Pair Constructions for High-Dimensional Dependence Models in Discrete and Continuous Time

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2013

vorgelegt
von

Dipl. Math. Oec. Stephan Nicklas

aus

Stuttgart

Referent:	Prof. Dr. Friedrich Schmid, Universität zu Köln
Korreferent:	Prof. Dr. Karl Mosler, Universität zu Köln
Tag der Promotion:	10. Mai 2013

Danksagung

Während meiner Zeit in Köln haben mich viele Personen in akademischer und privater Hinsicht unterstützt und geprägt. Ihnen allen bin ich zu großem Dank verpflichtet.

Zuerst möchte ich mich ganz herzlich bei Prof. Schmid für seine Bereitschaft bedanken, meine Promotion zu betreuen. Insbesondere möchte ich ihm für das mir entgegengebrachte Vertrauen danken und dafür, dass er mir die nötige Freiheit ließ, eigene Ideen zu verfolgen, diese Ideen aber regelmäßig in konstruktive Bahnen lenkte. Er sorgte stets für ein angenehmes und produktives Arbeitsumfeld am Lehrstuhl und hatte bei Fragen immer ein offenes Ohr. Ich habe von seiner Erfahrung und seinen Ratschlägen während meiner gesamten Zeit am Lehrstuhl sehr profitiert.

Prof. Mosler möchte ich dafür danken, dass er sich bereit erklärt hat, meine Arbeit zu begutachten. Seine Ratschläge und Hinweise waren immer hilfreich, konstruktiv und haben sehr zur besseren Verständlichkeit des Textes beigetragen.

Bei Dr. Oliver Grothe möchte ich mich für die vielen Gespräche während und nach der Arbeit bedanken. Sein Interesse an Lévy Copulas hat das Kapitel über die Pair-Lévy Copula Konstruktion und die darauf basierende Publikation erst möglich gemacht.

Den vielen Kollegen, die mich während meiner Promotion begleitet haben, möchte ich für die außergewöhnlich gute Zeit in Köln danken. Ich werde diesen Lebensabschnitt immer in besonders positiver Erinnerung behalten. Besonders hervorheben möchte ich Oliver Bantel, Felix Heinl, Carsten Körner, Dr. Duc Hung Tran, Philipp Immenkötter, Eugen Töws, Claudio Wewel, Dr. Thomas Blumentritt, Leila Samad-Tari, Dr. Julius Schnieders, Dr. Konstantin Glombek, Dr. Martin Ruppert und Dr. Tobias Wickern. Ein großer Dank gebührt auch den zahlreichen aktuellen und ehemaligen studentischen Hilfskräften am Lehrstuhl für Wirtschafts- und Sozialstatistik für ihre zuverlässige Arbeit. Insbesondere danke ich Julia Mindlina, Alexandra Lange, Lucas Welling und Alexander de Vivie, die mich über die gesamte Zeit begleitet haben.

Ganz besonders möchte ich mich auch bei Juli, Elisa, meinen Eltern und meinen Geschwistern für Ihre vielfältige und unablässige Unterstützung bedanken.

Contents

List of Symbols and Abbreviations

$\#$	number of elements of a set; 29
$*$	convolution; 65
$\langle \cdot, \cdot \rangle$	inner product $\langle x, y \rangle = \sum_{j=1}^{d} x_j y_j$; 65
$\vee$	maximum; 42
$\wedge$	minimum; 42
$\times$	Cartesian product; 6
$\otimes$	generalized product; 76
$f_{\#}\mu$	push forward measure of measurable function f and measure μ; 76
$\mathbb{1}(\cdot)$	indicator function; 18
$\mathcal{B}(\mathbb{R}^d)$	Borel σ-algebra of $\mathbb{R}^d$; 62
C	copula function; 6
c	copula density; 16
C_n	empirical Copula; 17
$\mathfrak{C}$	Lévy copula function; 72
$\mathfrak{c}$	Lévy copula density; 85
$\widehat{F}_j^n$	marginal empirical distribution function; 17
(γ, Σ, ν)	characteristic triplet; 66
λ_l	lower tail dependence coefficient; 15
λ_u	upper tail dependence coefficient; 16
M	Fréchet-Hoeffding upper bound; 7
$\mathbb{N}$	natural numbers; 63
$\overline{\mathbb{N}}$	extended natural numbers; 66
ν	Lévy measure; 65
$(\Omega, \mathcal{F}, P)$	probability space; 62

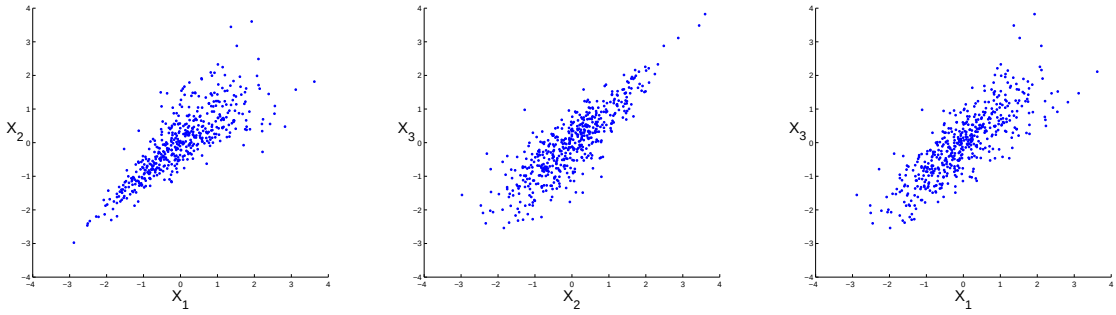
PCC	pair-copula construction; 3
φ_ξ	characteristic function of the probability measure ξ; 65
Π	independence copula; 7
PLCC	pair-Lévy copula construction; 4
PRM	Poisson random measure; 66
$\mathrm{Ran}(\cdot)$	range of a function; 6
$\mathbb{R}^d$	d-dimensional Euclidean space; 6
$\mathbb{R}^d_+$	non-negative elements of $\mathbb{R}^d$; 69
$\overline{\mathbb{R}}^d$	extended d-dimensional Euclidean space $[-\infty, \infty]^d$; 6
$\overline{\mathbb{R}}^d_+$	non-negative elements of $\overline{\mathbb{R}}^d$; 72
ρ_S	Spearman's rho; 15
S_d^+	symmetric non-negative definite $d \times d$ matrix; 65
τ	Kendall's tau; 15
U	tail integral; 71
$u_{i,j}^n$	pseudo observation; 17
$\mathscr{V}$	regular vine; 29
VaR_α	value at risk: for a random variable X we define the VaR_α by $P(X < \mathrm{VaR}_\alpha) = 1 - \alpha$; 52
W	Fréchet-Hoeffding lower bound; 7
$(X_t)_{t \in \mathbb{R}_+}$	stochastic process, where $X_t = X(t)$ denotes the respective random variable at time t; 62

Chapter 1

Introduction

Measuring, describing, and modeling the dependence structure between different random events is at the very heart of statistics. Therefore, a broad variety of different dependence concepts have been developed in the past. The most famous concept is the correlation coefficient of Bravais-Pearson, and in many applied fields it is even common to express the dependence of a multivariate random variable solely by its correlation. However, the correlation coefficient has various drawbacks and is only a measure of linear dependence. In order to model the full dependence structure of multivariate random variables, one needs to go beyond dependence measures. Copula theory offers the possibility to model the entire dependence structure of a multivariate random variable separately from its margins. This approach has been thoroughly investigated in recent publications, however, most work has been done in the 2-dimensional case and for static random variables. Using this concept in a high-dimensional or continuous-time setting is still challenging and the focus of this dissertation.

During the last two decades, we have observed a huge increase in scientific publications and conferences on dependence modeling. Embrechts (2009) explains this increasing attention by the application of the copula concept in the financial sector, driven by renewed regulatory guidelines and new financial products. The recent financial crisis has illustrated that risk management is another area where sound dependence modeling is essential. Throughout the crisis, it has become obvious that there are no risk free investments. Housing prices may decline, banks and insurance companies can file for bankruptcy, and even government bonds of developed countries and bank deposits are not perfectly save. Thus, common sense suggests to diversify a portfolio in order to avoid severe losses. However, the quantification of the risk exposure of a well-diversified portfolio requires a thorough modeling of the dependence structure between the single assets. That is why we need flexible and tractable high-dimensional dependence models that account for the dependence between the different assets. These models enable us to provide correct risk forecasts for the regulator and to base portfolio rebalancing decisions on correct risk scenarios. In particular, we need dependence concepts that are flexible enough to account for the joint behavior of assets in times of crises. Therefore, dependence modeling is an important field in financial risk management. Besides the numerous challenges, the enormously increased computational power as well as the improved quality and availability of financial data constantly improves the modeling



(a) (b) (c)

Figure 1.1: The three different plots show realizations of a 3-dimensional random variable with standard normally distributed margins. All bivariate margins have the same correlation coefficient; however, major differences in the joint behavior of the extremes are obvious. The underlying dependence structure is a pair-copula construction.

procedures. Overall, risk management is facing massive challenges and the constant improvement of the flexibility of the dependence models as well as the continuous adjustment of the high-dimensional dependence concepts is only a small part in a sound risk management process. Nevertheless, poor dependence models can lead to dramatic misjudgments on the risk exposure of well-diversified portfolios, and therefore, with this dissertation, we intend to provide new tools to further improve dependence modeling.

Outline and Summary

This dissertation is divided into two parts. In the first part (Chapter 2-4), we focus on modeling the dependence structure for static random variables in a high-dimensional setting. Figure 1.1 illustrates why the correlation coefficient concept is not sufficient to describe the full dependence structure of a multivariate random variable. All univariate margins of this 3-dimensional random variable are standard normally distributed, and all bivariate margins, which are visualized in the plot, have the same correlation coefficient; however, the joint behavior in the plots is quite different. In Figure 1.1a, extreme negative values in the first dimension tend to occur at the same time as extreme negative values in the second dimension, whereas in Figure 1.1b, we see a different behavior. Extreme positive values occur at the same time, and the dependence of the strong negative values is less distinct. Figure 1.1c shows neither a strong dependence in the extreme positive nor in the extreme negative values. These differences in the behavior of the extremes can have severe consequences in many applications on the overall risk exposure, and this problem is even more pronounced in higher dimensions as the correlation matrix only provides insufficient information on the bivariate margins. Copula theory offers the possibility to overcome most of the drawbacks of the correlation coefficient since this theory provides all information on the dependence structure. Unfortunately, most of

the existing high-dimensional parametric copulas are not flexible enough to model a dependence structure like in Figure 1.1. A very promising approach to create flexible and tractable high-dimensional parametric copulas in arbitrary dimensions is the pair-copula construction (PCC) of Aas *et al.* (2009). This theory is based on the seminal work of Joe (1997), and the structure of a pair-copula constructions is usually visualized by the concept of regular vines, which has been developed by Bedford and Cooke (2002). However, the implementation of estimation and simulation algorithms requires a different representation. Furthermore, the number of parameters in the pair-copula construction increases quadratically with the dimension and it is not evident that we need all of the parameters within the pair-copula construction. Therefore, the two main contributions of the dissertation to this field are

(i) the introduction of a new representation for pair-copula structures that is straightforward to interpret and can be easily implemented in pair-copula algorithms,

(ii) the development of a new parameter reduction technique that crucially reduces the number of parameters in the pair-copula construction in order to avoid overfitting and to keep the model tractable in high dimensions.

The second part of the dissertation (Chapter 5-6) focuses on dependence modeling for continuous-time processes. The increasing availability of high frequency data and non-equidistantly spaced observations demand for continuous-time concepts. Furthermore, it can be useful to model equidistant data by an underlying continuous-time process, which is only observed at a discrete time grid. Compared to the static situation, continuous-time dependence modeling is still in its infancy. The contributions of the dissertation to this area of dependence modeling are

(iii) the development of a flexible, high-dimensional, parametric dependence concept for continuous-time stochastic processes,

(iv) the derivation of simulation and estimation techniques for this new dependence model.

We transfer some of the ideas from classical dependence modeling to the continuous-time setting. In particular, we use the intuition of the pair-copula construction techniques and develop a generalization of this concept for the Lévy copula framework of Kallsen and Tankov (2006).

This dissertation is structured as follows:

In Chapter 2, we recall the basics of copula theory. We fix the notation for the following chapters and state Sklar's theorem, the starting point in copula theory. Furthermore, we recall the definitions of important parametric copula families. Copula based dependence measures of bivariate random variables are recapitulated in Section 2.3, and in the subsequent section, we discuss parametric and semi-parametric estimation techniques for copulas. In Section 2.5, we recall one of the most powerful goodness-of-fit tests for dependence structures. This test will play an important role in the parameter reduction procedure that we introduce in Chapter 4.

Chapter 3 reviews the pair-copula construction framework. In Section 3.1, we summarize two different ways to introduce and interpret pair-copula constructions. These two approaches help to clarify the role of the controversial simplifying assumption in Subsection 3.1.3. We describe two different procedures to represent the structure of a pair-copula in the following sections. Furthermore, we discuss the advantages and drawbacks of these two pair-copula representations. Finally, in Section 3.4, we propose a new representation that overcomes some of the drawbacks of the existing ones.

A discussion of model selection procedures for the pair-copula construction is given in Chapter 4. Section 4.1 presents different heuristics to fit a pair-copula model to a data set. In the following, we focus on the sequential model selection approach of Subsection 4.1.1. Since pair-copula models can get very complex with increasing dimensions, we introduce a new parameter reduction technique in Section 4.2, which helps to avoid redundant parameters in the pair-copula construction process. We show that parameter reduced pair-copula models offer a flexible and low-parametric framework for high-dimensional dependence modeling. In the subsequent section, we illustrate how this parameter reduction technique works for a multivariate set of equity return data.

In Chapter 5, we discuss dependence modeling for stochastic processes. Therefore, we review the concept of Lévy processes in Section 5.1. This short summary is by far not exhaustive, but we fix the notation and recall some of the most important theorems for Lévy processes. Dependence modeling for these processes is discussed in the second part of the chapter. There, we review the concept of Lévy copulas that we use in the subsequent chapter. Lévy copulas model the jump dependence of Lévy processes. This is particularly important in many applications since Lévy copulas allow to model the dependence of sudden large movements in stochastic processes.

In Chapter 6, we transfer the pair-copula construction concept of Chapter 3 to the continuous-time setting of Chapter 5. That is, we develop a pair construction concept for Lévy copulas. In this concept, we assemble bivariate Lévy copulas and bivariate distributional copulas to one high-dimensional Lévy copula. This construction approach is introduced in Section 6.1. In the following section, we develop simulation techniques for dependent Lévy processes, based on the pair-Lévy copula construction (PLCC). Furthermore, we provide maximum likelihood estimators for these continuous-time dependence structures. In Section 6.3, we evaluate these new procedures in a simulation study. In the following sections, we show how to apply the PLCC concept to multivariate Ornstein-Uhlenbeck processes and give an outlook on further applications.

Chapter 2

Dependence Modeling with Copulas

Modeling the dependence structure between different events is one of the fundamental challenges in statistics. In applications, one is seldom confronted with one single source of uncertainty and the dependence structure between the different sources can have a tremendous effect on the overall risk exposure. This is a well-recognized fact especially in financial and actuarial applications. See, e.g., Genest *et al.* (2009a), McNeil *et al.* (2005), Cherubini *et al.* (2004), and Panjer (2006). However, thorough dependence modeling is also important in many other areas like hydrology, engineering, operations research, economics, and biostatistics, see, e.g., Genest and Favre (2007) and Genest *et al.* (2009a).

The correlation coefficient of Bravais-Pearson is the most prominent dependence concept. However, as discussed, e.g., in Embrechts *et al.* (2002), the correlation coefficient has severe drawbacks for non-elliptical random variables and is only a measure of linear dependence. For a survey on dependence measures that are invariant under monotonic transformations and often more appropriate than the correlation, we refer to Schmid *et al.* (2010). In this chapter, we review the copula concept of Hoeffding (1940) and Sklar (1959) which has been rediscovered and expanded, e.g., by Joe (1997), Nelsen (2006), and references therein. Copulas go beyond dependence measures and provide a sound framework for general dependence modeling. They separate the margins and the dependence structure of any multivariate distribution. This result is important to estimate, understand, and interpret the dependence structure in a given set of data. Furthermore, this concept offers the possibility to combine arbitrary marginal distributions to a valid multivariate distribution function with a specific dependence structure. This is essential for multivariate modeling, since we can use sophisticated univariate models for any margin and combine them with a copula. Moreover, the separation of the dependence structure and the margins offers the possibility to estimate a parametric model for the multivariate distribution in two steps. That is, one estimates the parameters of the univariate marginal models in a first step and continues by estimating the dependence structure in a second step. This is a very important feature, since even the estimation of high-dimensional distributions becomes feasible in this sequential framework. In addition, simulation of random variables is straightforward within the copula concept. The standard references for copula theory are Joe (1997) and Nelsen (2006). Durante and Sempi (2010) give a historical introduction to the topic and provide an extensive

list of further references. The application of copula theory to finance is discussed, e.g., in McNeil *et al.* (2005), Cherubini *et al.* (2004), and Cherubini *et al.* (2012). Mai and Scherer (2012) give a broad overview on simulation techniques.

This chapter is structured as follows. In the first section, we give a short introduction to copula theory. Then, we discuss methods to create valid copula functions and give prominent examples of parametric copula families that we apply in the subsequent chapters. In the third section, we present dependence measures for bivariate random variables that are solely based on the underlying copula. In the fourth section, we discuss different estimation procedures for an underlying parametric copula. Finally, we recall goodness-of-fit procedures to decide whether we have found an adequate dependence model for a given set of data.

2.1 Definition, Properties, and Sklar's Theorem

In this section, we give a brief introduction to copula theory. We state some of the well-known properties of copula functions that we need in the subsequent chapters. For a more detailed introduction and the proofs of the given results, we refer to Joe (1997) and Nelsen (2006). Furthermore, with Sklar's theorem, we recall the central result in copula theory that explains the key role of copula functions in dependence modeling. In the next definition, we formally define copulas.

Definition 2.1 *A d-dimensional distribution function $C(u_1, \ldots, u_d) : [0,1]^d \to [0,1]$, where the margins satisfy $C_j(u_j) = C(1, \ldots, 1, u_j, 1, \ldots, 1) = u_j$ for all $u_j \in [0,1]$ and $j = 1, \ldots, d$, is called a copula.*

Obviously, the condition on the margins assures that the copula is a distribution function with uniform margins. Sklar's Theorem is the starting point in copula theory. It shows how we can decompose any multivariate distribution function into the marginal distribution functions and a copula that contains the dependence structure. The theorem was first given in Sklar (1959) and is also stated, e.g., in Nelsen (2006, Theorem 2.10.9).

Theorem 2.2 *Let $F_{1,\ldots,d}$ be a d-dimensional distribution function with marginal distribution functions $F_1, \ldots, F_d$. Then there exists a copula C such that for all $x \in \overline{\mathbb{R}}^d$,*

$$F_{1,\ldots,d}(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)). \tag{2.1}$$

If $F_1, \ldots, F_d$ are all continuous, then C is unique. Otherwise, C is uniquely determined on the Cartesian product of the ranges of the marginal distribution functions $Ran(F_1) \times \ldots \times Ran(F_d)$. Conversely, if C is a copula and $F_1, \ldots, F_d$ are distribution functions, then the function $F_{1,\ldots,d}$ defined by Equation (2.1) is a d-dimensional distribution function with margins $F_1, \ldots, F_d$.

In this work, we focus on continuous margins only. For a treatment of copula theory with discrete marginal distributions, we refer to Genest and Nešlehová (2007). Next, we discuss functions of special importance in copula theory. The Fréchet-Hoeffding upper bound M, the Fréchet-Hoeffding lower bound W, and the independence copula Π are given by

$$M(u) = \min(u_1, \ldots, u_d), \tag{2.2}$$
$$\Pi(u) = u_1 \cdot \ldots \cdot u_d, \tag{2.3}$$
$$W(u) = \max(u_1 + \ldots + u_d - d + 1, 0). \tag{2.4}$$

Note that the Fréchet-Hoeffding upper bound M and the independence copula Π are copulas in arbitrary dimensions, whereas the Fréchet lower bound W is only in the bivariate case a copula function. The Fréchet-Hoeffding bounds are pointwise bounds for any copula function. This is shown, for instance, in Nelsen (2006, Theorem 2.10.12).

Proposition 2.3 *If C is any copula function, then for every $u \in [0,1]^d$,*

$$W(u) \leq C(u) \leq M(u) \tag{2.5}$$

holds.

Note that even for $d > 2$, where the Fréchet-Hoeffding lower bound W is not a copula, these inequalities are best possible. Furthermore, M, Π, and W have a special interpretation, as given, e.g., in Nelsen (2006, Theorem 2.5.5, Theorem 2.10.14), whenever they are copulas.

Proposition 2.4 *For $d \geq 2$, let $X_1, \ldots, X_d$ be continuous random variables. Then*

(i) $X_1, \ldots, X_d$ are independent if and only if the copula of $X_1, \ldots, X_d$ is Π,

(ii) each of the random variables $X_1, \ldots, X_d$ is almost surely a strictly increasing function of any of the others if and only if the copula of $X_1, \ldots, X_d$ is M.

For $d = 2$, let X_1, X_2 be continuous random variables. Then

(iii) X_1 is almost surely a strictly decreasing function of X_2 if and only if the copula of X_1, X_2 is W.

In more than two dimensions, the properties of the Fréchet-Hoeffding upper and lower bound are very different. This is due to the fact that positive and negative dependencies are not comparable anymore. It is possible, for example, to have a collection of random variables $X_1, \ldots, X_d$ such that each of these random variables is an increasing function of any of the others. This is not possible for negative dependencies and more than two random variables. Suppose that X_1 is a decreasing function of X_2 and X_2 is a decreasing function of X_3 as well. Then, the variable X_1 has to be an *increasing* function of X_3. Another important and well-known property of any copula is given in the next proposition, see, e.g., Nelsen (2006, Theorem 2.10.7).

Proposition 2.5 *Let C be a d-dimensional copula. Then for every u and v in $[0,1]^d$,*

$$|C(v) - C(u)| \leq \sum_{j=1}^{d} |v_j - u_j|.$$

Hence, C is uniformly continuous on $[0,1]^d$.

There are different kinds of symmetry for copula functions. See, e.g., Nelsen (2006, Chapter 2.7). Here, we recall the definition of permutation symmetry, which is a necessary condition on the copula of exchangeable random variables.

Definition 2.6 *Let C be a d-dimensional copula. We say C is permutation-symmetric if*

$$C(u_1, \ldots, u_d) = C(u_{\tau(1)}, \ldots, u_{\tau(d)})$$

for any permutation τ and any $u_1, \ldots, u_d \in [0,1]^d$.

Thus, permutation-symmetry is closely related to the concept of exchangeability for random variables.

2.2 Parametric Copula Families

We are not only interested in theoretical properties of copula functions, but we want to apply this theory to a given data set as well. Therefore, we discuss different construction methods for d-dimensional, parametric dependence structures. Furthermore, we recall the definition of selected copula families that we need in the following chapters. Note that this is by no means exhaustive. Especially in two dimensions, there are plenty of different copulas families. For a thorough discussion, we refer to Joe (1997), Nelsen (2006), and references therein. However, there are only few copula models that are flexible enough to represent the dependence structure of high-dimensional data. The most prominent ones are the Gauss and t copula being easy to estimate and widely used in applications. In some cases, however, these dependence structures fail to model desirable dependence properties. Nevertheless, we use them as benchmarks for the high-dimensional copula model in Chapter 3. Another approach to build multivariate copulas is based on the extension of bivariate Archimedean copulas. Unfortunately, these copulas lack the desired flexibility in higher dimensions. In the third part of this section, we discuss a hierarchical procedure to overcome this problem. Therefore, we combine lower-dimensional Archimedean copulas such that the resulting structure is a valid d-dimensional copula.

2.2.1 Implicit Copula Families

Here, we present a way to define parametric copula families by extracting the dependence structure of a known multivariate distribution. The next corollary shows how to invert Sklar's theorem to define parametric copula families in d dimensions. This corollary is already stated, e.g., in Nelsen (2006, Corollary 2.10.10).

Corollary 2.7 *Let $F_{1,\dots,d}$ be a d-dimensional distribution function, where $F_1, \dots, F_d$ denote the continuous marginal distribution functions, and let C be the corresponding copula. We denote the marginal quasi-inverses, see Nelsen (2006, Definition 2.3.6), by $F_1^{-1}, \dots, F_d^{-1}$. Then, we have for any $u \in [0,1]^d$*

$$C(u_1, \dots, u_d) = F_{1,\dots,d}(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)).$$

This relation is particularly interesting for the class of elliptical distributions, since the univariate margins as well as the multivariate distribution function are well-known. See, e.g., Fang *et al.* (1990). Furthermore, the resulting copulas are quite flexible dependence models. These elliptical copulas can be used to combine arbitrary margins and create new multivariate distribution functions. In the next example, we give the definition of the best known elliptical copula.

Example 2.8 *Let $\Phi_{1,\dots,d}$ be the distribution function of the multivariate normal distribution with zero mean and correlation matrix R. We denote the distribution function of the univariate standard normal distribution by Φ, and define the d-dimensional Gauss copula by*

$$C(u_1, \dots, u_d) = \Phi_{1,\dots,d}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d)).$$

The Gauss copula extracts the dependence structure from the multivariate normal distribution. It is uniquely defined by the correlation matrix and thus has $d(d-1)/2$ parameters. Hence, it offers a certain flexibility and is widely applied. We use the Gauss copula as a benchmark model for the more advanced multivariate copula models that we discuss in Chapter 3. In financial applications, however, it is often observed that the dependence of the extreme events is stronger than suggested by the Gauss copula. Therefore, we state the definition of the t copula in the next example. For the definition of the multivariate t-distribution $t_{1,\dots,d}$, we refer to Fang *et al.* (1990, Example 2.5).

Example 2.9 *Let $t_{1,\dots,d}$ be the distribution function of a vector $X \sim t_{1,\dots,d}(\nu, 0, R)$, where R is a correlation matrix and $\nu > 0$. We denote the univariate standard t-distribution with ν degrees of freedom by t, and we define the d-dimensional t copula by*

$$C(u_1, \dots, u_d) = t_{1,\dots,d}(t^{-1}(u_1), \dots, t^{-1}(u_d)).$$

In contrast to the Gauss copula, the t copula has an additional parameter to control for the dependence of the extreme events. This makes the t copula more suitable for

financial application. However, the t copula is symmetric in the sense that extreme positive and extreme negative events are modeled equivalently. Furthermore, there is only one parameter for the dependence of all extreme events, which might not be enough for high-dimensional dependence structures. The definitions of these two copulas can be found, e.g., in McNeil *et al.* (2005, Chapter 5).

2.2.2 Archimedean Copulas

Besides the implicit copulas, the class of Archimedean copulas is one of the most popular families in parametric dependence modeling. In contrast to implicit copula families, we do not need well-known multivariate distribution functions to create a new copula family. Archimedean copulas are uniquely defined by a generating function. The generator of a 2-dimensional Archimedean copula is a convex, continuous, and strictly decreasing function $\varphi : [0,1] \mapsto [0,\infty]$, such that $\varphi(1) = 0$. The following theorem, stated, e.g., in Nelsen (2006, Section 4.1), shows how to use the generator to define a valid copula.

Theorem 2.10 *Let φ be a generator function and denote by $\varphi^{[-1]}$ the pseudo inverse*

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t) & 0 \leq t < \varphi(0), \\ 0 & \varphi(0) \leq t \leq \infty. \end{cases}$$

Then, the function $C : [0,1]^2 \mapsto [0,1]$

$$C(u,v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) \tag{2.6}$$

is a copula.

In the next example, we present three of the most popular Archimedean copulas that we are going to use in the subsequent chapters. For more information on these copulas and additional Archimedean families, we refer to Nelsen (2006).

Example 2.11 *In this example, we give the generator and the copula function of selected bivariate Archimedean copulas.*

- *Clayton copula (Clayton, 1978): For $\theta \in [-1,\infty)\backslash\{0\}$, the Clayton copula is given by*

  $$C(u,v) = \left(\max\{u^{-\theta} + v^{-\theta} - 1, 0\}\right)^{-\frac{1}{\theta}}, \tag{2.7}$$

  with generator

  $$\varphi(t) = \frac{1}{\theta}(t^{-\theta} - 1).$$

  For $\theta = 0$, we set $C = \Pi$.

- *Gumbel copula (Gumbel, 1960): For $\theta \in [1,\infty)$, the Gumbel copula is defined as*

  $$C(u,v) = \exp\left(-\left((-\log(u))^{\theta} + (-\log(v))^{\theta}\right)^{\frac{1}{\theta}}\right), \tag{2.8}$$

with generator

$$\varphi(t) = (-\log(t))^{\theta}.$$

- *Frank copula (Frank, 1979): For $\theta \in (-\infty, \infty)\backslash\{0\}$, the Frank copula is defined as*

$$C(u,v) = -\frac{1}{\theta}\log\left(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right), \tag{2.9}$$

with generator

$$\varphi(t) = -\log\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right).$$

Again, we set $C = \Pi$ for $\theta = 0$.

The concept of Archimedean copulas can be generalized to the multivariate case. The definition of the multivariate Archimedean copula is straightforward. However, we need to impose additional assumptions on the generator to guarantee for a valid copula function. This is formalized in Nelsen (2006, Theorem 4.6.2).

Theorem 2.12 *Let $\varphi : [0,1] \mapsto [0,\infty]$ be a continuous strictly decreasing function such that $\varphi(0) = \infty$ and $\varphi(1) = 0$, and let φ^{-1} denote the inverse of φ. We define a function $C : [0,1]^d \mapsto [0,1]$ by*

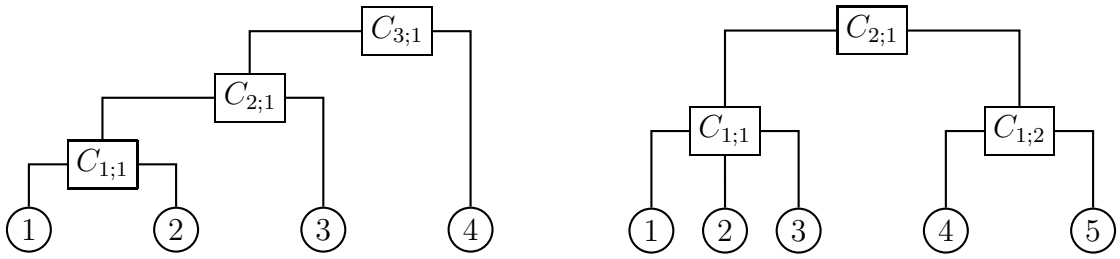
$$C(u) = \varphi^{-1}\left(\varphi(u_1) + \ldots + \varphi(u_d)\right).$$

The function C is a copula for all $d \geq 2$ if and only if φ^{-1} is completely monotonic on $[0,\infty)$, that is, it is continuous and has derivatives of all orders that alternate in sign for any $t \in (0,\infty)$.

For necessary and sufficient conditions on the generator for a fixed dimension d, we refer to McNeil and Nešlehová (2009). Note that these multivariate copulas are permutation-symmetric. This implies that all lower-dimensional margins of the copula have the same dependence structure. In particular, models based on multivariate Archimedean copulas have equicorrelated ranks.

2.2.3 Hierarchical Archimedean Copulas

Two-dimensional Archimedean copulas constitute a very important class in dependence modeling. Some of the best known copula families are Archimedean and this property has many theoretical and practical advantages. As shown in the preceding section, generalizations to the multivariate case ($d \geq 3$) are possible. However, the property of permutation-symmetry is a severe restriction in more than two dimensions. Usually, this symmetry is not tenable when dealing with a high-dimensional set of data. Joe (1997) introduces the idea of a hierarchical construction method to define multivariate copulas by nesting different lower-dimensional Archimedean copulas. With this approach, one can partially overcome the permutation-symmetry in high-dimensional copula models.



(a) (b)

Figure 2.1: This plot shows an example of two different hierarchical Archimedean copula structures

The idea in Joe (1997) is straightforward. First, one models the dependence between the first and second dimension with the bivariate Archimedean copula $C_{1;1}$. In the next step, one defines $z_{1,2} = C_{1;1}(u_1, u_2)$ and combines this new variable with the untransformed variable u_3 of the third dimension by the Archimedean copula $C_{2;1}$. Then, one defines $z_{1,2,3} = C_{2;1}(z_{1,2}, u_3)$ and we can iterate this procedure until all variables are included. This approach is visualized in Figure 2.1a with four variables. The dependence function is given by

$$\begin{aligned} &C_{1,2,3,4}(u_1, u_2, u_3, u_4) = \\ &\quad = C_{3;1}(C_{2;1}(C_{1;1}(u_1, u_2), u_3), u_4) \\ &\quad = \varphi_{3;1}^{[-1]}\left(\varphi_{3;1} \circ \varphi_{2;1}^{[-1]}\left(\varphi_{2;1} \circ \varphi_{1;1}^{[-1]}\big(\varphi_{1;1}(u_1) + \varphi_{1;1}(u_2)\big) + \varphi_{2;1}(u_3)\right) + \varphi_{3;1}(u_4)\right), \end{aligned}$$

where we denote the generator of $C_{i;j}$ by $\varphi_{i;j}$. A sufficient condition for $C_{1,2,3,4}$ to be a copula is that all inverse generator functions are completely monotonic, and furthermore, the composition of the generator functions $\varphi_{3;1} \circ \varphi_{2;1}^{[-1]}$ and $\varphi_{2;1} \circ \varphi_{1;1}^{[-1]}$ have to be completely monotonic as well, see Joe (1997, Chapter 4).

Different nestings strategies also lead to valid multivariate copulas. A 5-dimensional example is illustrated in Figure 2.1b. In this case, the copula is given by

$$\begin{aligned} C_{1,2,3,4,5}(u_1, \ldots, u_5) &= C_{2;1}(C_{1;1}(u_1, u_2, u_3), C_{1;2}(u_4, u_5)) \\ &= \varphi_{2;1}^{[-1]}\Big(\varphi_{2;1} \circ \varphi_{1;1}^{[-1]}\big(\varphi_{1;1}(u_1) + \varphi_{1;1}(u_2) + \varphi_{1;1}(u_3)\big) \\ &\quad + \varphi_{2;1} \circ \varphi_{1;2}^{[-1]}\big(\varphi_{1;2}(u_4) + \varphi_{1;2}(u_5)\big)\Big), \end{aligned}$$

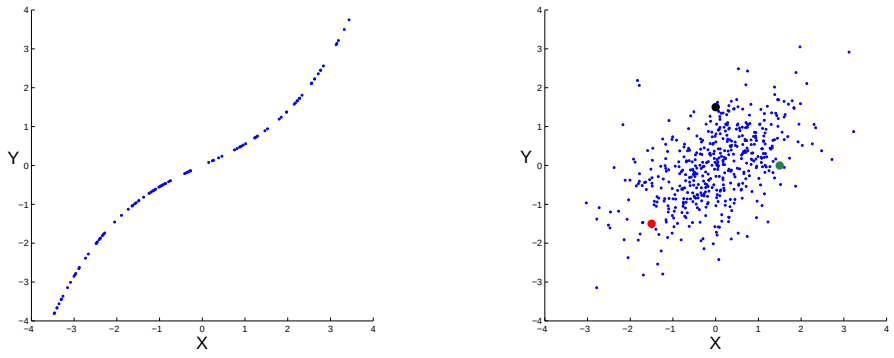
where all inverse generator functions as well as $\varphi_{2;1} \circ \varphi_{1;1}^{[-1]}$ and $\varphi_{2;1} \circ \varphi_{1;2}^{[-1]}$ have to be completely monotonic. This procedure can be extended easily to arbitrary dimensions. However, the notation gets involved, and therefore we refer to Savu and Trede (2010) for a general treatment. Choosing an adequate nesting structure is treated in Okhrin (2007), simulation techniques for hierarchical Archimedean copulas are given in Whelan

(2004), McNeil (2008), and Hofert (2008). The density for the general case is derived in Savu and Trede (2010).

Hierarchical Archimedean copulas are popular for several reasons. They overcome the problem of permutation-symmetry, the connection to other areas in probability theory like Laplace transforms is appealing, and the hierarchical structure often has a nice interpretation. In financial applications, for example, the dependence of the assets in one sector can be modeled on the first level and the dependence between the different sectors on a second. However, there are also severe drawbacks of this method. For any hierarchical structure and any selection of Archimedean copulas, the conditions on the composite generator functions have to be verified separately. Furthermore, these conditions can be very restrictive. In the hierarchy of Figure 2.1b, for example, it is not possible to use a Gumbel copula for $C_{1;1}$ and a Clayton copula for $C_{2;1}$, see Savu and Trede (2010). The conditions are only easy to verify if all Archimedean copulas in the hierarchy belong to a special Archimedean family. For instance, if all copulas in the structure are of Gumbel type, of Clayton type, or of Frank type one only has to check that the dependence parameters decrease with the hierarchy level (Aas and Berg, 2009). That is, for the hierarchy in Figure 2.1b, we need to guarantee that $\theta_{2;1} < \theta_{1;1}$ and $\theta_{2;1} < \theta_{1;2}$. However, the restriction to one copula family for all copulas in the hierarchy limits the applicability of the concept strongly. Finally, hierarchical Archimedean copulas are not the only high-dimensional dependence models that overcome permutation-symmetry. The multivariate Gauss and t copula, as well as the pair-copula construction of Chapter 3 are in general not permutation-symmetric. Furthermore, in the model comparison of Fischer *et al.* (2009) and Aas and Berg (2009), hierarchical Archimedean copula models perform worse than competing dependence structures.

2.3 Dependence Measures

The copula framework is a most general dependence concept. It captures all information on the dependence structure of a random variable. However, the copula is a d-dimensional cumulative distribution function on $[0,1]^d$ and therefore hard to interpret. Furthermore, comparing the magnitude of dependence for different copulas might be ambiguous. Therefore, it is often advantageous to summarize the dependence structure in a scalar valued dependence measure. The most famous dependence measure is the correlation coefficient ρ of Bravais-Pearson. But, as discussed, this measure has many drawbacks for non-elliptical dependence structures. In this section, we present three alternatives. These measures depend only on the underlying copula and not on the marginal distributions. This is a desirable property, since the margins do not influence the dependence structure. In addition, the first two dependence measures are more robust to estimate than the correlation coefficient. Here, we focus on the bivariate case. For generalizations to the d-dimensional case ($d \geq 3$), we refer to Gaißer (2010), Schmid *et al.* (2010), and references therein.



(a) (b)

Figure 2.2: The plot on the left shows a deterministic relation between X and Y. In particular, every pair of two realizations is concordant. In the plot on the right, we have a realization of a bivariate normal distribution with a positive correlation coefficient, and there are more concordant than discordant pairs.

2.3.1 Measures of Concordance

In this section, we discuss two well-known bivariate measures of association that do not depend on the marginal distributions. There exists a broad variety of different bivariate dependence measures. For an overview on these, we refer to Nelsen (2006). Here, we recall briefly two selected measures of association that are particularly important in the subsequent chapters. To introduce these measures, we need to discuss the concept of concordance. Let (X, Y) be a bivariate random variable. A pair of realizations (x_1, y_1) and (x_2, y_2) is concordant if $x_1 < x_2$ and $y_1 < y_2$, or $x_1 > x_2$ and $y_1 > y_2$. That is, if $(x_1 - x_2)(y_1 - y_2) > 0$. In Figure 2.2b, the red and the black points as well as the red and the green points are concordant. We say that (x_1, y_1) and (x_2, y_2) are discordant if $x_1 < x_2$ and $y_1 > y_2$, or $x_1 > x_2$ and $y_1 < y_2$. That is, if $(x_1 - x_2)(y_1 - y_2) < 0$. In Figure 2.2b, the black and the green observations are discordant. Figure 2.2 visualizes how we use the concordance concept to measure the dependence of random variables. Figure 2.2a shows a realization of a perfectly positively dependent random variable (X, Y). Here, all pairs are concordant. In Figure 2.2b, we have a non-deterministic dependence structure of a bivariate normal distribution with a positive correlation. In this case, we observe more concordant than discordant pairs. With this concept, we are able to introduce two different measures of association.

