Ordinary and Lévy Copulas in Finance

Models, Methods and Tools for Risk Management and Option Pricing

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

am Fachbereich C – Mathematik und Naturwissenschaften
der Bergischen Universität Wuppertal
genehmigte

Dissertation

von

M.Sc. Kai Tappe

aus Leverkusen

Tag der mündlichen Prüfung: 17. November 2008
Gutachter/Prüfer: Prof. Dr. Michael Günther
Gutachter/Prüfer: Prof. Dr. Peter Kloeden
Gutachter: Prof. Dr. Michael Nelles
Prüfer: Prof. Dr. Reinhard Michel
Prüfer: Jun.-Prof. Dr. Roland Pulch

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20090982
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20090982]

Meinem Neffen Lenn

Danksagung

Die vorliegende Dissertation entstand aus einer Kooperation der Bergischen Universität Wuppertal und der Sparkasse Leverkusen. Daher möchte ich an erster Stelle meinen Betreuern Prof. Dr. Michael Günther und Andreas Tangemann einen ausdrücklichen Dank für die intensive und erfolgereiche Zusammenarbeit aussprechen.

Ebenso möchte ich mich bei Manfred Herpolsheimer und Markus Vandrey redlich dafür bedanken, dass sie im Namen der Sparkasse Leverkusen durch die Vergabe eines außergewöhnlichen Forschungsstipendiums an und das damit verliehene Vertrauen in meine Person mir die willkommene Gelegenheit boten, meine Forschungstätigkeit auf eine praktische Anwendung zu richten. Ich achte die Gewährung eines Einblicks in die hausinternen Belange des Risikomanagments hoch.

Zugleich schulde ich Andreas Tangemann und Oliver Klenner für das gemeinsame Wirken und die freundschaftliche Zusammenarbeit aufrichtigen Dank. In den zumeist förderlichen Diskussionsrunden profitierte ich stark von ihrer Fachkompetenz und betriebswirtschaftlichen Sichtweise. Besonderer Dank gilt ihnen für das mir entgegengebrachte Zutrauen, meine Forschungsarbeiten im Rahmen des Promotionsprojektes weitgehend eigenverantwortlich zu gestalten. Es war mir eine große Freude, mit ihnen zusammenzuarbeiten.

Daneben möchte ich dem gesamten Sparkassen-Team 1300 meinen Dank sagen für die herzliche Aufnahme und das angenehm neckische Arbeitsumfeld.

Zweifellos möchte ich gegenüber Michael Günther einen besonderen Dank dafür äußern, dass er mein Promotionsvorhaben im Rahmen der Kooperation mit der Sparkasse Leverkusen ermöglichte und gewissenhaft betreut hat. Michael hat sich immer die Zeit genommen, meine fachlichen und organisatorischen Fragen zu beantworten. Zudem wertschätze ich das mir entgegengebrachte Vertrauen, die Schwerpunkte meiner Forschungstätigkeit frei zu gestalten.

Der gesamten Arbeitsgruppe Angewandte Mathematik/Numerische Analysis gilt mein Dank für ein gutes Forschungsumfeld und eine kameradschaftliche Umgangsweise. Ein besonderer Dank richtet sich an Cathrin von Emmerich, Christian Kahl und Patrick Deuß, die meine Arbeit mit Impulsen und Gegenmeinungen entscheidend vorantrieben und damit geprägt haben.

Große Dankbarkeit empfinde ich auch für den Rückhalt und das Verständnis seitens meiner Familie, meinen Freunden und Julia, die mir während des gesamten Studiums der Mathematik mit Geduld und Zuspruch beistanden. Ich danke meinen Eltern, dass sie mir ein Studium erst ermöglicht haben.

Contents

List of Figures

List of Tables

List of Algorithms

Preface

Dependence prepares to play an important role in contemporary finance. One can hardly imagine quantitative practices without the ground-breaking policies for interdependent business lines or the glorifying market launch of multi-line products. The latest interdisciplinary approaches to risk profiling and global trading have raised competence standards in the banking business by and by to staying abreast of dependence issues, which arise in the market these days. The identification of dependence structures has emerged as a key ingredient in simply all financial applications: portfolio management, risk assessment and derivative pricing, to name just a few. In these disciplines, complexity keeps growing at the pace of global market developments. The sound understanding of interrelations between different fields of action gains truly vital importance.

By tradition, actuarial and financial theory revolves very much around the assumption of normal markets, making the concept of linear correlation almost synonymous with that of dependence. Both theorists and practitioners have long resorted to Brownian motion and the Gaussian concept of dependence. In recent years however, authors have deviated from the assumption of joint normality, paying tribute to the insufficiencies of linear correlation in real-world applications. Empirical data was found to indicate a nontrivial degree of non-normal dependence, such as extreme co-movements and asymmetry, which contradicts the use of linear correlation as a summary statistic for complex dependence structures.

In a not necessarily normal world, copula functions prepare to cope with dependence structures between financial instruments in a very flexible way. Having made first appearance in Sklar [74], the notion of a copula is well known for some time in the stochastic literature. It has regained a paramount importance for the modelling of static and dynamic dependence in recent years. Copulas admit a universally valid approach to multidimensional dependence structures and facilitate the description of complex dependence concepts other than linear correlation.

This doctoral thesis is intended for a self-contained presentation of theoretical, numerical and empirical research on the use of copula methods in the fields of static and dynamic financial modelling. The motivation is precisely to offer methods of resolution for multidimensional dependence models with emphasize on practical implementations. We develop and combine techniques for the simulation, the estimation and the calibration of dependence phenomena in response to the growing need of truly practicable solutions for the increasingly complex financial instruments.

There are two parts in this thesis. Chapter 1 through Chapter 5 constitute the first part. They develop mostly fundamental results and techniques on the use of copula functions in modelling dependence between static random variables. The goal is to furnish a comprehensive scheme of how copula methods apply to high-dimensional dependence problems in multivariate probability distributions. Here we are keenly interested in the issue of scenario generation and model estimation for risk management purposes. Chapter 6 through Chapter 10 are the second part, which explains the use of copula methods in stochastic process modelling. We approach the notion of a dynamic dependence structure using characteristic functions of multi-dimensional Lévy processes. Our intents are to develop sophisticated simulation and calibration methods with regards to option pricing applications. At the same time, this thesis serves as an introduction to financial modelling with jump processes. The parts are widely independent from each other up to one of our main contributions, which combines the two copula notions.

Our firm intention is to treat dependence issues on their own. Hence we barely touch marginal distributions or processes in this thesis. The focus is knowingly set on the stand-alone association inasmuch as we exclude stochastic modelling in one dimension except for providing ourselves with the basics. Nor are dependence concepts or dependence measures a matter, albeit summary statistics and relevance orderings are a growing field of research in the scope of copula functions. As far as numerical examples are concerned, we do not go beyond enterprize risk assessment and European option pricing. Other applications of copula functions include portfolio management, credit derivatives and time series analysis.

Concerning the level of mathematical detail used to treat all the topics, we do not provide an in-depth study of multivariate distribution functions or general Lévy processes, as these already exist. Here we intend to elaborate on the mathematical tools necessary in the context of pure dependence modelling. We emphasize a mere subset of models (some of which are findings of our own) rather than surveying a catalogue of models, which can be found in the literature. Our selection is targeted on numerical feasibility and omits mathematically interesting topics other than (to us) practicable modelling tools.

This thesis grew out of a joint venture with Sparkasse Leverkusen, which aimed at the development of a new management tool for the adequate assessment of the firm-wide risk exposure. The achievement of these objectives are set down in Part I of this thesis. The contents therein are deliberately held moderate in mathematical abstraction and technicality in such a way as to serve in part as end-user support of the in-house implementation. Hence this part comprises fairly applied research, which is readily accessible also for an inexpert reader. Based on the application-oriented work, I began with the conceptually similar research field of dynamic dependence structures. This interest was on the spot drawn by Tankov's pioneering in this direction. Hereafter, the development of dependence methods for multidimensional Lévy processes paralleled the risk management concerns. Not least by attending the AMaMeF Conference 2007 and the NMF Conference 2008, the fascination for multidimensional Lévy processes was kept going inasmuch as it encouraged me to continue researching in the hugely popularized option pricing applications. These results are written down in Part II of this thesis. The contents therein are mathematically more challenging and contain this thesis' scientific research in large part.

Basic tools and notation

Sets, measurable spaces and measures $\mathbb{N}, \mathbb{R}$ and $\mathbb{C}$ are, respectively, the collections of all positive integers, all real numbers, and all complex numbers. For $z \in \mathbb{C}$, we denote the real part of z by $\Re z$ and the imaginary part of z by $\Im z$. $\mathbb{R}_\infty$ (or $\mathbb{R}_{\infty,+}$) is the collection of all real numbers (or all nonnegative real numbers) extending to $\{\infty\}$.

$\mathbb{R}^d$ is the d-dimensional Euclidian space. Its elements $x = (x_1, ..., x_d), y = (y_1, ..., y_d)$ are column vectors with d real components. The vector consisting entirely of ones is denoted by $\mathbf{1}$. The vector consisting entirely of zeros is denoted by $\mathbf{0}$. The inner product is $x.y = \sum_{i=1}^d x_i y_i$.

$\mathbb{R}^{r\times d}$ is the $(r \times d)$-dimensional space of real-valued matrices M with entries $m_{ij}, i = 1, ..., r, j = 1, ..., d$. The root $M^{1/2}$ of a regular matrix $M \in \mathbb{R}^{d\times d}$ is the matrix $L \in \mathbb{R}^{d\times d}$ such that $L'L = M$. The identity matrix is denoted by $\mathbf{1}$ (abusing notation).

$\mathbb{R}^d_\infty$ (or $\mathbb{R}^d_{\infty,+}$) is the d-dimensional space consisting of the elements $x = (x_1, ..., x_d)$ with $x_i \in \mathbb{R}_\infty$ (or $x_i \in \mathbb{R}_{\infty,+}$) for $i = 1, ..., d$. The word d-variate is used with the same meaning as d-dimensional.

For sets A and B, $A \subset B$ means that all elements of A belong to B. For $A \subset \mathbb{R}^d$, $z \in \mathbb{R}^d$, $A + z = \{x + z : x \in A\}$. $\sharp A$ is the number of elements of a set A.

$(E, \mathcal{E})$ denotes a measurable space, where $\mathcal{E}$ is a σ-algebra on E. For $E \subset \mathbb{R}^d$, $\mathcal{B}(E)$ denotes the Borel σ-algebra of E. An element $A \in \mathcal{E}$ is called a measurable set and, for a measure μ on $(E, \mathcal{E})$, $\mu(A)$ is called its measure.

A measure μ_2 is said to be absolutely continuous with respect to a measure μ_1 if for any measurable set A it holds $\mu_1(A) = 0 \Rightarrow \mu_2(A) = 0$.

If μ_2 is absolutely continuous with respect to μ_1 then there exists a measurable function $Z : E \to [0, \infty)$ such that for any measurable set A it holds $\mu_2(A) = \int_A Z d\mu_1 = \mu_1(Z1_A)$. The function Z is called the density or Radon-Nykodym derivative of μ_2 with respect to μ_1 and denoted as $d\mu_2/d\mu_1$.

The Lebesgue-measure of a measurable set $A \in \mathcal{B}(\mathbb{R}^d)$ is denoted as $\lambda(A)$. $\lambda(dx)$ is written dx.

The symbol δ_x represents the probability measure concentrated at x.

$\mu_{|B}$ denotes the restriction of μ to B.

Functions and operations The integral of a vector-valued function is a vector with componentwise integrals.

For $B \subset \mathbb{R}^d$, 1_B is the indicator function of set B, i.e. $1_B(x) = 1$ for $x \in B$ and

0 for $x \notin B$. The operation $a \wedge b$ denotes the minimum $\min\{a,b\}$, and $a \vee b$ denotes the maximkum $\max\{a,b\}$ of two real numbers. The expression $\text{sgn}(x)$ represents the sign function, i.e. $\text{sgn}(x) = 1$ if $x \geq 0$ or $\text{sgn}(x) = -1$ if $x < 0$.

Probability spaces, random variables and processes $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ is a filtered probability space, i.e. a complete probability space $(\Omega, \mathcal{F}, P)$ consisting of a measurable space $(\Omega, \mathcal{F})$ and a probability measure P, endowed with a right continuous filtration $\{\mathcal{F}_t : t \in [0,T]\}$ such that $\mathcal{F} = \mathcal{F}_T$ and all the null sets of $\mathcal{F}$ are contained in $\mathcal{F}_0$, where $T \geq 0$ is a terminal time. $\{\mathcal{F}_t\}$ is interpreted as the information flow.

$A \in \mathcal{F}$ (or $A \in \mathcal{F}_t$) is an event, $P[A]$ (or $P_{|\mathcal{F}_t}[A]$) is the probability of the event A, and $P[A|\mathcal{F}_s]$ (or $P_{|\mathcal{F}_t}[A|\mathcal{F}_s]$) is the probability of the event A conditional on the information $\mathcal{F}_s$.

For a probability space $(\Omega, \mathcal{F}, P)$, a mapping X from Ω into $\mathbb{R}^d$ is an $\mathbb{R}^d$-valued random variable (or random variable on $\mathbb{R}^d$), if it is $\mathcal{F}$-measurable, that is, $\{\omega \in \Omega : X(\omega) \in B\}$ is in $\mathcal{F}$ for each $B \in \mathcal{B}(\mathbb{R}^d)$, where $\mathcal{B}(\mathbb{R}^d)$ is the Borel σ-algebra of $\mathbb{R}^d$.

