$\Lambda_n(t) \sim F_n(t) \quad \text{and} \quad \Phi_\tau(t) \sim H_\tau(t)$$

have copulas $K(t)$ and $L(t)$, i.e., it holds

$$\begin{aligned} & F\left(\Delta\lambda_n^{d_1}(t), \ldots, \Delta\lambda_n^{d_E}(t), \Delta\lambda_n^{\tilde{d}_1}(t) = 0, \ldots, \Delta\lambda_n^{\tilde{d}_C}(t) = 0\right) \\ & \qquad = \quad K\left(F^{d_1}\left(\Delta\lambda_n^{d_1}(t)\right), \ldots, F^{d_E}\left(\Delta\lambda_n^{d_E}(t)\right), F^{\tilde{d}_1}(0,t), \ldots, F^{\tilde{d}_C}(0,t)\right) \quad , \end{aligned}$$

$$\begin{aligned} & F_n\big(\Delta\lambda_n^{d_1}(t), \ldots, \Delta\lambda_n^{d_E}(t), \Delta\lambda_n^{\tilde{d}_1}(t) = 0, \ldots, \Delta\lambda_n^{\tilde{d}_C}(t) = 0\big) \\ & \qquad = \quad K\left(F_n^{d_1}\left(\Delta\lambda_n^{d_1}(t)\right), \ldots, F_n^{d_E}\left(\Delta\lambda_n^{d_E}(t)\right), F_n^{\tilde{d}_1}(0,t), \ldots, F^{\tilde{d}_C}(0,t)\right) \end{aligned}$$

for the random change variables and

$$\begin{aligned} & H\left(\Delta\phi_\tau^{d_1}(t), \ldots, \Delta\phi_\tau^{d_E}(t), \Delta\phi_\tau^{\tilde{d}_1}(t) = 0, \ldots, \Delta\phi_\tau^{\tilde{d}_C}(t) = 0\right) \\ & \qquad = \quad L\left(H_\tau^{d_1}\left(\Delta\phi_\tau^{d_1}(t)\right), \ldots, H_\tau^{d_E}\left(\Delta\phi_\tau^{d_E}(t)\right), H_\tau^{\tilde{d}_1}(0,t), \ldots, H^{\tilde{d}_C}(0,t)\right) \quad , \end{aligned}$$

$$\begin{aligned} & H_\tau\left(\Delta\phi_\tau^{d_1}(t), \ldots, \Delta\phi_\tau^{d_E}(t), \Delta\phi_\tau^{\tilde{d}_1}(t) = 0, \ldots, \Delta\phi_\tau^{\tilde{d}_C}(t) = 0\right) \\ & \qquad = \quad L\left(H_\tau^{d_1}\left(\Delta\phi_\tau^{d_1}(t)\right), \ldots, H_\tau^{d_E}\left(\Delta\phi_\tau^{d_E}(t)\right), H_\tau^{\tilde{d}_1}(0,t), \ldots, H^{\tilde{d}_C}(0,t)\right) \end{aligned}$$

for the default change variables.

Proof

The proof is similar to the one of Proposition 10.11. For the first part, w.l.o.g. let $t \in (t_{n-1}, t_n)$ be fixed and $\mathcal{E}_t$, $\mathcal{C}_t$ and $\mathcal{H}_t$ be given. For the original random variable $\Lambda = \Lambda(t) \sim F(t)$ given $\Lambda(\mathcal{C}_t) = 0$, the next random change at t_n and under $\mathcal{H}_t$, we have

$$\begin{aligned} F\left(\Delta\lambda_n(t)\right) &= F\left(\Delta\lambda^{d_1}(t), \ldots, \Delta\lambda^{d_E}(t), \Delta\lambda^{\tilde{d}_1}(t) = 0, \ldots, \Delta\lambda^{\tilde{d}_C}(t) = 0\right) \\ &= \mathbb{P}\left[\Lambda(\mathcal{E}_t) \le \Delta\lambda_n(\mathcal{E}_t) \mid \{\Lambda(\mathcal{C}_t) = 0\} \vee \mathcal{H}_t\right] \\ &= \mathbb{P}\Big[\Lambda^{d_1}(t) \le \Delta\lambda_n^{d_1}(t), \ldots, \Lambda^{d_E}(t) \le \Delta\lambda_n^{d_E}(t) \mid \\ &\qquad \left\{\Lambda^{\tilde{d}_1}(t) = 0\right\} \vee \ldots \vee \left\{\Lambda^{\tilde{d}_C}(t) = 0\right\} \vee \mathcal{H}_t\Big] \quad . \\ &= \mathbb{P}\Big[F^{d_1}\left(\Lambda^{d_1}(t)\right) \le F^{d_1}\left(\Delta\lambda_n^{d_1}(t)\right), \ldots, F^{d_E}\left(\Lambda^{d_E}(t)\right) \le F^{d_E}\left(\Delta\lambda_n^{d_E}(t)\right) \mid \\ &\qquad \left\{F^{\tilde{d}_1}\left(\Lambda^{\tilde{d}_1}(t)\right) = F^{\tilde{d}_1}(0,t)\right\} \vee \ldots \vee \left\{F^{\tilde{d}_C}\left(\Lambda^{\tilde{d}_C}(t)\right) = F^{\tilde{d}_C}(0,t)\right\} \vee \mathcal{H}_t\Big] \end{aligned}$$

[13] Here as well, the argument t indicates filtration $\mathcal{H}_t \supset \mathcal{F}_t$ which in turn determines the defaulted positions $\mathcal{C}_t$.

As before, with $F^d\left(\Lambda^d\right) = F^d\left(\Lambda^d(t)\right) = V^d(t) = V^d$ this implies

$$\begin{aligned}
F\left(\Delta\lambda_n(t)\right) &= \mathbb{P}\Big[V^{d_1}(t) \le F^{d_1}\left(\Delta\lambda_n^{d_1}(t)\right), \ldots, V^{d_E}(t) \le F^{d_E}\left(\Delta\lambda_n^{d_E}(t)\right) \mid \\
&\qquad \left\{V^{\tilde{d}_1}(t) = F^{\tilde{d}_1}(0,t)\right\} \vee \ldots \vee \left\{V^{\tilde{d}_C}(t) = F^{\tilde{d}_C}(0,t)\right\} \vee \mathcal{H}_t\Big] \\
&= \frac{\partial^C}{\partial v(\mathcal{C}_t)} K\left(F^{d_1}\left(\Delta\lambda_n^{d_1}(t)\right), \ldots, F^{d_E}\left(\Delta\lambda_n^{d_E}(t)\right); v(\mathcal{C}_t)\right)\Big|_{v(\mathcal{C}_t)=\left(F^{\tilde{d}_1}(0,t),\ldots,F^{\tilde{d}_C}(0,t)\right)'} \\
&= K\left(F^{d_1}\left(\Delta\lambda_n^{d_1}(t)\right), \ldots, F^{d_E}\left(\Delta\lambda_n^{d_E}(t)\right), F^{\tilde{d}_1}(0,t), \ldots, F^{\tilde{d}_C}(0,t)\right) \quad .
\end{aligned}$$

For the model random variable $\Lambda_n = \Lambda_n(t) \sim F_n$ given $\Lambda_n(\mathcal{C}_t) = 0$, it holds

$$\begin{aligned}
&F_n\left(\Delta\lambda_n(t)\right) \\
&= F_n\left(\Delta\lambda^{d_1}(t), \ldots, \Delta\lambda^{d_E}(t), \Delta\lambda^{\tilde{d}_1}(t) = 0, \ldots, \Delta\lambda^{\tilde{d}_C}(t) = 0\right) \\
&= \mathbb{P}\left[\Lambda_n(\mathcal{E}_t) \le \Delta\lambda_n(\mathcal{E}_t) \mid \{\Lambda_n(\mathcal{C}_t) = 0\} \vee \mathcal{H}_t\right] \\
&= \mathbb{P}\left[\Lambda_n^{d_1}(t) \le \Delta\lambda_n^{d_1}(t), \ldots, \Lambda_n^{d_E}(t) \le \Delta\lambda_n^{d_E}(t) \mid \left\{\Lambda_n^{\tilde{d}_1}(t) = 0\right\} \vee \ldots \vee \left\{\Lambda_n^{\tilde{d}_C}(t) = 0\right\} \vee \mathcal{H}_t\right]
\end{aligned}$$

$$\begin{aligned}
&= \mathbb{P}\Big[\left(F^{d_1}\right)^{-1}\left[V_n^{d_1}(t)\left(\kappa_t^{d_1} - F^{d_1}\left(-\lambda_{t^-}^{d_1}|\mathcal{H}_t\right)\right) + F^{d_1}\left(-\lambda_{t^-}^{d_1}|\mathcal{H}_t\right)\right] \le \Delta\lambda_n^{d_1}(t), \ldots, \\
&\qquad \left(F^{d_E}\right)^{-1}\left[V_n^{d_E}(t)\left(\kappa_t^{d_E} - F^{d_E}\left(-\lambda_{t^-}^{d_E}|\mathcal{H}_t\right)\right) + F^{d_E}\left(-\lambda_{t^-}^{d_E}|\mathcal{H}_t\right)\right] \le \Delta\lambda_n^{d_E}(t) \mid \\
&\qquad \left\{\left(F^{\tilde{d}_1}\right)^{-1}\left[V_n^{\tilde{d}_1}(t)\left(\kappa_t^{\tilde{d}_1} - F^{\tilde{d}_1}\left(-\lambda_{t^-}^{\tilde{d}_1}|\mathcal{H}_t\right)\right) + F^{\tilde{d}_1}\left(-\lambda_{t^-}^{\tilde{d}_1}|\mathcal{H}_t\right)\right] = 0\right\} \vee \ldots \vee \\
&\qquad \left\{\left(F^{\tilde{d}_C}\right)^{-1}\left[V_n^{\tilde{d}_C}(t)\left(\kappa_t^{\tilde{d}_C} - F^{\tilde{d}_C}\left(-\lambda_{t^-}^{\tilde{d}_C}|\mathcal{H}_t\right)\right) + F^{\tilde{d}_C}\left(-\lambda_{t^-}^{\tilde{d}_C}|\mathcal{H}_t\right)\right] = 0\right\} \vee \mathcal{H}_t\Big]
\end{aligned}$$

$$\begin{aligned}
&= \mathbb{P}\Bigg[V_n^{d_1}(t) \le \frac{F^{d_1}(\Delta\lambda_n^{d_1}(t)) - F^{d_1}\left(-\lambda_{t^-}^{d_1}|\mathcal{H}_t\right)}{\kappa_t^{d_1} - F^{d_1}\left(-\lambda_{t^-}^{d_1}|\mathcal{H}_t\right)}, \ldots, V_n^{d_E}(t) \le \frac{F^{d_E}(\Delta\lambda_n^{d_E}(t)) - F^{d_E}\left(-\lambda_{t^-}^{d_E}|\mathcal{H}_t\right)}{\kappa_t^{d_E} - F^{d_E}\left(-\lambda_{t^-}^{d_E}|\mathcal{H}_t\right)} \mid \\
&\quad \left\{V_n^{\tilde{d}_1}(t) = \frac{F^{\tilde{d}_1}(0) - F^{\tilde{d}_1}\left(-\lambda_{t^-}^{\tilde{d}_1}|\mathcal{H}_t\right)}{\kappa_t^{\tilde{d}_1} - F^{\tilde{d}_1}\left(-\lambda_{t^-}^{\tilde{d}_1}|\mathcal{H}_t\right)}\right\} \vee \ldots \vee \left\{V_n^{\tilde{d}_C}(t) = \frac{F^{\tilde{d}_C}(0) - F^{\tilde{d}_C}\left(-\lambda_{t^-}^{\tilde{d}_C}|\mathcal{H}_t\right)}{\kappa_t^{\tilde{d}_C} - F^{\tilde{d}_C}\left(-\lambda_{t^-}^{\tilde{d}_C}|\mathcal{H}_t\right)}\right\} \vee \mathcal{H}_t\Bigg]
\end{aligned}$$

$$\begin{aligned}
&= \mathbb{P}\Big[V_n^{d_1}(t) \le F_n^{d_1}\left(\Delta\lambda_n^{d_1}(t)\right), \ldots, V_n^{d_E}(t) \le F_n^{d_E}\left(\Delta\lambda_n^{d_E}(t)\right) \mid \\
&\qquad \left\{V_n^{\tilde{d}_1}(t) = F_{\tau^{\tilde{d}_1}}^{\tilde{d}_1}(0,t)\right\} \vee \ldots \vee \left\{V_n^{\tilde{d}_C}(t) = F_{\tau^{\tilde{d}_C}}^{\tilde{d}_C}(0,t)\right\} \vee \mathcal{H}_t\Big] \quad .
\end{aligned}$$

For the last equation note that $F_n^{\tilde{d}_c}(0,t) = F_{\tau^{\tilde{d}_c}}^{\tilde{d}_c}(0,t)$ due to $\lambda_{t_n^-}^{\tilde{d}_c} = \lambda_{\tau^{\tilde{d}_c}}^{\tilde{d}_c}$ for $c = 1, \ldots, C$. Further, we receive

$$\begin{aligned}
&F_n\left(\Delta\lambda_n(t)\right) \\
&= \frac{\partial^C}{\partial v_n(\mathcal{C}_t)} K\left(F_n^{d_1}\left(\Delta\lambda_n^{d_1}(t)\right), \ldots, F_n^{d_E}\left(\Delta\lambda_n^{d_E}(t)\right); v_n(\mathcal{C}_t)\right)\Big|_{v_n(\mathcal{C}_t)=\left(F_{\tau^{\tilde{d}_1}}^{\tilde{d}_1}(0,t),\ldots,F_{\tau^{\tilde{d}_C}}^{\tilde{d}_C}(0,t)\right)'} \\
&= K\left(F_n^{d_1}\left(\Delta\lambda_n^{d_1}(t)\right), \ldots, F_n^{d_E}\left(\Delta\lambda_n^{d_E}(t)\right), F_{\tau^{\tilde{d}_1}}^{\tilde{d}_1}(0,t), \ldots, F_{\tau^{\tilde{d}_C}}^{\tilde{d}_C}(0,t)\right) \quad .
\end{aligned}$$

which proves the first part. For the second part the reader is referred to Appendix D.2.6. ∎

10.4.4 Remarks on the Default Trigger Level

The default trigger $U = \left(U^1, \ldots, U^D\right)'$ is not affected as the U^d, $d = 1, \ldots, D$, are supposed to be i.i.d.. In contrast, Schönbucher and Schubert [59] introduce a dependence structure for the default trigger level by means of a copula, i.e., $U \sim C$. In this framework the obvious advantage is that defaults have an immediate impact on the trigger level of the non-defaulted counterparties (via the derivative of C w.r.t. the defaulted indices) and thus on survival probabilities, see Proposition 9.17. As drawback, model intensities are not directly observable in markets and calibration is difficult.

Incorporating a copula-distributed default trigger level into our model is nevertheless possible. Algorithm A - 9.1 provides the idea of avoiding a "generalized" default level (constant on the whole time horizon) and the countdown process.

In a nutshell, let $U(t)$ be the default trigger at time $t \in [0, T]$. Initially, it holds $U(0) \sim C$. In each time step t_n, $n = 1, \ldots, N$, $U(t_n)$ is renewed and a default is observed if

$$\exp\left(-\int_{t_{n-1}}^{t_n} \lambda_s ds\right) \leq U(t_n) \quad .$$

We assume that for $\tilde{d}_c \in \mathcal{C}_{t_n}$ the trigger $U^{\tilde{d}_c}(t_n)$ is locked to its last realization $\tilde{u}^{\tilde{d}_c}(t_n)$ before default time $\tau^{\tilde{d}_c} \in (t_{n-1}, t_n]$. At the next step t_{n+1}, a conditional copula is used to generate the next default trigger level, i.e.,

$$U(t_{n+1}) \sim \frac{\partial^C}{\partial u(\mathcal{C}_{t_{n+1}})} C \text{ given } U\left(\mathcal{C}_{t_{n+1}}\right) = \tilde{u}\left(\mathcal{C}_{t_{n+1}}\right) \quad .$$

As done for the intensities, one can think of an additional change in the default trigger at each default time. A copula-distributed default level implies complex calibration. Changing to a conditional copula makes this issue even more challenging. Due to this reason, we omit a trigger with dependence structure.

Besides the disadvantage of an independent default trigger, the advantages of our model - compared to Schönbucher-Schubert framework - are that

- intensities are directly observable in market data,
- dependence structures are comprised by intensities, and
- calibration and implementation of the default trigger, see A - 9.2, are less expensive.

10.5 Model Summary

We summarize the main ideas and features of the previous sections. In an intensity-based setup we assume that a default is defined as first jump of components of a multivariate counting process N_t with stochastic intensity λ_t. This intensity is constructed as positive, càdlàg process which is driven by

- random changes Λ at fixed time instants and
- random default changes Φ at default times

which allow for dependency structures introduced by copulas K for Λ and L for Φ. In considerations of positivity maintenance and impact of defaults, the random changes Λ and Φ are time-varying. Therefore, they are revised to variables $\Lambda_n(t)$ and $\Phi_\tau(t)$ by conditional margins and conditional copulas $K(t)$ and $L(t)$.

Despite their altered design, these random variables still feature the identical copula as the original variables, see Proposition 10.20. Their dependence structures are preserved. For defaulted positions, we suppose that their intensity is locked and their random changes are fixed accordingly.

10.5.1 Representation of the Intensity Process

For $t \in [t_n, t_{n+1}),\ n = 0, \ldots, N$, the multivariate intensity is represented as

$$\lambda_t = \lambda_0 + \sum_{j=1}^{n} \Lambda_j(t) + \sum_{\tau \in \tau(\mathcal{C}_t) \leq t} \Phi_\tau(t)$$

with $\Lambda_j(t) \sim F_j(t)$, $j = 1, \ldots, n$ and $\Phi_\tau(t) \sim H_\tau(t)$. The one-dimensional case is specified as

$$\lambda_t^d = \lambda_0^d + \sum_{j=1}^{n} \Lambda_j^d(t) + \sum_{\tau \in \tau(\mathcal{C}_t) \leq t} \Phi_\tau^d(t)$$

with $\Lambda_j^d(t) = 0$ for $\tau^d \leq t_j \leq t$, $j = 1, \ldots, n$, and $\Phi_\tau^d(t) = 0$ for $\tau^d \leq \tau \leq t$ in case of a default for index d. Equivalently, our model construction holds

$$\begin{aligned} \lambda_t^d = \quad & \lambda_0^d + \sum_{j=1}^{n-1} \left(F^d\right)^{-1} \left[V_j^d(t) \left(\kappa_j^d - F^d\left(-\lambda_{t_j^-}^d\right)\right) + F^d\left(-\lambda_{t_j^-}^d\right)\right] \\ & + \sum_{\tau \in \tau(\mathcal{C}_t) \leq t} \left(H^d\right)^{-1} \left[W_\tau^d(t) \left(\mu_\tau^d - H^d\left(-\lambda_{\tau^-}^d\right)\right) + H^d\left(-\lambda_{\tau^-}^d\right)\right] \end{aligned}$$

with

- F^d, H^d, $d = 1, \ldots, D$, continuous and admissible margins,
- differentiable copulas K and L,

and

- $V_j(t) \sim K(t) \quad = \dfrac{\partial^C}{\partial v_n(\mathcal{C}_t)} K\left(v_n(\mathcal{E}_t); v_n(\mathcal{C}_t)\right)\Big|_{v_n(\mathcal{C}_t) = \tilde{v}_n(\mathcal{C}_t)}$ for $j = 1, \ldots, n,$ and $t \leq t_j$,
- $W_\tau(t) \sim L(t) \quad = \dfrac{\partial^C}{\partial w_\tau(\mathcal{C}_t)} L\left(w_\tau(\mathcal{E}_t); w_\tau(\mathcal{C}_t)\right)\Big|_{w_\tau(\mathcal{C}_t) = \tilde{w}_\tau(\mathcal{C}_t)}$ for $\tau \in (0, T]$ and $t \leq \tau$,
- $\kappa_j^d \quad = \begin{cases} 1 \\ F^d\left(-\lambda_{t_j^-}^d | \mathcal{H}_{t_j}\right) \end{cases}$ for $j = 1, \ldots, n$,
- $\mu_\tau^d \quad = \begin{cases} 1 \\ H^d\left(-\lambda_{\tau^-}^d | \mathcal{H}_\tau\right) \end{cases}$ for $\tau \in (0, T]$.

10.5.2 Survival Probabilities

For the evaluation of survival probabilities one has to consider that

- the underlying intensity process is path-dependent and
- the generation of random changes requires preceding default information.

In our model a separation of default and non-default information is not possible for the intensity process. Thus, straightforward computation of survival probabilities by conditioning on non-default-filtrations (e.g., see [40] or [59]) is not applicable. However, due to the càdlàg and path-dependent setup, this idea is transferable by an iterative procedure.

Let $i = 1, \ldots, I$ be the index for the number of jumps and denote corresponding jump times t_i, $i = 1, \ldots, I$, for λ_t, $t \in [0, T]$ [14]. Suppose that $I(t)$ defines the number of triggered changes up to time $t \in [0, T]$.

[14] It holds $I = N$ if no default has occurred on the whole period. Further, it yields $I = N + \tilde{D}$ for N random changes and $\tilde{D}$ defaults. If D defaults happen, no further random change will be triggered, thus $I \leq N + D - 1$.

10.21 Proposition (Survival Probabilities)
On the basis of the above presented intensity process in the modified information-based setup the survival probability from 0 to $t \in (0,T]$ for the single counterparty $d \in \mathcal{D}$ holds

$$S_d(0,t) := \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t\}}\right] = \mathbb{E}\left[\exp\left(-\int_0^t \lambda_s^d ds\right)\right] \quad . \tag{10.30}$$

Under filtration $\mathcal{H}_t$, the survival probability from t to T, $0 < t < T$, for counterparty $d \in \mathcal{D}$ is

$$S_d(t,T) := \mathbb{E}\left[\mathbf{1}_{\{\tau^d > T\}} | \mathcal{H}_t\right] = \mathbf{1}_{\{\tau^d > t\}} \mathbb{E}\left[\exp\left(-\int_t^T \lambda_s^d ds\right) | \mathcal{H}_t\right] \quad . \tag{10.31}$$

Proof

Let $I(t)$ and jump times t_i, $i = 1, \ldots, I(t)$, be given as defined above. As λ_t is constant in between changes, it holds

$$\begin{aligned}\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t\}}\right] &= \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t\}} | \mathcal{G}_{t_{I(t)}}\right]\right] = \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)}\}} \cdot \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t\}} | \mathcal{G}_{t_{I(t)}}\right]\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)}\}} \cdot \mathbb{P}\left[\{\tau^d > t\} | \mathcal{G}_{t_{I(t)}}\right]\right] = \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)}\}} \cdot \mathbb{P}\left[N_t^d = 0 | \mathcal{G}_{t_{I(t)}}\right]\right]\end{aligned}$$

due to conditional expectation properties and the definition of default times. Under $\mathcal{G}_{t_{I(t)}}$, the value of the process $N_{t_{I(t)}}$ is known. Additionally, we are located in the general Cox framework and can exploit the adaptivity feature of λ_t[15]. Therefore, it yields

$$\begin{aligned}\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t\}}\right] &= \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)}\}} \cdot \mathbb{P}\left[N_t^d - N_{t_{I(t)}}^d = 0 | \mathcal{G}_{t_{I(t)}}\right]\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)}\}} \cdot \mathbb{P}\left[U^d \leq \exp\left(-\int_{t_{I(t)}}^t \lambda_s^d ds\right) | \mathcal{G}_{t_{I(t)}}\right]\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)}\}} \cdot \exp\left(-\int_{t_{I(t)}}^t \lambda_s^d ds\right)\right] \quad .\end{aligned}$$

Using the property of iterated conditional expectation as above, we receive

$$\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t\}}\right] = \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)-1}\}} \cdot \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)}\}} \cdot \exp\left(-\int_{t_{I(t)}}^t \lambda_s^d ds\right) | \mathcal{G}_{t_{I(t)-1}}\right]\right] \quad .$$

Under filtration $\mathcal{G}_{t_{I(t)-1}}$ for the random variable $\mathbf{1}_{\{\tau^d > t_{I(t)}\}}$, we have

$$\begin{aligned}\mathbb{P}\left[\mathbf{1}_{\{\tau^d > t_{I(t)}\}} | \mathcal{G}_{t_{I(t)-1}}\right] &= \mathbb{P}\left[N_{t_{I(t)}} = 0 | \mathcal{G}_{t_{I(t)-1}}\right] = \mathbb{P}\left[N_{t_{I(t)}}^d - N_{t_{I(t)-1}}^d = 0 | \mathcal{G}_{t_{I(t)-1}}\right] \\ &= \mathbb{P}\left[U^d \leq \exp\left(-\int_{I(t)-1}^{I(t)} \lambda_s^d ds\right) | \mathcal{G}_{t_{I(t)-1}}\right] = \mathbb{P}\left[U^d \leq \hat{u}^d | \mathcal{G}_{t_{I(t)-1}}\right]\end{aligned}$$

with identical argumentation. The value $\lambda_{t_{I(t)}}$ is indifferent for evaluation of the right-hand side. Therefore, it is $\mathcal{G}_{t_{I(t)-1}}$-measurable. Moreover, the default trigger U^d is assumed to be independent

[15] Note that evaluation is independent of the realized value of $N_{t_{I(t)}}$.

of all other modeling variables. This implies the independence of $\mathbf{1}_{\{\tau^d > t_{I(t)}\}}$ and $\exp\left(-\int_{t_{I(t)}}^{t} \lambda_s^d ds\right)$ under filtration $\mathcal{G}_{t_{I(t)-1}}$. Hence, we receive

$$\begin{aligned}\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t\}}\right] &= \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)-1}\}} \cdot \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)}\}} \cdot \exp\left(-\int_{t_{I(t)}}^{t} \lambda_s^d ds\right) |\mathcal{G}_{t_{I(t)-1}}\right]\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)-1}\}} \cdot \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)}\}} | \mathcal{G}_{t_{I(t)-1}}\right] \cdot \mathbb{E}\left[\exp\left(-\int_{t_{I(t)}}^{t} \lambda_s^d ds\right) |\mathcal{G}_{t_{I(t)-1}}\right]\right] .\end{aligned}$$

Straight forward calculation leads to

$$\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t\}}\right] = \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)-1}\}} \cdot \exp\left(-\int_{t_{I(t)-1}}^{t_{I(t)}} \lambda_s^d ds\right) \cdot \mathbb{E}\left[\exp\left(-\int_{t_{I(t)}}^{t} \lambda_s^d ds\right) |\mathcal{G}_{t_{I(t)-1}}\right]\right] .$$

Applying rules of conditional expectation, one derives

$$\begin{aligned}&\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t\}}\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)-1}\}} \cdot \exp\left(-\int_{t_{I(t)-1}}^{t_{I(t)}} \lambda_s^d ds\right) \cdot \mathbb{E}\left[\exp\left(-\int_{t_{I(t)}}^{t} \lambda_s^d ds\right) |\mathcal{G}_{t_{I(t)-1}}\right] |\mathcal{G}_{t_{I(t)-2}}\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)-1}\}} \cdot \exp\left(-\int_{t_{I(t)-1}}^{t_{I(t)}} \lambda_s^d ds\right) \cdot \exp\left(-\int_{t_{I(t)}}^{t} \lambda_s^d ds\right) |\mathcal{G}_{t_{I(t)-1}}\right] |\mathcal{G}_{t_{I(t)-2}}\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)-1}\}} \cdot \exp\left(-\int_{t_{I(t)-1}}^{t} \lambda_s^d ds\right) |\mathcal{G}_{t_{I(t)-2}}\right]\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)-2}\}} \cdot \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)-1}\}} \cdot \exp\left(-\int_{t_{I(t)-1}}^{t} \lambda_s^d ds\right) |\mathcal{G}_{t_{I(t)-2}}\right]\right] .\end{aligned}$$

Due the independence of $\mathbf{1}_{\{\tau^d > t_{I(t)-1}\}}$ and $\exp\left(-\int_{t_{I(t)-1}}^{t} \lambda_s^d ds\right)$ under filtration $\mathcal{G}_{t_{I(t)-2}}$, it follows

$$\begin{aligned}\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t\}}\right] &= \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)-2}\}} \cdot \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)-1}\}} | \mathcal{G}_{t_{I(t)-2}}\right] \cdot \mathbb{E}\left[\exp\left(-\int_{t_{I(t)-1}}^{t} \lambda_s^d ds\right) |\mathcal{G}_{t_{I(t)-2}}\right]\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_{I(t)-2}\}} \cdot \exp\left(-\int_{t_{I(t)-2}}^{t_{I(t)-1}} \lambda_s^d ds\right) \cdot \mathbb{E}\left[\exp\left(-\int_{t_{I(t)-1}}^{t} \lambda_s^d ds\right) |\mathcal{G}_{t_{I(t)-2}}\right]\right] .\end{aligned}$$

Iteration for all jump times t_i, $i = 1, \ldots, I(t)$, yields

$$\begin{aligned}\mathbb{E}\left[\mathbf{1}_{\{\tau^d > t\}}\right] &= \mathbb{E}\left[\mathbf{1}_{\{\tau^d > t_1\}} \cdot \exp\left(-\int_{t_1}^{t_2} \lambda_s^d ds\right) \cdot \mathbb{E}\left[\exp\left(-\int_{t_2}^{t} \lambda_s^d ds\right) |\mathcal{G}_{t_1}\right]\right] \\ &= \mathbb{E}\left[\exp\left(-\int_{0}^{t_1} \lambda_s^d ds\right) \cdot \mathbb{E}\left[\exp\left(-\int_{t_1}^{t} \lambda_s^d ds\right) |\mathcal{G}_0\right]\right] = \mathbb{E}\left[\exp\left(-\int_{0}^{t} \lambda_s^d ds\right)\right] .\end{aligned}$$

For the second equation let $t \in (0,T)$. W.l.o.g.[16] denote t_j, $j \in \{1,\dots,I(T)\}$, as first random change on the interval $(t,T]$ with $0 = t_0 \leq t_{j-1} \leq t < t_j \leq t_{I(T)} \leq T$. Identical computations as above lead to

$$\begin{aligned}
\mathbb{E}\left[\mathbf{1}_{\{\tau^d>T\}}|\mathcal{H}_t\right] &= \mathbb{E}\left[\mathbf{1}_{\{\tau^d>t\}} \cdot \exp\left(-\int_t^{t_j} \lambda_s^d ds\right) \cdot \mathbb{E}\left[\exp\left(-\int_{t_j}^{T} \lambda_s^d ds\right)|\mathcal{G}_{t_j}\right]|\mathcal{H}_t\right] \\
&= \mathbb{E}\left[\mathbf{1}_{\{\tau^d>t\}} \cdot \mathbb{E}\left[\exp\left(-\int_t^{T} \lambda_s^d ds\right)|\mathcal{G}_{t_j}\right]|\mathcal{H}_t\right] \\
&= \mathbf{1}_{\{\tau^d>t\}} \cdot \mathbb{E}\left[\exp\left(-\int_t^{T} \lambda_s^d ds\right)|\mathcal{H}_t\right] \quad .
\end{aligned}$$

The last equation holds as $\mathbf{1}_{\{\tau^d>t\}}$ is $\mathcal{H}_t$-measurable and $\mathcal{H}_t \subset \mathcal{G}_{t_j}$ (iterated conditioning). ∎

Conditioning on different filtrations leads to an iterative evaluation which likewise reflects the path-dependence. Moreover, the proof exploits

- the respective measurability of random change variables in combination with the càdlàg construction of the intensity and
- the assumption that the default trigger level is independent of all other modeling variables.

10.22 Remark

1. As a specification of the generalized intensity-based setup, our model derives equivalent survival probabilities, i.e.,

$$S_d(0,t) = \mathbb{E}\left[\exp\left(-\int_0^t \lambda_s^d ds\right)\right] = P_d(0,t) \quad ,$$

   compare to Lemma 9.12. Insofar, the preceding result shows that our approach is in line with the generalized framework.

2. Survival probabilities $S_d(t,T)$ depend on filtration $\mathcal{H}_t$ by construction. For knowing $\mathbf{1}_{\{\tau^d>t\}}$ and λ_t in particular, all default- and non-default evolution is required. A separation of these information flows is not intended in our model. In the generalized framework, computation of survival probabilities $P_d(t,T)$ relies on the separation of filtrations. Therefore - under the assumption that position $d \in \mathcal{E}_t$ has not defaulted up to time t - results

$$\begin{aligned}
S_d(t,T) &= \mathbb{E}\left[\exp\left(-\int_t^T \lambda_s^d ds\right) \mid \mathcal{H}_t\right] \quad \text{and} \\
P_d(t,T) &= \mathbb{E}\left[\exp\left(-\int_t^T \lambda_s^d ds\right) \mid \mathcal{I}_t\right]
\end{aligned}$$

   naturally differ w.r.t. the underlying information. Here, $\mathcal{I}_t$ denotes the usual background filtration of Chapter 9.

3. Nevertheless, computations for $S_d(t,T)$ and $P_d(t,T)$ are related and reveal similar formulas, i.e., taking the expectation over all possible paths under a specified filtration. In this respect, our model and the generalized approach are consistent. Moreover, approaches are identical if evaluation is restricted to each iteration under the given information.

[16] For $t = T$, no further evaluation is needed. For $t_{I(T)} \leq t < T$, all positions are defaulted before t. Otherwise, a last random change would have been triggered at T. For both cases, survival probabilities are zero. For $t = 0$, we deal with the first part.

10.5.3 Algorithm

Taking over the direct default time simulation of A - 9.2, the above presented model can be implemented by the following

A - 10.1 Algorithm (Copula- and Default-Dependent Intensity Model)

Given

- time horizon with time grid t_n, $n = 0, \ldots, N$, $t_0 = 0$, $t_n = T$,
- marginal distributions F^d and H^d, $d = 1, \ldots, D$
- copula functions K and L

Step 1

- for $d = 1, \ldots, D$:
  - calibrate $\Lambda^d \sim F^d$ and $\Phi^d \sim H^d$ to market data on the given time grid
  - compute $(F^d)^{-1}$ and $(H^d)^{-1}$
  - set $\lambda_0^d > 0$ according to market data, κ_0^d and μ_0^d, $\gamma_0 = 1$
  - draw $u^d \sim U(0,1)$
- calibrate K and L to market data on the given time grid
- set $K_0 = K$ and $L_0 = L$
- set $\mathcal{E}_0 = \{1, \ldots, D\}$, $\mathcal{C}_0 = \emptyset$

Step 2 - for $n = 1, \ldots, N$:

- set $t_{min} = t_{n-1}$
- while $t_{min} \leq t_n$ do:
  - for $d \in \mathcal{E}_{n-1}$: compute $t_m^d = \min\left\{t > t_{n-1} | \gamma_{n-1}^d \cdot \exp(-\lambda_{n-1}^d (t - t_{n-1}) \leq u^d\right\}$
  - compute $(t_{min}, d_{min}) = \left(\min\limits_{d \in \mathcal{E}_{n-1}} \{t_m(\mathcal{E}_{n-1})\}, \{d^* \in \mathcal{E}_{n-1} | t_m^{d^*} = \min\limits_{d \in \mathcal{E}_{n-1}} \{t_m(\mathcal{E}_{n-1})\} \right)$
  - compare $t_{min} > t_n \quad \Rightarrow \quad$ no default on $(t_{n-1}, t_n]$!
    - set $\mathcal{E}_n = \mathcal{E}_{n-1}$, $\mathcal{C}_n = \mathcal{C}_{n-1}$ and $K_n = K_{n-1}$
    - draw $v_n \sim K_n$ given $V(\mathcal{C}_n) = \tilde{v}(\mathcal{C}_n)$
    - for $d \in \mathcal{E}_n$:
      - compute $\gamma_n^d = \gamma_{n-1}^d \cdot \exp(-\lambda_{n-1}^d (t_n - t_{n-1}))$
      - compute $\Delta\lambda_n^d = (F^d)^{-1}\left[v_n^d \left(\kappa_n^d - F^d\left(-\lambda_{n-1}^d\right)\right) + F^d\left(-\lambda_{n-1}^d\right)\right]$
      - compute $\lambda_n^d = \lambda_{n-1}^d + \Delta\lambda_n^d$
    - for $d \in \mathcal{C}_n$:
      - set $\Delta\lambda_n^d = 0$, $\lambda_n^d = \lambda_{n-1}^d$ and $\gamma_n^d = \gamma_{n-1}^d$
  - compare $t_{min} \leq t_n \quad \Rightarrow \quad$ default for position d_{min} at time $t_{min} \in (t_{n-1}, t_n]$!
    - compute

$$\tilde{v}^{d_{min}} = \frac{F^{d_{min}}(0) - F^{d_{min}}\left(-\lambda_{t_{n-1}}^{d_{min}}\right)}{\kappa_{n-1}^{d_{min}} - F^{d_{min}}\left(-\lambda_{t_{n-1}}^{d_{min}}\right)}$$

- compute
$$\tilde{w}^{d_{min}} = \frac{H^{d_{min}}(0) - H^{d_{min}}\left(-\lambda_{t_{n-1}}^{d_{min}}\right)}{\mu_{n-1}^{d_{min}} - H^{d_{min}}\left(-\lambda_{t_{n-1}}^{d_{min}}\right)}$$
- set $\tilde{v}(\mathcal{C}_{n-1}) = (\tilde{v}(\mathcal{C}_{n-1}), \tilde{v}^{d_{min}})$[17]
- set $\tilde{w}(\mathcal{C}_{n-1}) = (\tilde{w}(\mathcal{C}_{n-1}), \tilde{w}^{d_{min}})$[18]
- set $\mathcal{E}_{n-1} = \mathcal{E}_{n-1} \setminus \{d_{min}\}$, $\mathcal{C}_{n-1} = \mathcal{C}_{n-1} \cup \{d_{min}\}$, $t_{n-1} = t_{min}$
- compute $K_{n-1} = \frac{\partial}{\partial v^{d_{min}}} K_{n-1}$ given $V(\mathcal{C}_{n-1}) = \tilde{v}(\mathcal{C}_{n-1})$
- compute $L_{n-1} = \frac{\partial}{\partial w^{d_{min}}} L_{n-1}$ given $W(\mathcal{C}_{n-1}) = \tilde{w}(\mathcal{C}_{n-1})$
- draw $w_{n-1} \sim L_{n-1}$
- for $d \in \mathcal{E}_{n-1}$[19]:
  - compute $\gamma_{n-1}^d = \gamma_{n-1}^d \cdot \exp(-\lambda_{n-1}^d (t_{min} - t_{n-1}))$
  - compute
$$\Delta\phi_{n-1}^d = (H^d)^{-1}\left[w_{n-1}^d \left(\mu_{n-1}^d - H^d\left(-\lambda_{n-1}^d\right)\right) + H^d\left(-\lambda_{n-1}^d\right)\right]$$
  - compute $\lambda_{n-1}^d = \lambda_{n-1}^d + \Delta\phi_{n-1}^d$
- loop

Annotation of Algorithm A - 10.1

Initially, time horizon and time grid are fixed, relevant distributions for random changes (margins F^d, $d = 1, \ldots, D$, and copula K) and random default changes (margins H^d, $d = 1, \ldots, D$, and copula L) are specified. Their selection depends on various aspects (application issues, complexity of distributions, preferences for particular cdfs, goodness-of-fit performances etc.).

Subsequently, marginal distributions and copulas are calibrated to market data on the given time grid[20]. Margins are inverted and the type of truncation for the conditional cdfs (κ_0^d and μ_0^d, $d = 1, \ldots, D$) are defined. For each index $d \in \{1, \ldots, D\}$, the default trigger level is drawn, values and specifications are fixed for

- intensity (λ_0^d),
- copulas (K_0 and L_0), and
- index sets of defaulted and non-defaulted positions ($\mathcal{C}_0$ and $\mathcal{E}_0$).

Regarding the main part of the algorithm, in each time step t_n of the given time grid an auxiliary variable t_{min} is set to the previous time step t_{n-1}. Evaluations on the interval $(t_{n-1}, t_n]$ are controlled by t_{min}. While $t_{min} \leq t_n$, we perform the following loop:

For all non-defaulted positions $d \in \mathcal{E}_{n-1}$ at t_{n-1}, we determine the minimum t_m^d on $(t_{n-1}, t_n]$ such that the default process falls below the default trigger level as
$$t_m^d = \min\left\{t_{n-1} < t \leq t_n \middle| \exp\left(-\int_0^t \lambda ds\right) \leq u^d\right\} \quad .$$

[17] Not in permuted order.
[18] Not in permuted order.
[19] For $d \in \mathcal{C}_{n-1}$, λ_{n-1}^d and γ_{n-1}^d are already given.
[20] Adequacy of calibration should be verified by goodness-of-fit tests for all distributions.

Knowing that position $d \in \mathcal{E}_{n-1}$ survived up to time t_{n-1} and defining the default process by

$$\gamma_{n-1}^d := \exp\left(-\int_0^{t_{n-1}} \lambda^d ds\right) \quad ,$$

it can alternatively be checked if

$$\gamma_{n-1}^d \cdot \exp\left(-\lambda_{n-1}^d \cdot (t - t_{n-1})\right) \leq u^d \quad .$$

due to the exponential function properties. If no minimum - and therefore no default for index d on $(t_{n-1}, t_n]$ - is found, we set $t_m^d = \infty$. Next, we search for the minimum default time t_{min} and its corresponding index d_{min} for all non-defaulted positions at time t_{n-1}. In case that no default has happened (i.e., $t_{min} = \infty$ and $d_{min} = \emptyset$),

- sets for defaulted and non-defaulted counterparties and
- copulas

are identical to the previous time step (i.e., set $\mathcal{E}_n = \mathcal{E}_{n-1}$, $\mathcal{C}_n = \mathcal{C}_{n-1}$ and $K_n = K_{n-1}$). Then, the random change $\Delta\lambda_n^d$, $d \in \mathcal{E}_n$, is constructed via $v_n \sim K_n$ and the default process γ_n^d as well as intensity process λ_n^d are updated. The inner loop is left by increasing n to $n+1$ and due to $t_{min} = \infty$.

Otherwise, if a minimum default time exists (i.e., $t_{min} \leq t_n$ with $d_{min} = d^*$), the relevant realizations $\tilde{v}^{d_{min}}$ and $\tilde{w}^{d_{min}}$ are determined,

- vectors $\tilde{v}(\mathcal{C}_{n-1})$ and $\tilde{w}(\mathcal{C}_{n-1})$
- sets $\mathcal{E}_{n-1}$ and $\mathcal{C}_{n-1}$, time $t_{n-1} = t_{min}$, and
- copulas K_{n-1} and L_{n-1}

are refreshed[21]. Even though not used in this part, the relevant information for the random change (i.e., $\tilde{v}(\mathcal{C}_{n-1})$ and K_{n-1}) must be renewed to be able to use it in the non-default part of the algorithm. Moreover, the random default change $\Delta\phi_n^d$, $d \in \mathcal{E}_n$, is constructed via $w_n \sim L_n$ and the default process γ_n^d as well as intensity process λ_n^d are updated. Another loop is performed by initially comparing new minimum default times on $(t_{n-1} = t_{min}, t_n]$. An alternative algorithm leaving out the random default change is found in Appendix B.6.

10.23 Definition (Numerics for Instantly Occurring Defaults)
In the most unlikely numerical case that $t_{min} = t_n$, one has to define whether the random default change or the random change is triggered. In the case that two (or more) defaults happen at the same time instant t_{min}, we define that both (or all) defaults are reflected in the random default change, i.e., copula derivatives are taken w.r.t. to both (or all) defaulted indices.

10.24 Remark
Besides market calibration, time-dependent derivation of conditional copulas w.r.t. defaulted variables is the most complex and time-consuming issue of A - 10.1.

10.5.4 Interpretation of the Model

In contrast to the general intensity-based setup, our model exploits information of realized intensity paths which contain knowledge of detected defaults. A separation of non-defaulted and defaulted filtrations is therefore not possible on the **whole** period[22]. We believe that this is a realistic market feature as defaults of counterparties do influence default behavior of non-defaulted positions. The mathematical implementation of this idea is performed by locking variables for defaulted positions.

[21] We omit additional indexing by overwriting the subscription $n-1$. Note that now $t_{n-1} = t_{min} \neq$ time grid t_{n-1}.
[22] As modification of the information-based approach, our model yields equivalent results concerning survival probabilities, see Proposition 10.21.

Their further, one-dimensional evaluation is not considered anymore. This means that defaulted positions are not supposed to recover on the remaining time horizon.

Locking triggers a direct and an indirect impact. An immediate implication is modeled by the random default change variable. Economically seen, this variable causes a direct change for default probabilities of non-defaulted counterparties symbolizing a sort of market shock. An indirect effect is received by the differentiation of copulas which can be regarded as persistent market reaction. The conditional distribution leads to default contagion towards lower or higher default probabilities dependent on

- the chosen copula (i.e., the dependency structure),
- the respective dependence between defaulted and non-defaulted indices, and
- the level (i.e., the last value) of the defaulted intensities.

Due to the positivity premise (i.e., generation of admissible changes in sense of Definition 10.3), intensities are restricted by conditional margins. Mathematically, large negative jumps of λ_t towards non-positive values are not allowed. If both tails are cut ($\mu_\tau^d = H^d$, $\kappa_n^d = F^d$), also large positive changes of λ_t are prohibited. In either case probabilities for admissible points of the distributions are amplified. Thus, distributions H_τ^d and F_n^d are versions of H^d and F^d with **rescaled** probabilities.