Kendall's Tau

This dependence measure is introduced in Kendall (1938). Let (X_1, Y_1) and (X_2, Y_2) be independent and identically distributed random variables with continuous marginal

distribution functions $F_X(x)$, $F_Y(y)$ and copula C. We define Kendall's tau as

$$\begin{aligned}\tau &= P\left((X_1 - X_2)(Y_1 - Y_2) > 0\right) - P\left((X_1 - X_2)(Y_1 - Y_2) < 0\right) \\ &= 4 \int_{[0,1]^2} C(u,v) dC(u,v) - 1.\end{aligned} \tag{2.10}$$

That is, Kendall's tau gives the probability of concordance minus the probability of discordance. The second relation is proved, e.g., in Nelsen (2006).

Spearman's Rho

Spearman's rho is first mentioned in Spearman (1904) as a rank-based dependence measure. Nelsen (2006) introduces this measure in the following way. Let (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) be independent and identically distributed random variables with continuous marginal distribution functions $F_X(x)$, $F_Y(y)$ and copula C. We define Spearman's rho by

$$\begin{aligned}\rho_S &= 3\left(P\left((X_1 - X_2)(Y_1 - Y_3) > 0\right) - P\left((X_1 - X_2)(Y_1 - Y_3) < 0\right)\right) \\ &= 12 \int_{[0,1]^2} uv \, dC(u,v) - 3.\end{aligned} \tag{2.11}$$

Interestingly, it is easy to show that Spearman's rho is exactly the correlation coefficient between $F_X(X)$ and $F_Y(Y)$. That is,

$$\rho_S = \frac{\mathrm{Cov}(F_X(X), F_Y(Y))}{\sqrt{\mathrm{Var}(F_X(X))}\sqrt{\mathrm{Var}(F_Y(Y))}}.$$

2.3.2 Tail Dependence

Kendall's tau and Spearman's rho measure the dependence on the whole range of the bivariate random variable (X, Y). The tail dependence coefficient, on the contrary, is a measure for the dependence of extreme events. Again, we denote the marginal distribution functions by $F_X(x)$, $F_Y(y)$ and the copula by C. The lower tail dependence coefficient is given by

$$\lambda_l = \lim_{u \downarrow 0} P\left(Y \leq F_Y^{-1}(u) \middle| X \leq F_X^{-1}(u)\right) \tag{2.12}$$

$$= \lim_{u \downarrow 0} \frac{C(u,u)}{u}, \tag{2.13}$$

and the upper tail dependence coefficient is

$$\lambda_u = \lim_{u \uparrow 1} P\left(Y > F_Y^{-1}(u) \middle| X > F_X^{-1}(u)\right) \tag{2.14}$$

$$= 2 - \lim_{u \uparrow 1} \frac{1 - C(u,u)}{1-u}. \tag{2.15}$$

The representation of the tail dependence coefficient in terms of the copula in Equation (2.13) and (2.15) is derived, e.g., in Nelsen (2006, Section 5.4). See Frahm *et al.* (2005) for a survey of different estimators. Tail dependence is an important property in finance. The Gauss copula has a tail dependence of zero and is therefore not appropriate in many situations. The t copula can model upper and lower tail dependence, where $\lambda_u = \lambda_l$. The next example presents a copula that has both upper and lower tail dependence. Furthermore, this copula can have different values for the upper and lower tail dependence coefficient.

Example 2.13 *The BB1 copula family, introduced and discussed in Joe and Hu (1996, Example 5.1), is a parametric copula with different upper and lower tail dependence. In particular, the upper and lower tail dependence coefficients can be set independently of each other. The bivariate copula with parameters $\theta > 0$ and $\tau \geq 1$ is defined by*

$$C(u,v;\theta,\tau) = \left(1 + \left((u^{-\theta}-1)^\tau + (v^{-\theta}-1)^\tau\right)^{\frac{1}{\tau}}\right)^{-\frac{1}{\theta}}.$$

The lower tail dependence coefficient is $\lambda_l = 2^{-1/(\tau\theta)}$ and the upper tail dependence coefficient is $\lambda_u = 2 - 2^{1/\tau}$.

For further measures of tail dependence in arbitrary dimensions, we refer to Schmid and Schmidt (2007).

2.4 Estimation

There are different ways to estimate the underlying copula for a given set of i.i.d data. In this section, we give a very brief overview and discuss the estimation procedure that we use in the subsequent chapters. For a survey of the different estimation procedures, we refer to Choroś *et al.* (2010). We focus on maximum likelihood methods only, using that under absolute continuity assumptions the probability density $f_{1,\dots,d}$ can be decomposed into

$$f_{1,\dots,d}(x_1,\dots,x_d) = c(F_1(x_1),\dots,F_d(x_d)) \prod_{i=1}^{d} f_i(x_i), \tag{2.16}$$

where $f_1,\dots,f_d$ are the marginal densities and c is the copula density. The first possible method is the straightforward maximum likelihood estimation of the full model. This requires choosing a parametric model for the copula and the margins such that we can

calculate the full likelihood function with respect to all parameters. That is, we specify the parameter vectors $\alpha_1, \ldots, \alpha_d$ of the margins and the parameter vector θ of the copula altogether. The problem of this method is that the number of parameters can be very large, even for moderate dimensions. Therefore, the optimization algorithm might be very slow or even computationally infeasible.

To overcome this problem, or at least to provide good starting values for the full maximum likelihood method, Joe (1997) has suggested the inference functions for margins (IFM) approach. This is a sequential procedure where we estimate the marginal parameters separately for every dimension. In a second step, we transform the observations $(x_{i,1}, \ldots, x_{i,d})_{i=1,\ldots,n}$ with the estimated marginal distribution functions

$$u_{i,j}^{\mathrm{IFM}} = F_j(x_{i,j}; \hat{\alpha}_j),$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, d$. In the following, we treat the transformed variables $(u_{i,1}^{\mathrm{IFM}}, \ldots, u_{i,d}^{\mathrm{IFM}})_{i=1,\ldots,n}$ like observations from the underlying copula and estimate the parameter vector θ by the maximum likelihood approach. This is conceptually the same as plugging in the marginal estimators into the likelihood function that is based on the density in Equation (2.16) and maximizing this likelihood function with respect to the copula parameters. Of course, it is not possible to apply the standard maximum likelihood results for the asymptotic properties of this sequential estimator. However, Joe (1997) proves consistency and asymptotic normality under the usual regularity conditions. This stepwise procedure reduces the complexity of the problem and makes estimation feasible even for high dimensions and complex marginal models. However, since the marginal distributions are unknown, we cannot guarantee to select the correct model for the margins and, as shown in Kim *et al.* (2007), misspecified margins can have a crucial effect on the estimator of the copula parameter.

The third method is similar to the IFM approach but it is based on nonparametric estimates of the margins. This excludes the possibility of misspecified margins and is particularly appropriate if the dependence structure is in the focus of the analysis. We denote the one-dimensional empirical distribution function by $\widehat{F}_j^n$ and define the pseudo-observations as

$$u_{i,j}^n = \frac{n}{n+1} \widehat{F}_j^n(x_{i,j}), \tag{2.17}$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, d$. Note that we use the normalization $n/(n+1)$ to avoid problems at the boundaries of $[0,1]^d$. Thus, the pseudo-observations are simply the rank of the observation in its dimension, normalized by $1/(n+1)$. In the next step, we continue as in the IFM case. That is, we treat $(u_{i,1}^n, \ldots, u_{i,d}^n)_{i=1,\ldots,n}$ as observations from the underlying copula and estimate the parameter vector θ with a standard maximum likelihood approach. Genest *et al.* (1995) show that the semiparametric estimator is consistent and asymptotically normal. Kim *et al.* (2007) advocate the use of this estimation method if the marginal distributions are unknown.

A completely nonparametric way to estimate the copula is given in Deheuvels (1979). We use the pseudo-observation of Equation (2.17) and define the empirical copula on

$[0,1]^d$ as

$$C_n(u_1,\ldots,u_d) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(u_{i,1}^n \leq u_1,\ldots,u_{i,d}^n \leq u_d). \tag{2.18}$$

Note that C_n is not continuous and therefore not a copula function. Furthermore, one needs a large amount of data to get a good approximation of the underlying copula on $[0,1]^d$ by the empirical copula function. Thus, one of the main applications of the empirical copula is goodness-of-fit testing for parametric copula models.

2.5 Goodness-of-Fit

There are several different ways to evaluate the fit of a parametric copula for a given data set. A graphical analysis of the dependence structure should always be the first step. In two dimensions, scatter plots of the data can give a first hint on the dependence. However, in many cases, the influence of the marginal distributions conceals the underlying dependence structure. Therefore, it is advantageous to transform the original data and plot the pseudo-observations. Sometimes, these pseudo-observations are further transformed such that all univariate margins are standard normally distributed. In this case, all margins have the same distribution and do not affect the scatter plot asymmetrically. By construction, scatter plots are particularly appropriate for two dimensions. Nevertheless, in the multivariate case, we can apply the bivariate graphical methods on the 2-dimensional margins. A different graphical approach to evaluate the fit of a parametric copula is suggested by Genest and Rivest (1993). They transform the selected parametric copula C to the so called λ-function on $[0,1]$. Then, they compare this function with its nonparametric estimate from the data. For further information and the definition of the theoretical and empirical λ-function, we refer to Genest and Rivest (1993) and Genest *et al.* (2009b).

Graphical evaluations are often suitable to discover a bad fit of the selected dependence model. Though, it is substantially more difficult to distinguish between a moderate and an excellent fit by these approaches. Furthermore, it is desirable to have a statistical framework to validate the intuition that we get from the graphical procedures. Therefore, we need goodness-of-fit tests to decide whether the selected copula can represent the dependence structure in a given data set adequately. We denote a parametric copula family by $\mathcal{C} = \{C_\theta : \theta \in \mathcal{O}\}$, where $\mathcal{O}$ is an open subset in $\mathbb{R}^p$. The hypotheses for these goodness-of-fit tests are given by

$$H_0 : C \in \mathcal{C}, \text{ against } H_a : C \notin \mathcal{C}.$$

Recently, a large variety of different goodness-of-fit tests have been proposed. Fermanian (2005) bases his goodness-of-fit on the density of the copula. A different procedure is introduced in Breymann *et al.* (2003) and corrected in Dobrić and Schmid (2007). This approach is based on the Rosenblatt transformation, see Rosenblatt (1952). Genest *et al.* (2009b) and Berg (2009) introduce further approaches. Moreover, they conduct

extensive simulation studies on computer clusters to compare size and power of selected goodness-of-fit tests for different hypotheses and sample sizes. Genest *et al.* (2009b) focus on the bivariate case, whereas Berg (2009) applies the tests in $d = \{2, 4, 8\}$ dimensions. Here, we discuss one of the goodness-of-fit tests. This test performs particularly well in the power comparison of the simulation studies and we use this test in Section 4.2 in the bivariate case. We denote by θ_n an estimate of θ. The test is proposed in Genest and Rémillard (2008) and is based on the distance between the empirical copula C_n, see Equation (2.18), and the fitted copula $C_{\theta_n} \in \mathcal{C}$. More formally, the distance is measured by the empirical process

$$\mathbb{C}_n = \sqrt{n}\,(C_n - C_{\theta_n})$$

and the Cramér-von Mises test statistic

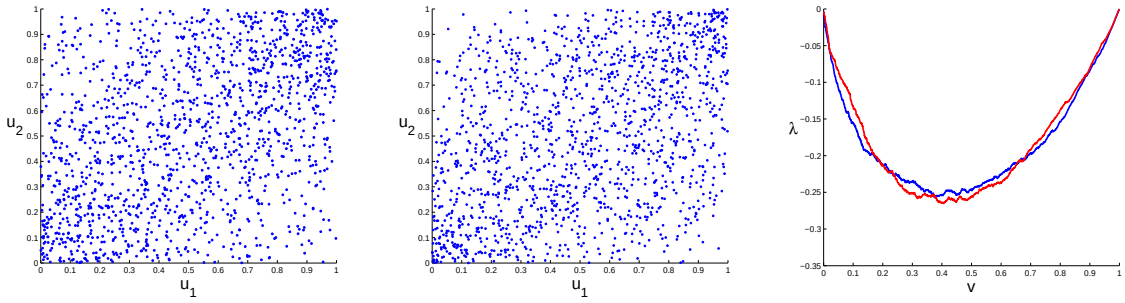
$$S_n = \int_{[0,1]^d} \mathbb{C}_n(u)^2 dC_n(u).$$

As shown in Genest and Rémillard (2008) and noted in Genest *et al.* (2009b), the asymptotic distribution of the test statistic is highly complex and depends on the copula family under the null hypothesis. Moreover, it is not possible to tabulate critical values since these values depend on the true parameter θ of the underlying copula. Therefore, we have to use a parametric bootstrap procedure to calculate the p-value of the goodness-of-fit test. The validity of this parametric bootstrap is derived in Genest and Rémillard (2008).

An example to illustrate the importance of goodness-of-fit tests is given in Figure 2.3. In the plot 2.3a and 2.3b, we see realizations of a Frank and Gauss copula, respectively. Both dependence structures have the same value of Kendall's tau $\tau = 0.3$. The corresponding empirical lambda functions are given in 2.3c. The blue empirical λ-function corresponds to the Frank copula. The red function corresponds to the Gauss copula. In this case, it is very difficult to distinguish between these dependence structures visually. The goodness-of-fit test, on the contrary, is able to reveal the correct underlying copula in our example. The null hypothesis that the copula in the plot on the left is Gaussian is correctly rejected at a 1% level (p-value = 0.004), and the null hypothesis that the underlying copula in the plot in the middle is a Frank copula is correctly rejected (p-value = 0.001) as well.

2.A Parametric Two-Level Bootstrap Goodness-of-Fit Algorithm

In this goodness-of-fit test, we compare the observed distance between the empirical and the estimated parametric copula with the distance that we expect under the null hypothesis. Therefore, we approximate the distribution of this distance under the null with a parametric bootstrap procedure. The calculation procedure is provided in Genest *et al.* (2009b, Appendix A), and the validity of the two-level bootstrap procedure is



(a) (b) (c)

Figure 2.3: Pseudo-observations of a Frank (a) and Gauss (b) copula. Both copulas have a Kendall's tau of $\tau = 0.3$. The plot on the right shows the empirical lambda functions. The blue function corresponds to the data in (a) and the red function to (b).

derived in Genest and Rémillard (2008). In particular, the estimation step in line 11 is time-consuming, since this estimation is conducted in every bootstrap iteration. On the other hand, the computation procedure is appropriate for parallelization of the bootstrap iterations. Note that, if the copula family under consideration can be evaluated efficiently on $u \in [0, 1]^2$, it is not necessary to approximate the parametric copula at lines 3-4 and 12-13 in the bootstrap algorithm.

Algorithm 1 Two-level parametric bootstrap goodness-of-fit test of Genest *et al.* (2009b)

Input: pseudo-observations $u_{i,j}^n$, $i = 1, \ldots, n$ and $j = 1, \ldots, d$, parametric copula family $\mathcal{C} = \{C_\theta : \theta \in \Theta\}$, number of primary bootstrap samples N, number of secondary bootstrap samples m

Output: approximation of the p-value

define $C_n(u) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(u_i^n \leq u)$
estimate θ with the maximum likelihood estimator θ_n
generate random sample $(y_1^*, \ldots, y_m^*)$ from distribution C_{θ_n}
approximate $C_{\theta_n}(u)$ by $B_m^*(u) = \frac{1}{m} \sum_{i=1}^m \mathbb{1}(y_i^* \leq u)$
compute $S_n = \sum_{i=1}^n \left(C_n(u_i^n) - B_m^*(u_i^n)\right)^2$
for $k = 1, \ldots, N$ **do**
    generate random sample $(y_{1,k}^*, \ldots, y_{n,k}^*)$ from distribution C_{θ_n}
    compute ranks $(r_{1,k}^*, \ldots, r_{n,k}^*)$ of $(y_{1,k}^*, \ldots, y_{n,k}^*)$
    compute pseudo-observations $(u_{1,k}^*, \ldots, u_{n,k}^*) = (\frac{r_{1,k}^*}{n+1}, \ldots, \frac{r_{n,k}^*}{n+1})$
    define $C_{n,k}^*(u) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(u_{i,k}^* \leq u)$
    estimate θ from $(u_{1,k}^*, \ldots, u_{n,k}^*)$ with the maximum likelihood estimator $\theta_{n,k}^*$
    generate random sample $(y_{1,k}^{**}, \ldots, y_{m,k}^{**})$ from distribution $C_{\theta_{n,k}^*}$
    approximate $C_{\theta_{n,k}^*}(u)$ by $B_m^{**}(u) = \frac{1}{m} \sum_{i=1}^m \mathbb{1}(y_{i,k}^{**} \leq u)$
    compute $S_{n,k}^* = \sum_{i=1}^n \left(C_{n,k}^*(u_{i,k}^*) - B_{m,k}^{**}(u_{i,k}^*)\right)^2$
end for
p-Value $= \sum_{k=1}^N \mathbb{1}(S_{n,k}^* > S_n)/N$

Chapter 3

Pair-Copula Construction

Dependence modeling with copulas has been a vibrant area of research for the last 20 years. The copula approach has improved the understanding of dependence between random variables considerably. Most work, though, was done for the 2-dimensional case, and therefore, a large variety of parametric bivariate copula families exists. Nelsen (2006) gives an excellent introduction to the topic and provides an overview of the most prominent bivariate copula families. However, modeling a high-dimensional dependence structure with a multivariate copula remains challenging. Nevertheless, in numerous applications one needs to find models for the dependence of many variables. See, e.g., Chavez-Demoulin *et al.* (2006) and Brechmann *et al.* (2012).

For most bivariate parametric copula families, generalizations to the d-dimensional case exist, but usually these multivariate copulas lack the desired flexibility. In Section 2.2, we review some of the more promising approaches, but, as discussed, all of these procedures come with some drawbacks. In this chapter, we present the pair-copula construction which is a relatively recent way to build multivariate copulas. This approach can model a broad variety of different dependence structures and it is in many aspects superior to the competing high-dimensional copula models. It uses a cascade of bivariate copulas as building blocks to create a high-dimensional dependence model. The idea goes back to a seminal paper of Joe (1996). Since there are numerous different ways to assemble a pair-copula, Bedford and Cooke (2001, 2002) have developed a graphical representation, called regular vines, to visualize the structure of pair-copula constructions. Due to this representation, pair-copulas are sometimes even called vine copulas. This graphical representation is an important tool in pair-copula modeling. However, it requires a certain amount of graph theoretical knowledge to work with these dependence models. Another drawback of the regular vine concept is that the implementation of simulation and estimation algorithms for pair-copulas requires a different representation. To overcome these problems, we introduce a new, matrix-based representation for pair-copulas. Using this representation, we present new algorithms for pair-copulas. In addition, we develop a tool to specify the structure of a pair-copula without any knowledge of graph theory. Thus, we hope to make the pair-copula concept accessible for a broader audience.

The remainder of this chapter is structured as follows. In Section 3.1, we introduce the concept of pair-copula constructions. In the next three sections, we present dif-

ferent ways to represent the structure of a pair-copula. In Section 3.2, we present a graphical approach, and in Section 3.3, we discuss a matrix based approach that has been developed for determining the number of regular vines, but can be used for the implementation of algorithms as well. In Section 3.4, we present the new, matrix-based representation for pair-copula structures. The matrix representation exploits the fact that we can store the structure of a pair-copula construction in an upper diagonal matrix. This is a consequence of Kurowicka and Cooke (2003, Lemma 3.5). Unfortunately, the proof that the authors have stated for the lemma is not correct. Therefore, we prove this lemma in Appendix 3.A. Algorithms that are based on this new matrix representation are given in Appendix 3.C.

3.1 Definition, Concept, and Properties

There are two different ways to introduce the pair-copula construction. In Section 3.1.1, we discuss the pair-copula composition. This integral-based approach has already been established in the first paper on pair-copulas by Joe (1996). The second approach, given in Section 3.1.2, is more popular in the recent literature. See, e.g., Aas *et al.* (2009) and Kurowicka and Joe (2011). There, one decomposes any multivariate distributional density into a sequence of bivariate copulas and univariate densities. The integral-based pair-copula composition is of special importance in the following since we develop a related approach for Lévy copulas in Chapter 6. This is due to the fact that it is not adequate to work with a probability density decomposition in the setting of Lévy processes. Finally, in Section 3.1.3, we discuss an important assumption for the applicability of pair-copula constructions. This assumption plays a crucial role in the estimation of parametric pair-copula models.

3.1.1 Pair-Copula Composition

In the first approach, we build high-dimensional copulas from bivariate ones. Since copulas are distribution functions, we can apply the well-known theory for conditional distributions on copulas. In Appendix 3.B we show how to evaluate the conditional distribution functions in the pair-copula setting. First, we fix some notation. Let $C_{i,j_1,\dots j_m}$ be a copula on the set of variables with indices $\{i, j_1, \dots, j_m\}$. Then, we denote by $F_{i|j_1,\dots,j_m}$ the corresponding conditional distribution function of variable i, given the set of variables $\{j_1, \dots, j_m\}$. Using this notation, we introduce the concept of pair-copula constructions as in Joe (1996).

We start in three dimensions and select any bivariate copula $C_{1,2}$ for the dependence between the first and second variable, and a second bivariate copula $C_{2,3}$ for the dependence between the second and third dimension. We use another bivariate copula, which

we denote by $C_{1,3|2}$, and define the 3-dimensional copula

$$C_{1,2,3}(u_1, u_2, u_3) = \int_{[0,u_2]} C_{1,3|2}(F_{1|2}(u_1|z_2), F_{3|2}(u_3|z_2))dz_2. \tag{3.1}$$

Note that $C_{1,2,3}$ is a well-defined copula function for any choice of $C_{1,2}$, $C_{2,3}$, and $C_{1,3|2}$. In Subsection 4.1.1, we discuss the interpretation of $F_{1|2}(u_1|u_2)$ and $F_{3|2}(u_3|u_2)$ as conditional pseudo-observations in detail. For $u_2 \in [0,1]$ the two equations $F_{1|2}(1|u_2) = 1$ and $F_{3|2}(1|u_2) = 1$ hold. Now, it is easy to check that the bivariate margins of this copula satisfy $C_{1,2,3}(u_1, u_2, 1) = C_{1,2}(u_1, u_2)$ and $C_{1,2,3}(1, u_2, u_3) = C_{2,3}(u_2, u_3)$, respectively. This shows that the choice of the copula $C_{1,3|2}$ only affects the bivariate dependence between the first and third variable as well as the 3-dimensional dependence structure. In order to construct a 4-dimensional copula on the variables $\{1, 2, 3, 4\}$, we need to define a second 3-dimensional copula $C_{2,3,4}$ on the set of variables $\{2, 3, 4\}$. It is important for the definition of the 4-dimensional pair-copula that $C_{2,3,4}$ has the same bivariate margin as $C_{1,2,3}$ on the variables $\{2, 3\}$. Therefore, we reuse the bivariate copula $C_{2,3}$ from the previous step. Additionally, we select the new bivariate copulas $C_{3,4}$ and $C_{2,4|3}$. Now, we can define the second 3-dimensional copula as

$$C_{2,3,4}(u_2, u_3, u_4) = \int_{[0,u_3]} C_{2,4|3}(F_{2|3}(u_2|z_3), F_{4|3}(u_4|z_3))dz_3. \tag{3.2}$$

Note that we use the bivariate copula $C_{2,3}$ to specify both the conditional distribution functions $F_{3|2}$ in Equation (3.1) and $F_{2|3}$ in Equation (3.2). Therefore, the 3-dimensional copulas $C_{1,2,3}$ and $C_{2,3,4}$ have the same dependence structure between the variables u_2 and u_3. For the 4-dimensional copula, we reuse both 3-dimensional copulas $C_{1,2,3}$ and $C_{2,3,4}$ to define $F_{1|2,3}$ and $F_{4|2,3}$, respectively. Furthermore, we select a new, bivariate copula, denoted by $C_{1,4|2,3}$, to define

$$C_{1,2,3,4}(u_1, u_2, u_3, u_4) = \int_{[0,u_2]\times[0,u_3]} C_{1,4|2,3}(F_{1|2,3}(u_1|z_2, z_3), F_{4|2,3}(u_4|z_2, z_3))dC_{2,3}(z_2, z_3). \tag{3.3}$$

Again, $C_{1,2,3,4}$ is a valid copula and we can continue this procedure up to arbitrary dimensions. Of course, any permutation of the indices in this example gives a copula function as well. It is important to note that interchanging the indices is not the only possibility to construct a valid 4-dimensional pair-copula. Alternatively, the bivariate copulas $C_{2,3}$, $C_{2,4}$ and $C_{3,4|2}$ can be used to define

$$C_{2,3,4}(u_2, u_3, u_4) = \int_{[0,u_2]} C_{3,4|2}(F_{3|2}(u_3|z_2), F_{4|2}(u_4|z_2))dz_2$$

instead of the copula in Equation (3.2). Together with the copula $C_{1,2,3}$ from the previous example, we define the 4-dimensional copula similar to Equation (3.3). Note that this pair-copula cannot be obtained from the first pair-copula by a permutation of the indices. There are many different ways to assemble the bivariate copulas in higher di-

mensions. Moreover, we can use any bivariate copula within the construction process. That is, we can choose the bivariate building blocks from a large number of different parametric families, and it is even possible to combine Archimedean and elliptical copulas. Therefore, we can model a broad variety of high-dimensional dependence structures. For these reasons, the pair-copula construction is such a flexible concept.

3.1.2 Density Decompositions

Decomposing the density of any high-dimensional distribution function into bivariate building blocks is a different approach to introduce pair-copulas. This procedure is more common in the recent literature, e.g., Aas *et al.* (2009), than the integral-based pair-copula composition. In Section 3.1.1, we sequentially build parametric copula models. Here, we start with a high-dimensional density and decompose it into bivariate functions of conditional distributions and univariate densities. First, we recall some well-known properties of multivariate density functions. Under the appropriate regularity conditions we have

$$f_{1,\dots,d}(x_1,\dots,x_d) = c_{1,\dots,d}\big(F_1(x_1),\dots,F_d(x_d)\big) f_1(x_1),\dots,f_d(x_d), \tag{3.4}$$

where $c_{1,\dots,d}$ is a copula density. The calculation rules for conditional densities give

$$\begin{aligned} f_{1,\dots,d}(x_1,\dots,x_d) = {} & f_{1|2,\dots,d}(x_1|x_2,\dots,x_d) \\ & f_{2|3,\dots,d}(x_2|x_3,\dots,x_d)\dots f_{d-1,d}(x_{d-1}|x_d) f_d(x_d). \end{aligned} \tag{3.5}$$

Again, any permutation of the indices gives a different decomposition. In the next example, we present the 3-dimensional case. We start with the density $f_{1,2,3}$ and apply the factorization of Equation (3.5). Moreover, we permute the second and the third index. This results in

$$f_{1,2,3}(x_1,x_2,x_3) = f_{1|2,3}(x_1|x_2,x_3) f_{3|2}(x_3|x_2) f_2(x_2). \tag{3.6}$$

Basic calculations give

$$f_{3|2}(x_3|x_2) = c_{2,3}(F_2(x_2),F_3(x_3)) f_3(x_3), \tag{3.7}$$

and for the first factor on the right in Equation (3.6)

$$\begin{aligned} f_{1|2,3}(x_1|x_2,x_3) &= \frac{f_{1,3|2}(x_1,x_3|x_2)}{f_{3|2}(x_3|x_2)} \\ &= c_{1,3|2}(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2)) f_{1|2}(x_1|x_2) \\ &= c_{1,3|2}(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2)) c_{1,2}(F_1(x_1),F_2(x_2)) f_1(x_1) \end{aligned} \tag{3.8}$$

holds. Using Equation (3.6) together with the Equations (3.7) and (3.8), we get the following decomposition

$$\begin{aligned} f_{1,2,3}(x_1, x_2, x_3) = & \, c_{1,3|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \\ & c_{1,2}(F_1(x_1), F_2(x_2)) c_{2,3}(F_2(x_2), F_3(x_3)) \\ & f_1(x_1) f_2(x_2) f_3(x_3). \end{aligned} \tag{3.9}$$

This example illustrates how we can represent a 3-dimensional density with bivariate building blocks together with the corresponding marginal distributions. The same procedure is possible for any other factor in Equation (3.5).

To point out the similarity between this approach and the integral-based representation from Section 3.1.1, we present the density of the 3-dimensional pair-copula from Equation (3.1)

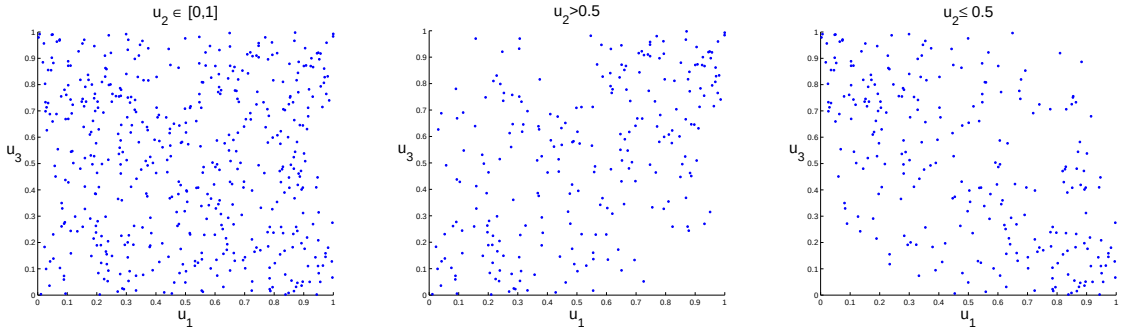
$$c_{1,2,3}(u_1, u_2, u_3) = c_{1,3|2}(F_{1|2}(u_1|u_2), F_{3|2}(u_3|u_2)) c_{1,2}(u_1, u_2) c_{2,3}(u_2, u_3). \tag{3.10}$$

Note that we work with the density of a copula, and therefore, the univariate marginal distribution functions are the identity function and the univariate marginal densities satisfy $f_1 \equiv f_2 \equiv f_3 \equiv 1$. Comparing Equation (3.9) and Equation (3.10) shows how the two approaches coincide.

3.1.3 Simplifying Assumption

The simplifying assumption is important to make the pair-copula constructions mathematically tractable. In particular, it is important for fast and robust estimation procedures. This assumption states that all bivariate copulas in the pair-copula construction are constant and depend on the conditional variables only through their arguments. In the decomposition, given in Equation (3.10), the simplifying assumption guarantees that the bivariate copula $C_{1,3|2}(F_{1|2}(u_1|u_2), F_{3|2}(u_3|u_2))$ does depend on the variable u_2 solely through the conditional distribution functions $F_{1|2}(u_1|u_2)$ and $F_{3|2}(u_3|u_2)$. The assumption assures that the copula $C_{1,3|2}$ itself does not change with the variable u_2. However, it is possible to construct probability distributions in such a way that the pair-copula decomposition leads to bivariate copulas that vary with the conditional variable. Here, we give a particularly easy example of a 3-dimensional pair-copula construction where the simplifying assumption does not hold. That is, the bivariate copula $C_{1,3|2}\left(F_{1|2}(u_1|u_2), F_{3|2}(u_3|u_2)|u_2\right)$ does depend directly on the variable u_2. We use the pair-copula from Equation (3.1), and set $C_{1,2}$ and $C_{2,3}$ to the independence copula Π. Furthermore, we specify

$$C_{1,3|2}(u_1, u_3|u_2) = \begin{cases} C^{\text{Gauss}}(u_1, u_3; \rho = 0.5) & \text{if } u_2 > 0.5, \\ C^{\text{Gauss}}(u_1, u_3; \rho = -0.5) & \text{if } u_2 \leq 0.5. \end{cases} \tag{3.11}$$



(a) (b) (c)

Figure 3.1: The plots show a 3-dimensional realization of a pair-copula, where $C_{1,2}$ and $C_{2,3}$ are set to the independence copula Π. The copula $C_{1,3|2}$ depends directly on u_2 and is specified in Equation (3.11). We present the bivariate margin $\{1,3\}$ for different values of the second variable.

This clearly violates the simplifying assumption, since the copula $C_{1,3|2}$ depends directly on the parameter u_2. In Figure 3.1, we illustrate the effect of this specification on the bivariate margin $\{1,3\}$. Looking only at (a), we cannot observe any dependence between the first and third variables, but (b) and (c) reveal the fact that these variables are dependent, conditional on the value of the second variable. For more examples on this topic, we refer to Hobæk Haff *et al.* (2010).

At the moment, there is a vivid debate on the necessity of the simplifying assumption for pair-copula constructions in the academic community, see Hobæk Haff *et al.* (2010) and Acar *et al.* (2012). Nevertheless, the simplifying assumption is essential for inference on pair-copulas in more than three dimensions. If the pair-copulas do not depend on the conditional variables directly, Aas *et al.* (2009) present estimators for the parameters of the pair-copula. If, on the other hand, the pair-copulas do depend on the conditional variables, it is not straightforward to estimate the parameters in the pair-copula construction. In a remarkable paper, Acar *et al.* (2012) give a kernel-based approach for the 3-dimensional case. Though, inference for higher dimensions without the simplifying assumption remains an open problem. In the following, we presume that the simplifying assumption holds. Hobæk Haff *et al.* (2010) show that any elliptical model with a positive definite scale matrix can be decomposed into a simplified pair-copula construction. In particular, the Gaussian and the t copula can both be represented by the simplified pair structure, and in many other cases, the simplified pair-copula construction gives good approximations, even if the simplifying assumption is not fulfilled.

In conclusion, one has to keep in mind that the simplifying assumption is a very natural assumption in the pair-copula composition in Section 3.1.1. In this case, we simply build a flexible, high-dimensional, parametric model from bivariate copulas. Allowing additionally that the bivariate copulas depend on the conditional variables directly creates a new class of extremely flexible but less tractable, high-dimensional copula models. Of course, parametric models are never perfect, but the simplified pair-copula construc-

tion has proven, e.g., in Aas and Berg (2011) or Czado *et al.* (2011), to be an excellent approach for dependence modeling in different applications.

3.2 Regular Vine Representation

Pair-copula constructions are flexible dependence models that offer the possibility to construct high-dimensional copulas by using bivariate building blocks. There are different ways to assemble the 2-dimensional copulas. The copula density decompositions, introduced in Section 3.1.2, for

$$c_{1,2,3,4} = c_{1,4|2,3} \cdot c_{1,3|2} \cdot c_{2,4|3} \cdot c_{1,2} \cdot c_{2,3} \cdot c_{3,4} \tag{3.12}$$

and

$$c_{1,2,3,4} = c_{3,4|1,2} \cdot c_{2,3|1} \cdot c_{2,4|1} \cdot c_{1,2} \cdot c_{1,3} \cdot c_{1,4} \tag{3.13}$$

are both valid pair-copula densities in four dimensions. For readers' convenience, we omit the unambiguous arguments of the copula densities. In Equation (3.12), we model the 2-dimensional margins $\{1,2\}$, $\{2,3\}$, and $\{3,4\}$ with a bivariate copula, and in Equation (3.13), we specify the margins $\{1,2\}$, $\{1,3\}$, and $\{1,4\}$ directly. Note that relabeling the dimensions does not account for these differences.

Whereas it is feasible to keep track of all possible pair-copula structures in four dimensions by using the density decomposition, this task becomes more difficult in higher dimensions, since the number of possible pair-copula constructions rises dramatically with the number of variables. There are already more than $4.8 \cdot 10^{14}$ possible structures for a 10-dimensional pair-copula construction. In general, we have $\frac{d!}{2} \times 2^{\binom{d-2}{2}}$ possible structures in d dimensions, see Morales-Nápoles (2011). Note that pair-copulas are designed for high-dimensional applications and that this approach can be applied for dependence modeling with 100 or even more variables. Therefore, we need a representation to visualize the structure of pair-copula constructions in order to interpret, compare, and choose the best pair-copula for a given application. For these reasons, Bedford and Cooke (2001) present a graphical model, called regular vines, in which they code every bivariate copula in the pair-copula construction by an edge in this vine. Thereby, they can represent any pair-copula construction within this graphical framework. These regular vines allow to keep track of the possible decompositions, they are easy to interpret, and they specify the structure of a pair-copula uniquely. The regular vine structures are based on an echelon of trees, i.e., connected graphs with no cycles. We define regular vines as in Bedford and Cooke (2002, Definition 4.1).

Definition 3.1 *$\mathscr{V}$ is a* regular vine *on d elements if all of the following conditions hold.*

(i) $\mathscr{V} = (T_1, \ldots, T_d)$.

(ii) *T_u is a tree with edge set E_u and node set $N_u = E_{u-1}$, with $\#N_u = d - (u-1)$ for $u = 1, \ldots, d$, where we # denotes the cardinality of a set. We specify* $N_1 = \{1, \ldots, d\}$.

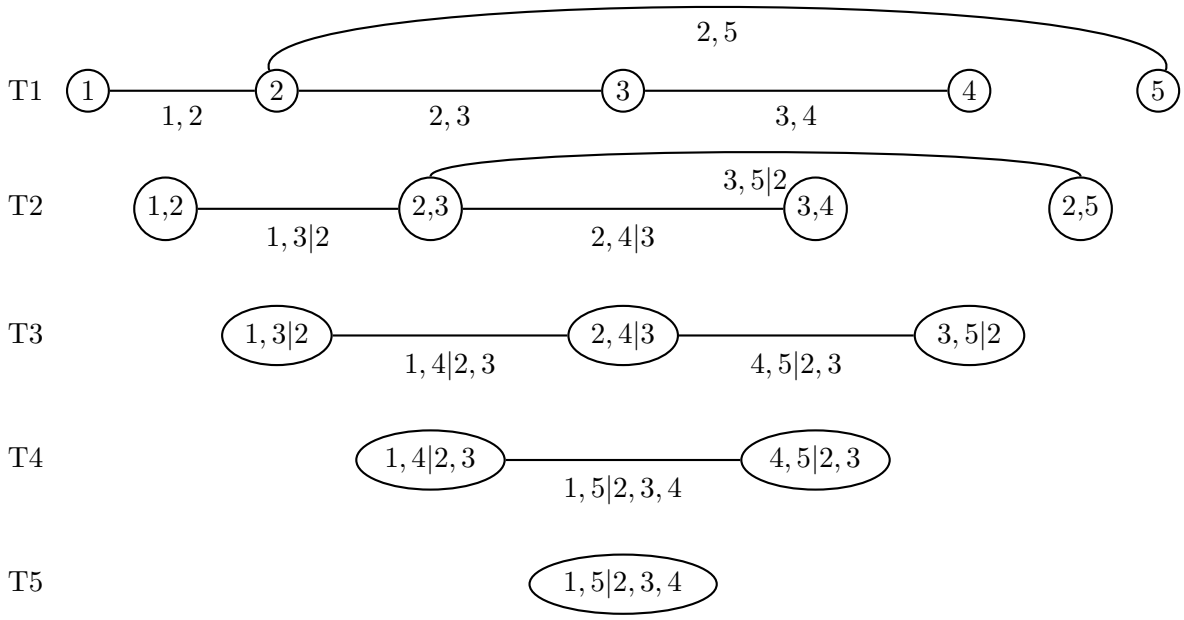


Figure 3.2: Regular vine $\mathscr{V} = (T_1, \ldots, T_5)$. The nodes in the first tree T_1 code the variables $u_1, \ldots, u_5$. Each of the $d(d-1)/2$ edges in the vine represents one bivariate copula in the pair-copula construction.

(iii) The proximity condition *holds: for* $u = 2, \ldots, d-1$, *if* $e^1 = \{i^1, j^1\}$ *and* $e^2 = \{i^2, j^2\}$ *are two nodes in* N_u *connected by an edge (recall* $i^1, i^2, j^1, j^2 \in N_{u-1}$*), then* $\#(e^1 \cap e^2) = 1$.

Figure 3.2 shows a regular vine. The proximity condition can be checked for every pair of connected nodes. For example, the nodes $\{1,2\}$ and $\{2,3\}$ in the second tree are connected by an edge. These two nodes contain both the node $\{2\}$ of the first tree, and therefore, satisfy the proximity condition $\#(\{1,2\} \cap \{2,3\}) = \#\{2\} = 1$. We will see that this condition is central for the applicability of the regular vine concept on pair-copula construction since it guarantees overlapping margins. In order to interpret the higher trees in Figure 3.2, we introduce some additional notation for regular vines in the following Definition. This notation is based on Bedford and Cooke (2002, Definition 4.2 and Definition 4.3) and Cooke *et al.* (2011, Definition 3.3).