For a $\mathbb{R}^d$-valued random variable $X = (X^1, ..., X^d)$ and an index set $I \subset \{1, ..., d\}$ of cardinality $|I|$, we call the random variable $X^I = (X^i)_{i \in I}$ on $\mathbb{R}^{|I|}$ the I-margin of X. The one-dimensional random variables $X^{\{i\}} = X^i$ are referred to as margins.

We write $P[\{\omega \in \Omega : X(\omega) \in B\}]$ as $P[X \in B]$. As a mapping of B, this is a probability measure on $\mathcal{B}(\mathbb{R}^d)$, which we denote by $P^X(B)$ and call the distribution of X.

If P^X is absolutely continuous with respect to the Lebesgue measure λ, then we call X a continuous random variable.

Two random variables X, Y on $\mathbb{R}^d$ are identically distributed, denoted by $X \overset{d}{=} Y$, if $P_X(B) = P_Y(B)$ for all $B \in \mathcal{B}(\mathbb{R}^d)$.

For a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, a family $\{X_t : t \in [0,T]\}$ of random variables on $\mathbb{R}^d$ is called a stochastic process and denoted by $\{X_t\}$ (or $(\{X_t\}, P)$ stressing the probability measure P). We assume that $\{X_t\}$ is $\{\mathcal{F}_t\}$-adapted, that is non-anticipating with respect to the information flow.

For any fixed $0 \leq t_1 < t_2, ..., t_n \leq T$, $P[X_{t_1} \in B_1, ..., X_{t_n} \in B_n]$ determines a probability measure on $\mathcal{B}((\mathbb{R}^d)^n)$. The family of the probability measures over all choices of n and $t_1, ..., t_n$ is called the system of finite-dimensional distributions of $\{X_t\}$. For any $t \in [0,T]$, in particular, X_t is a random variable on $\mathbb{R}^d$ and $P[X_t \in B]$ defines a probability measure on $\mathcal{B}(\mathbb{R}^d)$, which we denote by P^{X_t} and call the distribution of X_t.

If X is a random variable on $\mathbb{R}^d$, and h is a measurable function on $\mathbb{R}^d$ such that the integral $\int_{\mathbb{R}^d} h(x)P^X(dx) = \int_\Omega h(X(\omega))P(d\omega)$ exists, then $h(X)$ is said to be P-integrable and we call $\int_{\mathbb{R}^d} h(x)P^X(dx)$ its expectation (with respect to P) denoted by $E^P[h(X)]$. The expectation with respect to the probability measure P conditioned on information $\mathcal{F}_t$ is written as $E^P[h(X)|\mathcal{F}_t]$.

The choice $h(x_1, ..., x_d) = (x_1, ..., x_d)$ gives the expectation $E^P[X]$ of X. The choice $h(x) = e^{iz.x}$, seen as a function of $z \in \mathbb{R}^d$, gives the characteristic function of the distribution P^X of a random variable X, denoted by φ^X, that is $\varphi^X(z) = \int_{\mathbb{R}^d} e^{iz.x} P^X(dx) = E^P[e^{iz.X}]$. Allowing complex arguments $z \in \mathbb{C}$ we denote the extended characteristic function by $\varphi^X(z)$.

For equivalent measures P, Q, the restrictions $P_{|\mathcal{F}_t}$ and $Q_{|\mathcal{F}_t}$ are equivalent and there exists a positive stochastic process, called the density process of Q with respect to P and denoted by $\{dQ_{|\mathcal{F}_t}/dP_{|\mathcal{F}_t}\}$, such that, for any random variable Z on $\mathbb{R}^d$, $E^Q[Z|\mathcal{F}_t] = E^P[ZdQ/dP|\mathcal{F}_t]$.

Increasing, grounded and distribution functions For $a, b \in \mathbb{R}^d_\infty$ we write $a \leq b$ if $a_i \leq b_i, i = 1, ..., d$. In this case let $(a, b]$ denote the right-closed left-open interval $(a_1, b_1] \times ... \times (a_d, b_d]$ in $\mathbb{R}^d_\infty$.

To any $\mathbb{R}_\infty$-valued function on $S = S_1 \times ... \times S_d \subset \mathbb{R}^d_\infty$ and $a, b \in S$ with $a \leq b$ and $\overline{(a, b]} \subset S$, we associate the F-volume of $(a, b]$, denoted by $V_F((a, b])$,

$$V_F((a,b]) = \sum_{u \in \{a_1,b_1\} \times ... \times \{a_d,b_d\}} (-1)^{N(u)} F(u), \tag{1}$$

where $N(u) = \sharp\{k : u_k = a_k\}$.

F is said to be d-increasing, if $V_F((a, b]) \geq 0$ for all such intervals $(a, b] \subset \mathbb{R}^d_\infty$.

Given that S_i has a least element s_i for $i = 1, ..., d$, F is called grounded if $F(u) = 0$ for all $(u_1, ..., u_d) \in S_1 \times ... \times S_d$ such that $u_i = s_i$ for at least one $i \in \{1, ..., d\}$. In other words, a grounded function vanishes on the lower bounds of its domain.

For any non-empty index set $I \subset \{1, ..., d\}$, the I-margin of F is the function

$$F_I((u_i)_{i \in I}) = \sup_{a_i, b_i \in S_i : i \in I^c} \sum_{(u_j)_{j \in I^c} \in \prod_{j \in I^c} \{a_i, b_i\}} F(u_1, ..., u_d) \prod_{j \in I^c} \operatorname{sgn}(u_j),$$

where $I^c = \{1, ..., d\} \setminus I$.

Probability distributions are closely linked with certain increasing functions. We associate to every distribution P^X of a $\mathbb{R}^d$-valued random variable X a $[0, 1]$-valued cumulative distribution function (cdf) on $\mathbb{R}^d_\infty$, denoted by F (or F^X stressing the random variable),

$$F^X(x) = P^X((-\infty, x_1] \times ... \times (-\infty, x_d]).$$

The cdf F^X of a $\mathbb{R}^d$-valued random variable X is d-increasing as $V_{F^X}(B) = P^X(B) \geq 0$ for all $B \in \mathcal{B}(\mathbb{R}^d)$.

The Radon-Nykodym derivative of P^X with respect to λ (if it exists) is called the probability density function (pdf) of X and denoted by f (or f^X stressing the random variable). It holds

$$f^X(x_1, ..., x_d) = \frac{\partial^d F^X(x_1, ..., x_d)}{\partial x_1 ... \partial x_d}$$

in every point $x \in \mathbb{R}^d$, where P^X (or F^X or X) is continuous. In all other points, P^X (or F^X or X) is said to be singular. One can show that any cdf F^X is singular only on a set with zero Lebesgue measure, i.e. almost everywhere (a.e.).

The distribution function F^{X^I} of the I-margin X^I of a random variable X on $\mathbb{R}^d$ is given by

$$F^{X^I}((x_i)_{i \in I}) = F^X(y_1, ..., y_d),$$

where $y_i = x_i, i \in I$ and $y_i = \infty, i \in I^c$. With a view to increasing functions, we write F_I^X for the I-margins and F_i^X for the margins of F^X.

For two distinct index sets $I, J \subset \{1, ..., d\}$ we denote by $F_{I|J}^X$ the conditional cdf of $(X^i)_{i \in I}$ given $(X^i)_{i \in J}$, defined by

$$F_{I|J}((x_i)_{i \in I} | (x_i)_{i \in J}) = P(X^i \leq x_i, i \in I | X^i = x_i, i \in J).$$

If F^X (or f^X) depend on some set θ of parameters, we write $F^X(x; \theta)$ (or $f^X(x; \theta)$).

Standard probability distributions and transforms For a $\mathbb{R}$-valued random variable X, we write $X \sim UNF(a, b)$ if X is uniformly distributed on the interval $[a, b]$, $X \sim EXP(\lambda)$ if X is exponentially distributed with intensity $\lambda > 0$, $X \sim POIS(\lambda)$ if X has a Poisson distribution with intensity $\lambda > 0$, $X \sim GAM(\alpha, \beta)$ if X has a Gamma distribution with parameters α and β, or $X \sim CHI2(\nu)$ if X has a χ^2-distribution with $\nu > 0$ degrees of freedom.

For a $\mathbb{R}^d$-valued random variable X, we write $X \sim MVN(\mu, \Sigma)$ if X has a (multivariate) normal distribution with mean $\mu \in \mathbb{R}^d$ and covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$, $X \sim MVT(\mu, \Sigma, \nu)$ if X has a (multivariate) t-distribution with mean $\mu \in \mathbb{R}^d$, covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$ and ν degrees of freedom.

For a cdf F^X of a random variable X on $\mathbb{R}$, let $(F^X)^{-1}$ denote its generalized inverse (or quantile function), i.e. the strictly increasing function defined by $(F^X)^{-1}(y) = \inf\{x \in \mathbb{R} : F^X(x) \geq y\}$. If X is absolutely continuous, then there exists the inverse $(F^X)^{-1}$ in the usual sense.

If F is some cdf and $U \sim UNF(0, 1)$ has standard uniform distribution, then the random variable $F^{-1}(U)$ is F-distributed. This result is known as the quantile transformation.

If X is a random variable on $\mathbb{R}$ with absolutely continuous cdf F^X, then $F^X(X)$ has standard uniform distribution. This transformation is known as the probability integral transform.

Part I

Ordinary copula methods

Chapter 1

Introduction

Owing to a mutual consent to abandon the assumption of joint normality for the increasingly complex behavior of financial markets, copula based models have recently gained popularity among academics and practitioners in the field of finance. They are becoming the most significant new tool to handle the dependence between financial phenomena in a flexible way.

The persistent departure from normality in mathematical finance is the consequence of a growing evidence against joint normality of real markets. The markets exhibit smiles, skewness, excess kurtosis, jumps and extreme events, such as crashes or defaults. Normality seems unrealistic regarding the association between single phenomena, too. It is acknowledged that the co-movements in extreme situations, such as market crashes or credit crunches, is quite different from a normal behavior. The observable deviations include tail dependence, orthant dependence and asymmetry, to name just a few pieces of evidence against linear correlation as a summary statistic for complex dependence structures. Hence the empirical distributions are non-normal in two respects: as to the marginal factors and to the dependence structure. At the level of univariate phenomena, people have given competent answers, as they deviated from the classical principles. With a view to the emergence of multi-name financial products and multi-line business perspectives, academics and practitioners engage increasingly in the modelling of dependent distributions. New challenges are, for example, the profiling of risk exposures across all business lines of a financial institution or the pricing of derivatives written on a basket of names. This makes the handling of dependence structures among non-normally distributed variables unavoidable.

Here copula functions prove useful to decouple marginal aspects of a joint distribution from its dependence structure. Hence, they provide a flexible way to cope with non-normal marginal distributions and complex dependence structures in a separate manner. In other words, copula functions allow us to concentrate on the subject matter of dependence on its own. This is in tune with the univariate techniques already used at the level of single line perspectives. Our motivation is then to provide parsimonious and practicable approaches to multivariate copula modelling. We take account of the challenge that growing dimensions impose the controversial issue of a simple but rich dependence structure. Important ideas in this part include

- the use of ordinary copula functions as an appropriate summary of the dependence in a multivariate distribution separate from univariate margins,
- the parametric construction of copula based dependence models with emphasis on the elliptical and the pair type pattern,
- the structural properties of parametric copula families, which are conducive to simulation and estimation purposes,
- the development of model-intrinsic inference methods, especially the fitting of estimated parameters,
- the explicit documentation of the algorithmic handling as a walkthrough to implementations,
- the application of ordinary copula methods to the business assessment of the firm-wide risk exposure.

With this agenda we fall into line with a shower of publications on multivariate dependence structures and applied copula modelling.

The basic original reference on ordinary copula functions is Sklar [74]. A comprehensive summary of the history of copula functions (alias uniform representations [47] alias dependence functions [21]) is given in Schweizer [72]. Nelson's recent book [57] is an excellent monograph on copulas, featuring a comprehensive register of standard families and dependence properties. Joe [43] is a recurring reference and contains a more thorough classification of multivariate models with emphasis on distributional features of numerous construction patterns and on marginal perspectives. Orderings and measures of various dependence concepts is detailed there. Embrecht et al. [33] give a brief guide to the need of complex dependence structures, presenting fallacies from an illegitimate use of the normality assumption. Other standard literature comprising a qualitative inspection of empirical dependence phenomena is Embrecht et al. [32] and Mashal and Zeevi [53]. The fundamental textbooks by Cherubini et al. [16] and Embrecht et al. [31] give a comprehensive overview of copula based models, including parametric construction, simulation techniques, inference methods and applications. Related works are Bouyé et al. [13], Melchiori [55] and Lindskog [49]. It is interesting to see that some authors [e.g. 57, 16] make a sharp distinction between bivariate and multivariate dependence models, both in terms of definition and methodology, when deviating from elliptical copulas. Researchers and practitioners [e.g. 63] often restrict themselves to the bivariate case and leave the matter of multidimensional dependence concepts scarcely touched. Beyond this limitation, only a few authors have tackled the problem of high-dimensional dependence structures, answering to the growing need of a methodology for involved multivariate models. In this regard, Savu and Trede [70] give an elaborate extension of the Archimedean construction so as to incorporate hierarchical dependence structures. They extend the profound ideas in Joe [43]. But their findings do not provide the appropriate techniques for simulation and estimation of the model. The sampling problem is in turn taken up by Whelan [83], who provides some limited redress from this delicate task. But still, the literature on Archimedean copula functions shows a severe imbalance of available models and practical implementations. Aas et al. [1] contains a very practical alternative, that builds a multivariate model from families of conditional bivariate copulas. It pioneers the formulation of intrinsic sampling and inference methods for markedly ample dependence structures. Nelson [57] and Joe [43] make observations about copulas in the context of stochastic processes (discrete Markov chains). Patton [63] first introduced copulas in the conditional sense and thus paved the way for time varying dependence structures. Subsequent applications to discrete time series processes is found in Rockinger and Jondeau [65], Dias [22] and Palaro and Hotta [60]. Cherubini and Luciano [15] and Wakefield [81] use copula functions for multi-asset option pricing. Still, risk management is among the most frequently discussed fields of application. Along with the afore mentioned textbooks, Bouyé et al. [13] and Embrecht et al. [32, 33] provide with a plain mathematical formulation of risk measurement with copulas. Eberlein et al. [24] gives an overview of mathematical models and methods used in financial risk management. From there, Beck and Lesko [8], Beck et al. [9] and Schumacher et al. [71] give extensive studies on the copula approach in financial risk aggregation. The present part is most similar to those papers.