Economically, leaving out considerations on the default trigger, **survival** probabilities generally depend on the level of the intensity. They are either decreased by positive changes or increased by negative changes of the intensity process. Vice versa, **default** probabilities are increased / decreased by positive / negative changes. In particular, cutting the lower tail of the margin thus

- excludes high decreases of the intensity and
- amplifies decreases of the survival probability.

Truncation of the upper tail reduces the possibility high increases of the intensity. Hence, conditional margins imply a reduction of volatility for survival respectively default probabilities. Of course, the resulting impacts depend on the applied marginal distribution[23].

By construction the intensity process

- is càdlàg and therefore constant in between changes and
- has a finite number of jumps $I < N + D$.

Therefore, evaluation of survival probabilities is feasible on iterative intervals under previous realizations, see proof of Proposition 10.21. In other words, the generalized intensity-based setup of Chapter 9 is piecewise applicable. However, the proof evidences that N_t is not a Cox process in sense of Definition 9.7. Default and market information are not separable on the whole time horizon. It is rather a counting process with two stochastic components

- the jump part N_t and
- the intensity λ_t

which are mutually dependent.

11 Model Implementation

Generally, reduced-form models are used for pricing of credit risky securities or derivatives. Likewise, but more infrequently in practice, they can be applied to predict default probabilities in

[23] The truncated tail probability might be low such that the restriction to the conditional distribution has almost no effect. In this case, the shifts of probabilities are minimal.

frameworks of credit loan portfolios or - uncommonly - rating purposes. Being part of the intensity-based setup, our model is applicable to both cases.

We implement the mentioned portfolio perspective. As in the equity risk part (see Chapters 4 and 5), the model is calibrated for different specifications. Goodness-of-fit tests are applied to assess fitting quality. Moreover, simulation results are discussed with respect to convergence behavior, backtesting and benchmarking performance. As conclusion, numerical results are summarized.

11.1 Calibration

We implement our model in a real-world framework as prediction of single default probabilities in a portfolio setup. Here, we accept several simplifications which are questionable, but commonly assumed in practice. These constraints though provide an uncomplicated interpretation of model performance. Additionally, we are not interested in an assessment of a generalized portfolio loss as only DPs and **not** LGDs and EaDs are estimated.

11.1 Assumption (DP and Spread-Intensity-Relation)

1. The intensity λ_t is connected to the default probabilities DP via the relation

$$DP_T = 1 - \exp\left(-\int_0^T \lambda_s ds\right) \quad .$$

   In practice, time horizon for DP prediction is commonly one year for portfolio and rating models, i.e., $T = 1$ and

$$DP = DP_1 = 1 - \exp\left(-\int_0^1 \lambda_s ds\right) \quad .$$

2. The intensity can be extracted from spread quotes s by

$$\lambda \approx \frac{s}{LGD} \quad ,$$

   e.g., see [9], Subsection 21.3.6[1].

3. We suppose that $LGD = 1$ implying $\lambda \approx s$.

11.1.1 Market Implied Default Probabilities

Spreads are inferable from CDS, bonds or other credit risk-related quotations. The following approach is inspired by an intuitive example for a CDS calibration with constant intensities in [9], Section 22.3.

1. For the fitting of margins ($d = 1, \ldots, D$), we assume that yearly CDS spreads are constant. From these spreads $s^d(0, 1y),\ s^d(0, 2y),\ s^d(0, 3y),\ \ldots,\ s^d(0, Ny)$ relevant spreads

$$s^d(0, 1y),\ s^d(1y, 2y),\ s^d(2y, 3, y),\ \ldots,\ s^d((N-1)y, Ny)$$

   are received by bootstrapping.

2. The approximation $\lambda \approx \frac{s}{LGD}$ can be used to derive constant intensities

$$\lambda^d(0, 1y),\ \lambda^d(1y, 2y),\ \lambda^d(2y, 3, y),\ \ldots,\ \lambda^d((N-1)y, Ny) \quad ,$$

   for instance. The initial value λ_0 is set to $\lambda^d(0, 1y)$. This means that random changes in our model are triggered yearly, i.e., $t_1 = 1y,\ t_2 = 2y,\ \ldots,\ t_N = Ny$.

[1] Besides default risk, this approximation obviously includes risks of recovery. Also, it partly covers liquidity and market risk, as these risk factors are assumed to be incorporated in credit spreads as well.

3. Changes $\lambda^d(0,1y)$ to $\lambda^d(1y,2y)$, ..., $\lambda^d((N-2)y,(N-1)y)$ to $\lambda^d((N-1)y,Ny)$ are then suitable to calibrate the generalized marginal distributions F^d.

4. Afterward, parametrizing the copula K is straightforward.

Naturally, our model is enlargeable by an additional time dimension for margins. If sufficient market quotes are available, changes from t_1 to t_2, ..., t_{N-1} to t_N can be used to extend conditional distributions F_n^d to $F_{t_n,n}^d$, $n = 1,\dots,N$. The first index specifies the margin of the relevant year t_n. The second consistently indicates the marginal truncation at t_n. Accordingly, copula function K changes to time-indexed copulas K_{t_n}, $n = 1,\dots,N$.

Due to scarce information, calibration of the additional default jump components (i.e., margins H^d, $d = 1,\dots,D$, and copula L) is the most challenging task. If this impact is not deducible from market quotes, fitting might be performed on historical data or estimated by a global value assessed by experts. Alternatively, it can be left out as regular random changes partly include related default effects.

11.2 Remark
For pricing issues, calibration is performed on current market quotes. Thereby, certain requirements (e.g., validity of pricing formulas or martingale properties) or special characteristics (e.g., non-unique risk-neutral probability measures) have to be respected. Due these aspects, implementation for pricing is complex.

Due to the real-world setup, historical spread quotes can be used for calibration. Moreover, we restrict to one-year spreads to reduce complexity. Using the idea of the previous subsection, we slightly extend the time perspective for fitting:

- Historical spreads $s(0,1y,t_i)$ are used to generate intensities $\lambda(0,1y,t_i)$ for $i = 1,\dots,\hat{I}$ as stated above.
- Changes $\Delta\lambda(0,1y,t_i)$ are derived by $\lambda(0,1y,t_{i+1})$ and $\lambda(0,1y,t_i)$.
- Margins and copula are calibrated to $\Delta\lambda(0,1y,t_i)$, $i = 1,\dots,\hat{I}-1$.
- Now, we assume that random change is not triggered yearly but in dependence of time steps t_i. This implies a different time grid in comparison to the preceding approach.
- For a DP prediction of one year (i.e., 250 days[2]), the number of random changes N is computed as 250 days divided by the difference of t_{i+1} and t_i counted in days. As example, N is 250 for daily quotations, N is 50 for weekly[3] changes.

11.1.2 Model Specifications

The following specifications and assumptions are implemented for our model.

- **marginal distributions F^d and H^d:**

  Random changes $F_j^d(t)$, $d = 1,\dots,D$ are specified by

  - empirical distribution functions (compare Definition 3.2, Chapter 3)[4],
  - α-stable distributions (compare Appendix C), and
  - truncations κ_j^d defined as $F^d\left(-\lambda_{t_j^-}^d|\mathcal{H}_{t_j}\right)$.

  Default jumps are omitted, i.e., distributions $H_\tau^d(t)$, $d = 1,\dots,D$ are not employed.

[2] For quotation of credit spreads, one normally assumes a year fraction of $\Delta t = \frac{1}{200}$ or $\Delta t = \frac{1}{250}$. In contrast to the first part, we use 250 trading days here.
[3] 1 week = 5 days.
[4] The continuity of distributions is an analytical requirement. It is released here, as distributions are discretized for numerical implementation in any case.

- **copula functions K and L**:

  The dependence structure of random changes is described by

  - the Gaussian copula and
  - the t-copula

  for $K(t)$. As default jumps are skipped, copula $L(t)$ is left out.

- **default trigger U^d**:

  The default trigger variables U^d, $d = 1, \ldots, D$, are assumed to be independent.

11.1.3 Calibration and Simulation

Besides model settings, different calibrations base on the computation of changes and the underlying time-dimension. They are abbreviated as

- *abs* for a calibration to **absolute** changes,
- *rel* for a calibration to **relative** changes, and
- *log* for a calibration to **logarithmic** changes

and denoted by

- $K = 1$ for a calibration to **daily** changes and
- $K = 5$ for a calibration to **weekly** changes.

As calibration methods, the canonical maximum likelihood approach (CML, Subsection 3.1.3) and inference functions for margins (IFM) are applied. Gaussian and t-copula with empirical margins are fitted by CML (Algorithms A - 3.1 and A - 3.2).

Copulas with α-stable margins are calibrated by IFM. This simplifies fitting due to the separated maximization of log-likelihood functions (marginal cdfs and copula, see Subsection 3.1.2). High-dimensional optimization of the exact MLE (Subsection 3.1.1) induces barriers such as increasing computational time or problems of local maximizer and is therefore avoided. For implementation, a log-likelihood estimator for α-stable distributions from Mathworks is borrowed[5]. Copula calibration is then accordingly performed as for the CML approach. Observations are transformed by α-stable distributions with fitted parameters and are further used for copula parametrization.

Congruent to Chapter 4, the goodness-of-fit is tested by Algorithms A - 3.4 and A - 3.5 for all specifications.

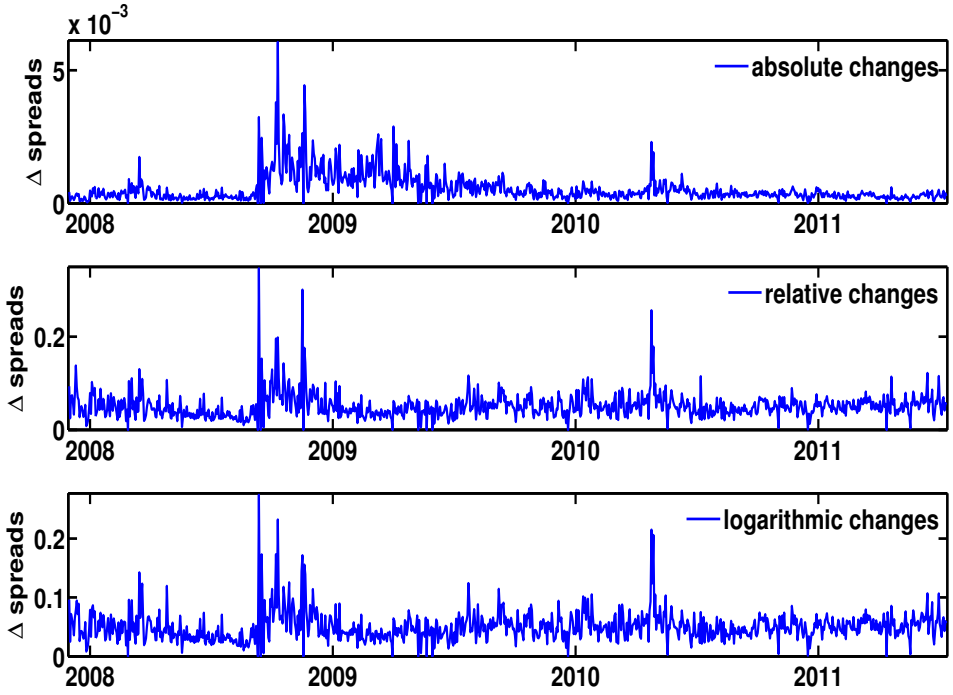
For simulation, Algorithm A - B.1 is implemented. Here, general explanatory comments on the model (Chapter 10) and statements on sampling conditional copulas (Section 2.3 and especially Subsection 2.3.3) are considered and applied.

11.1.4 Data Basis

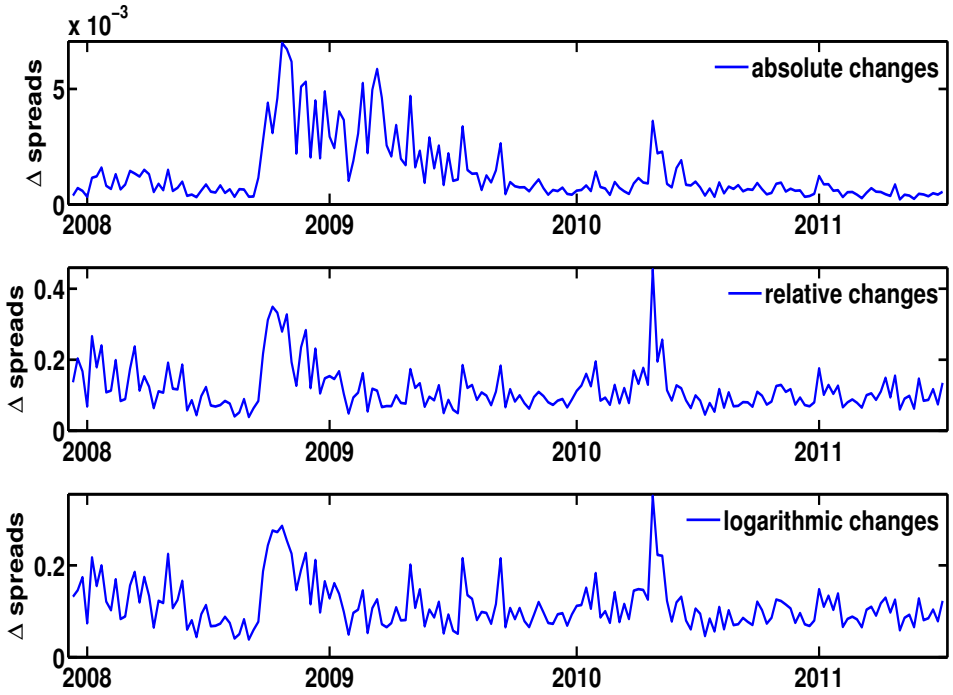
Interested in one-year default probabilities or more precisely the spread-connected intensities, we use one-year credit spreads $s^d(0, 1y)$ for $D = 24$ counterparties listed in Appendix B.4.

Moreover, we have $\hat{I} = 927$ daily quotations $s^d(0, 1y, t_i)$, $i = 1, \ldots, \hat{I}$ and $d = 1, \ldots, D$, from November 28, 2007 till July 27, 2011.

[5] Available on mathworks.com, package Master-stbl, 2012, author Mark Veillette.



F - 11.1 Figure: mean of absolute spread changes - $D = 24$ and $K = 1$



F - 11.2 Figure: mean of absolute spread changes - $D = 24$ and $K = 5$

Absolute, *relative* and *logarithmic* changes are computed for each counterparty $d \in \mathcal{D}$ and $i = 1, \ldots, I_{daily} = 926$ as

$$\begin{aligned}
abs: \ \Delta\lambda^d(t_i) &= s^d(0,1y,t_{i+1}) - s^d(0,1y,t_i) \ \approx \ \lambda^d(0,1y,t_{i+1}) - \lambda^d(0,1y,t_i) \quad , \\
rel: \ \Delta\lambda^d(t_i) &= \frac{s^d(0,1y,t_{i+1}) - s^d(0,1y,t_i)}{s^d(0,1y,t_i)} \approx \frac{\lambda^d(0,1y,t_{i+1}) - \lambda^d(0,1y,t_i)}{\lambda^d(0,1y,t_i)} \quad , \\
log: \ \Delta\lambda^d(t_i) &= \log\left(\frac{s^d(0,1y,t_{i+1})}{s^d(0,1y,t_i)}\right) \qquad \approx \log\left(\frac{\lambda^d(0,1y,t_{i+1})}{\lambda^d(0,1y,t_i)}\right) \quad .
\end{aligned}$$

Moreover, we derive $I_{weekly} = 184$ weekly changes from the above given daily spread data.

11.3 Remark
In symbolic stochastic differential terms, absolute changes assume an evolution of "$d\lambda = d\Lambda$", whereas logarithmic changes suppose a development of "$d\lambda = \lambda d\Lambda$". Relative changes are interpretable as first order approximation of logarithmic changes.

For the above given history and portfolio of $D = 24$ counterparties, the means of the absolute daily ($K = 1$) and weekly ($K = 5$) spread changes for each t_i, $i = 1, \ldots, I_{daily}$ (I_{weekly}, respectively), are computed as

$$\begin{aligned}
\Delta s_{abs}(t_i) &= \frac{1}{24}\sum_{d=1}^{24}\left|s^d(0,1y,t_{i+1}) - s^d(0,1y,t_i)\right| \ , \\
\Delta s_{rel}(t_i) &= \frac{1}{24}\sum_{d=1}^{24}\left|\frac{s^d(0,1y,t_{i+1}) - s^d(0,1y,t_i)}{s^d(0,1y,t_i)}\right| \ , \quad \text{and} \\
\Delta s_{log}(t_i) &= \frac{1}{24}\sum_{d=1}^{24}\left|\log\left(\frac{s^d(0,1y,t_{i+1})}{s^d(0,1y,t_i)}\right)\right|
\end{aligned}$$

and depicted in F - 11.1 and F - 11.2. These figures provide an indication of the volatility embedded in the portfolio. For both changes, high peaks are observed in times during the financial crisis (end of 2008 - end of 2009) and at the beginning of the EURO crisis (beginning of 2010).

11.1.5 Time Horizon for Calibration

For daily changes, the model is fitted on a horizon of **one year**, i.e., 250 data points. Further, another 250 observations are left out for a comparison of simulation results. This leads to 426 different calibrations[6].

Calibration horizon for weekly changes is extended to **two years** receiving a basis of 100 observations. Leaving out a simulation period of one year, we obtain 34 different calibrations[7].

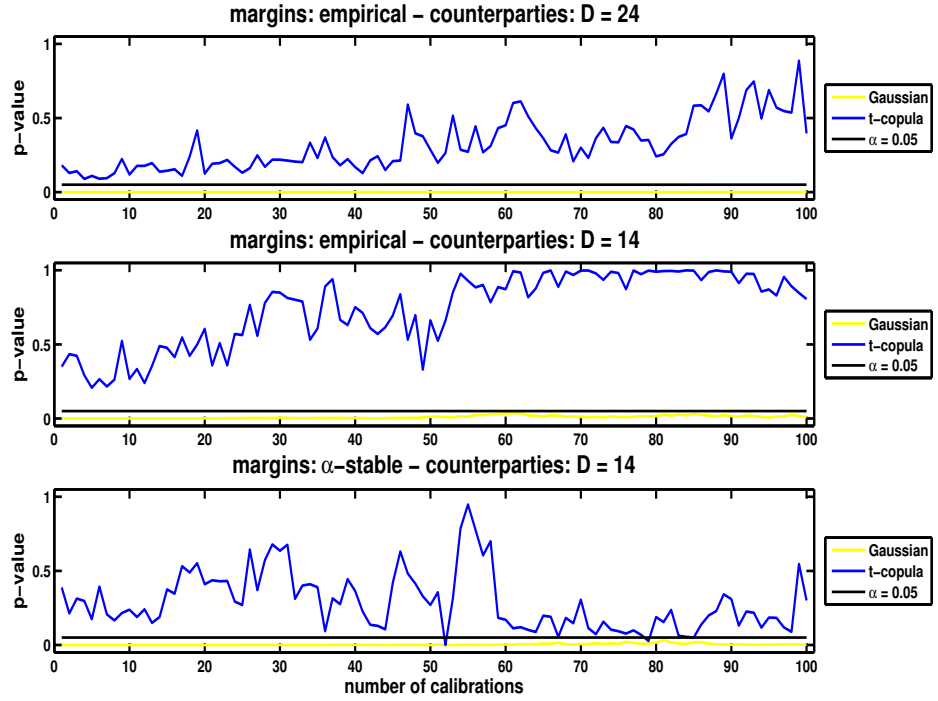
11.2 Numerical Results

Given presented specifications and described data basis, we calibrate our model. As a first result, we observe that fitting α-stable margins is computationally very intensive and exhibits numerical instabilities. For this reason, we restrict the analysis of models with α-stable margins to $D = 14$ counterparties and the last 100 available calibrations[8]. In addition, simulation is skipped. In order to have an comparison of results, models with empirical margins are adjusted to $D = 14$ besides $D = 24$ positions. We apply goodness-of-fit tests of Section 3.3 and compare results for the different specifications.

[6] Assuming a *rolling* or *moving* calibration procedure as described in Chapter 5. Calibration no. 1 corresponds to spread changes 1 - 250, no. 2 to 2 - 251, ..., no. 426 to 426 - 676. Last 250 days are reserved for backtesting simulation results.

[7] Calibration no. 1 corresponds to spread changes 1 - 100, no. 2 to 2 - 101, ..., no. 34 to 34 - 133. Last 50 weeks are reserved for backtesting simulation results.

[8] I.e., calibrations 327 to 426.



F - 11.3 Figure: goodness-of-fit test - $K = 1$ and *abs*

11.2.1 Goodness-of-Fit Tests for Daily Changes

Figures F - 11.3, F - 11.4, and F - 11.5 show that t-copula models have better p-values than Gaussian ones. They provide a significantly good fit, whereas hypotheses for the Gaussian copula are mainly rejected (p-value below the confidence level of 0.05).

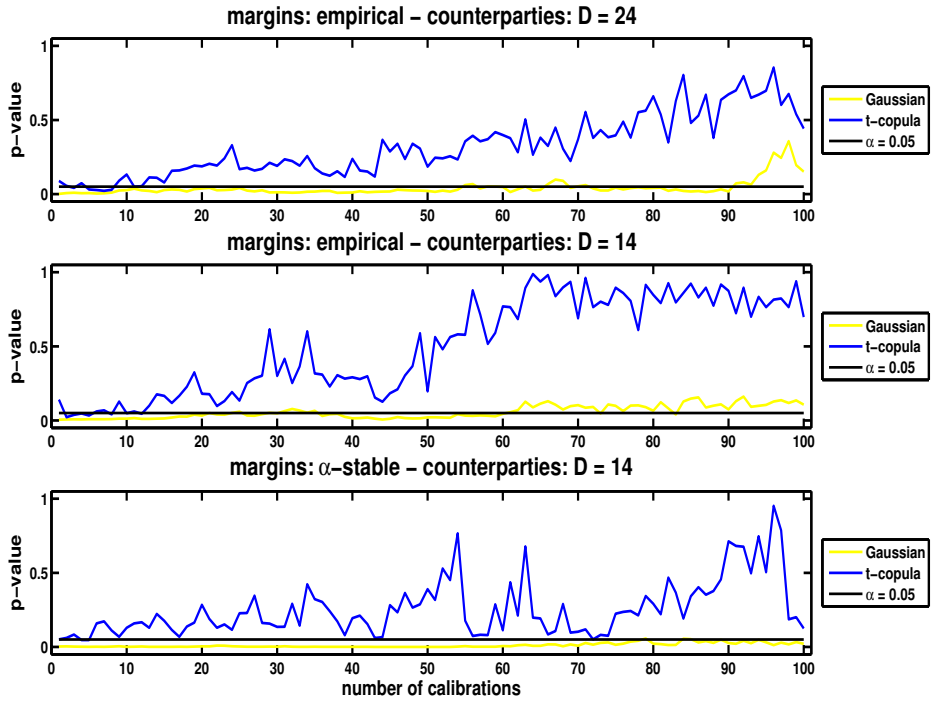
This results is independent of the portfolio size ($D = 24$ vs. $D = 14$ for empirical margins) and the chosen margins (empirical vs. α-stable margins for $D = 14$).

Extending the number of calibrations for models with empirical margins to the whole period from 100 to 426[9], the aforementioned statements still hold. However, in times of volatile markets the goodness-of-fit decreases. Often, hypotheses are rejected, compare Figures F - B.1, F - B.2 as well as F - B.3 in Appendix B.5.

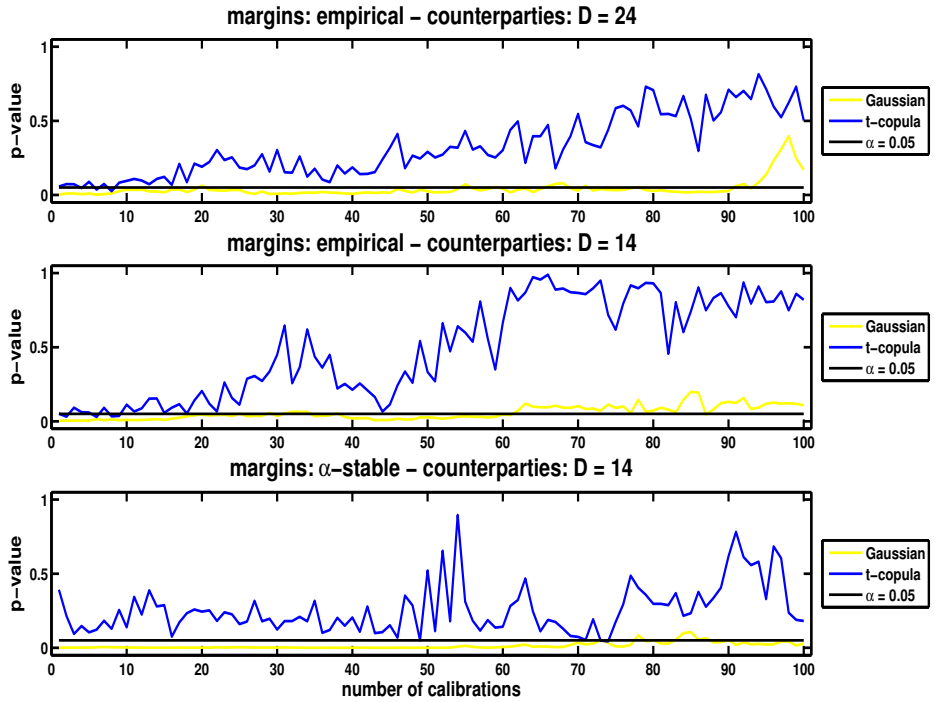
Comparing the goodness-of-fit for the t-copula in detail, several aspects are inferred from F - 11.6:

1. Calibration basis (*abs*, *rel*, *log*) does not to have large impacts on fitting results. Movements of p-values are alike.
2. Notably, smaller portfolios exhibit better fits than larger ones (p-values compared for empirical margins with $D = 24$ vs. $D = 14$ counterparties).
3. Results for models with empirical margins are better than for those with α-stable margins.

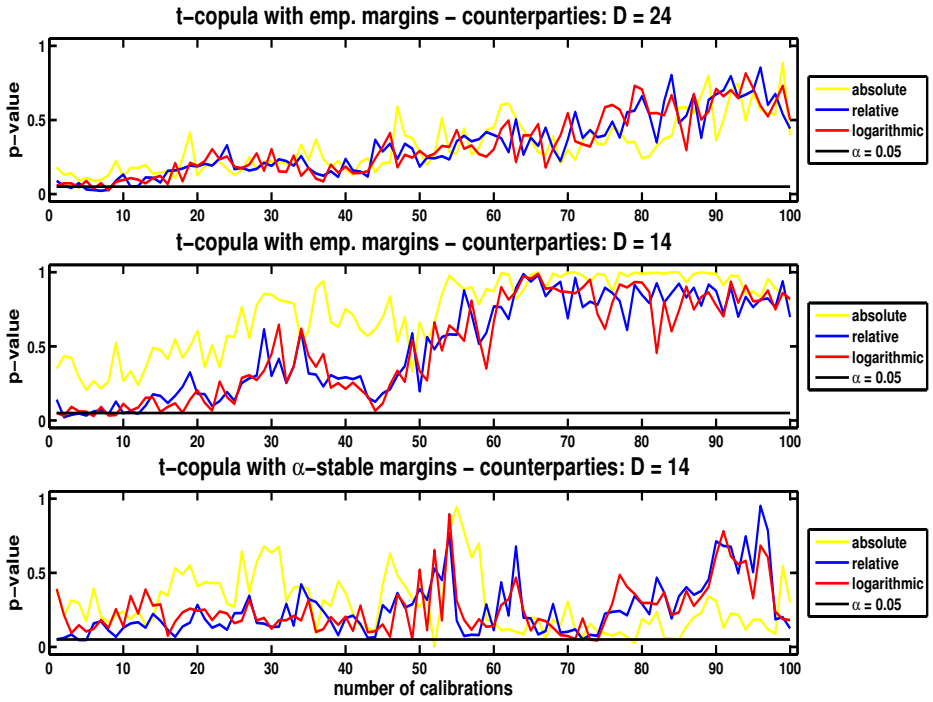
[9] I.e., including calibrations 1 to 326.



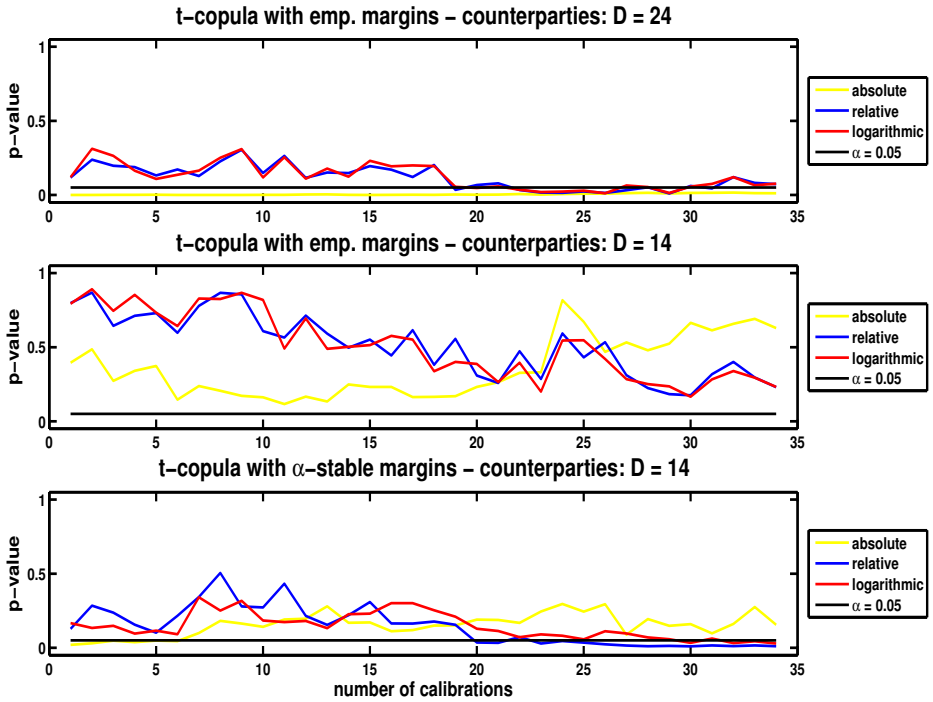
F - 11.4 Figure: goodness-of-fit test - $K = 1$ and *rel*



F - 11.5 Figure: goodness-of-fit test - $K = 1$ and *log*



F - 11.6 Figure: goodness-of-fit test - comparison for t-copula and $K = 1$



F - 11.7 Figure: goodness-of-fit test - comparison for t-copula and $K = 5$

11.2.2 Goodness-of-Fit Tests for Weekly Changes

Concerning the difference between Gaussian and t-copula models, results are confirmed for weekly changes, see Figures F - B.4, F - B.5, and F - B.6.

In contrast to $K = 1$, models provide acceptable fittings even in volatile markets (except for *abs* and $D = 24$, see F - B.4). For this finding, it has to be considered that calibration bases on a two-year horizon. This weakens the influence of volatile periods on the fitting.

Analyzing goodness-of-fit tests for the t-copula in detail, the following insights are received from Figure F - 11.7:

1. Movements of p-values are alike for *rel* and *log*. For *abs* they show slight different patterns. This differs from daily evaluation.
2. Smaller portfolios provide better fitting results.
3. Results for models with empirical margins are better than those with α-stable margins.

11.2.3 Simulation and Convergence Results

Due to the special form of our model (path-dependence and cut-off constructions), an analysis of convergence is important. For $K = 1$, $K = 5$ and for different numbers of Monte-Carlo runs ($M =$ 5.000, 10.000 and 20.000), the behavior of convergence is surveyed by comparison of simulated default rates (tabularly) and by evaluation of expected intensities (graphically).

Expected intensities and simulated default rates for $d = 1, \ldots, D$ are computed as

$$\begin{aligned} \lambda^d_{t_n}(sim) &= \frac{1}{M}\sum_{m=1}^{M} \lambda^d_{t_n}(m) \quad \text{and} \\ DP^d_{sim} &= \frac{1}{M}\sum_{m=1}^{M} \mathbf{1}_{\left\{\exp\left(\int_0^T \lambda^d_s(sim)ds\right) \leq U^d(m)\right\}} \end{aligned}$$

for each t_n, $n = 1, \ldots, N$, and with $m = 1, \ldots, M$ denoted as the number of Monte-Carlo run.

As first result (compare Tables T - B.1, T - B.2, T - B.3 and T - B.4 in Appendix B.5), we record that one-year simulated default rates for relative and logarithmic changes on daily and weekly calibration are not in line with market perceptions.

Discussion - Calibration Basis

This finding is reasonable as the intensity is calibrated to relative and logarithmic returns. Therefore, it is implicitly supposed to have an exponential evolution. If now the underlying distribution is neither centered nor symmetric which is generally the case for empirical distributions, the intensity receives a drift. If this drift term is positive, the exponential development for the intensity process is increased and additionally supported by term effects. As consequence, default probability of the single counterparty rises. Moreover, this impact is possibly strengthened by model properties (conditional distributions in case of defaults). This leads to default rates which are unrealistic from a historical point of view as well as from current market perspectives.

Subject to future research might be the analysis of default rates (and intensities) of *rel* and *log* models for which drift terms are deducted.

For models calibrated to absolute changes, simulated default rates are related to practical expectations in consideration of the made assumptions, see Tables T - 11.1 and T - 11.2[10]. Comparing $K = 1$ and $K = 5$, we generally observe slightly higher estimates for the daily setup as for a calibration to a weekly basis.

[10]For instance, the estimated DP_{sim} of Daimler AG, $d = 22$, is in between 0.2 and 0.5%.

	Gaussian copula			***t*-copula**		
cpty no.	$M = 5000$	$M = 10000$	$M = 20000$	$M = 5000$	$M = 10000$	$M = 20000$
1	0.440	0.520	0.450	0.440	0.480	0.450
2	0.280	0.260	0.300	0.260	0.450	0.365
3	0.680	0.650	0.595	0.660	0.760	0.625
4	0.040	0.060	0.065	0.080	0.030	0.065
5	1.200	1.070	1.285	1.200	1.220	1.160
6	0.200	0.180	0.185	0.200	0.240	0.160
7	4.120	4.180	4.125	4.260	4.210	4.275
8	0.480	0.450	0.390	0.400	0.640	0.610
9	0.140	0.270	0.210	0.120	0.290	0.240
10	0.420	0.670	0.590	0.500	0.720	0.620
11	0.240	0.220	0.255	0.280	0.240	0.245
12	0.120	0.170	0.205	0.100	0.180	0.200
13	0.260	0.370	0.365	0.280	0.370	0.320
14	0.240	0.260	0.285	0.280	0.230	0.225
15	0.660	0.730	0.725	0.720	0.760	0.725
16	0.860	0.470	0.600	0.680	0.530	0.610
17	0.440	0.490	0.515	0.420	0.320	0.420
18	0.960	0.900	0.890	1.060	0.960	1.085
19	0.340	0.370	0.385	0.320	0.310	0.350
20	0.260	0.210	0.150	0.320	0.190	0.155
21	1.200	1.150	1.120	0.960	0.910	1.020
22	0.320	0.190	0.270	0.400	0.210	0.215
23	0.380	0.420	0.490	0.480	0.450	0.540
24	0.860	0.720	0.625	1.320	1.280	1.190

T - 11.1 Table: simulated default rates for $K = 1$ and *abs* - in percent

Tables T - 11.1 and T - 11.2 indicate that convergence of predicted default rates is ambiguous. Evolutions of the expected intensity process reveal similar outcomes, see F - B.7 to F - B.52 in Appendix B.5. Differences for both, default rates and intensities, do not show specific patterns. This finding is independent of model specification, i.e., Gaussian or t-copula respectively daily or weekly calibration basis. Using the example of counterparty $d = 1$, we discuss simulation results and convergence behavior.

Discussion - Simulation Results and Convergence Behavior

Comparing Figures F - 11.8 and F - 11.9, we observe that the general level of the expected intensity for $K = 1$ is lower than for $K = 5$. Accordingly, this leads to lower DPs for daily based calibration ($\approx 0.5\%$ for $K = 1$ vs. $\approx 0.8\%$ and for $K = 5$). Moreover, intensities exhibit remarkable unequal developments. These effects are partly explainable by the different calibration horizons (one year for $K = 1$ vs. two years for $K = 5$).

For $K = 1$, the expected intensity path for $M = 10.000$ (blue line) is slightly larger than for $M = 5.000$ (yellow) and $M = 20.000$ (black). Corresponding default rates validate these outcomes, compare Table T - 11.1. For Gaussian and t-copula, we find similar DP levels as well as approximately equal evolutions and convergence behaviors of the corresponding intensities.

*As both models are calibrated to absolute returns and empirical distributions, one would assume that the mean of the empirical cdf reasons a **linear** development for λ. Here, we observe a negative, **non-linear** evolution which thins out and results in a flattening curve. We suppose that this is generated by two model features, the truncation of margins and the use of conditional copulas.*

	Gaussian copula			***t*-copula**		
cpty no.	$M = 5000$	$M = 10000$	$M = 20000$	$M = 5000$	$M = 10000$	$M = 20000$
1	0.940	0.770	0.785	0.900	0.630	0.720
2	0.400	0.320	0.400	0.380	0.360	0.400
3	0.820	1.060	0.680	0.760	0.770	0.820
4	0.060	0.050	0.040	0.080	0.090	0.035
5	1.600	1.700	1.695	1.660	1.990	1.620
6	0.080	0.150	0.210	0.140	0.210	0.160
7	5.100	5.160	4.690	4.820	4.730	4.730
8	0.880	1.160	0.965	0.800	0.950	0.975
9	0.240	0.170	0.275	0.200	0.210	0.275
10	0.500	0.460	0.475	0.580	0.540	0.490
11	0.240	0.270	0.230	0.320	0.210	0.240
12	0.140	0.060	0.180	0.140	0.090	0.155
13	0.300	0.270	0.255	0.400	0.260	0.230
14	0.400	0.240	0.330	0.340	0.230	0.270
15	0.940	0.900	1.000	1.020	0.800	0.925
16	0.660	0.400	0.535	0.700	0.420	0.530
17	0.600	0.570	0.505	0.420	0.440	0.495
18	2.260	2.250	2.470	2.240	2.270	2.610
19	0.820	0.630	0.650	0.720	0.730	0.720
20	0.520	0.360	0.235	0.400	0.520	0.320
21	0.380	0.510	0.545	0.600	0.600	0.620
22	0.340	0.450	0.350	0.260	0.420	0.345
23	0.920	0.690	0.850	0.680	0.490	0.610
24	0.980	0.940	0.865	1.180	1.190	1.170

T - 11.2 Table: simulated default rates for $K = 5$ and *abs* - in percent

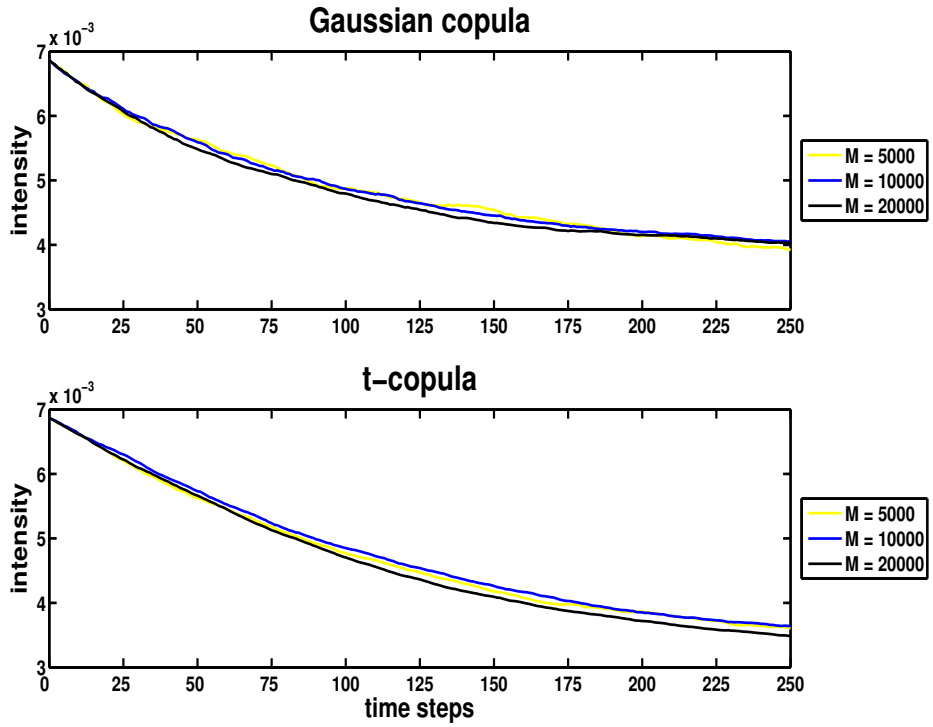
If the cut-off gets relevant (i.e., the current intensity level can fall below zero for the next random change drawn from the empirical cdf), the volatility of the corresponding margin decreases as both tails are truncated. Due to the changing mean[11] and lower volatilities of these new distributions, it is likely to receive a smoothing effect as depicted in Figure F - 11.8.

Additionally, the flattening curve might be influenced by conditional copulas. As time passes, the probability of defaults increases. Occurring defaults trigger the use of conditional copulas. In turn, these derivatives can shift probabilities of marginal distributions towards lower or higher values dependent on the associated dependence structure and copula model. The intensity level of the t-copula is slightly lower than for the Gaussian copula model. One reason can be the higher tail probability of the t-copula attended by higher probabilities for larger changes.

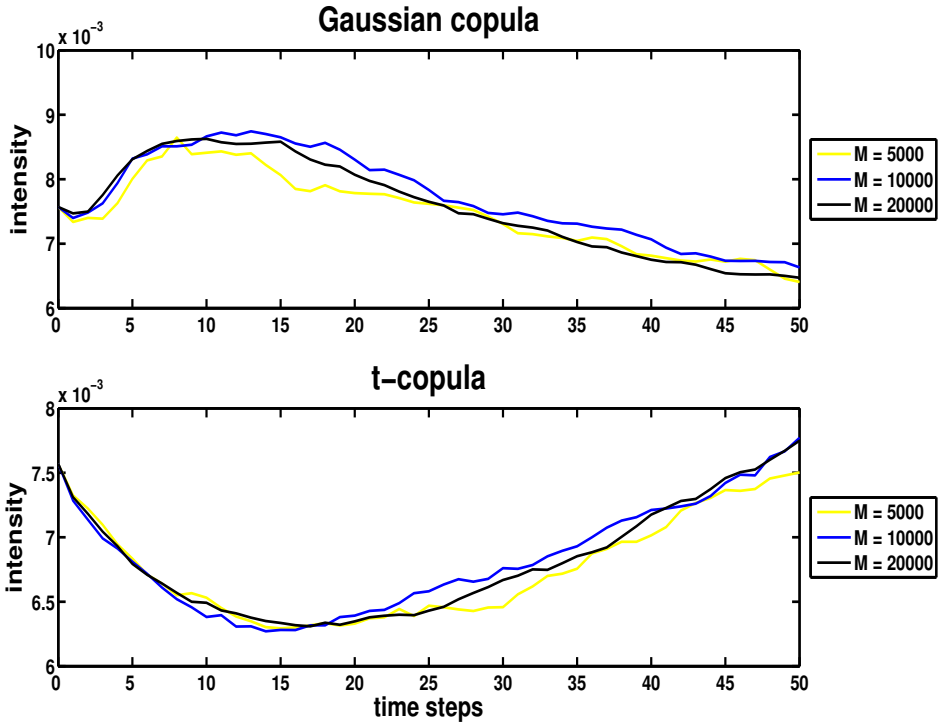
Summarized, we assume that the negative evolution (i.e., negative mean) thins out due to the combining effects of above mentioned aspects.

For $K = 5$, similar supposition apply. Here, we observe two remarkable features. First, we find different shapes and evolutions for the Gaussian and t-copula. This is generally not the case, compare Figures F - B.7 to F - B.52[12]. The behavior of convergence is not as smooth as for $K = 1$ which is partly reasonable due to the finer resolution for empirical margins of a daily calibration (250 observations for $K = 1$ vs. 100 for $K = 5$). Further, intensity levels cannot be connected to simulated default rates. In contrast to the intensity evolution, DP_{sim} show equal patterns.

[11] Note that the cdf is not necessarily centered.
[12] Exceptions are counterparties $d = 11$, 24 for $K = 1$ and $d = 1$, 12, 18 and 23 for $K = 5$.



F - 11.8 Figure: simulated intensity for counterparty $d = 1$ - $K = 1$ and *abs*



F - 11.9 Figure: simulated intensity for counterparty $d = 1$ - $K = 5$ and *abs*

Second, we remark a change of regime for both models. The intensity for the Gaussian copula first increases and decreases from step 15 on - vice versa for the t-copula model. Here, we assume that first a positive / negative mean drives the Gaussian / t-copula intensity. Then, above mentioned effects (truncation of margins and conditional copulas) presumably set in. Contrarily to $K = 1$, these impacts seem to change the positive / negative outlook into a negative / positive evolution.

This discussion reveals several topics for future research. From a numerical point of view, one could analyze how model features (truncation and conditional copulas) precisely impact intensity processes or to what extent their effects are separable and measurable for each individual aspect. Additionally, this matter can be linked with a more profound survey of regime changes as observed in Figure F - 11.9. Further, the behavior of convergence for default rates and intensities could be studied regarding these model properties as well as the number of required simulation runs.