Definition 3.2 *For a regular vine* $\mathscr{V}$*, we give the following definitions.*

(i) For any $e_u \in E_u$*, the* complete union *of* e_u *is the subset* $A_{e_u} = \{i \in N_1 : \exists e_k \in E_k, \text{with } i \in e_1, e_k \in e_{k+1} (k = 1, \ldots, u-1)\}$.

(ii) For a regular vine and an edge $e_u \in E_u$*, the* k-fold union *of* e_u $(0 < k \leq u)$ *is the subset* $U_{e_u}(k) = \{e_{u-k} \in E_{u-k} : \exists \text{ edges } e_l \in E_l \ (l = u-k+1, \ldots, u-1) \text{ with } e_l \in e_{l+1} \ (l = u-k, \ldots, u-1)\}$. *For* $k = 0$*, define* $U_{e_u}(0) = \{e_u\}$ *and for* $k = -1$*, define* $U_{e_u}(-1) = \emptyset$.

(iii) *For* $e = \{i, j\} \in E_u$, $u = 1, \ldots, d-1$, *the* conditioning set *associated with* e *is* $D_e = A_i \cap A_j$, *and the* conditioned sets *associated with* e *are* $C_{e,i} = A_i \backslash D_e$ *and* $C_{e,j} = A_j \backslash D_e$. *The* constraint set *for the edge* e *is defined as* $\mathscr{C}e = \{(C_{e,i}, C_{e,j}|D_e)\}$. *The constraint set of the full vine is given by* $\mathscr{C}\mathscr{V} = \{\mathscr{C}e : e \in E_u (u = 1, \ldots, d-1)\}$. *Note that* $A_e = A_i \cup A_j = C_{e,i} \cup C_{e,j} \cup D_e$.

(iv) *If node* e *is an element of node* f, *we say that* e *is a* child *of* f.

As an immediate consequence of Bedford and Cooke (2002, Lemma 4.3), the constraint set specifies the edge uniquely in the vine. Therefore, we denote the edges by their constraint set. Again, we illustrate this concept with the 5-dimensional regular vine in Figure 3.2. The different trees in this vine are denoted by $T_1, \ldots, T_5$. The edges in the vine are labeled by their corresponding constraint sets. Note that

$$e_{1,2} = \{1,2\}, \quad e_{1,3|2} = \big\{\{1,2\},\{2,3\}\big\}, \quad e_{1,4|2,3} = \Big\{\big\{\{1,2\},\{2,3\}\big\},\big\{\{2,3\}\{3,4\}\big\}\Big\}.$$

Since $1 \in e_{1,2} \in e_{1,3|2} \in e_{1,4|2,3}$, it is easy to see that $1 \in A_{e_{1,4|2,3}}$. Furthermore, the k-fold unions $U_{e_{1,4|2,3}}(k)$ for $k = 1, 2, 3$ are given by

$$\begin{aligned} U_{e_{1,4|2,3}}(1) &= \{e_{1,3|2}, e_{2,4|3}\}, \\ U_{e_{1,4|2,3}}(2) &= \{e_{1,2}, e_{2,3}, e_{3,4}\}, \\ U_{e_{1,4|2,3}}(3) &= A_{e_{1,4|2,3}} = \{1,2,3,4\}, \end{aligned}$$

and the regular vine in Figure 3.2 represents the pair-copula construction with density

$$c_{1,2,3,4,5} = c_{1,5|2,3,4} \cdot c_{1,4|2,3} \cdot c_{4,5|2,3} \cdot c_{1,3|2} \cdot c_{2,4|3} \cdot c_{3,5|2} \cdot c_{1,2} \cdot c_{2,3} \cdot c_{3,4} \cdot c_{2,5}.$$

All nodes that are connected by an edge in a tree satisfy the proximity condition of Definition 3.1 (iii). The nodes $e_{1,4|2,3} = \{e_{1,3|2}, e_{2,4|3}\}$ and $e_{4,5|2,3} = \{e_{2,4|3}, e_{3,5|2}\}$ in N_4, for example, share the common node $e_{2,4|3}$ in N_3.

The proximity condition is the reason why the graphical concept of regular vines is suitable for describing pair-copula constructions. The character of this condition becomes particularly clear in the pair-copula composition in Section 3.1.1. There, the condition guarantees that in each step, the u-dimensional copulas, which are combined to $(u+1)$-dimensional ones, have the same $(u-1)$-dimensional margin. This is important, since we build the $(u+1)$-dimensional copulas by integrating with respect to this joint $(u-1)$-dimensional margin.

The most prominent regular vine structures are C- and D-vines. C-vines or canonical-vines have one central node in the first tree. That is, all edges in the first tree share this common node. This structure is of special importance if the multivariate model has one outstanding variable with a leading influence on the other variables. A D-vine or drawable vine can be seen as the opposite of the C-vine structure. There, none of the nodes is an element in more than two edges. It is particularly easy to visualize the D-vine structure and in the first paper on pair-copula constructions by Joe (1996), only

D-vine structures were considered. The regular vine structure in Figure 3.2 is a D-vine, and the 4-dimensional copula in Equation (3.12) is an example of a D-vine pair-copula construction, whereas Equation (3.13) shows a C-vine decomposition.

3.3 Vine Array Representation

Morales-Nápoles (2011) introduces vine arrays which are a different approach to represent pair-copula constructions. This representation is based on a lower triangular matrix. Vine arrays have their origin in finding the number of possible regular vines for a given dimension d. The entries of a vine array represent the nodes of the trees in a vine. Therefore, we need to specify $d(d+1)/2$ entries for the complete vine array representation. This representation is based on a lower triangular matrix including the main diagonal. Before we can state the vine array representation, we need to introduce the concept of natural orderings for regular vines as given in Morales-Nápoles (2011).

Definition 3.3 *Let $\mathcal{V}$ be a d-dimensional regular vine. The natural order $NO(\mathcal{V})$ is a sequence $(A_d, A_{d-1}, \ldots, A_1)$ of the nodes in the first tree of the vine. We set the A_i in a sequential way. We specify A_d to be the smaller element and A_{d-1} to be the larger element of the two elements in the conditioned set of the single node in tree T_d. This node has two children in tree T_{d-1}, and A_{d-1} appears in one of the conditioned sets. We set A_{d-2} to be the second element in this conditioned set. Again, this node (with conditioned set $\{A_{d-1}, A_{d-2}\}$ has two children in tree T_{d-2}, where A_{d-2} appears in one of the conditioned sets. We iterate this procedure sequentially until A_1.*

The natural order of the regular vine in Figure 3.2 is $(1, 5, 4, 2, 3)$. This natural order does not specify the regular vine completely. Nevertheless, we need this ordering to introduce vine arrays, as in Morales-Nápoles (2011).

Definition 3.4 *A regular vine array $TA(\mathcal{V}) = \{A_{i,j}\}$ for $i, j = 1, \ldots, d$ and $j \geq i$ is a lower triangular matrix with elements in $\{1, \ldots, d\}$ indexed in "reverse order" (see Equation (3.14)), where $A_{j,j}$ equals the element in position j in $NO(\mathcal{V})$ and $A_{j-1,j}$ equals the element in position $j-1$ in the same natural order. The element $A_{i,j}$ codes the node $(A_{j,j}, A_{i,j}|A_{i-1,j}, \ldots, A_{1,j})$.*

To illustrate the concept of regular vine arrays, we present this representation for the 5-dimensional regular vine in Figure 3.2.

$$\mathrm{TA}(\mathcal{V}) = \begin{pmatrix} A_{5,5} & & & & \\ A_{4,5} & A_{4,4} & & & \\ A_{3,5} & A_{3,4} & A_{3,3} & & \\ A_{2,5} & A_{2,4} & A_{2,3} & A_{2,2} & \\ A_{1,5} & A_{1,4} & A_{1,3} & A_{1,2} & A_{1,1} \end{pmatrix} = \begin{pmatrix} 1 & & & & \\ 5 & 5 & & & \\ 4 & 4 & 4 & & \\ 3 & 3 & 2 & 2 & \\ 2 & 2 & 3 & 3 & 3 \end{pmatrix} \tag{3.14}$$

The single node in tree T_5 with constraint set $(1, 5|2, 3, 4)$ is coded by the element $A_{4,5}$ in the regular vine array. The element $A_{4,5}$ represents the node $(A_{5,5}, A_{4,5}|A_{3,5}, A_{2,5}, A_{1,5})$, which is the top node of the tree. Note that the representation of a constraint set is invariant with respect to reordering of the elements within the conditioned set and it is also invariant with respect to reordering of the elements within the conditioning set. The first two elements in the third row, i.e., $A_{3,5}$ and $A_{3,4}$, represent the nodes $(1, 4|2, 3)$ and $(4, 5|2, 3)$, respectively, in tree T_4. Any element $A_{i,j}$ that is not on the main diagonal codes the node in the vine in the following way. The conditioned set of the corresponding node is given by the element itself and by the element on the main diagonal in the same column. The conditioning set is given by all the elements in the same column below the element.

3.4 Matrix Representation

The graphical concept of regular vines has proven to be an important tool in pair-copula modeling. However, there is one drawback of this approach. It is difficult to implement a regular vine structure on a computer. Aas *et al.* (2009) propose to restrict the algorithms to special cases and suggest customized functions for C- and D-vines. This method is sufficient for small dimensions since there are only C- and D-vines in the 4-dimensional case. In higher dimensions, however, proceeding this way is very limiting. The vine array representation, on the other hand, can be used to implement algorithms for arbitrary vine structures. Though, it is not easy to interpret and compare the different dependence structures in this representation. To overcome these problems, we introduce a new representation that allows for the implementation of algorithms for arbitrary vine structures and is still very easy to interpret. To obtain a valid multivariate pair-copula, we need to specify the position of each bivariate copula in the pair composition, as in the regular vine representation. We use the fact that there are $d(d-1)/2$ edges in a vine structure that code the positions of the bivariate copulas in a pair-copula construction. An upper triangular matrix has $d(d-1)/2$ entries too. In this section, we show how we can represent any regular vine, and therefore any pair-copula structure, with an upper triangular matrix by coding the edges of the vine with corresponding matrix entries in the following way. There are two elements in the conditioned set $\{C_{e,i}, C_{e,j}\}$ of an edge e that specify the position of the corresponding entry in the matrix. Thus, we store the conditioning set D_e in the selected matrix entry. One advantage of this matrix representations is that this framework is still very easy to interpret. We present the algorithms, based on the matrix representation, in Appendix 3.C.

Before introducing the matrix representation for pair-copula constructions, we need a Lemma that is already stated in Kurowicka and Cooke (2003, Lemma 3.5). Unfortunately, the proof given there is not correct. Nevertheless, the statement is true and we provide a detailed proof of the Lemma in Appendix 3.A.

Lemma 3.5 *If the conditioned sets of edges i, j in a regular vine are equal, then $i = j$.*

Using this Lemma, the next theorem shows that a matrix representation is indeed possible and we have an unambiguous mapping from the edges of any regular vine to the entries of an upper triangular matrix.

Theorem 3.6 *Every regular vine can be represented in an upper triangular $d \times d$-matrix M, and the vine matrix entries $m_{u,v}$, $u < v$, are defined for every $e = \{i, j\} \in \bigcup_{u=1,\dots,d} E_u$ by*

$$m_{u,v} = D_e, \tag{3.15}$$

where $u = \min\{C_{e,i}, C_{e,j}\}$, and $v = \max\{C_{e,i}, C_{e,j}\}$.

Proof: A regular vine consists of $d(d-1)/2$ copulas, i.e., there are $d(d-1)/2$ edges in a regular vine. An upper triangular $d \times d$-matrix also has $d(d-1)/2$ entries and by Lemma 3.5, it is not possible that any entry can be used for more than one edge. □

Another advantage of the matrix representation for regular vines is the fact that the matrices are quite compact. Note that we omit the braces of the sets in the matrix representation. The regular vine presented in Figure 3.2, for example, can be represented by the following matrix.

	1	2	3	4	5
1		∅	2	2,3	2,3,4
2			∅	3	∅
3				∅	2
4					2,3

Recall that the edges in the first tree of the vine in Figure 3.2 are $e_{1,2} = \{1, 2\}$, $e_{2,3} = \{2, 3\}$, $e_{2,5} = \{2, 5\}$, and $e_{3,4} = \{3, 4\}$. These edges are represented by the matrix entries $m_{1,2} = m_{2,3} = m_{2,5} = m_{3,4} = \emptyset$, respectively. The second tree in the vine consists of the edges $e_{1,3|2}$, $e_{2,4|3}$, and $e_{3,5|2}$. Therefore we set the matrix entries $m_{1,3} = 2$, $m_{2,4} = 3$, and $m_{3,5} = 2$. In the third tree we represent the edges $e_{1,4|2,3}$ and $e_{4,5|2,3}$ by the entries $m_{1,4} = \{2, 3\}$ and $m_{4,5} = \{2, 3\}$. Finally, we store the single edge $e_{1,5|2,3,4}$ of the fourth tree in the last free entry of the upper triangular matrix M. That is, we set $m_{1,5} = \{2, 3, 4\}$.

The C- and D-vine copula, given in the decompositions in Equations (3.12) and (3.13), can be represented by the two matrices

	1	2	3	4
1		∅	2	2,3
2			∅	3
3				∅

,

	1	2	3	4
1		∅	∅	∅
2			1	1
3				1,2

,

respectively. These examples illustrate that the matrix representation is straightforward to implement on a computer and still very easy to interpret. We can instantly check which bivariate margins are modeled in the first tree and, in contrast to the other representations, we can immediately see that, in the first example, the dependence between the third and fifth variable is modeled conditional on the second variable.

In addition, we present an Excel tool which is based on the matrix representation for specifying regular vines. We do not presume any knowledge on graph theory, since we automatically check for the validity of the pair-copula structure. Figure 3.3 illustrates how to apply the Excel tool for easily creating a vine matrix. In the following example, we build the vine matrix M corresponding to the regular vine structure in Figure 3.2. We start with an empty upper triangular matrix. It is advantageous to specify the most important bivariate margins at a low level in the structure, since the dependence of bivariate margins in the first tree is modeled directly. In this way, we can choose from the broad variety of all bivariate copulas for this margins. Suppose that in this example, the edge $\{1, 2\}$ represents our most important bivariate dependence. Thus, we specify $m_{1,2}$ in (a) as the first entry of the vine matrix. Next, we select the entry $m_{2,3}$ in (b) that corresponds to the edge $\{2, 3\}$. Now, we cannot select the entry $m_{1,3}$ corresponding to the edge $\{1, 3\}$ in the first tree anymore. Remember that the selection of this edge leads to a circle in the first tree of the vine. We continue in (c) and (d) and select the vine matrix entries $m_{2,5}$ and $m_{3,4}$, respectively. Up to now, the selected matrix corresponds to the first tree T_1 in Figure 3.2. By selecting $m_{1,3}$ in (e), we specify the first edge in the second tree and the yellow fields give the possible entries of the vine matrix, so that the proximity conditions are satisfied in the current tree. We continue with the entry $m_{3,5}$ in (f). Now, the entry $m_{2,4}$ is the only possible choice for the last edge in the second tree. We can choose the edges for the third tree from three possibilities and we select $m_{1,4}$ and $m_{4,5}$. Finally, there is only one possible edge left in the fourth tree, so that we complete the selection procedure by choosing $m_{1,5}$ with conditioning set $\{2, 3, 4\}$.

3.5 Conclusion

Pair-copula constructions offer the possibility to reach a broad variety of different dependence structures in high dimensions. The pair-copula consists of $d(d-1)/2$ bivariate copulas and we can specify each of these copulas separately. This procedure allows for a direct modeling of $d-1$ of the 2-dimensional margins. In contrast to hierarchical Archimedean copulas, it is even possible to use different parametric families for each bivariate copula, and there are no restrictions on the parameters of these copulas either. Therefore, we can build high-dimensional dependence models with various different bivariate tail dependence coefficients, which is crucial in many financial applications. Simulation from pair-copulas is straightforward and the sequential structure allows for stepwise estimation algorithms, which is particularly important for high-dimensional applications. As we have seen, there is a huge number of possible ways to assemble the $d(d-1)/2$ bivariate copulas in order to build a valid d-dimensional pair-copula. Regular vines have proven to be an important tool for the visualization of these structures. Especially the first tree of the vine can be interpreted easily. Comparing the different representations in Section 3.2, 3.3 and 3.4, the graphical concept of regular vines is still superior from a didactic point of view. Therefore, we use this representation in the following chapters for explanatory purposes. However, implementing algorithms for pair-

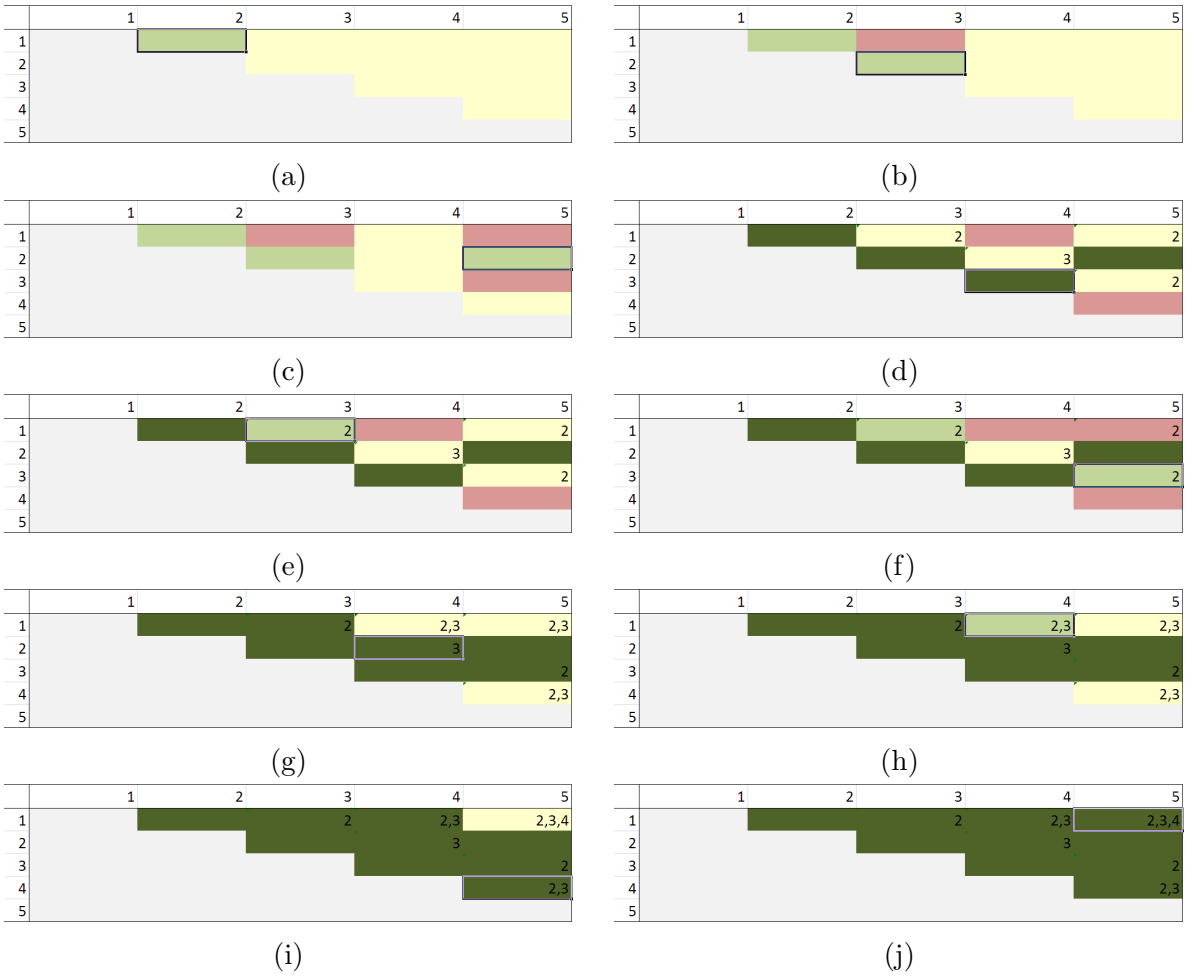


Figure 3.3: Selection process of a regular vine structure in the matrix representation with assistance of the Excel tool. Yellow fields can be chosen for the next vine matrix entry, light green entries are already selected in the current tree, and dark green fields are already selected in the previous trees. Red entries cannot be selected in the current tree, since this would lead to non-valid structures.

copula constructions requires different methods. Regular vine arrays are appropriate for simulation and estimation algorithms. They allow for very elegant algorithms and they are most parsimonious representations. On the other hand, they are harder to interpret than the competing approaches. The matrix representation, presented in Section 3.4, is also well-suited for the development of algorithms. Moreover, it is easy to interpret and, in addition, it can be used to develop tools that help the user to create a valid regular vine structure. Comparing all different representations, we base all algorithms in this work on the matrix representation.

3.A Proofs

In this section, we prove Lemma 3.5. To do so, we need the next remark, which is a direct consequence of the definitions.

Remark 3.7 *Let $\mathcal{V}$ be a regular vine and $u \in \{1, \dots, d-1\}$. Then*

(i) $\bigcup_{e \in E_u} A_e = \{1, \dots, d\}$,

(ii) $\bigcup_{e \in U_{e_u}(k)} A_e = A_{e_u}$.

This lemma is central for the proof.

Lemma 3.8 *Let $\mathcal{V}$ be a regular vine. Let $i = \{i_1, i_2\}$, $j = \{j_1, j_2\} \in E_u$ be connected by an edge e in E_{u+1} and w.l.o.g. suppose that $i_1 = j_1$. Then*

$$A_i \cup A_j = A_i \cup C_{j,j_2} = A_j \cup C_{i,i_2}$$

holds.

Proof:

$$A_i \cup A_j = A_i \cup A_{j_1} \cup A_{j_2} = (A_i \cup A_{j_1}) \cup (A_{j_2} \backslash A_{j_1}) = A_i \cup (A_{j_2} \backslash A_{j_1}) = A_i \cup C_{j,j_2}$$

Switching i and j completes the proof. □

Additionally, we need a more technical Lemma.

Lemma 3.9 *Let $u_1, u_2 > 0$, $i = \{i_1, i_2\} \in E_{u_1+u_2}$, and $j = \{j_1, j_2\} \in E_{u_1}$. If $j \in U_i(u_2)$, then the conditioned sets of i and j are different.*

Proof: We show that the conditioned sets of i and j cannot be equal if j is an element of $U_i(u_2)$. To do so, we first split the set $U_i(u_2)$ into two subsets such that each subset "contains" exactly one of the conditioned sets of i. Since j "contains" both of its conditioned sets and j has to be an element of one of the subsets, the conditioned sets of i and j cannot be equal. More formally: we split

$$U_i(u_2) = U_{i_1}(u_2 - 1) \cup U_{i_2}(u_2 - 1)$$

in two subsets. If j is an element of $U_i(u_2)$, then j is an element of $U_{i_1}(u_2 - 1)$ or $U_{i_2}(u_2 - 1)$. With Remark 3.7(ii), we know that $A_j \subset A_{i_1}$ or $A_j \subset A_{i_2}$. Now it immediately follows that j cannot have the same conditioned sets as i, because the conditioned sets $\{C_{j,j_1}, C_{j,j_2}\} \subset A_j$ and by the definition of the conditioned sets, the sets A_{i_1} and A_{i_2} contain exactly one of the conditioned sets $\{C_{i,i_1}, C_{i,i_2}\}$. Thus, $\{C_{j,j_1}, C_{j,j_2}\} \subset A_j \subset A_{i_1}$ or $\{C_{j,j_1}, C_{j,j_2}\} \subset A_j \subset A_{i_2}$ proves that, under the given conditions, i and j cannot have

the same conditioned sets. □

Now, we are able to state an alternative proof of Kurowicka and Cooke (2003, Lemma 3.5).

Lemma 3.10 *If the conditioned sets of edges i, j in a regular vine are equal, then $i = j$.*

Proof: We split the proof in two cases. First we consider the case where both i and j are elements of E_u, that is, i and j are edges in the same tree. We will prove this case by contradiction. We know that $\cup_{e \in E_u} A_e = \{1, \ldots, d\}$. The idea behind this proof is to show that when the two edges i and j have the same conditioned sets, the number of all elements in the set $\cup_{e \in E_u} A_e$ has to be smaller than d ↯. We show this contradiction by iteratively adding all elements of E_u to a working set. In the second part of the proof, we consider the case where i and j are edges in different trees of the vine.

First part: Suppose that $i, j \in E_u$ have the same conditioned sets and $i \neq j$. We start by counting the elements of E_u in order to know how many elements we have to add to the primarily empty working set, which we denote by M. With Bedford and Cooke (2002, Definition 4.1) we know that

$$\#E_u = d - u. \tag{3.16}$$

The first element that we add to our still empty working set M is $e^{(1)} := i$, such that $M = \{e^{(1)}\}$. With Bedford and Cooke (2002, Lemma 4.2) we get

$$\#A_{e^{(1)}} = u + 1. \tag{3.17}$$

Since T_{u+1} is a connected tree, there is at least one element of $E_u \backslash M$ which is connected to $e^{(1)}$ by an edge in E_{u+1}. We denote one of these elements by $e^{(2)}$ and add it to M. Since $e^{(1)} = \{i^{(1)}, j^{(1)}\}$ and $e^{(2)} = \{i^{(2)}, j^{(2)}\}$ are connected, we use the proximity condition and set w.l.o.g. $i^{(1)} = i^{(2)}$. With Lemma 3.8 we get

$$\#(A_{e^{(1)}} \cup A_{e^{(2)}}) = \#(A_{e^{(1)}} \cup C_{e^{(2)},j^{(2)}}) \leq u + 2. \tag{3.18}$$

In the special case, when we add j to the working set M, that is, $e^{(2)} = j$ and $M = \{i, j\} = \{e^{(1)}, e^{(2)}\}$, we get the following, stronger inequality

$$\#(A_{e^{(1)}} \cup A_{e^{(2)}}) = \#(A_{e^{(1)}} \cup C_{e^{(2)},j^{(2)}}) < u + 2, \tag{3.19}$$

since $C_{e^{(2)},j^{(2)}} \subset \{C_{j,j_1}, C_{j,j_2}\} \subset A_i = A_{e^{(1)}}$. This is the case because i and j have the same conditioned sets.

Now, we repeat this procedure and iteratively add all elements of E_u to the working set M. Finally, we get $M = E_u$. This is possible since T_{u+1} is a connected tree and there is always an element of $E_u \backslash M$ which is connected by an edge in E_{u+1} to an element in

M. Note that we started with $\#A_{e^{(1)}} = u+1$. In $d-u-1$ steps we have added at most one element to the set $\cup_{e\in M} A_e$. In one step we had to add j, which is different from i but has the same conditioned set, to the set M. Therefore, we did not add anything to $\cup_{e\in M} A_e$. This leads to

$$\#(\cup_{e\in E_u} A_e) < u+1+d-u-1 = d \tag{3.20}$$

in contradiction to $\#(\cup_{e\in E_u} A_e) = \#\{1,\dots,d\} = d$. $\lightning$

Second part: In the second case, we say w.l.o.g. that i is an edge in a higher tree of the vine than j, that is, $j \in E_{u_1}$ and $i \in E_{u_1+u_2}$. We also prove this case by contradiction. Again we know that $\cup_{e\in E_{u_1}} A_e = \{1,\dots,d\}$. We show that when the two edges i, j have the same conditioned sets, the number of all elements in the set $\cup_{e\in E_{u_1}} A_e$ has to be smaller than d $\lightning$. We show this contradiction by iteratively adding all elements of E_u to a working set. In contrast to the first case, we do not start with only one element in the working set, but with elements that can be easily deduced from i.

Suppose that $j \in E_{u_1}$ and $i \in E_{u_1+u_2}$ have the same conditioned sets and $u_1, u_2 > 0$. Since the edges i, j are not in the same tree anymore, we start with the working set $M := U_i(u_2)$. With Lemma 3.9, we can guarantee that $j \notin M$. This enables us to use the set E_{u_1}, in spite of the fact that the edges i, j are in different trees. By definition $M \subset E_{u_1}$ and with Bedford and Cooke (2002, Lemma 4.1) $\#M = u_2 + 1$. Furthermore we know that $\#E_{u_1} = d - u_1$. This leads to $\#(E_{u_1} \backslash M) = d - u_1 - u_2 - 1$ which is exactly the number of elements that we still have to add to M. Now, we use Remark 3.7(ii) and get

$$\#\bigcup_{e\in M} A_e = \#A_i = u_1 + u_2 + 1. \tag{3.21}$$

Since by Lemma 3.9 $j \notin M$, we have to add j in one of the following $d - u_1 - u_2 - 1$ steps to the working set. With the same argumentation as in the first case, we are now able to show that

$$\#(\cup_{e\in E_{u_1}} A_e) < u_1 + u_2 + 1 + d - u_1 - u_2 - 1 = d \tag{3.22}$$

in contradiction to $\#(\cup_{e\in E_{u_1}} A_e) = \#\{1,\dots,d\} = d$. $\lightning$
□

3.B Conditional Distributions

In Section 3.1, we discuss an approach to model high-dimensional dependence structures by a cascade of bivariate copulas. This procedure uses conditional distribution functions. In estimation and simulation routines, we need to evaluate these conditional distributions for every observation. Therefore, it is important that we can compute these functions

in an efficient way. Joe (1996) shows that this can be done by

$$F_{1|2,\dots,d}(x_1|x_2,\dots,x_d) = \frac{\partial C_{1,2|3,\dots,d}(F_{1|3,\dots,d}(x_1|x_3,\dots,x_d), F_{2|3,\dots,d}(x_2|x_3,\dots,x_d))}{\partial F_{2|3,\dots,d}(x_2|x_3,\dots,x_d)}. \tag{3.23}$$

Now, we have to apply this procedure sequentially to the lower-dimensional conditional distribution functions in Equation (3.23) until we have only unconditional univariate distribution functions left. Note that we might need to adjust the indices in order to use the appropriate bivariate copulas of the pair-copula construction. Aas *et al.* (2009) denote the partial derivative of the bivariate copula function $\partial C(u,v)/\partial v$ as h-function.

3.C Algorithms

In this section, we provide algorithms, written in pseudocode, for the simulation and estimation of pair-copula constructions. The algorithms are based on the matrix representation introduced in Section 3.4. In order to apply the algorithms to any R-vine structure, we make use of auxiliary functions that we define in Algorithms 5, 6, and 7. If we limit ourselves to C- and D-vine structures only, it is also possible to give customized simulation and estimation algorithms without these auxiliary functions. These algorithms, presented in Aas *et al.* (2009), are special cases of our algorithms. In addition to the flexibility, the auxiliary functions allow for very short and simple estimation and simulation routines.

The simulation procedure, presented in Algorithm 2, works sequentially. We start with the first variable and proceed step by step in the pair-copula structure. The auxiliary function in Algorithm 5 chooses the next variable and we simulate this variable, given the first. This procedure is based on Appendix 3.B and the h-function (H-FUNC$(\cdot)$) as well as the inverse of the h-function (H-INVERSE$(\cdot)$) are as in Aas *et al.* (2009). This sequential simulation procedure continues until all variables are simulated. Algorithm 3 evaluates the log-likelihood function for pair-copula constructions, and with Algorithm 4, we provide a routine for the stepwise maximum likelihood approach. This is particularly important for higher dimensions and in finding starting values for the full maximum likelihood approach. The auxiliary functions are used in the simulation or estimation procedures. The function in Algorithm 7 is a straightforward way to recursively evaluate the conditional distribution used in the pair-copula construction. In Algorithm 5, we define a function that selects the next variable given a set A of already simulated variables. Algorithm 6 makes sure that we follow the correct ordering when we use the already simulated variables.

Next, we fix the notation for the following algorithms. By A and B, we denote working sets of indices. As usual, $\#A$ denotes the number of indices in the set, and, as introduced in Section 3.4, the matrix M stores the structure of the pair-copula construction. The bivariate copula $C_{i,j}$ corresponds to the matrix entry $m_{i,j}$ and by $c_{i,j}$, we denote the copula density. Since M is an upper triangular matrix and in order to keep the

pseudocode simple, we use $m_{i,j}$ instead of $m_{i\wedge j,i\vee j}$, where $i \wedge j$ denotes the minimum and $i \vee j$ denotes the maximum of i and j.

Algorithm 2 Simulation algorithm for a pair-copula construction (R-vine structure).

Input: $\omega_1, \ldots, \omega_d \sim U([0,1])$, i.i.d., vine matrix M
Output: realization of a random variable $u_1, \ldots, u_d$ with a given pair-copula distribution and vine matrix M

1: $u_1 \leftarrow \omega_1$
2: $A \leftarrow \{1\}$
3: **for** $i \leftarrow 2, \ldots d$ **do**
4: $\quad j \leftarrow$ selectNextVariable(A)
5: $\quad u_j \leftarrow \omega_j$
6: $\quad B \leftarrow A$
7: $\quad$ **for** $h \leftarrow 1, \ldots i-1$ **do**
8: $\quad\quad k \leftarrow$ selectConditionedIndex(j, B)
9: $\quad\quad B \leftarrow B \backslash \{k\}$
10: $\quad\quad u_j \leftarrow$ h-Inverse$(u_j,$ calculateF$(u_k, B))$
11: $\quad$ **end for**
12: $\quad A \leftarrow A \cup \{j\}$
13: **end for**
14: **return** $u_1, \ldots, u_d$

Algorithm 3 Evaluation of the likelihood function of a pair-copula construction (full maximum likelihood).

Input: observations $(u_{m,1}, \dots, u_{m,d})_{m=1,\dots,n}$, vine matrix M
Output: value of the likelihood function ll for one observation

function CALCULATEFULLLIKELIHOOD
  $ll \leftarrow 0$
  for $i \leftarrow 1, \dots d-1$ **do**
    for $j \leftarrow i+1, \dots d$ **do**
      $ll \leftarrow ll + \sum_{m=1,\dots,n} \log\Big(c_{ij}\big(\text{CALCULATEF}(u_{m,i}, m_{i,j}), \dots \text{CALCULATEF}(u_{m,j}, m_{i,j})\big)\Big)$
    end for
  end for
  return ll
end function

Algorithm 4 Sequential estimation of the parameters in a pair-copula construction

Input: observations $(u_{m,1}, \dots, u_{m,d})_{m=1,\dots,n}$, vine matrix M
Output: upper triangular matrix C with stepwise maximum likelihood parameter estimates

function CALCULATESTEPWISELIKELIHOOD
  for $i \leftarrow 1, \dots d-1$ **do**
    for $j \leftarrow 1, \dots d-1$ **do**
      for $k \leftarrow j+1, \dots d$ **do**
        if $i = 1$ & $m_{j,k} = \emptyset$ **then**
          $(C)_{j,k} \leftarrow \arg\max \sum_{m=1,\dots,n} \log\big(c_{jk}(u_{m,j}, u_{m,k})\big)$
        else if $\#m_{j,k} = i-1$ **then**
          $(C)_{j,k} \leftarrow \arg\max \sum_{k=1,\dots,n} \log\Big(c_{j,k}\big(\text{CALCULATEF}(u_{m,j}, m_{j,k}), \dots \text{CALCULATEF}(u_{m,k}, m_{j,k})\big)\Big)$
        end if
      end for
    end for
  end for
  return C
end function

Algorithm 5 Auxiliary function for the simulation routine. This algorithm selects the index of the next variable in the simulation procedure.

Input: nonempty set A of already simulated indices, vine matrix M
Output: index j of the next variable in the simulation routine

```
1: function selectNextVariable(A)
2:     for i ∈ A do
3:         for j ∈ {1,...,d}\A do
4:             if #A = 1 & m_{i,j} = ∅ then
5:                 return j
6:             else if m_{i,j} = A\{i} then
7:                 return j
8:             end if
9:         end for
10:    end for
11: end function
```

Algorithm 6 Auxiliary function for simulation and estimation algorithms. This algorithms selects an index l in a given set B, such that $(k, l|B\backslash\{l\})$ is the constraint set of a node in a regular vine.

Input: index k, set of indices B such that $k \notin B$, vine matrix M
Output: index l such that $(k, l|B\backslash\{l\})$ is the constraint set of a node

```
1: function selectConditionedIndex(k,B)
2:     for l ∈ B do
3:         if m_{j,l} = B\{l} then
4:             return l
5:         end if
6:     end for
7: end function
```

Algorithm 7 Function to evaluate a conditional distribution function in a pair-copula construction setup.

Input: variable u_k, set of conditioning indices B and conditioning variables $\{u_j|j \in B\}$, vine matrix M
Output: value of the conditional distribution function $F_{k|B}(u_k|\{u_j|j \in B\})$

```
1: function calculateF(u_k, B)
2:     if B = ∅ then
3:         return u_k
4:     else
5:         j ← selectConditionedIndex(k, B)
6:         return h-Func(calculateF(u_k, B\{j}), calculateF(u_j, B\{j}), , θ_{k,j})
7:     end if
8: end function
```

Chapter 4

Model Selection for Pair-Copula Constructions

In the previous chapter, we have discussed the theoretical foundation of the pair-copula construction. This approach turned out to be a highly flexible framework for dependence modeling in arbitrary dimensions. In this chapter, we want to apply this promising method in a real world application. Therefore, we need to discuss in detail how to find an adequate pair-copula model for a given set of data. This is a challenging task, not because there are too few, but because there are too many competing models within the pair-copula framework. The number of parameters for pair-copula models increases quadratically with the dimension. Thus, parameter reduction techniques are important in this approach to keep the dependence model parsimonious and to prevent overfitting. In order to reduce the number of parameters, we introduce a new heuristic, which is based on goodness-of-fit tests. The heuristic selects and replaces specific parametric building blocks by the parameter-free independence copula.

In the first section of this chapter, we demonstrate how to choose an adequate vine structure, select appropriate bivariate copulas, and estimate the parameters of this selection. In Section 4.2, we introduce a new method to reduce the number of parameters in the model, and in Section 4.3, we apply our methods to multivariate financial return data.

4.1 Model Selection

All model selection techniques for pair-copulas that we discuss here base on pseudo-observations. Therefore, we are able to specify the dependence structure without any knowledge of the univariate margins. Thus, in a first step, we transform the data to the pseudo-observations. After that, fitting a pair-copula model to given data consists of three different tasks. Namely, we need to

(i) specify the vine structure for the model,

(ii) select an appropriate family for each bivariate copula in the structure,

(iii) estimate all parameters.

Since we can choose from many different bivariate parametric copulas, and since the vine structure allows for many different ways to assemble these building blocks, pair-copula modeling offers the possibility to model a broad variety of different dependence structures. In this section, we present different heuristics for finding an appropriate specification of the pair-copula construction. But note that none of the available model selection techniques can guarantee an optimal global fit. Generally, the model selection depends on the application, and therefore, customized techniques might be favorable in specific situations.

In the first part of this section, we discuss a sequential heuristic that is popular in the recent literature and is also used in the empirical example in Section 4.3. The next part of this section gives a short overview on other possible procedures for finding appropriate pair-copula models.

4.1.1 Sequential Model Selection

The sequential model selection approach for pair-copula constructions is suggested in Aas *et al.* (2009) for C- and D-vines. This procedure is generalized to all regular vine structures in Dißmann *et al.* (2011). Nikoloulopoulos *et al.* (2012) suggest a very similar procedure based on their experience with financial return data. The stepwise model selection proceeds tree by tree in the vine. That is, we specify the structure of the first tree, select an appropriate parametric family for all the bivariate copulas in this tree and estimate the parameters for the bivariate copulas. Then, we transform the pseudo-observations to conditional pseudo-observations by using the fitted copulas in the first tree. In the following steps, we apply the same procedure to the conditional pseudo-observations and iterate until we have found a complete specification of the pair-copula construction. After specifying the complete structure, one conducts a full maximum likelihood estimation, where we use the sequential parameter estimates as starting values to improve the fit of the model. A detailed discussion of this procedure is given in Dißmann *et al.* (2011). In the following, we discuss the different steps in the sequential model selection procedure.