We pursue the targets, which we outlined before, single-minded with regard to the business application at Sparkasse Leverkusen. The goal is to explain ordinary copula methods to a team of economists and to offer some unmitigated resolutions for the in-house risk assessment across a growing number of business lines. The status quo of risk management practices at Sparkasse Leverkusen is described by self-contained solutions for the separate assessment of various risk factors. Hence our only worry is the efficient determination of the overall capital adequacy in the sense of a top-down dependence structure. In this part we distance ourselves from

- the parametric modelling of marginal distributions as to the rejected assumption of normality,
- an exhaustive survey of existent copula models or the development of new dependence models,
- a guide to dependence measures and orderings.

As a consequence, we mainly reproduce existent copula methods here. Then our contribution is (aside from a revised parameter test) the business application of efficient practices to the measurement issue of the overall risk exposure at Sparkasse Leverkusen.

The part is organized as follows. In Chapter 2 we define an ordinary copula function as a specific multivariate distribution. We overlook its basic properties and give reasons for the integrity of ordinary copula functions in multivariate models. Then we review some state-of-the-art copula models. Here the focus is on the elliptical and the pair copula construction. The popular Archimedean family is described in Section A.1.

In Chapter 3 we treat the generation of random numbers from ordinary copula functions. The goal is to offer efficient sampling procedures with parsimonious computations. We argue in particular that the conditional sampling approach is the right choice for pair copula functions, while the transformation method is very effective for elliptical copula functions.

We discuss model estimation in Chapter 4. We describe how to extract model parameters from a set of sample data by the maximum likelihood method. This is applied to the elliptical and the pair copula models. Then we go into detail about testing the fitted parameters. We derive explicit goodness-of-fit tests for elliptical and pair copula functions.

In Chapter 5 the methods discussed in Chapters 2, 3 and 4 are applied to assessing a book of risk lines at Sparkasse Leverkusen. We explain the regulative risk drill in contemporary banking business and show how ordinary copula methods apply to the quantitative risk management. The closing section contains a documentation of the risk management tool CopRisk, which we personally have implemented for business use at Sparkasse Leverkusen.

For publication, the actual risk figures of Sparkasse Leverkusen were made anonymous to comply with existing non-disclosure agreements. This is done by normalizing all risk data in the graphics to actual values as of July 31st 2008, with the original wording retained in the explanation of such. The empirical findings described are then not perfectly consistent with the graphics but reflect actual results.

Chapter 2

Ordinary copula functions

In this chapter we model the dependence between random variables by means of copula functions. Then we analyze two common classes of copula functions.

Section 2.1 defines a copula function and discusses the fundamentals. Here the main result, which is due to Sklar [74], isolates the description of the dependence structure between associated random variables. Section 2.2 introduces the elliptical construction pattern as a direct consequence of Sklar's theorem. Then two standard examples of elliptical copulas are given. Section 2.3 develops the modern pair copula approach, which is based on the recursive use of building block copula functions. Then two particular recursions are detailed.

The presentation in this chapter mainly builds on the standard textbooks by Cherubini et al. [16], Embrecht et al. [31], Joe [43] and Nelson [57]. Aas et al. [1] is the major reference in the subject matter of pair copulas.

2.1 Definition and basic properties

Copula functions reopened recently as the primary tool in the analysis of multivariate dependence structures. We give definitions, properties and fundamental examples of copula functions.

Definition 2.1. A *d-dimensional copula*[1] function C (or d-copula or copula) is a mapping of the form $C : [0,1]^d \to [0,1]$ such that

(1) C is grounded

(2) C is d-increasing and

(3) $C(1, ..., 1, u_i, 1, ..., 1) = u_i$, for all $i \in \{1, ..., d\}, u_i \in [0,1]$

[1]The word *copula* is Latin for the English nouns *link*, *tie* or *bond*. It is used in grammar as a type of word, which connects a subject with its complements.

Properties (1) and (2) are necessary and sufficient conditions for a function C to be a multivariate cdf [cf. 43, Section 1.4.2]. Property (3) is the extra requirement of uniformity of the margins $C_i, i = 1, ..., d$. Hence copulas are joint cdfs on the unit hypercube $[0, 1]^d$ with standard uniform marginal distributions.[2] More precisely, a d-copula C is the cdf of a random variable $U = (U^1, ..., U^d)$ on $\mathbb{R}^d$,

$$C(u_1, ..., u_d) = P[U^1 \leq u_1, ..., U^d \leq u_d], \tag{2.1}$$

where $U^i \sim UNF(0, 1), i = 1, ..., d$ is standard uniformly distributed. We reserve the notation C (instead of F^U) for joint cdfs that are copulas. We index a d-copula $C = C_{1,...,d}$, whenever it is appropriate to stress dimensions.

The following result was proven by Sklar [74]. It shows, on the one hand, that all multivariate cdfs contain copula functions and, on the other hand, that copula functions may be used in conjunction with marginal cdfs to construct multivariate cdfs:

Theorem 2.2 ([57], Theorem 2.10.9)**.** *Let F be a joint cdf with margins $F_1, ..., F_d$. Then there exists a d-copula C such that, for all $x_1, ..., x_d \in \mathbb{R}_\infty$,*

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d)). \tag{2.2}$$

If the margins are continuous, then C is unique. Conversely, if C is a d-copula and $F_1, ..., F_d$ are univariate cdfs, then the function F defined by (2.2) *is a joint cdf with margins $F_1, ..., F_d$.*

Theorem 2.2 shows that copula functions describe the interrelations between the margins of a joint distributions. In turn, copula functions themselves are determined by the joint and marginal distributions:

Corollary 2.3 ([57], Theorem 2.10.10)**.** *Let F be a multivariate cdf with margins $F_1, ..., F_d$ and copula C. Then, for any $(u_1, ..., u_d) \in [0, 1]^d$,*

$$C(u_1, ..., u_d) = F(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)). \tag{2.3}$$

Corollary 2.3 shows how to isolate the copula function from a multivariate distribution with given margins. For a random variable $X = (X^1, ..., X^d)$ on $\mathbb{R}^d$ with joint cdf F and margins $F_1, ..., F_d$, the copula in (2.2) (or in (2.3)) is called the *copula of* X (or the copula of $X^1, ..., X^{d-1}$ and X^d) and is denoted by C_X (or $C_{X^1,...,X^d}$, whenever indexing is advantageous).

In the sense of a dependence structure in its own right, the copula of a random variable somehow ignores the marginal distributions:

Theorem 2.4 (cf. [57], Theorem 2.4.3)**.** *Let $X^1, ..., X^d$ be continuous random variables with copula $C_{X^1,...,X^d}$. If $T_i : \mathbb{R} \to \mathbb{R}, i = 1, ..., d$ are strictly increasing, then $T_1(X_1), ..., T_n(X_d)$ have copula $C_{X^1,...,X^d}$.*

[2] Copula functions belong to the Fréchet classes, i.e. to the classes of multivariate distributions with given margins [cf. 43].

Theorem 2.4 states that copulas of random variables are invariant under strictly increasing transformations of the margins. After all, copula functions are pure dependence structures of joint distributions.[3]

Understood as a joint cdf, the pdf c of C exists in all points $u \in \mathbb{R}^d$, where C is continuous [cf. 16, Theorem 2.10]. Using (2.2) and the chain rule, we then have, for a joint cdf F with margins $F_1, ..., F_d$, the *canonical representation* [cf. 16, Section 2.6]

$$f(x_1, ..., x_d) = c(F_1(x_1), ..., F_d(x_d)) \prod_{i=1}^{d} f_i(x_i) \tag{2.4}$$

of the pdf f of F in all points $x \in \mathbb{R}^d$, where F is continuous.

Fundamental copula functions Analogue to the Fréchet bounds for joint cdfs [cf. 31, 35, 43], copula functions are bounded from below and above:

Theorem 2.5 ([57], Theorem 2.10.12)**.** *If C is a d-copula, then, for every $(u_1, ..., u_d) \in [0,1]^d$,*

$$\max(u_1 + ... + u_d - d + 1, 0) \leq C(u_1, ..., u_d) \leq \min(u_1, ..., u_d). \tag{2.5}$$

The Fréchet upper bound, called the *maximum copula*, is a d-copula for all $d \geq 2$ [cf. 43, Theorem 3.2]. The lower bound is a copula, called the *minimum copula*, only in the case $d = 2$ [cf. 43, Theorem 3.3] but fails to satisfy property (2) of Definition 2.1 for $d > 2$ [cf. 31, Example 5.21]. The maximum copula is the cdf of the $[0,1]^d$-valued random variable $(U, ..., U)$, where $U \sim UNF(0,1)$ [cf. 31, Section 5.1.2]. The minimum copula is the cdf of the random variable $(U, 1-U)$ on $[0,1]^2$, where $U \sim UNF(0,1)$. This characterizes perfect positive and perfect negative dependence in some sense.[4]

The minimum and the maximum copula belong to the class of *fundamental copula functions*. Another fundamental copula is the d-variate product function

$$\Pi(u_1, ..., u_d) = u_1 \cdot ... \cdot u_d. \tag{2.6}$$

It follows from Theorem 2.2 that continuous random variables $X^1, ..., X^d$ are independent, if and only if the copula $C_{X^1,...,X^d}$ of $X^1, ..., X^d$ is (2.6). Hence Π is referred to as the *independence copula*.

Figure 2.1 shows the maximum, the independence and the minimum 2-copula over the unit square.

[3]Copula functions also go by the name dependence functions [21] or uniform representations [47].

[4]The technical term for perfect positive dependence (perfect negative dependence) is comonotonicity (countermonotonicity). The margins $F_i, i = 1, ..., d$ ($F_i, i = 1, 2$) are comonotone (countermonotone), if and only if they are coupled by the maximum copula (minimum copula) [cf. 16, Section 2.4]. See Joe [43] for a thorough discours on dependence concepts, measures and orderings.

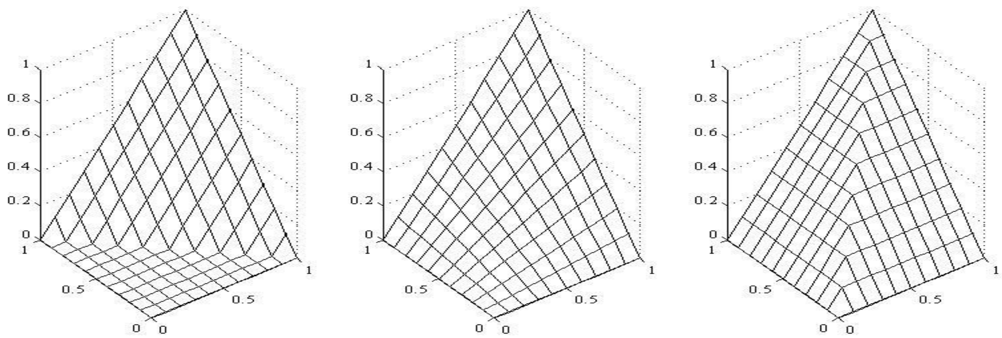

Figure 2.1: Surface of the minimum copula (left), the independence copula (center), and the maximum copula (right).

- The independence copula lays "in between[5]" the minimum and the maximum copula; the inequality (2.5) is satisfied.
- The independence copula is a smooth function; it is absolutely continuous on the unit square.
- Both the minimum and the maximum copula have a kink; they are singular on a non-empty set of points.

2.2 Elliptical copula functions

The copula functions most widely bespoken in contemporary literature are elliptical copulas. We motivate the general construction pattern and give the Gauss copula and the t copula as examples.

Using Corollary 2.3, elliptical copula functions are derived from multivariate elliptical distributions[6]:

Definition 2.6. Let $X = (X^1, ..., X^d) \sim ELL(\mu, \Sigma, \phi)$ be an elliptically distributed random variable on $\mathbb{R}^d$. Then

$$C(u_1, ..., u_d; \mu, \Sigma, \phi) = F^X((F_1^X)^{-1}(u_1), ..., (F_d^X)^{-1}(u_d)) \tag{2.7}$$

is a proper d-copula by Corollary 2.3. It is called an *elliptical copula*.

Basically, numerous copulas can be constructed by elliptical implication. The Gauss copula and the t copula have become generally accepted in this respect.

[5]Nelson [57] motivates graphically that, for a 2-copula C, the level sets $\{(u,v) \in [0,1]^2 | C(u,v) = t\}$ are contained in the region whose boundaries are the level sets determined by $\max(u+v-1, 0) = t$ and $\min(u,v) = t$.

[6]Elliptical distributions are described in Section A.3.

Gauss copula functions The Gauss copula is derived from the multivariate normal distribution with standard normal margins:

Definition 2.7. Let $\Sigma \in \mathbb{R}^{d \times d}$ be a symmetric, positiv-definite matrix with $\Sigma_{ii} = 1, i = 1, ..., n$ and Φ_Σ the centered joint normal cdf with covariance matrix Σ. Then the *Gauss copula* is defined by

$$C(u_1, ..., u_n; \Sigma) = \Phi_\Sigma(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n)), \tag{2.8}$$

where Φ^{-1} is the inverse of the standard normal cdf Φ.