11.2.4 Deviations - Backtesting and Benchmarking

Assessing model performance is difficult, as we simulate default probabilities DP_{sim} detached from complete portfolio perspective (with relevant LGD and EaD components) or pricing issues. Besides the default trigger, DP prediction bases on the underlying intensity. Therefore, backtesting and benchmarking is best performed by comparison of simulated intensities and default rates to standard approaches.

	$K = 1$			$K = 5$		
cpty no.	*const*	Gaussian copula	t-copula	constant	Gaussian copula	t-copula
1	0.422	**0.149**	0.163	0.522	0.501	**0.397**
2	0.744	**0.275**	0.403	0.879	**0.699**	0.722
3	0.742	**0.262**	0.275	0.708	**0.561**	0.592
4	**0.325**	0.586	0.548	**0.303**	0.575	0.560
5	0.693	0.516	**0.495**	**0.690**	0.759	0.758
6	0.438	**0.208**	0.224	0.780	**0.162**	0.207
7	1.146	**0.760**	0.785	1.209	1.087	**1.051**
8	1.463	0.239	**0.189**	1.483	1.026	**0.942**
9	0.596	**0.187**	0.253	0.526	**0.230**	0.237
10	0.320	0.196	**0.188**	0.387	0.405	**0.330**
11	0.702	**0.551**	0.676	0.709	**0.336**	0.370
12	**0.879**	1.036	1.006	0.718	0.728	**0.624**
13	0.519	0.537	**0.502**	0.436	**0.240**	0.242
14	**0.085**	0.182	0.285	0.170	0.086	**0.075**
15	0.262	0.222	**0.204**	**0.425**	0.461	0.445
16	0.263	0.127	**0.112**	0.331	**0.132**	0.155
17	0.285	0.142	**0.136**	0.370	0.183	**0.165**
18	**0.326**	0.537	0.396	0.419	**0.392**	0.480
19	**0.268**	0.346	0.441	0.242	0.243	**0.238**
20	**0.314**	0.572	0.593	**0.261**	0.326	0.336
21	1.209	0.819	**0.680**	1.221	**0.307**	0.386
22	0.725	0.493	**0.467**	1.643	**0.908**	1.019
23	0.326	0.191	**0.131**	0.386	0.612	**0.235**
24	1.252	**0.241**	1.536	1.626	**0.433**	1.140
mean	0.596	**0.391**	0.445	0.685	**0.475**	0.488
count	6	8	**10**	4	**11**	9

T - 11.3 Table: comparison of intensities: means of absolute relative differences

11.4 Assumption
For $N = 250$, $K = 1$ ($N = 50$, $K = 5$ respectively), $n = 0, \ldots, N$ and $d = 1, \ldots, D$ we set:

1. **Backtesting** - pseudo-real approach denoted as *real*:

$$\lambda^d_{t_n}(real) \;=\; \lambda^d_{t_n,real} \quad \text{and} \quad DP^d_{real} \;=\; 1 - \exp\left(-\int_0^1 \lambda^d_{t_n,real} dt_n\right)$$

   in which $\lambda^d_{t_n,real}$ are the realized intensities described in Subsection 11.1.4. We call this approach pseudo-real as the connection of DP and λ via $\lambda \approx \frac{s}{LGD} = s$ is a rough approximation. All the same we assume that knowing the whole path of λ (i.e., $\lambda_{t_n,real}$, $n = 0, \ldots, N$) leads to the "right" or "real" DP for the period from $t_0 = 0$ to $t_N = 1$.

2. **Benchmarking** - constant approach denoted as *const*:

$$\lambda^d_{t_n}(const) \;=\; \lambda^d_0 \quad \text{and} \quad DP^d_{const} \;=\; 1 - \exp\left(-\lambda^d_0\right)$$

   in which λ^d_0 is the initial value of the model. Using a constant intensity for one year DP estimations is commonly used in practice. Therefore, it serves as practical relevant and as simple benchmark.

3. Absolute relative differences for intensities and default rates are computed as

   a) intensities - for $d = 1, \ldots, D$ and $n = 1, \ldots, N$:

$$\Delta^d_{t_n}(x) = \left|\frac{\lambda^d_{t_n}(x) - \lambda^d_{t_n}(real)}{\lambda^d_{t_n}(real)}\right|$$

   b) mean - $d = 1, \ldots, D$:

$$\Delta^d(x) = \frac{1}{N}\sum_{n=1}^{N} \Delta^d_{t_n}(x)$$

   c) default probability - for $d = 1, \ldots, D$:

$$\Delta DP_x = \left|\frac{DP^d_x - DP^d_{real}}{DP^d_{real}}\right|$$

   in which $x \in \{sim,\ const\}$.

Due to the poor DP predictions revealed in the previous subsection, further analysis for models calibrated to relative and logarithmic changes is omitted. Sequent results are derived from models fitted on absolute changes with calibration basis $K = 1$ and $K = 5$ for the Gaussian and t-copula.

Table T - 11.3 contains a comparison of the means of absolute relative differences for the constant approach ($\Delta^d(const)$) as well as for the Gaussian and t-copula intensities ($\Delta^d(sim)$). On portfolio average, the Gaussian copula provides best results for both calibrations $K = 1$ and $K = 5$ followed by the t-copula and the constant benchmark. On absolute counts, the t-copula performs better for $K = 1$ and the Gaussian copula for $K = 5$. Generally, differences between the copula approaches are small. Gaps between model and benchmark are significant.

Table T - 11.4 compares "real" default rates DP^d_{real} with DP_{const} and DP_{sim}. They confirm results observed for the intensity. Models provide better backtesting results than the constant approach. In contrast, t-copula models perform better than the Gaussian ones on portfolio average and on absolute count.

For future research, further benchmark approaches could be examined. For instance, multivariate normal distributions or the independence copula could be used as dependence structure. Next, we analyze the quality of deviations.

$K = 1$							
		constant		**Gaussian copula**		***t*-copula**	
cpty no.	*real*	*const*	Δ *const*	*abs*	Δ *abs*	*abs*	Δ *abs*
1	0.515%	0.690%	0.340	0.450%	**0.126**	0.450%	**0.126**
2	0.272%	0.427%	0.570	0.300%	**0.103**	0.365%	0.342
3	0.630%	0.991%	0.573	0.595%	0.056	0.625%	**0.008**
4	0.174%	0.111%	**0.360**	0.065%	0.626	0.065%	0.626
5	1.155%	1.641%	0.421	1.285%	0.113	1.160%	**0.004**
6	0.163%	0.223%	0.365	0.185%	0.132	0.160%	**0.021**
7	2.796%	4.918%	0.759	4.125%	**0.475**	4.275%	0.529
8	0.649%	1.195%	0.841	0.390%	0.399	0.610%	**0.060**
9	0.232%	0.338%	0.454	0.210%	0.097	0.240%	**0.032**
10	0.768%	0.980%	0.276	0.590%	0.232	0.620%	**0.193**
11	0.186%	0.291%	0.568	0.255%	0.374	0.245%	**0.320**
12	0.089%	0.147%	**0.661**	0.205%	1.316	0.200%	1.259
13	0.234%	0.334%	0.426	0.365%	0.558	0.320%	**0.366**
14	0.292%	0.296%	**0.014**	0.285%	0.024	0.225%	0.229
15	0.717%	0.772%	0.076	0.725%	**0.011**	0.725%	**0.011**
16	0.534%	0.638%	0.194	0.600%	**0.123**	0.610%	0.141
17	0.485%	0.584%	0.205	0.515%	**0.062**	0.420%	0.134
18	1.780%	2.065%	**0.160**	0.890%	0.500	1.085%	0.390
19	0.737%	0.526%	**0.286**	0.385%	0.478	0.350%	0.525
20	0.513%	0.372%	**0.275**	0.150%	0.707	0.155%	0.698
21	0.759%	1.218%	0.604	1.120%	0.475	1.020%	**0.343**
22	0.197%	0.312%	0.586	0.270%	0.372	0.215%	**0.093**
23	0.558%	0.678%	0.215	0.490%	0.122	0.540%	**0.033**
24	0.644%	1.178%	0.830	0.625%	**0.029**	1.190%	0.849
mean			0.419		0.313		**0.306**
count			6		7		**13**

$K = 5$							
		constant		**Gaussian copula**		***t*-copula**	
cpty no.	*real*	*const*	Δ *const*	*abs*	Δ *abs*	*abs*	Δ *abs*
1	0.523%	0.757%	0.448	0.690%	**0.320**	0.720%	0.378
2	0.280%	0.474%	0.693	0.427%	0.525	0.400%	**0.428**
3	0.643%	0.993%	0.545	0.991%	0.542	0.820%	**0.276**
4	0.171%	0.112%	**0.345**	0.111%	0.351	0.035%	0.795
5	1.148%	1.658%	0.444	1.641%	0.430	1.620%	**0.411**
6	0.165%	0.280%	0.695	0.223%	0.350	0.160%	**0.031**
7	2.884%	5.375%	0.864	4.918%	0.705	4.730%	**0.640**
8	0.677%	1.261%	0.864	1.195%	0.766	0.975%	**0.441**
9	0.239%	0.332%	0.389	0.338%	0.414	0.275%	**0.151**
10	0.780%	1.052%	0.348	0.980%	**0.256**	0.490%	0.372
11	0.191%	0.299%	0.569	0.291%	0.527	0.240%	**0.259**
12	0.091%	0.135%	**0.485**	0.147%	0.617	0.155%	0.705
13	0.236%	0.319%	0.351	0.334%	0.414	0.230%	**0.026**
14	0.292%	0.339%	0.162	0.296%	**0.014**	0.270%	0.075
15	0.726%	0.967%	0.331	0.772%	**0.063**	0.925%	0.273
16	0.538%	0.689%	0.280	0.638%	0.185	0.530%	**0.016**
17	0.487%	0.639%	0.313	0.584%	0.200	0.495%	**0.017**
18	1.822%	2.378%	0.305	2.065%	**0.134**	2.610%	0.433
19	0.734%	0.671%	0.086	0.526%	0.283	0.720%	**0.019**
20	0.495%	0.465%	**0.061**	0.372%	0.249	0.320%	0.354
21	0.780%	1.259%	0.615	1.218%	0.562	0.620%	**0.205**
22	0.200%	0.492%	1.455	0.312%	**0.557**	0.345%	0.721
23	0.565%	0.731%	0.293	0.678%	0.199	0.610%	**0.079**
24	0.667%	1.422%	1.131	1.178%	0.765	1.170%	**0.753**
mean			0.503		0.393		**0.327**
count			3		6		**15**

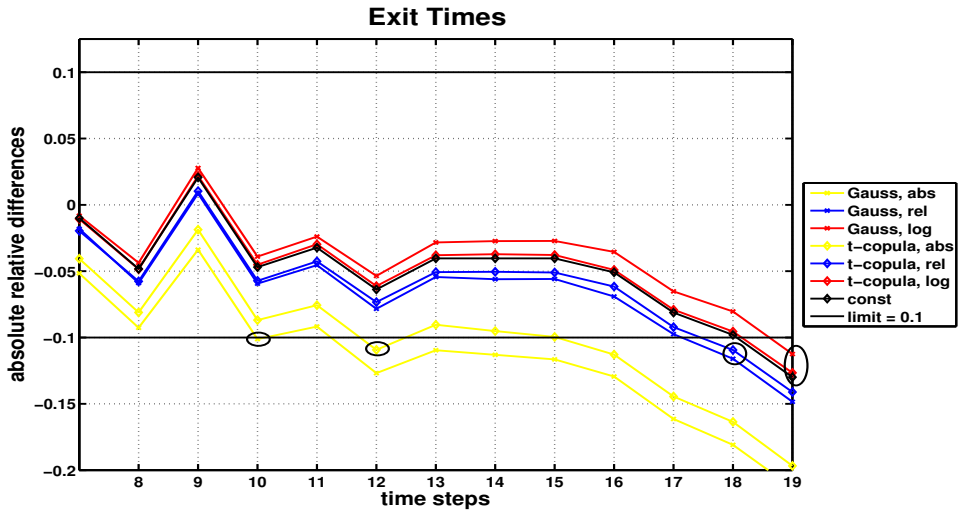
T - 11.4 Table: simulated vs predicted default rates for $K = 1$ and $K = 5$

11.2.5 Exit Times

We suppose that a perfect prediction is given by the "real" intensity. Thus, intensities $\lambda_{t_n}(sim)$ and $\lambda_{t_n}(const)$ are compared to the evolution of $\lambda_{t_n}(real)$, and deviations are computed. Moreover, upper and lower boundaries of predefined limits are set for these deviations. These work like a corridor incorporating time-dependence and evolution of the "real" intensity. If the compared intensities $\lambda_{t_n}(sim)$ or $\lambda_{t_n}(const)$ cross the limits, i.e., breaks out of the corridor, we denote the corresponding time as *exit time*. Mathematically seen, we compute

$$exit\ time\ (d,x) = \inf\left\{ n \in \{1,\dots,N\} : \left| \frac{\lambda^d_{t_n}(x) - \lambda^d_{t_n}(real)}{\lambda^d_{t_n}(real)} \right| = \left|\Delta^d_{t_n}(x)\right| > limit \right\}$$

for $x \in \{sim,\ const\}$. Economically seen, we allow an absolute (positive or negative) difference of the given limit between the comparative intensities $\lambda_{t_n}(sim)$ or $\lambda_{t_n}(const)$ and the "real" intensity $\lambda_{t_n}(real)$. If this difference is above the limit, we interpret the deviation as too large. Model or benchmark prediction is no longer regarded to be adequate. This analysis focuses a more qualitative and short-term perspective for model performance. As example, Figure F - 11.10 depicts exit times for counterparty $d = 1$.



F - 11.10 Figure: exit times for counterparty $d = 1$ and $K = 1$

Table T - 11.5 (for $K = 1$ with limit $= 0.1$ and for $K = 5$ with limit 0.25) shows exit times for all model specifications. For $K = 1$, the constant approach outperforms all simulations on portfolio average and portfolio count. Concerning models, t-copula calibrated to absolute changes provides best results on average. On count, t-copula fitted on logarithmic changes has the highest value. For $K = 5$, Table T - 11.5 shows that the t-copula calibrated to absolute changes performs best on average and on count followed by the Gaussian copula fitted on absolute returns. In contrast to $K = 1$, these models outperform the constant benchmark. Due to the focus on short-term evaluation, models with *rel* and *log* changes provide better results than observed for simulation.

It turns out that exit times depend on the specific counterparty. For instance, we observe ranges from 1 to 33 for $K = 1$ and *const*, 2 to 35 for $K = 5$ and *const*. Often, exit times for the different bases (*const*, *abs*, *rel* and *log*) are on a similar level, e.g., $d = 16$ for $K = 1$ or $d = 2$ for $K = 5$. Concerning the different horizons, exit times are mainly congruent, e.g., $d = 15$ with low exit times for $K = 1$ and $K = 5$. However, the last two aspects do not hold in general, e.g., see $d = 21$ for $K = 1$ and $K = 5$. Therefore, exit time results should be connected with intensity evolutions under consideration of previous discussions which may be subject to future research. Moreover, one could analyze the dependence of exit times with respect to the chosen limit.

	$K = 1$							$K = 5$						
		Gaussian copula			***t*-copula**				**Gaussian copula**			***t*-copula**		
cpty	*const*	*abs*	*rel*	*log*	*abs*	*rel*	*log*	*const*	*abs*	*rel*	*log*	*abs*	*rel*	*log*
1	**19**	10	18	**19**	12	18	**19**	10	9	3	3	**12**	9	8
2	**32**	20	6	6	**32**	6	6	**9**	**9**	8	8	**9**	**9**	8
3	**4**	3	2	2	3	2	2	9	**12**	9	7	**12**	9	7
4	6	6	5	**10**	6	5	**10**	**5**	**5**	**5**	**5**	**5**	**5**	**5**
5	23	**24**	23	23	**24**	23	23	**8**	**8**	**8**	4	**8**	**8**	4
6	16	**20**	10	9	17	11	10	2	**6**	2	2	**6**	2	2
7	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**11**	**11**	9	8	**11**	8	8
8	**4**	3	3	3	3	3	2	12	**15**	9	3	**15**	11	3
9	1	**7**	**7**	**7**	**7**	**7**	**7**	17	**27**	11	11	**27**	11	11
10	**9**	**9**	**9**	**9**	**9**	**9**	**9**	**12**	4	**12**	**12**	4	**12**	**12**
11	23	11	**31**	29	23	29	27	10	**16**	10	10	**16**	9	9
12	**33**	32	27	25	**33**	27	25	**15**	**15**	13	9	**15**	12	9
13	1	3	3	3	3	3	**5**	14	**16**	14	14	**16**	13	12
14	**9**	**9**	**9**	**9**	**9**	**9**	**9**	16	16	13	12	**17**	11	3
15	**3**	**3**	**3**	**3**	**3**	**3**	**3**	2	2	2	2	**3**	2	2
16	**20**	19	**20**	**20**	19	**20**	**20**	13	**38**	12	9	**38**	12	9
17	17	17	17	17	**21**	17	17	12	36	13	12	**40**	12	5
18	**19**	13	18	**19**	14	**19**	**19**	**36**	**36**	13	12	16	3	3
19	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**2**	**2**	**2**	**2**	**2**	**2**	**2**
20	**2**	**2**	**2**	**2**	**2**	**2**	**2**	2	**3**	2	2	**3**	2	2
21	**18**	4	4	3	16	3	3	16	6	**18**	16	6	12	11
22	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**
23	**20**	19	19	19	19	**20**	**20**	12	12	12	10	**33**	12	4
24	**3**	2	2	2	2	2	2	3	**12**	**12**	9	**12**	9	2
mean	**12.1**	10.2	10.3	10.3	11.9	10,3	10,3	10.4	13.3	8.9	7.7	**13.7**	8.2	6.0
count	**17**	10	10	12	12	11	14	9	17	7	4	**21**	6	4

T - 11.5 Table: exit time comparison - $K = 1$, limit $= 0.1$ / $K = 5$, limit $= 0.25$

11.3 Conclusion

The application presented in Section 11.1 relies on one main assumption. Intensities are connected to credit spreads s through the approximation $\lambda \approx s$ using $LGD = 1$[13]. This strong supposition bears more than just credit risk components (e.g., market or liquidity risk) and is therefore discussable. However, it is widely used in practice.

In total, the model can be calibrated and simulated by Gaussian and t-copulas with empirical and α-stable margins. Corresponding goodness-of-fit tests offer the following

11.5 Summary (Goodness-Of-Fit Test)

1. t-copula models perform by far better than Gaussian ones.
2. Calibration basis (*abs*, *rel* and *log* returns) does not have large influence on fitting results, whereas changes and horizons - daily and one year vs. weekly and two years - affect outcomes.
3. Smaller portfolios exhibit better fits than larger ones. In times of volatile markets, goodness-of-fit decreases for daily changes. For weekly returns, this effect is not observable - possibly due to the longer calibration horizon.

In comparison, models with empirical margins outperform the ones with α-stable marginal cdfs. Thus, model variability does not seem to secure better performance for goodness-of-fit tests.

[13] Besides, the fundamental condition of the intensity-based approach is that survival and default probabilities are modeled by the first jump of a Cox process and are hence dependent on the intensity process λ and the default trigger variable U.

Summarized, the t-copula model produces good fitting results especially in less volatile periods. For the Gaussian copula, outputs are worse but may be acceptable under certain conditions. Simulation results and analysis of convergence behavior are collected in

11.6 Summary (Simulation and Convergence Results)

1. Model effects are visible but not precisely assignable to the model features (marginal truncation and conditional copulas).
2. For one-year DP prediction, models calibrated to *rel* and *log* changes are not suitable. Due to the assumed exponential evolution and term effects, outcomes overestimate default probabilities compared to market perceptions.
3. For absolute returns, simulated DP are realistic.
4. Due to model construction, a large number of simulation runs is essential and required to receive convergence.

Backtesting and benchmarking is carried out under the assumption that the "real" i.e., observable intensity predicts the "real" default rate. Detached from a portfolio perspective, simulated outcomes ($DP_{sim},\ \lambda(sim)$ for *abs*) are backtested with realizations ($DP_{real},\ \lambda(real)$) and benchmarked with a constant approach ($DP_{const},\ \lambda(const)$).

11.7 Summary (Deviations - Backtesting and Benchmarking)

1. For all specifications, benchmark outputs are worse than simulation results on portfolio average and count.
2. Concerning deviations of intensities, the Gaussian copula model predominantly provides better results. For default rates, the t-copula model exhibits improved predictions. These facts hold for both daily and weekly calibration basis.

Under Assumptions 11.1 and 11.4 and in consideration of a long-run perspective (i.e., a one-year time horizon), modeled intensities and default rates deviate less from "reality" compared to the constant benchmark approach.

This finding must be linked with the presented exit time analysis. This evaluation focuses model performance for a short period and therefore from a more qualitative point of view measured in terms of deviations.

11.8 Summary (Exit Times)

1. For daily calibration ($K = 1$), constant intensities show better performances than the modeled ones.
2. For weekly changes ($K = 5$), the t-copula model with *abs* are best on portfolio average and counts followed by the Gaussian model with *abs*. The constant benchmark is outperformed.
3. Both approaches (model and benchmark) fail to work for several counterparties independent of calibration basis and horizon. This underlines the known fact that prediction of "real" intensities - be it through modeling or other methods - is challenging.

To what extent the model is applicable for pricing issues including surveying numerical results and practical implications may be subject to future research. Most likely, implementation is possible. Calibration under risk-neutral measures is complex whereas simulation is congruent to the real-world framework.

12 Summary

At this point we summarize main results and highlight essential as well as academically new insights of this thesis. To our knowledge, the intensity model construction of Chapter 10 has not been developed before.

One new aspect is the truncation of margins to preserve positivity of the underlying intensity. Additionally, the copula approach has not been applied for modeling dependencies of intensities in a reduced-form model in this particular way. The idea of copula derivatives for the default trigger and locked processes is borrowed from Schönbucher and Schubert [59]. However, it is firstly assigned to the intensity implying a direct default dependence.

The model incorporates a finite number of jumps (random and default changes). The intensity is hence constructed as discrete-time stochastic process in a fully continuous-time framework. This representation has been set up before. Though, a direct impact of defaults on the intensity through default changes is rarely used and analyzed.

In this context, we deal rather with a counting process with two stochastic components. These are linked and are not separable as for the usual Cox process. Due to this construction and resulting features (special form of underlying filtrations, marginal truncation, conditional copulas and their recursive exploitation, default jump component and "non-separable" counting and intensity process), our model is highly path-dependent. The general differentiation between default and non-default information is not possible anymore. This new idea is substantially different to the normal information-based approach of reduced-form models. Nevertheless, as the intensity is defined as càdlàg process, our model still fits into the generalized continuous-time intensity-based setup - regarding survival probabilities for instance, see Proposition 10.21.

In the next sections, we assess advantages and drawbacks of the model from a methodological and a numerical point of view.

12.1 Model - Pros and Cons

We recapitulate the central statements and findings of Chapter 10. Model construction focuses on a high degree of freedom concerning applicable distributions, specifications and time grid as well as implementation of copula features. Leading benefits are:

Degree of model freedom

+ The model uses arbitrary distributions. This aspect is restricted in so far as margins are assumed to be continuous and have to fulfill certain properties dependent on the scope of application. Further, the copula must be differentiable.
+ The underlying time grid (i.e., the number of predefined random changes N) is universally valid.
+ Random default changes can be omitted.
+ Truncation of margins - if necessary - is possible for the lower tail or both tails.

Robustness

+ Dependence structures are maintained on the whole time horizon due to the invariance of copulas. This also holds in the case of differentiation. Thus, recalibration is not needed.
+ Marginal distributions are initially set and are sustained. A direct evaluation of conditional, i.e., truncated distributions is not required.

Economic reasonable interpretation and realistic features

+ The model incorporates default contagion. Effects are twofold. On the one hand, the general setup changes due to conditional copulas and locked processes. On the other hand, a random default change is instantaneously triggered. Thus, defaults directly affect intensities.

+ Default and non-default information is used simultaneously and go into the model as one relevant filtration. As such a separation of information is not observable in markets, this is a realistic feature from an economic point of view.

+ The discrete-time setup for the intensity reduces complexity and facilitates implementation. As side effect, several defaults can possibly occur in one interval.

In parts, the above mentioned advantages are mutually dependent. However, they also cause drawbacks:

- Truncation of tails lead to rescaled probabilities and hence to a reduced volatility. This could not be desired or realistic, respectively.

- Due to the special model features (truncation of margins, differentiation of copulas in case of defaults), analytical in sense of closed-form or unique solutions do not exist in general. They depend on the applied distributions (i.e., copulas and margins). It may happen that requirements on the intensity (e.g., validity of pricing formulas, martingale properties) can hardly or not at all be fulfilled under certain conditions.

- Model construction bases on the fact that default and non-default information are simultaneously exploited. Hence, the main pro of the general information-based setup (i.e., modeling and evaluating default-free and defaulted variables separately) is suspended.

The last aspect leads to a sort of "certain" default prediction which in turn triggers the default change. Nevertheless, this framework is comparable to the general setup. In our model, the perspective on time is extended or - in other words - more granular:

- In the general model, the intensity is evaluated first. Second, defaults are simulated afterwords on the whole time horizon.

- In our model, intensities and defaults are simultaneously analyzed on one interval. This evaluation produces default times. Interpreting this interval as complete time horizon, this matches the common model up to this point. Now, these default times effect intensities through default changes and conditional copulas. Thus, it can be regarded as extension of the general setup by a component of default time. In particular, this allows to model contagion effects on dependence structures of intensities and is not incorporated in the generic intensity-based framework.

12.2 Implementation - Pros and Cons

Concerning numerical implementation, our model provides following advantages:

+ Due to the inverse modeling approach, only copula-distributed random variables are used.

+ Truncation of marginal distributions can easily be implemented. A particular evaluation of conditional margins is not required.

+ Implementation is possible and provides reasonable outcomes. Model effects are observable.

+ Calibration has to be performed only once. A recalibration due to defaults is not necessary as their effects are considered.

Numerical drawbacks are mainly provoked by the special model construction. Partly, these cons depend on specifications and can therefore be diminished.

- Inversion of marginal distributions is necessary - even though only once.

- Numerical differentiation of copulas is complex - even though closed-from solutions exits, e.g., for the Gaussian copula.
- The model is path-dependent. Sequent simulations outputs rely on all former realizations.
- For default prediction and identification, all defaulted positions must be memorized.
- Due to the previous facts, implementation is computationally intensive.
- Calibration is only valid on the predetermined time grid. A sort of time-scaling as for the Brownian motion for instance is not possible. Moreover, default changes is most likely difficult to fit due to scarce data.

Summarized, we come to the conclusion that implementation of the developed model is possible. For the given specifications and under the fixed assumptions, fit tests provide good results. Model peculiarities are reflected in numerical outcomes. The model works out as whole period valuation, it outperforms a benchmark approach in backtesting from a long-term view. Short-term results are worse. The high complexity is attended by

- higher computing times and
- a difficult numerical implementation.

These disadvantages overthrow the gain on performance. Therefore, an application as DP prediction for a portfolio perspective is probably not workable in practice.

12.3 Outlook and Literature

For future research, the exact **numerical** interpretation and understanding of model details can be focused. Here, an in-depth analysis of a one-dimensional problem may reveal impacts of the marginal truncation. Moreover, an additional benchmark approach with the independence copula may offer dues to the effects of conditional copulas.

As further problem, analytical and numerical outcomes may be investigated with respect to their dependence of the time grid variable N, i.e., the number of predetermined random changes.

As possible central aspect for future research, it could be analyzed if our model can be calibrated for pricing problems and if it is of practical use for this scope of application. Concerning calibration, requirements on the intensity restrict the set of applicable distributions. Moreover, the underlying probability measure may not be unique and / or techniques of change of numeraires must be considered. Therefore, model fitting is exceptionally challenging. If calibration is reasonable, the model could be benchmarked with a standard intensity model specified by mean-reversion processes, for instance.

Concerning literature for copulas, the reader is referred to the monographs mentioned in Section 6.3, especially [12], [15], [37], [49] and [52].

For a general introduction to the world of credit risk, "(An Introduction to) Credit Risk Modeling" by Bluhm et al. [7] provides a comprehensive survey. Detailed descriptions of credit risk models and applications are found in

- Bielecki and Rutkowski's "Credit Risk: Modeling, Valuation and Hedging" [4],
- Brigo and Mercurio's "Interest Rate Models - Theory and Practice" [9] or
- Schönbucher's "Credit Derivatives Pricing Models" [58], for example.

For the mathematical treatment of probability measures in combination with Cox process, "Convergence of probabilities" [5] by Billingsley and "Doubly Stochastic Poisson Processes" [32] by Grandell are useful.

Part III

Appendix

A Appendix to Part I

A.1 Copula Related Topics

Let $S = S_1 \times \ldots \times S_D \subset \bar{\mathbb{R}}^D$ and $F : S \to \bar{\mathbb{R}}$. Suppose that S has a least element $s = (s_1, \ldots, s_D)'$.

A.1 Definition (grounded)
A function F is *grounded* if $F(x_1, \ldots, x_D) = 0$ for at least **one** x_d, $d \in \{1, \ldots, D\}$, with $x_d = s_d$.

A.2 Definition (D-increasing)
A function F is *D-increasing* if for all D-dimensional rectangles $[x_{1,1} \times x_{1,2}], \ldots, [x_{D,1} \times x_{D,2}] \in S$ with $x_{\cdot,1} \leq x_{\cdot,2}$ it holds

$$\sum_{i_1=1}^{2} \ldots \sum_{i_D=1}^{2} (-1)^{i_1+\ldots+i_D} F(x_{1,i_1}, \ldots, x_{D,i_D}) \geq 0 \quad .$$

The rectangle inequality is **one** necessary and sufficient condition for the function F to be a multivariate distribution function. It states that for each non-empty D-dimensional rectangle the function F has a positive value. For example, for the bivariate case with $[a, b] \times [c, d]$ with $a < b$ and $c < d$ it is

$$F(b, d) + F(a, c) - F(a, d) - F(b, c) \geq 0 \quad .$$

A.3 Definition (completely monotonic)
A decreasing function F is *completely monotonic* on an interval $[a, b] \subset \mathbb{R}$ if it holds

$$(-1)^d \frac{\partial^d}{(\partial x)^d} F(x) \geq 0, \ d \in \mathbb{N}, \ x \in [a, b] \quad .$$

A.4 Definition (pseudo-inverse, compare [49], Definition 5.41)
Assume that the function $\phi : [0, 1] \to [0, \infty]$ is strictly decreasing and continuous with $\phi(1) = 0$ and $\phi(0) \leq \infty$. The *pseudo-inverse* $\phi^{[-1]} : [0, \infty] \to [0, 1]$ of ϕ is defined as

$$\phi^{-1}(x) = \begin{cases} \phi^{[-1]}(x) & \text{for } 0 \leq x \leq \phi(0) \\ 0 & \text{for } \phi(0) < x \leq \infty \end{cases} \quad .$$

The pseudo-inverse function is used as generator to construct Archimedean copulas in Subsection 2.2.3. The generator is called *strict* if $\phi(\infty) = 0$. For a strict generator, it is $\phi^{[-1]} = \phi^{-1}$.

A.2 Multivariate Distribution Functions

Let $X = (X_1, \ldots, X_D)$, $D \in \mathbb{N}$, be a $\mathbb{R}^D$-valued random variable. The multivariate cumulative distribution function $F_X : \mathbb{R} \to [0, 1]$ describes the multivariate probability that $X \leq x$ for a given vector $x = (x_1, \ldots, x_D)' \in \mathbb{R}^D$, i.e.,

$$F_X(x) = \mathbb{P}[X \leq x] = \mathbb{P}[X_1 \leq x_1, \ldots, X_D \leq x_D] \quad .$$

If F_X has a multivariate density function f_X and provided that F_X is differentiable, it holds

$$\frac{\partial^D}{\partial x_1 \ldots \partial x_D} F_X(x) = f_X(x)$$

which is equivalent to

$$F_X(x) = \int_{-\infty}^{x_D} \ldots \int_{-\infty}^{x_1} f_X(s_1, \ldots, s_D) ds_1 \ldots ds_D \quad .$$

The probability that **one** random variable X_d is smaller or equal to x_d, $d \in \{1, \ldots, D\}$, can be associated to the multivariate probability

$$\mathbb{P}[X_d \leq x_d] = \mathbb{P}[X_1 < \infty, \ldots, X_{d-1} < \infty, X_d \leq x_d, X_{d+1} < \infty, \ldots, X_D < \infty]$$

and defines the margin or marginal cdf $F_d : \mathbb{R} \to [0, 1]$ as

$$F_d(x_d) = F_{X_d}(x_d) = F_X(\infty, \ldots, \infty, x_d, \infty, \ldots, \infty) = \int_{-\infty}^{x_d} f_d(s_d) ds_d$$

given that F_d is an one-dimensional distribution function with marginal density f_d.

A right-continuous function F on $\bar{\mathbb{R}}^D$ is a multivariate cdf if the following conditions are fulfilled[1]:

1. $\lim_{x_d \to -\infty} F(x) = 0$, $d = 1, \ldots, D$,
2. $\lim_{x_d \to \infty \ \forall d} F(x) = 1$,
3. F is D-increasing.

A.5 Definition (generalized inverse, compare [12], p. 49)
The *generalized inverse* $F^{-1} : [0, 1] \to [-\infty, \infty]$ of a distribution function F is defined as

$$F^{-1}(u) = \inf\{x \in \bar{\mathbb{R}} | F(x) \geq u, \ 0 < u < 1\} \quad .$$

It is obvious that F^{-1} is the normal inverse function if F is strictly increasing. Do not mix up with Definition A.4 of the pseudo-inverse function ϕ^{-1} for the copula generator ϕ.

A.3 Conditional Multivariate Distributions

We want to determine the conditional multivariate probabilities

$$\mathbb{P}\left[X^{-d} \leq x^{-d} | X_d \leq x_d\right] \quad \text{and} \quad \mathbb{P}\left[X^{-d} \leq x^{-d} | X_d = x_d\right]$$

for $d \in \{1, \ldots, D\}$ subject to multivariate cdfs. Due to the definitions of conditional probabilities, the conditional multivariate cdfs are specified for $X_d \leq x_d$ and $X_d = x_d$ as

$$F_{X^{-d}|X_d \leq x_d}\left(x^{-d}\right) = \mathbb{P}\left[X^{-d} \leq x^{-d} | X_d \leq x_d\right] = \frac{F_X(x)}{F_d(x_d)}$$

$$F_{X^{-d}|X_d = x_d}\left(x^{-d}\right) = F_{X^{-d}|x_d}\left(x^{-d}\right) = \mathbb{P}\left[X^{-d} \leq x^{-d} | X_d = x_d\right] = \frac{\frac{\partial}{\partial x_d} F_X(x)}{\frac{\partial}{\partial x_d} F_d(x_d)} \quad .$$

W.l.o.g. we set $d = D$. It holds

$$\begin{aligned}
&F_{X^{-d}|x_d}\left(x^{-d}\right) = \mathbb{P}[X_1 \leq x_1, \ldots, X_{D-1} \leq x_{D-1} | X_D = x_D] \\
&= \lim_{\epsilon \to 0} \left\{\mathbb{P}[X_1 \leq x_1, \ldots, X_{D-1} \leq x_{D-1} | x_D \leq X_D \leq x_D + \epsilon]\right\} \\
&= \lim_{\epsilon \to 0} \left\{\frac{\mathbb{P}[X_1 \leq x_1, \ldots, X_{D-1} \leq x_{D-1}, x_D \leq X_D \leq x_D + \epsilon]}{\mathbb{P}[x_D \leq X_D \leq x_D + \epsilon]}\right\} \\
&= \lim_{\epsilon \to 0} \left\{\frac{\mathbb{P}[X_1 \leq x_1, \ldots, X_D \leq x_D + \epsilon] - \mathbb{P}[X_1 \leq x_1, \ldots, X_D \leq x_D]}{\mathbb{P}[X_D \leq x_D + \epsilon] - \mathbb{P}[X_D \leq x_D]}\right\} \\
&= \lim_{\epsilon \to 0} \left\{\frac{F_X(x_1, \ldots x_{D-1}, x_D + \epsilon) - F_X(x)}{F_D(x_D + \epsilon) - F_D(x_D)}\right\} \\
&= \lim_{\epsilon \to 0} \left\{\frac{\frac{1}{\epsilon}\left(F_X(x_1, \ldots, x_D + \epsilon) - F_X(x)\right)}{\frac{1}{\epsilon}\left(F_D(x_D + \epsilon) - F_D(x_D)\right)}\right\} = \frac{\frac{\partial}{\partial x_D} F_X(x)}{\frac{\partial}{\partial x_D} F_D(x_D)} \quad .
\end{aligned}$$

[1] Compare [37], 1.4.2 or [52], Definition 2.10.8, for instance.

We pass over to the multivariate distributions with **several** conditions. For that purpose let

- $\mathcal{D} = \{1, \ldots, D\}$, $\#\mathcal{D} = D$,
- $\mathcal{E} = \{d_1, \ldots, d_E \mid d_e \in \mathcal{D},\ e = 1, \ldots, E\}$, $\#\mathcal{E} = E$ and $\mathcal{E} \subset \mathcal{D}$,
- $\mathcal{C} = \{d_1, \ldots, d_C \mid d_c \in \mathcal{D},\ c = 1, \ldots, C\}$, $\#\mathcal{C} = C$ and $\mathcal{C} \subset \mathcal{D}$,
- $\mathcal{E} \cup \mathcal{C} = \mathcal{D}$ and $\mathcal{E} \cap \mathcal{C} = \emptyset$,
- $x_\mathcal{E} = (x_{d_1}, \ldots, x_{d_E})$,
- $x_\mathcal{C} = (x_{d_1}, \ldots, x_{d_C})$,
- $F_\mathcal{E} = F_{X_\mathcal{E}} = F_{X_{d_1}, \ldots, X_{d_E}}$ is the $\mathcal{E}$th marginal distribution of F_X.

For an arbitrary margin $F_\mathcal{E}$, it is $F_\mathcal{E}(x_\mathcal{E}) = F_X(y)$ with y defined as

$$y_d = \begin{cases} \infty & \text{for } d \notin \mathcal{E} \\ x_d & \text{for } d \in \mathcal{E} \end{cases} .$$

For the probability $\mathbb{P}[X_\mathcal{C} \leq x_\mathcal{C} | X_\mathcal{E} \leq x_\mathcal{E}]$, we receive the following conditional multivariate cdf

$$F_{X_\mathcal{C} | X_\mathcal{E} \leq x_\mathcal{E}}(x_\mathcal{C}) = \frac{F_X(x)}{F_\mathcal{E}(x_\mathcal{E})} = \frac{F_X(x)}{F_X(y)}$$

with y defined as above. For the expression $\mathbb{P}[X_\mathcal{C} \leq x_\mathcal{C} | X_\mathcal{E} = x_\mathcal{E}]$, we get with $\#\mathcal{E} = E$

$$F_{X_\mathcal{C} | x_\mathcal{E}}(x_\mathcal{C}) = F_{\mathcal{C}|\mathcal{E}}(x_\mathcal{C}) = \frac{\frac{\partial^E}{\partial x_\mathcal{E}} F_X(x)}{\frac{\partial^E}{\partial x_\mathcal{E}} F_\mathcal{E}(x_\mathcal{E})} = \frac{\frac{\partial^E}{\partial x_\mathcal{E}} F_X(x)}{\frac{\partial^E}{\partial x_\mathcal{E}} F_X(y)}$$

in which

$$\frac{\partial^E}{\partial x_\mathcal{E}} F_X(x) = \frac{\partial^E}{\partial x_{d_1} \ldots \partial x_{d_E}} F_X(x) \quad .$$

A.4 Random Number Generation

> Anyone attempting to generate random numbers by deterministic means is, of course, living in a state of sin.
>
> John von Neumann

Many monographs have been written about pseudo-random number generation, e.g., see [18]. Creating standard normally distributed random variables out of uniformly distributed ones can be performed for instance by the Box-Muller method or the polar algorithm, refer to [33], Subsection 5.2.2. Our aim is to create **correlated** normally distributed random variables.

Therefore, let $Z = (Z_1, \ldots, Z_D)' \in \mathbb{R}^D$ be a vector of independent, standard normally distributed variables Z_d, $d = 1, \ldots, D$. To receive a vector X which is $N(\mu, \Sigma)$-distributed with mean vector $\mu \in \mathbb{R}^D$ and covariance matrix $\Sigma \in \mathbb{R}^{D \times D}$, we perform a Cholesky decomposition of Σ:

$$\Sigma = L'L$$

and compute

$$X = \mu + L \cdot Z \quad .$$

This implies that $X \sim N(\mu, \Sigma)$, see [33], Subsection 5.2.3.

Let $Z_1, \ldots, Z_D$ be independent standard normal variables. The random variable

$$S^2 = \sum_{d=1}^{D} Z_d^2$$

is said to be $\chi^2(D)$-distributed with D degrees of freedom.

Let $S^2 \sim \chi^2(\nu)$, $Y \sim N(0, P)$ with correlation matrix P, and $Z = (Z_1, \ldots, Z_D)'$ be i.i.d with $Z_d \sim N(0, 1)$. Suppose that the Cholesky decomposition of P holds $P = L'L$. Then, the variable

$$X = \sqrt{\frac{\nu}{S}} \cdot Y = \sqrt{\frac{\nu}{S}} \cdot L \cdot Z$$

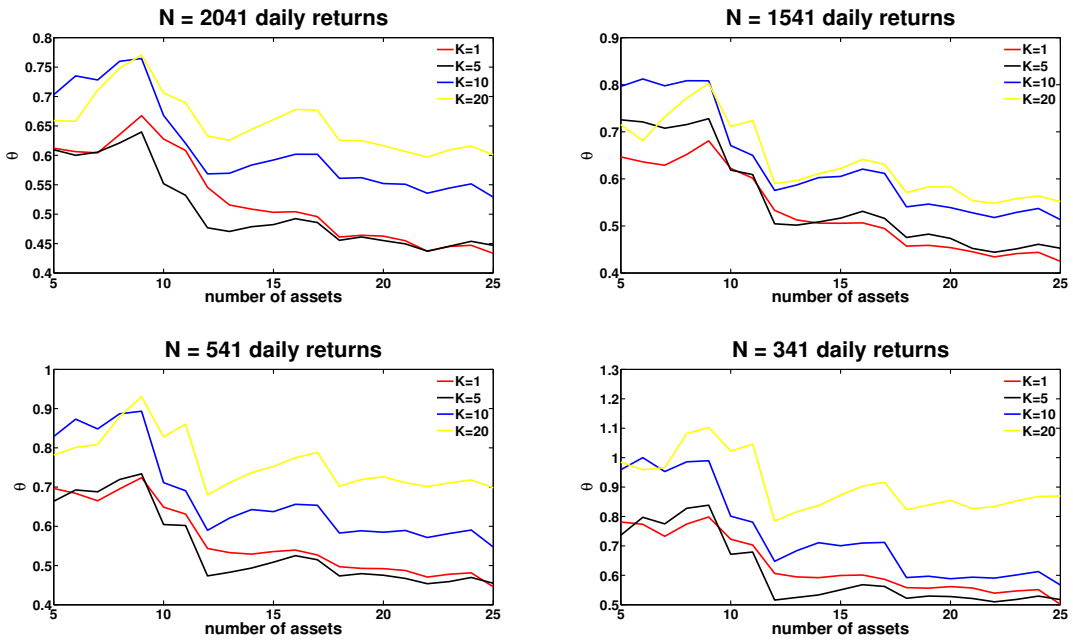
is multivariate t-distributed with ν degrees of freedom and correlation matrix P, e.g., see [39], Section 1.2.