First Tree

We start with the pseudo-observations and calculate Kendall's tau for each of the $d(d-1)/2$ possible bivariate margins. That is, we measure the strength of all 2-dimensional dependencies within our data set. We specify the first tree in the vine by using a maximum spanning tree algorithm, where the absolute value of Kendall's tau is the weight of the edge. Thereby, we try to capture as much information as possible on the dependence structure in the first level of the vine. Now we have to choose appropriate copula families for the bivariate building blocks that correspond to the edges in the first tree. Therefore, we estimate the parameters separately for all bivariate copula families under consideration and calculate the Akaike information criterion (AIC), see Akaike

(1974). Finally, we select the parametric copula family with the smallest AIC.

Transformation of the Pseudo-Observations

In the first tree, we are able to work with the pseudo-observations directly. Unfortunately, this is no longer possible in the higher trees, since the bivariate copulas in the higher layers of the vine specify conditional dependencies. This is particularly easy to see in pair-copula representation of Section 3.1.1. To illustrate this, we recall the 3-dimensional example of a pair-copula in Equation (3.1)

$$C_{1,2,3}(u_1, u_2, u_3) = \int_{[0,u_2]} C_{1,3|2}(F_{1|2}(u_1|z_2), F_{3|2}(u_3|z_2))dz_2.$$

Obviously, the conditional distribution of the variables u_1 and u_3 given u_2 is specified by $C_{1,3|2}(F_{1|2}(u_1|u_2), F_{3|2}(u_3|u_2))$. Using the simplifying assumption, we know that the bivariate copula function $C_{1,3|2}$ does not change with u_2. Therefore, we transform the pseudo-observations to the conditional pseudo-observations $u_{1|2} = F_{1|2}(u_1|u_2)$ and $u_{3|2} = F_{3|2}(u_3|u_2)$, and then, we use the well-known procedures from copula theory to estimate $C_{1,3|2}$. Remember that we utilize the estimated parametric copula $C_{1,2}$ and $C_{2,3}$ from the first tree to calculate $F_{1|2}$ and $F_{3|2}$, respectively, and thus to compute the conditional pseudo-observations $u_{1|2}$ and $u_{3|2}$.

Subsequent Trees

In the second tree, we proceed as follows. All selected edges in the first tree of the vine turn to nodes in the second tree. Next, we check all possible edges for the proximity condition. If this condition holds, we transform those variables which correspond to the nodes of the edge to conditional pseudo-observations. Then we calculate Kendall's tau for this edge. Again, we build a maximum spanning tree from all edges that satisfy the proximity condition, and thereby, we select the structure of the second tree in the vine. Once more, we try to capture as much dependence in the lower trees as possible. In the next step, we continue by selecting appropriate copula models by the AIC. This procedure is repeated for all subsequent trees until the complete pair-copula model is specified.

4.1.2 Alternative Model Selection Techniques

The model selection technique, presented in the previous section, is not the only possible heuristic to specify the pair-copula construction. One of the competing approaches is given in Kurowicka (2011). There, the model selection proceeds bottom-up. That is, one specifies the edge in the highest tree firstly and continues sequentially until the structure of the first tree is chosen. This model selection procedure is particulary appropriate if the dependence structure is close to a Gaussian copula. A different approach to choose an

appropriate pair-copula construction specifies the complete structure for a small subset of variables at first. Then, it sequentially adds the remaining variables to this model, c.f. Morales-Nápoles (2011). Schnieders (2012, Chapter 4.3) introduces such an approach based on goodness-of-fit tests for C- or D-vines. The different model selection techniques can be classified into

(i) top-down (Section 4.1.1 and Aas *et al.* (2009)),

(ii) bottom-up (Kurowicka (2011)),

(iii) side-to-side (Schnieders (2012))

procedures. Note that many variations are possible within these model selection approaches. In Section 4.1.1, for example, we use Kendall's tau to choose a tree structure in the top-down approach, whereas Vaz de Melo Mendes *et al.* (2010) implicitly use the tail dependence coefficient for defining the structure. Note that the model selection procedure simplifies considerably if we allow for C- and D-vine structures only, as in Aas *et al.* (2009), Nikoloulopoulos *et al.* (2012), and Schnieders (2012). If there exists one extraordinary variable such that most of the dependence is captured by the bivariate margins including this pilot variable, it is reasonable to restrict the regular vines to the C-vine case. Furthermore, a Bayesian approach for model selection, as suggested in Czado and Min (2011), is also possible.

4.2 Parameter Reduction Approach

In this section, we introduce a new approach to check if all parametric building blocks are necessary to characterize the complete dependence model. To the best knowledge of the author, this is the first approach that uses the observations directly and checks all building blocks separately for their relevance.

The pair-copula construction is a d-dimensional parametric model that is built from $d(d-1)/2$ bivariate ones. Thus, the number of building blocks increases quadratically with the dimension. Using, for instance, only one-parametric bivariate copula families as building blocks, like the Gauss or Frank copula, we already have $d(d-1)/2$ parameters that specify the dependence model. If we select bivariate copula families with more than one parameter, like the t or BB1 copula, this number increases even more. Remember that pair-copula models only account for the dependence structure and do not include any information on the margins. Thus, the number of parameters for the full model can be considerably larger. Working with such high-parametric models might be appropriate in situations where we have a sufficient number of observations. However, in many cases, we need to model a high-dimensional dependence structure while we only have a limited number of observations. In financial applications, for example, we can only use a short time horizon to calibrate the model, since the dependence structure might change over time, and therefore, the number of relevant observations is low. Thus, it is important to have a parsimonious dependence model to avoid overfitting. One advantage of the pair-copula construction is that we can choose the number of parameters that we want to use

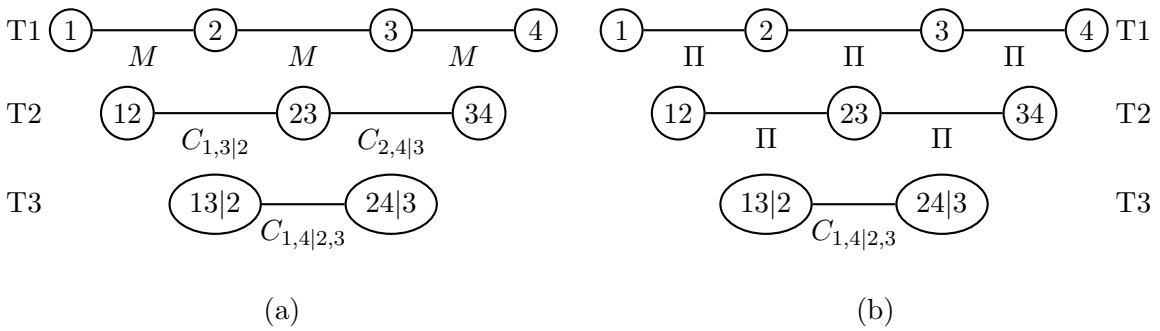


Figure 4.1: Two D-vine structures in four dimensions, where all bivariate copulas in the first tree are the Fréchet upper bound M (a), and all copulas in the first and second tree are bivariate independence copulas Π (b).

for an application, since it is possible to set specific building blocks to the independence copula Π, and therefore, reduce the number of parameters. For instance, if we set all bivariate copulas from the second to the last tree to the independence copula, we get a Markov tree model, which is build from $d-1$ bivariate parametric copulas, see, e.g., Bedford and Cooke (2002). But note that Markov tree models are often too restrictive and have too few parameters. The fully specified pair-copula construction, on the other hand, might include unnecessary parametric copula families. Thus, we need another step in the model selection procedure to decide whether we need to specify a certain parametric bivariate copula in the pair-copula construction, or whether we can simply replace it with the independence copula. Furthermore, we need to set the structure of the vine in a way such that many bivariate building blocks can be replaced by the independence copula. In particular, we want to replace the bivariate copulas in the higher trees of the pair-copula construction by the independence copula. Therefore, we try to capture as much information on the dependence structure as possible in the lower trees of the vine. This consideration is one of the reasons why we use maximum spanning trees in the model selection procedure to specify the structure of the trees in the vine, as discussed in Section 4.1.1.

In the next step, we discuss how to decide whether we can set specific bivariate copulas to the independence copula. One way to avoid redundant parametric copula families in the pair-copula construction is already suggested in Dißmann *et al.* (2011). They conduct a test of independence for all conditional pseudo-observations that are connected by an edge in the vine structure. That is, before selecting a parametric family for the bivariate building blocks, they check if they can reject the null hypothesis of independence for this conditional dependence structure. Only if they can reject this hypothesis, they continue and select a parametric copula family for this bivariate dependence structure. This is the first heuristic to reduce the number of parameters in the vine. However, this parameter reduction strategy does not account for an important characteristic of the pair-copula construction. That is, the influence of a bivariate copula in the pair-copula construction depends on the pair-copulas in the lower trees. This is illustrated in the next example.

Example 4.1 *In this example, we use the 4-dimensional D-vine structure of Figure 4.1 to illustrate the influence of the bivariate building blocks in the higher trees of the pair-copula construction.*

- *In Figure 4.1a, the bivariate copulas in the first tree are all set to the Fréchet upper bound. That is, we have a deterministic relation between the first and the second, the second and the third, and between the third and the fourth variable. Therefore, the resulting 4-dimensional pair-copula construction is the 4-dimensional Fréchet upper bound. This is true whatever bivariate copula we use in the second and the third tree. Thus, the copulas $C_{1,3|2}$, $C_{2,4|3}$, and $C_{1,4|2,3}$ have no influence at all.*
- *In Figure 4.1b, all bivariate copulas in the first and second tree are set to the independence copula. Therefore, the influence of the copula $C_{1,4|2,3}$ is not affected by the building blocks in the lower trees. That is, the unconditional dependence between the first and the fourth variable is completely modeled by the copula $C_{1,4|2,3}$.*

The pair-copula constructions in Figure 4.1 illustrate two extreme cases where the copulas in the higher trees have no influence, or the influence is completely passed through to the unconditional variables. Choosing less extreme copulas in the lower trees will, of course, lead to less extreme results. However, independence tests on the conditional pseudo-observations cannot decide whether the bivariate copula in the pair-copula construction have a relevant influence on the unconditional variables. This is further illustrated in the next example.

Example 4.2 *We analyze a 4-dimensional pair-copula with a D-vine structure, as in Figure 4.1. The bivariate building blocks are all set to Gauss copulas. The parameters of the copulas in the first tree are $\rho_{1,2} = 0.9$, $\rho_{2,3} = -0.7326$, and $\rho_{3,4} = -0.9$. In the second tree, we set the parameters of $C_{1,3|2}$ and $C_{2,4|3}$ to $\rho_{1,3|2} = 0.9$ and $\rho_{2,4|3} = -0.9$, respectively. Finally, we use different values for the parameter of the copula $C_{1,4|2,3}$ to illustrate the influence of this copula on the 4-dimensional dependence structure. We use $\rho_{1,4|2,3} \in \{-0.5, 0, 0.5\}$ to model the conditional dependence of the first and the fourth variable, given the second and the third.*

Using only bivariate Gauss copulas in the pair-copula construction results in a 4-dimensional Gauss copula. This copula is completely specified by the covariance matrix Σ. Now, we can illustrate the effect of the parameter value $\rho_{1,4|2,3}$ on the 4-dimensional dependence structure by analyzing the covariance matrix.

$$\Sigma = \begin{pmatrix} 1 & 0.9 & -0.3923 & -0.0180/0/0.0180 \\ 0.9 & 1 & -0.7326 & 0.3923 \\ -0.3923 & -0.7326 & 1 & -0.9 \\ -0.0180/0/0.0180 & 0.3923 & -0.9 & 1 \end{pmatrix}$$

We see that large variations in the parameter $\rho_{1,4|2,3}$ have only a very limited influence on the complete dependence structure. Furthermore, the influence of $C_{1,4|2,3}$ is limited to the unconditional dependence between the first and the fourth parameter.

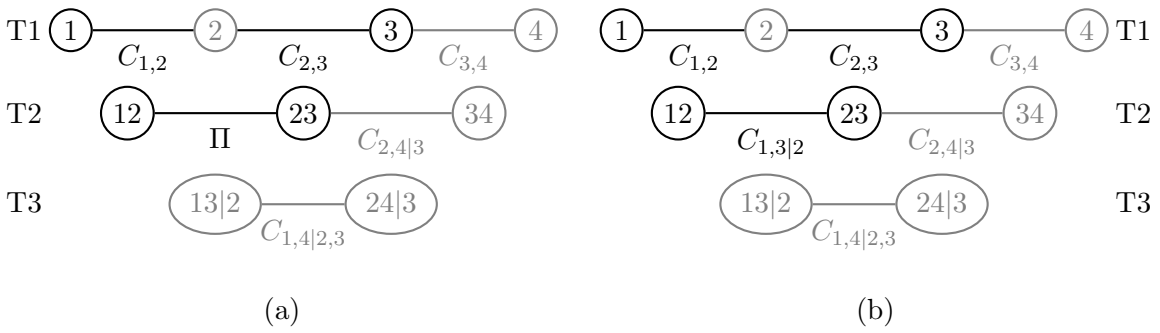


Figure 4.2: In these 4-dimensional vine structures, we focus only on the dependence structure of the bivariate margin between the first and the third variable. This margin is completely specified by $C_{1,2}$ and $C_{2,3}$ in (a) and by $C_{1,2}$, $C_{2,3}$, and $C_{1,3|2}$ in (b).

In general, every bivariate copula in the current tree of the sequential model selection procedure affects only one bivariate margin. Therefore, we propose to conduct an independence test on the untransformed variables within the sequential model selection procedure of Section 4.1.1. These bivariate independence tests are based on the Cramér-von Mises test statistic, as discussed in Section 2.5. Thus, our parameter reduction technique is similar to the one in Dißmann *et al.* (2011). However, we work with the untransformed variables.

Here, we illustrate this procedure in four dimensions. Suppose that we have already specified the complete vine structure of a 4-dimensional pair-copula construction as in Figure 4.1. In order to avoid redundant building blocks in the first tree, we test the bivariate margins $\{1,2\}$, $\{2,3\}$ and $\{3,4\}$ for independence. In the next step, we choose a parametric family for the copulas $C_{1,2}$, $C_{2,3}$, and $C_{3,4}$ whenever we reject the null hypothesis of independence for these margins. Otherwise, we set these bivariate copulas to the independence copula. We continue in the second tree by analyzing if the bivariate dependence structure $\{1,3\}$ is already adequately modeled by the copulas $C_{1,2}$ and $C_{2,3}$. Dißmann *et al.* (2011) suggest to transform the pseudo-observations and test if $u_{1|2} = F(u_1|u_2)$ and $u_{3|2} = F(u_3|u_2)$ are independent. Unfortunately, this procedure neglects the fact that, as seen in Example 4.2, the influence of the bivariate copula in a vine structure strongly depends on the position in the vine and on the bivariate copulas in the lower trees. Therefore, we propose to test if the dependence between the untransformed pseudo-observations u_1 and u_3 is already adequately modeled. This is visualized in Figure 4.2. We use the goodness-of-fit test in Appendix 2.A and compare the distance between the empirical bivariate copula of u_1 and u_3 with the corresponding bivariate marginal copula in the pair-copula model. Note that this bivariate marginal copula between the first and the third variable is completely specified by $C_{1,2}$ and $C_{2,3}$, which are fitted in the first tree, and $C_{1,3|2}$, which is set to the independence copula for this test, see Figure 4.2a. Only if we can reject this test, we precede as in Figure 4.2b and select the most appropriate bivariate copula for $C_{1,3|2}$.

We proceed with the next bivariate building block in the second tree. In order to check if we can set $C_{2,4|3}$ to the independence copula, we test if the dependence structure between the second and fourth variable is already appropriately modeled by $C_{2,3}$ and $C_{3,4}$. Finally, we conduct the goodness-of-fit test on the pseudo-observations u_1 and u_4 to see if this bivariate margin is already implicitly well-fitted by the copulas $C_{1,2}, C_{2,3}, C_{3,4}, C_{1,3|2}, C_{2,4|3}$ of the lower trees. If not, we need to select a parametric copula family for $C_{1,4|2,3}$. Note that we choose the goodness-of-fit test in Appendix 2.A, since it is not practical to evaluate all marginal copulas or their densities directly in the pair-copula construction. However, it is fast to simulate from a pair-copula, and therefore, it is computationally feasible to conduct two-level goodness-of-fit tests.

Alternative procedures for parameter reduction are the truncation and simplification approach, discussed in Brechmann *et al.* (2012) and Heinen and Valdesogo (2009). There, one sequentially specifies the vine structure up to a certain level and replaces all copulas in higher trees by independence copulas (truncation) or by bivariate Gauss copulas (simplification). These procedures are particularly appropriate in situations where rough parameter reduction procedures are needed. Furthermore, these procedures can be used for very high-dimensional problems, since they stay computationally feasible. On the contrary, our parameter reduction approach, based on the two-level parametric goodness-of-fit test, is an in-depth approach that checks every bivariate copula in the pair-copula construction individually. Note that this accuracy comes with the price of a higher computational effort. However, for less then ten dimensions, our approach is still feasible on a personal computer, and in a higher-dimensional setting, we can apply parallel computing techniques. Therefore, our approach is a valid alternative to the rough truncation or simplification procedures and helps to prevent the specification of redundant building blocks in the pair-copula construction.

4.3 Empirical Example

In this empirical study, we evaluate the performance of different dependence models for a data set of seven major German stocks in a copula-GARCH setting. That is, we specify the marginal GARCH models for each univariate time series separately, and in the next step, we fit the competing copula models to the residuals. This procedure is well-established in recent articles on high-dimensional dependence modeling for financial data. See, e.g., Aas *et al.* (2009) and Kurowicka and Joe (2011). In order to evaluate the different high-dimensional copula models, we use rolling window estimators to generate out-of-sample value at risk (VaR) forecasts from the different models for an equally weighted portfolio. These out-of-sample forecasts vary only in the choice of the underlying dependence model. Finally, we compare the VaR forecasts with the observed values of the equally weighted portfolio. In contrast to in-sample evaluation techniques, this procedure allows us to check these highly complex models for overfitting. Furthermore, the results from the out-of-sample comparison are very important and easy to interpret in a risk management framework. We use the fGarch package by Wuertz and Chalabi

(2009) within the statistical software package R 2.12.0 (R Development Core Team, 2010) to estimate and evaluate the univariate GARCH model of Bollerslev (1986). In particular, we use the GARCH(1,1) model

$$y_t = \mu + \varepsilon_t, \tag{4.1}$$
$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{4.2}$$

where ε_t is skewed t-distributed with $E[\varepsilon_t] = 0$ and $Var[\varepsilon_t] = \sigma_t^2$. In contrast to the original model, we allow for non-normally distributed error terms, as already suggested in Bollerslev (1987). The estimation and simulation algorithms for the different copula models have been implemented in MATLAB R2010a.

Our data set consists of daily log returns of Allianz, Deutsche Bank, E.ON, Munich RE, RWE, Siemens, and ThyssenKrupp from January 2003 to December 2012. The stock price time series are available at http://finance.yahoo.com. In a preliminary analysis we fit different GARCH models to each of the univariate log return time series on the full time range. Overall, the GARCH(1,1) model with a skewed t- error distribution shows a satisfactory fit for all univariate time series and varying time periods. The Ljung-Box tests in Table 4.7 indicate that we cannot reject the null hypothesis that there is no autocorrelation left in the residuals and squared residuals for nearly all time series and different lag sizes. A visual analysis of the QQ-plots shows that the skewed t-distribution provides a good fit for the distribution of the residuals. Table 4.6 gives an overview on the estimated parameters on the full time range. Figure 4.3 shows the log return time series of an equally weighted portfolio. In our data set we have 2441 daily log returns, and we use rolling window estimators with a lag size of 250 observations. Thus, we generate $\text{VaR}_{0.99}$ forecasts of an equally weighted portfolio on $t = 251, \ldots, 2441$. In this study we use the Gauss copula (Gauss), t copula (t), pair-copula construction (PCC), and the parameter reduced pair-copula construction (PRPCC) as competing dependence models. We denote the out-of-sample forecasts by $(\text{VaR}_{0.99,t}^{\text{Gauss}})$, $(\text{VaR}_{0.99,t}^{\text{t}})$, $(\text{VaR}_{0.99,t}^{\text{PCC}})$, and $(\text{VaR}_{0.99,t}^{\text{PRPCC}})$ and generate these forecasts on $t = [251, \ldots, 2441]$ with the following procedure.

1. Univariate rolling window estimation:
   a) Estimate the GARCH(1,1)-Skewed-t parameters on $[t - 250, t - 1]$ for all 7 time series separately.
   b) Transform the standardized residuals with the empirical distribution function to $u_{i,j}^n$, $i = t - 250, \ldots, t - 1$, $j = 1, \ldots, 7$, as in Equation (2.17).
2. Fit the dependence models to the pseudo-observations $u_{i,j}^n$.
3. Simulate $u_{i,j}^{\text{fc}}$, $i = 1, \ldots, 10\,000$, $j = 1, \ldots, 7$ from the estimated copulas.
4. Approximation of the univariate log return distributions for time t and $j = 1, \ldots, 7$:
   a) For $i = 1, \ldots, 10\,000$, transform the $u_{i,j}^{\text{fc}}$ with the quantile function of the skewed t-distribution with the estimated skewness and shape parameter to $x_{i,j}^{\text{fc}}$. (mean = 0, variance = 1)

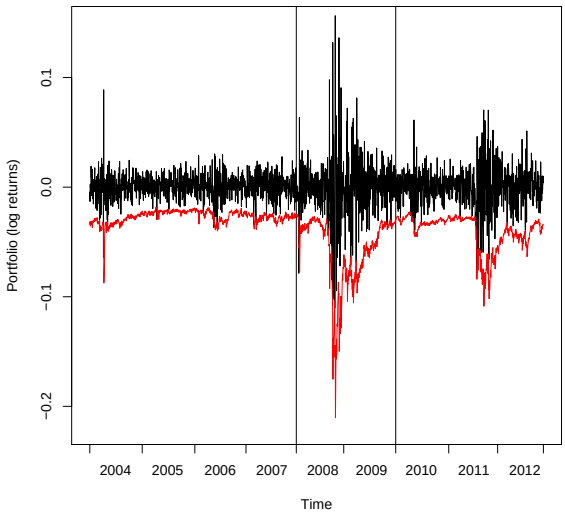


Figure 4.3: This plot shows the log returns of an equally weighted portfolio of seven major German stocks. The red line gives the one day $\text{VaR}_{0.99}$ forecasts from the parameter reduced PCC model.

b) Forecast the variance σ_j^{fc} with Equation (4.2) and the estimated parameters.

c) For $i = 1, \ldots, 10\,000$, forecast $y_{i,j}^{\text{fc}} = \hat{\mu}_j + \sigma_j^{\text{fc}} x_{i,j}^{\text{fc}}$.

5. We get 10 000 dependent forecasts for each time series. Generate 10 000 possible log returns for the equally weighted portfolio by

$$y_i^{\text{pfc}} = \log\left(\sum_{j=1}^{7} \frac{\exp\left(y_{i,j}^{\text{fc}}\right)}{7}\right).$$

6. Use a nonparametric estimator to generate a $\text{VaR}_{0.99,t}$ forecast for the portfolio on day t.

In step 2 of the rolling window procedure, we fit the four different dependence models to the residuals. Due to the computational complexity, we cannot conduct the parameter reduction procedure in the PRPCC model in every iteration. Therefore, we reduce the parameters only every 50 steps and reuse this parameter reduced model for the subsequent forecasts. In order to evaluate the different dependence models, we compare the exceedances of the portfolio log return time series with the VaR forecasts. If the selected model provides a good fit for the data, we expect that only one percent of the log returns are below the $\text{VaR}_{0.99}$ forecasts. Furthermore, we do not expect these exceedances to appear in clusters. To check the fit of the different models, we apply the test by Christoffersen (1998), see Appendix 4.A, that accounts for these two features.

We also evaluate the influence of the different dependence models on the portfolio VaR

Time Period	Gauss (21)	t (22)	PCC (30)	PRPCC (ø21.2)
2004–2012	**0.031**	0.19	0.42	0.19
2004–2007	0.26	0.23	0.18	0.26
2008–2012	**0.0052**	0.072	0.30	0.12
2008–2009	**0.040**	0.19	0.36	0.092
2010–2012	0.11	0.34	0.71	0.71

Table 4.1: P-values of the conditional coverage test for the different dependence models and different time periods. We provide the number of parameters of the parametric copula model within brackets. P-values smaller than 5% are bold.

forecasts for different time periods, see Figure 4.3. The findings of this model evaluation are summarized in Table 4.1. More detailed results are given in Appendix 4.B. Table 4.1 gives the p-values of conditional coverage test, see Appendix 4.A, and the first row in the table indicates that we can reject the Gaussian dependence model on the full time range at a 5% level. None of the other dependence models can be rejected. In the subsequent rows of Table 4.1, we list the p-values of this test for different time periods. The poor performance of the Gaussian copula on the whole time range is mainly due to the bad fit of the dependence structure during the financial crisis $2008 - 2009$. Table 4.2 indicates that the portfolio VaR forecasts in this period are too optimistic. This is a severe problem since in times of crises the Gaussian dependence model is not cautious enough. Comparing the expected and the observed exceedances under the VaR forecasts in Tables 4.3-4.5, the pair-copula construction gives slightly better forecasts than the t copula and the parameter reduced pair-copula construction on this set of data. This slight improvement on forecast quality comes along with the cost of additional parameters. The t copula dependence model and the parameter reduced pair-copula construction are comparable in the number of parameters and the quality of VaR forecasts. Both cannot be rejected at any time period under consideration, and they clearly outperform the Gaussian dependence model. This example illustrates that it is possible to decisively reduce the number of parameters in the pair-copula construction from 30 to an average of 21.2 by the method presented in Section 4.2 without severe negative effects on the forecast quality. Still, the parameter reduced model is a pair-copula construction and this is advantageous compared to the t copula, e.g., in economically interpreting the dependence structure, and building time varying dependence models as in Almeida *et al.* (2012).

4.4 Conclusion

The pair-copula construction framework is a flexible and tractable concept to build parametric copula models in arbitrary dimensions. Throughout the last two chapters, we have seen how this method transfers the flexibility of bivariate copula families to

higher dimensions. Here, we discuss model selection techniques in detail. Model selection heuristics are particularly important in this concept since there are too many possible ways to create a valid pair-copula. One drawback of the pair-copula construction concept is that the number of bivariate parametric copulas increases quadratically with the dimension. Furthermore, it is not obvious that we actually need all of these parametric copulas in the model. To overcome this problem, we present a parameter reduction technique that evaluates the impact of an additional copula in the multivariate dependence structure. This heuristic helps to distinguish between necessary and redundant copulas in the model. We apply this parameter reduction technique in an empirical study and observe that it is possible to reduce the number of parameters for financial return data seriously without a severe impact on the forecast quality.

4.A Conditional Coverage Test

Here, we recall a test proposed by Christoffersen (1998) to evaluate the appropriateness of a fitted model. That is, we compare a given time series $(y_t)_{t\in\mathbb{T}}$ with out-of-sample VaR_p forecasts $(\mathrm{VaR}_{p,t})_{t\in\mathbb{T}}$. Therefore, we define an auxiliary time series $(I_t)_{t\in\mathbb{T}}$ for all $t \in \mathbb{T}$ by

$$I_t = \mathbb{1}(y_t \geq \mathrm{VaR}_{p,t}).$$

If the model provides appropriate out-of-sample forecasts, the sequence (I_t) is independent and Bernoulli(p) distributed. Christoffersen's test consists of two different parts. The unconditional coverage (UC) part tests the null hypothesis that the exceedance ratio of the y_t under the $\mathrm{VaR}_{p,t}$ forecasts is $1-p$. This results in the likelihood ratio test statistic

$$\mathrm{LR}_{\mathrm{UC}} = -2\log\left(\frac{(1-p)^{n_0}(p)^{n_1}}{(1-\hat{\pi})^{n_0}\hat{\pi}^{n_1}}\right), \tag{4.3}$$

where n_0 is the number of exceedances under the VaR_p, n_1 is the number of y_t that are larger than the corresponding $\mathrm{VaR}_{p,t}$ forecast, and $\hat{\pi} = \frac{n_1}{n_0+n_1}$.

The independence (Ind) part checks if the exceedances appear in clusters. Therefore, we test the null that $(I_t)_{t\in\mathbb{T}}$ is independent against a first-order Markov alternative. Let $n_{i,j}$ be the number of observations with value i followed by j, where $i, j \in \{0, 1\}$. The test statistic compares the transition matrix under the independence assumption

$$\Pi_{\mathrm{Ind}} = \begin{pmatrix} 1-\pi_2 & \pi_2 \\ 1-\pi_2 & \pi_2 \end{pmatrix}$$

with the general transition matrix

$$\Pi_{\mathrm{Gen}} = \begin{pmatrix} 1-\pi_{0,1} & \pi_{0,1} \\ 1-\pi_{1,1} & \pi_{1,1} \end{pmatrix}.$$

This results in the likelihood ratio test statistic

$$\mathrm{LR}_{\mathrm{Ind}} = -2\log\left(\frac{(1-\hat{\pi}_2)^{(n_{0,0}+n_{1,0})}\hat{\pi}_2^{(n_{0,1}+n_{1,1})}}{(1-\hat{\pi}_{0,1})^{n_{0,0}}\hat{\pi}_{0,1}^{n_{0,1}}(1-\hat{\pi}_{1,1})^{n_{1,0}}\hat{\pi}_{1,1}^{n_{1,1}}}\right), \tag{4.4}$$

where $\hat{\pi}_{0,1} = n_{0,1}/(n_{0,0}+n_{0,1})$, $\hat{\pi}_{1,1} = n_{1,1}/(n_{1,0}+n_{1,1})$, and $\hat{\pi}_2 = (n_{0,1}+n_{1,1})/(n_{0,0}+n_{0,1}+n_{1,0}+n_{1,1})$.

Finally, combining these two parts gives the conditional coverage (CC) test statistic

$$\mathrm{LR}_{\mathrm{CC}} = \mathrm{LR}_{\mathrm{UC}} + \mathrm{LR}_{\mathrm{Ind}}, \tag{4.5}$$

and Christoffersen (1998) shows that $\mathrm{LR}_{\mathrm{CC}}$ is asymptotically $\chi^2(2)$ distributed under the null hypothesis that the sequence (I_t) is independent and Bernoulli(p) distributed.

4.B Supplementary Results for the Empirical Example

Time Periode	EE	OE	$n_{0,0}$	$n_{0,1}$	$n_{1,0}$	$n_{1,1}$	UC	Ind	CC
2004–2012	21.91	35	1	34	34	2121	**0.0097**	0.59	**0.031**
2004–2007	9.97	11	1	10	10	975	0.75	0.11	0.26
2008–2012	11.94	24	0	24	24	1145	**0.0020**	0.32	**0.0052**
2008–2009	4.8	11	0	11	11	457	**0.015**	0.47	**0.040**
2010–2012	7.14	13	0	13	13	687	**0.048**	0.49	0.11

Table 4.2: Results for the Gauss copula model in the empirical example in Section 4.3. The expected exceedances (EE) and the observed exceedances (OE) under the $\text{VaR}_{0.99}$ forecasts are given in the first two columns. The definition of $n_{0,0}, n_{0,2}, n_{1,0}, n_{1,1}$ is given in Appendix 4.A. The last three columns provide p-values for the unconditional coverage (UC), independence (Ind), and conditional coverage (CC) tests.

Time Periode	EE	OE	$n_{0,0}$	$n_{0,1}$	$n_{1,0}$	$n_{1,1}$	UC	Ind	CC
2004–2012	21.91	30	1	29	29	2131	0.10	0.43	0.19
2004–2007	9.97	10	1	9	9	977	0.99	0.085	0.23
2008–2012	11.94	20	0	20	20	1153	**0.033**	0.41	0.072
2008–2009	4.8	9	0	9	9	461	0.086	0.56	0.19
2010–2012	7.14	11	0	11	11	691	0.18	0.56	0.34

Table 4.3: Results for the t copula model in the empirical example in Section 4.3. The expected exceedances (EE) and the observed exceedances (OE) under the $\text{VaR}_{0.99}$ forecasts are given in the first two columns. The definition of $n_{0,0}, n_{0,2}, n_{1,0}, n_{1,1}$ is given in Appendix 4.A. The last three columns provide p-values for the unconditional coverage (UC), independence (Ind), and conditional coverage (CC) tests.

Time Periode	EE	OE	$n_{0,0}$	$n_{0,1}$	$n_{1,0}$	$n_{1,1}$	UC	Ind	CC
2004–2012	21.91	26	1	25	25	2139	0.39	0.32	0.42
2004–2007	9.97	9	1	8	8	979	0.75	0.066	0.18
2008–2012	11.94	17	0	17	17	1159	0.17	0.48	0.30
2008–2009	4.8	8	0	8	8	463	0.18	0.60	0.36
2010–2012	7.14	9	0	9	9	695	0.50	0.63	0.71

Table 4.4: Results for the pair-copula construction model in the empirical example in Section 4.3. The expected exceedances (EE) and the observed exceedances (OE) under the $\text{VaR}_{0.99}$ forecasts are given in the first two columns. The definition of $n_{0,0}, n_{0,2}, n_{1,0}, n_{1,1}$ is given in Appendix 4.A. The last three columns provide p-values for the unconditional coverage (UC), independence (Ind), and conditional coverage (CC) tests.

Time Periode	EE	OE	$n_{0,0}$	$n_{0,1}$	$n_{1,0}$	$n_{1,1}$	UC	Ind	CC
2004–2012	21.91	30	1	29	29	2131	0.10	0.43	0.19
2004–2007	9.97	11	1	10	10	975	0.75	0.11	0.26
2008–2012	11.94	19	0	19	19	1155	0.059	0.43	0.12
2008–2009	4.8	10	0	10	10	459	**0.037**	0.51	0.092
2010–2012	7.14	9	0	9	9	695	0.50	0.63	0.71

Table 4.5: Results for the parameter reduced pair-copula construction model in the empirical example in Section 4.3. The expected exceedances (EE) and the observed exceedances (OE) under the $\text{VaR}_{0.99}$ forecasts are given in the first two columns. The definition of $n_{0,0}, n_{0,2}, n_{1,0}, n_{1,1}$ is given in Appendix 4.A. The last three columns provide p-values for the unconditional coverage (UC), independence (Ind), and conditional coverage (CC) tests.

Stock	μ	ω	α_1	β_1	skew	shape
Allianz	$7.05e-04$	$3.86e-06$	$7.97e-02$	0.91	$-5.03e-02$	2.81
Deutsche Bank	$4.49e-04$	$2.72e-06$	$7.68e-02$	0.92	$1.49e-02$	2.68
E.ON	$7.79e-04$	$1.03e-05$	$1.18e-01$	0.85	1.02	5.22
Munich RE	$4.85e-04$	$3.28e-06$	$7.86e-02$	0.91	$-1.55e-03$	1.93
RWE	$4.71e-04$	$6.11e-06$	$8.37e-02$	0.89	$-3.57e-02$	2.09
Siemens	$6.79e-04$	$2.12e-06$	$5.32e-02$	0.94	1.02	6.60
ThyssenKrupp	$8.96e-04$	$7.15e-06$	$1.11e-01$	0.88	$9.69e-01$	6.07

Table 4.6: Univariate parameter estimates for the GARCH(1,1)-Skewed-t model on the whole time range.

Stock	$Q_R(10)$	$Q_R(15)$	$Q_R(20)$	$Q_{R^2}(10)$	$Q_{R^2}(15)$	$Q_{R^2}(20)$
Allianz	0.056	**0.029**	0.082	0.091	0.16	0.30
Deutsche Bank	0.13	0.16	0.11	0.66	0.73	0.63
E.ON	0.90	0.88	0.79	1	1	1
Munich RE	0.11	0.31	0.50	0.91	0.90	0.72
RWE	0.24	0.38	0.43	0.19	0.55	0.78
Siemens	0.55	0.37	0.51	0.76	0.95	0.98
ThyssenKrupp	0.98	0.98	0.83	1	1	1

Table 4.7: P-values of the Ljung-Box tests for different lag sizes for the residuals (R) and the squared residuals (R^2) of the univariate GARCH(1,1)-Skewed-t models.

Chapter 5

Dependence Modeling for Lévy Processes

The previous chapters contain a discussion on dependence modeling for random variables. In particular, the pair-copula construction has proven to be a flexible and applicable method to deal with high-dimensional problems. In practice, however, it is rarely sufficient to model the dependence between the dimensions for a given time t only, because many observations also show a time structure. Therefore, it is important to find models that account for the dependence between the different dimensions and include the dependence in time as well. Up to now, we removed the time series facet of the data in a univariate preprocessing step. This procedure is well-known for time series data on an equidistant time grid as in the empirical example in Section 4.3. However, the increasing availability of high frequency and non-equidistantly spaced data demands for more general methods. Therefore, we introduce a continuous-time approach to cope with these current problems. In this chapter, we focus on Lévy processes that constitute an important class of continuous-time stochastic processes. These Lévy processes are of special importance due to several reasons. Firstly, they are used for modeling purposes in many applications directly. Secondly, Lévy processes are building blocks in most continuous-time models, and thirdly, models based on Lévy processes can be applied to equidistant and non-equidistant spaced data.

In the first section of this chapter, we discuss Lévy processes in d dimensions. The second section introduces a concept to separate the dependence structure and the margins of the Lévy process.

5.1 Lévy Processes

Lévy processes are continuous-time stochastic processes with the main feature that their increments are independent and identically distributed. They have been extensively studied in Sato (1999) or Kallenberg (2002) and references therein. The application of Lévy processes in finance is discussed in Cont and Tankov (2004). The class of Lévy processes includes the Brownian motion, the compound Poisson process and also processes with infinitely many jumps on any bounded interval. Especially the jumps

of the Lévy processes make them relevant for applications in any area where normality assumptions are violated. The well-known results on Lévy processes, which we state in the following, make this class flexible and tractable. Lévy processes can be used in many different applications. The arrival process of insurance claims or losses in the operational risk context, for instance, are often supposed to be modeled adequately with a pure jump Lévy process like the compound Poisson process. In finance, stochastic processes defined by an Itô-integral are usually driven by a Brownian motion. Moreover, the important class of Ornstein-Uhlenbeck processes are built upon so-called background driving Lévy processes.

Firstly, we fix some notation which will be used throughout the following chapters. The well-known definitions, theorems, propositions, and remarks are close to Sato (1999), where we adjust the notation for consistency.

5.1.1 Definition and Properties

Stochastic processes are intended to describe random objects throughout time. It is possible to think of a stochastic processes as a special random variable, where the realization of this random variable is a function of time. Here, we focus on continuous-time processes and follow Sato (1999, Definition 1.4) in defining stochastic processes.

Definition 5.1 *A family* $\{X_t : t \geq 0\}$ *of random variables on* $\mathbb{R}^d$ *with parameter* $t \in [0, \infty)$ *defined on a common probability space* $(\Omega, \mathcal{F}, P)$ *is called a stochastic process. It is written as* $(X_t)_{t\in\mathbb{R}_+}$. *A stochastic process* $(Y_t)_{t\in\mathbb{R}_+}$ *is called a modification of a stochastic process* $(X_t)_{t\in\mathbb{R}_+}$, *if*

$$P(X_t = Y_t) = 1 \quad \textit{for } t \in [0, \infty).$$

For any fixed $0 \leq t_1 < t_2 < \ldots < t_n$,

$$P(X(t_1) \in B_1, \ldots, X(t_n) \in B_n)$$

determines a probability measure on the Borel σ*-algebra* $\mathcal{B}((\mathbb{R}^d)^n)$. *The family of probability measures over all possible choices of* n *and* $t_1, \ldots, t_n$ *is called the system of finite-dimensional distributions. Two stochastic processes* $(X_t)_{t\in\mathbb{R}_+}$ *and* $(Y_t)_{t\in\mathbb{R}_+}$ *are identical in law, written as*

$$(X_t)_{t\in\mathbb{R}_+} \stackrel{\mathscr{D}}{=} (Y_t)_{t\in\mathbb{R}_+},$$

if the systems of their finite-dimensional distributions are identical. For any fixed value of ω, X_t *as a function of* t *is called a sample path.*

Lévy processes are stochastic processes with independent increments. We follow Sato (1999, Definition 1.6) and define Lévy processes in this thesis in the following way.