As a consequence [cf. 16, Section 4.8.1], the Gauss copula is absolutely continuous on $[0, 1]^d$ with pdf

$$c(u_1, ..., u_d; \Sigma) = |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}x.(\Sigma^{-1} - \mathbf{1})x\right), \tag{2.9}$$

where $x = (x_1, ..., x_d) = (\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d))$ and $|\Sigma|$ is the determinant of Σ. Regarding Theorem 2.2, the Gauss copula may then be used in conjunction with marginal cdfs $F_1, ..., F_d$ to generate a joint cdf F with a Gauss dependence structure.[7]

The fundamental copulas are special cases of the Gauss copula [cf. 31, Section 5.1.2]. If $\Sigma \in \mathbb{R}^{d \times d}$ is equal to the identity matrix $\mathbf{1}$, then the Gauss copula degenerates to the independence copula (2.6). If else $\Sigma \in \mathbb{R}^{d \times d}$ tends to the matrix consisting entirely of plus ones, the maximum copula (2.5) is obtained in the limit. Conversely, if $\Sigma \in \mathbb{R}^{2 \times 2}$ tends to the matrix, whose off-diagonal elements are negative ones, then the minimum copula (2.5) is obtained in the limit. Hence, in two dimensions, the Gauss copula interpolates between the Fréchet bounds.

Example 2.8. Consider a random variable $U = (U^1, U^2, U^3)$ on $\mathbb{R}^3$ with uniform margins $U^i \sim UNF(0, 1), i = 1, 2, 3$. Assume that the copula C of U is the Gauss copula with correlation matrix

$$\Sigma = \begin{pmatrix} 1 & 0.7 & 0.2 \\ 0.7 & 1 & -0.5 \\ 0.2 & -0.5 & 1 \end{pmatrix}.$$

Then the joint cdf of U is given by (2.8) and the joint pdf of U is given by (2.9). Moreover, it can be shown (which we do at a later time) that the Gauss copula is closed under the taking of margins. Hence the marginal pdfs of (U^i, U^j) are given by

$$c(u_i, u_j; \rho_{ij}) = \frac{1}{\sqrt{1 - \rho_{ij}^2}} \exp\left(\frac{\rho_{ij}^2(x_i^2 + x_j^2) - 2\rho_{ij}x_ix_j}{(1 - \rho_{ij}^2)}\right), \quad i, j \in \{1, ..., 3\}, i \neq j, \tag{2.10}$$

where $x_i = \Phi^{-1}(u_i), x_j = \Phi^{-1}(u_j)$ and $\rho_{ij} = \Sigma_{ij}$. Then we may just as well view this example from a bivariate perspective, only.

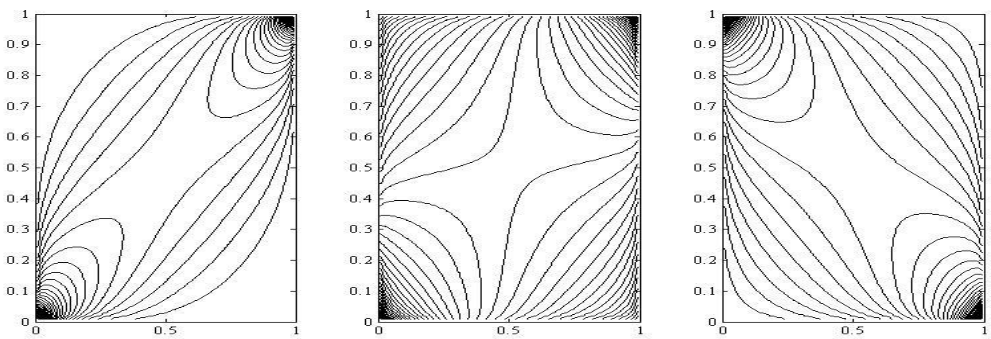

Figure 2.2: Contours of the Gauss copula density using parameters $\rho = 0.7$ (left), $\rho = 0.2$ (center), and $\rho = -0.5$ (right).

Figure 2.2 shows the contour lines of the Gauss 2-copula pdf according to Example 2.8 using various correlation coefficients ρ.[8]

- The contours of c follow, for $\rho < 0$, the diagonal with slope -1 and, for $\rho < 0$, the diagonal with slope 1; the correlation coefficient describes the sign of the dependence.
- The higher the absolute value of ρ the narrower the level curves of c; the correlation coefficient is responsible for the degree of the dependence.
- The contours of c are mirror imaged; the dependence is symmetric.

Hence the Gauss copula describes the familiar concept of linear dependence.

t copula functions The t copula function is deduced from the multivariate Student t distribution with standard Student t-distributed margins:

Definition 2.9. Let $\Sigma \in \mathbb{R}^{d \times d}$ be a symmetric, positiv-definite matrix with $\Sigma_{ii} = 1, i = 1, ..., d$ and $t_{\Sigma,\nu}$ the standard joint Student t cdf with covariance matrix Σ and ν degrees of freedom. Then the *t copula* is defined by

$$(2.11) \qquad C(u_1, ..., u_d; \Sigma, \nu) = t_{\Sigma,\nu}(t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_d)),$$

where t_ν^{-1} is the inverse of the standard Student t cdf t_ν with ν degrees of freedom.

[7] It follows from Definition 2.7 and Theorem 2.2 that the Gauss copula C generates the standard joint normal distribution $F = \Phi_\Sigma$, if and only if the margins $F_i = \Phi, i = 1, ..., d$ are taken standard normal [cf. 16, Proposition 4.1].

[8] Figure 2.3 is comparable to Bouyé et al. [13, Figure 13], who plotted the contours of the Gauss copula density using both uniform and non-uniform marginal distributions.

As a consequence [cf. 16, Section 4.8.2], the t copula function is absolutely continuous on $[0,1]^d$ with pdf

$$c(u_1,...,u_d;\Sigma,\nu) = |\Sigma|^{-\frac{1}{2}} \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})} \left[\frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+1}{2})}\right]^d \frac{(1+\frac{x.\Sigma^{-1}x}{\nu})^{-\frac{\nu+d}{2}}}{\prod_{i=1}^d (1+\frac{x_i^2}{\nu})^{-\frac{\nu+1}{2}}}, \tag{2.12}$$

where $x = (t_\nu^{-1}(u_1),...,t_\nu^{-1}(u_d))'$. Regarding Theorem 2.2, the t copula may then be used in conjunction with marginal cdfs $F_1,...,F_d$ to generate a joint cdf F with a t dependence structure.[9]

Similar to the Gauss copula, if $\Sigma \in \mathbb{R}^{d\times d}$ tends to the matrix, that consists entirely of ones, then the limit is the maximum copula (2.5). But the choice $\Sigma = \mathbf{1}$ does not yield the independence copula. This is due to the fact that uncorrelated multivariate t-distributed random variables are not independent [cf. 31, Lemma 3.5]. The minimum copula is unattainable by the same argument.[10]

Example 2.10. Consider a random variable $U = (U^1, U^2, U^3)$ on $\mathbb{R}^3$ with uniform margins $U^i \sim UNF(0,1), i = 1,2,3$. Assume that the copula C of U is the t copula with ν degrees of freedom and correlation matrix

$$\Sigma = \begin{pmatrix} 1 & 0.7 & 0.2 \\ 0.7 & 1 & -0.5 \\ 0.2 & -0.5 & 1 \end{pmatrix}.$$

Then the joint cdf of U is given by (2.11) and the joint pdf of U is given by (2.12). Moreover, it can be shown (which we do at a later time) that the t copula is closed under the taking of margins. Hence the marginal pdfs of (U^i, U^j) are given by

$$c(u_1,u_2;\rho_{ij},\nu) = \frac{\Gamma(\frac{\nu+2}{2})\Gamma(\frac{\nu}{2})}{\sqrt{1-\rho_{ij}^2}\,\Gamma(\frac{\nu+1}{2})^2} \frac{\left(1+\frac{x_i^2+x_j^2-2\rho_{ij}x_ix_j}{\nu(1-\rho_{ij}^2)}\right)^{-\frac{\nu+2}{2}}}{[(1+\frac{x_i^2}{\nu})(1+\frac{x_j^2}{\nu})]^{-\frac{\nu+1}{2}}}, \quad i,j \in \{1,...,3\}, i \neq j,$$

where $x_i = t_\nu^{-1}(u_i), x_j = t_\nu^{-1}(u_j)$ and $\rho_{ij} = \Sigma_{ij}$. Then we may just as well view this example from a bivariate perspective, only.

Figure 2.3 shows the contours of the t 2-copula pdf according to Example 2.10 using various correlation coefficients ρ and degrees of freedom ν.[11]

- The correlation coefficient describes linear dependence to the same effect as in Figure 2.2.

[9] It follows from Definition 2.9 and Theorem 2.2 that the t copula C generates the standard joint t distribution $F = t_{\Sigma,\nu}$, if and only if the margins $F_i = t_\nu, i = 1,...,d$ are standard t-distributed.

[10] Bouyé et al. [13] give graphical support to the convergence of the t copula to the Gauss copula in the limit $\nu \to \infty$.

[11] Figure 2.3 is comparable to Bouyé et al. [13, Figure 14], who plotted the contours of the t copula densities using both uniform and non-uniform marginal distributions.

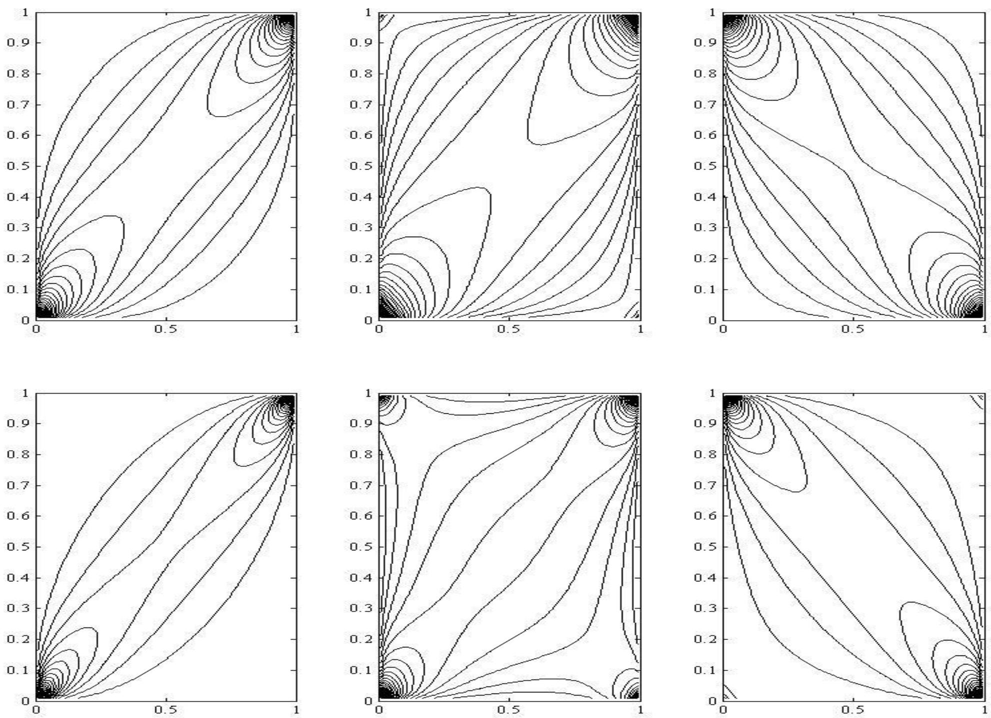

Figure 2.3: Contours of the t copula density using parameters $\rho = 0.7, \nu = 10$ (upper left),$\rho = 0.2, \nu = 10$ (upper center), $\rho = -0.5, \nu = 10$ (upper right), $\rho = 0.7, \nu = 3$ (lower left),$\rho = 0.2, \nu = 3$ (lower middle), and $\rho = -0.5, \nu = 3$ (lower right).

- The lower the ν the denser the contours of c in the corners; the degrees of freedom determine the tail dependence[12].

Hence the t copula allows to describe symmetric linear and extreme dependence.

2.3 Pair copula functions

The pair copula approach represents a radically new way of modelling multivariate copulas with highly complex dependence structures. We elaborate the involved building block pattern and go into details about two specific pair copula constructions.

The decomposition of a multivariate copula into simple blocks allows for a flexible but parsimonious dependence structure. We indicate the pattern in the three dimensional case:

[12]Joe [43] says that, if a bivariate copula C is such that $\lim_{u\to 1} C(u,u)/(1-u) = \lambda$ exists, then C has upper tail dependence for $\lambda \in (0,1]$ and no upper tail dependence for $\lambda = 0$. Similarly, if $\lim_{u\to 0} C(u,u)/u = \lambda$ exists, then C has lower dependence for $\lambda \in (0,1]$ and no upper tail dependence for $\lambda = 0$. And Bouyé et al. [13] gives graphical support to tail dependence of the t copula relating to the degrees of freedom ν.

Example 2.11 (cf. [1], Section 2)**.** Consider a random variable $X = (X^1, X^2, X^3)$ on $\mathbb{R}^3$ with joint pdf f. Then

$$f(x_1, ..., x_d) = f(x_3)f(x_2|x_3)f(x_1|x_2, x_3).$$

It follows from (2.4) that

$$f(x_2|x_3) = c_{2,3}(F_2(x_2), F_3(x_3))f_2(x_2)$$

for some bivariate copula density $c_{2,3}$. Similarly, we have

$$f(x_1|x_2, x_3) = c_{1,3|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2))f_{1|2}(x_1|x_2),$$

where $c_{1,3|2}$ another 2-copula density. The conditional pdf $f_{1|2}(x_1|x_2)$ admits a pair copula representation similar to (2.13). Altogether

$$\begin{aligned} f(x_1, x_2, x_3) &= f(x_1)f(x_2)f(x_3) \\ &\cdot\ c_{1,2}(F(x_1), F(x_2))c_{2,3}(F(x_2), F(x_3)) \\ &\cdot\ c_{1,3|2}(F(x_1|x_2), F(x_3|x_2)). \end{aligned} \tag{2.13}$$

Hence the joint density decomposes into bivariate building blocks, so-called *pair copulas*.