A.5 Portfolio Composition - Part I

25 DAX Companies, December 31, 2009

adidas	BMW	EON	MAN	RWE
Allianz	Commerzbank	Fresenius Med.	Lufthansa	SAP
BASF	Daimler	Henkel	MAN	Siemens
Bayer	Deutsche Bank	K+S	Merck	ThyssenKrupp
Beiersdorf	Telekom	Linde	Münchener Rück	VW

A.6 Further Numerical Results - Part I



F - A.1 Figure: Clayton copula parameter θ - dependence on sample sizes

	period	**Gaussian copula**	***t*-copula**	**Clayton copula**	**BS model**
$\alpha = 0.95$ $D = 5$	$K = 1$	-0.024554	-0.02394	-0.024522	-0.026172
	$K = 5$	-0.051873	-0.051776	-0.052392	-0.049921
	$K = 10$	-0.0765	-0.078109	-0.082219	-0.075298
	$K = 20$	-0.11192	-0.11121	-0.11486	-0.10601
$\alpha = 0.99$ $D = 5$	$K = 1$	-0.041427	-0.043573	-0.049405	-0.035861
	$K = 5$	-0.084143	-0.089634	-0.09897	-0.070546
	$K = 10$	-0.12153	-0.13802	-0.15564	-0.10318
	$K = 20$	-0.17983	-0.19141	-0.20844	-0.1408
$\alpha = 0.995$ $D = 5$	$K = 1$	-0.049377	-0.055723	-0.058008	-0.040386
	$K = 5$	-0.095313	-0.10549	-0.12926	-0.078266
	$K = 10$	-0.14186	-0.15508	-0.17817	-0.11512
	$K = 20$	-0.20316	-0.21277	-0.23227	-0.16553
$\alpha = 0.95$ $D = 10$	$K = 1$	-0.026178	-0.026742	-0.026575	-0.028166
	$K = 5$	-0.055717	-0.056059	-0.055754	-0.056166
	$K = 10$	-0.084385	-0.084334	-0.08817	-0.0837
	$K = 20$	-0.12278	-0.11691	-0.12592	-0.11761
$\alpha = 0.99$ $D = 10$	$K = 1$	-0.045398	-0.043685	-0.04782	-0.039075
	$K = 5$	-0.090782	-0.09872	-0.10696	-0.077422
	$K = 10$	-0.13532	-0.15015	-0.16512	-0.11628
	$K = 20$	-0.19769	-0.2145	-0.22884	-0.16257
$\alpha = 0.995$ $D = 10$	$K = 1$	-0.051766	-0.060055	-0.06201	-0.042937
	$K = 5$	-0.1033	-0.11303	-0.126	-0.084946
	$K = 10$	-0.16525	-0.1783	-0.20497	-0.12766
	$K = 20$	-0.22614	-0.23665	-0.26853	-0.17406
$\alpha = 0.95$ $D = 15$	$K = 1$	-0.022741	-0.022756	-0.023131	-0.025161
	$K = 5$	-0.050345	-0.050245	-0.04847	-0.04913
	$K = 10$	-0.075107	-0.073455	-0.075859	-0.075226
	$K = 20$	-0.10672	-0.10201	-0.1104	-0.10569
$\alpha = 0.99$ $D = 15$	$K = 1$	-0.038609	-0.039543	-0.041434	-0.035836
	$K = 5$	-0.078084	-0.089626	-0.087991	-0.070433
	$K = 10$	-0.1253	-0.12777	-0.14519	-0.10276
	$K = 20$	-0.1739	-0.17865	-0.20147	-0.14288
$\alpha = 0.995$ $D = 15$	$K = 1$	-0.043759	-0.047951	-0.051952	-0.038873
	$K = 5$	-0.089194	-0.10354	-0.11735	-0.078657
	$K = 10$	-0.14422	-0.15614	-0.17915	-0.11374
	$K = 20$	-0.19334	-0.21427	-0.22555	-0.15758
$\alpha = 0.95$ $D = 20$	$K = 1$	-0.022788	-0.022873	-0.022204	-0.024878
	$K = 5$	-0.049028	-0.048912	-0.048361	-0.050337
	$K = 10$	-0.071871	-0.071868	-0.074193	-0.073241
	$K = 20$	-0.10182	-0.10064	-0.1057	-0.10205
$\alpha = 0.99$ $D = 20$	$K = 1$	-0.037131	-0.040931	-0.040967	-0.033464
	$K = 5$	-0.077337	-0.081054	-0.092059	-0.068057
	$K = 10$	-0.11424	-0.12384	-0.13796	-0.10107
	$K = 20$	-0.16448	-0.17575	-0.20172	-0.14304
$\alpha = 0.995$ $D = 20$	$K = 1$	-0.042265	-0.051125	-0.050634	-0.037874
	$K = 5$	-0.086024	-0.099022	-0.10753	-0.075312
	$K = 10$	-0.13394	-0.14886	-0.16967	-0.1123
	$K = 20$	-0.19092	-0.20053	-0.22292	-0.15953

T - A.1 Table: VaR_α simulation results of A - 4.1 and A - 4.2 for $D = 5$, 10, 15, 20

B Appendix to Part II

B.1 Probability Theory

B.1 Definition (compare [35], p. 2)
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A *filtration* or an *information flow* $(\mathcal{F}_t)_{t\in[0,T]}$ is an increasing and right-continuous family of sub-σ-fields of $\mathcal{F}$. If $\mathcal{F}_0$ additionally contains all $\mathbb{P}$-null-sets, $\mathcal{F}_t$ fulfills the *usual conditions*.

B.2 Definition
The *σ-algebra generator* $\sigma(A)$ produces the smallest possible σ-field of the set A or a random variable A. For a stochastic process X_t, $\sigma(X_u, u \leq t)$ generates a filtration.

B.3 Theorem (Conditional Expectation Calculus, compare [60], Theorem 2.3.2)
Assume that every used random variable is integrable. The following statements hold for the conditional expectation:

1. *Taking out what is known*

   If X is $\mathcal{H}$-measurable, it holds

   $$\mathbb{E}[XY|\mathcal{H}] = X\mathbb{E}[Y|\mathcal{H}] \quad .$$

2. *Iterated conditioning or tower law*

   If $\mathcal{G} \subseteq \mathcal{H}$ ($\mathcal{G}$ contains less information than $\mathcal{H}$), we have

   $$\mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}] \quad .$$

   As a special case, it holds $\mathbb{E}[\mathbb{E}[X|\mathcal{H}]] = \mathbb{E}[X]$.

3. *Independence*

   If X is independent of $\mathcal{H}$, it is

   $$\mathbb{E}[X|\mathcal{H}] = \mathbb{E}[X] \quad .$$

For a sketch of proof see [60], pp. 70-73. ∎

For the next statements and their proofs compare [3], §5, Definition 10.1 and Korollar 12.2.

B.4 Definition (Convergence of Random Variables and Distributions)
Let X_n, $n \in \mathbb{N}$ be a sequence of random variables and $\mathbb{P}_{X_n}$, $n \in \mathbb{N}$, their corresponding distributions (i.e. probability measures).

1. X_n *converges $\mathbb{P}-$almost surely* to a random variable X if it holds for all $\epsilon > 0$

   $$\mathbb{P}\left[\limsup_{n\to\infty} (|X_n - X| > \epsilon)\right] = 0 \quad . \tag{B.1}$$

2. X_n *converges in probability* to a random variable X if it holds for $\epsilon > 0$

   $$\lim_{n\to\infty} \mathbb{P}\left[(|X_n - X| > \epsilon)\right] = 0 \quad . \tag{B.2}$$

3. $\mathbb{P}_{X_n}$ *converges in distribution* to $\mathbb{P}_X$ if it holds for all continuous functions $f \in \mathbb{R}$

   $$\lim_{n\to\infty} \int f d\mathbb{P}_{X_n} = \int f d\mathbb{P}_X \quad .$$

B.5 Definition (Strong and Weak Law of Large Numbers)

1. A sequence of integrable random variables X_n, $n \in \mathbb{N}$, obeys the *strong law of large numbers* if it holds
$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} (X_j - \mathbb{E}[X_j]) = 0$$
in sense of Equation (B.1).
2. A sequence of integrable random variables X_n, $n \in \mathbb{N}$, obeys the *weak law of large numbers* if it holds
$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} (X_j - \mathbb{E}[X_j]) = 0$$
in sense of Equation (B.2).

B.6 Theorem (Kolmogorov's Law of Large Numbers)
Any sequence of real, integrable and i.i.d. random variables obeys the strong law of large numbers, i.e., for i.i.d. X_n, $n \in \mathbb{N}$ with $\mathbb{E}[X_1] = \mu$, it is

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} (X_j - \mathbb{E}[X_j]) = 0 \quad \Leftrightarrow \quad \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} X_j = \mu \quad .$$

B.7 Notation (Central Limit Theorem)
Let $(X_n)_{n\in\mathbb{N}}$, a independent sequence of real, square-integrable ($\mathbb{E}[X_n^2] < \infty$) random variables. The *Central Limit Theorem* holds if standardized sum variables S_n converges in distribution to the standard normal distribution $N(0,1)$, i.e., it holds

$$S_N := \frac{\sum_{n=1}^{N} (X_n - \mathbb{E}[X_n])}{\sqrt{\mathbb{V}[X_1 + \ldots + X_N]}} \quad \to \quad N(0,1)$$

for $N \to \infty$. For $(X_n)_{n\in\mathbb{N}}$ i.i.d., real-valued and square-integrable random variables, we have

$$\tilde{S}_N := \frac{\sum_{n=1}^{N} X_n - N \cdot \mathbb{E}[X_1]}{\sqrt{N} \cdot \sqrt{\mathbb{V}[X_1]}} \quad \to \quad N(0,1) \quad \text{for } N \to \infty \quad .$$

B.2 Stochastic Analysis

B.8 Definition (compare [35], Definition 1.20)
A stochastic process X_t is $\mathcal{F}_t$-*adapted* if X_t is $\mathcal{F}_t$-measurable for all $t \in [0,T]$. In literature, one also finds the expression *progressively measurable* as notation for $\mathcal{F}_t$-adaptivity.

B.9 Lemma
Under certain regularity conditions, see [59], Section 2.3, one can derive the intensity as

$$\lambda_t = -\frac{\partial}{\partial} |_{T=t} \mathbb{Q}\left[\tau > T | \mathcal{G}\right]$$

if no default has occurred up to time T and $\mathbb{Q}\left[\tau > T|\mathcal{G}\right]$ is differentiable from the right with respect to T at $T = t$.

B.3 Asset Pricing Theory

B.10 Definition (Arbitrage)
Compare [60], Definition 5.4.6. An *arbitrage opportunity* is a stochastic process X_t in a market model satisfying $X_0 = 0$ and for some time $T > 0$

$$\mathbb{P}[X_T \leq 0] = 1 \quad \text{and} \quad \mathbb{P}[X_T > 0] > 0 \quad .$$

B.11 Theorem (First Fundamental Theorem of Asset Pricing)
Compare [60], Theorem 5.4.7. If a market model has a risk-neutral probability measure, then it does not admit arbitrage.

A proof can be found in Shreve [60], p. 231, for instance. ∎

B.12 Theorem (Second Fundamental Theorem of Asset Pricing)
Compare [60], Theorem 5.4.7. A market model is complete, i.e., every security can be hedged if and only if it has a unique risk-neutral probability measure.

Shreve sketches a proof in [60] on pp. 232-233. ∎

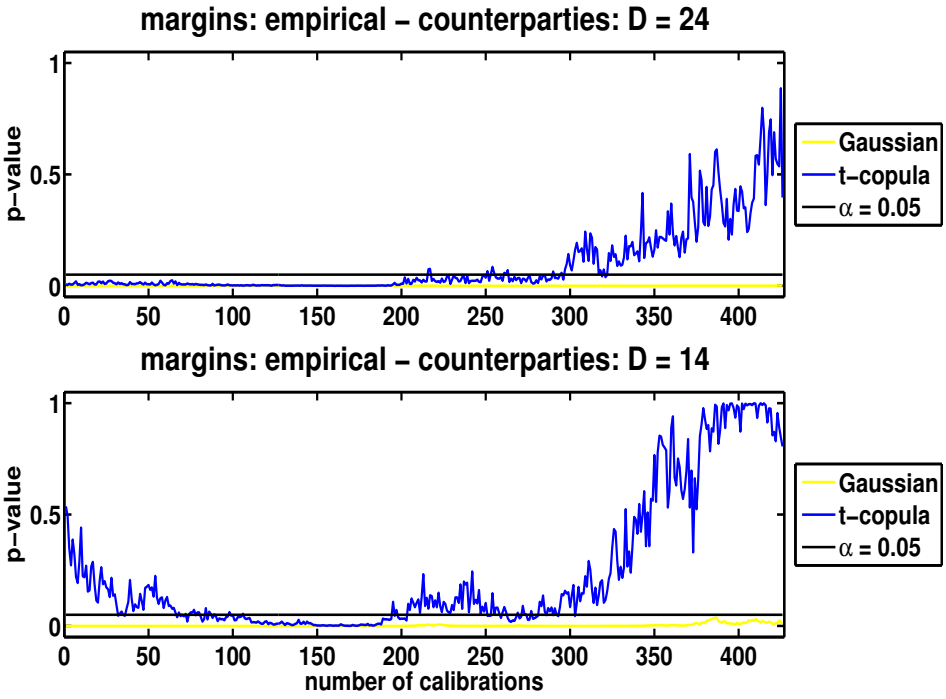
B.4 Portfolio Composition - Part II

One-year Credit Spreads Data Basis, $s(0, 1y)$

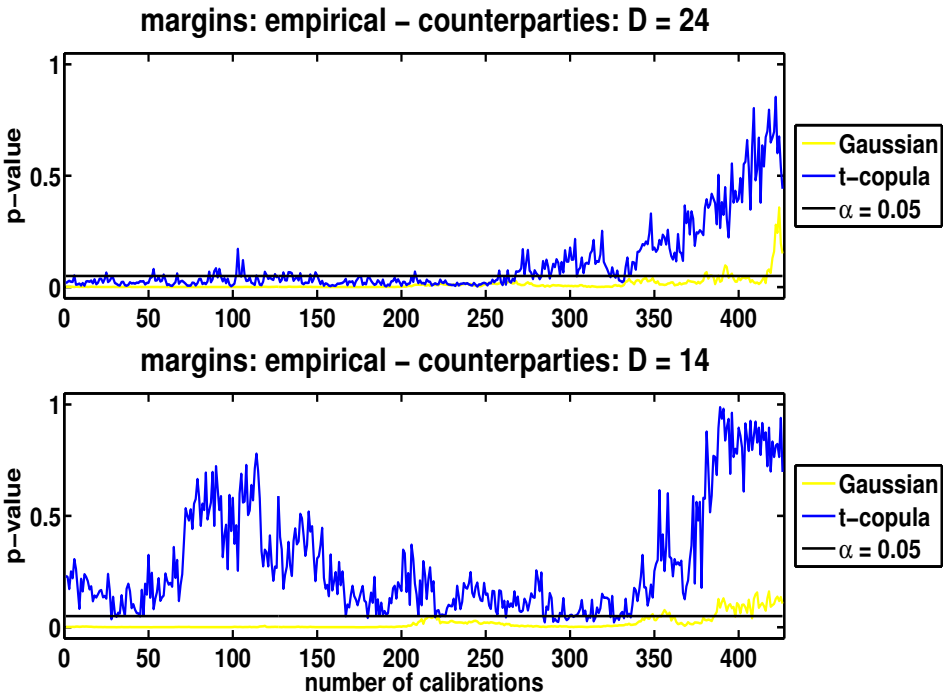
Hilton Hotels Corp.	Citigroup Inc. NY	Fed. Republic of Brazil
Ryder System Inc.	Univision Communication	Thailand
Kroger Company	Alliance Boots PLC	Rabobank Nederland Utrecht
R.R. Donnelley + So	ITV PLC London	Dresdner Bank AG
Mohawk Industries	Bae Systems PLC	Barclays Bank Plc
Conagra Foods Inc.	QBE INSURANCE Group Sydney	Daimler AG
Heinz Company.	Countrywide Home Loans Inc.	Republic of Portugal
First Data Corp.	Credit Agricole Paris	United Mexican States

Credit spreads are extracted from Bloomberg, quotations from November 28, 2007 till July 27, 2011. Above given portfolio selection is chosen w.r.t. data availability and consistency.

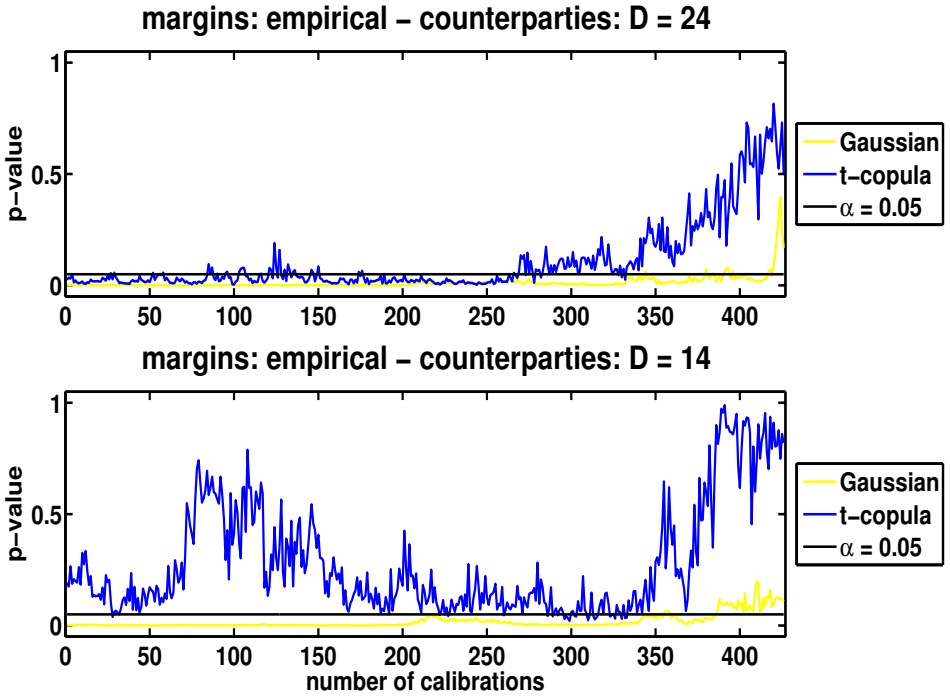
B.5 Further Numerical Results - Part II



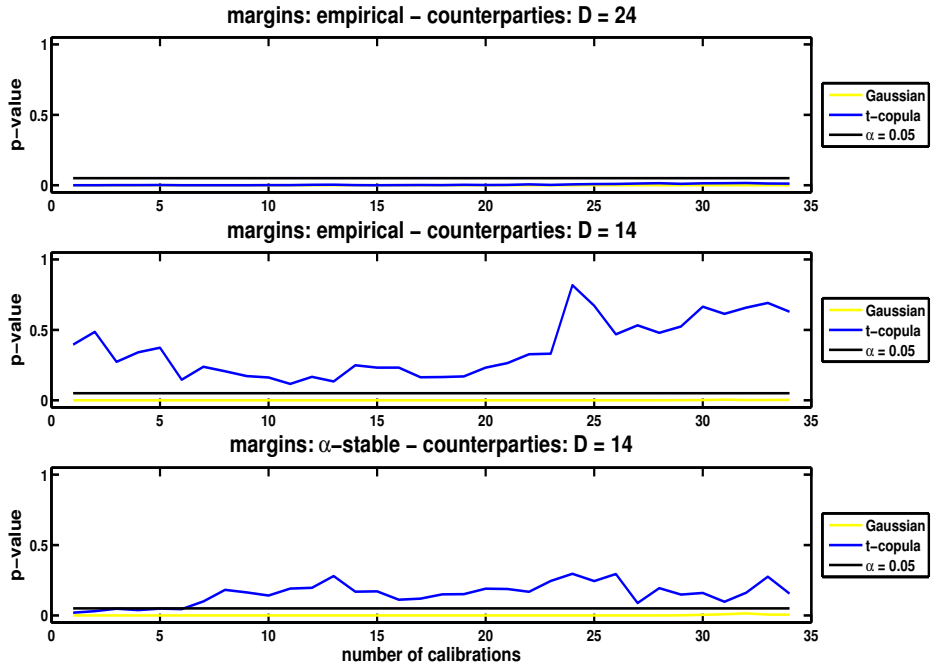
F - B.1 Figure: goodness-of-fit test - $K = 1$ and *abs*, complete period



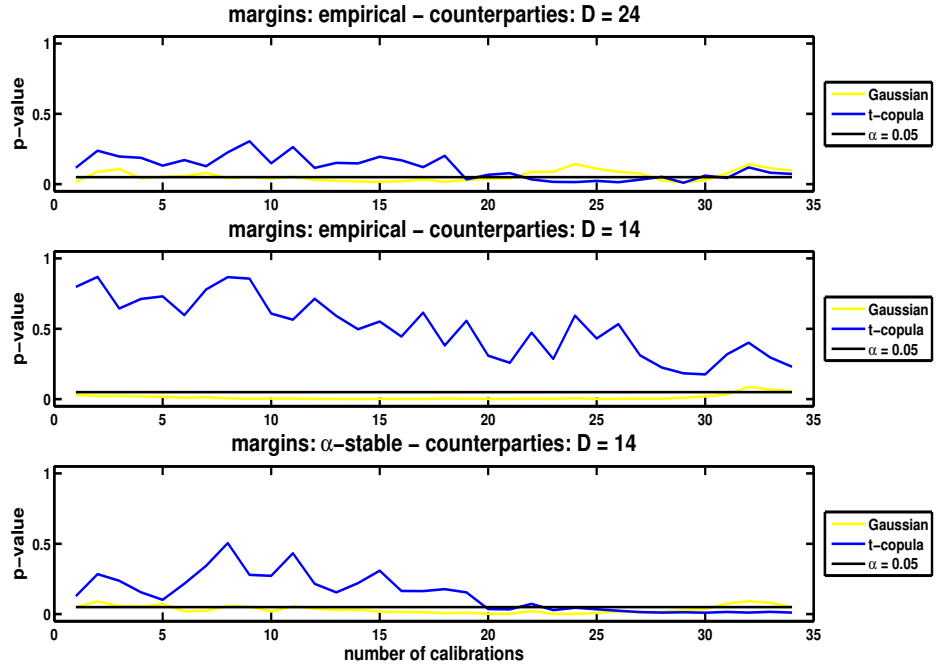
F - B.2 Figure: goodness-of-fit test - $K = 1$ and *rel*, complete period



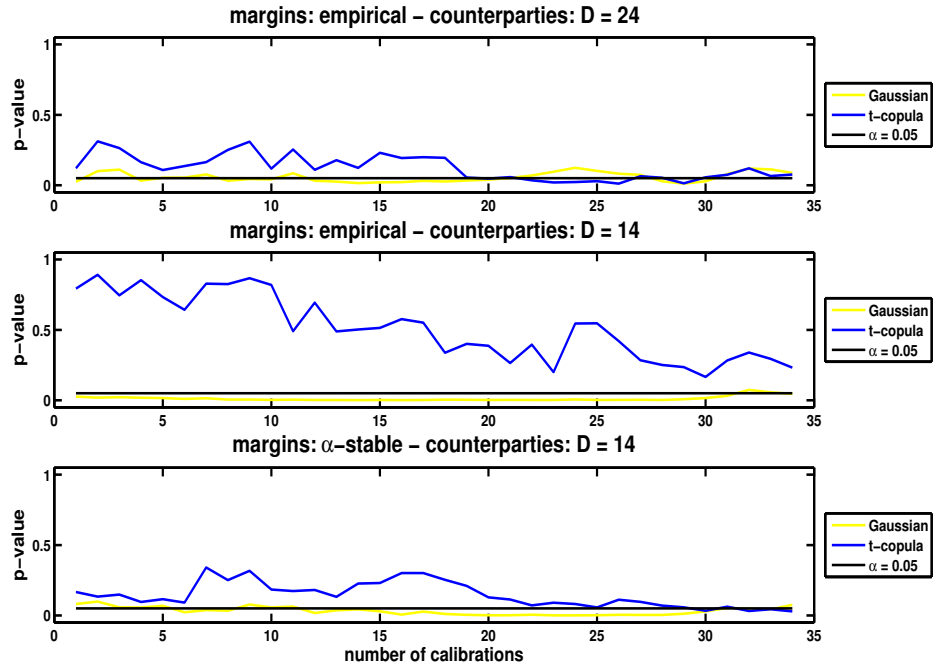
F - B.3 Figure: goodness-of-fit test - $K = 1$ and *log*, complete period



F - B.4 Figure: goodness-of-fit test - $K = 5$ and *abs*



F - B.5 Figure: goodness-of-fit test - $K = 5$ and *rel*



F - B.6 Figure: goodness-of-fit test - $K = 5$ and *log*

	Gaussian copula			***t*-copula**		
cpty no.	$M = 5000$	$M = 10000$	$M = 20000$	$M = 5000$	$M = 10000$	$M = 20000$
1	1.080	1.370	1.300	2.080	1.880	1.890
2	0.560	0.550	0.535	1.360	1.260	1.245
3	7.700	8.220	8.245	8.240	8.340	8.005
4	0.080	0.010	0.065	0.260	0.170	0.125
5	1.480	1.350	1.520	4.220	4.750	4.665
6	2.020	2.210	2.145	2.620	2.800	2.740
7	7.520	7.460	7.365	8.700	8.360	8.635
8	1.540	1.740	1.890	2.660	2.460	2.450
9	1.040	0.770	0.745	4.420	4.430	4.545
10	1.420	1.820	1.895	4.440	4.700	4.625
11	0.740	0.790	0.775	3.420	3.120	3.070
12	0.240	0.290	0.300	1.000	1.090	0.945
13	1.480	1.490	1.510	4.860	4.610	4.475
14	0.620	0.710	0.750	1.100	1.020	1.115
15	2.380	2.390	2.495	3.680	3.970	3.860
16	1.100	0.880	1.030	1.820	1.480	1.505
17	0.740	0.690	0.720	1.280	1.120	1.160
18	1.600	1.580	1.610	3.440	3.210	3.360
19	1.000	1.230	1.245	2.420	2.250	2.480
20	0.540	0.530	0.620	2.120	2.010	1.890
21	2.040	2.220	2.340	3.460	3.630	3.695
22	0.760	0.560	0.675	2.480	2.330	2.375
23	0.800	0.810	0.905	1.580	1.670	1.560
24	1.940	1.840	2.050	8.680	8.940	9.090

T - B.1 Table: simulated default rates for $K = 1$ and *rel* - in percent

	Gaussian copula			***t*-copula**		
cpty no.	$M = 5000$	$M = 10000$	$M = 20000$	$M = 5000$	$M = 10000$	$M = 20000$
1	1.600	1.540	1.710	2.640	2.450	2.355
2	0.800	0.820	0.825	1.920	2.000	1.900
3	11.200	11.520	10.880	10.500	10.860	10.945
4	1.480	1.200	1.225	4.220	5.050	4.495
5	2.500	2.510	2.495	8.220	8.500	8.305
6	2.720	3.030	2.950	3.720	3.810	3.830
7	9.180	9.230	9.615	10.340	10.520	10.480
8	2.260	2.210	2.275	3.160	3.320	3.320
9	1.160	1.310	1.295	7.180	7.200	7.335
10	2.400	2.530	2.600	6.540	6.360	6.675
11	0.880	1.110	1.030	3.960	4.130	4.075
12	0.260	0.380	0.335	1.340	1.400	1.480
13	2.220	2.370	2.335	6.480	6.770	6.400
14	0.760	0.800	0.815	1.320	1.340	1.385
15	3.020	3.130	3.240	5.320	5.080	5.270
16	1.420	0.970	1.210	2.060	1.820	1.910
17	0.900	0.710	0.810	1.380	1.250	1.390
18	1.940	1.820	1.945	4.740	4.590	4.695
19	1.480	1.550	1.545	3.100	3.160	3.305
20	0.780	0.840	0.805	3.160	3.160	3.185
21	2.620	2.840	2.730	4.900	4.590	4.805
22	1.000	0.800	0.925	3.560	3.370	3.470
23	0.940	0.850	0.975	1.640	1.750	1.895
24	2.840	2.400	2.600	12.560	12.050	12.565

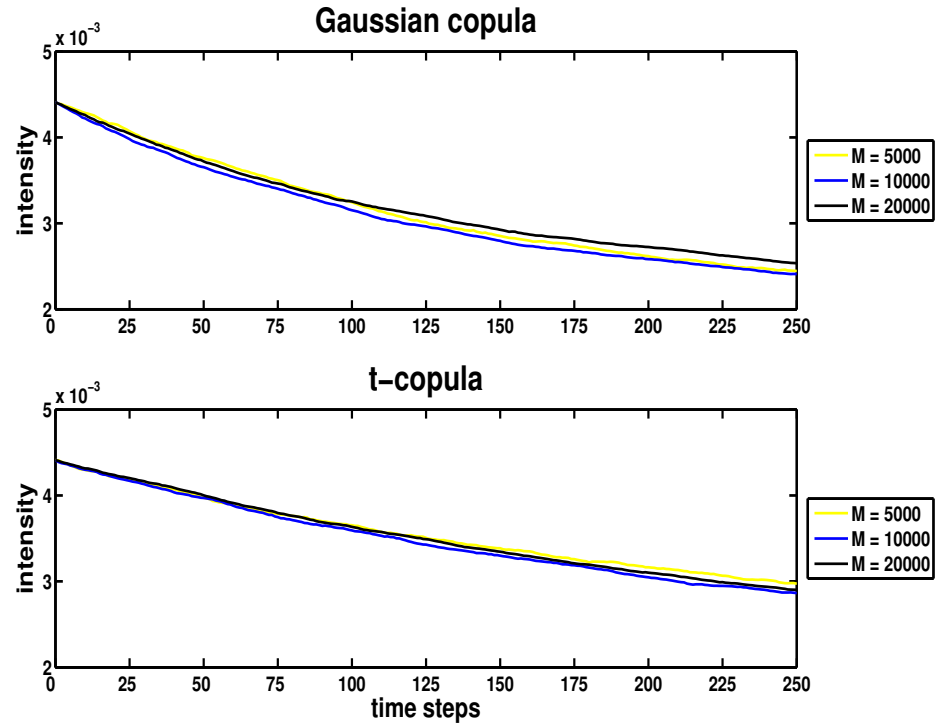
T - B.2 Table: simulated default rates for $K = 1$ and *log* - in percent

	Gaussian copula			***t*-copula**		
cpty no.	$M = 5000$	$M = 10000$	$M = 20000$	$M = 5000$	$M = 10000$	$M = 20000$
1	3.140	2.560	2.760	3.160	2.690	2.665
2	0.720	0.780	0.850	1.040	1.110	0.965
3	2.440	2.360	2.245	2.560	2.940	2.315
4	0.160	0.130	0.120	0.340	0.270	0.245
5	2.080	2.330	2.085	2.460	2.550	2.380
6	0.420	0.470	0.540	0.680	0.630	0.745
7	7.660	7.870	7.335	9.800	10.260	9.575
8	2.420	2.430	2.410	2.200	2.630	2.270
9	0.820	0.640	0.785	1.180	1.260	1.060
10	1.400	1.320	1.410	2.040	1.840	1.755
11	0.680	0.680	0.660	0.980	0.980	0.885
12	0.380	0.260	0.460	0.980	0.950	1.025
13	0.900	0.810	0.770	1.380	1.370	1.180
14	0.980	0.740	0.835	1.980	1.830	1.605
15	2.440	2.360	2.455	3.520	3.770	3.540
16	1.360	1.290	1.390	1.800	1.690	1.940
17	1.000	1.050	1.055	1.720	1.960	1.755
18	3.460	3.490	3.790	6.520	6.240	6.225
19	1.540	1.640	1.705	2.160	2.330	1.995
20	0.980	0.900	0.800	1.380	1.470	1.170
21	1.520	1.850	1.680	2.120	2.500	2.225
22	0.700	0.840	0.825	1.040	1.140	1.005
23	1.680	1.410	1.605	2.400	2.410	2.335
24	1.600	1.650	1.495	2.340	2.360	2.055

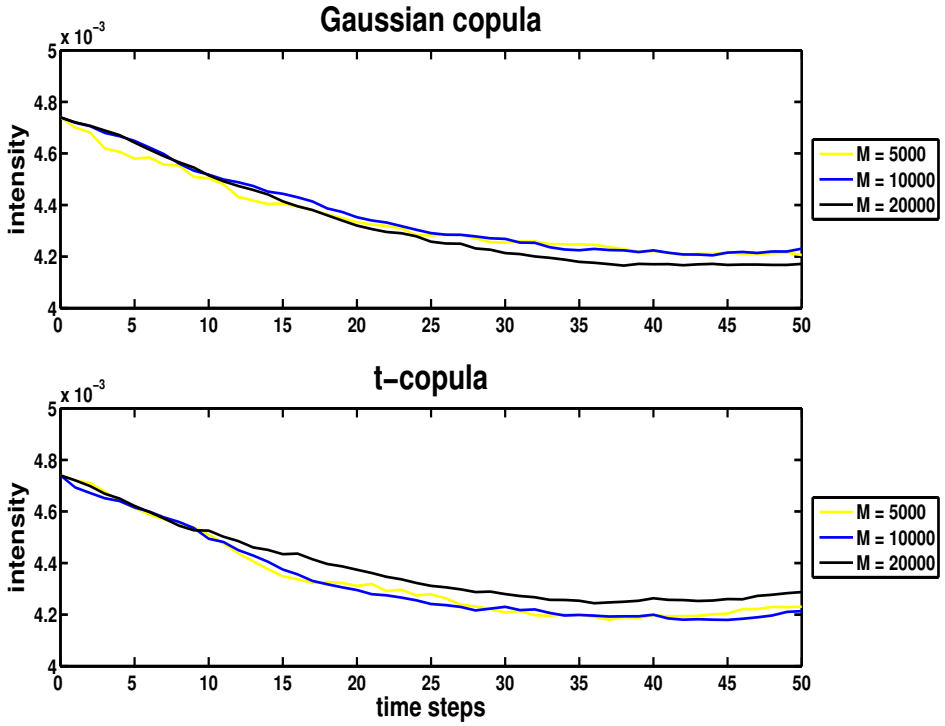
T - B.3 Table: simulated default rates for $K = 5$ and *rel* - in percent

	Gaussian copula			***t*-copula**		
cpty no.	$M = 5000$	$M = 10000$	$M = 20000$	$M = 5000$	$M = 10000$	$M = 20000$
1	4.200	4.290	4.085	4.820	4.120	4.215
2	1.140	1.060	1.110	1.740	1.720	1.680
3	2.980	3.080	3.165	3.720	3.750	3.670
4	0.180	0.160	0.150	0.660	0.730	0.595
5	8.040	7.800	7.970	8.960	9.050	8.870
6	1.000	1.010	1.050	2.260	1.930	2.145
7	10.160	10.130	9.910	13.980	14.320	14.210
8	6.080	7.360	7.100	7.540	7.600	7.485
9	1.000	0.850	0.985	2.360	2.000	2.100
10	1.980	1.560	1.690	3.160	2.570	2.665
11	0.860	0.890	0.895	1.660	1.400	1.655
12	0.580	0.570	0.770	2.100	1.860	2.095
13	1.060	1.020	1.070	2.780	2.090	2.175
14	1.460	1.190	1.260	3.460	2.910	3.095
15	3.500	3.290	3.425	6.600	5.600	5.980
16	2.840	2.370	2.865	5.980	4.970	5.555
17	1.880	1.890	1.795	5.240	4.430	4.730
18	4.200	4.280	4.615	8.980	8.320	8.910
19	1.980	2.110	2.195	3.360	3.160	3.120
20	1.740	1.590	1.510	3.380	3.070	3.210
21	2.220	2.530	2.620	4.100	3.580	4.035
22	1.540	1.750	1.630	4.780	4.660	4.820
23	2.660	2.520	2.840	7.460	7.010	7.360
24	3.000	3.040	3.250	8.680	8.290	8.295

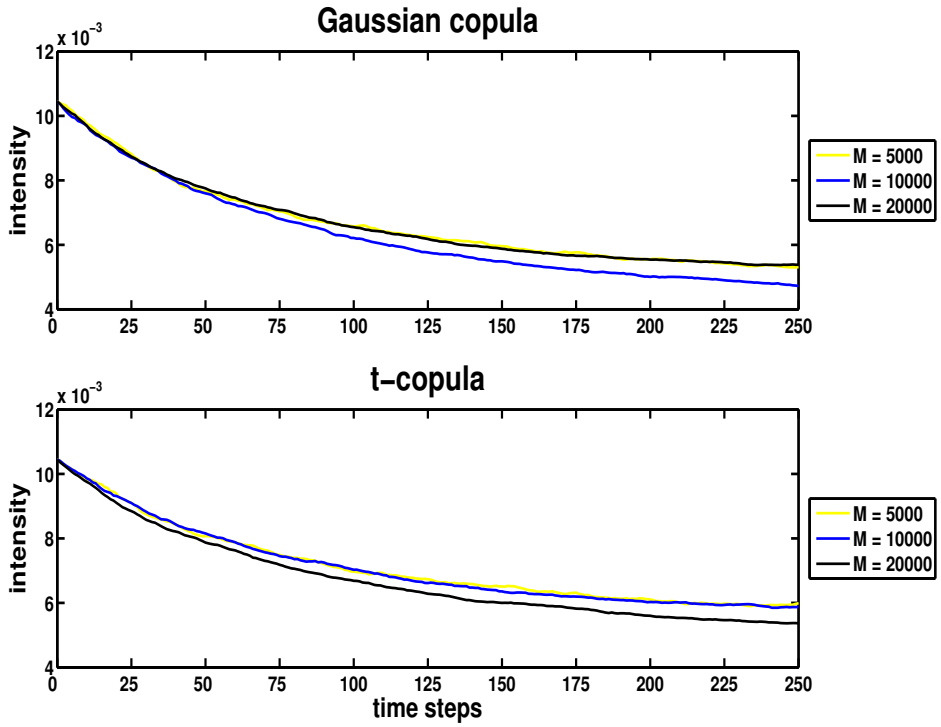
T - B.4 Table: simulated default rates for $K = 5$ and *log* - in percent



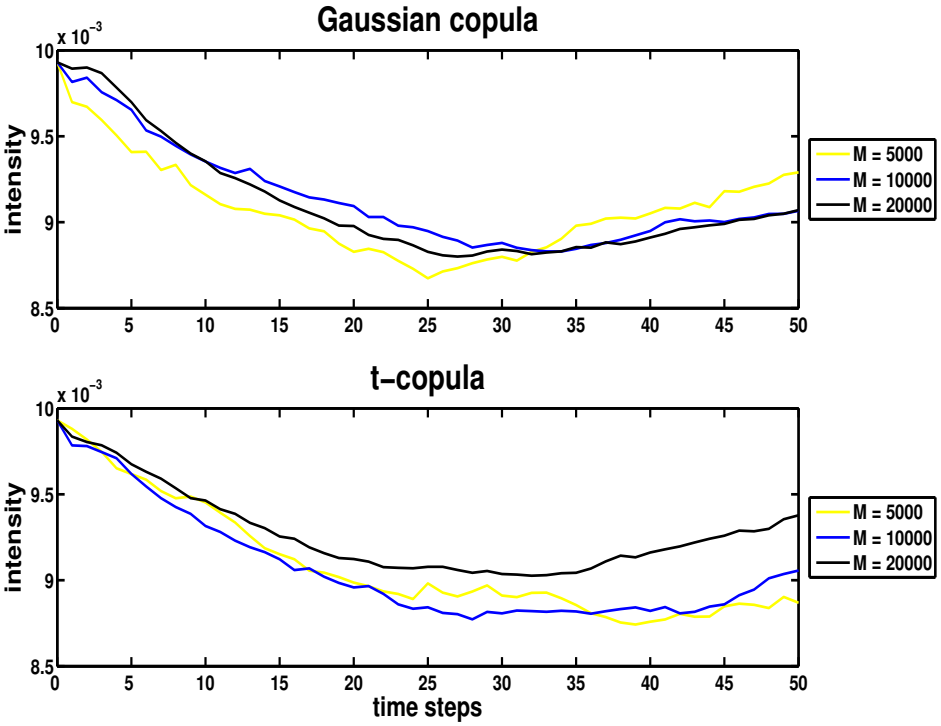
F - B.7 Figure: simulated intensity for counterparty $d = 2$ - $K = 1$ and *abs*



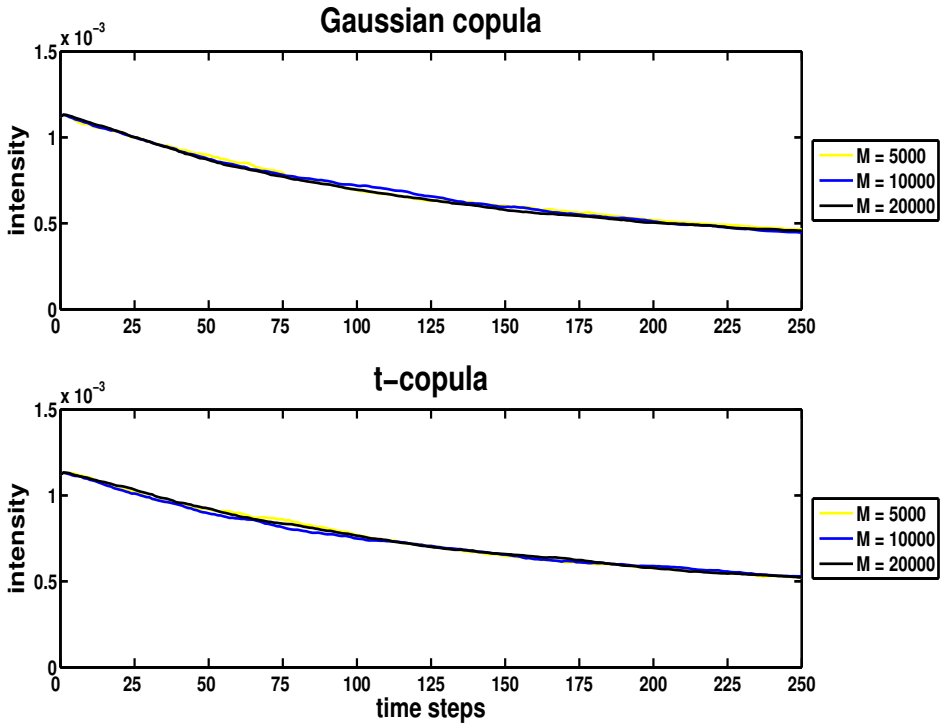
F - B.8 Figure: simulated intensity for counterparty $d = 2$ - $K = 5$ and *abs*



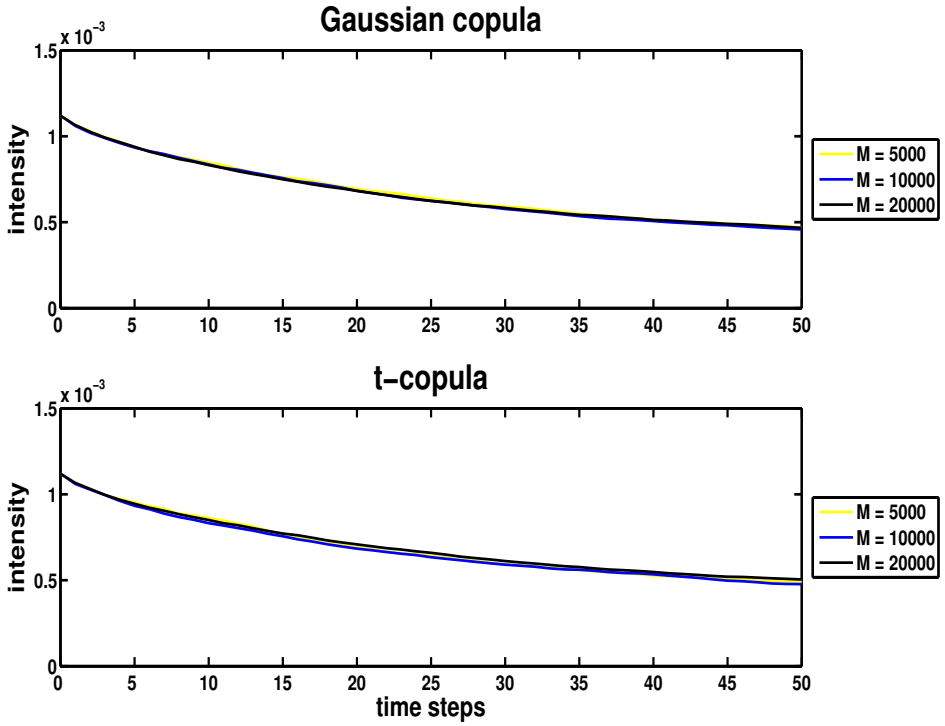
F - B.9 Figure: simulated intensity for counterparty $d = 3$ - $K = 1$ and *abs*



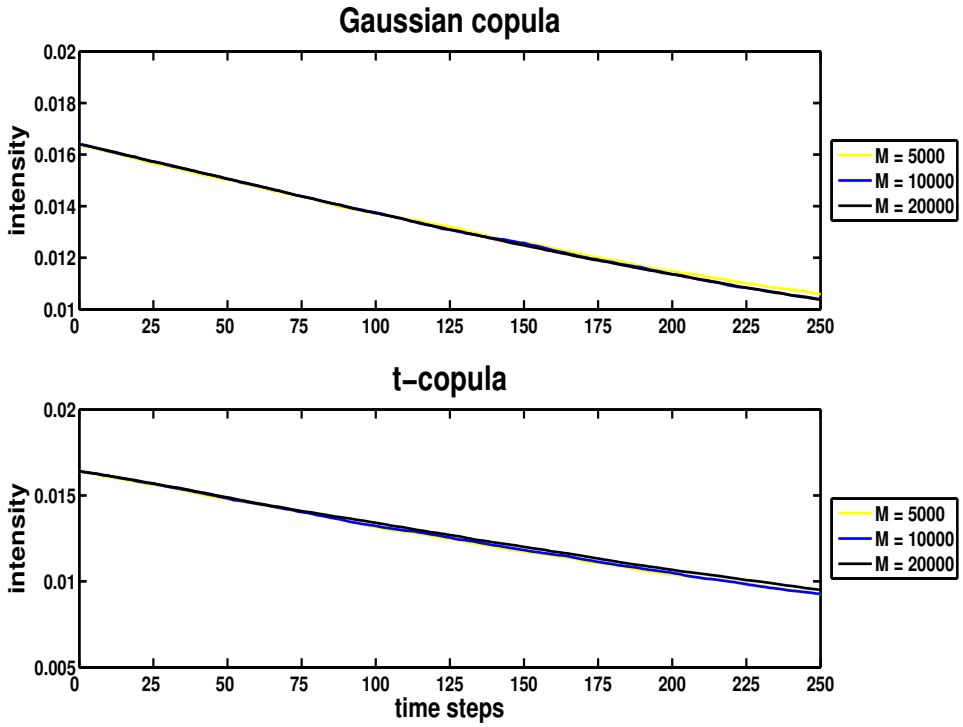
F - B.10 Figure: simulated intensity for counterparty $d = 3$ - $K = 5$ and *abs*