Definition 5.2 (Lévy Process) *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space. Then an* $\mathbb{R}^d$*-valued stochastic process* $(L_t)_{t\in\mathbb{R}_+}$ *is called a Lévy process, if*

(i) $L_0 = 0$ *a.s.*

(ii) *For any choice* $n \in \mathbb{N}$ *and* $0 \leq t_0 < t_1 < \cdots < t_n$ *the random variables*

$$L_{t_0}, L_{t_1} - L_{t_0}, \ldots, L_{t_n} - L_{t_{n-1}}$$

are independent (independent increments).

(iii) $L_{t+h} - L_t \stackrel{\mathcal{D}}{=} L_{s+h} - L_s \quad \forall h, s, t \in \mathbb{R}_+$ *(stationary increments).*

(iv) *The process* L *is continuous in probability, i.e.,*

$$L_t - L_s \xrightarrow{P} 0 \quad \textit{for} \quad t \to s \quad \forall s \in \mathbb{R}_+$$

(stochastically continuous).

(v) *There is an* $\Omega_0 \in \mathcal{F}$ *with* $P(\Omega_0) = 1$ *such that, for every* $\omega \in \Omega_0$, $L_t(\omega)$ *is right-continuous in* $t \geq 0$ *and has left limits in* $t > 0$ *(càdlàg path).*

It is possible to define Lévy processes without explicitly stating condition (v), since from Sato (1999, Theorem 11.5) follows that every process that satisfies the conditions (i)-(iv) has a modification such that (i)-(v) holds. Therefore, we just use condition (v) for convenience.

Some of the best known stochastic processes satisfy the Lévy properties. We recall some processes in the following example. In order to display the path behavior of these Lévy processes, we give one realization of such Lévy Processes on the interval $[0, 1]$ in Figure 5.1.

Example 5.3 *(i) Standard Brownian motion: A stochastic process* $(W_t)_{t \in \mathbb{R}_+}$ *on* $\mathbb{R}^d$ *is a standard Brownian motion if it is a Lévy process and if, for any* $t > 0$, *the random variable* W_t *has a Gaussian distribution with mean* 0 *and covariance matrix* tI_d, *where* I_d *is the identity matrix.*

(ii) Poisson process: Let $(\tau_i)_{i \in \mathbb{N}}$ *be a sequence of independent, exponential random variables with parameter* λ *and* $T_n = \sum_{i=1}^{n} \tau_i$. *The process* $(N_t)_{t \in \mathbb{R}_+}$ *defined by*

$$N_t = \sum_{n \geq 1} \mathbb{1}(t \geq T_n)$$

is called a Poisson process with intensity λ. *At any time* $t \geq 0$, *the random variable* N_t *follows a Poisson distribution with parameter* λt.

(iii) Compound Poisson process : Let $(N_t)_{t \in \mathbb{R}_+}$ *be a Poisson process with intensity* λ. *Let* $(Y_i)_{i \in \mathbb{N}}$ *be a sequence of independent and identically distributed random variables that are independent from* (N_t). *Then*

$$X_t = \sum_{i=1}^{N_t} Y_i$$

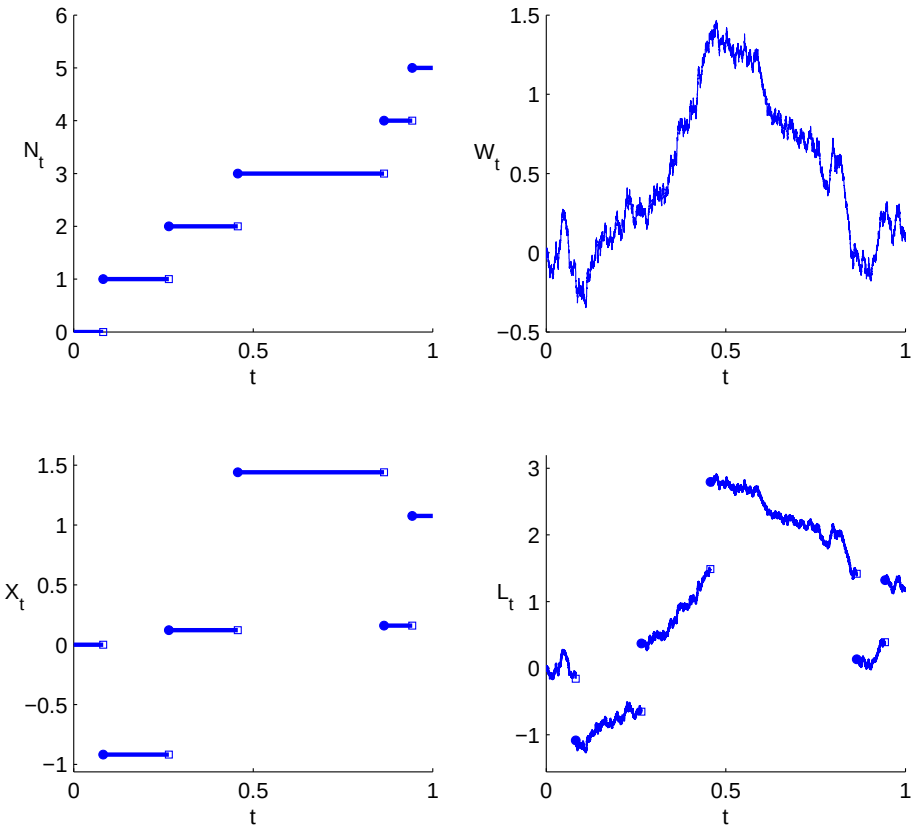


Figure 5.1: Sample path of a Poisson process N_t (upper left), a Brownian motion W_t (upper right), a compound Poisson process X_t (lower left), and a sum of a Brownian motion with a compound Poisson process L_t(lower right).

is called a compound Poisson process.

(iv) Sums of Lévy processes: An immediate consequence of Definition 5.2 is that the sum of two Lévy processes is again a Lévy process.

For an in-depth discussion of these processes we refer to Sato (1999) and Cont and Tankov (2004).

Due to the special properties of Lévy processes, it is possible to characterize each Lévy process $(L_t)_{t\in\mathbb{R}_+}$ completely by its distribution at any point t up to identity in law. We will see that these distributions are all infinitely divisible. The definition of infinitely divisible distributions can be found, for example, in Sato (1999, Definition 7.1).

Definition 5.4 *A probability measure ξ on $\mathbb{R}^d$ is infinitely divisible if, for any positive integer n, there is a probability measure ξ_n on $\mathbb{R}^d$ such that*

$$\xi = \underbrace{\xi_n * \ldots * \xi_n}_{n-times}.$$

That is ξ is the n-fold convolution of ξ_n.

Now we are able to state the connection between Lévy processes and infinitely divisible distributions in a more formal way. Therefore, we recall a special case of Sato (1999, Theorem 9.1, Corollary 11.6).

Proposition 5.5 *If $(L_t)_{t\in\mathbb{R}_+}$ is a Lévy process on $\mathbb{R}^d$, then, for every t, the distribution of L_t is infinitely divisible. On the other hand for every infinitely divisible probability measure ξ on $\mathbb{R}^d$, there is a Lévy process $(L_t)_{t\in\mathbb{R}_+}$ such that L_1 is distributed like ξ and $(L_t)_{t\in\mathbb{R}_+}$ is unique up to identity in law.*

Since there is this one-to-one correspondence between Lévy processes and infinitely divisible distributions, any representation of infinitely divisible distributions can be used to characterize Lévy processes. This representation is given in the famous Lévy-Khintchine formula that is stated, for example, in Sato (1999, Theorem 8.1).

Theorem 5.6 (Lévy-Khintchine representation) *(i) Let ξ be an infinitely divisible distribution on $\mathbb{R}^d$. Then there exists $\gamma \in \mathbb{R}^d$, a symmetric non-negative definite matrix $\Sigma \in S_d^+$ and a measure ν on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}^d}(|x|^2 \wedge 1)\nu(dx) < \infty$, such that the characteristic function of ξ*

$$\begin{aligned}\varphi_\xi(z) &= \exp\big\{i\langle\gamma, z\rangle - \tfrac{1}{2}\langle z, \Sigma z\rangle \\ &\quad + \textstyle\int_{\mathbb{R}^d}(e^{i\langle z,x\rangle} - 1 - i\langle z, x\rangle \mathbb{1}_{[0,1]}(|x|))\nu(dx)\big\},\end{aligned} \tag{5.1}$$

where $\langle ., . \rangle$ denotes the inner product.

(ii) The representation of φ_ξ by (γ, Σ, ν) is unique.

(iii) Let (γ, Σ, ν) be as in (i). Then there exists an infinitely divisible distribution ξ with φ_ξ given as in (i).

We will see later in this section that γ and Σ can be interpreted as the drift and the covariance matrix of a Brownian motion, whereas the Lévy measure ν controls the jumps of the Lévy process.

By looking at the distribution of the Lévy process $(L_t)_{t\in\mathbb{R}_+}$ only at time $t = 1$, which is an infinitely divisible distribution, it is possible to characterize the Lévy process with γ, Σ and ν used in Theorem 5.6. This is a remarkable result, since we have seen that the class of Lévy processes is a rich class with a great diversity of possible processes, and we only assume that this processes satisfy the conditions (i)-(v) in Definition 5.2. In the following, we call (γ, Σ, ν) the characteristic triplet.

In this chapter, we want to introduce a general dependence modeling method for multivariate Lévy processes. Since we know from Proposition 5.5 that a Lévy process is uniquely defined by the distribution of L_t at any time t, in particular by L_1, we can simply use the copula theory to model the dependence of the random variable L_1 and thereby specify the dependence within the Lévy process completely. This is theoretically true but, as mentioned in Kallsen and Tankov (2006), this procedure is not applicable in practice due to several reasons. A d-dimensional probability distribution constructed of univariate, infinitely divisible distributions and a copula does not have to be infinitely divisible. It is furthermore unclear under which conditions, on the copula and the margins, such a construction leads to a multivariate, infinitely divisible distribution. Another drawback of specifying the distribution of L_1 is that it is not apparent how to simulate the Lévy process on the real line. It is only straightforward to simulate realizations of this Lévy process at times $t = 1, 2, 3, \ldots$. In the next section, we introduce a dependence modeling technique to resolve these shortcomings, but therefore, we need to discuss more characteristics of Lévy processes.

Up to now, we have focused on the properties of the Lévy process L_t at a given time t which is simply a random variable. Since Lévy processes are stochastic objects in time, it is also interesting to analyze the sample path $L(t, \omega)$ for a given ω as a function of time t. Therefore, we introduce the Poisson random measure as in Sato (1999, Definition 19.1).

Definition 5.7 *Let $(\Theta, \mathcal{B}, \rho)$ be a σ-finite measure space. A family of $\overline{\mathbb{N}}$-valued random variables $\{N(B) : B \in \mathcal{B}\}$ is called a Poisson random measure (PRM) on Θ with intensity measure ρ if the following hold:*

(i) for every ω, $N_\omega(.)$ is a measure on Θ,

(ii) for every B, $N(B)$ has Poisson distribution with mean $\rho(B)$,

(iii) if $B_1, \ldots, B_n$ are disjoint, then $N(B_1), \ldots, N(B_n)$ are independent.

We introduce the notion of Poisson random measures as a mathematical tool for describing the jump behavior of Lévy processes. The following lemma reveals how the jumps of

a Lévy process define a PRM, and therefore, the connection between these two concepts. It is important that $N_\omega(.)$ is a measure on Θ for every $\omega \in \Omega$. Thus, for every $\omega \in \Omega$, we can integrate with respect to this measure. In the next lemma, we establish the connection between the jumps of Lévy processes and Poisson random measures. This result is given in Sato (1999, Theorem 19.2)

Lemma 5.8 *Let $(L_t)_{t\in\mathbb{R}_+}$ be a Lévy process on $\mathbb{R}^d$ defined on the probability space $(\Omega, \mathcal{F}, P)$ with characteristic triplet (γ, Σ, ν) and define the measure $\tilde{\nu}$ on $H = (0,\infty) \times (\mathbb{R}^d)\backslash\{0\}$ by $\tilde{\nu}((0,t] \times B) = t\nu(B)$ for $B \in \mathcal{B}(\mathbb{R}^d)$. Using Ω_0 from Definition 5.2 of a Lévy process, define for $B \in \mathcal{B}(H)$,*

$$J(B,\omega) = \begin{cases} \#\{s : (s, L_s(\omega) - L_{s-}(\omega)) \in B\} & \text{for } \omega \in \Omega_0, \\ 0 & \text{else.} \end{cases} \tag{5.2}$$

Then $\{J(B) : B \in \mathcal{B}(H)\}$ is a Poisson random measure on H with intensity measure $\tilde{\nu}$.

Up to now the Lévy measure ν was a theoretical object arising from the Lévy-Khintchine formula which we needed to describe all possible Lévy processes. With the result from Lemma 5.8 we have an intuitive interpretation of the Lévy measure. Let $B \in \mathcal{B}(\mathbb{R}^d)$, then $\nu(B) = \tilde{\nu}((0,1] \times B)$ is the expected number of jumps of the Lévy process with values in B within a unit time interval. With the definition of $J(B,\omega)$ in Equation (5.2) we can now state the Lévy-Itô decomposition as in Sato (1999, Theorem 19.2).

Theorem 5.9 *Let $(L_t)_{t\in\mathbb{R}_+}$, $\tilde{\nu}$ and $J(B,\omega)$ be as in Lemma 5.8. Then the following hold.*

(i) There is an $\Omega_1 \in \mathcal{F}$ with $P(\Omega_1) = 1$ such that, for any $\omega \in \Omega_1$,

$$L_t^1(\omega) = \lim_{\epsilon\downarrow 0} \int_{(0,t]\times\{x:\varepsilon<|x|\le 1\}} \{xJ(d(s,x),\omega) - x\tilde{\nu}(d(s,x))\} \tag{5.3}$$

$$+ \int_{(0,t]\times\{x:|x|>1\}} xJ(d(s,x),\omega) \tag{5.4}$$

is defined for all $t \in [0,\infty)$ and the convergence is uniform in t on any bounded interval. The process $(L_t^1)_{t\in\mathbb{R}_+}$ is a Lévy process on $\mathbb{R}^d$ with characteristic triplet $(0,0,\nu)$.

(ii) Define

$$L_t^2(\omega) = L_t(\omega) - L_t^1(\omega) \quad \text{for } \omega \in \Omega_1.$$

There is an $\Omega_2 \in \mathcal{F}$ with $P(\Omega_2) = 1$ such that, for any $\omega \in \Omega_2$, $L_t^2(\omega)$ is continuous in t. The process $(L_t^2)_{t\in\mathbb{R}_+}$ is a Lévy process with characteristic triplet $(\gamma, \Sigma, 0)$.

(iii) The two processes $(L_t^1)_{t\in\mathbb{R}_+}$ and $(L_t^2)_{t\in\mathbb{R}_+}$ are independent.

This theorem states that we can decompose any Lévy process into a Jump process $(L_t^1)_{t\in\mathbb{R}_+}$ and a Brownian motion $(L_t^2)_{t\in\mathbb{R}_+}$, i.e., a Lévy process with characteristic triplet $(\gamma, \Sigma, 0)$. Again, this is a remarkable result since we only presume that the process satisfies the Conditions (i)-(v) in Definition 5.2. The second part of the jump process $(L_t^1)_{t\in\mathbb{R}_+}$, given in (5.4), is simply the sum of all big jumps. The first part of this process, given in (5.3), is a bit more difficult to interpret. There, we sum over all small jumps, but to ensure that the limit exists, we need to compensate this sum with the expected number of the small jumps. To avoid this compensation in Theorem 5.9, we can impose the stronger condition $\int_{|x|\leq 1} |x|\nu(dx) < \infty$ on the Lévy measure ν. That is, we state the Lévy-Itô decomposition for a subclass of Lévy processes, as in Sato (1999, Theorem 19.3). This condition guarantees that the jump part of the Lévy process has finite variation on $(0, t]$ for any $t \in (0, \infty)$, see Sato (1999, Theorem 21.9). This assumption is not very strict and we will see that it holds for most parametric families of Lévy processes that we use in the following.

Theorem 5.10 *Let $(L_t)_{t\in\mathbb{R}_+}$ be a Lévy process with characteristic triplet (γ, Σ, ν) as in Lemma 5.8. Let the Lévy measure ν satisfy $\int_{|x|\leq 1} |x|\nu(dx) < \infty$ and let $\tilde{\gamma}$ be the drift of $(L_t)_{t\in\mathbb{R}_+}$. Then the following hold.*

(i) There is $\Omega_3 \in \mathcal{F}$ with $P(\Omega_3) = 1$ such that, for any $\omega \in \Omega_3$,

$$L_t^3(\omega) = \int_{(0,t]\times\{x:|x|>0\}} xJ(d(s,x),\omega)$$

is defined for all $t \in [0, \infty)$. The process $(L_t^3)_{t\in\mathbb{R}_+}$ is a Lévy process on $\mathbb{R}^d$ with

$$\varphi_{L_1^3}(z) = \exp\left(\int_{\mathbb{R}^d} (e^{i\langle z,x\rangle} - 1)\nu(dx)\right).$$

(ii) Define

$$L_t^4(\omega) = L_t(\omega) - L_t^3(\omega) \quad \text{for } \omega \in \Omega_3.$$

Then, for any $\omega \in \Omega_2 \cap \Omega_3$, $L_t^4(\omega)$ is a continuous Lévy process with characteristic triplet $(\tilde{\gamma}, \Sigma, 0)$, thus a Brownian motion.

(iii) The two processes $(L_t^3)_{t\in\mathbb{R}_+}$ and $(L_t^4)_{t\in\mathbb{R}_+}$ are independent.

Comparing the jump parts L_t^1 in Theorem 5.9 and L_t^3 in Theorem 5.10, we see how the stronger assumption in the second theorem simplifies the decomposition clearly.

5.1.2 Lévy Subordinators

In this work, Lévy subordinators, which are increasing Lévy processes, play a crucial role. They are of special importance in many applications, e.g., in summing up losses, and we will see in the following chapters that this restriction facilitates the notation, used

for the dependence modeling, considerably. We follow Barndorff-Nielsen *et al.* (2001a, Section 3) and introduce Lévy subordinators by their path properties.

Definition 5.11 *A Lévy Subordinator is a Lévy process $(L_t)_{t\in\mathbb{R}_+}$ in $\mathbb{R}^d_+$ which is increasing in each coordinate.*

Using this special property simplifies the Lévy-Khintchine formula substantially, and facilitates the interpretation as well. The Lévy-Khintchine representation for subordinators is stated in Barndorff-Nielsen *et al.* (2001a, Proposition 3.1).

Proposition 5.12 *(i) Let $(L_t)_{t\in\mathbb{R}_+}$ be a d-dimensional Lévy subordinator. The characteristic function φ_{L_1} can be represented as*

$$\varphi_{L_1}(z) = \exp\left\{ i\langle\tilde{\gamma}, z\rangle + \int_{\mathbb{R}^d_+} (e^{i\langle z,x\rangle} - 1)\nu(dx) \right\}, \tag{5.5}$$

where $\tilde{\gamma} \in \mathbb{R}^d_+$ and ν is a σ-finite measure on $\mathbb{R}^d$, which is concentrated on $\mathbb{R}^d_+\backslash\{0\}$ and satisfies $\int_{|x|\leq 1} |x|\nu(dx) < \infty$.

(ii) Conversely, let $\tilde{\gamma} \in \mathbb{R}^d_+$ and ν be a σ-finite measure on $\mathbb{R}^d$ which is concentrated on $\mathbb{R}^d_+\backslash\{0\}$ and satisfies $\int_{\mathbb{R}^d_+} |x| \wedge 1\nu(dx) < \infty$. Then there exists a d-dimensional subordinator $(L_t)_{t\in\mathbb{R}_+}$ such that φ_{L_1} satisfies (5.5).

The first difference to Theorem 5.6 is the lack of the covariance matrix Σ in Equation (5.5). This is no surprise because we know from the Lévy-Itô decomposition that the covariance matrix Σ specifies the Brownian motion part of the Lévy process. Since the Brownian motion is in- and also decreasing, subordinators cannot have a Brownian part. The second difference lies in the integral part. In the Lévy-Khintchine representation for general Lévy processes, we need the truncation function $i\langle z, x\rangle \mathbb{1}_{[0,1]}(|x|))$ in order to guarantee that the integral exists. With the restriction to subordinators, we do not need this truncation function anymore. Since the Lévy measure for subordinators satisfies $\int_{\mathbb{R}^d_+} |x| \wedge 1\nu(dx) < \infty$, we can use the easier version of the Lévy-Itô decomposition, stated in Theorem 5.10. In conclusion, Lévy subordinators are pure jump processes with a non-negative drift. In the next example we define parametric families of univariate subordinators.

Example 5.13 *(i) Compound Poisson processes with positive jumps: Let $(X_t)_{t\in\mathbb{R}_+}$ be a compound Poisson process as defined in Example 5.3. If we only allow for positive jumps, (X_t) is a Lévy subordinator. These stochastic processes are of special importance in the field of loss modeling.*

(ii) Stable subordinator (Basawa and Brockwell, 1978): The Lévy measure is defined for any $0 < \alpha < 1$ and $\beta > 0$ by

$$\nu(B) = \int_{\mathbb{R}_+} \mathbb{1}_B(z) \frac{\alpha\beta}{z^{\alpha+1}} dz.$$

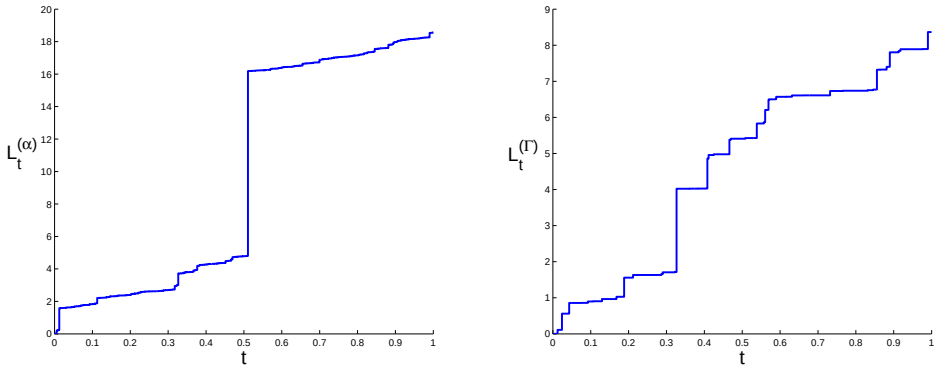


Figure 5.2: Sample path of a stable Subordinator $L_t^{(\alpha)}$ (left) and a Gamma process $L_t^{(\Gamma)}$ (right), as introduced in Example 5.13.

This Lévy process has heavy tails and is therefore suggested in many financial applications. One realization of this process is given in Figure 5.2.

(iii) Tempered stable subordinators (Tankov, 2005, p.115): This Lévy subordinator has three parameters α, c, λ, where $c, \lambda > 0$ and $0 \leq \alpha < 1$. The Lévy measure is given by

$$\nu(B) = \int_{\mathbb{R}_+} \mathbb{1}_B(z) \frac{ce^{-\lambda z}}{z^{\alpha+1}} dz.$$

The idea behind the tempering concept is to multiply the density of a Lévy measure with the term $ce^{-\lambda z}$, where $c, \lambda > 0$. This changes the intensity measure of the Poisson random measure in the way that the probability of large jumps is reduced. Since the term $ce^{-\lambda z}$ declines exponentially, this has decisive effect on Lévy processes with a high probability of large jumps, and therefore, this concept can be used to enrich the class of stable Lévy processes. The following two examples are important Lévy subordinators within the class of tempered stable processes.

(iv) Gamma process (Tankov, 2005, p.116): The gamma process is a tempered stable process with $\alpha = 0$. The Lévy measure is given by

$$\nu(B) = \int_{\mathbb{R}_+} \mathbb{1}_B(z) \frac{ce^{-\lambda z}}{z} dz,$$

where $c, \lambda > 0$. One realization of this process is given in Figure 5.2.

(v) Inverse Gaussian process (Tankov, 2005, p.116): The inverse Gaussian process is a tempered stable process with $\alpha = 1/2$.

$$\nu(B) = \int_{\mathbb{R}_+} \mathbb{1}_B(z) \frac{ce^{-\lambda z}}{z^{3/2}} dz,$$

where $c, \lambda > 0$.

5.2 Lévy Copulas

In this section, we discuss modeling the dependence structure between the dimensions of multivariate Lévy processes. Recall that we can decompose every Lévy process into a Brownian motion and a pure jump process. Since these two parts are independent, we are able to model each process separately. The dependence structure of the Brownian motion is completely specified by the covariance matrix Σ of the characteristic triplet. However, we are more interested in the jump part since the dependence of the jumps is crucial in many applications. Insurance companies, for instance, may model their losses for different risk types with univariate compound Poisson processes, which are basic Lévy processes. It is indisputably important to have adequate models for the margins, but the overall risk exposure of the insurance company and the extent to which diversification of risk is possible depends decisively on the interaction between the risks.

The fundamental work on dependence modeling for Lévy processes are the seminal publications of Tankov (2004), and Kallsen and Tankov (2006), where the concept of Lévy copulas is introduced. This concept transfers the idea of distributional copulas to the context of pure jump Lévy processes. As shown in Chapter 2, distributional copulas (normally just referred to as copulas) are functions that connect the marginal distribution functions of a random variable to the joint distribution. They contain the entire dependence information of the random variable. In the same sense, the theory of Lévy copulas enables us to model a multivariate Lévy process by its marginal Lévy processes and to choose a suitable Lévy copula for the dependence structure separately.

5.2.1 Definition and Properties

In this section, we illustrate how to decompose the Lévy measure of a Lévy subordinator into marginal parts and a dependence component, the Lévy copula. In this context, it is not advantageous to work with the Lévy measure directly. Therefore, we need to introduce the tail integral which captures all the information of the Lévy measure and is not defined on the σ-algebra $\mathcal{B}(\mathbb{R}_+^d)$, but on $\mathbb{R}_+^d$. The tail integral is related to the Lévy measure in a comparable way as the distribution function is related to probability measures. The tail integral is defined as follows (see, for example, Definition 3.1 in Esmaeili and Klüppelberg (2010)).

Definition 5.14 (Tail Integral) *Let ν be a Lévy measure on $\mathbb{R}_+^d$. The tail integral is a function $U : \overline{\mathbb{R}}_+^d \mapsto \overline{\mathbb{R}}_+$ defined by*

$$U(x_1, \ldots, x_d) = \begin{cases} \nu([x_1, \infty) \times \ldots \times [x_d, \infty)) & \text{if } (x_1, \ldots, x_d) \in \mathbb{R}_+^d \backslash \{0\}, \\ 0 & \text{if } x_i = \infty \text{ for at least one } i, \\ \infty & \text{if } (x_1, \ldots, x_d) = 0. \end{cases}$$

The tail integral U of a Lévy subordinator uniquely determines its Lévy measure ν. We define the marginal tail integrals U_k for any dimension $k = 1, \ldots, d$ of the multivariate Lévy subordinator in a similar way. For one-dimensional Lévy measures ν on $\mathbb{R}_+$, the tail integral $U(x) = \nu([x, \infty))$ is the expected number of jumps per unit of time with jump sizes larger or equal to x.

For most of the one-dimensional subordinators in Example 5.13, there are no explicit formulas for the tail integrals. One exception is the univariate stable subordinator. For this Lévy process, the tail integral can be explicitly calculated and inverted.

Example 5.15 *Let (L_t) be a stable subordinator as in Example 5.13. Then, for any $x > 0$, its tail integral is given by*

$$U(x) = \int_{[x,\infty)} \frac{\alpha\beta}{z^{\alpha+1}} dz = \beta x^{-\alpha}.$$

The inverse of the tail integral is needed for the simulation of the process and can also be calculated explicitly by

$$U^{-1}(u) = \left(\frac{u}{\beta}\right)^{-\frac{1}{\alpha}}$$

for any $u > 0$.

The tail integrals of the other subordinators may be calculated and inverted numerically. The dependence in the jumps of a multivariate Lévy process can be described by a Lévy copula which couples the marginal tail integrals to the joint one.

Definition 5.16 *A d-dimensional measure defining function $\mathfrak{C}(u_1, \ldots, u_d) : \overline{\mathbb{R}}_+^d \to \overline{\mathbb{R}}_+$, where the margins satisfy $\mathfrak{C}_k(u_k) := \mathfrak{C}(\infty, \ldots, \infty, u_k, \infty, \ldots, \infty) = u_k$ for all $u_k \in \overline{\mathbb{R}}_+$, and $k = 1, \ldots, d$, is called a Lévy copula.*

For a detailed introduction of measure defining functions, we refer to Kingman and Taylor (1966), and after Assumption 5.18, we discuss the connection between positive measures and Lévy copulas.

Now, we are able to state Sklar's theorem for Lévy copulas, as in Cont and Tankov (2004, Theorem 5.6).

Theorem 5.17 *Let U denote the tail integral of a d-dimensional Lévy subordinator whose components have the tail integrals $U_1, \ldots, U_d$. Then, there exists a Lévy copula $\mathfrak{C}$ such that for all $(x_1, \ldots, x_d) \in \overline{\mathbb{R}}_+^d$*

$$U(x_1, \ldots, x_d) = \mathfrak{C}(U_1(x_1), \ldots, U_d(x_d)). \tag{5.6}$$

Conversely, if $\mathfrak{C}$ is a Lévy copula and $U_1, \ldots, U_d$ are marginal tail integrals of a Lévy subordinator, Equation (5.6) defines the tail integral of a d-dimensional Lévy subordinator and $U_1, \ldots, U_d$ are the tail integrals of its components.

In this dissertation, we focus on Lévy copulas for which the following assumption holds.

Assumption 5.18 *Let $\mathfrak{C}$ be a Lévy copula on $\overline{\mathbb{R}}_+^d$ such that for every nonempty set $I \subset \{1, \ldots, d\}$*

$$\lim_{(u_i)_{i\in I}\to\infty} \mathfrak{C}(u_1, \ldots, u_d) = \mathfrak{C}(u_1, \ldots, u_d)|_{(u_i)_{i\in I}=\infty} \tag{5.7}$$

holds.

This is a rather weak assumption on the Lévy copula and is assumed in many papers, e.g., in Tankov (2005). It means that the Lévy copula has no new information at the points $u_i = \infty$ which is not already contained in the limit for $u_i \to \infty$. We need it since it ensures a bijection between a Lévy copula on $\overline{\mathbb{R}}_+^d$ and a positive measure μ on $\mathcal{B}(\mathbb{R}_+^d)$ with one-dimensional Lebesgue margins. This measure is given by

$$\mu((a, b]) = V_{\mathfrak{C}}([a, b]), \tag{5.8}$$

where $a, b \in \mathbb{R}_+^d$ with $a \leq b$, component-wise, and $V_{\mathfrak{C}}$ refers to the $\mathfrak{C}$-volume of the d-box $[a, b]$ which is defined as

$$V_{\mathfrak{C}}([a, b]) = \sum \operatorname{sgn}(c)\mathfrak{C}(c).$$

The sum is taken over all vertices c of $[a, b]$ and

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

Furthermore, any positive measure μ on $\mathbb{R}_+^d$ with Lebesgue margins uniquely defines a Lévy copula on $\overline{\mathbb{R}}_+^d$ that satisfies Assumption 5.18 by

$$\mathfrak{C}(u_1, \ldots, u_d) := \mu([0, u_1] \times \ldots \times [0, u_d])$$

and by specifying

$$\mathfrak{C}(u_1, \ldots, u_d)|_{(u_i)_{i\in I}=\infty} := \lim_{(u_i)_{i\in I}\to\infty} \mu([0, u_1] \times \ldots \times [0, u_d]).$$

These results are proved, e.g., in Section 4.5 in Kingman and Taylor (1966).

In this work, we focus on the theory of Lévy copulas for subordinators. Therefore, it is sufficient to define the Lévy copula on $\overline{\mathbb{R}}_+^d$. This is no restriction at all since it is straightforward to define a Lévy copula for every orthant separately. In two dimensions, for example, we need to define four Lévy copulas. Proceeding this way, we can use the techniques for Lévy copulas on $\overline{\mathbb{R}}_+^d$ to construct Lévy processes with positive and negative jumps in all dimensions. For the theoretical background on general Lévy copulas on $\overline{\mathbb{R}}^d$ we refer to Kallsen and Tankov (2006).

5.2.2 Parametric Lévy Copula Families

In the following chapter, we use bivariate Lévy copulas to construct Lévy copulas in arbitrary dimensions. Therefore, we are interested in parametric families of bivariate Lévy copulas. There are several ways to define these 2-dimensional dependence functions. The first method that we present uses distributional copulas and a function to adjust for the domain and the image of the copula. This procedure is suggested in Cont and Tankov (2004, Proposition 5.5).

Proposition 5.19 *Let C be a distributional bivariate copula and $f : [0,1] \mapsto \overline{\mathbb{R}}_+$ an increasing, convex function. Then*

$$\mathfrak{C}(u,v) = f(C(f^{-1}(u), f^{-1}(v)))$$

defines a 2-dimensional Lévy copula.

A different class of Lévy copulas is defined similarly to Archimedean distributional copulas, which is shown in Cont and Tankov (2004, Proposition 5.6).

Proposition 5.20 *Let $\phi : \overline{\mathbb{R}}_+ \mapsto \overline{\mathbb{R}}_+$ be a strictly decreasing convex function such that $\phi(0) = \infty$ and $\phi(\infty) = 0$. Then*

$$\mathfrak{C}(u,v) = \phi^{-1}(\phi(u) + \phi(v))$$

defines a 2-dimensional Lévy copula.

An example for such a Lévy copula, which is used later in this thesis, is the Clayton-Lévy copula (Cont and Tankov, 2004, Example 5.5).

Example 5.21 *The Clayton-Lévy copula on $\overline{\mathbb{R}}_+^2$ for 2-dimensional Lévy subordinators is given by*

$$\mathfrak{C}(u,v) = \left(u^{-\theta} + v^{-\theta}\right)^{-1/\theta}. \tag{5.9}$$

Here, $\theta > 0$ determines the dependence of the jump sizes, where larger values of θ indicate a stronger dependence.

Chapter 6

The Pair-Lévy Copula Construction

To the best knowledge of the author, all papers involving Lévy copulas focus on rather small dimensions since higher-dimensional flexible Lévy copulas are difficult to construct. A similar effect has been observed during the first years of literature on distributional copulas, where mainly 2-dimensional distributional copulas have been analyzed. One solution regarding distributional copulas has been the development of very flexible pair constructions of copulas, see Chapter 3. In the pair-copula construction, a d-dimensional copula is constructed from $d(d-1)/2$ bivariate copulas, where $d-1$ of the bivariate copulas model the dependence of bivariate margins, and the remaining bivariate copulas model certain conditional distributions, such that the entire d-dimensional dependence structure is specified. The Lévy copula concept is conceptually different from distributional copulas. While d-dimensional distributional copulas are distribution functions on a $[0,1]^d$ hypercube, d-dimensional Lévy copulas are defined on $\overline{\mathbb{R}}_+^d$ and relate to Radon measures, see Section 5.2. Therefore, the pair construction idea for distributional copulas is not directly transferable to Lévy copulas and up to now it has not been clear whether it is possible at all. In this chapter, we show that a pair-copula construction of Lévy copulas is indeed possible. It also consists of $d(d-1)/2$ bivariate dependence functions but only $d-1$ of them are Lévy copulas, while the remaining ones are distributional copulas. For statistical inference, we derive sequential maximum likelihood estimators for an arbitrary pair construction of Lévy copulas as well as a simulation algorithm. We analyze the applicability of the concept in a simulation study. Moreover, we show in detail how the pair-Lévy copula construction (PLCC) concept can be applied to model the dependence structure of multivariate Ornstein-Uhlenbeck processes and we give an outlook on further applications. This chapter is joint work with Oliver Grothe and several parts of it are accepted for publication (Grothe and Nicklas, 2013).

6.1 Pair-Lévy Copulas

In this section, we present the theory of the pair construction for d-dimensional Lévy copulas. In particular, we show that analogously to the pair construction of distributional copulas, $d(d-1)/2$ functions of bivariate dependence may be arranged such that they define a d-dimensional Lévy copula. In Section 6.1.2, we provide illustrating ex-

amples how to construct multivariate pair-Lévy copulas. Readers not interested in the technical parts may read these examples first.

6.1.1 Technical Part

The central theorem for the construction is Theorem 6.4. It states that two $(d-1)$-dimensional Lévy copulas with overlapping $(d-2)$-dimensional margins may be coupled to a d-dimensional Lévy copula by a new, 2-dimensional distributional copula. Ensured by vine constructions (see Bedford and Cooke (2002)) and starting at $(d-1)=2$, Theorem 6.4 therefore enables to sequentially construct Lévy copulas out of 2-dimensional dependence functions, i.e., 2-dimensional distributional copulas and Lévy copulas. Before we state the theorem, for convenience, we recall some definitions which can be found, e.g., in Ambrosio *et al.* (2000, Chapter 1 & 2).

Definition 6.1 *Let $(X,\mathcal{E})$ and $(Y,\mathcal{F})$ be measure spaces, and let μ be a positive measure on $(X,\mathcal{E})$.*

(i) We say that $N\subset X$ is μ-negligible if there exists $E\in\mathcal{E}$ such that $N\subset E$ and $\mu(E)=0$.

(ii) Let $\mathcal{E}_\mu$ be the collection of all the subsets of X of the form $F=E\cup N$, with $E\in\mathcal{E}$ and N μ-negligible; then $\mathcal{E}_\mu$ is a σ-algebra which is called the μ-completion of $\mathcal{E}$.

(iii) A function $f:X\mapsto Y$ is said to be $\mathcal{E}$-measurable if $f^{-1}(A)\in\mathcal{E}$ for every $A\in\mathcal{F}$.

(iv) The function f is said to be μ-measurable if it is $\mathcal{E}_\mu$-measurable.

(v) A positive measure on $(\mathbb{R}_+^d,\mathcal{B}(\mathbb{R}_+^d))$ that is finite on compact sets is called a positive Radon measure.

(vi) Let $f:X\to Y$ be a μ-measurable function. We define the push forward measure $f_\#\mu$ in $(Y,\mathcal{F})$ by

$$f_\#\mu(K):=\mu\left(f^{-1}(K)\right)\quad\forall K\in\mathcal{F}.$$

(vii) Let μ be a positive Radon measure on $\mathbb{R}_+^d$ and $x\mapsto\xi_x$ a function which assigns a finite Radon measure ξ_x on $\mathbb{R}_+^m$ to each $x\in\mathbb{R}_+^d$. We say this map is μ-measurable if $x\mapsto\xi_x(B)$ is μ-measurable for any $B\in\mathcal{B}(\mathbb{R}_+^m)$.

Definition 6.2 *Let μ be a positive Radon measure on $\mathbb{R}_+^d$ and $x\mapsto\xi_x$ a μ-measurable function which assigns a probability measure ξ_x on $\mathbb{R}_+^m$ to each $x\in\mathbb{R}_+^d$. We denote by $\mu\otimes\xi_x$ the Radon measure on $\mathbb{R}_+^{d+m}$ defined by*

$$\mu\otimes\xi_x(B):=\int_{\mathbb{R}_+^d}\left(\int_{\mathbb{R}_+^m}\mathbb{1}_B(x,y)d\xi_x(y)\right)d\mu(x)\quad\forall B\in\mathcal{B}(K\times\mathbb{R}_+^m),$$

where $K\subset\mathbb{R}_+^d$ is any compact set.

We also need a theorem which states that a Radon measure may be decomposed into a projection onto some of its dimensions and a probability measure. For a proof see Theorem 2.28 in Ambrosio *et al.* (2000) and also the sentence after Corollary 2.29 there.