For general random variables X on $\mathbb{R}^d$, the pdf f of X allows of a similar decomposition. This involves the conditional pdfs

$$f(y|z) = c_{y,z_j|z_{-j}}(F(y|z_{-j}), F(z_j|z_{-j}))f(y|z_{-j}),$$

where y is a single component and z is some conditioning vector.[13] Here z_j is an arbitrary component of z and z_{-j} denotes the vector z excluding component z_j. Each conditional pdf consists of a pair copula pdf $c_{y,z_j|z_{-j}}$ and two conditional marginal cdfs of the form

$$F(y|z) = \frac{\partial C_{y,z_j|z_{-j}}(F(y|z_{-j}), F(z_j|z_{-j}))}{\partial F(z_j|z_{-j})}, \tag{2.14}$$

where $C_{y,z_j|z_{-j}}$ are again bivariate copulas. By recursive use of formula (2.14), the conditioning argument goes one-dimensional in the end. This leads to the (unconditional) margin representations

$$F(y|z) = \frac{\partial C_{y,z}(F(y), F(z))}{\partial F(z)}, \tag{2.15}$$

where $C_{y,z}$ is a bivariate copula.

The iterative design depends on the choice of the vector z and the (dropped) component z_j in each cascade. This offers many different ways to factorize the original density. These can be tracked by graph theoretic tools, called *vines*[14]. For appropriate choices (vines), the pair copula decomposition becomes very tractable:

[13] Joe [cf. 43, Chapter 9] gives a similar construction via conditional density specifications detailing distributions in exponential families without the use of copula functions.

[14] Bedford and Cooke [10] introduce the concept of vines, which is taken up by Aas et al. [1] for the pair copula construction pattern.

Definition 2.12. Let $X = (X^1, ..., X^d)$ be a random variable on $\mathbb{R}^d$ with pdf f. The *decomposition of f corresponding to a D-vine* (or D-vine decomposition of f) is given by [cf. 1, Section 2]
(2.16)
$$f(x_1, ..., x_d) = \prod_{k=1}^{d} f(x_k) \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,i+j}(F(x_i|x_{i+1}, ..., x_{i+j-1}), F(x_{i+j}|x_{i+1}, ..., x_{i+j-1})),$$

where the conditional marginal cdfs are evaluated according to

$$(2.17) \quad F(x_j|x_i, ..., x_{j-1}) = \frac{\partial C_{j,i|i+1,...,j-1}(F(x_j|x_{i+1}, ..., x_{j-1}), F(x_i|x_{i+1}, ..., x_{j-1}))}{\partial F(x_i|x_{i+1}, ..., x_{j-1})}$$
$$\text{and } F(x_j|x_{j+1}, ..., x_i) = \frac{\partial C_{j,i|j+1,...,i-1}(F(x_j|x_{j+1}, ..., x_{i-1}), F(x_i|x_{j+1}, ..., x_{i-1}))}{\partial F(x_i|x_{j+1}, ..., x_{i-1})}.$$

In the case $d = 3$, the D-vine decomposition of f coincides with (2.13), while f admits the decomposition [cf. 1, p.6]

$$\begin{aligned} f(x_1, x_2, x_3, x_4) &= f(x_1)f(x_2)f(x_3)f(x_4) \\ &\cdot\ c_{1,2}(F(x_1), F(x_2))c_{2,3}(F(x_2), F(x_3))c_{3,4}(F(x_3), F(x_4)) \\ &\cdot\ c_{1,3|2}(F(x_1|x_2), F(x_3|x_2))c_{2,4|3}(F(x_2|x_3), F(x_4|x_3)) \\ &\cdot\ c_{1,4|2,3}(F(x_1|x_2, x_3), F(x_4|x_2, x_3)) \end{aligned}$$

in the case $d = 4$.

Definition 2.13. Let $X = (X^1, ..., X^d)$ be a random variable on $\mathbb{R}^d$ with pdf f. The *decomposition of f corresponding to a canonical vine* (or canonical vine decomposition of f) is given by [cf. 1, Section 2]

$$(2.18) \quad f(x_1, ..., x_d) = \prod_{k=1}^{d} f(x_k) \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{j,j+i}(F(x_j|x_1, ..., x_{j-1}), F(x_{j+i}|x_1, ..., x_{j-1})),$$

where the marginal conditional cdfs are evaluated according to

$$(2.19) \quad F(x_j|x_1, ..., x_{i-1}) = \frac{\partial C_{j,i-1|1,...,i-2}(F(x_j|x_1, ..., x_{i-2}), F(x_{i-1}|x_1, ..., x_{i-2}))}{\partial F(x_1|x_2, ..., x_{i-1})}$$

In the case $d = 3$, the canonical vine decomposition of f coincides with (2.13), while f admits the decomposition [cf. 1, Section 2]

$$\begin{aligned} f(x_1, x_2, x_3, x_4) &= f(x_1)f(x_2)f(x_3)f(x_4) \\ &\cdot\ c_{1,2}(F(x_1), F(x_2))c_{1,3}(F(x_1), F(x_3))c_{1,4}(F(x_1), F(x_4)) \\ &\cdot\ c_{2,3|1}(F(x_2|x_1), F(x_3|x_1))c_{2,4|1}(F(x_2|x_1), F(x_4|x_1)) \\ &\cdot\ c_{3,4|1,2}(F(x_3|x_1, x_2), F(x_4|x_1, x_2)) \end{aligned}$$

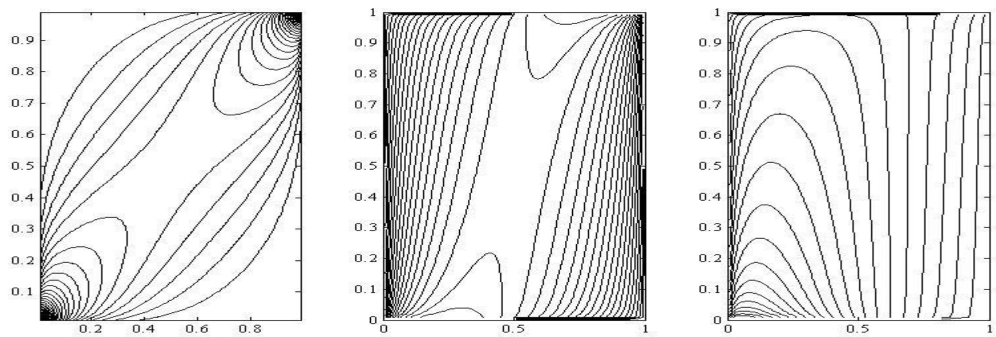

Figure 2.4: Contours of the (conditional) pair copula density using parameters $\rho_{1,2} = 0.7$ (left), $\rho_{1,3|2} = 0.2, u_2 = 0.5$ (middle), and $\rho_{1,3|2} = 0.2, u_2 = 0.9$ (right).

in the case $d = 4$. In each case, (2.17) and (2.19) reallocate the pair copulas used, respectively, in (2.16) and (2.18).

Having discussed the decomposition of general (continuous) distributions, we concern the pair copula construction on its own right. By (2.1), copula functions are recognized as joint cdfs with uniform margins. Hence, if $F(u) = u$ in (2.15), then (2.16) and (2.18) produce copula pdfs corresponding to a D-vine and a canonical vine decomposition, respectively. We write

$$h(y, z) = \frac{\partial C_{y,z}(y, z)}{\partial z} \tag{2.20}$$

instead of (2.15), whenever the margins are uniform. For some bivariate copula functions, (2.20) admits a closed form representation [cf. 1, Appendix B ff.]:

Example 2.14. Consider a random variable $U = (U^1, U^2, U^3)$ on $\mathbb{R}^3$ with uniform margins $U^i \sim UNF(0, 1), i = 1, 2, 3$. Regarding (2.13) and (2.20), the pdf c of U decomposes into

$$c_{1,2,3}(u_1, u_2, u_3) = c_{1,2}(u_1, u_2)c_{2,3}(u_2, u_3)c_{1,3|2}(h(u_1, u_2), h(u_2, u_3)). \tag{2.21}$$

Assume that all bivariate building blocks have a Gauss 2-copula pdf as defined in (2.10), that is

$$c_{12}(u_1, u_2) = c(u_1, u_2; \rho_{12}), \quad c_{23}(u_2, u_3) = c(u_2, u_3; \rho_{23}), \quad c_{13|2}(u_1, u_3) = c(u_1, u_3; \rho_{13|2}).$$

Then it follows from (2.8) [cf. 1, Section C.1] that

$$h(u_1, u_2) = \Phi\left(\frac{\Phi^{-1}(u_1) - \rho_{12}\Phi^{-1}(u_2)}{1 - \rho_{12}^2}\right) \quad \text{and} \quad h(u_2, u_3) = \Phi\left(\frac{\Phi^{-1}(u_2) - \rho_{23}\Phi^{-1}(u_3)}{1 - \rho_{23}^2}\right).$$

Figure 2.4 shows the contours of the copula pdfs of the building blocks, that are involved in Example 2.14.[15]

[15]Figure 2.4 is comparable to Aas et al. [1, Figure 9], who plotted the contours of the pair copula densities using empirical data of the Norwegian stock index, the MSCI world stock index, the Norwegian bond index and the SSBWG hedged bond index.

- The contours of $c_{1,2}$ (or $c_{2,3}$, which we omit in good faith) are identical with Figure 2.2; the unconditioned building blocks are Gauss copulas.
- The level curves of $c_{1,3|2}$ follow a dislocated diagonal; the conditioned building block describes conditional linear dependence.
- The level curves of $c_{1,3|2}$ conditioned on a centered u_2 are mirror imaged; the Gauss copula blocks produce symmetric dependence.

2.4 Summary

In this chapter we argued for the convenience of ordinary copula functions in decoupling the dependence structure and the margins of a joint distribution. Then we analyzed the elliptical and the pair copula family.

We showed at the beginning that copula functions, which are multivariate cdfs with standard uniform margins, cope with isolating the dependence structure between some univariate distributions. Moreover, the characterization of the dependence structure of a multivariate random variable turned out to be independent of the distributional shape of its margins.

Then we derived the Gauss and the t copula function implicitly from their eponymous multivariate elliptical distributions. Both copulas were characterized by their parametric pdfs, which are available in closed form due to construction. By graphical illustration of the pdf contours, we recognized that the Gauss copula can handle linear dependence only, while the t copula allows us to model tail dependence.

After motivating a conditional decomposition of general pdfs, we formulated two specific patterns, that can be used to construct parametric multivariate copula pdfs from cascades of some bivariate building blocks. Here we have gone into details about the complex but parsimonious allocation of the cascades. We illustrated the pdf contours of the two decompositions using Gauss building block copulas. Here the concept of linear dependence was rediscovered, although the association of the Gauss parameters were not so obvious.

Chapter 3

Simulation of copula functions

In this chapter we explain the basic techniques for the simulation of random scenarios from an ordinary copula set up. These are subsequently applied to the elliptical and the pair copula fucntions.

Section 3.1 describes the transformation method for elliptical copulas. Here the approach is to use stochastic representations of elliptically distributed random variables. Section 3.2 gives the conditional sampling method, which is generally applicable. The recursive pattern used to construct pair copulas is very much suited to this approach. Detailed algorithms are given and implemented.

The procedures discussed in this chapter follow closely along the methods described in Aas et al. [1], Cherubini et al. [16] and Embrecht et al. [31].

3.1 Transformation method

The implicit construction of an elliptical copula makes the sampling from it very efficient. The prevailing approach uses stochastic representations of elliptical distributions. We describe the method by instance of the Gauss and the t copula.

Elliptically distributed random variables can be represented by affine transformations of a uncorrelated random variables:[1]

Lemma 3.1 (cf. [34], Section 2.5)**.** *Let $X \sim ELL(\mu, \Sigma, \phi)$ with Σ nonsingular. Then X can be represented as*

$$X \overset{d}{=} \mu + \Sigma^{1/2} Y,$$

where $Y \sim SPH(\phi)$.

Lemma 3.1 shows that a random variable X on $\mathbb{R}^d$ with an elliptical distribution is reducible to some spherically distributed random variables $Y^1, ..., Y^d$. Hence univariate

[1]For our purposes, we introduced elliptical distributions using non-singular matrices $\Sigma \in \mathbb{R}^{d \times d}$ only. [34] is a thorough study on more general elliptical distributions based on potentially singular covariance matrices.

number generators can be used to draw from the dependent random variable. As to (2.7), the $\mathbb{R}^d$-valued random variable $U = (U^1, ..., U^d)$, which obtains from $X \sim ELL(\mu, \Sigma, \phi)$ by $U_i = F_i^X(X^i), i = 1, ..., d$, has an elliptical copula distribution. Then the simulation of elliptical copula distributions returns to a transformation of spherically distributed random draws.

Gauss copula simulation Lemma 3.1 applies to a multivariate normally distributed random variable:

Corollary 3.2. *Let $X \sim MVN(\mathbf{0}, \Sigma)$ be a normally distributed random variable on $\mathbb{R}^d$. Then*

$$X \stackrel{d}{=} \Sigma^{1/2} Z,$$

where Z is a random variable on $\mathbb{R}^d$ with independent standard normal margins $Z_i \sim MVN(0,1), i = 1, ..., d$.