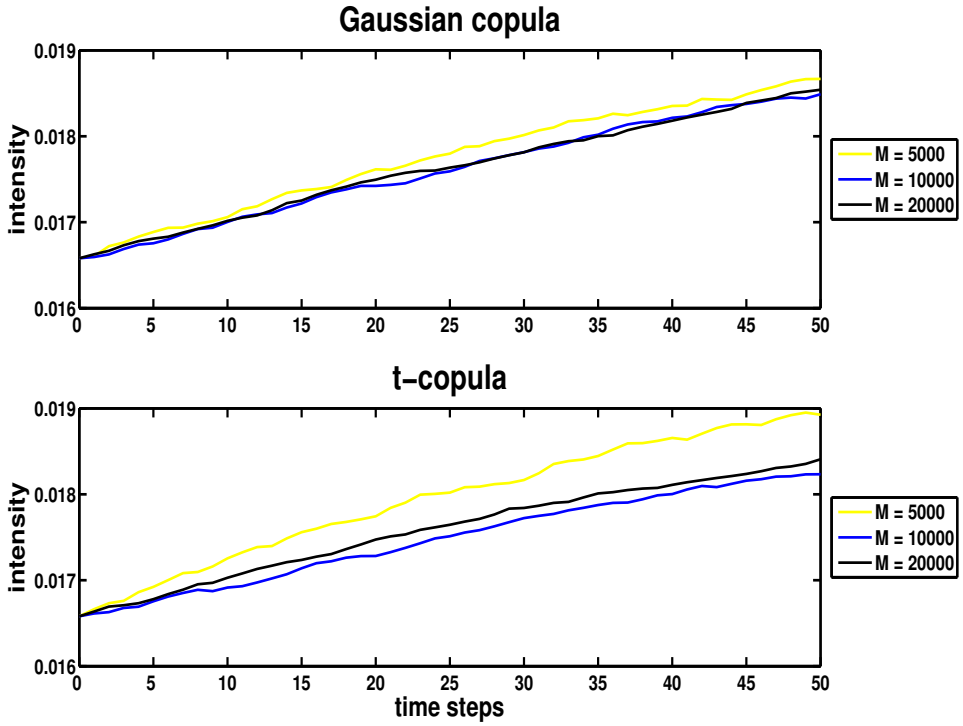
F - B.11 Figure: simulated intensity for counterparty $d = 4$ - $K = 1$ and *abs*



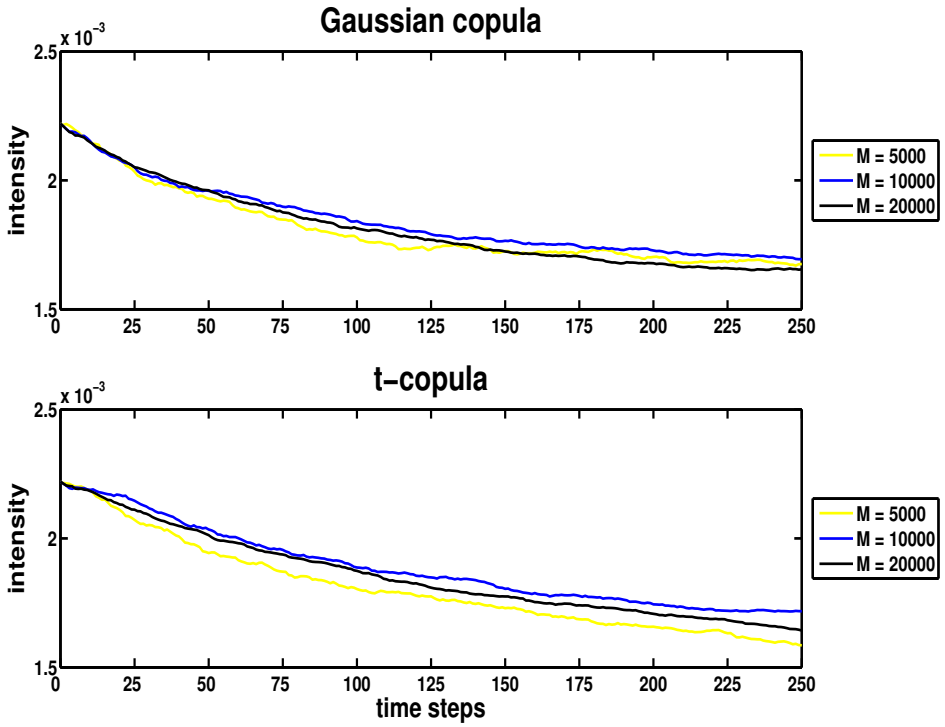
F - B.12 Figure: simulated intensity for counterparty $d = 4$ - $K = 5$ and *abs*



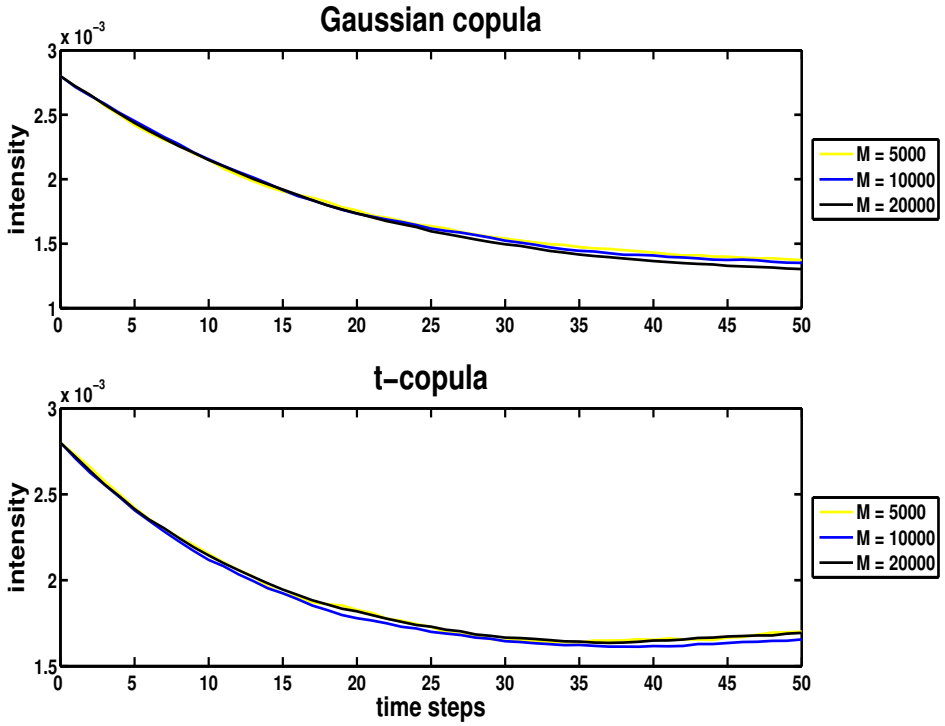
F - B.13 Figure: simulated intensity for counterparty $d = 5$ - $K = 1$ and *abs*



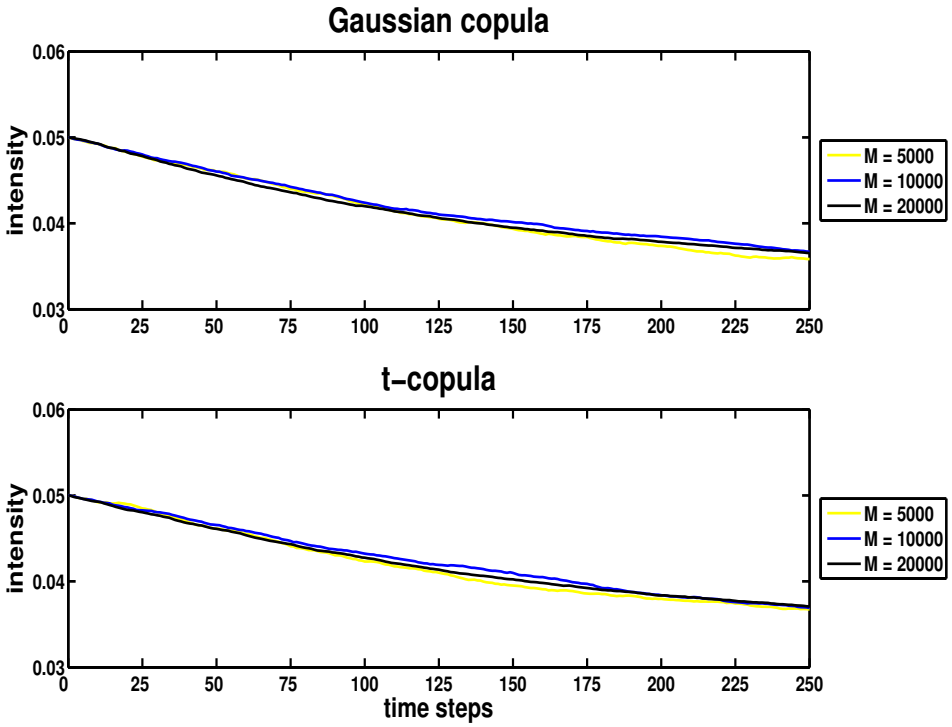
F - B.14 Figure: simulated intensity for counterparty $d = 5$ - $K = 5$ and *abs*



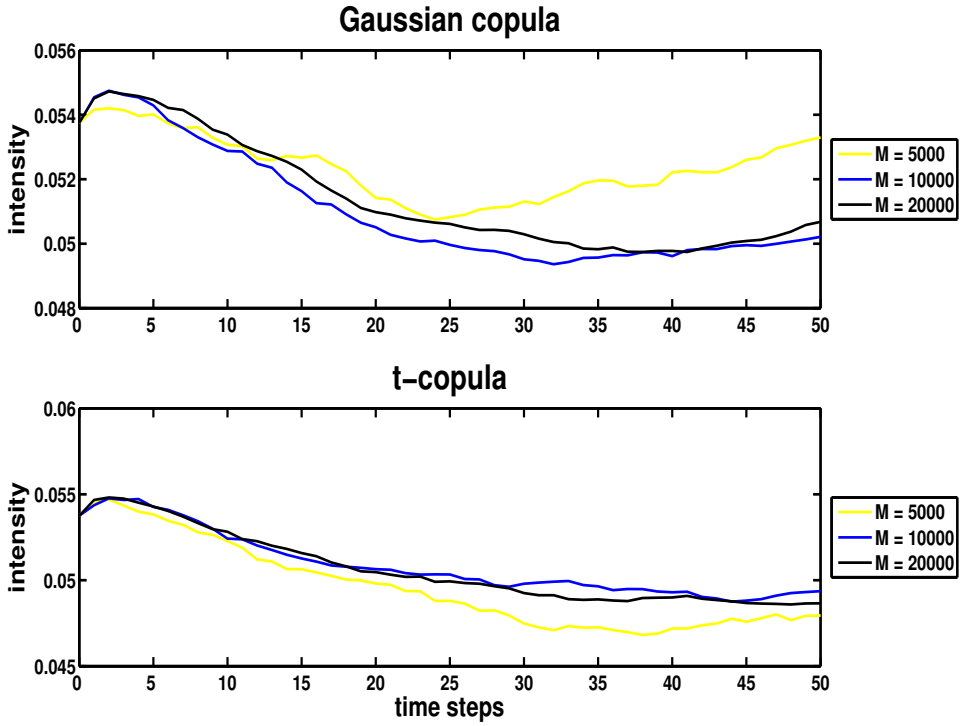
F - B.15 Figure: simulated intensity for counterparty $d = 6$ - $K = 1$ and *abs*



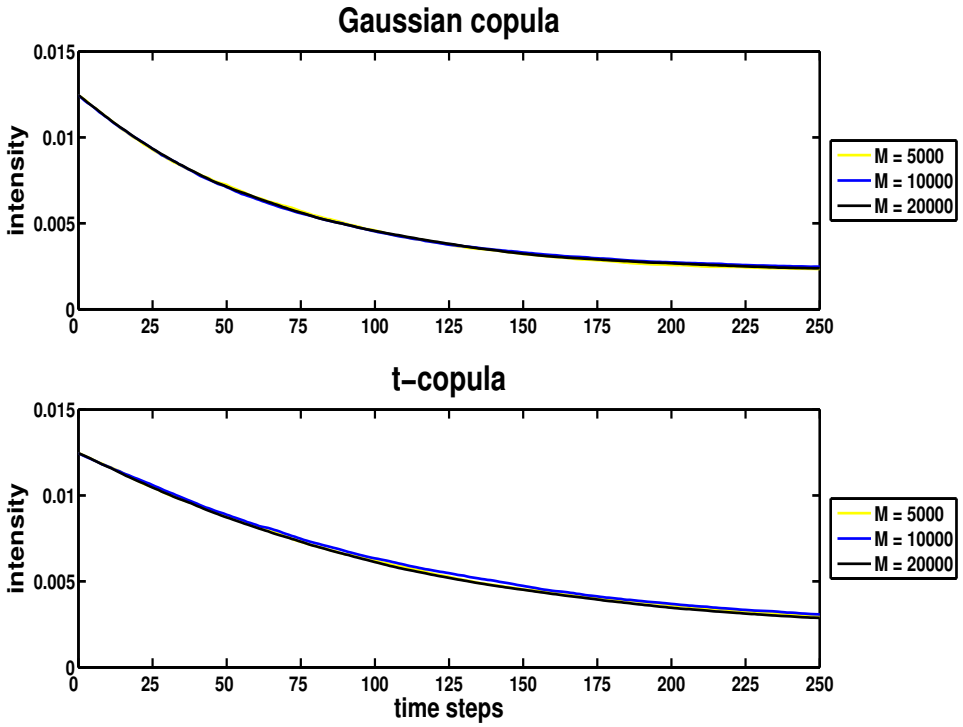
F - B.16 Figure: simulated intensity for counterparty $d = 6$ - $K = 5$ and *abs*



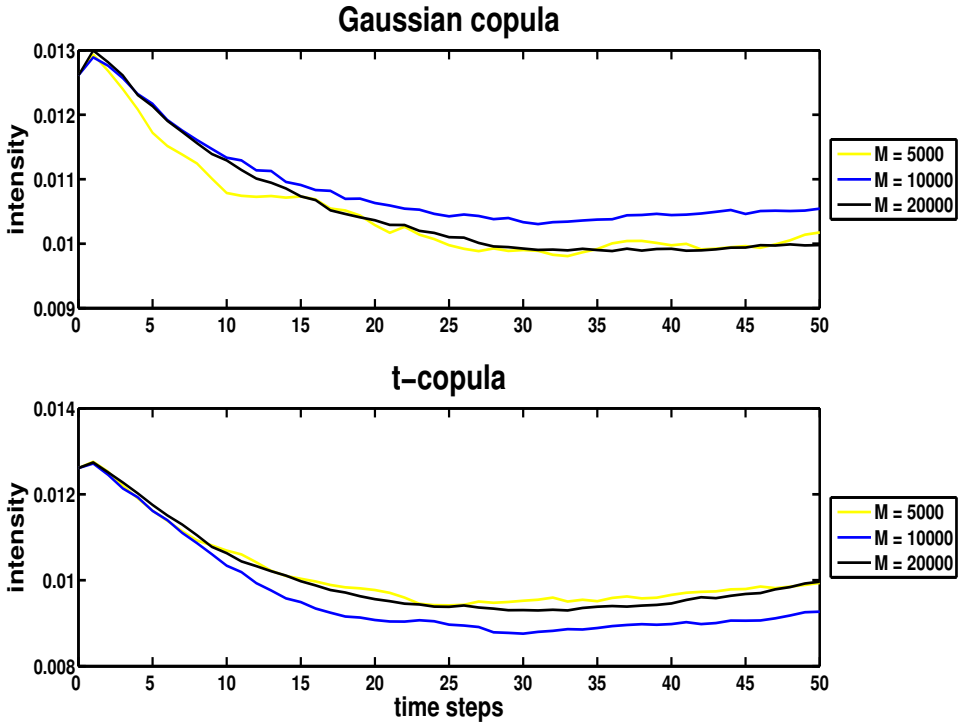
F - B.17 Figure: simulated intensity for counterparty $d = 7$ - $K = 1$ and *abs*



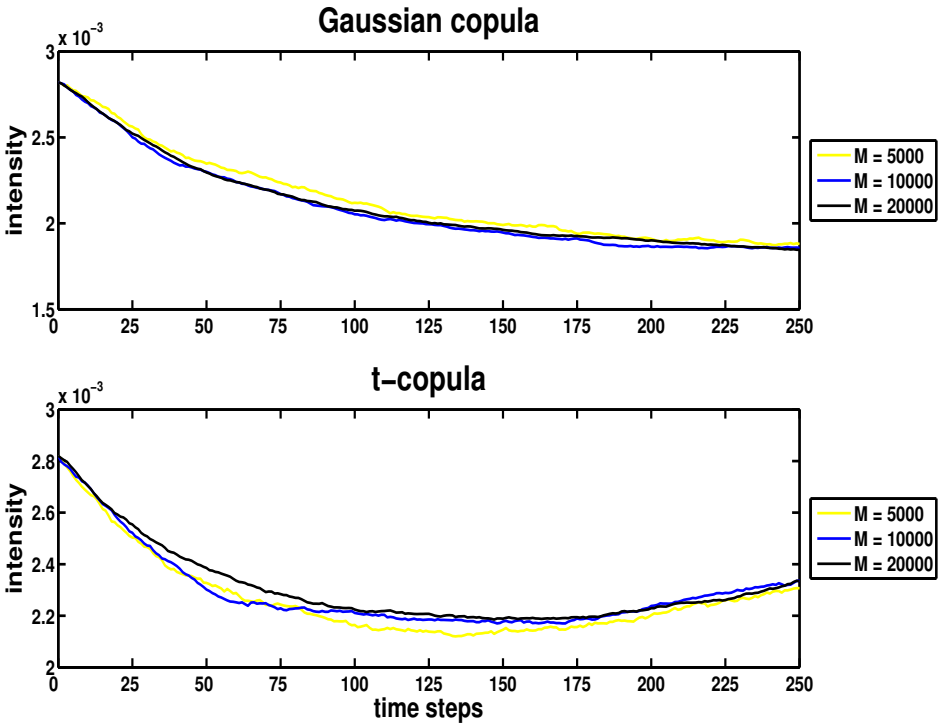
F - B.18 Figure: simulated intensity for counterparty $d = 7$ - $K = 5$ and *abs*



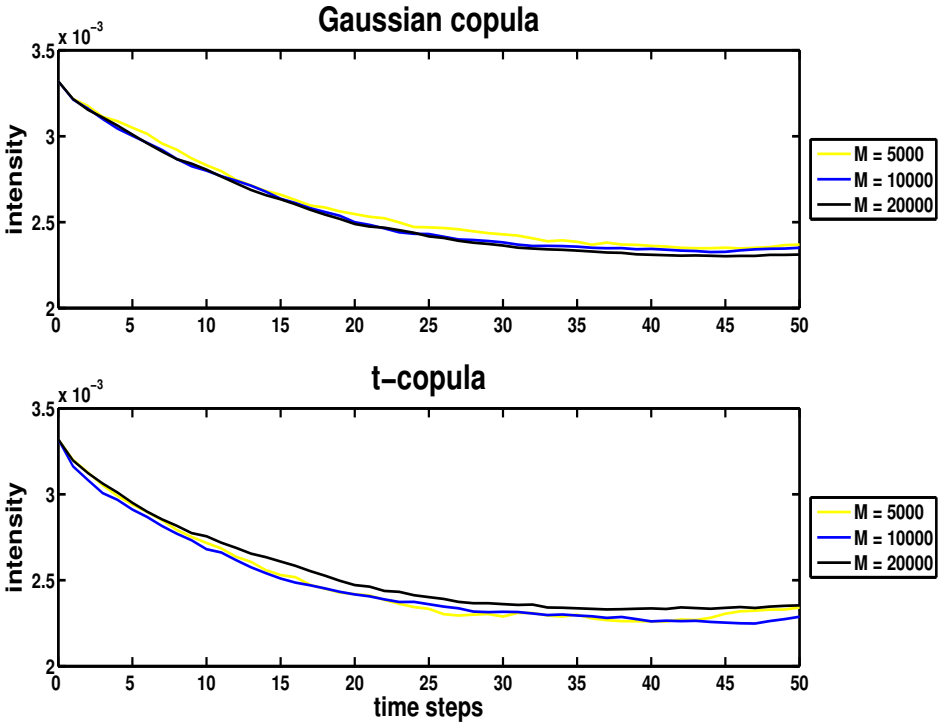
F - B.19 Figure: simulated intensity for counterparty $d = 8$ - $K = 1$ and *abs*



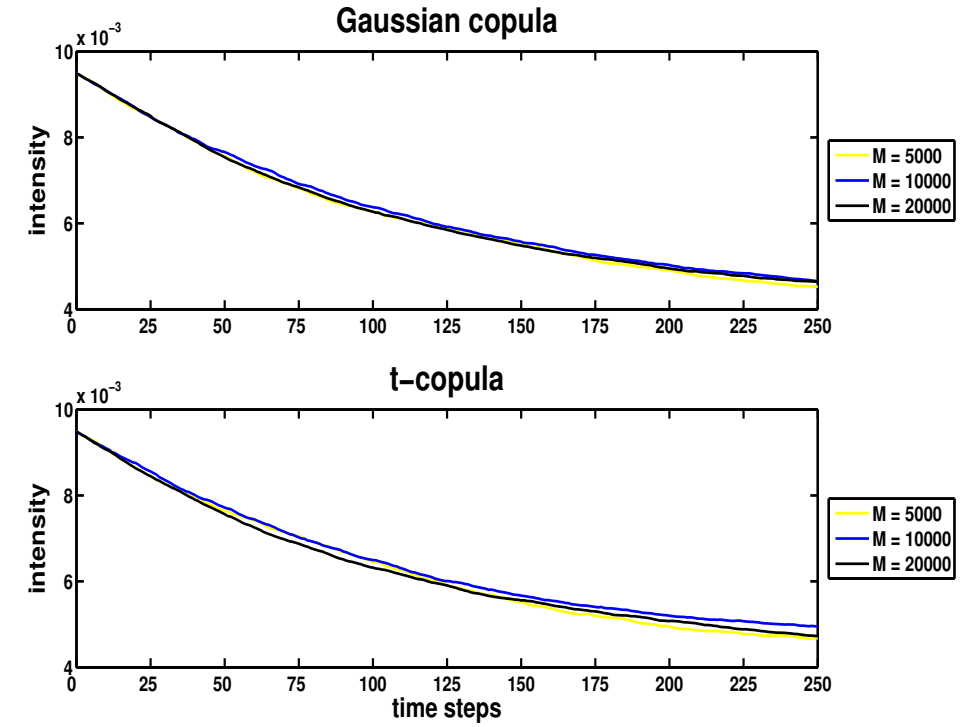
F - B.20 Figure: simulated intensity for counterparty $d = 8$ - $K = 5$ and *abs*



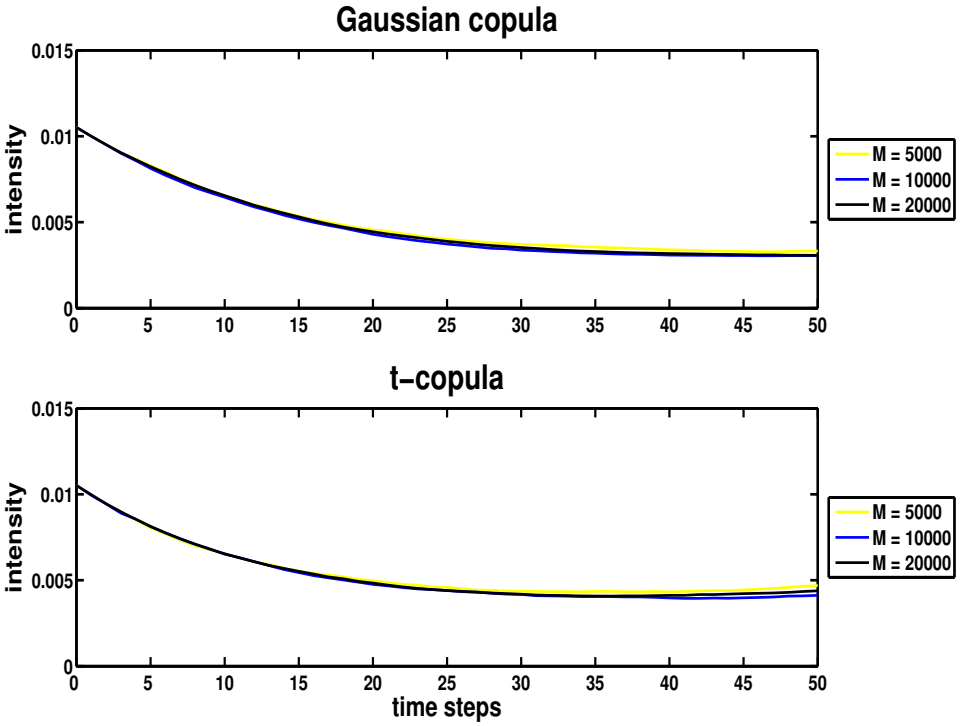
F - B.21 Figure: simulated intensity for counterparty $d = 9$ - $K = 1$ and *abs*



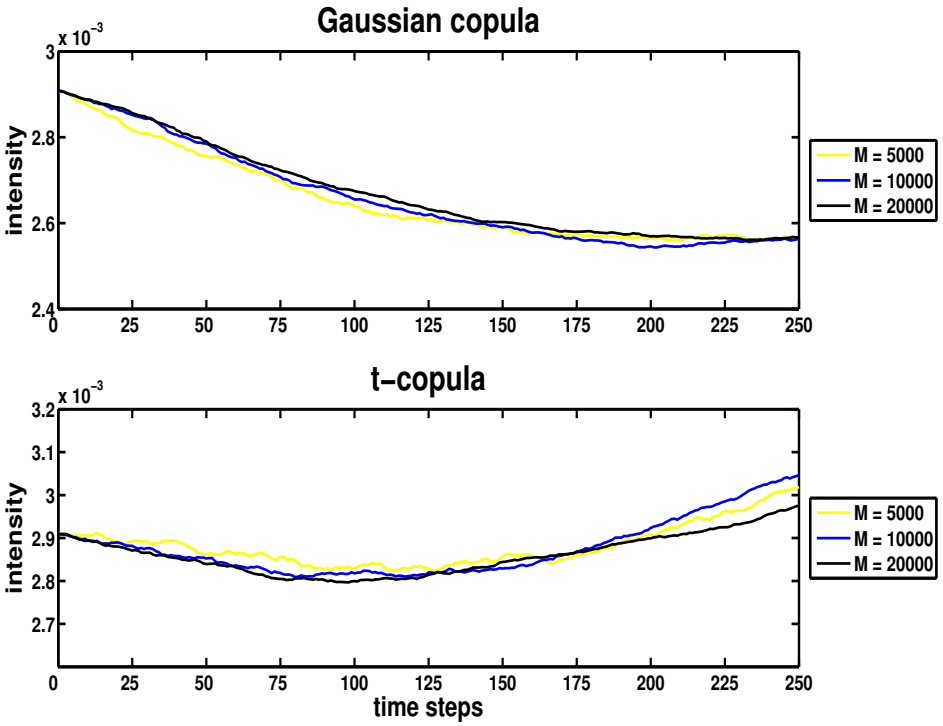
F - B.22 Figure: simulated intensity for counterparty $d = 9$ - $K = 5$ and *abs*



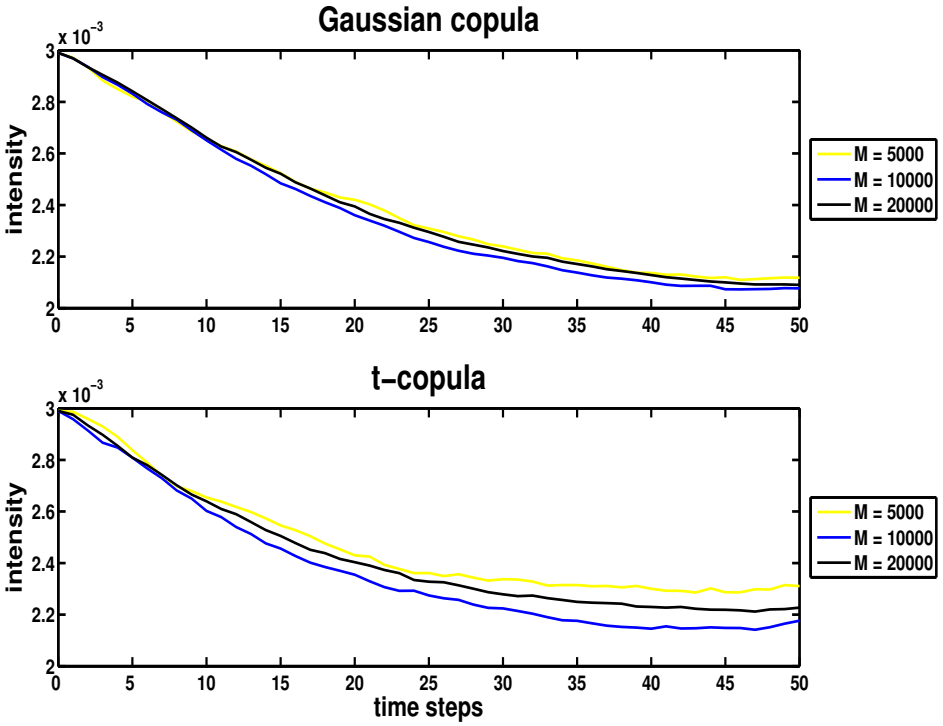
F - B.23 Figure: simulated intensity for counterparty $d = 10$ - $K = 1$ and *abs*



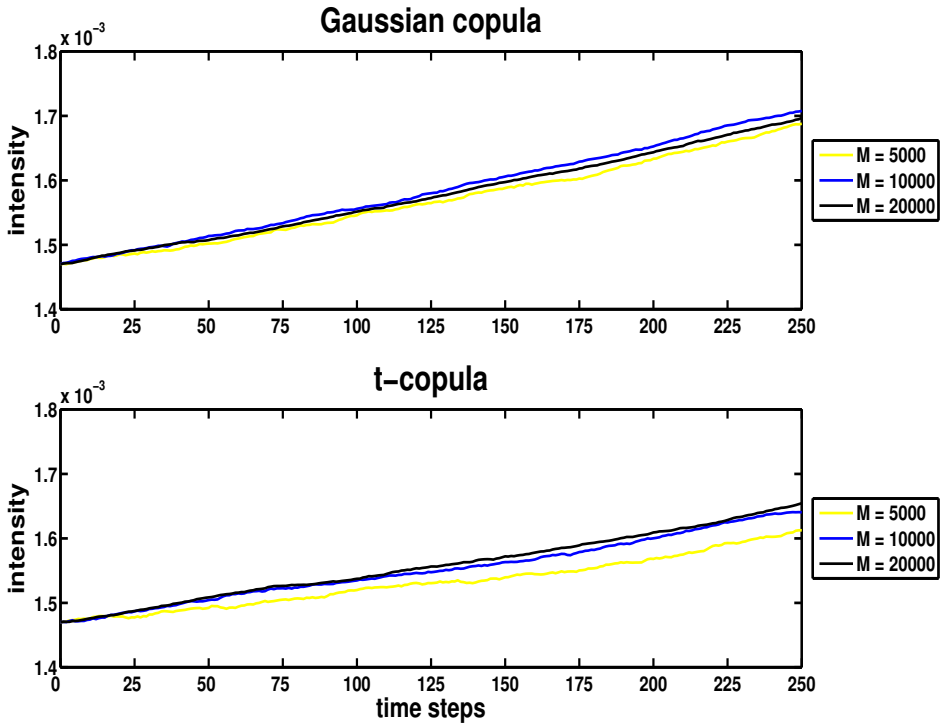
F - B.24 Figure: simulated intensity for counterparty $d = 10$ - $K = 5$ and *abs*



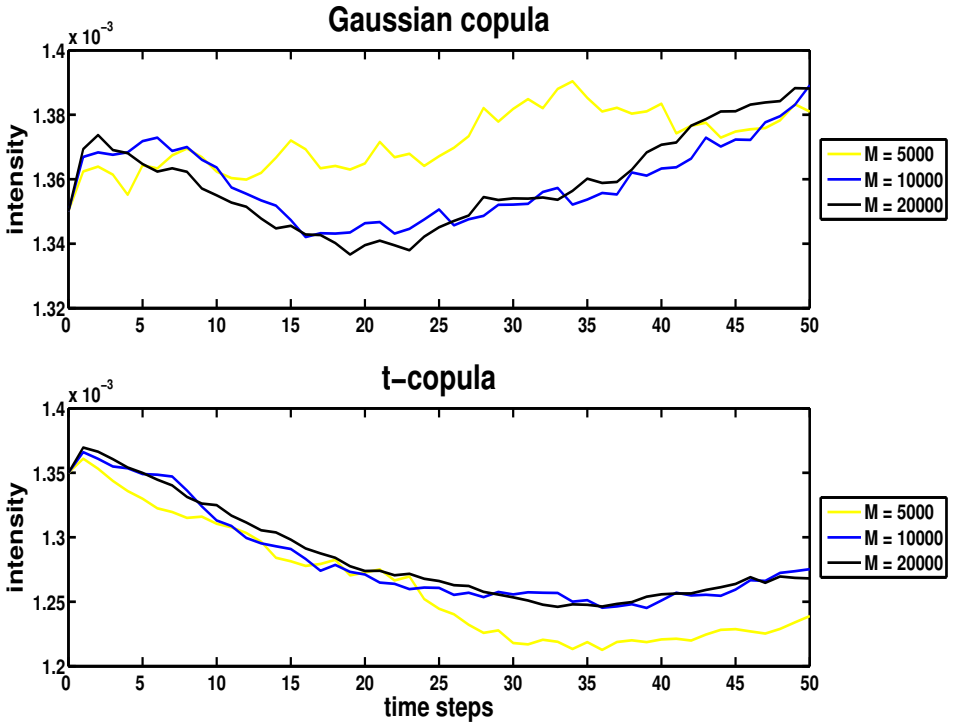
F - B.25 Figure: simulated intensity for counterparty $d = 11$ - $K = 1$ and *abs*



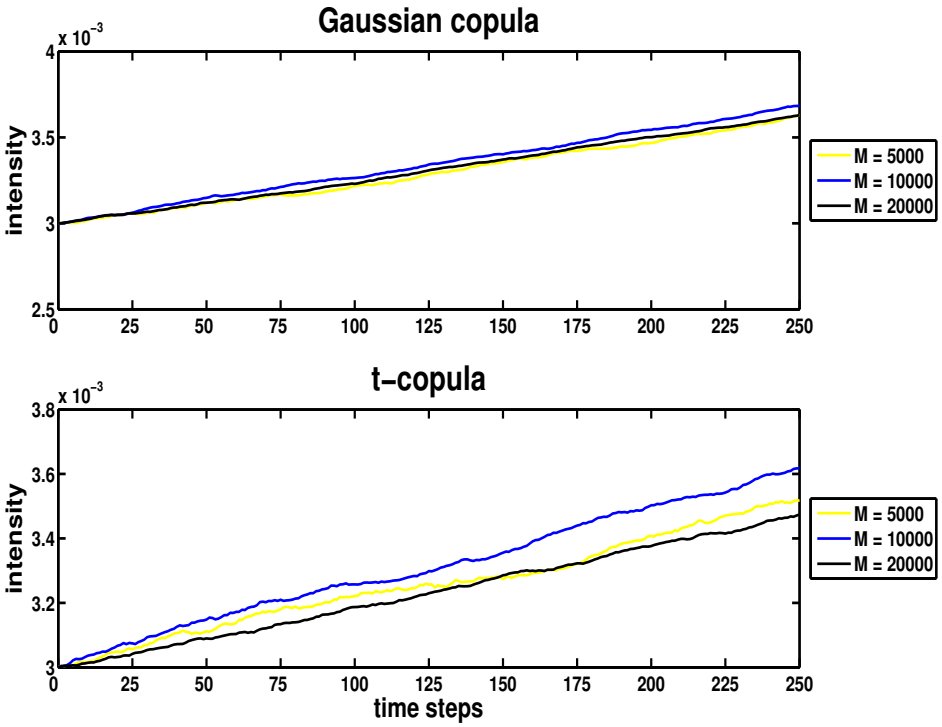
F - B.26 Figure: simulated intensity for counterparty $d = 11$ - $K = 5$ and *abs*



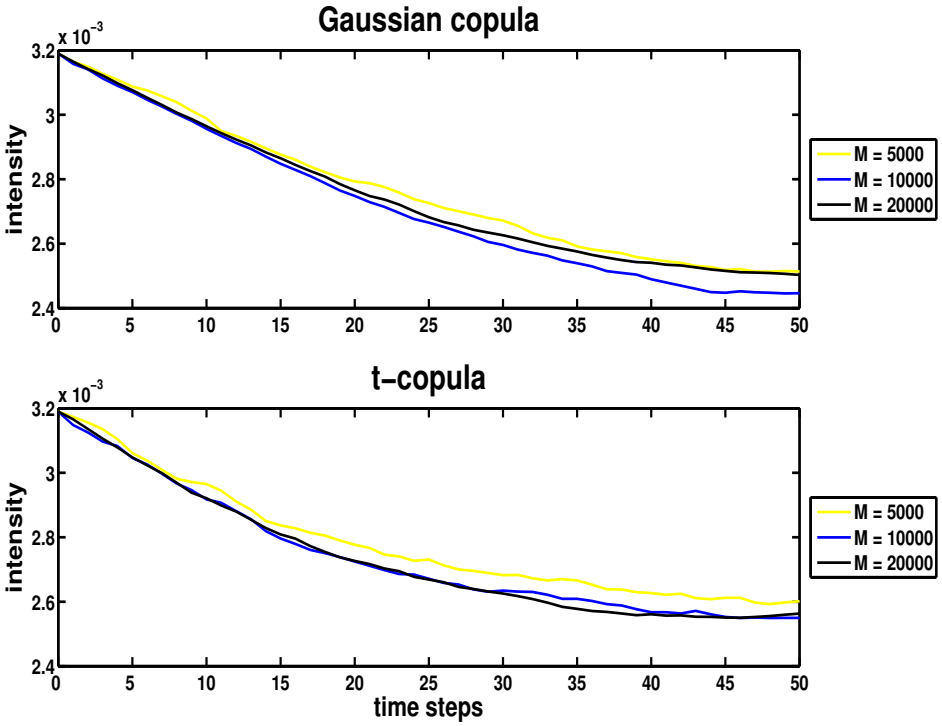
F - B.27 Figure: simulated intensity for counterparty $d = 12$ - $K = 1$ and *abs*



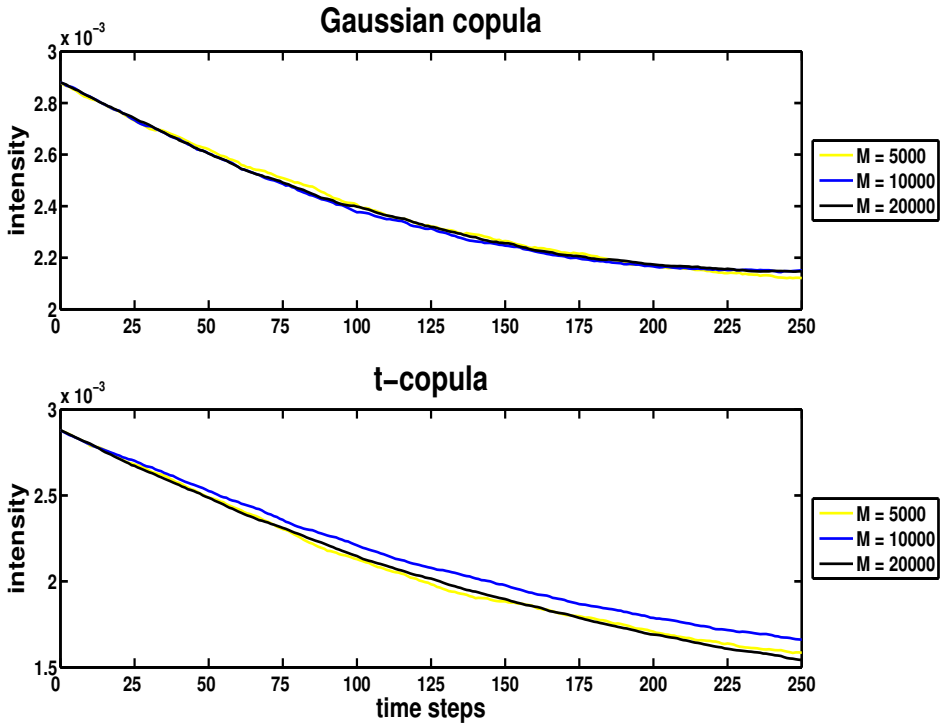
F - B.28 Figure: simulated intensity for counterparty $d = 12$ - $K = 5$ and *abs*



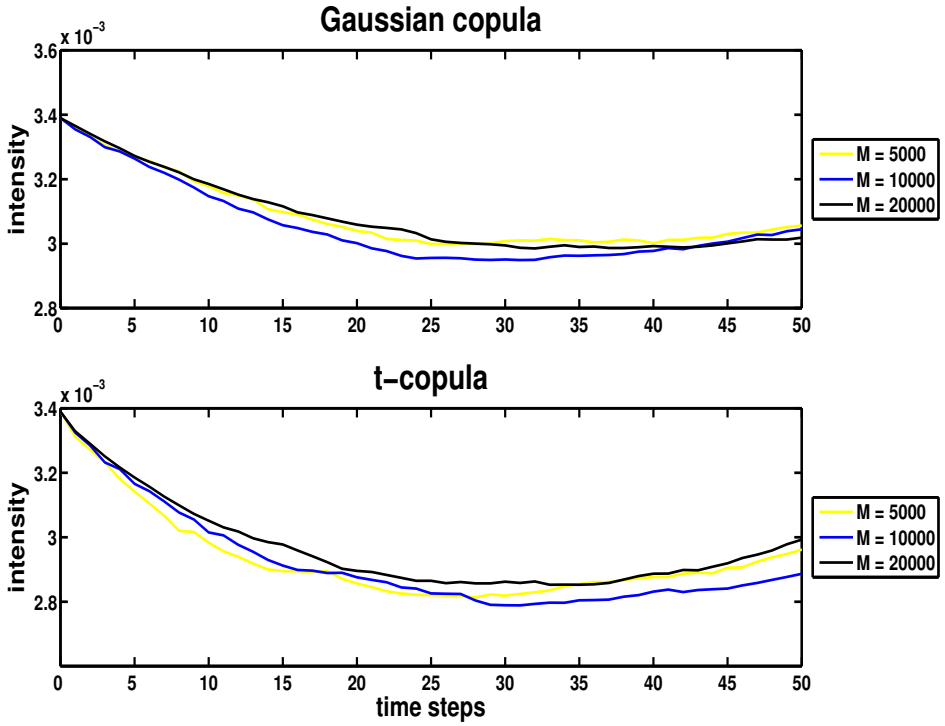
F - B.29 Figure: simulated intensity for counterparty $d = 13$ - $K = 1$ and *abs*



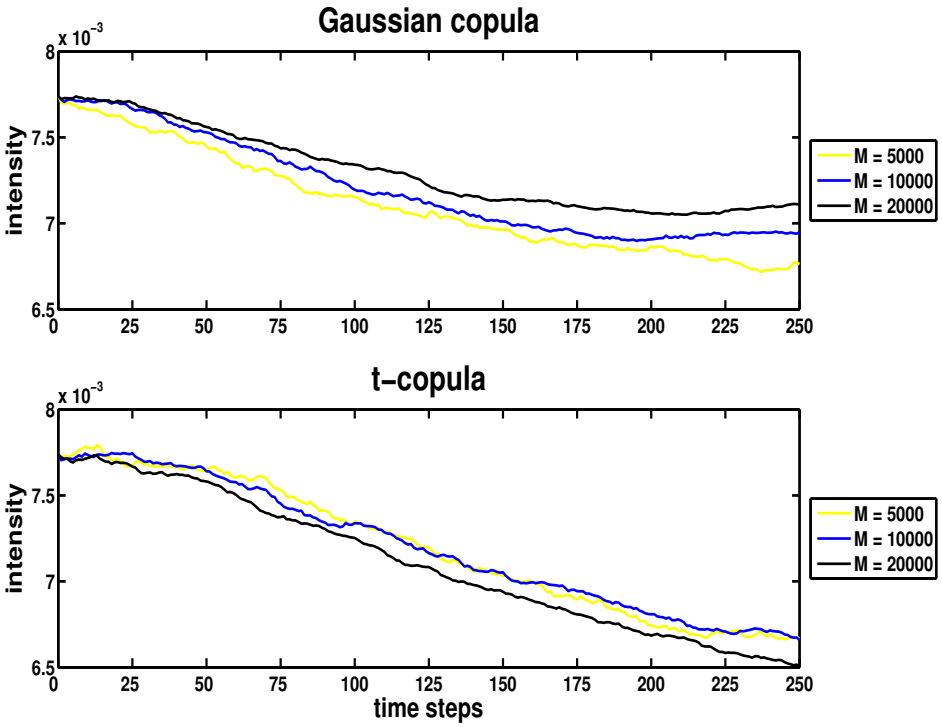
F - B.30 Figure: simulated intensity for counterparty $d = 13$ - $K = 5$ and *abs*