Theorem 6.3 *Let $\mu_{1,\dots,d+m}$ be a Radon measure on $\mathbb{R}_+^{d+m}$, $\pi : \mathbb{R}_+^{d+m} \mapsto \mathbb{R}_+^d$ the projection on the first d variables and $\mu_{1,\dots,d} = \pi_{\#}\mu_{1,\dots,d+m}$. Let us assume that $\mu_{1,\dots,d}$ is a positive Radon measure, i.e., that $\mu_{1,\dots,d+m}(K \times \mathbb{R}_+^m) < \infty$ for any compact set $K \subset \mathbb{R}_+^d$. Then, there exists a finite measure ξ_x in $\mathbb{R}_+^m$ such that $x \mapsto \xi_x$ is $\mu_{1,\dots,d}$-measurable, ξ_x is a probability measure almost everywhere in $\mathbb{R}_+^d$, and for any $B \in \mathcal{B}(K \times \mathbb{R}_+^m)$, where $K \subset \mathbb{R}_+^d$ is any compact set*

$$\int_{\mathbb{R}_+^{d+m}} \mathbb{1}_B(x,y) d\mu_{1,\dots,d+m}(x,y) = \int_{\mathbb{R}_+^d} \left(\int_{\mathbb{R}_+^m} \mathbb{1}_B(x,y) d\xi_x(y) \right) d\mu_{1,\dots,d}(x).$$

This is $\mu_{1,\dots,d+m}(B) = \mu_{1,\dots,d} \otimes \xi_x(B)$.

We are now able to state the main theorem that shows how we can construct high-dimensional, parametric Lévy copulas.

Theorem 6.4 *Let $\mathfrak{C}_{1,\dots,d-1}$ and $\mathfrak{C}_{2,\dots,d}$ be two Lévy copulas on $\overline{\mathbb{R}}_+^{d-1}$ where $\mathfrak{C}_{1,\dots,d-1}$ is a Lévy copula on the variables $u_1, \dots, u_{d-1}$ and $\mathfrak{C}_{2,\dots,d}$ is a Lévy copula on the variables $u_2, \dots, u_d$. Denote the corresponding measures on $\mathbb{R}_+^{d-1}$ by $\mu_{1,\dots,d-1}$ and $\mu_{2,\dots,d}$, respectively. Suppose that the two measures have an identical $(d-2)$-dimensional margin $\mu_{2,\dots,d-1}$ on the variables $u_2, \dots, u_{d-1}$. Then, we can define a Lévy copula on $\mathbb{R}_+^d$ by*

$$\mathfrak{C}_{1,\dots,d}(u_1,\dots,u_d) := \int\limits_{[0,u_2]\times\dots\times[0,u_{d-1}]} C(F_{1|z_2,\dots,z_{d-1}}(u_1), F_{d|z_2,\dots,z_{d-1}}(u_d)) d\mu_{2,\dots,d-1}(z_2,\dots,z_{d-1}),$$

where $F_{1|u_2,\dots,u_{d-1}}$ is the one-dimensional distribution function corresponding to the probability measure $\xi_{1|u_2,\dots,u_{d-1}}$ from the decomposition of $\mu_{1,\dots,d-1}$ into

$$\mu_{1,\dots,d-1} = \mu_{2,\dots,d-1} \otimes \xi_{1|u_2,\dots,u_{d-1}},$$

$F_{d|u_2,\dots,u_{d-1}}$ is the one-dimensional distribution function corresponding to the probability measure $\xi_{d|u_2,\dots,u_{d-1}}$ from the decomposition of $\mu_{2,\dots,d}$ into

$$\mu_{2,\dots,d} = \mu_{2,\dots,d-1} \otimes \xi_{d|u_2,\dots,u_{d-1}},$$

and C is a distributional copula. Since Lévy copulas are functions on $\overline{\mathbb{R}}_+^d$, we set for every $I \subset \{1,\dots,d\}$ nonempty,

$$\mathfrak{C}_{1,\dots,d}(u_1,\dots,u_d)|_{(u_i)_{i\in I}=\infty} := \lim_{(u_i)_{i\in I}\to\infty} \mathfrak{C}_{1,\dots,d}(u_1,\dots,u_d). \tag{6.1}$$

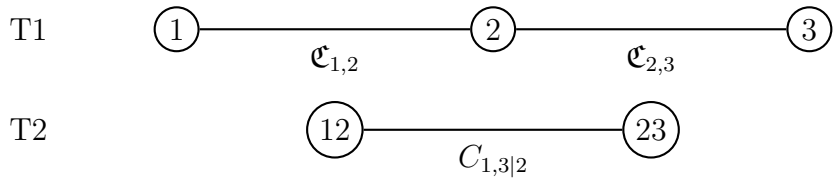


Figure 6.1: Pair-Lévy copula construction of a 3-dimensional Lévy copula out of $3(3-1)/2 = 3$ bivariate dependence functions. The functions $\mathfrak{C}_{1,2}$ and $\mathfrak{C}_{2,3}$ in the first tree are Lévy copulas, while $C_{1,3|2}$ in the second tree is a distributional copula.

The theorem, which is proved in Appendix 6.B, illustrates how to construct a d-dimensional Lévy copula from two $(d-1)$-dimensional Lévy copulas with a common margin. Applying the theorem recursively, these $(d-1)$-dimensional Lévy copulas can be constructed from $(d-2)$-dimensional ones. This can be repeated down to construct 3-dimensional Lévy copulas from bivariate ones.

6.1.2 Pair-Lévy Copula Construction

In higher dimensions, the pair-copula construction method offers many ways to build multivariate Lévy copulas from bivariate dependence functions. This is due to permutations of the dimensions and numerous possible pairwise combinations within each step. We visualize the different structures of the pair construction by the concept of regular vines, see Section 3.2. Regular vines also help to construct pair-Lévy copulas top-down. This means to start with $d-1$ bivariate Lévy copulas and to combine them successively to $3, 4, 5, \ldots, d$-dimensional Lévy copulas. The regular vine approach ensures that at each step the involved Lévy copulas have sufficiently overlapping margins, and that therefore the theorem can be applied. To illustrate this approach, we give two detailed examples. The first example refers to the most simple case, a 3-dimensional Lévy copula. The second, 4-dimensional example then illustrates how to sequentially add dimensions to the pair-Lévy copula construction.

Example 6.5 *A 3-dimensional example can be constructed applying Theorem 6.4 to combine two 2-dimensional Lévy copulas by a distributional copula. As in the usual pair-copula construction for distributional copulas, in Figure 6.1 we use the vine concept to visualize the resulting dependence structure. The bivariate dependence structures in the first tree are Lévy copulas, whereas the copula in the second tree is a distributional copula. From Theorem 6.4 follows that*

$$\mathfrak{C}_{1,2,3}(u_1, u_2, u_3) = \int_{[0,u_2]} C_{1,3|2}(F_{1|z_2}(u_1), F_{3|z_2}(u_3)) d\mu_2(z_2) \tag{6.2}$$

is a Lévy copula, where $F_{1|u_2}(u_1)$ is the one-dimensional distribution function corre-

sponding to the probability measure $\xi_{1|u_2}$ from the decomposition of $\mu_{1,2}$ into

$$\mu_{1,2} = \mu_2 \otimes \xi_{1|u_2}, \tag{6.3}$$

and $F_{3|u_2}$ is the one-dimensional distribution function corresponding to the probability measure $\xi_{3|u_2}$ from the decomposition of $\mu_{2,3}$ into

$$\mu_{2,3} = \mu_2 \otimes \xi_{3|u_2}. \tag{6.4}$$

Remember that $\mu_{1,2}$ is the Radon measure corresponding to $\mathfrak{C}_{1,2}$. With Theorem 6.3 and the considerations after Assumption 5.18 we see that μ_2 in Equation (6.3) *is the Lebesgue measure. Analogously, $\mu_{2,3}$ is the Radon measure corresponding to $\mathfrak{C}_{2,3}$ and therefore μ_2 in Equation* (6.4) *is the Lebesgue measure as well. To check whether $\mathfrak{C}_{1,2,3}(u_1, u_2, u_3)$ has the desired margins, we calculate*

$$\begin{aligned}
\mathfrak{C}_{1,2,3}(u_1, u_2, \infty) &= \int_{[0,u_2]} C_{1,3|2}(F_{1|z_2}(u_1), F_{3|z_2}(\infty))dz_2 \\
&= \int_{[0,u_2]} C_{1,3|2}(F_{1|z_2}(u_1), 1)dz_2 \\
&= \int_{[0,u_2]} F_{1|z_2}(u_1)dz_2 \\
&= \int_{[0,u_2]} \left(\int_{[0,u_1]} d\xi_{1|z_2}(z_1) \right) dz_2 \\
&= \int_{[0,u_1]\times[0,u_2]} d\mu_{1,2}(z_1, z_2) \\
&= \mathfrak{C}_{1,2}(u_1, u_2).
\end{aligned}$$

A similar procedure shows that

$$\mathfrak{C}_{1,2,3}(\infty, u_2, u_3) = \mathfrak{C}_{2,3}(u_2, u_3).$$

As expected, we do not get such a direct representation of the third bivariate margin

$$\mathfrak{C}_{1,2,3}(u_1, \infty, u_3) = \int_{[0,\infty)} C_{1,3|2}(F_{1|z_2}(u_1), F_{3|z_2}(u_3))dz_2$$

because this margin is not only influenced by the distributional copula $C_{1,3|2}$ but also by $\mathfrak{C}_{1,2}$ and $\mathfrak{C}_{2,3}$. Note that, by changing $C_{1,3|2}$, we can modify the bivariate margin of the first and third dimension without any influence the other two bivariate margins.

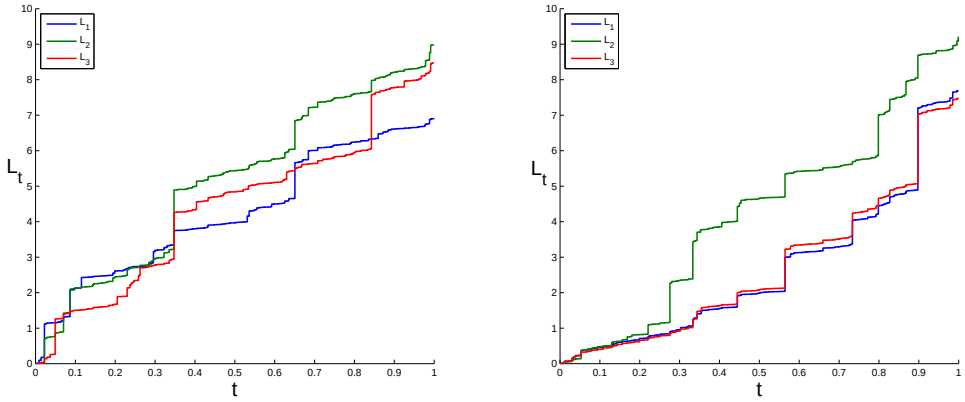


Figure 6.2: Realizations of 3-dimensional Lévy processes with stable margins. The dependence structure is modeled by the Lévy copula given in Example 6.5. In both cases we use Clayton-Lévy copulas with parameter $\theta = 1$ in the first tree. In the second tree, we use a Gaussian copula. For the process on the left, we set the parameter of $C_{1,3|2}$ to $\rho = -0.99$ and for the process on the right we set $\rho = 0.99$.

In Appendix 6.A, we give an example for a 3-dimensional pair-Lévy copula. In this special case, we combine two Clayton-Lévy copulas and one distributional Clayton copula and give an explicit formula for the resulting PLCC. In Figure 6.2, we illustrate the flexibility of the PLCC concept. There, we see two realizations of 3-dimensional Lévy processes. All univariate processes are stable with parameters $\alpha = 0.8$ and $\beta = 1.25$. To visualize the influence of the distributional copula in the second tree, we use the same Clayton-Lévy copulas with parameter $\theta = 1$ in the first tree for both processes. In the second tree we use a Gaussian distributional copula. For the process on the left, we set the parameter of the copula $C_{1,3|2}$ to $\rho = -0.99$ and for the process on the right, we set $\rho = 0.99$. Omitting the third dimension for a moment and only looking at the processes L_1 and L_2 in Figure 6.2, we see joint jumps in the realization of these stochastic processes. These joint jumps occur, for example, at small values of t on the left hand side and around $t = 0.56$ on the right. We also observe jumps in the first dimension without noticeable jumps in the second dimension in both realizations. We do not see much qualitative difference in the jump dependence of L_1 and L_2 between the path on the left and the path on the right in Figure 6.2. This is no surprise since the dependence between the first and the second component is modeled by the same bivariate Lévy copula $\mathfrak{C}_{1,2}$ for both processes. The same holds true if we focus on the jump dependence between L_2 and L_3 since this dependence is also modeled by the same Lévy copula $\mathfrak{C}_{2,3}$ for both processes. As we know from the theory, the jump dependence between L_1 and L_3 is the only bivariate marginal dependence that is affected by the choice of $C_{1,3|2}$. This effect can be observed if we compare both graphs in Figure 6.2. In the graph on the right, the processes L_1 and L_3 always jump at the same time and when they jump, the jump size is almost the same. This effect is due to the distributional copula $C_{1,3|2}$ with a very

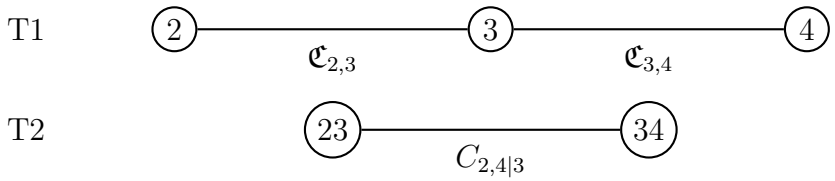


Figure 6.3: Pair construction of the second three dimensions of a 4-dimensional Lévy copula out of $3(3-1)/2 = 3$ bivariate dependence functions. The functions $\mathfrak{C}_{2,3}$ and $\mathfrak{C}_{3,4}$ in the first tree are Lévy copulas, while $C_{2,4|3}$ in the second tree is a distributional copula. The Lévy copula $\mathfrak{C}_{2,3}$ is the same Lévy copula as in Figure 6.1 which refers to the pair construction of the first three dimensions.

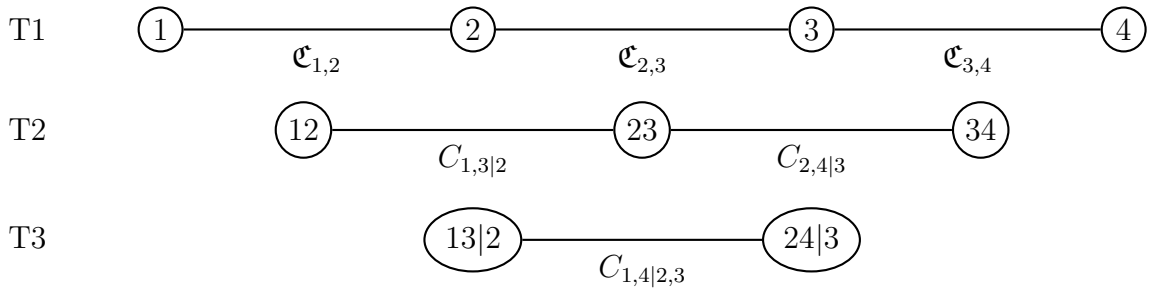


Figure 6.4: Combination of the first three dimensions and the second three dimensions to a pair construction of a 4-dimensional Lévy copula. It consists of $4(4-1)/2 = 6$ bivariate dependence functions. The functions in the first tree are Lévy copulas, while the functions in the second and third tree are distributional copulas.

high value of $\rho = 0.99$ in the PLCC. In the graph on the left, the processes L_1 and L_3 rarely jump at the same time and in cases they do jump together, we do not observe similar jump sizes. We achieve this effect by using a distributional copula with negative dependence in the second tree.

Example 6.6 *Considering* 4 *dimensions, we need two* 3*-dimensional Lévy copulas with an identical* 2*-dimensional margin. Here, we reuse the Lévy copula from Example 6.5 for the first three dimensions. The second* 3*-dimensional Lévy copula is constructed in the same way and has the vine representation shown in Figure 6.3.*

Notice that the Lévy copula $\mathfrak{C}_{2,3}$ *is used in both* 3*-dimensional pair-Lévy copulas. Therefore, the marginal Lévy copulas*

$$\mathfrak{C}_{1,2,3}(\infty, u_2, u_3) = \mathfrak{C}_{2,3}(u_2, u_3) = \mathfrak{C}_{2,3,4}(u_2, u_3, \infty)$$

are the same and we can apply Theorem 6.4 to construct a 4*-dimensional Lévy copula*

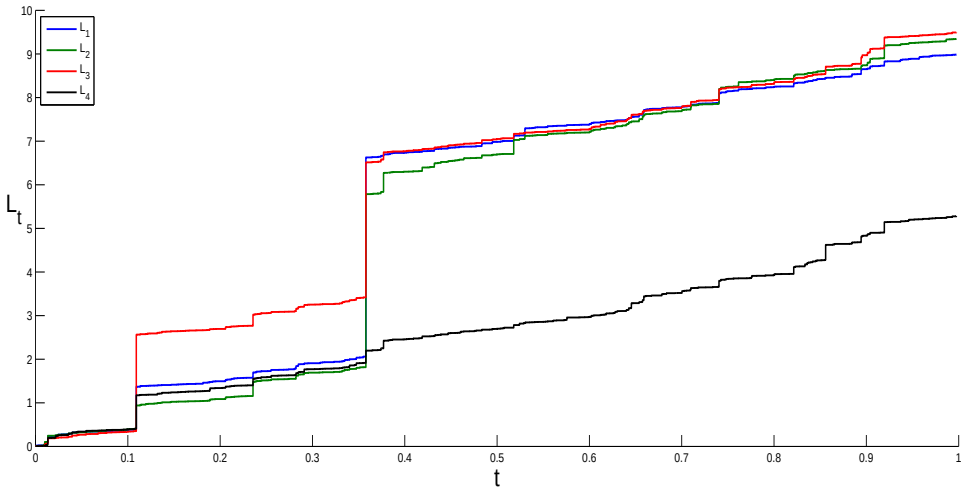


Figure 6.5: Realization of a 4-dimensional Lévy process with stable margins. The dependence structure is modeled by the pair-Lévy copula construction given in Example 6.6. We use three bivariate Clayton-Lévy copulas with parameter $\theta = 2$ in the first tree. In the second and third tree, we use Gaussian copulas with parameter $\rho = 0.3$.

with the vine representation shown in Figure 6.4 and

$$\mathfrak{C}_{1,2,3,4}(u_1, u_2, u_3, u_4) = \int_{[0,u_2]\times[0,u_3]} C_{1,4|2,3}(F_{1|z_2,z_3}(u_1), F_{4|z_2,z_3}(u_4))d\mu_{2,3}(z_2, z_3),$$

where $F_{1|u_2,u_3}$ *is the one-dimensional distribution function corresponding to the probability measure* $\xi_{1|u_2,u_3}$ *from the decomposition of* $\mu_{1,2,3}$ *from the first pair-Lévy copula* $\mathfrak{C}_{1,2,3}$ *into*

$$\mu_{1,2,3} = \mu_{2,3} \otimes \xi_{1|u_2,u_3}.$$

The one-dimensional distribution function $F_{4|u_2,u_3}$ *corresponds to the probability measure* $\xi_{4|u_2,u_3}$ *from the decomposition of* $\mu_{2,3,4}$ *from the second pair-Lévy copula* $\mathfrak{C}_{2,3,4}$ *into*

$$\mu_{2,3,4} = \mu_{2,3} \otimes \xi_{4|u_2,u_3}.$$

In Figure 6.5, we show a realization of a 4-dimensional PLCC. All univariate processes are stable with parameters $\alpha = 0.8$ and $\beta = 1.25$. The structure of the pair-Lévy copula construction is as visualized in Figure 6.4. We use three Clayton-Lévy copulas with dependence parameter $\theta = 2$ in the first tree. The distributional copulas in the second and third tree are all Gaussian with the same dependence parameter $\rho = 0.3$. This setting represents a medium magnitude of dependence. We see that the large jumps in the processes tend to occur at the same time, i.e. at $t = 0.11$ and $t = 0.36$. However, the jump sizes do not coincide perfectly between the four dimensions.

6.2 Simulation and Estimation

In this section we discuss the simulation of multivariate Lévy processes as well as the maximum likelihood estimation of the pair-Lévy copula. We need the following assumption which is fulfilled by the common parametric families of the bivariate (Lévy) copulas.

Assumption 6.7 *In the following, we assume that all bivariate distributional and Lévy copulas are continuously differentiable.*

6.2.1 Simulation

The simulation of multivariate Lévy processes built upon Lévy copulas is based on a series representation for Lévy processes and the following theorem.

Theorem 6.8 *Let ν be a Lévy measure on $\mathbb{R}_+^d$, satisfying $\int_{\mathbb{R}_+^d}(|x| \wedge 1)d\nu(x) < \infty$, with marginal tail integrals U_j, $j = 1, \dots, d$ and Lévy copula $\mathfrak{C}_{1,\dots,d}$ with corresponding measure $\mu_{1,\dots,d}$. Let $(V_i)_{i\in\mathbb{N}}$ be a sequence of independent and uniformly $[0,1]$ distributed random variables and $(\Gamma_i^1, \dots, \Gamma_i^{d-1})_{i\in\mathbb{N}}$ be a Poisson point process on $\mathbb{R}_+^{d-1}$ with intensity measure $\mu_{1,\dots,d-1}$ from the decomposition of*

$$\mu_{1,\dots,d} = \mu_{1,\dots,d-1} \otimes \xi_{d|u_1,\dots,u_{d-1}},$$

with $\xi_{d|u_1,\dots,u_{d-1}}$ being a probability measure. For any value of $\Gamma_i^1, \dots, \Gamma_i^{d-1}$, we suppose that Γ_i^d is a random variable with probability measure $\xi_{d|\Gamma_i^1,\dots,\Gamma_i^{d-1}}$. Then, the process $(L_t^1, \dots, L_t^d)_{t\in[0,1]}$ defined by

$$L_t^j = \sum_{i=1}^{\infty} U_j^{-1}(\Gamma_i^j)\mathbb{1}_{[0,t]}(V_i), \quad j = 1, \dots, d$$

is a d-dimensional Lévy process $(L_t)_{t\in[0,1]}$ without a Brownian component and drift. The Lévy measure of (L_t) is ν.

Proof: The proof is similar to the proof of Tankov (2005, Theorem 4.3).

In practical simulations, the sum cannot be evaluated up to infinity and one omits very small jumps. The sequence $(\Gamma_i^1)_{i\in\mathbb{N}}$ is therefore only simulated up to a sufficiently large N, resulting in a large value of Γ_N^1 (see Rosiński (2001) for this approximation). Note that large values of Γ_i^1 correspond to small values of the jumps $U_1^{-1}(\Gamma_i^1)$, since the tail integral is decreasing.

Based on the pair-copula construction of the Lévy copula, $\Gamma_i^2, \dots, \Gamma_i^d$ can be drawn conditionally on Γ_i^1 in a sequential way. For convenience, assume that the pair-Lévy

copula has a D-vine structure and that the dimensions are ordered from left to right. The dependence between Γ_i^1 and Γ_i^2 is then determined in the first tree of the pair construction by the bivariate Lévy copula $\mathfrak{C}_{1,2}$, and the distribution function $F_{2|\Gamma_i^1}$ of Γ_i^2 given Γ_i^1 is derived in the following Proposition.

Proposition 6.9 *Let $\mathfrak{C}_{1,2}$ be a 2-dimensional Lévy copula with corresponding measure $\mu_{1,2}$. Then, we can decompose*

$$\mu_{1,2} = \mu_1 \otimes \xi_{2|u_1},$$

where $\xi_{2|u_1}$ is a probability measure and the distribution function for almost all $u_1 \in [0,\infty)$ is given by

$$F_{2|u_1}(u_2) = \frac{\partial \mathfrak{C}_{1,2}(u_1,u_2)}{\partial u_1}.$$

Proof: This is a special case of Tankov (2005, Lemma 4.2).

Inverting this distribution function allows the simulation of Γ_i^2. Now suppose that we have already simulated the variables $\Gamma^1,\dots,\Gamma^{d-1}$, $d \geq 3$ and we want to simulate the last variable Γ^d. We already know from Theorem 6.3 that the distribution of the last variable, given the first $d-1$, is a probability distribution and therefore we are interested in the corresponding distribution function $F_{d|u_1,\dots,u_{d-1}}$. Having found $F_{d|u_1,\dots,u_{d-1}}$, we can again invert it and easily simulate a realization of a random variable with this distribution function. The next proposition provides $F_{d|u_1,\dots,u_{d-1}}$ within the pair construction of the Lévy copula.

Proposition 6.10 *Let $d \geq 3$ and $\mathfrak{C}_{1,\dots,d}$ be a pair-Lévy copula, $\mu_{1,\dots,d}$ the corresponding measure, π the projection on the first $d-1$ variables, and $\mu_{1,\dots,d-1} = \pi \# \mu_{1,\dots,d}$ the push forward measure. Then, we can decompose*

$$\mu_{1,\dots,d} = \mu_{1,\dots,d-1} \otimes \xi_{d|u_1,\dots,u_{d-1}},$$

where $\xi_{d|u_1,\dots,u_{d-1}}$ is a probability measure on $\mathbb{R}_+$ with distribution function

$$F_{d|u_1,\dots,u_{d-1}}(u_d) = \frac{\partial C_{1,d|2,\dots,d-1}(F_{1|u_2,\dots,u_{d-1}}(u_1), F_{d|u_2,\dots,u_{d-1}}(u_d))}{\partial F_{1|u_2,\dots,u_{d-1}}(u_1)}$$

$\mu_{1,\dots,d-1}$-almost everywhere. Moreover, $F_{d|u_1,\dots,u_{d-1}}$ is continuously differentiable.

The proposition is proved in the appendix. Similar to Aas *et al.* (2009), it shows how we can iteratively evaluate and invert the distribution function $F_{d|u_1,\dots,u_{d-1}}$.

6.2.2 Maximum Likelihood Estimation

It is usually not possible to track Lévy processes in continuous time. Therefore, we have to choose a more realistic observation scheme. In the context of inference for pure

jump Lévy processes, a common assumption is that it is possible to observe all jumps of the processes larger than a given ε (see, e.g., Basawa and Brockwell (1978, 1980) or Esmaeili and Klüppelberg (2010)). Following Esmaeili and Klüppelberg (2011b), we estimate the marginal Lévy processes separately from the dependence structure. That is, for dimension $j \in \{1, \dots, d\}$, we use all $N_j^{(\varepsilon)}$ observed jumps with jump sizes $x_{1,j}, \dots, x_{N_i^{(\varepsilon)},j}$ that are larger than ε in dimension j, and estimate the parameters of the one-dimensional Lévy process. For the estimation of the dependence structure, i.e., the Lévy copula, we can use the fact that the process, consisting of all $N_{1,\dots,d}^{(\varepsilon)}$ jumps where the jump sizes $x_{i,j}$ are larger than ε in all dimensions, is a compound Poisson process. Note that, $N_{1,\dots,d}^{(\varepsilon)} \leq \min\{N_1^{(\varepsilon)}, \dots, N_d^{(\varepsilon)}\}$. We suppose that all densities $f_1, \dots, f_d$ of the marginal Lévy measures exist and we denote the parameter vectors of the Lévy copula and the marginal Lévy measures by $\delta, \gamma_1, \dots, \gamma_d$, respectively. The likelihood function is given by

$$L^{\varepsilon}(\delta, \gamma_1, \dots, \gamma_d) = \tag{6.5}$$
$$e^{-\lambda_{1,\dots,d}^{(\varepsilon)} t} \prod_{i=1}^{N_{1,\dots,d}^{(\varepsilon)}} \left[f_1(x_{i,1}, \gamma_1) \cdot \ldots \cdot f_d(x_{i,d}, \gamma_d) \mathfrak{c}_{1,\dots,d}(U_1(x_{i,1}, \gamma_1), \dots, U_d(x_{i,d}, \gamma_d), \delta) \right],$$

where $\lambda_{1,\dots,d}^{(\varepsilon)}(\delta) = \mathfrak{C}_{1,\dots,d}(U_1(\varepsilon, \gamma_1), \dots, U_d(\varepsilon, \gamma_d), \delta)$ and $\mathfrak{c}_{1,\dots,d}$ is the density of $\mathfrak{C}_{1,\dots,d}$. This result also holds for m-dimensional marginal Lévy processes with $m < d$ and is already stated in Esmaeili and Klüppelberg (2011a) for two dimensions.

To use the above likelihood for pair-Lévy copula constructions, we have to know how to calculate the density $\mathfrak{c}_{1,\dots,d}$ of a pair-Lévy copula.

Proposition 6.11 *Let $\mathfrak{C}_{1,\dots,d}$ be a pair-Lévy copula of the following form*

$$\mathfrak{C}_{1,\dots,d}(u_1, \dots, u_d) = \int_{[0,u_2]\times\dots\times[0,u_{d-1}]} C(F_{1|z_2,\dots,z_{d-1}}(u_1), F_{d|z_2,\dots,z_{d-1}}(u_d)) d\mu_{2,\dots,d-1}(z_2, \dots, z_{d-1})$$

and $\mu_{1,\dots,d}$ the corresponding measure and suppose that the density $f_{2,\dots,d-1}$ of $\mu_{2,\dots,d-1}$ exists. Then the density of $\mu_{1,\dots,d}$ exists as well and has the form

$$\begin{aligned} f_{1,\dots,d}(u_1, \dots, u_d) &= c(F_{1|u_2,\dots,u_{d-1}}(u_1), F_{d|u_2,\dots,u_{d-1}}(u_d)) \\ &\quad \cdot \frac{\partial F_{1|u_2,\dots,u_{d-1}}(u_1)}{\partial u_1} \frac{\partial F_{d|u_2,\dots,u_{d-1}}(u_d)}{\partial u_d} \\ &\quad \cdot f_{2,\dots,d-1}(u_2, \dots, u_{d-1}). \end{aligned}$$

This proposition is proved in the appendix and states that we can iteratively decompose the pair-Lévy copula into bivariate building blocks and therefore evaluate the density function in an efficient manner.

A straightforward estimation approach would be maximizing the full likelihood function to estimate the dependence structure. This, however, is disadvantageous because of

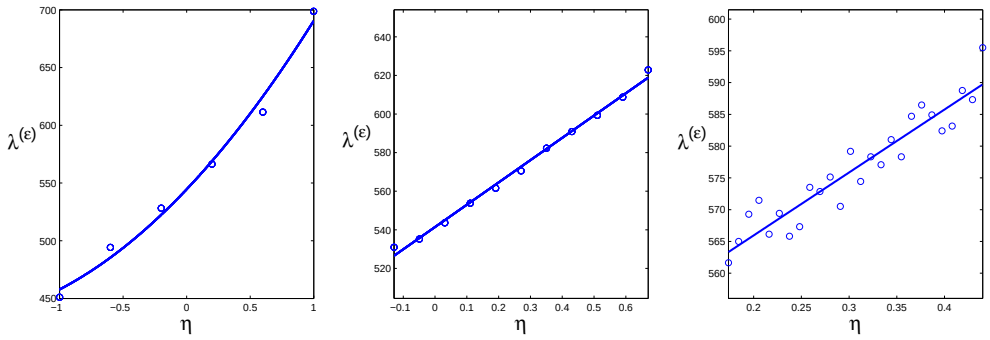


Figure 6.6: Evaluation of a pair-Lévy copula for different values of η by Monte Carlo methods, where η is the parameter of the bivariate distributional copula used in the last tree of the PLCC. This methods are used in the sequential estimation algorithm.

two reasons. The first reason is a numerical one. The likelihood function is not easy to evaluate if more than one parameter is unknown. The second reason is more conceptual. Since we can only use jumps larger than ε in all d dimensions, we waste a tremendous part of the information about the dependence structure, especially if the dependence structure is weak. For weak dependence structures, the probability that two jumps are both larger than a threshold (conditioned that at least one jump exceeds the threshold) is lower than for strong dependence.

For both reasons, we estimate the parameters of the bivariate Lévy and distributional copulas of the vine structure sequentially. This is also common for pair-copula constructions of distributional copulas (see, e.g., Hobæk Haff (2012)). We make use of the estimated marginal parameters and start in the first tree, using all observations larger than ε in the first and second components to estimate the parameters of $\mathfrak{C}_{1,2}$. Then, we use all observations larger than ε in the second and third components to estimate the parameters of $\mathfrak{C}_{2,3}$. We continue this procedure for all other Lévy copulas in the first tree. To estimate the parameter of $C_{1,3|2}$, we use all observations larger than ε in dimensions one, two, and three as well as the previously estimated marginal parameters of the first three dimensions and the parameters of $\mathfrak{C}_{1,2}$ and $\mathfrak{C}_{2,3}$. This means that we proceed tree by tree and within the tree, copula by copula or Lévy copula by Lévy copula, respectively. In each step, we make use of the estimated parameters from the preceding steps. In contrast to the computation of the density, it is not easy to evaluate a pair-Lévy copula itself. This is no real drawback since the value of $\mathfrak{C}_{1,\dots,d}$ is not needed in most cases. For the normalizing constant $\lambda^{(\varepsilon)}_{1,\dots,d}(\delta)$ of the likelihood, however, $\mathfrak{C}_{1,\dots,d}$ has to be evaluated. In the bivariate case we do not face any problems since usually explicit formulas to evaluate the Lévy copula exist. In higher dimensions we have to evaluate the pair-Lévy copula, and therefore we need to solve high-dimensional integrals. Even worse, since $\lambda^{(\varepsilon)}_{1,\dots,d}(\delta)$ is part of the likelihood function, see Equation (6.5), we need to solve these integrals in every step of the maximum likelihood optimization algorithm. Thus, using numerical integration methods seems out of reach. Straightforward Monte Carlo

integration cannot be used either, since the Monte Carlo error causes severe problems in the optimization routines. This is because randomly occurring large Monte Carlo errors might lead to strong positive deviations from the true value of the function. This causes the optimization routine to stop far away from the maximum, since other evaluations in this area give smaller values and indicate spuriously a local maximum.

The solution, we propose, makes use of Monte Carlo methods but still allows for the usual optimization routines. To simplify the notation, we present this estimation procedure for a one-parametric bivariate distributional copula in the highest tree of a pair-Lévy copula construction. Note that this procedure can be applied to any other distributional copula in the vine structure as well. Since we focus on the last tree of the vine structure, we suppose that all parameters of the parameter vector δ have already been sequentially estimated except for the parameter of the bivariate distributional copula in the last tree. We denote this parameter under consideration by η, and to emphasize that all other parameters are treated as constants in this sequential estimation step, we denote the normalizing constant by $\lambda^{(\varepsilon)}(\eta)$. In a first step, we evaluate $\lambda^{(\varepsilon)}(\eta)$ for different values of η and approximate the function $\lambda^{(\varepsilon)}(\eta)$ by the polynomial $\tilde{\lambda}^{(\varepsilon)}(\eta)$. We illustrate this procedure on the left hand side of Figure 6.6. Then, we replace $\lambda^{(\varepsilon)}(\eta)$ in the likelihood function by the smooth polynomial $\tilde{\lambda}^{(\varepsilon)}(\eta)$ and maximize the likelihood function. To enhance the speed and the accuracy, we proceed as follows. In a first step, we use only few evaluations of $\lambda^{(\varepsilon)}(\eta)$ on the whole range of possible η to fit a first polynomial $\tilde{\lambda}^{(\varepsilon)}(\eta)$ and maximize the likelihood function. We denote the maximum likelihood estimate of this first step by $\hat{\eta}_1$. In the next step we use more evaluations of $\lambda^{(\varepsilon)}(\eta)$ in a region around $\hat{\eta}_1$ to fit the second polynomial. Again, we maximize the likelihood function in this region where we replace $\lambda^{(\varepsilon)}(\eta)$ by the second, more accurate polynomial. This procedure may be iterated until the desired accuracy is reached. In a simulation study, we found that a three step approach performs well in terms of accuracy and speed. This procedure is illustrated in Figure 6.6, where the distributional copula under consideration is a Gauss copula, however, any other distributional copula that satisfies Assumption 6.7 could also be used. In the first step we evaluate the likelihood function $\lambda^{(\varepsilon)}(\eta)$ for $\eta \in \{-0.99, -0.6, -0.2, 0.2, 0.6, 0.99\}$ to get a rough approximation of $\lambda^{(\varepsilon)}(\eta)$ on the whole range of possible dependence parameters η. The evaluations of the likelihood function are plotted as dots and the fitted polynomial of order two is visualized as the solid line on the left in Figure 6.6. The optimization of the approximated likelihood function (remember that we replace $\lambda^{(\varepsilon)}(\eta)$ by $\tilde{\lambda}^{(\varepsilon)}(\eta)$) gives an optimal value of $\hat{\eta}_1 = 0.27$. Since we use an approximated likelihood function, we do not believe that $\hat{\eta}_1$ is the true optimal parameter value, but we know that the optimal value is somewhere close to $\hat{\eta}_1$. Therefore, we calculate $\lambda^{(\varepsilon)}(\eta)$ for more values of η in the interval $[-0.13, 0.67]$ around $\hat{\eta}_1$ and fit a new polynomial of order one. This is illustrated in the second plot of Figure 6.6. We use the fitted polynomial to approximate the likelihood function in the interval $[-0.13, 0.67]$ and the maximization routine gives the optimal value $\hat{\eta}_2 = 0.31$. We repeat this procedure a third time on the smaller interval $[0.17, 0.45]$ around $\hat{\eta}_2$ and get the final optimal value $\hat{\eta} = \hat{\eta}_3 = 0.30$ in this maximum likelihood approach.

Scenario	Clayton Parameters θ	Gaussian Parameters ρ
High dependence (H)	5	0.8
Medium dependence (M)	2	0.3
Low dependence (L)	1	-0.2

Table 6.1: Parameters of the PLCC for scenarios H, M and L.

6.3 Simulation Study

In order to evaluate the estimators, we conduct a simulation study with a 5-dimensional PLCC. To make the results comparable, all marginal Lévy processes are chosen to be stable Lévy processes with parameters $\alpha = 0.5$ and $\beta = 1$. All bivariate Lévy copulas in the first tree are Clayton-Lévy copulas, see Example 5.21, and all distributional copulas in the higher trees are bivariate Gaussian copulas, see Example 2.8.

We analyze three different scenarios of dependence structures: high dependence (H), medium dependence (M) and low dependence (L). In scenarios H and M, we choose a D-vine structure for the PLCC and in scenario L a C-vine. The D-vine structure refers to a structure where all dimensions in the lowest tree form a line and are each connected to the nearest neighbors, whereas the dimensions in a C-vine structure are connected to only one central dimension (see, e.g., Aas *et al.* (2009)). Within a scenario, all Clayton-Lévy copulas have the same parameter θ and all Gaussian copulas have the same parameter ρ. The parameter values are summarized in Table 6.1.