Proof. Corollary 3.2 follows directly from Lemma 3.1. □

By (2.8), the cdf of the random variable $U = (U^1, ..., U^d)$ on $\mathbb{R}^d$ with $U^i = \Phi(X^i), i = 1, ..., d$, where X is as in Corollary 3.2, is the Gauss copula function with covariance matrix Σ. This is exploited in Algorithm 1, which can be used to generate a draw from the Gauss copula.

```
Algorithm 1: Sampling from Gauss copula
Samples (u_1, ..., u_d) from the Gauss copula with correlation matrix Σ.;
  Sample z_i, i = 1, ..., d independent normal;
  x = Σ^{1/2} z;
  for i = 1 to d do
      u_i = Φ(x_i);
  end
```

We implement Algorithm 1 in the context of Example 2.8. Figure 3.1 shows the marginal pairs of 5000 simulated draws from the Gauss 3-copula.[2]

- The sampled pairs follow, for $\rho < 0$, the diagonal with slope -1 and, for $\rho > 0$, the diagonal with slope 1; the concept of linear dependence is preserved in the simulation.
- There is only moderate clustering of points in the respective corners; the draws are not tail dependend.

Hence the analysis of the Gaussian copula density of Section 2.2 is reaffirmed.

[2]Figure 3.1 is comparable with Embrecht et al. [31, Figure 5.3 (a)], who also plots Gauss copula draws using $\rho = 0.7$

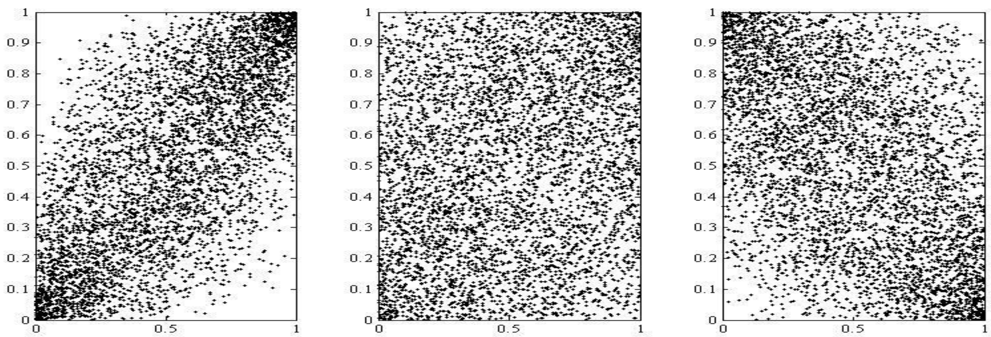

Figure 3.1: 5000 samples from the Gauss 3-copula function using parameters $\rho_{1,2} = 0.7, \rho_{2,3} = -0.5, \rho_{1,3} = 0.2$ in cross section of $(1,2)$-margin (left), $(1,3)$-margin (center), and $(2,3)$-margin (right).

t copula simulation A multivariate t-distributed random variable can be represented as a transformed multivariate normally distributed random variable:

Lemma 3.3 (cf. Fang et al. [34], Example 2.5)**.** *Let $X \sim MVT(\mu, \Sigma, \nu)$ be a t-distributed random variable on $\mathbb{R}^d$. Then*

$$X \stackrel{d}{=} \sqrt{\frac{\nu}{S}} Y,$$

where $Y \sim MVN(\mu, \Sigma)$ is a normally distributed random variable on $\mathbb{R}^d$, and $S \sim CHI2(\nu)$ is chi-squared distributed and independent of Y.

Then a multivariate t-distributed random variable is reducible to independent standard normally distributed random variables:

Corollary 3.4. *Let $X \sim MVT(\nu, \mathbf{0}, \Sigma)$ be a t-distributed random variable on $\mathbb{R}^d$. Then*

$$X \stackrel{d}{=} \sqrt{\frac{\nu}{S}} \Sigma^{1/2} Z,$$

where Z is a random variable on $\mathbb{R}^d$ whose margins $Z_i \sim MVN(0,1), i = 1, ..., d$ are independent standard normal, and $S \sim CHI2(\nu)$ is chi-squared distributed and independent of $Z_i, i = 1, ..., d$.

Proof. Corollary 3.4 follows directly from Lemma 3.1 and Lemma 3.3. □

Regarding (2.11), the cdf of the random variable $U = (U^1, ..., U^d)$ on $\mathbb{R}^d$ with $U^i = t_\nu(X^i), i = 1, ..., d$, where X is as in Corollary 3.4, is the t copula function with covariance matrix Σ and ν degrees of freedom. This is exploited in Algorithm 2, that can be used to generate a random draw from the t copula.

Algorithm 2: Sampling from t copula

Samples $(u_1, ..., u_d)$ *from the t copula with correlation matrix* Σ *and* ν *degrees of freedom.*;

Sample $z_i, i = 1, ..., d$ independent t;
Sample s from $CHI2$ independent of $z_1, ..., z_d$;
$y = \Sigma^{1/2} z$;
$x = \frac{\sqrt{\nu}}{\sqrt{s}} y$;
for $i = 1$ *to* d **do**
  $u_i = t_\nu(x_i)$;
end

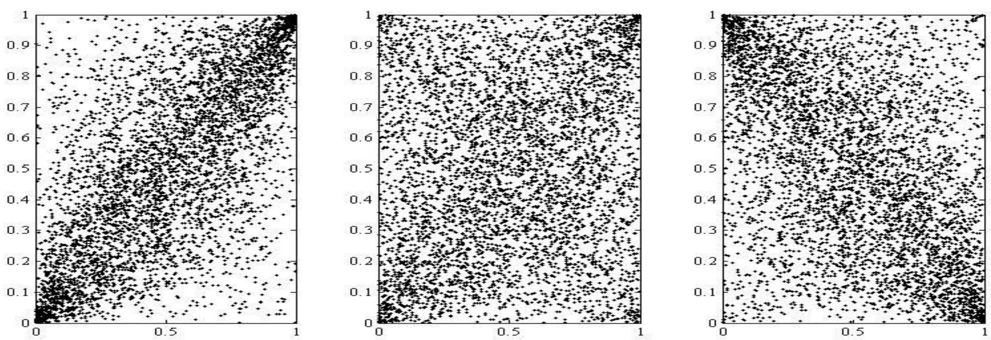

Figure 3.2: 5000 samples from the t 3-copula function using parameters $\nu = 3, \rho_{1,2} = 0.7, \rho_{2,3} = -0.5, \rho_{1,3} = 0.2$ in cross section of $(1,2)$-margin (left), $(1,3)$-margin (center), and $(2,3)$-margin (right).

We implement Algorithm 2 in the context of Example 2.10. Figure 3.2 shows the marginal pairs of 5000 simulated draws from the t 3-copula.[3]

- The sampled pairs follow, for $\rho < 0$, the diagonal with slope -1 and, for $\rho > 0$, the diagonal with slope 1; the concept of linear dependence is preserved in the simulation.
- There is heavy clustering of points in the respective corners; the draws are severely tail dependend.
- The clustering of points is still present in the case $|\rho| \ll 1$; the dependence is strong even though correlation is not.

These results coincide with the observations about the t copula density in Section 2.2.

[3]Figure 3.2 is comparable with Embrecht et al. [31, Figure 5.3 (d)], who plots t copula draws using $\rho = 0.71$ and $\nu = 4$.

3.2 Conditional sampling

In the context of copulas, the most concentraded simulation method is conditional sampling. We state the general procedure, following which we point out why conditional sampling is very well suited to the pair copula approach.

The conditional sampling approach involves subsequent probability transformations by conditional margin distributions:

Proposition 3.5 (cf. [66]). *Let $X = (X^1, ..., X^d)$ be a random variable on $\mathbb{R}^d$ with absolutely continuous cdf F and $F_{i|1,...,i-1}$ denote the conditional cdf of X_i given $X^1, ..., X^{i-1}$ for all $i = 1, ..., d$. Consider the d transformations $T_i : \mathbb{R} \to \mathbb{R}$ defined by*

$$\begin{aligned} T_1(x_1) &= F_1(x_1) \\ T_2(x_2) &= F_{2|1}(x_2|x_1) \\ \vdots &= \vdots \\ T_d(x_d) &= F_{d|1,...,d-1}(x_d|x_1, ..., x_{d-1}). \end{aligned}$$

Then the random variables $Z^i = T_i(X^i), i = 1, ..., d$ are uniformly and independently distributed on $[0, 1]^d$.

Conversely, if $Z^i, i = 1, ..., d$ are uniformly and independently distributed random variables and F is a continuous cdf on $\mathbb{R}^d$, then the random variable $X = (X^1, ..., X^d)$, which is successively defined by $X^i = T_i^{-1}(Z^i), i = 1, ..., d$, has cdf F.

Proposition 3.5 can be seen as a multivariate extension of the probability integral transform. This involves the conditional cdfs of the original distribution. For a $\mathbb{R}^d$-valued random variable U with continuous (copula) distribution C, the conditional cdf of U^i given $U^1, ..., U^{i-1}$ is [cf. 16, Section 6.3]

$$C_{i|1,...,i-1}(u_i|u_1, ..., u_{i-1}) = \frac{[\partial^{i-1}C_{1,...,i}(u_1, ..., u_i)]/[\partial u_1...\partial u_{i-1}]}{[\partial^{i-1}C_{1,...,i-1}(u_1, ..., u_{i-1})]/[\partial u_1...\partial u_{i-1}]}. \tag{3.1}$$

We assume the nominator and denominator in (3.1) to exist and the denominator to be nonzero. Then Proposition 3.5 (2nd part) and (3.1) prepare to generate random numbers from a copula distribution in the following way:

(1) Sample $z_1, ..., z_d$ independent uniform on $[0, 1]$.

(2) Define $u_1 = z_1$.

(3) For $i = 2, ..., d$, set $u_2 = C^{-1}_{i|1,...,i-1}(z_i|u_1, ..., u_{i-1})$.

The conditional approach is very elegant, but it may not be possible to compute $C^{-1}_{i|1,...,i-1}$ analytically.[4] In this case, numerical root finders are required. This may be computationally intensive, for (3.1) involves high order derivatives.[5]

[4]Cherubini et al. [16] makes the conditional sampling procedure explicit for Archimedean copula families including the Clayton copula, the Gumbel copula and the Frank copula.

[5]Whelan [83] finds an elegant way to draw from multidimensional Archimedean copula functions in

Pair copula sampling The pair copula construction allows us to manage the simulation without computing high order derivatives:

Example 3.6 (cf. [1])**.** Reconsider the random variable (U^1, U^2, U^3) on $\mathbb{R}^3$ as specified in Example 2.14. The goal is to sample a realization (u_1, u_2, u_3) from (U^1, U^2, U^3) by the conditional method. This requires the following steps. First, sample z_1, z_2, z_3 independent uniform on $[0, 1]$. Then set $u_1 = z_1$. By (2.14), we have $C_{2|1}(u_2|u_1) = h(u_2, u_1)$. Hence put $u_2 = h^{-1}(z_2, u_1)$, where [cf. 1, Section B.1]

$$h^{-1}(u_2, u_1) = \Phi\left(\Phi^{-1}(u_2)\sqrt{1-\rho_{12}^2} + \rho_{12}\Phi^{-1}(u_1)\right).$$

It holds $C_{3|1,2}(u_3|u_1, u_2) = h(h(u_3, u_1), h(u_2, u_1))$ by recursion of (2.14). Hence put $u_3 = h^{-1}(h^{-1}(z_3, h(u_2, u_1)), u_1)$, where [cf. 1, Section B.1]

$$\begin{aligned} h^{-1}(u_3, u_1) &= \Phi\left(\Phi^{-1}(u_3)\sqrt{1-\rho_{13|2}^2} + \rho_{13|2}\Phi^{-1}(u_1)\right) \\ \text{and} \quad h^{-1}(u_3, u_2) &= \Phi\left(\Phi^{-1}(u_3)\sqrt{1-\rho_{23}^2} + \rho_{23}\Phi^{-1}(u_2)\right). \end{aligned}$$

For general random variables U on $\mathbb{R}^d$ with cdf C, the pair copula decomposition reduces the conditional cdfs (2.14) (or (3.1)) to some cascades of bivariate conditional cdfs (2.20). Hence one way to invert the conditional copula cdf (3.1) is by recursive inversion of the pair building blocks (2.20). This facilitates the conditional sampling of random numbers.

The recursive pattern is again dependent on which cascade (vine) is used. In the case of a D-vine decomposition of the copula pdf as to Definition 2.12, the recursion is initiated at

$$(3.2) \qquad C_{i|1,\ldots,i-1}(u_i|u_1, \ldots, u_{i-1}) = \frac{\partial C_{1,i|2,\ldots,i-1}(F(x_i|x_2, \ldots, x_{i-1}), F(x_1|x_2, \ldots, x_{i-1}))}{\partial F(x_1|x_2, \ldots, x_{i-1})}$$

and followed according to (2.17). Aas et al. [cf. 1, Algorithm 2] propose Algorithm 3, which can be used to randomly draw from the pair copula function corresponding to a D-vine decomposition.

In the case of a canonical vine decomposition of the copula density as to Definition 2.13, the start is at

$$(3.3) \qquad C_{i|1,\ldots,i-1}(u_i|u_1, \ldots, u_{i-1}) = \frac{\partial C_{i,i-1|1,\ldots,i-2}(F(x_i|x_1, \ldots, x_{i-2}), F(x_{i-1}|x_1, \ldots, x_{i-2}))}{\partial F(x_{i-1}|x_2, \ldots, x_{i-1})}$$

and the recursion is followed according to (2.19). Aas et al. [cf. 1, Algorithm 1] have formulated Algorithm 4, which can be used to sample from the pair copula function corresponding to a canonical vine decomposition. Algorithms 3 and 4 make extensive use of the reallocation of pair copulas used for factorization.

that he partitions and scales one-dimensional draws to the right multivariate distribution. His paper complements that by Savu and Trede [70] on hierarchical Archimedean copulas.