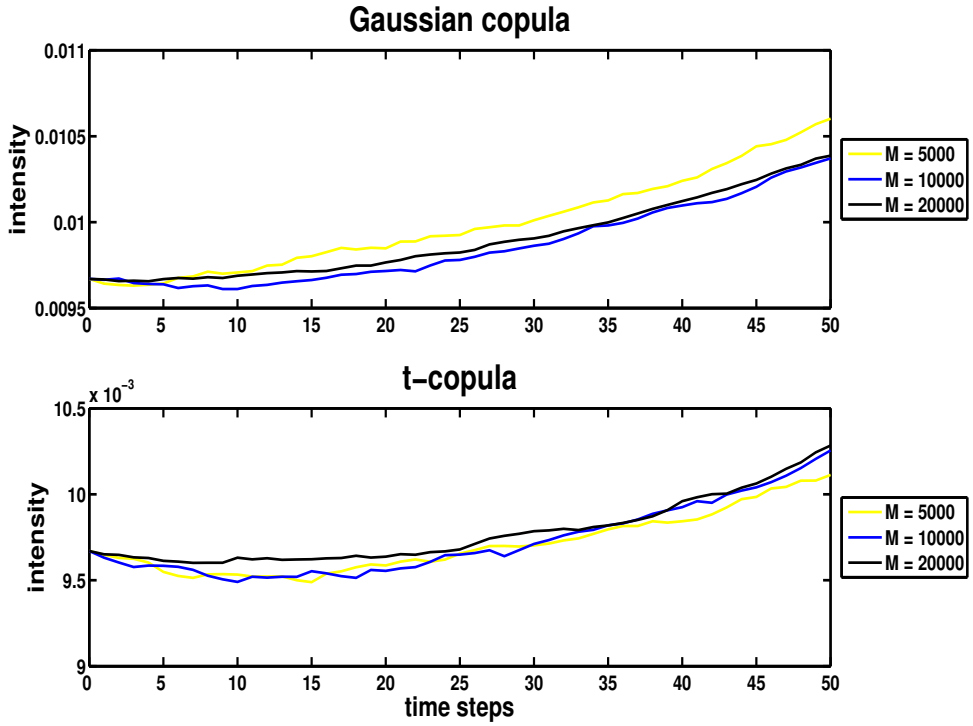
F - B.31 Figure: simulated intensity for counterparty $d = 14$ - $K = 1$ and *abs*



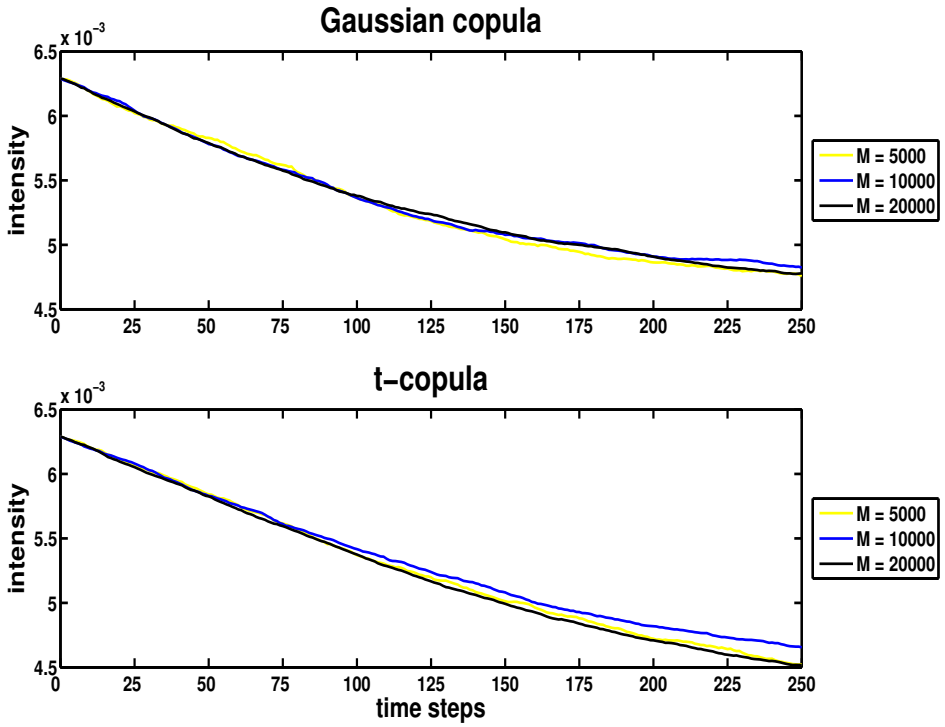
F - B.32 Figure: simulated intensity for counterparty $d = 14$ - $K = 5$ and *abs*



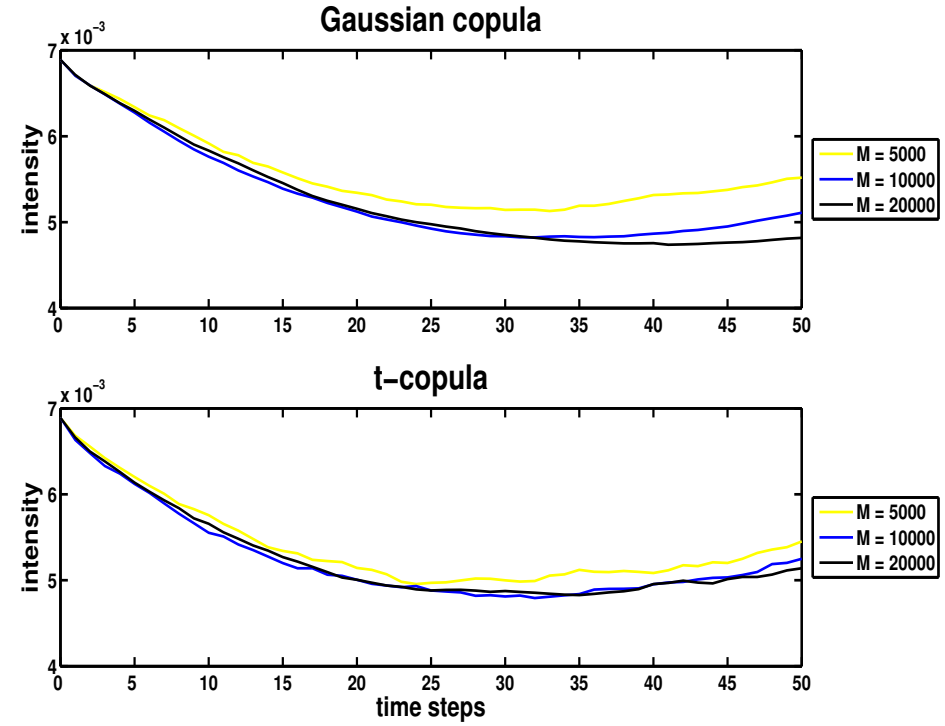
F - B.33 Figure: simulated intensity for counterparty $d = 15$ - $K = 1$ and *abs*



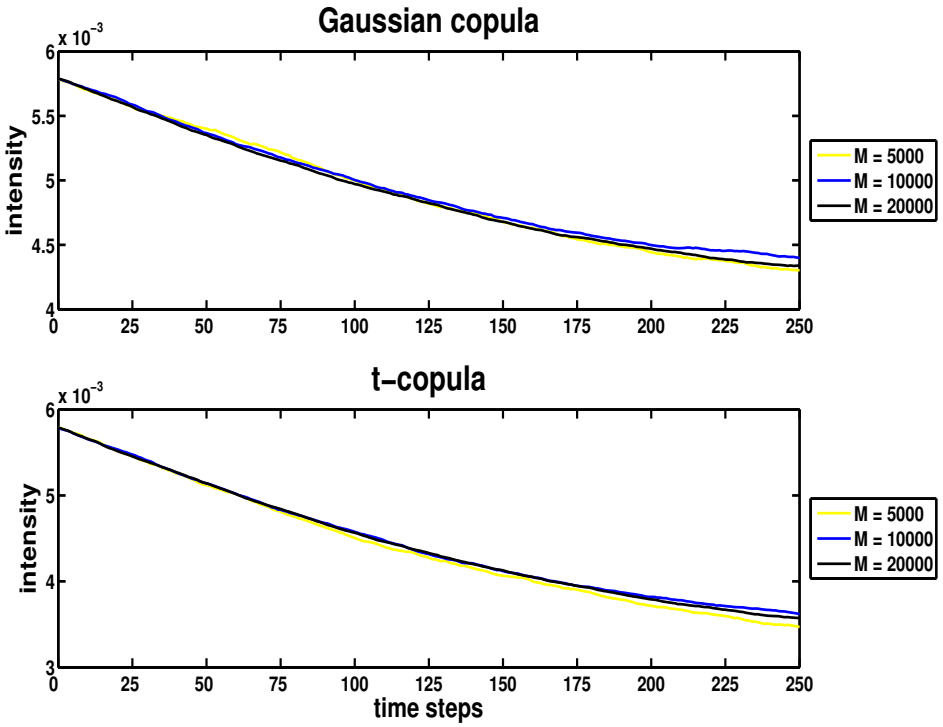
F - B.34 Figure: simulated intensity for counterparty $d = 15$ - $K = 5$ and *abs*



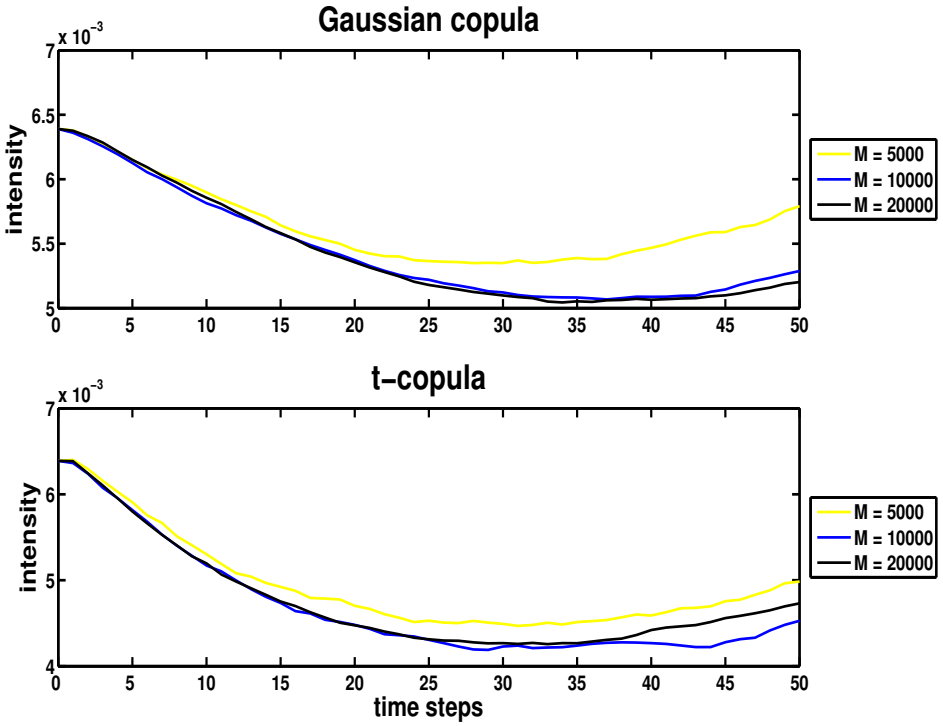
F - B.35 Figure: simulated intensity for counterparty $d = 16$ - $K = 1$ and *abs*



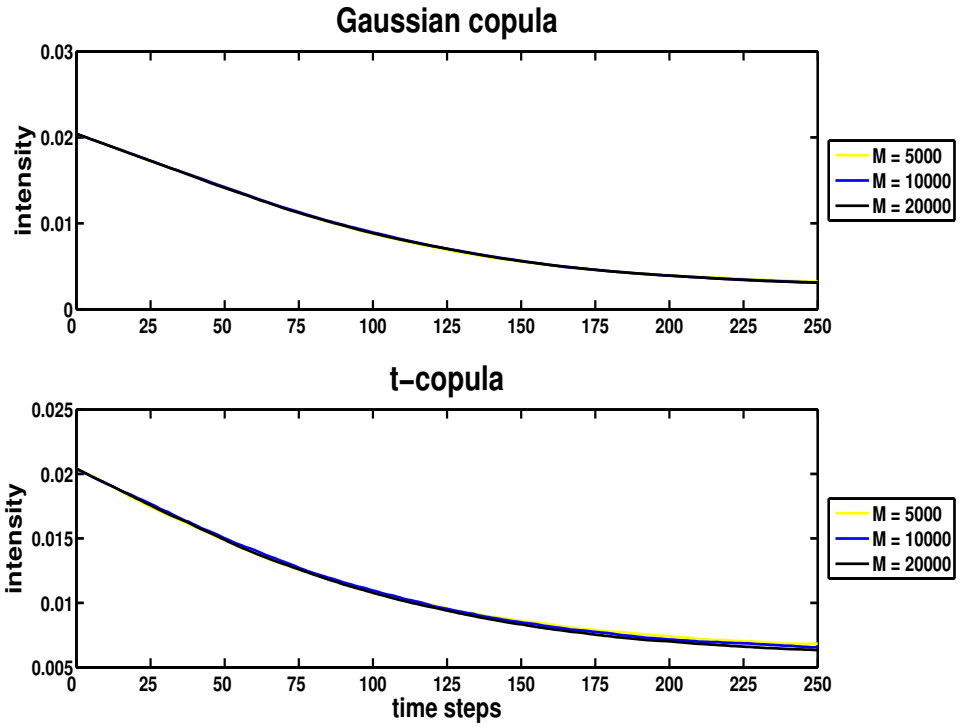
F - B.36 Figure: simulated intensity for counterparty $d = 16$ - $K = 5$ and *abs*



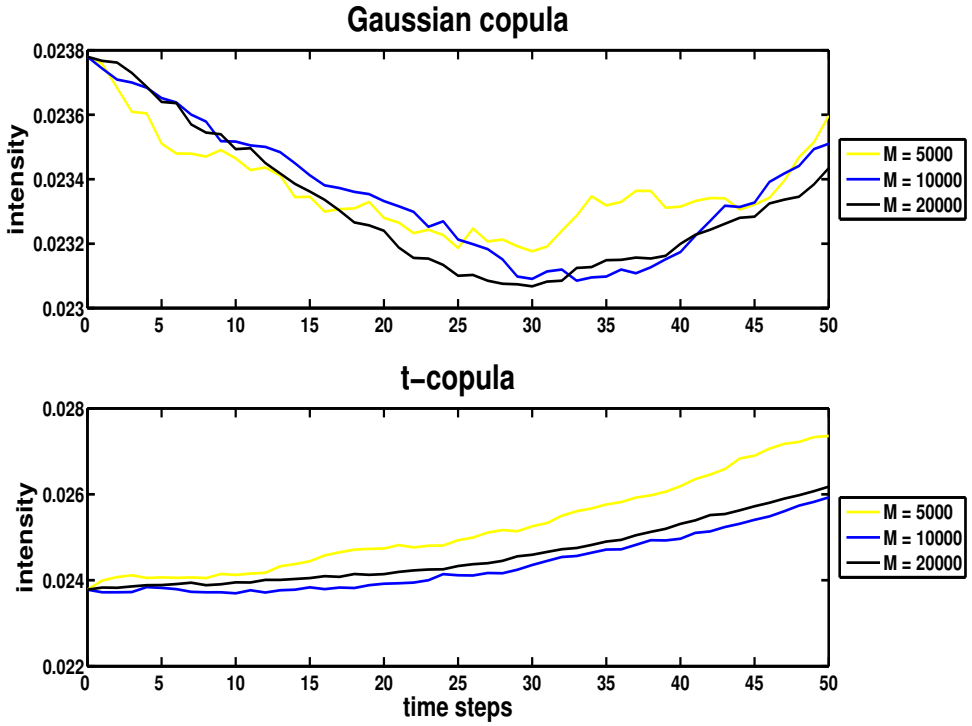
F - B.37 Figure: simulated intensity for counterparty $d = 17$ - $K = 1$ and *abs*



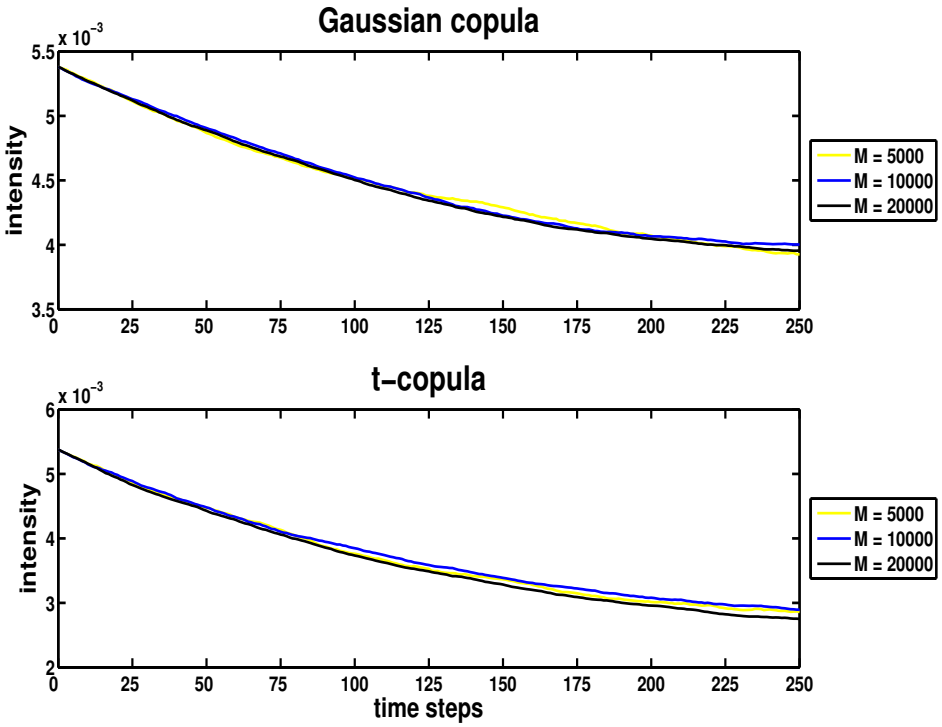
F - B.38 Figure: simulated intensity for counterparty $d = 17$ - $K = 5$ and *abs*



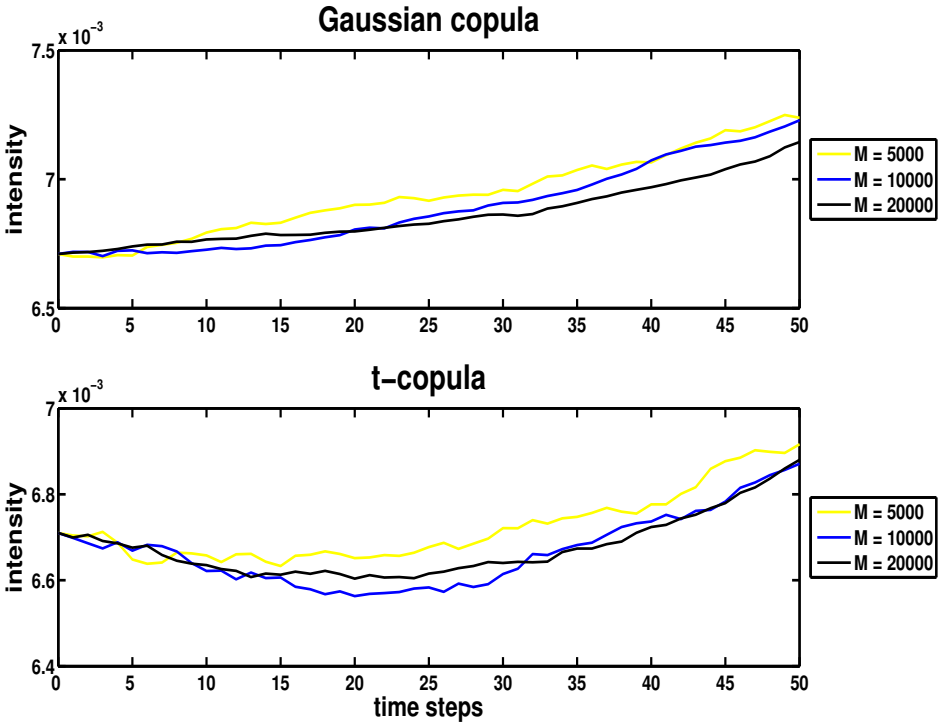
F - B.39 Figure: simulated intensity for counterparty $d = 18$ - $K = 1$ and *abs*



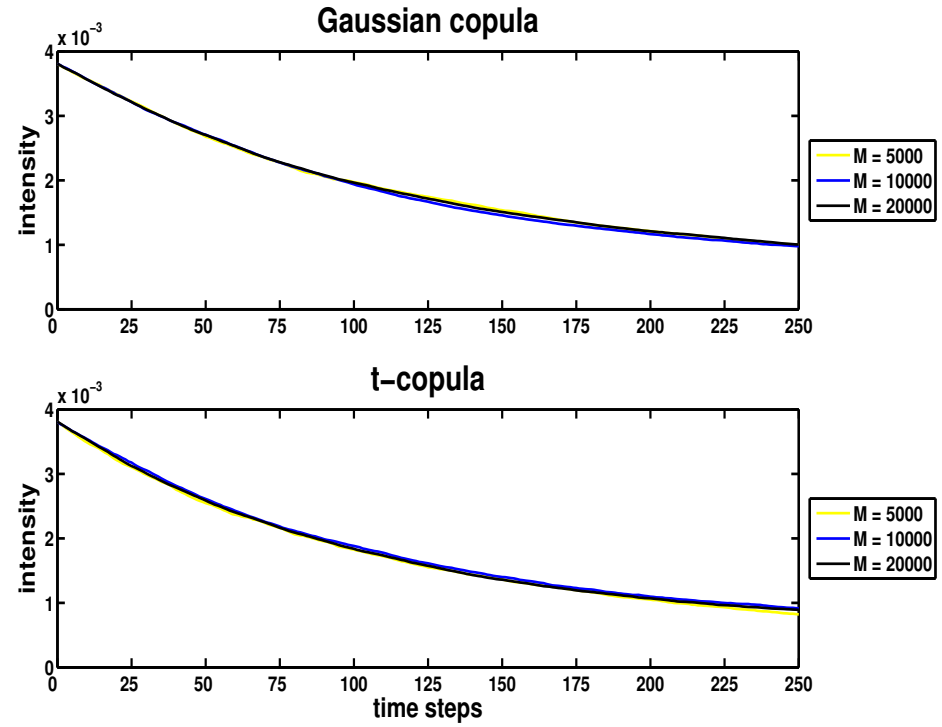
F - B.40 Figure: simulated intensity for counterparty $d = 18$ - $K = 5$ and *abs*



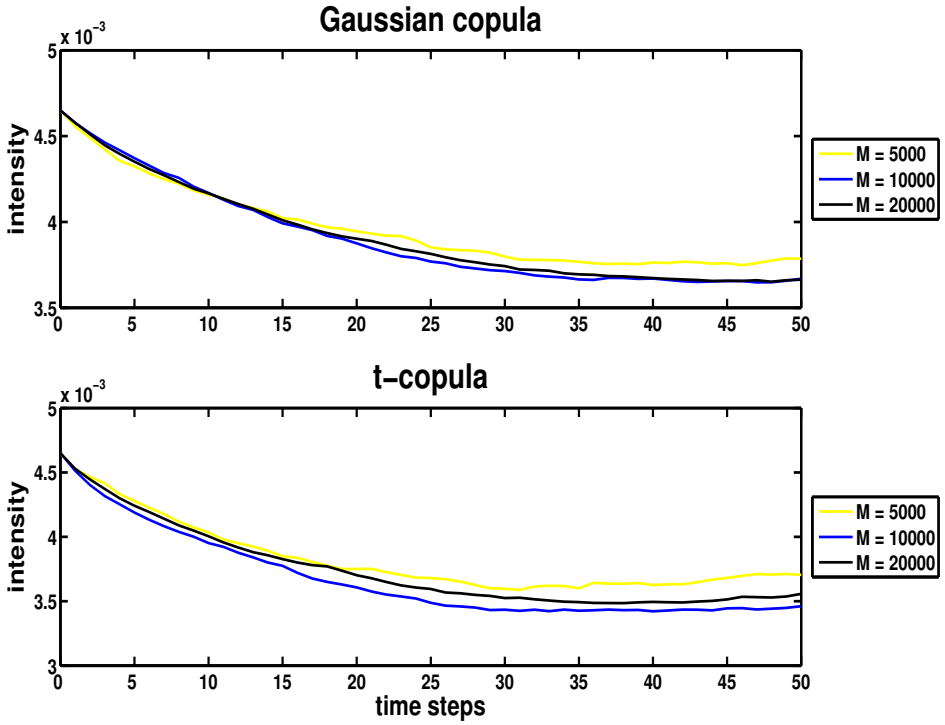
F - B.41 Figure: simulated intensity for counterparty $d = 19$ - $K = 1$ and *abs*



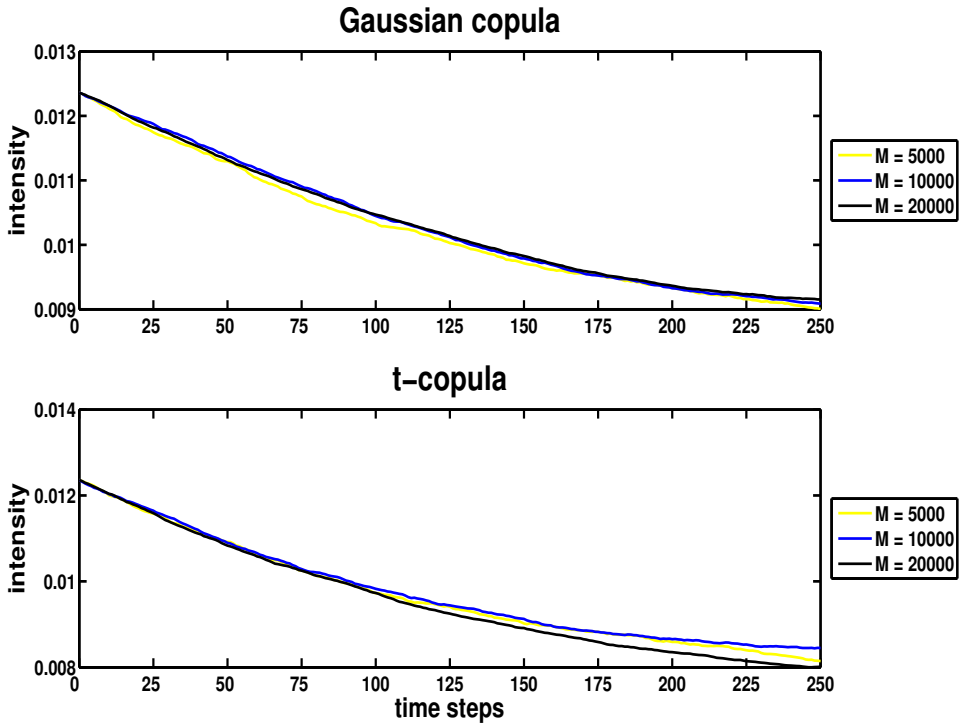
F - B.42 Figure: simulated intensity for counterparty $d = 19$ - $K = 5$ and *abs*



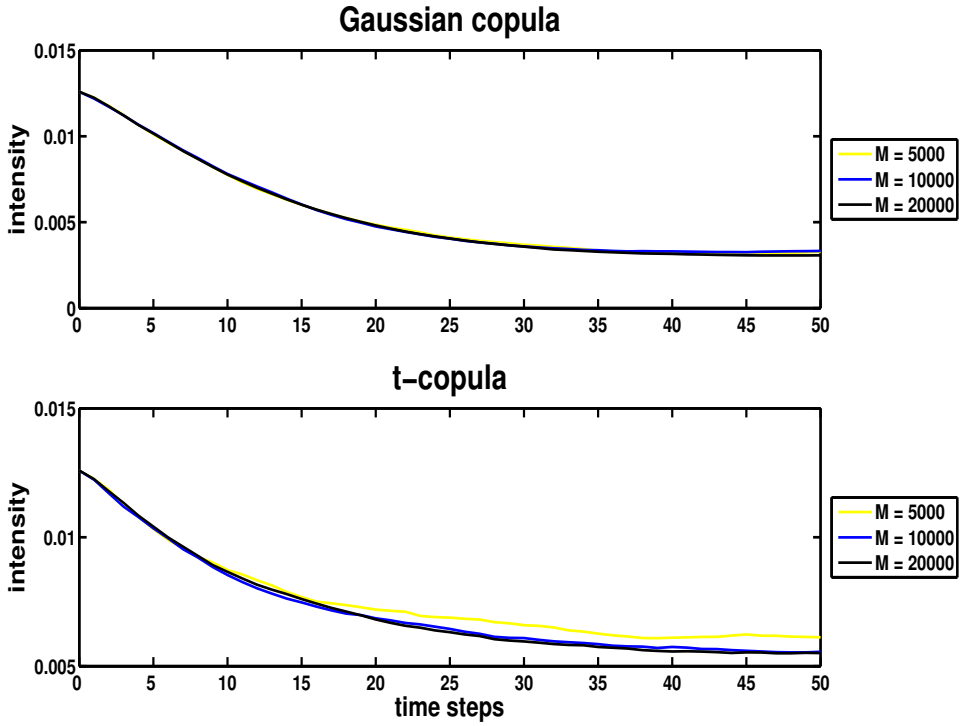
F - B.43 Figure: simulated intensity for counterparty $d = 20$ - $K = 1$ and *abs*



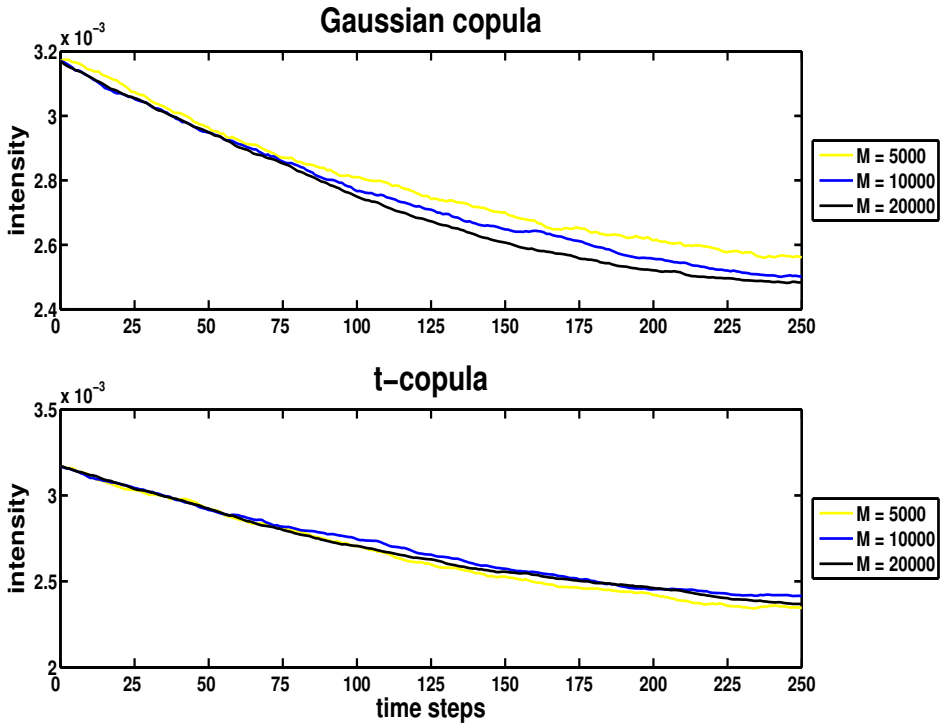
F - B.44 Figure: simulated intensity for counterparty $d = 20$ - $K = 5$ and *abs*



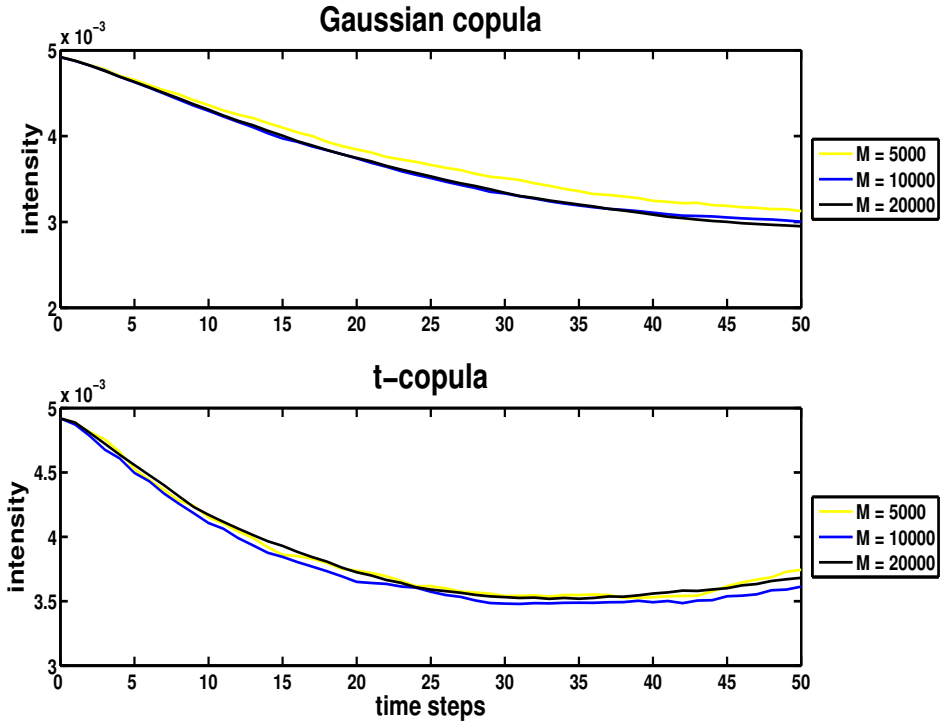
F - B.45 Figure: simulated intensity for counterparty $d = 21$ - $K = 1$ and *abs*



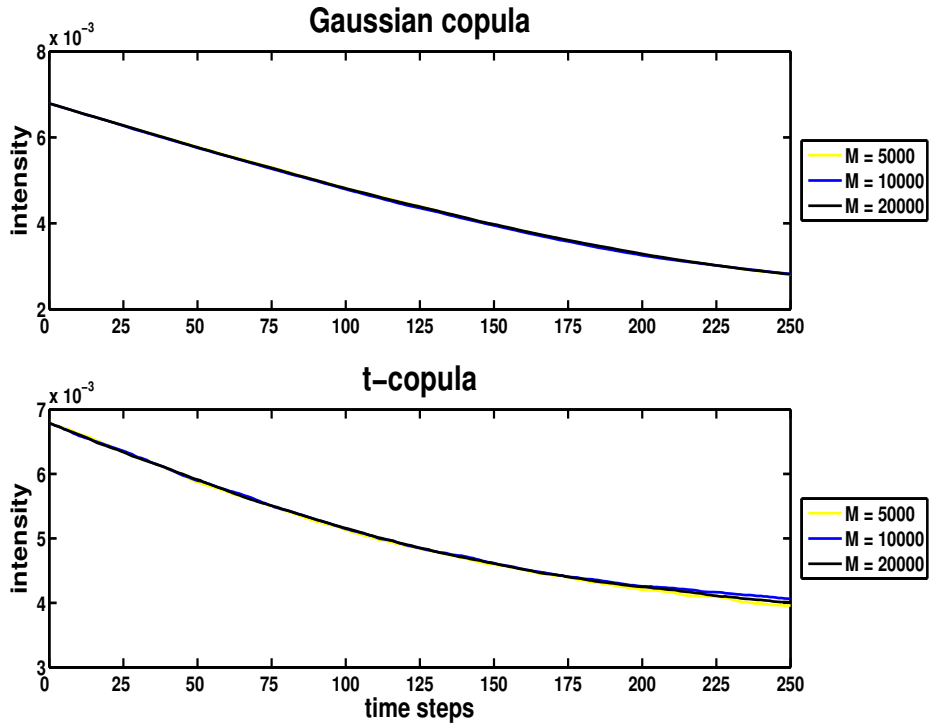
F - B.46 Figure: simulated intensity for counterparty $d = 21$ - $K = 5$ and *abs*



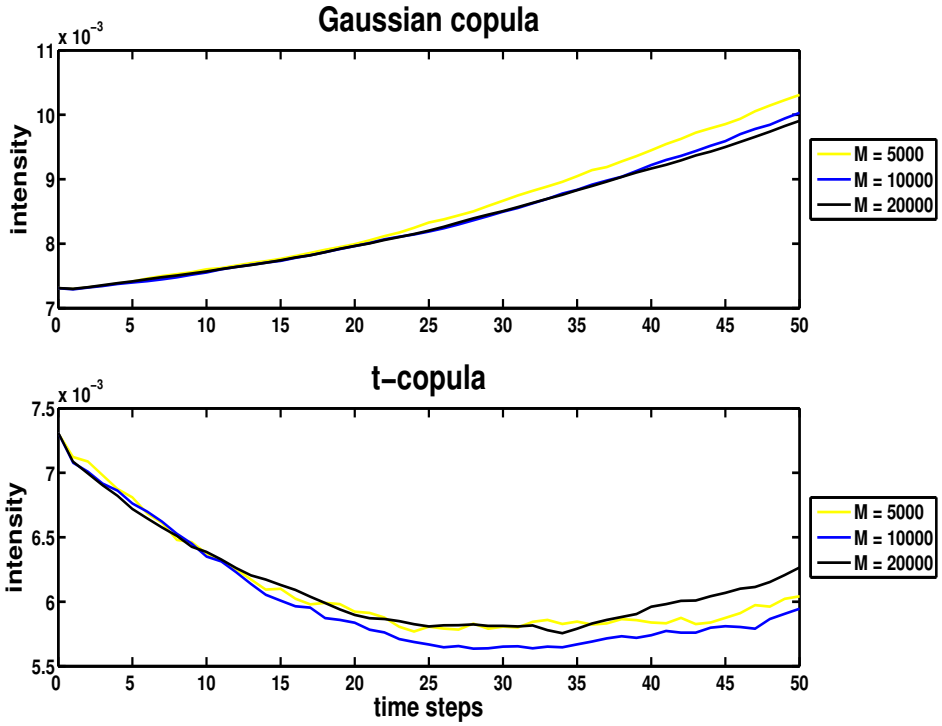
F - B.47 Figure: simulated intensity for counterparty $d = 22$ - $K = 1$ and *abs*



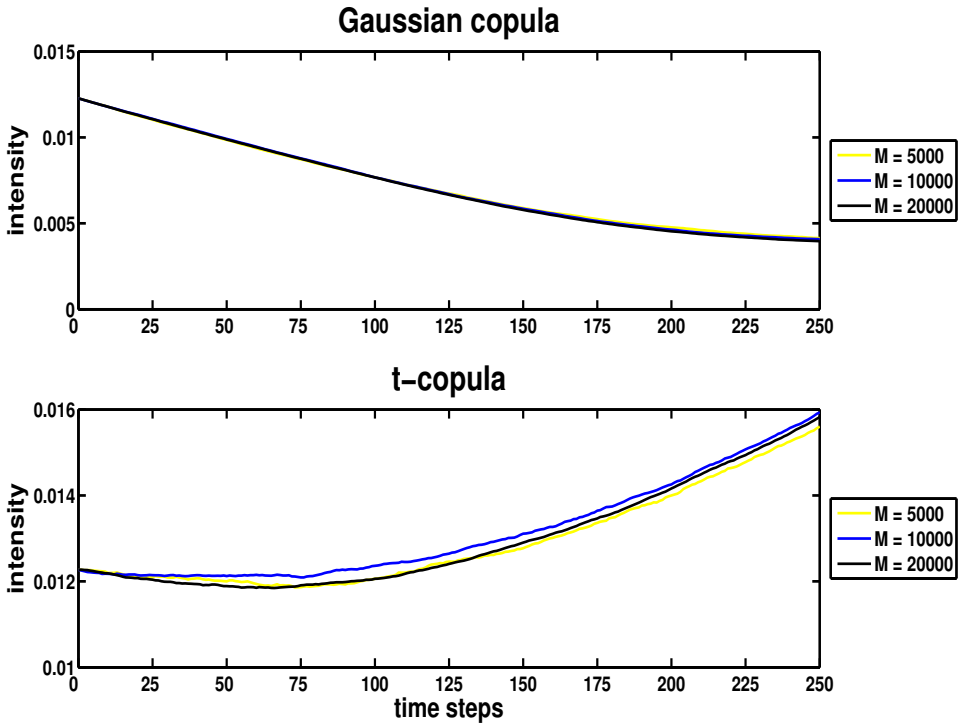
F - B.48 Figure: simulated intensity for counterparty $d = 22$ - $K = 5$ and *abs*



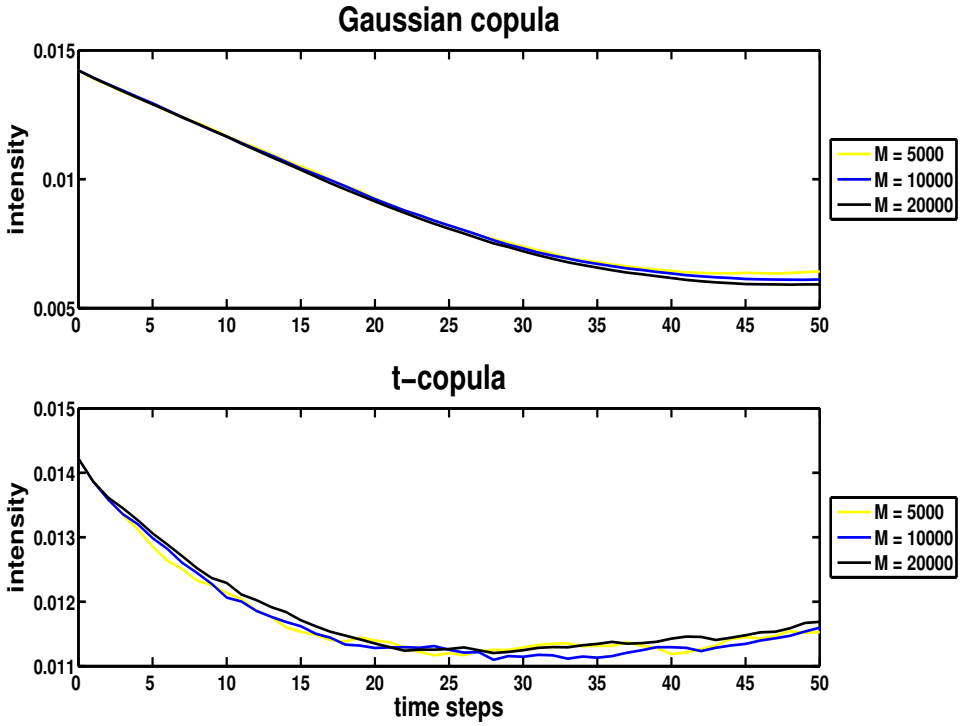
F - B.49 Figure: simulated intensity for counterparty $d = 23$ - $K = 1$ and *abs*



F - B.50 Figure: simulated intensity for counterparty $d = 23$ - $K = 5$ and *abs*



F - B.51 Figure: simulated intensity for counterparty $d = 24$ - $K = 1$ and *abs*



F - B.52 Figure: simulated intensity for counterparty $d = 24$ - $K = 5$ and *abs*

B.6 An Alternative Algorithm

A - B.1 Algorithm (Alternative Copula- and Default-Dependent Intensity Model)

Given

- time horizon with time grid t_n, $n = 0, \ldots, N$, $t_0 = 0$, $t_n = T$
- copula function K and marginal distributions F^d, $d = 1, \ldots, D$

Step 1

- for $d = 1, \ldots, D$:
  - calibrate $\Lambda^d \sim F^d$ to market data on the given time grid
  - compute $(F^d)^{-1}$
  - set $\lambda_0^d > 0$ according to market data, κ_0^d, $\gamma_0 = 1$
  - draw $u^d \sim U(0,1)$
- calibrate K to market data on the given time grid
- set $K_0 = K$
- set $\mathcal{E}_0 = \{1, \ldots, D\}$, $\mathcal{C}_0 = \emptyset$

Step 2 - for $n = 1, \ldots, N$:

- for $d \in \mathcal{E}_{n-1}$:
  - compare $\gamma_{n-1}^d \cdot \exp(-\lambda_{n-1}^d (t_n - t_{n-1}) \leq u^d$
  - true $\Rightarrow$ default for position $\tilde{d}$
    - set $\mathcal{E}_{n-1} = \mathcal{E}_{n-1} \setminus \{\tilde{d}\}$, $\mathcal{C}_{n-1} = \mathcal{C}_{n-1} \cup \{\tilde{d}\}$
    - compute
$$\tilde{v}_n^{\tilde{d}} = \frac{F^{\tilde{d}}(0) - F^{\tilde{d}}\left(-\lambda_{t_{n-1}}^{\tilde{d}}\right)}{\kappa_{n-1}^{\tilde{d}} - F^{\tilde{d}}\left(-\lambda_{t_{n-1}}^{\tilde{d}}\right)}$$
    - set $\tilde{v}(\mathcal{C}_{n-1}) = (\tilde{v}(\mathcal{C}_{n-1}), \tilde{v}_n^{\tilde{d}})$[1]
    - next $d \in \mathcal{E}_{n-1}$
  - false $\Rightarrow$ no default: next $d \in \mathcal{E}_{n-1}$
- set $\mathcal{E}_n = \mathcal{E}_{n-1}$ and $\mathcal{C}_n = \mathcal{C}_{n-1}$
- compute $K_n = \frac{\partial}{\partial \tilde{v}(\mathcal{C}_n)} K_{n-1}$ given $V(\mathcal{C}_n) = \tilde{v}(\mathcal{C}_n)$
- draw $v_n \sim K_n$
- for $d \in \mathcal{E}_n$:
  - compute $\Delta\lambda_n^d = (F^d)^{-1}\left[v_n^d\left(\kappa_n^d - F^d\left(-\lambda_{n-1}^d\right)\right) + F^d\left(-\lambda_{n-1}^d\right)\right]$
  - compute $\lambda_n^d = \lambda_{n-1}^d + \Delta\lambda_n^d$
  - compute $\gamma_n^d = \gamma_{n-1}^d \cdot \exp(-\lambda_{n-1}^d (t_n - t_{n-1}))$
- for $d \in \mathcal{C}_n$:
  - set $\Delta\lambda_n^d = 0$, $\lambda_n^d = \lambda_{n-1}^d$ and $\gamma_n^d = \gamma_{n-1}^d$
- set $n = n + 1$

[1] Not in permuted order.