For each scenario, we simulate a realization of a 5-dimensional Lévy process over a time horizon $[0, T]$. We then estimate the parameters of the marginal processes and the dependence parameters from the simulated data using our estimation approach. We choose two different thresholds $\varepsilon = 10^{-4}$ and $\varepsilon = 10^{-6}$ for jump sizes we can observe, i.e., we neglect jumps smaller than $\varepsilon = 10^{-4}$ or $\varepsilon = 10^{-6}$, respectively. Each simulation/estimation step is repeated 1000 times. The estimation results are reported in Tables 6.2, 6.3, and 6.4. We give the true values of the parameters, the mean of the estimates of the 1000 repetitions and resulting estimates for bias and root mean square error (RMSE). Since the parameters in the different trees rely on different numbers of observations (the higher the tree, the more dimensions have to exceed the threshold at the same time), we also report the mean numbers of available jumps per tree. As the marginal parameters influence the estimation procedure of the dependence parameters, we conduct another simulation study where we do not estimate the marginal parameters. In this setting, the parameters of the univariate marginal processes are taken from the simulation setup. The estimation results are given in Table 6.5. Comparing the tables, we see that the lower threshold leads to a higher number of jumps. We also find that weaker dependence leads to less co-jumps available for the estimation of higher trees than a stronger dependence. The impact of the estimation error in the marginal parameters on the estimators for the dependence structure is rather small, at least in our estimation setting. We observe that the estimator for the dependence structure in Tree 4 of the

	Tree	# Jumps	True Value	Mean	Bias	RMSE
High Dep.	1	870.61	5	5.0038	$3.78 \cdot 10^{-3}$	$2.33 \cdot 10^{-1}$
	2	833.51	0.8	0.7987	$-1.28 \cdot 10^{-3}$	$1.33 \cdot 10^{-2}$
	3	814.39	0.8	0.7980	$-1.97 \cdot 10^{-3}$	$1.34 \cdot 10^{-2}$
	4	798.46	0.8	0.7890	$-1.10 \cdot 10^{-2}$	$2.19 \cdot 10^{-2}$
Med. Dep.	1	707.18	2	2.0010	$1.02 \cdot 10^{-3}$	$9.65 \cdot 10^{-2}$
	2	573.56	0.3	0.2983	$-1.67 \cdot 10^{-3}$	$4.58 \cdot 10^{-2}$
	3	498.45	0.3	0.2983	$-1.72 \cdot 10^{-3}$	$4.97 \cdot 10^{-2}$
	4	451.69	0.3	0.3001	$1.31 \cdot 10^{-4}$	$5.11 \cdot 10^{-2}$
Low Dep.	1	500.10	1	1.0016	$1.63 \cdot 10^{-3}$	$4.46 \cdot 10^{-2}$
	2	267.36	-0.2	-0.1987	$1.31 \cdot 10^{-3}$	$4.98 \cdot 10^{-2}$
	3	163.22	-0.2	-0.1992	$7.96 \cdot 10^{-4}$	$7.12 \cdot 10^{-2}$
	4	113.91	-0.2	-0.2004	$-3.76 \cdot 10^{-4}$	$9.50 \cdot 10^{-2}$

Table 6.2: Results for a time horizon T=1 and a threshold $\varepsilon = 10^{-6}$ for three scenarios from low dependence to high dependence. The columns refer to the number of jumps used in the estimation of parameters within a certain tree, the true value of the parameters, the mean of the estimated parameters, estimated bias and RMSE from 1000 Monte Carlo repetitions. The first three trees contain more than one dependence function and we report the mean values of the estimators in these cases.

high dependence structure is slightly more symmetric if the marginal parameters are known, see Figure 6.7 and 6.13. In all cases, the bias is very small. We find, however, that the RMSE is affected by the number of jumps available in certain trees. Comparing the RMSEs in the second, third and fourth tree, the RMSE increases with decreasing number of jumps. The RMSE in the first tree is not directly comparable to the RMSEs in the subsequent trees since in the first tree, we estimate the parameter of a bivariate Lévy copula instead of a distributional copula, and therefore, we estimate a different parameter in a different parameter space. This effect is illustrated in Figures 6.7, 6.11, 6.12, and 6.13 in terms of histograms of the estimates. Overall the results of the simulation study are very satisfying. The simulation and estimation algorithms work well within the simulation study in terms of accuracy. Furthermore, the estimators are well distributed around the true value.

6.4 Non-Gaussian Ornstein-Uhlenbeck Processes

In the previous sections, we have shown how to use the pair-Lévy copula construction to create dependent Lévy processes in arbitrary dimensions. Here, we discuss how to transfer this concept to more advanced models which use Lévy processes as building

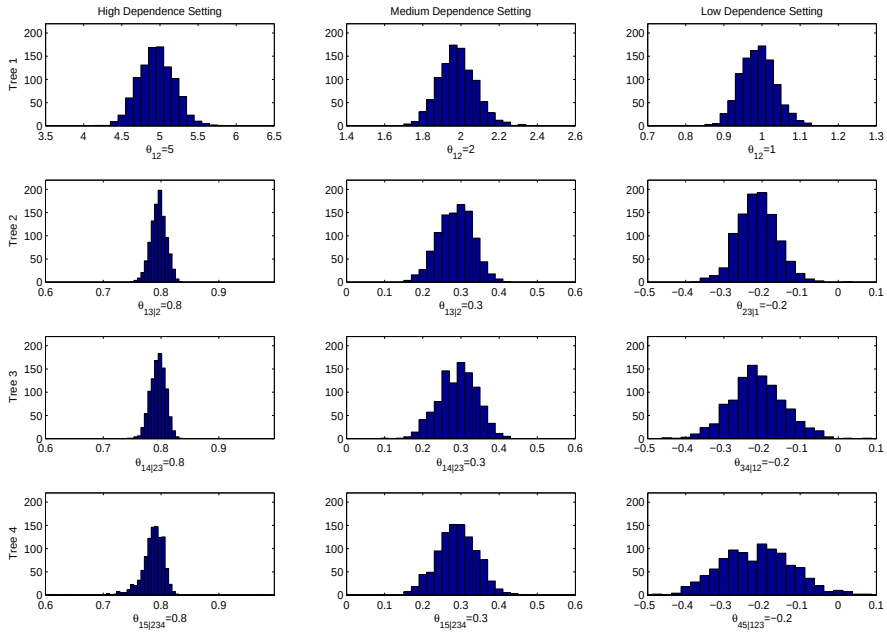


Figure 6.7: Histograms of the estimation results for a time horizon T=1 and a threshold $\varepsilon = 10^{-6}$. Each column refers to one scenario, the rows refer to the estimated parameters in the first to fourth tree.

blocks. That is, we show how to apply the pair-Lévy copula construction to model the dependence of stochastic processes that do not necessarily satisfy the conditions of a Lévy process but are more appropriate in certain applications. Important examples of such processes are Ornstein-Uhlenbeck processes that are driven by Lévy subordinators. Barndorff-Nielsen and Shephard (2001) introduce these processes as the solution to a stochastic differential equation to model the volatility of financial time series, and they already address a multivariate extension of their model. Here, we show how the flexibility and tractability of the pair-Lévy copula concept can be transferred to this highly relevant area of multivariate financial and econometric research. Before we start with the multivariate case, we recall the univariate non-Gaussian Ornstein-Uhlenbeck model.

6.4.1 Univariate Ornstein-Uhlenbeck Processes

We follow Brockwell *et al.* (2007) and introduce Ornstein-Uhlenbeck processes that are driven by univariate, second order Lévy subordinators. That is, Lévy subordinators with $E(L_1^2) < \infty$. To avoid identifiability problems, we suppose that the Lévy subordinator $(L_t)_{t\in\mathbb{R}_+}$ is scaled, such that $\mathrm{Var}(L_t) = t$. The univariate Ornstein-Uhlenbeck processes, which we consider here, are stationary solutions to the stochastic differential equation

$$dy_t = -\alpha\ y_t + \sigma\ dL_t, \tag{6.6}$$

where $\alpha, \sigma > 0$ and $(L_t)_{t\in\mathbb{R}_+}$ is a second order Lévy subordinator. We denote the process $(L_t)_{t\in\mathbb{R}_+}$ as the background driving Lévy process of $(y_t)_{t\in\mathbb{R}_+}$. The stochastic process, defined by

$$y_t = \exp(-\alpha t)y_0 + \sigma \int_0^t \exp(-\alpha(t-u))\, dL_u, \tag{6.7}$$

is a stationary solution to the stochastic differential equation (6.6) if y_0 is independent of $(L_t)_{t\in\mathbb{R}_+}$ and has the distribution of $\sigma \int_0^\infty e^{-\alpha u} dL_u$, see Brockwell *et al.* (2007). Jumps in the background driving Lévy process $(L_t)_{t\in\mathbb{R}_+}$ result in jumps in $(y_t)_{t\in\mathbb{R}_+}$. After a jump of $(y_t)_{t\in\mathbb{R}_+}$, the process decreases exponentially until the next jump in the Lévy subordinator causes again a jump in the Ornstein-Uhlenbeck process. This is illustrated in Figure 6.8, where we see an Ornstein-Uhlenbeck process with parameters $\alpha = 1/2$ and $\sigma = 15$. This process is driven by a gamma process, see Example 5.13, with parameters $c = 1/2$ and $\lambda = \sqrt{1/2}$. Figure 6.8 visualizes clearly the connection between the Ornstein-Uhlenbeck process and the background driving Lévy process and shows how the jumps in the background driving Lévy process $(L_t)_{t\in\mathbb{R}_+}$ carry over to the process $(y_t)_{t\in\mathbb{R}_+}$. For example, the large jumps of $(L_t)_{t\in\mathbb{R}_+}$ around $t = 9.5$ and $t = 15.1$ cause the large jumps of $(y_t)_{t\in\mathbb{R}_+}$ at these points in time. The time interval $[15.5, 17.5]$ is particularly illustrative for the path behavior of these processes. All jumps of the background driving Lévy process in this interval are very small, causing an exponential decline of the Ornstein-Uhlenbeck process within this interval.

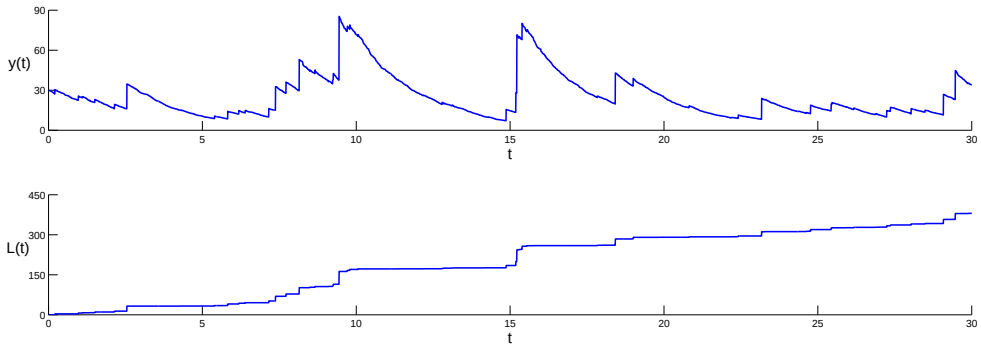


Figure 6.8: This figure shows a realization of an Ornstein-Uhlenbeck process in the upper plot. The lower plot shows the corresponding background driving Lévy process.

Please note that the condition on the background driving Lévy process for the existence of a stationary solution to Equation (6.6) can be relaxed from a second order Lévy subordinator to $E(0 \vee \log |L_1|) < \infty$, see Brockwell and Lindner (2012, Example 1).

6.4.2 Multivariate Ornstein-Uhlenbeck Processes

The univariate case illustrates the effect of the jumps in the background driving Lévy process on the Ornstein-Uhlenbeck process. This makes dependence modeling of the dimensions of a multivariate Ornstein-Uhlenbeck process straightforward. The idea is to use different univariate Ornstein-Uhlenbeck processes and specify the dependence in these processes by modeling the dependence of the background driving Lévy process with a Lévy copula. Therefore, we can apply the theory of pair-Lévy copula constructions to model the dependence structure of high-dimensional Ornstein-Uhlenbeck processes. We follow the idea of Barndorff-Nielsen and Shephard (2001, Section 6.4) and define the multivariate Ornstein-Uhlenbeck process $(y_t)_{t\in\mathbb{R}_+}$ in $\mathbb{R}_+^d$ by the following equations

$$y_t^j = \exp(-\alpha_j t) y_0^j + \sigma_j \int_0^t \exp(-\alpha_j (t-u))\, dL_u^j, \quad \text{for } j = 1, \ldots, d, \tag{6.8}$$

where $(L_t)_{t\in\mathbb{R}_+}$ is a multivariate subordinator, see Definition 5.11, and the marginal processes satisfy $\mathrm{Var}(L_t^j) = t$, for $j = 1, \ldots, d$. This procedure is illustrated in Figure 6.9. There we see a realization of a 3-dimensional Ornstein-Uhlenbeck process. The parameters of the univariate Ornstein-Uhlenbeck processes are $\alpha_j = 1/2$ and $\sigma_j = 15$ for all three margins ($j = 1, 2, 3$). The univariate Lévy subordinators are all gamma processes with parameters $c = 1/2$ and $\lambda = \sqrt{1/2}$. We use the same marginal processes for all three time series, since our focus is on dependence modeling. The dependence structure of the 3-dimensional Lévy process is modeled via the background driving Lévy process by a 3-dimensional pair-Lévy copula construction as in Figure 6.1, where we use

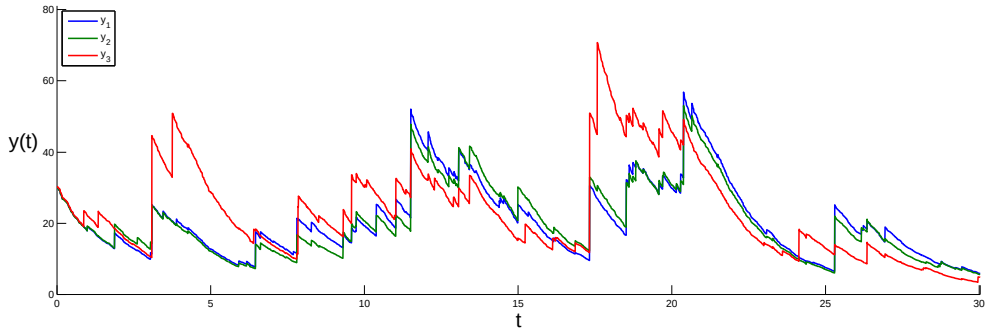


Figure 6.9: This figure shows a realization of a 3-dimensional Ornstein-Uhlenbeck process, where the dependence structure in the background driving Lévy process is modeled by a PLCC.

a bivariate Clayton-Lévy copula with parameter $\theta = 5$ for $\mathfrak{C}_{1,2}$ and a bivariate Clayton-Lévy copula with parameter $\theta = 1$ for $\mathfrak{C}_{2,3}$. The bivariate distributional copula $C_{1,3|2}$ is a Gaussian copula with parameter $\rho = 0.2$. Due to the strong dependence between the first and second dimension and the weak dependence between the second and third, most of the dependence is already captured in the first tree of the pair-Lévy copula by the bivariate Lévy copulas. Figure 6.9 visualizes how the dependence structure of the background driving Lévy processes is carried over to the Ornstein-Uhlenbeck process. The strong dependence between the Lévy subordinator in the first and second dimension is directly visible in the path behavior of the first and the second dimension of the Ornstein-Uhlenbeck process. These two dimensions are almost identical. The weaker dependence between the third dimension and the first two in the background driving Lévy process is also visible in the Ornstein-Uhlenbeck paths. We observe joint jumps in Figure 6.9 in all three dimensions, however, the dependence in the jump sizes is much stronger between the first and second dimension, compared to the other bivariate margins. This effect is an immediate consequence of the jump dependence in the background driving Lévy process, and this example illustrates the potential of the pair-Lévy copula construction framework for modeling the dependence structure of high-dimensional Ornstein-Uhlenbeck processes. In this example, the univariate driving background subordinators are identical in law, however, due to the flexibility of the pair-Lévy copula construction, it is also possible to define multivariate Ornstein-Uhlenbeck models, where each marginal background driving Lévy process belongs to a different parametric family.

6.5 Overview of Further Applications

In this section, we give an outlook on four possible applications of the pair-Lévy copula construction framework. In the first two subsections, Lévy processes are used directly in the modeling procedures, whereas we apply the Ornstein-Uhlenbeck processes of Section 6.4 in the last two subsections.

6.5.1 Operational Risk Modeling

A possible application of the pair-Lévy copula construction framework is operational risk modeling. For an introduction to the topic, we refer to McNeil *et al.* (2005) and Panjer (2006). The Basel II accord requires banks to implement a sound operational risk management (Basel Committee on Banking Supervision, 2004), and under the advanced measurement approach in the Basel II framework, banks are allowed to develop their own model that accounts for their operational risk exposure. Within this framework the banks are encouraged to classify their losses to 8 business lines and 7 event types and to model the dependencies between these 56 risk cells. Böcker and Klüppelberg (2008, 2010) model the losses in the single cells by a compound Poisson process and suggest to apply the Lévy copula concept for dependence modeling. In order to implement such a theoretical risk modeling approach in a bank, tractable and flexible dependence concepts for 56 dimensions are needed. Up to now, the pair-Lévy copula construction is the only Lévy copula framework that offers these properties. Therefore, the pair-Lévy copula construction approach permits the implementation of Lévy copula modeling within the operational risk framework.

6.5.2 Subordination of Lévy Processes

A different area of research, where the pair-Lévy copula construction can be applied to, is subordination of Lévy processes. Please note that, in order to facilitate the reading, we use the notation $(X(t))_{t\in\mathbb{R}_+}$ instead of $(X_t)_{t\in\mathbb{R}_+}$ for stochastic processes in this subsection. In the univariate case, subordination is a popular way to define new families of Lévy processes. Therefore, one combines an arbitrary univariate Lévy process $(L(t))_{t\in\mathbb{R}_+}$ with an independent univariate Lévy subordinator $(T(t))_{t\in\mathbb{R}_+}$ by

$$X(t) = L(T(t)).$$

This procedure defines a new Lévy process $(X(t))_{t\in\mathbb{R}_+}$ and is extensively studied, e.g., in Sato (1999). Barndorff-Nielsen *et al.* (2001b) extend this procedure to the multivariate case, where they set $(L(t))_{t\in\mathbb{R}_+} = (L^1(t), \dots, L^d(t))_{t\in\mathbb{R}_+}$ to be a Lévy process consisting of d independent marginal processes. Furthermore, they set $(T(t))_{t\in\mathbb{R}_+}$ to be a d-dimensional Lévy subordinator and define

$$X(t) = (L^1(T^1(t)), \dots, L^d(T^d(t))).$$

Again, this defines a d-dimensional Lévy process. Since this construction method uses multivariate subordinators, we can apply the pair-Lévy copula construction to model the dependence between the dimensions of the process $(T(t))_{t\in\mathbb{R}_+}$. Thus, we use the pair-Lévy copula construction to introduce a dependence structure implicitly to $(X(t))_{t\in\mathbb{R}_+}$.

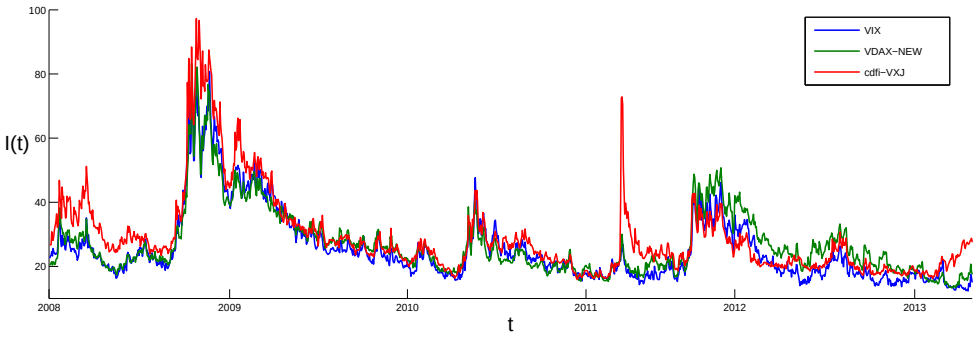


Figure 6.10: The plot shows the path of three volatility indices between January 2008 and March 2013. The blue line shows a volatility measure for the S&P 500 index (VIX), the green line is a measure for the volatility of the DAX (VDAX-NEW), and the red line visualizes a volatility process connected to the NIKKEI 225 (csfi-VXJ).

6.5.3 Modeling of Multivariate Volatility Indices

A further possible application of the pair-Lévy copula construction is dependence modeling for volatility indices. These indices exist for a broad variety of different markets and are used as underlyings for financial derivatives, see, e.g., Rhoads (2011). In Figure 6.10 we show three different volatility indices, which are all based on well-known equity indices. The volatility indices in the chart are the Chicago Board Options Exchange market volatility index (VIX) on the S&P 500, a volatility index calculated by the Deutsche Börse on the DAX (VDAX-NEW), and a volatility index for the Nikkei 225 (csfi-VXJ) provided by the Center for the Study of Finance and Insurance at Osaka University. The data is provided directly by the respective institutions. Distinct features of these times series are the sudden jumps and the exponential decline, as well as the strong dependence between the dimensions. These features can be well captured by multivariate Ornstein-Uhlenbeck processes, where the background driving Lévy process is a PLCC subordinator.

6.5.4 Stochastic Volatility Modeling

Multivariate stochastic volatility modeling is another research area, where we can apply the pair-Lévy copula construction. The idea in the univariate case goes back to Barndorff-Nielsen and Shephard (2001), where they already mention a possible multivariate generalization of their stochastic volatility model. The multivariate model, which we use here, is defined in Barndorff-Nielsen *et al.* (2002) by the stochastic differential equation

$$dx_t = (\mu + \Sigma_t \beta)dt + \Sigma_t^{1/2} dW_t + \rho dL_t, \tag{6.9}$$

where $(W_t)_{t\in\mathbb{R}_+}$ is a d-dimensional standard Brownian motion, $(\Sigma_t)_{t\in\mathbb{R}_+}$ is a $(d \times d)$-dimensional time varying covariance matrix, $(L_t)_{t\in\mathbb{R}_+}$ is a $(d+k)$-dimensional Lévy subordinator with independent marginal processes, $\mu, \beta \in \mathbb{R}^d$, and $\rho \in \mathbb{R}^{d\times(d+k)}$. Barndorff-Nielsen *et al.* (2002) model the dependence in the volatility processes $(\Sigma_t)_{t\in\mathbb{R}_+}$ by a factor model. That is, they define an Ornstein-Uhlenbeck process $(y_t)_{t\in\mathbb{R}_+}$ on $\mathbb{R}_+^{(d+k)}$ driven by the $(d+k)$-dimensional Lévy subordinator $(L_t)_{t\in\mathbb{R}_+}$, i.e., they construct $d+k$ independent univariate Ornstein-Uhlenbeck processes. The first d Ornstein-Uhlenbeck processes are responsible for the idiosyncratic properties of the volatility model and the subsequent k processes represent the common factors. That is, they define two diagonal matrices

$$I_t = \operatorname{diag}(y_t^1, \ldots, y_t^d),$$
$$J_t = \operatorname{diag}(y_t^{d+1}, \ldots, y_t^{d+k}),$$

and a factor loadings matrix $B \in \mathbb{R}^{d\times k}$. Altogether, the stochastic volatility matrix process is defined by

$$\Sigma_t = I_t + BJ_tB'. \tag{6.10}$$

The term ρdL_t in Equation (6.9) accounts for the leverage effect, which is often observed in financial data, see, e.g., Cont (2001). Note that the Lévy subordinator $(L_t)_{t\in\mathbb{R}_+}$ appears not only in the leverage term, but also in the definition of the $(d+k)$-dimensional Ornstein-Uhlenbeck process $(y_t)_{t\in\mathbb{R}_+}$, which is used to define the volatility processes $(\Sigma_t)_{t\in\mathbb{R}_+}$. Therefore, large jumps in the Lévy process cause a jump in the volatility process and also influence the stock price process $(x_t)_{t\in\mathbb{R}_+}$ directly. Thus, this model is capable to account for the leverage effect.

The factor approach in Equation (6.10) is not the only possibility to model the dependence structure in multivariate stochastic volatility models. A different framework for dependence modeling is already suggested in Barndorff-Nielsen and Shephard (2001). They propose to generalize such a model by allowing for dependent Ornstein-Uhlenbeck processes. As we have shown in Subsection 6.4.2, the pair-Lévy copula construction is particularly suitable for modeling the dependence of these volatility processes in higher dimensions. This framework also offers the possibility to simplify the multivariate stochastic volatility model in Equation (6.9) decisively. Since we do not need to rely on factors anymore, it is possible to use d instead of $d+k$ Ornstein-Uhlenbeck processes. Moreover, Σ can be defined as

$$\Sigma_t = I_t^{1/2} B I_t^{1/2}, \tag{6.11}$$

where $I_t^{1/2} = \operatorname{diag}\left(\sqrt{y_t^1}, \ldots, \sqrt{y_t^d}\right)$ and B is a constant correlation matrix. Furthermore, we choose $\rho \in \mathbb{R}^d$ instead of $\mathbb{R}^{d\times(d+k)}$ to account for the leverage effect. Note that even if we select a constant B, rolling window estimates of the correlation coefficient between the dimensions of a realization of $(x_t)_{t\in\mathbb{R}_+}$ on a discrete time grid may show time varying correlations. This is due to the dependence of the driving background

Lévy process and the leverage effect. If the dependence between the dimensions in $(L_t)_{t\in\mathbb{R}_+}$ is strong, exceptional large jumps in one dimension tend to occur simultaneously with exceptional large jumps in the other dimensions. Allowing for the leverage effect, i.e. setting $\rho \neq 0$ componentwise, results in simultaneous joint jumps in the different dimensions of $(x_t)_{t\in\mathbb{R}_+}$. Therefore, this model is able to account for an increase in the absolute value of the correlation in times of high volatility.

A further extension of the multivariate stochastic volatility model is to use non-negative CARMA processes (Brockwell *et al.*, 2011) instead of Ornstein-Uhlenbeck processes to describe the volatility process. A multivariate generalization of the CARMA model can be done analogously to the Ornstein-Uhlenbeck framework via Lévy copulas.

6.6 Conclusion

Lévy copulas describe the jump dependence of Lévy processes in a multivariate setting. For more than two dimensions, however, known parametric Lévy copulas are very restrictive. In this chapter, we develop a multidimensional pair construction of Lévy copulas from 2-dimensional dependence functions, which are either Lévy copulas or distributional copulas. The resulting parametric Lévy copula is very flexible, since every regular vine and every bivariate (Lévy) copula can be used in the PLCC. Therefore, the pair-Lévy copula construction opens the way to flexible high-dimensional continuous-time modeling. In a 5-dimensional simulation study, we evaluate the proposed pair-Lévy copula concept. Overall, the results of this simulation study are very satisfying. In particular, the bias of the estimators is small throughout all the trees in the pair-Lévy copula construction. The pair-Lévy copula concept is not only theoretically appealing, it is also important in numerous applications.

6.A Explicit 3-Dimensional Clayton PLCC

Here, we present a special case, where we can give an explicit formula for a 3-dimensional pair-Lévy copula. We use the setting of Example 6.5 and set the Lévy copulas $\mathfrak{C}_{1,2}$ and $\mathfrak{C}_{2,3}$ to be Clayton-Lévy copulas (5.9) with the same parameter θ. Furthermore we set $C_{1,3|2}$ to be a bivariate distributional Clayton copula (2.7) with parameter $\tilde{\theta} > 0$. This simplifies the formula for this distributional copula to

$$C(u,v) = (u_1^{-\tilde{\theta}} + u_2^{-\tilde{\theta}} - 1)^{-\frac{1}{\tilde{\theta}}},$$

where we set the parameter $\tilde{\theta} = \frac{\theta}{\theta+1}$. Using Proposition 6.9, we get

$$F_{1|u_2}(u_1) = \frac{\partial \mathfrak{C}_{1,2}(u_1,u_2)}{\partial u_2} = (u_1^{-\theta} + u_2^{-\theta})^{-\frac{\theta+1}{\theta}} u_2^{-(\theta+1)},$$

and

$$F_{3|u_2}(u_3) = \frac{\partial \mathfrak{C}_{2,3}(u_2,u_3)}{\partial u_2} = (u_3^{-\theta} + u_2^{-\theta})^{-\frac{\theta+1}{\theta}} u_2^{-(\theta+1)}.$$

Finally, we construct a 3-dimensional Lévy copula as suggested in Equation (6.2)

$$\begin{aligned}
\mathfrak{C}_{1,2,3}(u_1,u_2,u_3) &= \int_{[0,u_2]} C_{1,3|2}(F_{1|z_2}(u_1), F_{3|z_2}(u_3))dz_2 \\
&= \int_{[0,u_2]} \Bigg(\left((u_1^{-\theta} + z_2^{-\theta})^{-\frac{\theta+1}{\theta}} z_2^{-(\theta+1)}\right)^{-\frac{\theta}{\theta+1}} \\
&+ \left((u_3^{-\theta} + z_2^{-\theta})^{-\frac{\theta+1}{\theta}} z_2^{-(\theta+1)}\right)^{-\frac{\theta}{\theta+1}} - 1\Bigg)^{-\frac{\theta+1}{\theta}} dz_2 \\
&= \int_{[0,u_2]} \left((u_1^{-\theta} + z_2^{-\theta})z_2^{\theta} + (u_3^{-\theta} + z_2^{-\theta})z_2^{\theta} - 1\right)^{-\frac{\theta+1}{\theta}} dz_2 \\
&= \int_{[0,u_2]} \left(u_1^{-\theta} z_2^{\theta} + 1 + u_3^{-\theta} z_2^{\theta}\right)^{-\frac{\theta+1}{\theta}} dz_2 \\
&= \int_{[0,u_2]} \left(u_1^{-\theta} + z_2^{-\theta} + u_3^{-\theta}\right)^{-\frac{\theta+1}{\theta}} z_2^{-(\theta+1)} dz_2 \\
&= \left(u_1^{-\theta} + u_2^{-\theta} + u_3^{-\theta}\right)^{-\frac{1}{\theta}}
\end{aligned}$$

and get a 3-dimensional Clayton-Lévy copula.

6.B Proofs

For readers convenience, we provide the more complicated proofs of this chapter in the appendix.

Proof of Theorem 6.4

For the proof of Theorem 6.4, we need a lemma which we state first.

Lemma 6.12 *Let μ be a positive Radon measure on $\mathbb{R}_+^d$, $f_1 : x \mapsto \xi_x^1$ and $f_2 : x \mapsto \xi_x^2$ μ-measurable measure-valued maps, where ξ_x^1 and ξ_x^2 are probability measures on $\mathbb{R}_+$ with corresponding distribution functions F_x^1 and F_x^2. Let C be a 2-dimensional distributional copula and let ξ_x^C be the probability measure defined by the distribution function $C(F_x^1, F_x^2)$ on $\mathbb{R}_+^2$. Then, the map $x \mapsto \xi_x^C$ is μ-measurable.*

Proof: By definition, the maps $x \mapsto \xi_x^1(B_1)$ and $x \mapsto \xi_x^2(B_2)$ are μ-measurable for any $B_1, B_2 \in \mathcal{B}(\mathbb{R}_+)$. This holds in particular for the intervals $[0, b] \in \mathcal{B}(\mathbb{R}_+)$. Therefore, the maps $x \mapsto F_x^1(b_1)$ and $x \mapsto F_x^2(b_2)$ are μ-measurable for any $b_1, b_2 \in \mathbb{R}_+$. By definition of ξ_x^C, we have

$$\xi_x^C(B) = C(F_x^1(b_1), F_x^2(b_2))$$

for any rectangle $B \in \{[0, b_1] \times [0, b_2] : b_1, b_2 \in \mathbb{R}_+\}$. Since C is a copula, it is continuous and therefore measurable. We get that $x \mapsto \xi_x^C(B)$ is a composition of μ-measurable functions and therefore μ-measurable for any rectangle $B \in \{[0, b] : b \in \mathbb{R}_+^2\}$. Now that we have shown that $x \mapsto \xi_x^C(B)$ is μ-measurable for any $B \in \{[0, b] : b \in \mathbb{R}_+^2\}$, we use the same argumentation as in the proof of Ambrosio *et al.* (2000, Proposition 2.26), to show that $x \mapsto \xi_x^C(B)$ is μ-measurable for any $B \in \mathcal{B}(\mathbb{R}_+^2)$. Note that the set of intervals $B \in \{[0, b] : b \in \mathbb{R}_+^2\}$ is closed under finite intersection, it is a generator of the σ-algebra $\mathcal{B}(\mathbb{R}_+^2)$, and there exists a sequence (B_h) of these intervals such that $\mathbb{R}_+^2 = \cup_h B_h$. Denote the family of Borel sets such that $x \mapsto \xi_x^C(B)$ is μ-measurable by $\mathcal{M}$. Obviously, $\mathcal{M} \supset \{[0, b] : b \in \mathbb{R}_+^2\}$. In order to use Ambrosio *et al.* (2000, Remark 1.9), we have to show that the following conditions hold:

(i) $(E_h) \in \mathcal{M}$, $E_h \uparrow E \Rightarrow E \in \mathcal{M}$,

(ii) E,F, $E \cup F \in \mathcal{M} \Rightarrow E \cap F \in \mathcal{M}$,

(iii) $E \in \mathcal{M} \Rightarrow \mathbb{R}_+^2 \backslash E \in \mathcal{M}$.

This is already shown in the first part in the proof of Ambrosio *et al.* (2000, Proposition 2.26). □

Now we are able to prove Theorem 6.4.

We show that the integral is well-defined in the first step. From Theorem 6.3 follows that $(u_2, \ldots, u_{d-1}) \mapsto \xi_{1|u_2,\ldots,u_{d-1}}$ is $\mu_{2,\ldots,d-1}$-measurable. By the definition of measure-valued maps, $(u_2, \ldots, u_{d-1}) \mapsto \xi_{1|u_2,\ldots,u_{d-1}}(B)$ is $\mu_{2,\ldots,d-1}$-measurable for any $B \in \mathcal{B}(\mathbb{R}_+)$ and especially for any $B \in \{[0, b] : b \in \mathbb{R}_+\}$. Therefore,

$$\xi_{1|u_2,\ldots,u_{d-1}}([0, b]) = F_{1|u_2,\ldots,u_{d-1}}(b)$$

is $\mu_{2,\ldots,d-1}$-measurable. With the same arguments, we see immediately that $F_{d|u_2,\ldots,u_{d-1}}(b)$ is $\mu_{2,\ldots,d-1}$-measurable for any $b \in \mathbb{R}_+$. Since every copula is continuous, we can use the

same arguments as in the proof of Lemma 6.12 to show that

$$(u_2, \dots, u_{d-1}) \mapsto C(F_{1|u_2,\dots,u_{d-1}}(u_1), F_{d|u_2,\dots,u_{d-1}}(u_d))$$

is $\mu_{2,\dots,d-1}$-measurable and that the integral is well-defined. To show that $\mathfrak{C}_{1,\dots,d}$ is indeed a Lévy copula, we have to check the properties of Tankov (2005, Definition 3.3). We start by showing that $\mathfrak{C}_{1,\dots,d}$ is d-increasing. In a first step, we show this property for any d-box B where all vertices lie in $\mathbb{R}_+^d$. For every $(u_2, \dots, u_{d-1}) \in \mathbb{R}_+^{d-2}$ let $\xi^C_{1,d|u_2,\dots,u_{d-1}}$ be the probability measure on $\mathbb{R}_+^2$ defined by the distribution function $C(F_{1|u_2,\dots,u_{d-1}}(u_1), F_{d|u_2,\dots,u_{d-1}}(u_d))$. With Lemma 6.12 we know that $(u_2, \dots, u_{d-1}) \mapsto \xi^C_{1,d|u_2,\dots,u_{d-1}}$ is $\mu_{2,\dots,d-1}$-measurable. By definition of $\mathfrak{C}_{1,\dots,d}$

$$\begin{aligned}\mathfrak{C}_{1,\dots,d}(u_1, \dots, u_d) &= \int_{[0,u_2]\times\dots\times[0,u_{d-1}]} C(F_{1|z_2,\dots,z_{d-1}}(u_1), F_{d|z_2,\dots,z_{d-1}}(u_d)) d\mu_{2,\dots,d-1}(z_2, \dots, z_{d-1}) \\ &= \int_{[0,u_2]\times\dots\times[0,u_{d-1}]} \left(\int_{[0,u_1]\times[0,u_d]} d\xi^C_u \right) d\mu_{2,\dots,d-1}(z_2, \dots, z_{d-1})\end{aligned}$$

holds, and therefore

$$\begin{aligned}\mathfrak{C}_{1,\dots,d}(u_1, \dots, u_d) &= \mu_{2,\dots,d-1} \otimes \xi^C_{1,d|u_2,\dots,u_{d-1}}([0, u_1] \times \dots \times [0, u_d]) \\ &= \mu_{1,\dots,d}([0, u_1] \times \dots \times [0, u_d]).\end{aligned}$$

Since $\mu_{2,\dots,d-1} \otimes \xi^C_{1,d|u_2,\dots,u_{d-1}}$ is a positive and well-defined measure,

$$V_{\mathfrak{C}_{1,\dots,d}}(B) = \mu_{2,\dots,d-1} \otimes \xi^C_u(B) \geq 0.$$

In the next step, we denote $u_I := \{u_i : i \in I\}$ and show that the limit in Equation (6.1) exists for any $I \subset \{1, \dots, d\}$ nonempty, $I \neq \{1, \dots, d\}$. First, suppose that $\{1, d\} \subset I$. Since $I \neq \{1, \dots, d\}$, we say w.l.o.g. that $\{2\} \notin I$. Since $\mathfrak{C}_{1,\dots,d}$ is non-decreasing in

every component, it suffices to show that

$$\begin{aligned}
& \lim_{u_I \to \infty} \mathfrak{C}_{1,\dots,d}(u_1, \dots, u_d) \\
= & \lim_{u_I \to \infty} \int_{[0,u_2]\times\dots\times[0,u_{d-1}]} C(F_{1|z_2,\dots,z_{d-1}}(u_1), F_{d|z_2,\dots,z_{d-1}}(u_d)) d\mu_{2,\dots,d-1}(z_2, \dots, z_{d-1}) \\
= & \lim_{u_{I\setminus\{1,d\}} \to \infty} \int_{[0,u_2]\times\dots\times[0,u_{d-1}]} d\mu_{2,\dots,d-1}(z_2, \dots, z_{d-1}) \\
= & \lim_{u_{I\setminus\{1,d\}} \to \infty} \int_{[0,u_2]\times\dots\times[0,u_{d-1}]} \int_{[0,\infty)} d\xi_{1|z_2,\dots,z_{d-1}} d\mu_{2,\dots,d-1}(z_2, \dots, z_{d-1}) \\
= & \lim_{u_{I\setminus\{1,d\}} \to \infty} \int_{[0\times\infty)\times[0,u_2]\times\dots\times[0,u_{d-1}]} d\mu_{1,\dots,d-1}(z_1, \dots, z_{d-1}) \\
= & \lim_{u_{I\setminus\{1,d\}} \to \infty} \mathfrak{C}_{1,\dots,d-1}(\infty, u_2, \dots, u_{d-1}) \\
\leq & \ \mathfrak{C}_{1,\dots,d-1}(\infty, u_2, \infty, \dots, \infty) = u_2
\end{aligned}$$

to prove that the limit exists. We use the dominated convergence theorem (stated, e.g., in Ambrosio *et al.* (2000, Theorem 1.21)) and the fact that for every distributional copula $C(u_1, u_2) \leq 1$ holds. For the inequality, we use the fact that Assumption 5.18 holds for the Lévy copula $\mathfrak{C}_{1,\dots,d-1}$. Now, suppose that at least one element of $\{1, d\}$ is not in I. W.l.o.g. $\{1\} \notin I$ then we have

$$\begin{aligned}
& \lim_{u_I \to \infty} \mathfrak{C}_{1,\dots,d}(u_1, \dots, u_d) \\
= & \lim_{u_I \to \infty} \int_{[0,u_2]\times\dots\times[0,u_{d-1}]} C(F_{1|z_2,\dots,z_{d-1}}(u_1), F_{d|z_2,\dots,z_{d-1}}(u_d)) d\mu_{2,\dots,d-1}(z_2, \dots, z_{d-1}) \\
\leq & \lim_{u_{I\setminus\{d\}} \to \infty} \int_{[0,u_2]\times\dots\times[0,u_{d-1}]} F_{1|z_2,\dots,z_{d-1}}(u_1) d\mu_{2,\dots,d-1}(z_2, \dots, z_{d-1}) \\
= & \lim_{u_{I\setminus\{d\}} \to \infty} \int_{[0,u_2]\times\dots\times[0,u_{d-1}]} \int_{[0,u_1]} d\xi_{1|z_2,\dots,z_{d-1}} d\mu_{2,\dots,d-1}(z_2, \dots, z_{d-1}) \\
= & \lim_{u_{I\setminus\{d\}} \to \infty} \int_{[0,u_1]\times[0,u_2]\times\dots\times[0,u_{d-1}]} d\mu_{1,\dots,d-1}(z_2, \dots, z_{d-1}) \\
\leq & \ \mathfrak{C}_{1,\dots,d-1}(u_1, \infty, \dots, \infty) = u_1.
\end{aligned}$$

Now that we have shown that the limit exists, it follows immediately that $\mathfrak{C}_{1,\dots,d}$ is d-increasing on $\overline{\mathbb{R}}_+^d$. To show that the Lévy copula $\mathfrak{C}_{1,\dots,d}$ has Lebesgue margins, we can again use the same equations as before and replace "$\leq$" by "$=$" since in this case $|I| = d - 1$ and therefore we can directly use Assumption 5.18. □