Algorithm 3: Sampling from pair copula based on D-vine decomposition

Samples $(u_1, ..., u_d)$ *from the pair copula based on a D-vine density decomposition with (conditional) correlation matrix* Σ*.*;

Sample $w_i, i = 1, ...,$ independent uniform on $[0, 1]$;
$u_1 = z_{1,1} = w_1$;
$u_2 = z_{2,1} = h^{-1}(w_2, z_{1,1}; \theta_{1,1}$;
$z_{2,2} = h(z_{1,1}, z_{2,1}; \theta_{1,1})$;
for *i = 3 to d* **do**
    $z_{i,1} = w_i$;
    for *k = i-1 to 2* **do**
        $z_{i,1} = h^{-1}(z_{i,1}, z_{i-1,2k-2}; \theta_{k,i-k})$;
    end
    $z_{i,1} = h^{-1}(z_{i,1}, z_{i-1,1}; \theta_{1,i-1})$;
    $u_i = z_{i,1}$;
    if *i=d* **then**
        Stop;
    end
    $z_{i,2} = h(z_{i-1,1}, z_{i,1}; \theta_{1,i-1})$;
    $z_{i,3} = h(z_{i,1}, z_{i-1,1}; \theta_{1,i-1})$;
    if $i > 3$ **then**
        for *j=2 to i-2* **do**
            $z_{i,2j} = h(z_{i-1,2j-2}, z_{i,2j-1}; \theta_{j,i-j})$;
            $z_{i,2j+1} = h(z_{i,2j-1}, z_{i-1,2j-2}; \theta_{j,i-j})$;
        end
    end
    $z_{i,2i-2} = h(z_{i-1,2i-4}, z_{i,2i-3}; \theta_{i-1,1})$;
end

Algorithm 4: Sampling from pair copula based on canonical vine decomposition

Samples $(u_1, ..., u_d)$ *from the pair copula based on a canonical vine density decomposition with (conditional) correlation matrix* Σ.;

```
Sample w_i, i = 1, ..., independent uniform on [0,1];
u_1 = z_{1,1} = w_1;
for i = 2 to d do
    z_{i,1} = w_i;
    for k = i-1 to 1 do
        z_{i,1} = h^{-1}(z_{i,1}, z_{k,k}; θ_{k,i-k});
    end
    u_i = z_{i,1};
    if i=d then
        Stop;
    end
    for j=1 to i-1 do
        z_{i,j+1} = h(z_{i,j}, z_{j,j}; θ_{j,i-j});
    end
end
```

We implement Algorithm 3 in the context of Example 3.6. Figure 3.3 shows the scattered (unconditional) marginal pairs of 5000 random draws.[6]

- The clustering of points is moderate and symmetric; the Gauss building blocks induce normal dependence only.
- The $(1, 2)$-margin and the $(2, 3)$-margin pairs are comparable with Figure 3.1.
- The $(1, 3)$-margin draws can not be associated exactly with Figure 3.1.

In conclusion, the pair copula construction (using Gauss building blocks) describes a dependence structure that is very similar to the Gauss copula in terms of symmetry and extremes. However it reveals a more complex association between the copula parameters and the resulting dependence structure.

3.3 Summary

In this chapter we developed explicit methods for the convenient simulation of elliptical and pair copula distributed random variables.

[6] Figure 3.3 is comparable with the scatter plots of some empirical data sets provided by Aas et al. [cf. 1, Figure 10].

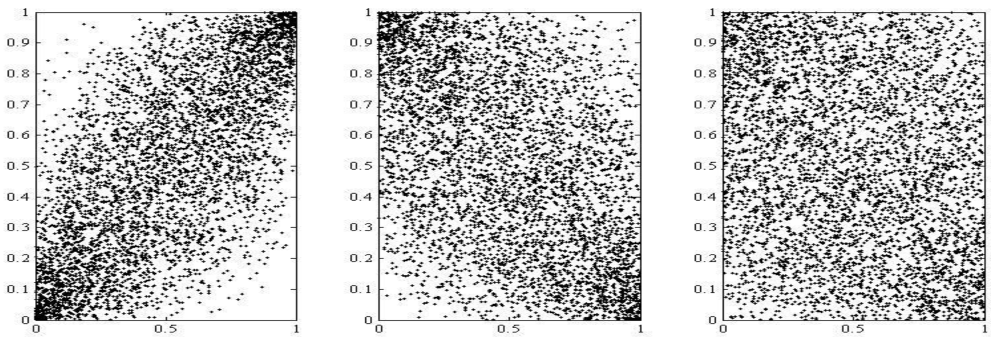

Figure 3.3: 5000 samples from the pair 3-copula function using parameters $\rho_{1,2} = 0.7, \rho_{2,3} = -0.5, \rho_{1,3|2} = 0.2$ in cross section of $(1,2)$-margin (left), $(2,3)$-margin (center), and $(1,3)$-margin (right).

Regarding the implicit construction of elliptical copula functions, we proved a transformation method to be useful to sample from the Gauss and the t copula. This relied on stochastic representations of normally and Student t distributed random variables.

Then we sampled from the two elliptical copulas using the derived algorithms and rediscovered the properties of the copula pdf in both cases. Specifically, we observed from the simulations that the t copula function gives rise to scenarios, which are more severe in the extremes as opposed to the Gauss copula draws.

We found out that the conditional sampling approach is the method of choice for the generation of random draws from a pair copula distribution. In this respect we argued that the building block principle suits very well a conditional approach. The involved but explicit sampling algorithms were written out.

Then we revisited the example of a pair 3-copula function constructed from Gauss building blocks. Here the properties of the copula pdf turned out to be reinforced by the simulation. Specifically, we observed that the samples have an evidently ample but moderate dependence structure whose association with the Gauss parameters is somewhat hidden.

Chapter 4

Inference for copula functions

This chapter is devoted to the problem of estimating the model parameters from a set of realizations. We are keenly interested in the testing of these estimators, too.

Section 4.1 describes the maximum likelihood estimation for a copula based model in its canonical version. This is elaborated for the elliptical and the pair copula functions. Then Section 4.2 deals with the testing of some given estimates. The goal is to develop efficient inference methods to assess how good an estimated model fits the data. This is straight forward in the case of pair copulas but involved in the case of elliptical copulas. Detailed algorithms are given and implemented.

The methods treated in this chapter follow the standard references Cherubini et al. [16] and Embrecht et al. [31] as to parameter estimation. The goodness-of-fit test is refined on the basis of Fang et al. [34], Embrecht et al. [33].

4.1 Maximum likelihood method

Maximum likelihood is the method of choice to estimate the parameters of a copula based model. We elaborate the general principle and particularize a canonical version by instance of the elliptical and the pair copulas.

Let $\aleph = \{x_{1,t}, ..., x_{d,t}\}_{t=1}^{T}$ denote a *sample data matrix*, which contains independent observations from the $\mathbb{R}^d$-valued random variable $X = (X^1, ..., X^d)$. Hence, for each $t \in \{1, ...T\}$, $(x_{1,t}, ..., x_{d,t})$ is a realization of $(X^1, ..., X^d)$ and independent from $(x_{1,s}, ..., x_{d,s})$, $s \neq t$. This assumption will be referred to as the *usual assumptions*. For the cdf F of X, assume a copula based model of the form

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d)),$$

where F_i is the cdf of X^i for each $i = 1, ..., d$ and C is a copula function. Introduce the set $\Theta = \Theta_0 \times \Theta_1 \times ... \times \Theta_d$ of all admissible model parameters θ. Hence the copula function and the marginal distribution functions depend on θ:

$$(4.1) \qquad C(u_1, ..., u_d) = C(u_1, ..., u_d; \theta_0) \quad \text{and} \quad F_i(x_i) = F_i(x_i; \theta_i), i = 1, ..., d.$$

By (2.4), we have

$$f(x_1,...,x_d;\theta) = c(F_1(x_1;\theta_1),...,F_d(x_d;\theta_d);\theta_0)\prod_{i=1}^{d} f_i(x_i;\theta_i).$$

It follows from the independence of the $(x_{1,t},...,x_{d,t}), t = 1,...,T$ that the *likelihood* (or likelihood function) $L : \Theta \to \mathbb{R}$ of the observed data $\aleph$ under model (4.1) with parameter θ is obtained as

$$L(\theta) = \prod_{t=1}^{T} f(x_{1,t},...,x_{d,t};\theta).$$

The likelihood function measures the probability of the event $\aleph$, if θ was the true set of parameters. Then a maximum likelihood estimator tries to find the optimal set of parameters $\theta \in \Theta$, which maximizes $L(\theta)$. It is convenient to consider the logarithmic transformation of the likelihood function:

Definition 4.1. Let $X = (X^1,...,X^d)$ be a random variable on $\mathbb{R}^d$ with joint pdf $f(\cdot;\theta)$, where $\theta \in \Theta$ an admissible set of model parameters. Further let $\aleph = \{x_{1,t},...,x_{d,t}\}_{t=1}^T$ be a sample data matrix of X satisfying the usual assumptions. Then the *log-likelihood* (or log-likelihood function) $l : \Theta \to \mathbb{R}$ is defined as

$$l(\theta) = \sum_{t=1}^{T} \ln f(x_{1,t},...,x_{d,t};\theta).$$

Regarding (2.4), there exists a canonical representation of the log-likelihood function:

Proposition 4.2 (cf. [16], Section 5.2)**.** *Let $X = (X^1,...,X^d)$ be a random variable on $\mathbb{R}^d$ with cdf $F(\cdot;\theta)$, marginal cdf's $F_i(\cdot;\theta), i = 1,...,d$ and copula $C(\cdot;\theta)$, where $\theta \in \Theta$ is an admissible set of model parameters. Further let $\aleph = \{x_{1,t},...,x_{d,t}\}_{t=1}^T$ be a sample data matrix of X satisfying the usual assumptions. Then*

$$l(\theta) = \sum_{t=1}^{T} \ln c(F_1(x_{1,t};\theta_1),...,F_d(x_{d,t};\theta_d);\theta_0) + \sum_{t=1}^{T}\sum_{i=1}^{d} \ln f_i(x_{i,t};\theta_i). \tag{4.2}$$

The logarithmic function is strictly increasing. Thus the log-likelihood obtains its maximum wherever the likelihood does. The *exact maximum likelihood* estimator is defined by

$$\hat{\theta}_{EML} = \arg\max_{\theta\in\Theta} l(\theta).$$

This is the optimal set of parameters with respect to the (log-)likelihood.[1] The estimator is said to be exact, because optimization is over the whole set of model parameters Θ.

[1] We assume in this section that the maximum likelihood estimators exist, are consistent and asymptotically efficient [cf. 73, 43].

Inference for margins (4.2) distinguishes a copula density term and d marginal density terms. This motivates a split optimization strategy, which is known as the *inference for margins* method [cf. 84].

The second term on the right-hand side of (4.2) is the sum of the marginal log-likelihood functions

$$l_i(\theta_i) = \sum_{t=1}^{T} \ln f_i(x_{i,t}; \theta_i)$$

over all $i \in \{1, ..., d\}$. These are maximized, in a first step and separately, over all $\theta_i \in \Theta_i$:

$$\hat{\theta}_i = \arg \max_{\theta_i \in \Theta_i} \sum_{t=1}^{T} \ln f_i(x_{i,t}; \theta_i).$$

In a second step, the term on the right-hand side of (4.2) containing the copula density is maximized. This involves the values $F_i(x_{i,t}; \theta_i), i = 1, ..., d, t = 1, ..., T$. Here we employ the former estimates $\hat{\theta}_i$ as an approximation to the optimal θ_i. This allows to maximize the likelihood over θ_0 alone:

$$\hat{\theta}_0 = \arg \max_{\theta_0 \in \Theta_0} \sum_{t=1}^{T} \ln c(F_1(x_{1,t}; \hat{\theta}_1), ..., F_d(x_{d,t}; \hat{\theta}_d); \theta_0).$$

Then the overall[2] estimate is given by

$$\hat{\theta}_{IFM} = (\hat{\theta}_0, \hat{\theta}_1, ..., \hat{\theta}_d).$$

Canonical maximum likelihood Empirical margins are used in (4.2) whenever a parametric specification of the marginal distributions is to be avoided. This is referred to as the *canonical maximum likelihood* method.

Dropping the parametric models $F_i(\cdot; \theta_i), i = 1, ..., d$ for the marginal cdfs, we consider the empirical margins, denoted by $\hat{F}_i$,

$$\hat{F}_i(y) = \frac{1}{T} \sum_{t=1}^{T} \mathbf{1}_{x_{i,t} \leq y}.$$

These are determined by the columns of the sample data matrix $\aleph$ alone. As a consequence, (4.2) is independent of $\theta_i \in \Theta_i, i = 1, ..., d$. Hence

$$l(\theta_0) = \sum_{t=1}^{T} \ln c(\hat{u}_{1,t}, ..., \hat{u}_{d,t}); \theta_0),$$

[2] Bouyé et al. [13] refers to $\theta_1, ..., \theta_d$ as specific parameters and θ_0 as common parameters.

where $\hat{u}_{i,t} = \hat{F}_i(x_{i,t}), i = 1, ..., d, t = 1, ..., T$ are the so-called *pseudo-observations*. Then the canonical maximum likelihood estimator is given by

$$\hat{\theta}_{CML} = \arg\max_{\theta_0 \in \Theta_0} \sum_{t=1}^{T} \ln c(\hat{u}_{1,t}, ..., \hat{u}_{d,t}; \theta_0). \tag{4.3}$$

Canonical maximum likelihood estimation is sufficient for our purposes.[3] We assume the pseudo-observations $\hat{u}_{i,t}, i = 1, ..., d, t = 1, ..., T$ to be given from the sample data matrix $\aleph$ and consider the estimation of the copula model on its own.