C α-stable Distributions

The application of mathematical models naturally generates issues such as calibration to data, simulation or evaluation. Especially when components of time are involved (as for stochastic processes for instance), outcomes depend on the relevant time horizon and the specified time grid for numerical implementation. Once this setup is fixed, all modeling variables are fitted to this very partition. Afterward, a further subsequent separation is generally not possible.

In the presented càdlàg framework with generalized (i.e., not truncated) distributions, the problem of being restricted to one predetermined time grid can be bypassed - to some extents - if marginal distributions follow an **α-stable law**.

For outlining the idea, assume that the general distributions F^d, $d = 1, \dots, D$, are calibrated to the time grid $[t_n, t_{n+1})$, $n = 0, \dots, N$[1]. Now, it can be desirable to use a finer resolution: Say $[t_n, t_{n+1})$ has $M \in \mathbb{N}$ sub-intervals $[t_{n,m-1}, t_{n,m})$, $m = 1, \dots, M$, with $t_{n,0} = t_n$ and $t_{n,M} = t_{n+1}$.

The direct way is to fit F^d to the new setup. Avoiding a new calibration, one can chose distributions which offer certain stability features, i.e., the sum of the random variables on the sub-interval Λ_m^d are connected to the former random variable Λ^d in terms of distributions, mathematically

$$\sum_{m=0}^{M-1} \Lambda_m^d \stackrel{d}{=} \Lambda^d \quad .$$

So called *α-stable* distributions feature this sum-stability property[2]. The mathematical framework - comprising equivalent definitions, characteristic of α-stable laws and a multivariate extension in combination with copulas - are presented in the following Sections C.1 and C.2.

C.1 One-dimensional α-stable Distributions

The analysis of α-stable laws - α indicates the *index of stability* - as sums of i.i.d. random variables goes back to the beginning of probability theory. As a pioneer, Paul Lévy advanced developments on this sub-domain. Early comprehensive and quite general monographs on the topic are "Limit Distributions For Sums of Independent Random Variables" [30] by Gnedenko and Kolmogorov in 1954[3] and "An Introduction to Probability Theory and Its Applications" [24] by Feller[4].

In particular, profound monographs on α-stable distributions are sparse, advisable are "Stable Non-Gaussian Random Processes" [57] by Samorodnitsky and Taqqu and "One-Dimensional Stable Distributions" [66] by Zolotarev[5].

The subsequent results are primarily borrowed from [57], particularly with regard to notation and parametrization[6]. Due to the fact that **four** parameters determine α-stable random variables, their distributions provide a great flexibility for modeling issues, for instance heavy-tailed or skewed laws. Though, density functions do not have closed-forms in general.

[1] Distributions F^d, $d = 1, \dots, D$, are not truncated as conditional cdfs in each time step as for our model, see Chapter 10.

[2] In literature, the reader will also find different types of stability under diverse mathematical operations such as max, min amongst others. We focus on sum-stable distributions.

[3] Revised in 1968. The Russian edition appeared in 1946.

[4] First edition 1966, second edition 1971. As a German book "Wahrscheinlichkeitstheorie" by Bauer [3] is recommendable.

[5] Further information is found in the books of Christoph and Werner [13] and Nolan [53].

[6] Throughout history of stable distributions there are confusions about

- the notation of stability (strictly stable, stable and quasi stable) and
- the description of parameters. Dependent exploited features, a different parametrization are used.

C.1.1 Equivalent Definitions and Properties of α-stable Laws

Samorodnitsky and Taqqu show in [57], Chapter 1, that the following definitions of α-stable laws are equivalent. he first and commonly most intuitive specification points out the stability feature. It indicates that the distribution of sums of i.i.d. random variables is equivalent to the distribution of the original random variable under certain conditions.

C.1 Definition (compare Definition 1.1.4 in [57])
A random variable X has *stable distribution* with *index of stability* or *characteristic exponent* α if for any $N \in \mathbb{N}, N \geq 2$, there exist a positive number C_N and a real number D_N such that

$$\sum_{n=1}^{N} X_n \stackrel{d}{=} C_N X + D_N$$

holds in which $X_1, \ldots, X_N$ are i.i.d. copies of X. Feller showed in [24], Section VI.1., Theorem 1., that necessarily it holds $C_N = N^{1/\alpha}$ for some $\alpha \in (0, 2]$.

C.2 Notation (α-stable)
A stable distribution with index of stability α is usually denoted as *α-stable distribution*. If $D_N = 0$ for all N, the random variable X is called *strictly stable*.

The second definition is related to the central limit theorem[7].

C.3 Definition (compare Definition 1.1.5 in [57])
A random variable X has stable distribution if it has a *domain of attraction*, i.e., for sequences of i.i.d. random variables $Y_1, \ldots, Y_N$, $N \in \mathbb{N}$, with positive numbers c_N and real numbers d_N

$$\frac{1}{c_N} \sum_{n=1}^{N} Y_n + d_N \stackrel{d}{\to} X \tag{C.1}$$

holds. Note that the random variables Y_n, $n \in \mathbb{N}$, must not necessarily be stable. If the sequence in the denominator of (C.1) holds

$$c_N = N^{1/\alpha} \quad ,$$

then the Y_n's belong to the *normal domain of attraction*. Generally, the sequence can be set as

$$c_N = N^{1/\alpha} \cdot h(N)$$

in which $h(x)$, $x > 0$, is a slowly varying function at infinity[8] which means that

$$\lim_{x \to \infty} \frac{h(u \cdot x)}{h(x)} = 1 \quad \forall\ u > 0 \quad .$$

The next example shows in how far Definition (C.3) is connected the Central Limit Theorem.

C.4 Example (Central Limit Theorem)
Let $Y_1, \ldots, Y_N$ be a sequence of i.i.d. random variables which are square integrable, i.e., $\mathbb{E}[Y_1^2] < \infty$, such that Y_1 has finite variance σ^2. Choose $c_N = \sqrt{N} \cdot \sigma$. W.lo.g. let $\mathbb{E}[Y_1] = 0$ and set $d_N = 0$. Then, due to the Central Limit Theorem B.7, it holds

$$\frac{Y_1 + \ldots + Y_N}{\sqrt{N}\sigma} \stackrel{d}{\to} X \sim N(0, 1) \quad .$$

Thus, Equation (C.1) can be regarded as a sort of expanded version of the Central Limit Theorem for i.i.d. random variables.

[7] Compare Appendix B.1, Theorem B.7.
[8] Compare [24], Section XVII.5.

The last definition is the most useful one for analysis and calculation as it comprises the characteristic function and all four defining parameters of stable distributions.

C.5 Notation (Parameters of Stable Laws)
An α-stable distribution is uniquely specified by four parameters:

1. $\alpha \in (0, 2]$ - the *index of stability* or *characteristic exponent*
2. $\sigma \geq 0$ - the *scale parameter*
3. $\beta \in [-1, 1]$ - the *skewness parameter*
4. $\mu \in \mathbb{R}$ - the *shift* or *location parameter*

This notation becomes clear in Section C.1.2 dealing with properties of stable laws.

C.6 Definition (compare Definition 1.1.6, [57])
A random variable X has stable distribution if its characteristic function ϕ_X holds

$$\begin{aligned}\phi_X(\theta) &= \mathbb{E}\left[\exp\left(i\theta X\right)\right] \\ &= \begin{cases} \exp\left(-\sigma^\alpha|\theta|^\alpha\left(1 - i\beta\,\mathrm{sgn}(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\right) + i\mu\theta\right) & \text{for } \alpha \neq 1 \\ \exp\left(-\sigma|\theta|\left(1 + i\beta\frac{2}{\pi}\,\mathrm{sgn}(\theta)\ln\left(|\theta|\right)\right) + i\mu\theta\right) & \text{for } \alpha = 1 \end{cases} .\end{aligned} \tag{C.2}$$

Samorodnitsky and Taqqu show in [57] that Definitions C.1, C.3 and C.6 are equivalent.

C.7 Notation
If X has stable distribution, we denote $X \sim S_\alpha(\sigma, \beta, \mu)$ with the parametrization of (C.2)[9].

It is easy to check that parameter β vanishes for $\alpha = 2$, and Equation (C.2) states

$$\phi_Z(\theta) = \exp\left(-\sigma^2|\theta|^2 + i\mu\theta\right)$$

which is unmistakable the characteristic function for a **Gaussian** random variable $X \sim N(\mu, 2\sigma^2)$. Moreover, for $\alpha = 1$, the term $\tan\left(\frac{\pi}{2}\right) = \infty$ causes difficulties and explains the separation of (C.2). Without going into further details, slight different equations and features need to be analyzed for this case.

C.1.2 Parameters and Properties of Stable Laws

Due to Equation (C.2), it is obvious that the most influencing parameter is the index of stability α. Moreover, we want to reason the denotation of the other parameters, see Notation C.5. Some properties of stable laws can directly be assigned, compare [57], Properties 1.2.2 - 1.2.5, 1.2.8, Corollaries 1.2.7 and 1.2.9.

C.8 Proposition (Properties and Parameters)
Let $X \sim S_\alpha(\sigma, \beta, \mu)$.

1. Let $k \in \mathbb{R}$. Then, $X + k \sim S_\alpha(\sigma, \beta, \mu + k)$ justifying the notation **shift** parameter.
2. Let $k \in \mathbb{R} \setminus \{0\}$. Then,

$$\begin{aligned}&\text{for } \alpha \neq 1: \quad k \cdot X \sim S_\alpha(|k| \cdot \sigma, \mathrm{sgn}(k) \cdot \beta, k \cdot \mu) \quad \text{and} \\ &\text{for } \alpha = 1: \quad k \cdot X \sim S_1\left(|k| \cdot \sigma, \mathrm{sgn}(k) \cdot \beta, k \cdot \left(\mu - \frac{2}{\pi} \cdot \ln(|k|)\sigma\beta\right)\right)\end{aligned}$$

   justifying the notation **scale** parameter for σ.

[9] For a discussion of the different parametrization and notation of parameters, the reader is referred to [66].

3. Let $\mu = 0$ and $\alpha \in (0, 2)$. Then,

$$X \sim S_\alpha(\sigma, \beta, 0) \quad \Leftrightarrow \quad -X \sim S_\alpha(\sigma, -\beta, 0)$$

4. $X \sim S_\alpha(\sigma, \beta, \mu)$ is symmetric about μ if and only if $\beta = 0$.

   Properties 3 and 4 justify the notation **skewness** parameter for β.

5. Let $X \sim S_\alpha(\sigma, \beta, \mu)$ and $\alpha \neq 1$. Then, $X - \mu$ is strictly stable.

   Let $X \sim S_1(\sigma, \beta, \mu)$. Then, X is strictly stable if and only if $\beta = 0$.

6. Let $X_1, \ldots, X_N$ be i.i.d. with $X_1 \sim S_\alpha(\sigma, \beta, \mu)$. Then,

$$\text{for } \alpha \neq 1: \quad \underbrace{\frac{1}{N^{1/\alpha}}}_{c_N} \sum_{n=1}^{N} X_n \quad + \quad \underbrace{-\frac{1}{N^{1/\alpha}} \left(\mu \left(N - N^{1/\alpha} \right) \right)}_{d_N} \quad \stackrel{d}{=} X_1 \quad \text{and}$$

$$\text{for } \alpha = 1: \quad \underbrace{\frac{1}{N}}_{c_N} \sum_{n=1}^{N} X_n \quad + \quad \underbrace{-\frac{2}{\pi} \ln(N) \sigma \beta}_{d_N} \quad \stackrel{d}{=} X_1 \quad .$$

A proof can be found in [57], Sections 1.2. ∎

The following proposition is Property 1.2.16 in [57] and sheds light on the moments (expectation and variance) of stable random variables.

C.9 Proposition (Moments of α-stable Laws)

1. For $X \sim S_\alpha(\sigma, \beta, \mu)$ and $\alpha \in (0, 2)$, it holds

$$\mathbb{E}\left[|X|^p\right] < \infty \text{ for } 0 < p < \alpha \quad \text{and} \quad \mathbb{E}\left[|X|^p\right] = \infty \text{ for } p \geq \alpha \quad .$$

2. For $\alpha \in (1, 2]$ and $X \sim S_\alpha(\sigma, \beta, \mu)$, it holds $\mathbb{E}[X] = \mu$.

Sketch of Proof

A derivation is found in [57], pp. 16 - 18, which exploits the tail behavior of α-stable distributions. For $X \sim S_\alpha(\sigma, \beta, \mu)$ and $0 < \alpha < 2$, we have

$$\lim_{\lambda \to \infty} \lambda^\alpha \mathbb{P}[X > \lambda] = \zeta_\alpha \frac{1 + \beta}{2} \sigma^\alpha \qquad \text{and}$$
$$\lim_{\lambda \to \infty} \lambda^\alpha \mathbb{P}[X < -\lambda] = \zeta_\alpha \frac{1 - \beta}{2} \sigma^\alpha$$

for some α-depending constant ζ_α. As $\mathbb{E}[|X|^p] = \int_0^\infty \mathbb{P}[|X|^p > \lambda] d\lambda$, the statement follows. ∎

Summarized, the advantageous features of α-stable random variables are

+ stability in distribution
+ closed-form characteristic functions
+ the possibility to generate a great variety of distributions (skewed, heavy-tailed etc.).

As drawback - especially as a consequence of Proposition C.9 - they do not always possess an expectation (for $\alpha \leq 1$) and exhibit infinite variance (for $\alpha < 2$). Thus, assumptions for several classical probability theorems are (partly) not fulfilled and therefore not applicable as they are for the Gaussian case $\alpha = 2$.

The following statement serves as support for the next section.

C.10 Lemma (Comparison of Parameters)
Assume that

- $X \sim S_\alpha(\sigma, \beta, \mu)$,
- $Y_1, \dots, Y_N$ are i.i.d. with $Y_1 \sim S_\alpha(\sigma_N, \beta_N, \mu_N)$ for $N \in \mathbb{N}$, and
  - $\sigma_N = N^{-1/\alpha}\sigma$,
  - $\beta_N = \beta$,
  - $\mu_n = \frac{1}{N}\mu$
- $Y := \sum_{n=1}^{N} Y_n$ for $N \in \mathbb{N}$, and
- $Y_L := \sum_{n=1}^{L} Y_n$ with $Y_L \sim S_\alpha(\sigma_L, \beta_L, \mu_L)$ for $L \in \mathbb{N}$, and
  - $\sigma_L = \left(\frac{L}{N}\right)^{1/\alpha} \sigma$
  - $\beta_L = \beta$,
  - $\mu_L = \frac{L}{N}\mu$.

Then, it holds

1. $Y_1 \sim S_\alpha\left(N^{-1/\alpha}\sigma, \beta, \frac{1}{N}\mu\right)$
2. $X \stackrel{d}{=} Y$
3. $Y_L \sim S_\alpha\left(\left(\frac{L}{N}\right)^{1/\alpha}\sigma, \beta, \frac{L}{N}\mu\right)$
4. $Y_1 \stackrel{d}{=} c_N X - d_N$ with

$$\begin{array}{llll} c_N := N^{-1/\alpha} & \text{and} & d_N := \mu\left(N^{-1/\alpha} - N^{-1}\right) & \text{for} \quad \alpha \neq 1. \\ c_N := \frac{1}{N} & \text{and} & d_N := \frac{1}{N}\frac{2}{\pi}\ln(N)\sigma\beta & \text{for} \quad \alpha = 1 \end{array}$$

Proof

a) By definition, the characteristic function of X holds

$$\Phi_X(\theta) = \mathbb{E}\left[\exp\left(i\theta X\right)\right] = \begin{cases} \exp\left(-\sigma^\alpha|\theta|^\alpha\left(1 - i\beta\,\mathrm{sgn}(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\right) + i\mu\theta\right) & \text{for } \alpha \neq 1 \\ \exp\left(-\sigma|\theta|\left(1 + i\beta\frac{2}{\pi}\,\mathrm{sgn}(\theta)\ln\left(|\theta|\right)\right) + i\mu\theta\right) & \text{for } \alpha = 1 \end{cases}. \tag{C.3}$$

b) For the characteristic function of Y, we have

$$\begin{aligned} \Phi_Y(\theta) &= \Phi_{\sum_{n=1}^{N} Y_n}(\theta) = \Phi_{Y_1}(\theta)\dots\Phi_{Y_N}(\theta) = \left(\Phi_{Y_1}(\theta)\right)^N \quad \text{as } Y_n,\ n = 1, \dots, N, \text{ are i.i.d.} \\ &= \begin{cases} \left[\exp\left(-\sigma_N^\alpha|\theta|^\alpha\left(1 - i\beta_N\,\mathrm{sgn}(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\right) + i\mu_N\theta\right)\right]^N & \text{for } \alpha \neq 1 \\ \left[\exp\left(-\sigma_N|\theta|\left(1 + i\beta_N\frac{2}{\pi}\,\mathrm{sgn}(\theta)\ln\left(|\theta|\right)\right) + i\mu_N\theta\right)\right]^N & \text{for } \alpha = 1 \end{cases} \\ &= \begin{cases} \exp\left(-N\sigma_N^\alpha|\theta|^\alpha\left(1 - i\beta_N\,\mathrm{sgn}(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\right) + iN\mu_N\theta\right) & \text{for } \alpha \neq 1 \\ \exp\left(-N\sigma_N|\theta|\left(1 + i\beta_N\frac{2}{\pi}\,\mathrm{sgn}(\theta)\ln\left(|\theta|\right)\right) + iN\mu_N\theta\right) & \text{for } \alpha = 1 \end{cases}. \end{aligned} \tag{C.4}$$

$$\text{With} \quad \begin{array}{rclcrcl} \sigma^\alpha & = & N\sigma_N^\alpha & \Rightarrow & \sigma_N & = & N^{-1/\alpha}\sigma \\ \beta & = & \beta_N & & & & \\ \mu & = & N\mu_N & \Rightarrow & \mu_N & = & \frac{1}{N}\mu \end{array}$$

it holds (C.3) = (C.4) which proofs

1) $Y_1 \sim S_\alpha\left(N^{-1/\alpha}\sigma, \beta, \frac{1}{N}\mu\right)$ and implies

2) $X \stackrel{d}{=} Y = \sum_{n=1}^{N} Y_n$ due to the equivalence of characteristic functions.

c) For the characteristic function of Y_L, we have

$$\begin{aligned}\Phi_{Y_L}(\theta) &= (\Phi_{Y_1}(\theta))^L \\ &= \begin{cases}\exp\left(-L\sigma_N^\alpha|\theta|^\alpha\left(1 - i\beta_N \operatorname{sgn}(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\right) + iL\mu_N\theta\right) & \text{for } \alpha \neq 1 \\ \exp\left(-L\sigma_N|\theta|\left(1 + i\beta_N\frac{2}{\pi}\operatorname{sgn}(\theta)\ln(|\theta|)\right) + iL\mu_N\theta\right) & \text{for } \alpha = 1\end{cases} . \end{aligned} \tag{C.5}$$

$$\begin{array}{lccccccc} \text{Set} & L\sigma_N^\alpha & = & L\left(N^{-1/\alpha}\sigma\right)^\alpha & = & \frac{L}{N}\sigma^\alpha & = & \left(\left(\frac{L}{N}\right)^{1/\alpha}\sigma\right)^\alpha \\ & \beta_N & = & \beta & & & & \\ & L\mu_N & = & \frac{L}{N}\mu & & & & \end{array}$$

which yields assertion 3) $Y_L = \sum_{n=1}^{L} Y_n \sim S_\alpha\left(\left(\frac{L}{N}\right)^{1/\alpha}\sigma, \beta, \frac{L}{N}\mu\right)$.

d) From Proposition C.8, number 6), we know that

$$\frac{1}{N^{1/\alpha}}\left(\sum_{n=1}^{N} X_n - \mu\left(N - N^{1/\alpha}\right)\right) \stackrel{d}{=} X$$

holds for $X \sim S_\alpha(\sigma, \beta, \mu)$ with $\alpha \neq 1$ and $X_1, \ldots, X_N$, $N \in \mathbb{N}$, as i.i.d. copies of X.

Straightforward calculation yields

$$\begin{aligned}\frac{1}{N^{1/\alpha}}\left(\sum_{n=1}^{N} X_n - \mu\left(N - N^{1/\alpha}\right)\right) &= \frac{1}{N^{1/\alpha}}\left(\sum_{n=1}^{N}\left(X_n - \mu\left(1 - \frac{N^{1/\alpha}}{N}\right)\right)\right) \\ &= \sum_{n=1}^{N}\left(\frac{1}{N^{1/\alpha}}X_n - \frac{1}{N^{1/\alpha}}\mu\left(1 - \frac{N^{1/\alpha}}{N}\right)\right) \\ &= \sum_{n=1}^{N}\left(N^{-1/\alpha}X_n - \mu\left(N^{-1/\alpha} - N^{-1}\right)\right) .\end{aligned}$$

Exploiting properties 2) and 1) of Proposition C.8, we receive the distribution of Y_1 as

$$\begin{aligned} & N^{-1/\alpha}X_1 \sim S_\alpha(N^{-1/\alpha}\sigma, \beta, N^{-1/\alpha}\mu) \text{ and} \\ & N^{-1/\alpha}X_1 - \mu\left(N^{-1/\alpha} - N^{-1}\right) \sim S_\alpha\left(N^{-1/\alpha}\sigma, \beta, N^{-1/\alpha}\mu - \mu\left(N^{-1/\alpha} - N^{-1}\right)\right) \\ \Leftrightarrow \quad & N^{-1/\alpha}X_1 - \mu\left(N^{-1/\alpha} - N^{-1}\right) \sim S_\alpha\left(N^{-1/\alpha}\sigma, \beta, \frac{1}{N}\mu\right) . \end{aligned}$$

e) From Proposition C.8, number 6), we know that

$$\frac{1}{N}\sum_{n=1}^{N} X_n - \frac{2}{\pi}\ln(N)\sigma\beta \stackrel{d}{=} X$$

holds for $X \sim S_1(\sigma, \beta, \mu)$ and $X_1, \ldots, X_N$, $N \in \mathbb{N}$, as i.i.d. copies of X.

Transformation of the sum implies

$$\sum_{n=1}^{N}\left(\frac{1}{N}X_n - \frac{1}{N}\frac{2}{\pi}\ln(N)\sigma\beta\right) \stackrel{d}{=} X .$$

Exploiting properties 2) and 1) of Proposition C.8, we receive that

$$\begin{aligned} & \frac{1}{N}X_1 \sim S_1\left(\frac{1}{N}\sigma, \beta, \frac{1}{N}\mu - \frac{1}{N}\frac{2}{\pi}\ln\left(\frac{1}{N}\right)\sigma\beta\right) \\ \Leftrightarrow \quad & \frac{1}{N}X_1 \sim S_1\left(\frac{1}{N}\sigma, \beta, \frac{1}{N}\mu + \frac{1}{N}\frac{2}{\pi}\ln(N)\sigma\beta\right) \end{aligned}$$

and therefore

$$\begin{aligned} & \frac{1}{N}X_n - \frac{1}{N}\frac{2}{\pi}\ln(N)\sigma\beta \sim S_1\left(\frac{1}{N}\sigma, \beta, \frac{1}{N}\mu + \frac{1}{N}\frac{2}{\pi}\ln(N)\sigma\beta - \frac{1}{N}\frac{2}{\pi}\ln(N)\sigma\beta\right) \\ \Leftrightarrow \quad & \frac{1}{N}X_n - \frac{1}{N}\frac{2}{\pi}\ln(N)\sigma\beta \sim S_\alpha\left(\frac{1}{N}\sigma, \beta, \frac{1}{N}\mu\right) \end{aligned}$$

which is the distribution of Y_1. ∎

Note that the first results are independent of the choice of $\alpha \in (0, 2]$. The implication is similar to number 6) of Proposition C.8. There, however, the sum of random variables is scaled to receive in turn the very cdf, i.e., $X = c_N \cdot \sum_{n=1}^{N} X_n - d_N$ with $X \stackrel{d}{=} X_n$, $n = 1, \ldots, N$.

In contrast, the intention is to describe a given distribution (of X) by means of a non-scaled sum of random variables, i.e., $X = \sum_{n=1}^{N} Y_n$. Lemma C.10 shows that this is possible if parameters of each summand $Y_{n=1,\ldots,N}$ are scaled versions of the parameters of the initial variable X. The result is that distributions are equivalent if the scaling factors (c_N, d_N) and the scaled parameters are chosen as given in Lemma C.10.

C.2 A Multivariate Extension

A general treatment on multivariate α-stable distributions is given in Samorodnitsky and Taqqu [57], Chapter 2. We analyze multivariate dependency structures, i.e., multivariate random variables with α-stable margins connected by copula functions.

For the one-dimensional case, the previous subsection shows that sums of α-stable variables are compoundable to single ones, Proposition C.8 and Lemma C.10. This property is extendable to a multivariate setup.

C.11 Notation

1. $X^d \sim F^d := S_{\alpha_d}(\sigma_d, \beta_d, \mu_d)$ for $d = 1, \ldots, D$,
2. $X := (X^1, \ldots, X^D)' \sim F$ with copula C in sense of Notation 2.4,
3. Y_n^d, $n = 1, \ldots, N$, as sequence of i.i.d. random variables for each $d = 1, \ldots, D$ with

$$Y_1^d \sim G^d := S_{\alpha_d}\left(N^{-1/\alpha_d}\sigma_d, \beta_d, \frac{1}{N}\mu_d\right) \quad ,$$

4. $Y_1 := (Y_1^1, \ldots, Y_1^D)' \sim G$,
5. Z_n^d, $n = 1, \ldots, N$, as sequence of i.i.d. random variables for each $d = 1, \ldots, D$ with

$$Z_1^d := c_N^d X^d + d_N^d$$

   in which c_N^d and d_N^d are given as in Lemma C.10, item 4, and
6. $Z_1 := (Z_1^1, \ldots, Z_1^D)' \sim H$.

Part I

1. As shown in Lemma C.10, it holds for each $d = 1, \dots, D$

   a) $Z_1^d \stackrel{d}{=} Y_1^d \sim G^d$ and

   b) $\sum_{n=1}^{N} Y_n^d \stackrel{d}{=} \sum_{n=1}^{N} Z_n^d \stackrel{d}{=} X^d \sim F^d$.

2. Due to their continuity, margins $X^1, \dots, X^D$ have an unique copula C, see Sklar's Theorem 2.2. As
$$T^d(X^d) := c_N^d X^d - d_N^d = Z_1^d \sim G^d, \ d = 1, \dots, D$$
is a strictly increasing transformation of X^d, margins $Z_1^1, \dots, Z_1^D$ also have copula C (invariance property, see Proposition 2.5). Thus, it holds
$$H(x^1, \dots, x^D) = C(G^1(x^1), \dots, G^D(x^D))$$
which is generally not equal to
$$F(x^1, \dots, x^D) = C(F^1(x^1), \dots, F^D(x^D)) \quad .$$

3. Further, $Y_1 = \left(Y_1^1, \dots, Y_1^D\right)'$ has a multivariate distribution function G with α-stable margins $G^1, \dots, G^D$. As they are continuous, their copula C_{Y_1} is unique, Theorem 2.2, i.e.,
$$G(x^1, \dots, x^D) = C_{Y_1}(G^1(x^1), \dots, G^D(x^D)) \quad .$$
Choosing $C_{Y_1} = C$, we receive an equivalence in distribution for Y_1 and Z_1, i.e.,
$$G(x^1, \dots, x^D) = C(G^1(x^1), \dots, G^D(x^D)) = H(x^1, \dots, x^D) \quad .$$
If and only if Y_1 has copula C, it holds $Y_1 \stackrel{d}{=} Z_1$ or equivalently $G = H = C(G^1, \dots, G^D)$.

Part II

We denote $Y_N := \left(\sum_{n=1}^{N} Y_n^1, \dots, \sum_{n=1}^{N} Y_n^D\right)' \sim G_N$ and $Z_N := \left(\sum_{n=1}^{N} Z_n^1, \dots, \sum_{n=1}^{N} Z_n^D\right)' \sim H_N$.

The sequences $\left(Y_n^1, \dots, Y_n^D\right)'$ and $\left(Z_n^1, \dots, Z_n^D\right)'$, $n = 1, \dots, N$, are i.i.d. with Y_1, $Z_1 \sim G$ and copula C. Sums of i.i.d. $\mathbb{R}^D$-valued random variables are added up component-wise
$$\sum_{n=1}^{N} \left(Y_n^1, \dots, Y_n^D\right)' = \left(\sum_{n=1}^{N} Y_n^1, \dots, \sum_{n=1}^{N} Y_n^D\right)' \quad .$$

Due to the stability property for α-stable random variables, we have
$$\sum_{n=1}^{N} Y_n^d \stackrel{d}{=} X^d \sim F^d$$
for $d = 1, \dots, D$. Again, G_N is a multivariate distribution function with continuous margins $F^1, \dots, F^D$, therefore its copula C_{Y_N} is unique.

If and only if we set $C_{Y_N} = C$, this implies an equivalence in distribution $G_N = F$, i.e.,
$$G_N(x^1, \dots, x^D) = C_{Y_N}((F^1(x^1), \dots, F^D(x^D)) = C(F^1(x^1), \dots, F^D(x^D)) = F(x^1, \dots, x^D)$$

or equivalently

$$\left(\sum_{n=1}^{N} Y_n^1, \ldots, \sum_{n=1}^{N} Y_n^D\right)' \stackrel{d}{=} (X^1, \ldots, X^D)' \quad .$$

The same argument applies for Z_N. Choosing $C_{Z_N} = C$, we receive $H_N = F$, i.e.,

$$H_N(x^1, \ldots, x^D) = C_{Z_N}(F^1(x^1), \ldots, F^D(x^D)) = C(F^1(x^1), \ldots, F^D(x^D)) = F(x^1, \ldots, x^D)$$

or equivalently

$$\left(\sum_{n=1}^{N} Z_n^1, \ldots, \sum_{n=1}^{N} Z_n^D\right)' \stackrel{d}{=} (X^1, \ldots, X^D)' \quad .$$

C.12 Remark
The distribution of sums of i.i.d. $\mathbb{R}^D$-valued random variables is the convolution of the distributions of its summands[10], i.e.,

$$G_N = G * \ldots * G \quad .$$

As $G = C(G^1, \ldots, G^D)$ - Part I with $C_{Y_1} = C$, and $G_N = F$ - Part II with $C_{Y_N} = C$, it must hold

$$G_N = G * \ldots * G = C(G^1, \ldots, G^D) * \ldots * C(G^1, \ldots, G^D) = C(F^1, \ldots, F^D) = F$$

and therefore with $F^d = G^d * \ldots * G^d$ - Part I, number 1

$$C(G^1, \ldots, G^D) * \ldots * C(G^1, \ldots, G^D) = C(G^1 * \ldots * G^1, \ldots, G^D * \ldots * G^D) \quad .$$

The same applies for random variables Z_1 and Z_N. Summarized, using identical dependence structures (i.e., the same copula) for the respective variables, equivalent outcomes can be produced

- by drawing N-times copula distributed random variables and aggregating via sums or
- by drawing N-times random variables with given margins and aggregating via copulas.

Moreover, due to their α-stability, it is insignificant whether random variables are transformed (Z_1^d) or have a different parametrization (Y_1^d) as long as they are in line with the superior variable X^d for each $d = 1, \ldots, D$.

D Proofs and Derivations

D.1 Proofs and Derivations - Part I

D.1.1 Proof of Proposition 2.5

Let C_X be the copula of a random variable $X = (X_1, \ldots, X_D)' \in \mathbb{R}^D$ with continuous margins $F_1, \ldots, F_D$. Further, let $T_1, \ldots, T_D$ be strictly increasing transformations. Then, the transformed random variable

$$T(X) := (T(X_1), \ldots, T_D(X_D))' \in \mathbb{R}^D$$

has copula C_X as well, i.e., it holds

$$C_X = C_{T(X)} \quad .$$

[10] See Satz 8.4. in [3], for instance.

Proof

Denote $X \sim F$ and $T(X) \sim G$ with multivariate distribution functions F and G. Let $T_d(X_d) \sim G_d$ and remember that $X_d \sim F_d$, $d = 1, \ldots, D$. It holds

$$G_d(x_d) = \mathbb{P}[T_d(X_d) \leq x_d] = \mathbb{P}[X_d \leq T_d^{-1}(x_d)] = F_d\left(T_d^{-1}(x_d)\right) \quad \text{for } d = 1, \ldots, D \quad .$$

As T_d are strictly increasing and F_d are continuous, margins G_d are continuous for $d = 1, \ldots, D$.

Due to Sklar's Theorem 2.2 - continuity of the margins implies uniqueness of the copula - copula $C_{T(X)}$ of $T(X) \sim G$ is unique. Precisely, we have

$$\begin{aligned} C_{T(X)}(G_1(x_1), \ldots, G_D(x_D)) &= G(x_1, \ldots, x_D) = \mathbb{P}\left[T_1(X_1) \leq x_1, \ldots, T_D(X_D) \leq x_D\right] \\ &= \mathbb{P}\left[X_1 \leq T_1^{-1}(x_1), \ldots, X_D \leq T_D^{-1}(x_D)\right] \\ &= F\left(T_1^{-1}(x_1), \ldots, T_D^{-1}(x_D)\right) \\ &= C_X\left(F_1\left(T_1^{-1}(x_1)\right), \ldots, F_D\left(T_D^{-1}(x_D)\right)\right) \\ &= C_X(G_1(x_1), \ldots, G_D(x_D)) \end{aligned}$$

or shortened $C_{T(X)} = C_X$. ∎

D.1.2 Clayton Copula Density - Equation (3.13)

For $D \geq 2$, the Clayton Copula holds

$$C_\theta^{Cl}(u_1, \ldots, u_D) = \left(\sum_{d=1}^{D}\left(u_d^{-\theta} - 1\right) + 1\right)^{-\frac{1}{\theta}} \quad .$$

The multivariate density function for the Clayton Copula is

$$c^{Cl}(u_1, \ldots, u_D, \theta) = \prod_{d=1}^{D}\left[(1 + (d-1)\theta) \cdot u_d^{-\theta-1}\right] \cdot \left(\sum_{d=1}^{D} u_d^{-\theta} - D + 1\right)^{-\frac{1}{\theta}-D} \quad . \tag{D.1}$$

Proof

Assertion: The dth derivative of the D-dimensional Clayton copula is

$$\frac{\partial^d C(u_1, \ldots, u_D)}{\partial u_1 \ldots u_d} = \prod_{i=1}^{d}\left[(1 + (d-1)\theta) \cdot u_i^{-\theta-1}\right] \cdot \left(\sum_{d=1}^{D} u_d^{-\theta} - D + 1\right)^{-\frac{1}{\theta}-d} \quad . \tag{D.2}$$

Proof by induction.

$d = 1$:

W.l.o.g. we take the derivative with respect to variable u_1.

$$\begin{aligned} \frac{\partial C(u_1, \ldots, u_D)}{\partial u_1} &= -\frac{1}{\theta}\left(\sum_{i=1}^{D} u_i^{-\theta} - D + 1\right)^{-\frac{1}{\theta}-1} \cdot (-\theta) u_1^{-\theta-1} \\ &= \prod_{i=1}^{1}\left[(1 + (d-1)\theta) \cdot u_i^{-\theta-1}\right] \cdot \left(\sum_{i=1}^{D} u_i^{-\theta} - D + 1\right)^{-\frac{1}{\theta}-1} \quad . \end{aligned}$$

$\underline{d-1 \to d}$:

W.l.o.g. we take the derivatives with respect to the variables $u_1, \ldots, u_d$.

$$
\begin{aligned}
&\frac{\partial}{\partial u_d}\left(\frac{\partial^{d-1} C(u_1,\ldots,u_D)}{\partial u_1 \ldots \partial u_{d-1}}\right) \\
=&\frac{\partial}{\partial u_d}\left(\prod_{i=1}^{d-1}\left[(1+(i-1)\theta)u_i^{-\theta-1}\right]\left(\sum_{i=1}^{D} u_i^{-\theta}-1\right)^{-\frac{1}{\theta}-(d-1)-1}\right) \\
=&\prod_{i=1}^{d-1}\left[(1+(i-1)\theta)u_i^{-\theta-1}\right]\left(\sum_{i=1}^{D} u_i^{-\theta}-1\right)^{-\frac{1}{\theta}-(d-1)-1}\cdot\left(-\frac{1}{\theta}-(d-1)\right)\cdot(-\theta)\cdot u_d^{-\theta-1} \\
=&\prod_{i=1}^{d}\left[(1+(i-1)\theta)\cdot u_i^{-\theta-1}\right]\cdot\left(\sum_{i=1}^{D} u_i^{-\theta}-D+1\right)^{-\frac{1}{\theta}-d} \quad .
\end{aligned}
$$

Set $d = D$ in (D.1) to receive the multivariate Clayton copula density function (D.2). ∎

D.1.3 Conditional Clayton Copula Distribution - Equation (2.10)

For the Clayton copula (D.1), we have

$$
C^{Cl}_{d|1,\ldots,d-1}(u_d|u_1,\ldots,u_{d-1}) = \left(1+\frac{u_d^{-\theta}-1}{\sum_{i=1}^{d-1} u_i^{-\theta}-(d-1)+1}\right)^{-\frac{1}{\theta}-(d-1)} \quad . \tag{D.3}
$$

$\underline{\text{Proof}}$

First, see that

$$
C(u_1,\ldots,u_d,1,\ldots,1) = \left(\sum_{i=1}^{d} u_i^{-\theta}+\sum_{i=d+1}^{D} 1-D+1\right)^{-\frac{1}{\theta}} = C(u_1,\ldots,u_d) \quad .
$$

Thus, it is

$$
\begin{aligned}
&C_{d|1,\ldots,d-1}(u_d,1,\ldots,1|u_1,\ldots,u_{d-1}) \\
=&\frac{\frac{\partial^{d-1} C(u_1,\ldots,u_d,1,\ldots,1)}{\partial u_1,\ldots\partial u_{d-1}}}{\frac{\partial^{d-1} C(u_1,\ldots,u_{d-1},1,\ldots,1)}{\partial u_1\ldots\partial u_{d-1}}} = \frac{\frac{\partial^{d-1} C(u_1,\ldots,u_d)}{\partial u_1,\ldots\partial u_{d-1}}}{\frac{\partial^{d-1} C(u_1,\ldots,u_{d-1})}{\partial u_1\ldots\partial u_{d-1}}} = \frac{\frac{\partial^{d-1} C(u_1,\ldots,u_d)}{\partial u_1,\ldots\partial u_{d-1}}}{c(u_1,\ldots,u_{d-1})} \\
=&\frac{\prod_{i=1}^{d-1}\left[(1+(i-1)\theta)\cdot u_i^{-\theta-1}\right]\cdot\left(\sum_{i=1}^{d} u_i^{-\theta}-d+1\right)^{-\frac{1}{\theta}-(d-1)}}{\prod_{i=1}^{d-1}\left[(1+(i-1)\theta)\cdot u_i^{-\theta-1}\right]\cdot\left(\sum_{i=1}^{d-1} u_i^{-\theta}-(d-1)+1\right)^{-\frac{1}{\theta}-(d-1)}} \\
=&\left(\frac{\sum_{i=1}^{d} u_i^{-\theta}-d+1}{\sum_{i=1}^{d-1} u_i^{-\theta}-(d-1)+1}\right)^{-\frac{1}{\theta}-(d-1)} = \left(1+\frac{u_d^{-\theta}-1}{\sum_{i=1}^{d-1} u_i^{-\theta}-(d-1)+1}\right)^{-\frac{1}{\theta}-(d-1)} \quad .
\end{aligned}
$$

∎

D.1.4 Inverted Conditional Clayton Copula - Equation (2.11)

For $d = 2, \ldots, D$, the conditional distribution functions $C^{Cl}_{d|1,\ldots,d-1}$ for the Clayton copula hold

$$C^{Cl}_{d|1,\ldots,d-1}(u_d|u_1,\ldots,u_{d-1};\theta) = \left(1 + \frac{u_d^{-\theta} - 1}{\sum_{i=1}^{d-1} u_i^{-\theta} - (d-1) + 1}\right)^{-\frac{1}{\theta}-(d-1)} .$$

Proof

Derivation of Equation (2.11) is the inversion of Equation (2.10) = Equation (D.3). We have to show that for each $d \in \{2, \ldots, D\}$ and given z_d and $u_1, \ldots, u_{d-1}$ it holds

$$u_d = \left(\left(z_d^{-\frac{\theta}{1+\theta(d-1)}} - 1\right) \cdot \left(\sum_{i=1}^{d-1} u_i^{-\theta} - (d-1) + 1\right) + 1\right)^{-\frac{1}{\theta}}$$

for the inversion of Equation (D.3)

$$z_d = C^{Cl}_{d|1,\ldots,d-1}(u_d|u_1,\ldots,u_{d-1};\theta) = \left(1 + \frac{u_d^{-\theta} - 1}{\sum_{i=1}^{d-1} u_i^{-\theta} - (d-1) + 1}\right)^{-\frac{1}{\theta}-(d-1)} .$$

Therefore, for any $d \in \{2, \ldots, D\}$ we have

$$\begin{aligned}
& z_d = \left(1 + \frac{u_d^{-\theta} - 1}{\sum_{i=1}^{d-1} u_i^{-\theta} - (d-1) + 1}\right)^{-\frac{1+\theta(d-1)}{\theta}} \\
\Leftrightarrow\ & z_d^{-\frac{\theta}{1+\theta(d-1)}} - 1 = \frac{u_d^{-\theta} - 1}{\sum_{i=1}^{d-1} u_i^{-\theta} - (d-1) + 1} \\
\Leftrightarrow\ & u_d^{-\theta} = \left(z_d^{-\frac{\theta}{1+\theta(d-1)}} - 1\right) \cdot \left(\sum_{i=1}^{d-1} u_i^{-\theta} - (d-1) + 1\right) + 1 \\
\Leftrightarrow\ & u_d = \left(z_d^{-\frac{\theta}{1+\theta(d-1)}} - 1\right) \cdot \left(\sum_{i=1}^{d-1} u_i^{-\theta} - (d-1) + 1\right)^{-\frac{1}{\theta}} .
\end{aligned}$$

■

D.2 Proofs and Derivations - Part II

D.2.1 Proof of Lemma 9.12

To show:

$$P(t,T) := \mathbb{E}\left[\mathbf{1}_{\{\tau>t\}}|\mathcal{H}_t\right] = \mathbb{E}\left[\exp\left(-\int_t^T \lambda_s ds\right)|\mathcal{G}_t\right]$$

Proof

$$P(t,T) := \mathbb{E}\left[\mathbf{1}_{\{\tau>T\}}|\mathcal{H}_t\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\{\tau>T\}}|\mathcal{G}\vee\mathcal{F}_t\right]|\mathcal{H}_t\right] = \mathbb{E}\left[\mathbb{Q}[N_T - N_t = 0|\mathcal{G}\vee\mathcal{F}_t]|\mathcal{H}_t\right]$$
$$= \mathbb{E}\left[\exp\left(-\int_t^T \lambda_s ds\right)|\mathcal{H}_t\right] = \mathbb{E}\left[\exp\left(-\int_t^T \lambda_s ds\right)|\mathcal{G}_t\vee\mathcal{F}_t\right]$$

Note that we also have

$$\mathbb{E}\left[\mathbf{1}_{\{\tau>T\}}|\mathcal{G}\vee\mathcal{F}_t\right] = \mathbb{Q}\left[\{\tau>T\}\cup\{\tau>t\}|\mathcal{G}\right] = \mathbb{Q}\left[\gamma_T > U|\{\gamma_t > U\}\vee\mathcal{G}\right]$$
$$= \frac{\mathbb{Q}\left[\gamma_T > U|\mathcal{G}\right]}{\mathbb{Q}\left[\gamma_t > U|\mathcal{G}\right]} = \frac{\mathbb{Q}\left[U\le\gamma_T|\mathcal{G}\right]}{\mathbb{Q}\left[U\le\gamma_t|\mathcal{G}\right]} = \frac{\gamma_T}{\gamma_t} = \exp\left(-\int_t^T \lambda_s ds\right)$$

which gives the same result. Here, we have exploited the content of information $\mathcal{F}_t$ and the construction of the default time, see Equation (9.2). Moreover, we have to get rid of the conditioning filtration $\mathcal{F}_t$. The default trigger U is independent of the filtration $\mathcal{G}$. This in turns means that U is independent of the σ-field $\sigma\left(\exp\left(-\int_t^T \lambda_s ds\right)\right)\vee\mathcal{G}_t$. Williams [65] showed that

$$\mathbb{E}\left[\exp\left(-\int_t^T \lambda_s ds\right)|\mathcal{G}_t\vee\sigma(U)\right] = \mathbb{E}\left[\exp\left(-\int_t^T \lambda_s ds\right)|\mathcal{G}_t\right] \quad .$$

With the inclusion $\mathcal{G}_t\subset\mathcal{G}_t\vee\mathcal{F}_t\subset\mathcal{G}_t\vee\sigma(U)$, it yields

$$\mathbb{E}\left[\exp\left(-\int_t^T \lambda_s ds\right)|\mathcal{G}_t\vee\mathcal{F}_t\right] = \mathbb{E}\left[\exp\left(-\int_t^T \lambda_s ds\right)|\mathcal{G}_t\right] \quad . \blacksquare$$

D.2.2 Proof of Proposition 9.17

1. Assume that up to time $t\in[0,T)$ no default has occurred, i.e., $\mathcal{E}_t = \mathcal{D}$ and $\tau(\mathcal{E}_t) > t$.