Proof of Proposition 6.10

Suppose that $F_{1|u_2,\dots,u_{d-1}}$ and $F_{d|u_2,\dots,u_{d-1}}$ are continuously differentiable. For any rectangle $B = ([0,u_1] \times \ldots \times [0,u_d])$, we get by Theorem 6.3

$$\int_{\mathbb{R}_+^d} \mathbb{1}_B(z_1,\dots,z_d) d\mu_{1,\dots,d}(z_1,\dots,z_d) = \int_{[0,u_1]\times\ldots\times[0,u_{d-1}]} F_{d|u_1,\dots,u_{d-1}}(u_d) d\mu_{1,\dots,d-1}(z_1,\dots,z_{d-1}).$$

By the definition of the pair-Lévy copula we see that

$$\begin{aligned}
&\int_{\mathbb{R}_+^d} \mathbb{1}_B(z_1,\dots,z_d) d\mu_{1,\dots,d}(z_1,\dots,z_d) \\
=& \int_{\mathbb{R}_+^{d-2}} \left(\int_{\mathbb{R}_+^2} \mathbb{1}_B(z_1,\dots,z_d) d\xi^C_{1,d|u_2,\dots,u_{d-1}} \right) d\mu_{2,\dots,d-1}(z_2,\dots,z_{d-1}) \\
=& \int_{[0,u_2]\times\ldots\times[0,u_{d-1}]} \left(C(F_{1|z_2,\dots,z_{d-1}}(u_1), F_{d|z_2,\dots,z_{d-1}}(u_d)) \right) d\mu_{2,\dots,d-1}(z_2,\dots,z_{d-1}) \\
=& \int_{[0,u_2]\times\ldots\times[0,u_{d-1}]} \\
& \left(\int_{[0,u_1]} \frac{\partial C(F_{1|z_2,\dots,z_{d-1}}(z_1), F_{d|z_2,\dots,z_{d-1}}(u_d))}{\partial F_{1|z_2,\dots,z_{d-1}}(z_1)} \frac{\partial F_{1|z_2,\dots,z_{d-1}}(z_1)}{\partial z_1} dz_1 \right) d\mu_{2,\dots,d-1}(z_2,\dots,z_{d-1}) \\
=& \int_{[0,u_2]\times\ldots\times[0,u_{d-1}]} \\
& \left(\int_{[0,u_1]} \frac{\partial C(F_{1|z_2,\dots,z_{d-1}}(z_1), F_{d|z_2,\dots,z_{d-1}}(u_d))}{\partial F_{1|z_2,\dots,z_{d-1}}(z_1)} d\xi_{1|z_2,\dots,z_{d-1}}(z_1) \right) d\mu_{2,\dots,d-1}(z_2,\dots,z_{d-1}) \\
=& \int_{[0,u_1]\times\ldots\times[0,u_{d-1}]} \frac{\partial C(F_{1|z_2,\dots,z_{d-1}}(z_1), F_{d|z_2,\dots,z_{d-1}}(u_d))}{\partial F_{1|z_2,\dots,z_{d-1}}(z_1)} d\mu_{1,\dots,d-1}(z_1,\dots,z_{d-1}),
\end{aligned}$$

and therefore

$$F_{d|u_1,\dots,u_{d-1}}(u_d) = \frac{\partial C(F_{1|u_2,\dots,u_{d-1}}(u_1), F_{d|u_2,\dots,u_{d-1}}(u_d))}{\partial F_{1|u_2,\dots,u_{d-1}}(u_1)}$$

holds $\mu_{1,\dots,d-1}$-almost everywhere. The fact that this result does not only hold for fixed values of u_d but for all $u_d \in \mathbb{R}_+$ is already shown in the proof of Tankov (2005, Lemma 4.2). Since $F_{1|u_2,\dots,u_{d-1}}$, $F_{d|u_2,\dots,u_{d-1}}$ are continuously differentiable and C is by Assumption 6.7 also continuously differentiable, we get immediately that $F_{d|u_1,\dots,u_{d-1}}$ is differentiable and

$$\frac{\partial F_{d|u_1,\dots,u_{d-1}}(u_d)}{\partial u_d} = \frac{\partial^2 C(F_{1|u_2,\dots,u_{d-1}}(u_1), F_{d|u_2,\dots,u_{d-1}}(u_d))}{\partial F_{1|u_2,\dots,u_{d-1}}(u_1) \partial F_{d|u_2,\dots,u_{d-1}}(u_d)} \frac{\partial F_{d|u_2,\dots,u_{d-1}}(u_d)}{\partial u_d}$$

is a composition of continuous functions and therefore continuous. Finally, all bivariate Lévy copulas are by Assumption 6.7 continuously differentiable and therefore, the proposition follows by complete induction. □

Proof of Proposition 6.11

This statement follows from the definition of the pair-Lévy copula construction, since

$$
\begin{aligned}
&\mathfrak{C}_{1,\dots,d}(u_1,\dots,u_d) = \\
&\qquad \int_{[0,u_2]\times\dots\times[0,u_{d-1}]} C(F_{1|z_2,\dots,z_{d-1}}(u_1), F_{d|z_2,\dots,z_{d-1}}(u_d)) d\mu_{2,\dots,d-1}(z_2,\dots,z_{d-1}) \\
&\qquad = \int_{[0,u_2]\times\dots\times[0,u_{d-1}]} \Bigg(\int_{[0,u_1]\times[0,u_d]} c(F_{1|z_2,\dots,z_{d-1}}(z_1), F_{d|z_2,\dots,z_{d-1}}(z_d)) \\
&\qquad \frac{\partial F_{1|z_2,\dots,z_{d-1}}(z_1)}{\partial z_1} \frac{\partial F_{d|z_2,\dots,z_{d-1}}(z_d)}{\partial z_d} d(z_1,z_d) \Bigg) d\mu_{2,\dots,d-1}(z_2,\dots,z_{d-1}) \\
&\qquad = \int_{[0,u_2]\times\dots\times[0,u_{d-1}]} \Bigg(\int_{[0,u_1]\times[0,u_d]} c(F_{1|z_2,\dots,z_{d-1}}(z_1), F_{d|z_2,\dots,z_{d-1}}(z_d)) \\
&\qquad \frac{\partial F_{1|z_2,\dots,z_{d-1}}(z_1)}{\partial z_1} \frac{\partial F_{d|z_2,\dots,z_{d-1}}(z_d)}{\partial z_d} d(z_1,z_d) \Bigg) \\
&\qquad f_{2,\dots,d-1}(z_2,\dots,z_{d-1}) d(z_2,\dots,z_{d-1}) \\
&\qquad = \int_{[0,u_1]\times\dots\times[0,u_d]} c(F_{1|z_2,\dots,z_{d-1}}(z_1), F_{d|z_2,\dots,z_{d-1}}(z_d)) \\
&\qquad \frac{\partial F_{1|z_2,\dots,z_{d-1}}(z_1)}{\partial z_1} \frac{\partial F_{d|z_2,\dots,z_{d-1}}(z_d)}{\partial z_d} f_{2,\dots,d-1}(z_2,\dots,z_{d-1}) d(z_1,\dots,z_d)
\end{aligned}
$$

as stated. □

6.C Results of the Simulation Study

Here, we present additional results from the simulation study with varying thresholds and time horizons. The tables are in the same format as Table 6.2. We also provide the corresponding histograms similar to Figure 6.7.

	Tree	# Jumps	True Value	Mean	Bias	RMSE
High Dep.	1	87.26	5	5.0403	$4.03 \cdot 10^{-2}$	$7.06 \cdot 10^{-1}$
	2	83.63	0.8	0.7933	$-6.76 \cdot 10^{-3}$	$4.61 \cdot 10^{-2}$
	3	81.69	0.8	0.7810	$-1.90 \cdot 10^{-2}$	$5.67 \cdot 10^{-2}$
	4	80.10	0.8	0.7086	$-9.14 \cdot 10^{-2}$	$1.46 \cdot 10^{-1}$
Med. Dep.	1	70.82	2	2.0312	$3.12 \cdot 10^{-2}$	$3.19 \cdot 10^{-1}$
	2	57.47	0.3	0.2970	$-3.00 \cdot 10^{-3}$	$1.50 \cdot 10^{-1}$
	3	50.00	0.3	0.2844	$-1.56 \cdot 10^{-2}$	$1.59 \cdot 10^{-1}$
	4	45.37	0.3	0.2797	$-2.03 \cdot 10^{-2}$	$1.63 \cdot 10^{-1}$
Low Dep.	1	50.21	1	1.0246	$2.46 \cdot 10^{-2}$	$1.55 \cdot 10^{-1}$
	2	26.88	-0.2	-0.2019	$-1.87 \cdot 10^{-3}$	$1.67 \cdot 10^{-1}$
	3	16.42	-0.2	-0.1859	$1.41 \cdot 10^{-2}$	$2.57 \cdot 10^{-1}$
	4	11.47	-0.2	-0.1378	$6.22 \cdot 10^{-2}$	$3.44 \cdot 10^{-1}$

Table 6.3: Results for a time horizon T=1 and a threshold $\varepsilon = 10^{-4}$ for three scenarios from low dependence to high dependence. The columns refer to the number of jumps used in the estimation of parameters within a certain tree, the true value of the parameters, the mean of the estimated parameters, estimated bias and RMSE from 1000 Monte Carlo repetitions. The first three trees contain more than one dependence function and we report the mean values of the estimators in these cases. Compared to Table 6.2, the higher threshold ε results in fewer observed jumps and in higher RMSE of the estimates.

	Tree	# Jumps	True Value	Mean	Bias	RMSE
High Dep.	1	870.89	5	4.9921	$-7.85 \cdot 10^{-3}$	$2.27 \cdot 10^{-1}$
	2	833.89	0.8	0.7995	$-4.79 \cdot 10^{-4}$	$1.33 \cdot 10^{-2}$
	3	814.79	0.8	0.7978	$-2.23 \cdot 10^{-3}$	$1.34 \cdot 10^{-2}$
	4	798.91	0.8	0.7890	$-1.10 \cdot 10^{-2}$	$2.21 \cdot 10^{-2}$
Med. Dep.	1	707.03	2	2.0027	$2.74 \cdot 10^{-3}$	$9.50 \cdot 10^{-2}$
	2	573.47	0.3	0.2994	$-6.18 \cdot 10^{-4}$	$4.87 \cdot 10^{-2}$
	3	498.36	0.3	0.2968	$-3.17 \cdot 10^{-3}$	$5.05 \cdot 10^{-2}$
	4	451.75	0.3	0.2999	$-1.46 \cdot 10^{-4}$	$5.18 \cdot 10^{-2}$
Low Dep.	1	500.98	1	1.0041	$4.14 \cdot 10^{-3}$	$4.50 \cdot 10^{-2}$
	2	268.03	-0.2	-0.1984	$1.56 \cdot 10^{-3}$	$5.18 \cdot 10^{-2}$
	3	163.82	-0.2	-0.1978	$2.16 \cdot 10^{-3}$	$7.11 \cdot 10^{-2}$
	4	114.17	-0.2	-0.1929	$7.09 \cdot 10^{-3}$	$9.36 \cdot 10^{-2}$

Table 6.4: Results for a time horizon $T = 10$ and a threshold $\varepsilon = 10^{-4}$ for three scenarios from low dependence to high dependence. The columns refer to the number of jumps used in the estimation of parameters within a certain tree, the true value of the parameters, the mean of the estimated parameters, estimated bias and RMSE from 1000 Monte Carlo repetitions. The first three trees contain more than one dependence function and we report the mean values of the estimators in these cases. The results are comparable to Table 6.2.

	Tree	# Jumps	True Value	Mean	Bias	RMSE
High Dep.	1	869.93	5	5.0084	$8.43 \cdot 10^{-3}$	$1.65 \cdot 10^{-1}$
	2	832.90	0.8	0.7991	$-9.13 \cdot 10^{-4}$	$1.30 \cdot 10^{-2}$
	3	813.85	0.8	0.7996	$-4.15 \cdot 10^{-4}$	$1.27 \cdot 10^{-2}$
	4	798.03	0.8	0.7984	$-1.64 \cdot 10^{-3}$	$1.27 \cdot 10^{-2}$
Med. Dep.	1	706.83	2	2.0020	$1.97 \cdot 10^{-3}$	$8.10 \cdot 10^{-2}$
	2	573.26	0.3	0.2980	$-1.98 \cdot 10^{-3}$	$4.64 \cdot 10^{-2}$
	3	498.19	0.3	0.2981	$-1.94 \cdot 10^{-3}$	$5.10 \cdot 10^{-2}$
	4	451.45	0.3	0.2972	$-2.75 \cdot 10^{-3}$	$5.16 \cdot 10^{-2}$
Low Dep.	1	500.08	1	1.0003	$2.56 \cdot 10^{-4}$	$4.59 \cdot 10^{-2}$
	2	267.83	-0.2	-0.1982	$1.83 \cdot 10^{-3}$	$5.22 \cdot 10^{-2}$
	3	163.83	-0.2	-0.1968	$3.24 \cdot 10^{-3}$	$7.36 \cdot 10^{-2}$
	4	114.23	-0.2	-0.1973	$2.67 \cdot 10^{-3}$	$9.67 \cdot 10^{-2}$

Table 6.5: Results for a time horizon $T = 1$ and a threshold $\varepsilon = 10^{-6}$ for three scenarios from low dependence to high dependence. In this simulation study we only estimate the dependence parameters. The marginal parameters are not estimated but taken from the simulation setup. The columns refer to the number of jumps used in the estimation of parameters within a certain tree, the true value of the parameters, the mean of the estimated parameters, estimated bias and RMSE from 1000 Monte Carlo repetitions. The first three trees contain more than one dependence function and we report the mean values of the estimators in these cases. The results are comparable to Table 6.2.

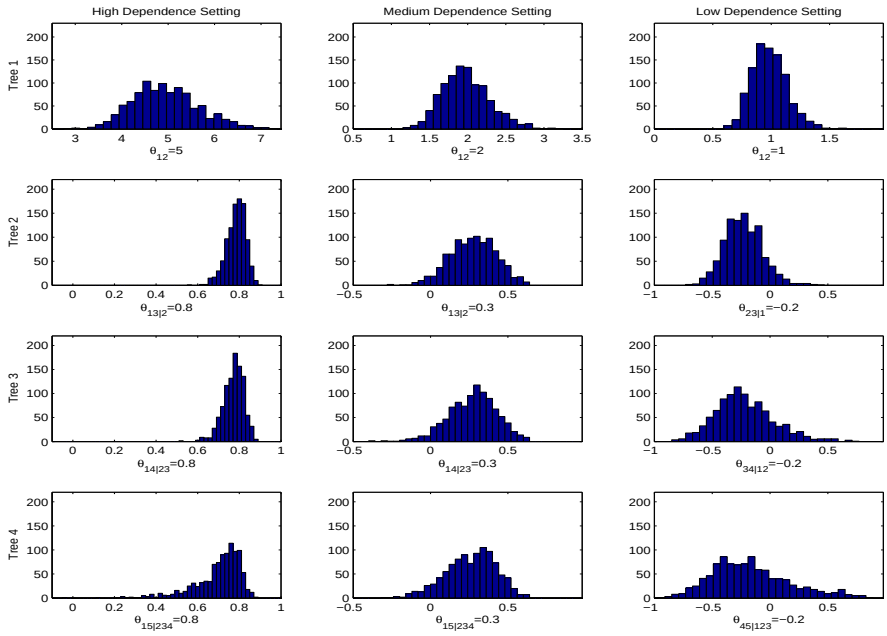


Figure 6.11: Histograms of the estimation results for a time horizon T=1 and a threshold $\varepsilon = 10^{-4}$. Each column refers to one scenario, the rows refer to the estimated parameters in the first to fourth tree. Compared to Figure 6.7, the higher threshold ε and fewer jumps results in a higher variation of the estimator.

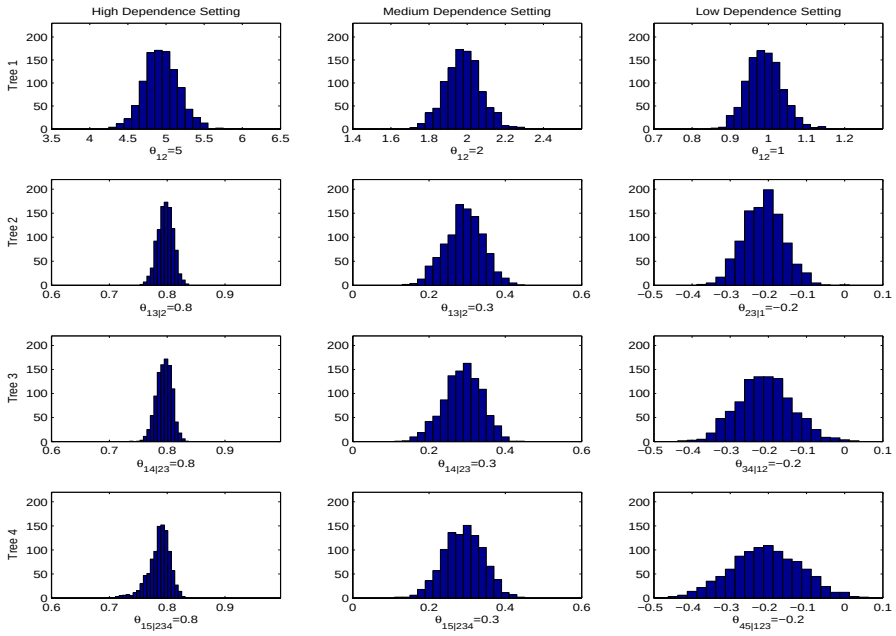


Figure 6.12: Histograms of the estimation results for a time horizon T=10 and a threshold $\varepsilon = 10^{-4}$. Each column refers to one scenario, the rows refer to the estimated parameters in the first to fourth tree. The results are comparable to Figure 6.7.

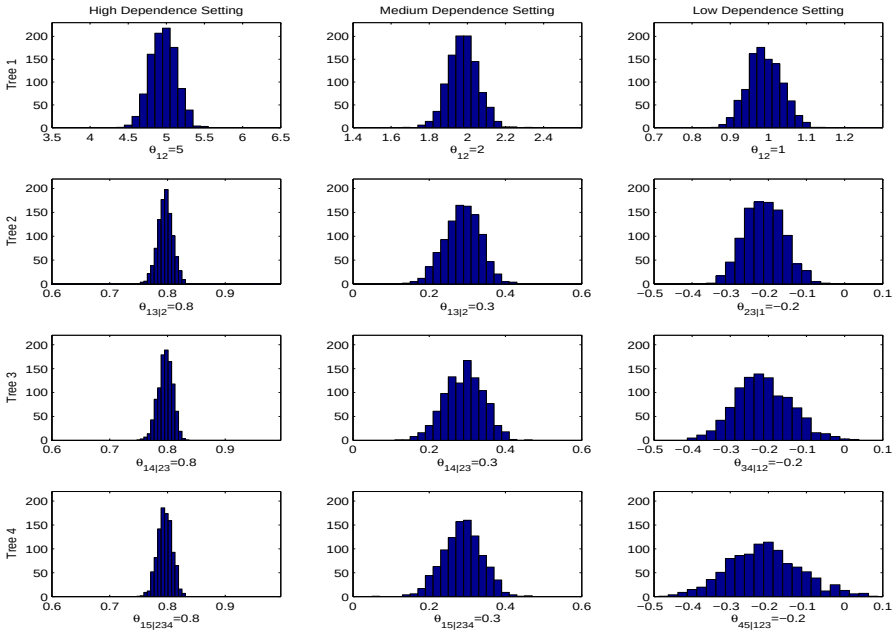


Figure 6.13: Histograms of the estimation results for a time horizon T=1 and a threshold $\varepsilon = 10^{-6}$. Each column refers to one scenario, the rows refer to the estimated parameters in the first to fourth tree. Here, the marginal parameters are known from the simulation setup, and the results are slightly more symmetric than in Figure 6.7, in particular for Tree 4 in the high dependence setting.

6.D Algorithms

In this section, we provide algorithms for the pair-Lévy copula concept. The structure of distributional pair copulas, discussed in Chapter 3, and pair-Lévy copulas can both be visualized by regular vines. We exploit these structural similarities and modify the algorithms, presented in Appendix 3.C, slightly so that we can apply them to pair-Lévy copulas. Firstly, we need to adapt the h-function (H-FUNC$(\cdot)$) to

$$h(x_i, x_j, \theta) = \begin{cases} \frac{\partial \mathfrak{C}_{i,j}(x_i,x_j,\theta)}{\partial x_j} & \text{if } m_{i,j} = \emptyset, \\ \frac{\partial C_{i,j|m_{i,j}}(x_i,x_j,\theta)}{\partial x_j} & \text{else,} \end{cases}$$

where we use the same notation as in Appendix 3.C. The inverse of the h-function (H-INVERSE$(\cdot)$, inverse with respect to the first variable) has to be modified similarly.

As discussed in Section 6.2.2, we use every available jump for the estimation procedure and transform it with the marginal tail integral to the so called pseudo-jumps $\Gamma^1, \dots, \Gamma^d$. Now, we can reuse Algorithm 4 with the adapted h-function to estimate a pair-Lévy copula from the pseudo-jumps. Note that in these Algorithms $c_{i,j}$ can either be the density of an ordinary copula or the density of a Lévy copula, depending on position in the vine structure.

The simulation Algorithm 8 is based on Theorem 6.8. Therefore, we create pseudo-jumps $\Gamma^1, \dots, \Gamma^d$ which we then transform, in a second step, with the inverse of the marginal tail integrals to the jumps of the Lévy process. As discussed in Section 6.2.1, we do not simulate single pseudo-jumps, but a sequence $(\Gamma_i^1, \dots, \Gamma_i^d)_{i=1,\dots,N}$ of N pseudo-jumps.

Algorithm 8 Simulate a Poisson point process from a pair-Lévy copula (R-Vine)

Input: Let $(\Gamma_l^1)_{l=1,\dots,N}$ be the jump times of a Poisson process with intensity 1 and N large. $(\omega_l^2,\dots,\omega_l^d)_{l=1,\dots,N}$ an independent sequence of i.i.d random variables with $\omega_l^i \sim U([0,1])$.

for $l = 1,\dots,N$ **do**
  $A \leftarrow \{1\}$
  for $i \leftarrow 2,\dots d$ **do**
    $j \leftarrow$ selectNextVariable(A)
    $\Gamma_l^j \leftarrow \omega_l^j$
    $B \leftarrow A$
    for $h \leftarrow 1,\dots i-1$ **do**
      $k \leftarrow$ selectConditionedIndex(j, B)
      $B \leftarrow B \backslash \{k\}$
      calculateF(Γ_l^k, B)
      $\Gamma_l^j \leftarrow$ hInverse$(\Gamma_l^j,$ calculateF$(\Gamma_l^k, B), \theta_{j,k})$
    end for
    $A \leftarrow A \cup \{j\}$
  end for
end for

Chapter 7

Summary

The overall topic of this dissertation is dependence modeling in $d > 2$ dimensions. In the first part of this dissertation, we summarize the copula concept to lay the theoretical foundation for dependence modeling. We review the different state of the art multivariate dependence modeling techniques, in particular, we discuss high-dimensional parametric copula models, like Archimedean, implicit, and hierarchical Archimedean copulas. Still, the main focus is on the pair-copula construction concept, which is a very flexible and tractable concept to transfer the advantages of bivariate copula modeling to higher dimensions. We contribute to this framework by providing a new representation for the structure of the pair-copula construction. This new representation is easy to interpret and particularly useful for the implementation of simulation and estimation algorithms. Furthermore, we develop a new parameter reduction technique that crucially reduces the number of parameters in the pair-copula construction approach. In an empirical study, we evaluate the performance of the different high-dimensional dependence models and show that a parameter reduced pair-copula is competitive in terms of complexity and forecast quality.

In the second part of this dissertation, we discuss continuous-time dependence modeling. Up to now, dependence modeling for Lévy processes focuses mainly on the Brownian motion. However, the Brownian motion has no jumps and is therefore not appropriate in many applications. The recent introduction of Lévy copulas provides a new concept for more flexible dependence models, since this mathematical tool allows to model the dependence in the jumps of multivariate Lévy processes. Unfortunately, the generalization of the existing parametric Lévy copula families to $d > 2$ dimensions is not flexible enough. Thus, we develop a new concept to construct adaptable high-dimensional Lévy copulas. The idea behind the pair-Lévy copula construction is similar to the pair-copula construction for random variables. However, the technical details are substantially more challenging since the Lévy copula is defined on $\overline{\mathbb{R}}_+^d$. In order to construct a pair-Lévy copula in d dimensions, we need to assemble bivariate Lévy copulas and bivariate distributional copulas. Simulation and estimation techniques are essential for the applicability of this new concept, thus, we develop a fast sequential simulation procedure and derive maximum likelihood estimators for pair-Lévy copulas. In a simulation study, we evaluate the performance of these simulation and estimation procedures. Furthermore, we give outlook on possible applications of this high-dimensional dependence concept.

Topics for Further Research

Finally, we want to point out some further research topics in dependence modeling for stochastic processes. Up to now, continuous-time dependence modeling for pure jump processes is not a very active research area but we hope that we have convinced the reader that this is an important and interesting part in dependence modeling. One topic for further research is the development of new bivariate parametric Lévy copula families. Some construction techniques for building new Lévy copulas are well-known, however, only few parametric models are explicitly defined and in use. Further bivariate Lévy copulas that provide a broad range of different dependence structures are needed. As we show in Chapter 6, these bivariate Lévy copulas can also be used in the multivariate setting of a pair-Lévy copula construction. The development of new multivariate Lévy copula families as benchmark models for the pair-Lévy copula construction is also an interesting topic. Moreover, the adjustment of the presented PLCC techniques to particular applications should be addressed in further research.

List of Tables

List of Figures

Bibliography

Aas, K. and Berg, D. (2009). Models for construction of multivariate dependence – a comparison study. *The European Journal of Finance*, **15** (7-8), 639–659.

— and — (2011). Modeling dependence between financial returns using pair-copula constructions. In D. Kurowicka and H. Joe (eds.), *Dependence Modeling: Vine Copula Handbook*, Singapore: World Scientific, pp. 305–328.

—, Czado, C., Frigessi, A. and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, **44** (2), 182 – 198.

Acar, E. F., Genest, C. and Nešlehová, J. (2012). Beyond simplified pair-copula constructions. *Journal of Multivariate Analysis*, **110** (0), 74 – 90.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19** (6), 716–723.

Almeida, C., Czado, C. and Manner, H. (2012). Modeling high dimensional time-varying dependence using D-vine SCAR models. *Submitted for publication*.

Ambrosio, L., Fusco, N. and Pallara, D. (2000). *Functions of Bounded Variation and Free Discontinuity Problems*. New York: Clarendon Press.

Barndorff-Nielsen, O. E., Nicolato, E. and Shephard, N. (2002). Some recent developments in stochastic volatility modelling. *Quantitative Finance*, **2** (1), 11–23.

—, Pedersen, J. and Sato, K.-I. (2001a). Multivariate subordination, self-decomposability and stability. *Advances in Applied Probability*, **33** (1), 160–187.

—, — and — (2001b). Multivariate subordination, self-decomposability and stability. *Advances in Applied Probability*, **33** (1), pp. 160–187.

— and Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **63** (2), 167–241.

Basawa, I. V. and Brockwell, P. J. (1978). Inference for gamma and stable processes. *Biometrika*, **65** (1), 129–133.

— and — (1980). A note on estimation for gamma and stable processes. *Biometrika*, **67** (1), 234–236.

Basel Committee on Banking Supervision (2004). *International Convergence of Capital Measurement and Capital Standards*. Bank for International Settlements, Basel.

Böcker, K. and Klüppelberg, C. (2008). Modeling and measuring multivariate operational risk with Lévy copulas. *The Journal of Operational Risk*, **3** (2), 3–27.

— and — (2010). Multivariate models for operational risk. *Quantitative Finance*, **10** (8), 855–869.

Bedford, T. and Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*, **32** (1–4), 245–268.

— and — (2002). Vines – a new graphical model for dependent random variables. *The Annals of Statistics*, **30** (4), 1031–1068.

Berg, D. (2009). Copula goodness-of-fit testing: an overview and power comparison. *The European Journal of Finance*, **15** (7-8), 675–701.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31** (3), 307 – 327.

— (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, **69** (3), pp. 542–547.

Brechmann, E. C., Czado, C. and Aas, K. (2012). Truncated regular vines in high dimensions with application to financial data. *The Canadian Journal of Statistics*, **40** (1), 68 – 85.

Breymann, W., Dias, A. and Embrechts, P. (2003). Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance*, **3** (1), 1–14.

Brockwell, P. J., Davis, R. A. and Yang, Y. (2007). Estimation for nonnegative Lévy-driven Ornstein-Uhlenbeck processes. *Journal of Applied Probability*, **44** (4), 977–989.

—, — and — (2011). Estimation for non-negative Lévy-driven CARMA processes. *Journal of Business & Economic Statistics*, **29** (2), 250–259.

— and Lindner, A. (2012). Lévy-driven time series models for financial data. In T. Subba Rao, S. Subba Rao and C. R. Rao (eds.), *Handbook of Statistics*, vol. 30, Amsterdam: Elsevier, pp. 543–563.

Chavez-Demoulin, V., Embrechts, P. and Nešlehová, J. (2006). Quantitative models for operational risk: Extremes, dependence and aggregation. *Journal of Banking & Finance*, **30** (10), 2635 – 2658.

Cherubini, U., Gobbi, F., Mulinacci, S. and Romagnoli, S. (2012). *Dynamic Copula Methods in Finance*. Chichester: John Wiley & Sons.

—, Luciano, E. and Vecchiato, W. (2004). *Copula methods in finance*. Chichester: John Wiley & Sons.

Choroś, B., Ibragimov, R. and Permiakova, E. (2010). Copula estimation. In P. Jaworski, F. Durante, W. Härdle and T. Rychlik (eds.), *Copula Theory and Its Applications – Proceedings of the Workshop Held in Warsaw*, Berlin: Springer, pp. 77–91.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, **39** (4), pp. 841–862.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, **65** (1), 141–151.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, **1** (2), 223–236.

— and Tankov, P. (2004). *Financial Modelling with Jump Processes*. London: Chapman & Hall/CRC.

Cooke, R. M., Joe, H. and Aas, K. (2011). Vines arise. In D. Kurowicka and H. Joe (eds.), *Dependence Modeling: Vine Copula Handbook*, Singapore: World Scientific, pp. 37–72.

Czado, C., Gärtner, F. and Min, A. (2011). Analysis of Australian electricity loads using joint bayesian inference of d-vines with autoregressive margins. In D. Kurowicka and H. Joe (eds.), *Dependence Modeling: Vine Copula Handbook*, Singapore: World Scientific, pp. 265–280.

— and Min, A. (2011). Bayesian inference for D-vines: Estimation and model selection. In D. Kurowicka and H. Joe (eds.), *Dependence Modeling: Vine Copula Handbook*, Singapore: World Scientific, pp. 249–264.

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés: un test non paramétrique d'indépendance. *Académie Royale Belgique. Bulletin de la Classes des Sciences (5e Série)*, **65**, 274–292.

Dissmann, J., Brechmann, E. C., Czado, C. and Kurowicka, D. (2011). Selecting and estimating regular vine copulae and application to financial returns, to appear in: Computational Statistics and Data Analysis.

Dobrić, J. and Schmid, F. (2007). A goodness of fit test for copulas based on Rosenblatt's transformation. *Computational Statistics & Data Analysis*, **51** (9), 4633 – 4642.

Durante, F. and Sempi, C. (2010). Copula theory: An introduction. In P. Jaworski, F. Durante, W. Härdle and T. Rychlik (eds.), *Copula Theory and Its Applications – Proceedings of the Workshop Held in Warsaw*, Berlin: Springer, pp. 3–31.

Embrechts, P. (2009). Copulas: A personal view. *Journal of Risk and Insurance*, **76** (3), 639–650.

—, McNeil, A. J. and Straumann, D. (2002). Correlation and dependence in risk management: Properties and pitfalls. In M. A. H. Dempter (ed.), *Risk Management: Value at Risk and Beyond*, Cambridge: Cambridge University Press.

Esmaeili, H. and Klüppelberg, C. (2010). Parameter estimation of a bivariate compound Poisson process. *Insurance: Mathematics and Economics*, **47** (2), 224 – 233.

— and — (2011a). Parametric estimation of a bivariate stable Lévy process. *Journal of Multivariate Analysis*, **102** (5), 918 – 930.

— and — (2011b). Two-step estimation of a multivariate Lévy process. *Preprint*.

Fang, K., Kotz, S. and Ng, K. W. (1990). *Symmetric multivariate and related distributions*. London: Chapman & Hall.

Fermanian, J.-D. (2005). Goodness-of-fit tests for copulas. *Journal of Multivariate Analysis*, **95** (1), 119 – 152.

Fischer, M., Köck, C., Schlüter, S. and Weigert, F. (2009). An empirical analysis of multivariate copula models. *Quantitative Finance*, **9** (7), 839–854.

Frahm, G., Junker, M. and Schmidt, R. (2005). Estimating the tail-dependence coefficient: Properties and pitfalls. *Insurance: Mathematics and Economics*, **37** (1), 80 – 100.

Frank, M. J. (1979). On the simultaneous associativity of $f(x, y)$ and $x + y - f(x, y)$. *Aequationes Mathematicae*, **19**, 194–226.

Gaisser, S. C. (2010). *Statistics for Copula-based Measures of Multivariate Association – Theory and Applications to Financial Data*. Ph.D. thesis, University of Cologne.

Genest, C. and Favre, A.-C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, **12** (4), 347–368.

—, Gendron, M. and Bourdeau-Brien, M. (2009a). The advent of copulas in finance. *The European Journal of Finance*, **15** (7-8), 609–618.

—, Ghoudi, K. and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, **82** (3), 543–552.

— and Nešlehová, J. (2007). A primer on copulas for count data. *ASTIN Bulletin*, **37** (2), 475–515.

— and Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*, **88** (423), 1034–1043.

— and Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'Institut Henri Poincaré: Probabilités et Statistiques*, **44** (6), 1096–1127.

—, — and Beaudoin, D. (2009b). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, **44** (2), 199 – 213.

Grothe, O. and Nicklas, S. (2013). Vine constructions of Lévy copulas. *Journal of Multivariate Analysis*, **119**, 1–15.

Gumbel, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, **55** (292), pp. 698–707.

Heinen, A. and Valdesogo, A. (2009). Asymmetric CAPM dependence for large dimensions: the canonical vine autoregressive model, CORE Discussion Papers 2009/69, Université catholique de Louvain, Center for Operations Research and Econometrics.

Hobæk Haff, I. (2012). Comparison of estimators for pair-copula constructions. *Journal of Multivariate Analysis*, forthcoming.

—, Aas, K. and Frigessi, A. (2010). On the simplified pair-copula construction – simply useful or too simplistic? *Journal of Multivariate Analysis*, **101** (5), 1296 – 1310.

Hoeffding, W. (1940). Maßstabsinvariante Korrelationstheorie. *Schriften des mathematischen Instituts und des Instituts für angewandte Mathematik der Universität Berlin*, **5** (3), 181–233.

Hofert, M. (2008). Sampling Archimedean copulas. *Computational Statistics & Data Analysis*, **52** (12), 5163 – 5174.

Joe, H. (1996). Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. In L. Rüschendorf, B. Schweizer and M. Taylor (eds.), *Distributions with Fixed Marginals and Related Topics*, Hayward, CA: IMS, pp. 120–141.

— (1997). *Multivariate Models and Dependence Concepts*. London: Chapman & Hall.

— and Hu, T. (1996). Multivariate distributions from mixtures of max-infinitely divisible distributions. *Journal of Multivariate Analysis*, **57** (2), 240 – 265.

Kallenberg, O. (2002). *Foundation of Modern Probability*. New York: Springer.

Kallsen, J. and Tankov, P. (2006). Characterization of dependence of multidimensional Lévy processes using Lévy copulas. *Journal of Multivariate Analysis*, **97** (7), 1551 – 1572.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, **30** (1/2), 81–93.

Kim, G., Silvapulle, M. J. and Silvapulle, P. (2007). Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics & Data Analysis*, **51** (6), 2836 – 2850.

Kingman, J. F. C. and Taylor, S. J. (1966). *Introduction to measure and probability*. Cambridge: Cambridge University Press.

Kurowicka, D. (2011). Optimal truncation of vines. In D. Kurowicka and H. Joe (eds.), *Dependence Modeling: Vine Copula Handbook*, Singapore: World Scientific, pp. 233–247.

— and Cooke, R. (2003). A parameterization of positive definite matrices in terms of partial correlation vines. *Linear Algebra and its Applications*, **372** (0), 225 – 251.

— and Joe, H. (eds.) (2011). *Dependence Modeling: Vine Copula Handbook*. Singapore: World Scientific.

Mai, J.-F. and Scherer, M. (2012). *Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications*. Series in Quantitative Finance, London: Imperial College Press.

McNeil, A. J. (2008). Sampling nested Archimedean copulas. *Journal of Statistical Computation and Simulation*, **78** (6), 567–581.

—, Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management – Concepts, Techniques, and Tools*. Princeton: Princeton University Press.

— and Nešlehová, J. (2009). Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions. *The Annals of Statistics*, **37** (5B), 3059–3097.

Morales-Nápoles, O. (2011). Counting vines. In D. Kurowicka and H. Joe (eds.), *Dependence Modeling: Vine Copula Handbook*, Singapore: World Scientific, pp. 189–218.

Nelsen, R. B. (2006). *An Introduction to Copulas*. New York: Springer, second edition edn.

Nikoloulopoulos, A. K., Joe, H. and Li, H. (2012). Vine copulas with asymmetric tail dependence and applications to financial return data. *Computational Statistics & Data Analysis*, **56** (11), 3659 – 3673.

Okhrin, O. (2007). *Hierarchical Archimedean Copulas: Structure Determination, Properties, Applications*. Ph.D. thesis, Europa-Univarsität Viadrina, Frankfurt (Oder).

Panjer, H. H. (2006). *Operational Risks – Modeling Analytics*. Hoboken: John Wiley & Sons.

R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Rhoads, R. (2011). *Trading VIX Derivatives: Trading and Hedging Strategies Using VIX Futures, Options, and Exchange Traded Notes*. Hoboken: John Wiley & Sons.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The Annals of Mathematical Statistics*, **23** (3), 470–472.

Rosiński, J. (2001). Series representations of Lévy processes from the perspective of point processes. In O. E. Barndorff-Nielsen, T. Mikosch and S. I. Resnick (eds.), *Lévy Processes: Theory and Applications*, Boston: Birkhäuser.

Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge: Cambridge University Press.

Savu, C. and Trede, M. (2010). Hierarchies of Archimedean copulas. *Quantitative Finance*, **10** (3), 295–304.

Schmid, F. and Schmidt, R. (2007). Multivariate conditional versions of Spearman's rho and related measures of tail dependence. *Journal of Multivariate Analysis*, **98** (6), 1123 – 1140.

—, —, Blumentritt, T., Gaisser, S. and Ruppert, M. (2010). Copula-based measures of multivariate association. In P. Jaworski, F. Durante, W. Härdle and T. Rychlik (eds.), *Copula Theory and Its Applications – Proceedings of the Workshop Held in Warsaw*, Berlin: Springer, pp. 209–236.

Schnieders, J. (2012). *Analyzing and Modeling Multivariate Association – Statistical Measures and Pair-Copula Constructions*. Ph.D. thesis, University of Cologne.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, **8** (3), 229–231.

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, **15** (1), 72–101.

Tankov, P. (2004). *Lévy Processes in Finance: Inverse Problems and Dependence Modelling*. Ph.D. thesis, Ecole Polytechnique, France.

— (2005). Simulation and option pricing in Lévy copula models. *Unpublished manuscript*.

Vaz de Melo Mendes, B., Mendes Semeraro, M. and Câmara Leal, R. P. (2010). Pair-copulas modeling in finance. *Financial Markets and Portfolio Management*, **24** (2), 193–213.

Whelan, N. (2004). Sampling from Archimedean copulas. *Quantitative Finance*, **4** (3), 339–352.

Wuertz, D. and Chalabi, Y. (2009). *fGarch: Rmetrics - Autoregressive Conditional Heteroskedastic Modelling*. R package version 2110.80.