   The survival probability from t to T for **a single obligor** $d\in\{1,\dots,D\}$ is given by

$$Q_d(t,T) := \mathbb{E}\left[\mathbf{1}_{\{\tau^d>T\}}|\mathcal{H}_t\right] = \mathbb{E}\left[\frac{C\left(\gamma_t^1,\dots,\gamma_t^{d-1},\gamma_T^d,\gamma_t^{d+1},\dots,\gamma_t^D\right)}{C(\gamma_t)}|\mathcal{H}_t\right] \quad .$$

   The survival probability from t to T for **all obligors** is given by

$$Q(t,T) := \mathbb{E}\left[\mathbf{1}_{\{\tau>T\}}|\mathcal{H}_t\right] = \mathbb{E}\left[\frac{C(\gamma_T)}{C(\gamma_t)}|\mathcal{H}_t\right] \quad .$$

2. Assume that the set of defaulted obligors up to time $t\in[0,T)$ is $\mathcal{C}_t$, i.e., $\tau(\mathcal{C}_t)\le t$.

   The survival probability from t to T of **a single obligor** $d\in\mathcal{E}_t$ is given by

$$R_d(t,T) := \mathbb{E}\left[\mathbf{1}_{\{\tau^{d\in\mathcal{E}_t}>T\}}|\mathcal{H}_t\right] = \mathbb{E}\left[\frac{\frac{\partial^C}{\partial\lambda_t(\mathcal{C}_t)}C\left(\gamma_t(\mathcal{E}_t\setminus\{d\}),\gamma_T^d;\gamma_\tau(\mathcal{C}_t)\right)}{\frac{\partial^C}{\partial\lambda_t(\mathcal{C}_t)}C(\gamma_t(\mathcal{E}_t);\gamma_\tau(\mathcal{C}_t))}|\mathcal{H}_t\right] \quad .$$

The survival probability from t to T of **all non-defaulted obligors** is given by

$$R(t,T) := \mathbb{E}\left[\mathbf{1}_{\{\tau(\mathcal{E}_t)>T\}}|\mathcal{H}_t\right] = \mathbb{E}\left[\frac{\frac{\partial^C}{\partial\lambda_t(\mathcal{C}_t)}C\left(\gamma_T(\mathcal{E}_t);\gamma_\tau(\mathcal{C}_t)\right)}{\frac{\partial^C}{\partial\lambda_t(\mathcal{C}_t)}C(\gamma_t(\mathcal{E}_t);\gamma_\tau(\mathcal{C}_t))}|\mathcal{H}_t\right] \quad .$$

Proof

Assume that no default has occurred up to time t. For $Q_d(t,T)$, we have

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}_{\{\tau^d>T\}}|\mathcal{H}_t\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\{\tau^d>T\}}|\mathcal{G}\vee\mathcal{F}_t\right]|\mathcal{H}_t\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\{\tau^d>T\}\cap\{\tau>t\}}|\mathcal{G}\vee\mathcal{F}_t\right]|\mathcal{H}_t\right]\\
=&\mathbb{E}\left[\mathbb{Q}\left[\{\tau^d>T\}\cap\{\tau>t\}|\mathcal{G}\vee\mathcal{F}_t\right]|\mathcal{H}_t\right]\\
=&\mathbb{E}\left[\frac{\mathbb{Q}\left[(\tau^1,\ldots,\tau^{d-1},\tau^d,\tau^{d+1},\ldots,\tau^D)'>(t,\ldots,t,T,t\ldots,t)'|\mathcal{G}\vee\mathcal{F}_t\right]}{\mathbb{Q}\left[\tau>t|\mathcal{G}\vee\mathcal{F}_t\right]}|\mathcal{H}_t\right]\\
=&\mathbb{E}\left[\frac{\mathbb{Q}\left[(U^1,\ldots,U^{d-1},U^d,U^{d+1},\ldots,U^D)'<(\gamma_t^1,\ldots,\gamma_t^{d-1},\gamma_T^d,\gamma_t^{d+1},\ldots,\gamma_t^D)'|\mathcal{G}\vee\mathcal{F}_t\right]}{\mathbb{Q}\left[U<\gamma_t|\mathcal{G}\vee\mathcal{F}_t\right]}|\mathcal{H}_t\right]\\
=&\mathbb{E}\left[\frac{\mathbb{Q}\left[(U^1,\ldots,U^{d-1},U^d,U^{d+1},\ldots,U^D)'\leq(\gamma_t^1,\ldots,\gamma_t^{d-1},\gamma_T^d,\gamma_t^{d+1},\ldots,\gamma_t^D)'|\mathcal{G}\vee\mathcal{F}_t\right]}{\mathbb{Q}\left[U\leq\gamma_t|\mathcal{G}\vee\mathcal{F}_t\right]}|\mathcal{H}_t\right]\\
=&\mathbb{E}\left[\frac{C\left(\gamma_t^1,\ldots,\gamma_t^{d-1},\gamma_T^d,\gamma_t^{d+1},\ldots,\gamma_t^D\right)}{C(\gamma_t)}|\mathcal{H}_t\right] = \frac{\mathbb{E}\left[C\left(\gamma_t^1,\ldots,\gamma_t^{d-1},\gamma_T^d,\gamma_t^{d+1},\ldots,\gamma_t^D\right)|\mathcal{H}_t\right]}{C(\gamma_t)} \quad .
\end{aligned}$$

For $Q(t,T)$, we have

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}_{\{\tau>T\}}|\mathcal{H}_t\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\{\tau>T\}}|\mathcal{G}\vee\mathcal{F}_t\right]|\mathcal{H}_t\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\{\tau>T\}\cap\{\tau>t\}}|\mathcal{G}\vee\mathcal{F}_t\right]|\mathcal{H}_t\right]\\
=&\mathbb{E}\left[\mathbb{Q}\left[\{\tau>T\}\cap\{\tau>t\}|\mathcal{G}\vee\mathcal{F}_t\right]|\mathcal{H}_t\right] = \mathbb{E}\left[\frac{\mathbb{Q}\left[\{\tau>T\}|\mathcal{G}\vee\mathcal{F}_t\right]}{\mathbb{Q}\left[\tau>t|\mathcal{G}\vee\mathcal{F}_t\right]}|\mathcal{H}_t\right]\\
=&\mathbb{E}\left[\frac{\mathbb{Q}\left[U<\gamma_T|\mathcal{G}\vee\mathcal{F}_t\right]}{\mathbb{Q}\left[U<\gamma_t|\mathcal{G}\vee\mathcal{F}_t\right]}|\mathcal{H}_t\right] = \mathbb{E}\left[\frac{\mathbb{Q}\left[U\leq\gamma_T|\mathcal{G}\vee\mathcal{F}_t\right]}{\mathbb{Q}\left[U\leq\gamma_t|\mathcal{G}\vee\mathcal{F}_t\right]}|\mathcal{H}_t\right] = \frac{\mathbb{E}\left[C(\gamma_T)|\mathcal{H}_t\right]}{C(\gamma_t)} \quad .
\end{aligned}$$

Assume that the set of defaulted obligors up to time $t\in[0,T)$ is $\mathcal{C}_t$. Let $\tau(\mathcal{C}_t)=t(\mathcal{C}_t)$, i.e.,

$$\left(\tau^{\tilde{d}_1},\ldots,\tau^{\tilde{d}_C}\right) = \left(t^{\tilde{d}_1},\ldots,t^{\tilde{d}_C}\right)'$$

with $t^{\tilde{d}_c}\leq t$, $\tilde{d}_c\in\mathcal{C}_t$. For $R_d(t,T)$, we have

$$\begin{aligned}
\mathbb{E}\left[\mathbf{1}_{\{\tau^d>T\}}|\mathcal{H}_t\right] &= \mathbb{E}\left[\mathbf{1}_{\{\tau(\mathcal{E}_t\setminus\{d\})>t\}\cap\{\tau^d>T\}\cap\{\tau(\mathcal{C}_t)=t(\mathcal{C}_t)\}}|\mathcal{H}_t\right]\\
&= \mathbb{E}\left[\frac{\mathbb{Q}\left[\{\tau(\mathcal{E}_t\setminus\{d\})>t\}\cap\{\tau^d>T\}\cap\{\tau(\mathcal{C}_t)=t(\mathcal{C}_t)\}|\mathcal{G}\vee\mathcal{F}_t\right]}{\mathbb{Q}\left[\{\tau>t\}\cap\{\tau(\mathcal{C}_t)=t(\mathcal{C}_t)\}|\mathcal{G}\vee\mathcal{F}_t\right]}|\mathcal{H}_t\right]\\
&= \mathbb{E}\left[\frac{\mathbb{Q}\left[U(\mathcal{E}_t\setminus\{d\})<\gamma_t(\mathcal{E}_t\setminus\{d\}),U^d<\gamma_T^d;U(\mathcal{C}_t)=\gamma_\tau(\mathcal{C}_t)|\mathcal{G}\vee\mathcal{F}_t\right]}{\mathbb{Q}\left[U(\mathcal{E}_t)<\gamma_t(\mathcal{E}_t);U(\mathcal{C}_t)=\gamma_\tau(\mathcal{C}_t)|\mathcal{G}\vee\mathcal{F}_t\right]}|\mathcal{H}_t\right]\\
&= \mathbb{E}\left[\frac{\mathbb{Q}\left[U(\mathcal{E}_t\setminus\{d\})\leq\gamma_t(\mathcal{E}_t\setminus\{d\}),U^d\leq\gamma_T^d;U(\mathcal{C}_t)=\gamma_\tau(\mathcal{C}_t)|\mathcal{G}\vee\mathcal{F}_t\right]}{\mathbb{Q}\left[U(\mathcal{E}_t)\leq\gamma_t(\mathcal{E}_t);U(\mathcal{C}_t)=\gamma_\tau(\mathcal{C}_t)|\mathcal{G}\vee\mathcal{F}_t\right]}|\mathcal{H}_t\right]\\
&= \mathbb{E}\left[\frac{\frac{\partial^C}{\partial\lambda_t(\mathcal{C}_t)}C\left(\gamma_t(\mathcal{E}_t\setminus\{d\}),\gamma_T^d;\gamma_\tau(\mathcal{C}_t)\right)}{\frac{\partial^C}{\partial\lambda_t(\mathcal{C}_t)}C\left(\gamma_t(\mathcal{E}_t);\gamma_\tau(\mathcal{C}_t)\right)}|\mathcal{H}_t\right] \quad .
\end{aligned}$$

For $R(t,T)$,we have

$$\begin{aligned}
\mathbb{E}\left[\mathbf{1}_{\{\tau(\mathcal{E}_t)>T\}}|\mathcal{H}_t\right] &= \mathbb{E}\left[\mathbf{1}_{\{\tau(\mathcal{E}_t)>T\}\cap\{\tau(\mathcal{C}_t)=t(\mathcal{C}_t)\}}|\mathcal{H}_t\right] \\
&= \mathbb{E}\left[\frac{\mathbb{Q}\left[\{\tau(\mathcal{E}_t)>T\}\cap\{\tau(\mathcal{C}_t)=t(\mathcal{C}_t)\}|\mathcal{G}\vee\mathcal{F}_t\right]}{\mathbb{Q}\left[\{\tau(\mathcal{E}_t)>t\}\cap\{\tau(\mathcal{C}_t)=t(\mathcal{C}_t)\}|\mathcal{G}\vee\mathcal{F}_t\right]}|\mathcal{H}_t\right] \\
&= \mathbb{E}\left[\frac{\mathbb{Q}\left[U(\mathcal{E}_t)<\gamma_T(\mathcal{E}_t);U(\mathcal{C}_t)=\gamma_\tau(\mathcal{C}_t)|\mathcal{G}\vee\mathcal{F}_t\right]}{\mathbb{Q}\left[U(\mathcal{E}_t)<\gamma_t(\mathcal{E}_t);U(\mathcal{C}_t)=\gamma_\tau(\mathcal{C}_t)|\mathcal{G}\vee\mathcal{F}_t\right]}|\mathcal{H}_t\right] \\
&= \mathbb{E}\left[\frac{\mathbb{Q}\left[U(\mathcal{E}_t)\leq\gamma_T(\mathcal{E}_t);U(\mathcal{C}_t)=\gamma_\tau(\mathcal{C}_t)|\mathcal{G}\vee\mathcal{F}_t\right]}{\mathbb{Q}\left[U(\mathcal{E}_t)\leq\gamma_t(\mathcal{E}_t);U(\mathcal{C}_t)=\gamma_\tau(\mathcal{C}_t)|\mathcal{G}\vee\mathcal{F}_t\right]}|\mathcal{H}_t\right] \\
&= \mathbb{E}\left[\frac{\frac{\partial^C}{\partial\lambda_t(\mathcal{C}_t)}C\left(\gamma_T(\mathcal{E}_t);\gamma_\tau(\mathcal{C}_t)\right)}{\frac{\partial^C}{\partial\lambda_t(\mathcal{C}_t)}C(\gamma_t(\mathcal{E}_t);\gamma_\tau(\mathcal{C}_t))}|\mathcal{H}_t\right] \quad .
\end{aligned}$$

■

D.2.3 Proof of Lemma 10.6

Under Assumption 10.5 each marginal intensity process λ_t^d is $\mathbb{P}$-almost sure positive.

Proof

Let $d\in\{1,\ldots,D\}$ be fixed. As λ_t^d, $t\in[0,T]$, is constant on each interval $[t_n,t_{n+1})$, it is sufficient to prove that $\lambda_{t_n}^d$ is positive for each n. As λ_t is $\mathcal{G}_t$-adapted, conditioning on filtration $\mathcal{G}_{t_{n-1}}$ yields

$$\begin{aligned}
\mathbb{P}\left[\lambda_{t_n}^d\leq 0|\mathcal{G}_{t_{n-1}}\right] &= \mathbb{P}\left[\lambda_{t_{n-1}}^d+\Lambda_n^d\leq 0|\mathcal{G}_{t_{n-1}}\right] = \mathbb{P}\left[\Lambda_n^d\leq -\lambda_{t_{n-1}}^d|\mathcal{G}_{t_{n-1}}\right] \\
&= F_n^d\left(-\lambda_{t_{n-1}}^d|\mathcal{G}_{t_{n-1}}\right) \\
&= \frac{F^d\left(-\lambda_{t_{n-1}}^d|\mathcal{G}_{t_{n-1}}\right)-F^d\left(-\lambda_{t_{n-1}}^d|\mathcal{G}_{t_{n-1}}\right)}{1-F^d\left(-\lambda_{t_{n-1}}^d|\mathcal{G}_{t_{n-1}}\right)}=0 \quad .
\end{aligned}$$

for $n=1,\ldots,N$. We naturally have $F^d\left(-\lambda_{t_{n-1}}^d|\mathcal{G}_{t_{n-1}}\right)<1$. Otherwise, the distribution is not admissible in sense of Definition 10.3. ■

D.2.4 Proof of Positivity of Equation (10.12)

Under Assumption 10.5 each marginal intensity λ_t^d constructed by Equation (10.12) is $\mathbb{P}$-almost sure positive.

Proof

Define $a_n^d:=-\lambda_{t_{n-1}}^d$, $\mathcal{G}_n:=\mathcal{G}_{t_{n-1}}$. Remember that $\Lambda_n^d=\left(F^d\right)^{-1}\left[V_n^d\left(1-F^d(-a_n^d)\right)+F^d(-a_n^d)\right]$. As in D.2.3, it is sufficient to show that $\lambda_{t_n}>0$ under conditioning on $\mathcal{G}_n$ for each $n=0,\ldots,N$.

$$\begin{aligned}
\mathbb{P}[\lambda_{t_n}^d\leq 0|\mathcal{G}_n] &= \mathbb{P}[\lambda_{t_{n-1}}^d+\Lambda_n^d\leq 0|\mathcal{G}_n] \\
&= \mathbb{P}\left[\left(F^d\right)^{-1}\left[V_n^d\left(1-F^d(-a_n^d)\right)+F^d(-a_n^d)\right]\leq -a_n^d|\mathcal{G}_n\right] \\
&= \mathbb{P}\left[V_n^d\left(1-F^d(-a_n^d)\right)+F^d(-a_n^d)\leq F^d(-a_n^d)|\mathcal{G}_n\right] \\
&= \mathbb{P}\left[V_n^d\leq\frac{0}{1-F^d(-a_n^d)}|\mathcal{G}_n\right]=0
\end{aligned}$$

as F^d is admissible and V_n^d, $n=1,\ldots,N$, are continuously uniformly distributed.

■

D.2.5 Proof of Lemma 10.8

Under Assumptions 10.5, 10.7, and relevant filtrations $\mathcal{G}_m$ and $\mathcal{G}_n$, the increments

$$\Lambda_m^d = \left(F^d\right)^{-1}\left[V_m^d\left(\kappa_m^d - F^d(-a_m^d)\right) + F^d(-a_m^d)\right]$$

and

$$\Lambda_n^d = \left(F^d\right)^{-1}\left[V_n^d\left(\kappa_n^d - F^d(-a_n^d)\right) + F^d(-a_n^d)\right]$$

derived as in Equation (10.12) are *pseudo-independent* for all $m < n$, i.e., it holds

$$\mathbb{P}\left[\Lambda_m^d \leq x_1, \Lambda_n^d \leq x_2 | \mathcal{G}_m\right] = \mathbb{P}\left[\Lambda_m^d \leq x_1 | \mathcal{G}_m\right] \cdot \mathbb{E}\left[\mathbb{P}\left[\Lambda_n^d \leq x_2 | \mathcal{G}_n\right] | \mathcal{G}_m\right] \quad .$$

Proof

Remember that it holds $\mathbb{P}[X \leq x|\mathcal{G}] = \mathbb{E}\left[\mathbf{1}_{\{X \leq x\}}|\mathcal{G}\right]$. Conditioning on filtration $\mathcal{G}_m$ provides

$$\begin{aligned}
& \mathbb{P}\left[\Lambda_m^d \leq x_1, \Lambda_n^d \leq x_2 | \mathcal{G}_m\right] = \mathbb{E}\left[\mathbf{1}_{\{\Lambda_m^d \leq x_1, \Lambda_n^d \leq x_2\}} | \mathcal{G}_m\right] \\
= \; & \mathbb{E}\left[\mathbf{1}_{\left\{(F^d)^{-1}[V_m^d(\kappa_m^d - F^d(-a_m^d)) + F^d(-a_m^d)] \leq x_1, (F^d)^{-1}[V_n^d(\kappa_n^d - F^d(-a_n^d)) + F^d(-a_n^d)] \leq x_2\right\}} | \mathcal{G}_m\right] \\
= \; & \mathbb{E}\left[\mathbf{1}_{\left\{V_m^d \leq \frac{F^d(x_1) - F^d(-a_m^d)}{\kappa_m^d - F^d(-a_m^d)}, V_n^d \leq \frac{F^d(x_2) - F^d(-a_n^d)}{\kappa_n^d - F^d(-a_n^d)}\right\}} | \mathcal{G}_m\right]
\end{aligned}$$

As $V_m \perp V_n$, we receive

$$\begin{aligned}
\mathbb{P}\left[\Lambda_m^d \leq x_1, \Lambda_n^d \leq x_2 | \mathcal{G}_m\right] &= \mathbb{E}\left[\mathbf{1}_{\left\{V_m^d \leq \frac{F^d(x_1) - F^d(-a_m^d)}{\kappa_m^d - F^d(-a_m^d)}\right\}} \cdot \mathbf{1}_{\left\{V_n^d \leq \frac{F^d(x_2) - F^d(-a_n^d)}{\kappa_n^d - F^d(-a_n^d)}\right\}} | \mathcal{G}_m\right] \\
&= \mathbb{E}\left[\mathbf{1}_{\left\{V_m^d \leq \frac{F^d(x_1) - F^d(-a_m^d)}{\kappa_m^d - F^d(-a_m^d)}\right\}} | \mathcal{G}_m\right] \cdot \mathbb{E}\left[\mathbf{1}_{\left\{V_n^d \leq \frac{F^d(x_2) - F^d(-a_n^d)}{\kappa_n^d - F^d(-a_n^d)}\right\}} | \mathcal{G}_m\right] \quad .
\end{aligned}$$

Due to iterated conditioning and $\mathcal{G}_m \subseteq \mathcal{G}_n$, this can be rewritten as

$$\begin{aligned}
& \mathbb{P}\left[\Lambda_m^d \leq x_1, \Lambda_n^d \leq x_2 | \mathcal{G}_m\right] \\
= \; & \mathbb{P}\left[V_m^d \leq \frac{F^d(x_1) - F^d(-a_m^d)}{\kappa_m^d - F^d(-a_m^d)} | \mathcal{G}_m\right] \cdot \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\left\{V_n^d \leq \frac{F^d(x_2) - F^d(-a_n^d)}{\kappa_n^d - F^d(-a_n^d)}\right\}} | \mathcal{G}_n\right] | \mathcal{G}_m\right] \\
= \; & \mathbb{P}\left[\Lambda_m^d \leq x_1 | \mathcal{G}_m\right] \cdot \mathbb{E}\left[\mathbb{P}\left[\Lambda_n^d \leq x_2 | \mathcal{G}_n\right] | \mathcal{G}_m\right]
\end{aligned}$$

which proves the statement. Note that the random variable $\mathbf{1}_{\left\{V_m^d \leq \frac{F^d(x_1) - F^d(-a_m^d)}{\kappa_m^d - F^d(-a_m^d)}\right\}}$ is $\mathcal{G}_n$-measurable. Therefore, conditioning on $\mathcal{G}_n$ implies

$$\begin{aligned}
\mathbb{P}\left[\Lambda_m^d \leq x_1, \Lambda_n^d \leq x_2 | \mathcal{G}_n\right] &= \mathbf{1}_{\left\{V_m^d \leq \frac{F^d(x_1) - F^d(-a_m^d)}{\kappa_m^d - F^d(-a_m^d)}\right\}} \cdot \mathbb{E}\left[\mathbf{1}_{\left\{V_n^d \leq \frac{F^d(x_2) - F^d(-a_n^d)}{\kappa_n^d - F^d(-a_n^d)}\right\}} | \mathcal{G}_n\right] \\
&= \mathbf{1}_{\{\Lambda_m^d \leq x_1\}} \cdot \mathbb{P}\left[\Lambda_n^d \leq x_2 | \mathcal{G}_n\right] \quad .
\end{aligned}$$

■

D.2.6 Proof of Proposition 10.20 - Part II

Let $t \in (0, T]$, $\mathcal{E}_t$, $\mathcal{C}_t$ and $\mathcal{G}_n$ be given. The original random variables

$$\begin{aligned}
& \Lambda = \Lambda(t) \sim F \text{ with } \Lambda(t) = (\Lambda(\mathcal{E}_t), \Lambda(\mathcal{C}_t)) \text{ given } \Lambda(\mathcal{C}_t) = 0 \\
\text{and} \quad & \Phi = \Phi(t) \sim H \text{ with } \Phi(t) = (\Phi(\mathcal{E}_t), \Phi(\mathcal{C}_t)) \text{ given } \Phi(\mathcal{C}_t) = 0
\end{aligned}$$

and the random change variables

$$\Lambda_n(t) \sim F_n(t) \quad \text{and} \quad \Phi_\tau(t) \sim H_\tau(t)$$

have copulas $K(t)$ and $L(t)$, i.e., it holds

$$\begin{aligned} & F\left(\Delta\lambda_n^{d_1}(t), \ldots, \Delta\lambda_n^{d_E}(t), \Delta\lambda_n^{\tilde{d}_1}(t) = 0, \ldots, \Delta\lambda_n^{\tilde{d}_C}(t) = 0\right) \\ & \qquad = \quad K\left(F^{d_1}\left(\Delta\lambda_n^{d_1}(t)\right), \ldots, F^{d_E}\left(\Delta\lambda_n^{d_E}(t)\right), F^{\tilde{d}_1}(0,t), \ldots, F^{\tilde{d}_C}(0,t)\right) \quad , \end{aligned}$$

$$\begin{aligned} & F_n\big(\Delta\lambda_n^{d_1}(t), \ldots, \Delta\lambda_n^{d_E}(t), \Delta\lambda_n^{\tilde{d}_1}(t) = 0, \ldots, \Delta\lambda_n^{\tilde{d}_C}(t) = 0\big) \\ & \qquad = \quad K\left(F_n^{d_1}\left(\Delta\lambda_n^{d_1}(t)\right), \ldots, F_n^{d_E}\left(\Delta\lambda_n^{d_E}(t)\right), F_n^{\tilde{d}_1}(0,t), \ldots, F^{\tilde{d}_C}(0,t)\right) \end{aligned}$$

for the random change variables and

$$\begin{aligned} & H\left(\Delta\phi_\tau^{d_1}(t), \ldots, \Delta\phi_\tau^{d_E}(t), \Delta\phi_\tau^{\tilde{d}_1}(t) = 0, \ldots, \Delta\phi_\tau^{\tilde{d}_C}(t) = 0\right) \\ & \qquad = \quad L\left(H_\tau^{d_1}\left(\Delta\phi_\tau^{d_1}(t)\right), \ldots, H_\tau^{d_E}\left(\Delta\phi_\tau^{d_E}(t)\right), H_\tau^{\tilde{d}_1}(0,t), \ldots, H^{\tilde{d}_C}(0,t)\right) \quad , \end{aligned}$$

$$\begin{aligned} & H_\tau\left(\Delta\phi_\tau^{d_1}(t), \ldots, \Delta\phi_\tau^{d_E}(t), \Delta\phi_\tau^{\tilde{d}_1}(t) = 0, \ldots, \Delta\phi_\tau^{\tilde{d}_C}(t) = 0\right) \\ & \qquad = \quad L\left(H_\tau^{d_1}\left(\Delta\phi_\tau^{d_1}(t)\right), \ldots, H_\tau^{d_E}\left(\Delta\phi_\tau^{d_E}(t)\right), H_\tau^{\tilde{d}_1}(0,t), \ldots, H^{\tilde{d}_C}(0,t)\right) \end{aligned}$$

for the default change variables.

Proof - Part II

For the second part, let - w.l.o.g. - $t \in (0,\tau)$ be fixed and $\mathcal{E}_t$, $\mathcal{C}_t$ and $\mathcal{H}_t$ be given. For the original random variable $\Phi = \Phi(t) \sim F(t)$ given $\Phi(\mathcal{C}_t) = 0$, the next random change at τ, and under $\mathcal{H}_t$, we have

$$\begin{aligned} & H\left(\Delta\phi_n(t)\right) \\ & = H\left(\Delta\phi^{d_1}(t), \ldots, \Delta\phi^{d_E}(t), \Delta\phi^{\tilde{d}_1}(t) = 0, \ldots, \Delta\phi^{\tilde{d}_C}(t) = 0\right) \\ & = \mathbb{P}\left[\Phi(\mathcal{E}_t) \leq \Delta\phi_n(\mathcal{E}_t) \mid \{\Phi(\mathcal{C}_t) = 0\} \vee \mathcal{H}_t\right] \\ & = \mathbb{P}\left[\Phi^{d_1}(t) \leq \Delta\phi_n^{d_1}(t), \ldots, \Phi^{d_E}(t) \leq \Delta\phi_n^{d_E}(t) \mid \left\{\Phi^{\tilde{d}_1}(t) = 0\right\} \vee \ldots \vee \left\{\Phi^{\tilde{d}_C}(t) = 0\right\} \vee \mathcal{H}_t\right] \\ & = \mathbb{P}\Bigg[H^{d_1}\left(\Phi^{d_1}(t)\right) \leq H^{d_1}\left(\Delta\phi_n^{d_1}(t)\right), \ldots, H^{d_E}\left(\Phi^{d_E}(t)\right) \leq H^{d_E}\left(\Delta\phi_n^{d_E}(t)\right) \mid \\ & \qquad \left\{H^{\tilde{d}_1}\left(\Phi^{\tilde{d}_1}(t)\right) = H^{\tilde{d}_1}(0,t)\right\} \vee \ldots \vee \left\{H^{\tilde{d}_C}\left(\Phi^{\tilde{d}_C}(t)\right) = H^{\tilde{d}_C}(0,t)\right\} \vee \mathcal{H}_t\Bigg] \quad . \end{aligned}$$

As before with $H^d\left(\Phi^d\right) = H^d\left(\Phi^d(t)\right) = W^d(t) = W^d$, this implies

$$\begin{aligned} & = \mathbb{P}\Bigg[W^{d_1}(t) \leq H^{d_1}\left(\Delta\phi_n^{d_1}(t)\right), \ldots, W^{d_E}(t) \leq H^{d_E}\left(\Delta\phi_n^{d_E}(t)\right) \mid \\ & \qquad \left\{W^{\tilde{d}_1}(t) = H^{\tilde{d}_1}(0,t)\right\} \vee \ldots \vee \left\{W^{\tilde{d}_C}(t) = H^{\tilde{d}_C}(0,t)\right\} \vee \mathcal{H}_t\Bigg] \\ & = \frac{\partial^C}{\partial v(\mathcal{C}_t)} K\left(H^{d_1}\left(\Delta\phi_n^{d_1}(t)\right), \ldots, H^{d_E}\left(\Delta\phi_n^{d_E}(t)\right); v(\mathcal{C}_t)\right)\Big|_{v(\mathcal{C}_t)=\left(H^{\tilde{d}_1}(0,t),\ldots,H^{\tilde{d}_C}(0,t)\right)'} \\ & = K\left(H^{d_1}\left(\Delta\phi_n^{d_1}(t)\right), \ldots, H^{d_E}\left(\Delta\phi_n^{d_E}(t)\right), H^{\tilde{d}_1}(0,t), \ldots, H^{\tilde{d}_C}(0,t)\right) \quad . \end{aligned}$$

For the model random variable $\Phi_n = \Phi_n(t) \sim H_n$ given $\Phi_n(\mathcal{C}_t) = 0$, it holds

$$
\begin{aligned}
& H_n\left(\Delta\phi_n(t)\right) \\
&= H_n\left(\Delta\phi^{d_1}(t), \ldots, \Delta\phi^{d_E}(t), \Delta\phi^{\tilde{d}_1}(t) = 0, \ldots, \Delta\phi^{\tilde{d}_C}(t) = 0\right) \\
&= \mathbb{P}\left[\Phi_n(\mathcal{E}_t) \leq \Delta\phi_n(\mathcal{E}_t) \mid \{\Phi_n(\mathcal{C}_t) = 0\} \vee \mathcal{H}_t\right] \\
&= \mathbb{P}\left[\Phi_n^{d_1}(t) \leq \Delta\phi_n^{d_1}(t), \ldots, \Phi_n^{d_E}(t) \leq \Delta\phi_n^{d_E}(t) \mid \left\{\Phi_n^{\tilde{d}_1}(t) = 0\right\} \vee \ldots \vee \left\{\Phi_n^{\tilde{d}_C}(t) = 0\right\} \vee \mathcal{H}_t\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \mathbb{P}\Bigg[\left(H^{d_1}\right)^{-1}\left[W_n^{d_1}(t)\left(\kappa_t^{d_1} - H^{d_1}\left(-\phi_{t^-}^{d_1}|\mathcal{H}_t\right)\right) + H^{d_1}\left(-\phi_{t^-}^{d_1}|\mathcal{H}_t\right)\right] \leq \Delta\phi_n^{d_1}(t), \ldots, \\
&\qquad \left(H^{d_E}\right)^{-1}\left[W_n^{d_E}(t)\left(\kappa_t^{d_E} - H^{d_E}\left(-\phi_{t^-}^{d_E}|\mathcal{H}_t\right)\right) + H^{d_E}\left(-\phi_{t^-}^{d_E}|\mathcal{H}_t\right)\right] \leq \Delta\phi_n^{d_E}(t) \mid \\
&\qquad \left\{\left(H^{\tilde{d}_1}\right)^{-1}\left[W_n^{\tilde{d}_1}(t)\left(\kappa_t^{\tilde{d}_1} - H^{\tilde{d}_1}\left(-\phi_{t^-}^{\tilde{d}_1}|\mathcal{H}_t\right)\right) + H^{\tilde{d}_1}\left(-\phi_{t^-}^{\tilde{d}_1}|\mathcal{H}_t\right)\right] = 0\right\} \vee \ldots \vee \\
&\qquad \left\{\left(H^{\tilde{d}_C}\right)^{-1}\left[W_n^{\tilde{d}_C}(t)\left(\kappa_t^{\tilde{d}_C} - H^{\tilde{d}_C}\left(-\phi_{t^-}^{\tilde{d}_C}|\mathcal{H}_t\right)\right) + H^{\tilde{d}_C}\left(-\phi_{t^-}^{\tilde{d}_C}|\mathcal{H}_t\right)\right] = 0\right\} \vee \mathcal{H}_t\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
&= \mathbb{P}\Bigg[W_n^{d_1}(t) \leq \frac{H^{d_1}(\Delta\phi_n^{d_1}(t)) - H^{d_1}\left(-\phi_{t^-}^{d_1}|\mathcal{H}_t\right)}{\kappa_t^{d_1} - H^{d_1}\left(-\phi_{t^-}^{d_1}|\mathcal{H}_t\right)}, \ldots, \\
&\qquad\qquad W_n^{d_E}(t) \leq \frac{H^{d_E}(\Delta\phi_n^{d_E}(t)) - H^{d_E}\left(-\phi_{t^-}^{d_E}|\mathcal{H}_t\right)}{\kappa_t^{d_E} - H^{d_E}\left(-\phi_{t^-}^{d_E}|\mathcal{H}_t\right)} \mid \\
&\qquad \left\{W_n^{\tilde{d}_1}(t) = \frac{H^{\tilde{d}_1}(0) - H^{\tilde{d}_1}\left(-\phi_{t^-}^{\tilde{d}_1}|\mathcal{H}_t\right)}{\kappa_t^{\tilde{d}_1} - H^{\tilde{d}_1}\left(-\phi_{t^-}^{\tilde{d}_1}|\mathcal{H}_t\right)}\right\} \vee \ldots \vee \\
&\qquad\qquad \left\{W_n^{\tilde{d}_C}(t) = \frac{H^{\tilde{d}_C}(0) - H^{\tilde{d}_C}\left(-\phi_{t^-}^{\tilde{d}_C}|\mathcal{H}_t\right)}{\kappa_t^{\tilde{d}_C} - H^{\tilde{d}_C}\left(-\phi_{t^-}^{\tilde{d}_C}|\mathcal{H}_t\right)}\right\} \vee \mathcal{H}_t\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
&= \mathbb{P}\Bigg[W_n^{d_1}(t) \leq H_n^{d_1}\left(\Delta\phi_n^{d_1}(t)\right), \ldots, W_n^{d_E}(t) \leq H_n^{d_E}\left(\Delta\phi_n^{d_E}(t)\right) \mid \\
&\qquad \left\{W_n^{\tilde{d}_1}(t) = H_{\tau^{\tilde{d}_1}}^{\tilde{d}_1}(0,t)\right\} \vee \ldots \vee \left\{W_n^{\tilde{d}_C}(t) = H_{\tau^{\tilde{d}_C}}^{\tilde{d}_C}(0,t)\right\} \vee \mathcal{H}_t\Bigg] \quad .
\end{aligned}
$$

For the last equation note that $H_n^{\tilde{d}_c}(0,t) = H_{\tau^{\tilde{d}_c}}^{\tilde{d}_c}(0,t)$ due to $\phi_{t_n^-}^{\tilde{d}_c} = \phi_{\tau^{\tilde{d}_c}}^{\tilde{d}_c}$ for $c = 1, \ldots, C$.

Further, we get

$$
\begin{aligned}
&= \frac{\partial^C}{\partial w_n(\mathcal{C}_t)} K\left(H_n^{d_1}\left(\Delta\phi_n^{d_1}(t)\right), \ldots, H_n^{d_E}\left(\Delta\phi_n^{d_E}(t)\right); w_n(\mathcal{C}_t)\right)\Big|_{w_n(\mathcal{C}_t) = \left(H_{\tau^{\tilde{d}_1}}^{\tilde{d}_1}(0,t), \ldots, H_{\tau^{\tilde{d}_C}}^{\tilde{d}_C}(0,t)\right)'} \\
&= K\left(H_n^{d_1}\left(\Delta\phi_n^{d_1}(t)\right), \ldots, H_n^{d_E}\left(\Delta\phi_n^{d_E}(t)\right), H_{\tau^{\tilde{d}_1}}^{\tilde{d}_1}(0,t), \ldots, H_{\tau^{\tilde{d}_C}}^{\tilde{d}_C}(0,t)\right) \quad .
\end{aligned}
$$

which proves the second part. ∎

Bibliography

[1] Kjersti Aas, Henrik Bakken, Claudia Czado, and Arnoldo Frigessi. Pair-Copula Constructions of Multiple Dependence. Technical report, Norsk Regnesentral, Oslo, August 2006.

[2] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions*, volume 55. Dover Publications, 1972.

[3] Heinz Bauer. *Wahrscheinlichkeitstheorie*. de Gruyter, 2002.

[4] Tomasz R. Bielecki and Marek Rutkowski. *Credit Risk: Modeling, Valuation and Hedging*. Springer, 2002.

[5] P. Billingsley. *Convergence of probabilities*. J. Wiley & Sons, 1969.

[6] Fisher Black and Myron Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[7] Christian Bluhm, Ludger Overbeck, and Christoph Wagner. *(An Introduction to) Credit Risk Modeling*. Chapman & Hall/CRC, 2003.

[8] Eric Bouyé. Copulas for Finance - A Reading Guide and Some Applications. Technical report, Financial Econometrics Research Centre City University Business School, London, July 2000.

[9] Damiano Brigo and Fabio Mercurio. *Interest Rate Models - Theory and Practice*. Springer, 2nd edition, 2006.

[10] Deutsche Bundesbank. Validierungsansätze für interne Ratingsysteme. Technical Report Monatsbericht September 2003, Deutsche Bundesbank, Franfurt, September 2003.

[11] Christian Cech. Die IRB Formel. Technical report, University of Applied Science of bfi Vienna, March 2004.

[12] Umberto Cherubini, Elisa Luciano, and Walter Vecchiato. *Copula Methods in Finance*. Wiley, 2004.

[13] Gerd Christoph and Werner Wolf. *Convergence Theorems with a Stable Limit Law*, volume 70 of *Mathematical Research*. Akademie Verlag, 1992.

[14] Rama Cont and Peter Tankov. *Financial Modelling With Jump Processes*. Chapman & Hall / CRC, 2004.

[15] Alexandra da Costa Dias. *Copula Inference For Finance And Insurance*. PhD thesis, ETH Zürich, 2004. Doctoral Thesis ETH No. 15283.

[16] Paul Deheuvels. La Fonction de Dépendence empirique et ses Propriétés. Un Test non paramétrique d'indépendence. *Académie Royale de Belgique, Bulletin de la Classe des Sciences*, 65(5):274–292, 1979.

[17] Paul Deheuvels. A Nonparametric Test for Independence. Technical Report No.2, Institut de Statistique, Université de Paris, 1981.

[18] Luc Devroye. *Non-Uniform Random Variate Generation*. Springer, 1986.

[19] Darrell Duffie and Kenneth J. Singleton. Modeling term structures of defaultable bonds. *The Review of Financial Studies*, 12(4):687–720, 1999.

[20] Darrell Duffie and Kenneth J. Singleton. *Credit Risk, Pricing, Measurement and Management*. Princeton, 2003.

[21] Scott R. Eliason. *Maximum Likelihood Estimation, Logic and Practice*. University Paper Series on Quantitative Applications in the Social Science, series no. 07-069. Sage Publication, Inc., Newbury Park, CA : Sage, 1993.

[22] Paul Embrechts, Filip Lindskog, and Alexander J. McNeil. Modelling Dependence with Copulas and Applications to Risk Management. Technical report, Department of Mathematics, ETHZ, September 2001.

[23] K.-T. Fang, S. Kotz, and K.-W. Ng. *Symmetric Multivariate and Related Distributions*. Chapman & Hall / CRC, 1990.

[24] Wiliam Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. John Wiley & Sons, 2 edition, 1971.

[25] Jean-David Fermanian and Olivier Scaillet. Nonparametric Estimation of Copulas for Times Series. *Journal of Risk*, 5(4), 2003.

[26] Christian Genest. Frank's family of bivariate distributions. *Biometrika*, 74:549–555, 1987.

[27] Christian Genest, Kilani Ghoudi, and Louis-Paul Rivest. A Semiparametric Estimation Procedure of Dependence in Multivariate Families of Distributions. *Biometrika*, 82(2):543–552, 1995.

[28] Christian Genest and Louis-Paul Rivest. Statistical Inference Procedures for Bivariate Archimedean Copulas. *Journal of the American Statistical Association*, 88:1034–1043, 1993.

[29] Christian Genest and Louis-Paul Rivest. On the multivariate probability integral transformation. *Statistics and Probability Letters*, 53:391–399, 2001.

[30] Boris Vladimirovich Gnedenko and Andrey Nikolaevich Kolmogorov. *Limit Distributions For Sums of Independent Random Variables*. Addison-Wesley Publishing Company, revised from 1954 edition, 1968. translated by K.L. Chung, Stanford University.

[31] Michael B. Gordy. A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules. *Journal of Financial Intermediation*, 12(3):192–232, 2003.

[32] Jan Grandell. *Doubly Stochastic Poisson Processes*. Lecture Notes In Mathematics 529. Springer, 1976.

[33] Michael Günther and Ansgar Jüngel. *Finanzderivate mit MATLAB*. Vieweg, Wiesbaden, 2003.

[34] H. Hult and Filip Lindskog. Multivariate Extremes, Aggregation and Dependence in Elliptical Distributions. Technical report, 2002.

[35] Jean Jacod and Albert N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, 2nd edition, 2002.

[36] Robert A. Jarrow and Stuart M. Turnbull. Pricing Derivatives on Financial Securities Subject to Credit Risk. *Journal of Finance*, 50:53–85, 1995.

[37] Harry Joe. *Multivariate Models and Dependence Concepts*. Chapman & Hall/CRC, 1997.

[38] B. M. Golam Kibra and Anwar H. Joarder. A Short Review of Multivariate t-Distribution. *Journal of Statistical Research*, 40, No. 1:59–72, 2006.

[39] Samuel Klotz and Saralees Nadarajah. *Multivariate t-Distributions and Their Applications*. Cambridge University Press, 2004.

[40] David Lando. On Cox Processes and Credit Risky Securities. *Review of Derivatives Research*, 2(2/3):99–120, 1998.

[41] David Li. On Default Correlation: A Copula Function Approach. *Journal of Fixed Income*, 9:43–54, 2000.

[42] Filip Lindskog. Linear Correlation Estimation. Working paper, Risklab, ETHZ, 2000.

[43] Filip Lindskog, Alexander J. McNeil, and Uwe Schmock. Kendall's Tau for Elliptical Distributions. *Credit Risk: Review of Derivatives Research*, pages 144–156, 2003.

[44] Robert S. Liptser and Albert N. Shiryaev. *Theory of Martingales*. Kluwer Academic Publishers, 1986.

[45] Dilip B. Madan and Haluk Unal. Pricing the Risk of Default. *Review of Derivatives Research*, 2:121–160, 1998.

[46] Jan R. Magnus and Heinz Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons, New York, 1980.

[47] A.W. Marshall and Ingram Olkin. Families of Multivariate Distributions. *Journal of Amercian Statistical Association*, 83:834–841, 1988.

[48] Marcus R. W. Martin, Stefan Reitz, and Carsten S. Wehn. *Kreditderivate und Kreditrisikomodelle*. Vieweg, Wiesbaden, 2006.

[49] Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative Risk Management*. Princeton University Press, 2005.

[50] Robert C. Merton. Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

[51] Robert C. Merton. On the Pricing of Corporate Debt: the Risk Structure of Interest Rates. *Journal of Finance*, 29:449–470, 1974.

[52] Roger B. Nelsen. *An Introduction to Copulas*. Springer, 2nd edition, 2006.

[53] J P. Nolan. *Stable Distributions - Models for Heavy Tailed Data*. Birkhäuser, Boston, 2nd edition, 2011. In progress, Chapter 1 online at academic2.american.edu/∼jpnolan.

[54] Andrew Patton. Modelling Time-Varying Exchange Rate Dependence Using the Conditional Copula. Technical Report Discussion Paper 2001-09, Department of Economics, University of California, San Diego, June 2001.

[55] Robert Rauhmeier and Bernd Engelmann. *The Basel II Risk Parameters*. Springer, 2006.

[56] Muray Rosenblatt. Remarks on a Multivariate Transformation. *Annals of Mathematical Statistics*, 23(3):470–472, 1952.

[57] Gennady Samorodnitsky and Murad S. Taqqu. *Stable Non-Gaussian Random Processes*. Chapman & Hall/CRC, 1994.

[58] Philipp J. Schönbucher. *Credit Derivatives Pricing Models*. Wiley, 2003.

[59] Philipp J. Schönbucher and Dirk Schubert. Copula-Dependent Default Risk in Intensity Models. Technical report, Department of Statistics, Bonn University, December 2001.

[60] Steven E. Shreve. *Stochastic Calculus for Finance II - Continuous-Time Models*. Springer, 2nd edition, 2004.

[61] Albert Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, 8:229–231, 1959.

[62] Peter Tankov. *Lévy Processes in Finance: Inverse Problems and Dependence Modelling*. PhD thesis, École Polytechnique, 2004.

[63] Kai Tappe. *Ordinary and Lévy Copulas in Finance*. PhD thesis, Bergische Universität Wuppertal, 2008.

[64] Niall Whelan. Sampling from Archimedean Copulas. *Quantitative Finance*, 4(3):339–352, 2004.

[65] David W. Williams. *Probabilities With Martingales*. Cambridge University Press, 1991.

[66] Vladimir M. Zolotarev. *One-Dimensional Stable Distributions*. American Mathematical Society, 1986.