Gauss copula estimation Using representation (2.9) of the pdf of the Gauss copula, the (canonical) log-likelihood function has a closed form representation:

Lemma 4.3 (cf. [31], Example 5.58)**.** *For the Gauss copula with correlation matrix Σ, the canonical log-likelihood function is given by*

$$l(\Sigma) = \sum_{t=1}^{T} \ln \frac{1}{\sqrt{|\Sigma|}} e^{-\frac{1}{2}\xi_t.(\Sigma^{-1}-\mathbf{1})\xi_t} = -\frac{T}{2}\ln|\Sigma| - \frac{1}{2}\sum_{t=1}^{T} \xi_t(\Sigma^{-1} - \mathbf{1}).\xi_t, \tag{4.4}$$

where $\xi_t = (\Phi^{-1}(\hat{u}_{1,t}), ..., \Phi^{-1}(\hat{u}_{d,t}))'$.

In this case, there exists an analytical solution to (4.3). The optimal symmetric positive definite correlation matrix $\hat{\Sigma}$ is

$$\hat{\Sigma}_{CML} = \frac{1}{T}\sum_{t=1}^{T} \xi_t.\xi_t\ .^4$$

where ξ_t is as in Lemma 4.4. This is exploited in Algorithm 5, that can be used to calibrate the Gauss copula to a sample data matrix.

We run Algorithm 5 on the 5000 Gaussian 3-copula draws sampled in Section 3.1. The resulting estimate is

$$\hat{\Sigma} = \begin{pmatrix} 1 & 0.7034 & -0.4946 \\ 0.7034 & 1 & 0.2009 \\ -0.4946 & 0.2009 & 1 \end{pmatrix}.$$

That is to say that the correlation matrix, which was used for simulation, is adequately regained.

[3] Bouyé et al. [13] performs a Monte-Carlo study to test the three methods on the basis of a bivariate Gauss copula with exponential and gamma distributed margins.

[4] The explicit estimate (4.5) formulates the usual linear correlation matrix (or Pearson's linear correlation) of the pseudo-sample $(\Phi^{-1}(\hat{u}_{1,t}), ..., \Phi^{-1}(\hat{u}_{d,t})), t = 1, ..., T$, Embrecht et al. [cf. 31, Example 5.53] shows further how to calibrate a Gauss copula using Spearman's rho.

Algorithm 5: CML for Gauss copula parameters

Estimates the correlation matrix Σ *of the Gauss copula for a random variable* $(X^1, ..., X^d)$ *by means of the sample* $(x_{1,t}, ..., x_{d,t})$, $t = 1, ..., T$.;

for t = *1 to* T **do**
  for i = *1 to* d **do**
    $\hat{u}_{i,t} = \hat{F}_i(x_{i,t})$;
    $\hat{\xi}_{i,t} = \Phi^{-1}(\hat{u}_{i,t})$;
  end
end
$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} \hat{\xi}_t.\hat{\xi}_t$;

t copula estimation Using representation (2.12) of the pdf of the t copula, the (canonical) log-likelihood has a closed form representation:

Lemma 4.4 (cf. [31], Example 5.59)**.** *In the case of t copula with correlation matrix* Σ *and* ν *degrees of freedom, the canonical log-likelihood function is given by*

$$l(\Sigma, \nu) \quad = \quad \sum_{t=1}^{T} \ln \frac{1}{\sqrt{|\Sigma|}} \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})} \left[\frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+1}{2})}\right]^d \frac{(1 + \frac{\xi_t.\Sigma^{-1}\xi_t}{\nu})^{-\frac{\nu+d}{2}}}{\prod_{i=1}^{d}(1 + \frac{\xi_{i,t}^2}{\nu})^{-\frac{\nu+1}{2}}}, \tag{4.5}$$

where $\xi_t = (t_\nu^{-1}(\hat{u}_{1,t}), ..., t_\nu^{-1}(\hat{u}_{d,t}))'$.

The log-likelihood function (4.5) is more involved than (4.4) and the additional degrees of freedom parameter prevents us solving (4.3) analytically. Thus we have to resort to numerical optimization procedures [cf. 31, Example 5.59]. Here the notion of Kendall's tau proves very useful:

Definition 4.5. For a random variable $X = (X^1, ..., X^d)$, the *Kendall's tau* matrix[5] is given by

$$\tau(X) = Cov(\text{sgn}(X - \tilde{X})),$$

where $\tilde{X}$ an independent copy of X (that is a second random variable on $\mathbb{R}^d$ with the same distribution but independent of the first).

In the case of elliptically distributed random variables, Kendall's tau can be related pairwise to Pearson's linear correlation coefficient:

Proposition 4.6 (cf. [31], Proposition 5.37)**.** *Let* $X \sim ELL(0, \Sigma, \psi)$ *be an elliptically distributed random variable on* $\mathbb{R}^2$ *with correlation matrix* Σ*, whose off-diagonal element is* ρ*, and assume that* $P[X = 0] = 0$*. Then it holds*

$$\rho = \sin(\pi\tau/2). \tag{4.6}$$

[5]For a $\mathbb{R}^d$-valued random variable X, Kendall's tau can be understood as a measure of concordance [cf. 31] between components.

Hence Proposition 4.6 enables us to extract the linear correlation matrix of an elliptical distribution from the Kendall's tau values, regardless of the margins (or the elliptical generator, in fact). For a sample data matrix $\aleph = \{x_{1,t}, ..., x_{d,t}\}_{t=1}^{T}$, the empirical version of Kendall's tau is given pairwise by

$$\hat{\tau}_{i,j} = \frac{2}{T(T-1)} \sum_{t=1}^{T} \sum_{s>t} A_{ts}^{ij}, \quad i, j = 1, ..., d, \tag{4.7}$$

where $A_{ts}^{ij} = \text{sgn}(x_{i,t} - x_{i,s})(x_{j,t} - x_{j,s})$. This gives an estimator[6] of the linear correlation matrix, which can be computed analytically from the sample data by (4.6) and (4.7). Then the canonical maximum likelihood method for the t copula may be described as follows:

(1) compute $\hat{\Sigma}$ from the Kendall's tau matrix of the data $\aleph$ by (4.6),

(2) estimate $\hat{\nu}$ by numerical maximization of (4.5), using the preestimate $\hat{\Sigma}$.

This is formulated in Algorithm 6, that can be used to calibrate the t copula to a sample data matrix. Here we did not dwell the numerical optimization routine.

Algorithm 6: CML for t copula parameters

Estimates the correlation matrix Σ *and the degrees of freedom* ν *of the t copula for a random variable* $(X^1, ..., X^n)$ *by means of the sample* $(x_{1,t}, ..., x_{d,t})$, $t = 1, ..., T$.;

```
for t = 1 to T do
    for i = 1 to d do
        û_{i,t} = F̂_i(x_{i,t});
    end
    for i = 1 to d do
        for j = 1 to d do
            τ̂_{i,j} = 2/(T(T−1)) Σ_{t=1}^{T} Σ_{s>t} sgn(û_{i,t} − û_{i,s})(û_{j,t} − û_{j,s});
            Σ̂_{ij} = sin(π τ̂_{ij}/2);
        end
    end
end
Maximize the log-likelihood l(Σ̂, ν) over ν numerically;
```

We run Algorithm 6 on the 5000 Student 3-copula draws sampled in Section 3.1. The resulting estimates are

$$\hat{\Sigma} = \begin{pmatrix} 1 & 0.6959 & -0.5121 \\ 0.6959 & 1 & 0.1925 \\ -0.5121 & 0.1925 & 1, \end{pmatrix} \quad \text{and} \quad \hat{\nu} = 2.9823.$$

Hence we are able to win back the correlation matrix and the degrees-of-freedom.

[6]It is not guaranteed that the transformation (4.6) retains a positive definite linear correlation matrix [see 31, Section 5.5.1, for a workaround].

Pair copula estimation The pair copula construction is premised on the decomposition of the copula density into cascades of simple building blocks. This facilitates the computation of the log-likelihood function.

In the case of a D-vine decomposition as to Definition 2.12, the log-likelihood function may be evaluated by parsimonious recursions:

Lemma 4.7 (cf. [1], Section 5.1)**.** *For a pair copula corresponding to the D-vine decomposition as in* (2.16)*, the log-likelihood function can be evaluated as to the following rule:*

$$\sum_{j=1}^{d-1}\sum_{i=1}^{d-j}\sum_{t=1}^{T} log\left(c_{i,i+j}(F(x_{i,t}|x_{i+1,t},...,x_{i+j-1,t}),F(x_{i+j,t}|x_{i+1,t},...,x_{i+j-1,t}))\right). \tag{4.8}$$

If the conditional cdf's in (4.8) are computed according to (2.17), then the reallocation of bivariate copula blocks spares much of the computational costs. This is exploited in Algorithm 7 [cf. 1, Algorithm 3], which can be used to evaluate the log-likelihood function at the sample data matrix $\aleph = \{x_{1,t},...,x_{d,t}\}_{t=1}^{T}$ in the case of a D-vine decomposition.

Algorithm 7 can now be readily used in numerical schemes in order to get the canonical maximum likelihood estimator (4.3). This is applied to the sample drawn in Section 3.2. We get

$$\rho_{1,2} = 0.7051, \quad \rho_{2,3} = 0.2036 \quad \text{and} \quad \rho_{1,3|2} = 0.5058.$$

Hence we regain the starting configuration.

In a similar way, the density decomposition according to a canonical vine as in Definition 2.12 prepares an explicit computation of the log-likelihood function:

Lemma 4.8 (cf. [1], Section 5.2)**.** *In the case of a pair copula corresponding to the canonical vine decomposition as in* (2.18)*, the log-likelihood function can be evaluated as to the following rule:*

$$\sum_{j=1}^{d-1}\sum_{i=1}^{d-j}\sum_{t=1}^{T} \log\left(c_{j,j+i}(F(x_{j,t}|x_{1,t},...,x_{j-1,t}),F(x_{j+i,t}|x_{1,t},...,x_{j-1,t}))\right).$$

If the conditional cdf's in (4.9) are computed according to (2.19), then the reallocation of bivariate copula blocks spares again much of the computational costs. This is written out in Algorithm 7 [cf. 1, Algorithm 3], which can be used to evaluate the log-likelihood function at the sample data matrix $\aleph = \{x_{1,t},...,x_{d,t}\}_{t=1}^{T}$ in the case of a canonical vine decomposition.

Algorithm 7 can now be readily used in numerical schemes in order to get the canonical maximum likelihood estimator (4.3). This is applied to the sample drawn in Section 3.2. The result is

$$\rho_{1,2} = 0.7051, \quad \rho_{2,3} = 0.2076 \quad \text{and} \quad \rho_{1,3|2} = 0.5045.$$

Hence we obtain (almost) the starting configuration.

Algorithm 7: Likelihood for pair copula based on D-vine decomposition

Computes the log-likelihood of the sample $(x_{1,t}, ..., x_{d,t})$, $t = 1, ..., T$ *for the pair copula based on a D-vine density decomposition with the set* θ *of (conditional) parameters.*;

```
L = 0;
for i = 1 to d do
    z_{0,i} = x_i;
end
for i = 1 to d − 1 do
    L = L + ln(c_{1,i}(z_{0,i}, z_{0,i+1}; θ_{1,i}));
end
z_{1,1} = h(z_{0,1}, z_{0,2}; θ_{1,1});
for k = 1 to d-3 do
    z_{1,2k} = h(z_{0,k+2}, z_{0,k+1}; θ_{1,k+1});
    z_{1,2k+1} = h(z_{0,k+1}, z_{0,k+2}; θ_{1,k+1});
end
z_{1,2d−4} = h(z_{0,d}, z_{0,d−1}; θ_{1,d−1});
for j=1 to d-j do
    for i = 1 to d-j do
        L = L + ln(c_{j,i}(z_{j−1,2i−1}, z_{j−1,2i}; θ_{j,i}));
    end
    if j=d-1 then
        Stop;
    end
    z_{j,1} = h(z_{j−1,1}, z_{j−1,2}; θ_{j,1});
    if d > 4 then
        for i=1 to d-j-2 do
            z_{j,2i} = h(z_{j−1,2i+2}, z_{j−1,2i+1}; θ_{j,i+1});
            z_{j,2i+1} = h(z_{j−1,2i+1}, z_{j−1,2i+2}; θ_{j,i+1});
        end
    end
    z_{j,2d−2j−2} = h(z_{j−1,2d−2j}, z_{j−1,2d−2j−1}; θ_{j,d−j});
end
```

Algorithm 8: Likelihood for pair copula based on canonical vine decomposition

Computes the log-likelihood of the sample $(x_{1,t}, ..., x_{d,t})$, $t = 1, ..., T$ *for the pair copula based on a D-vine density decomposition with the set* θ *of (conditional) parameters.*;

L = 0;
for *i = 1 to d* **do**
  $z_{0,1} = \mathbf{x}_i$;
end
for *j = 1 to d-1* **do**
  for *i=1 to d-j* **do**
    $L = L + \ln(c_{j,i}(z_{j-1,1}, z_{j-1,i+1}; \theta_{j,i}))$;
  end
  if *j=d-1* **then**
    Stop;
  end
  for *i=1 to d-j* **do**
    $z_{j,i} = h(z_{j-1,i+1}, z_{j-1,1}; \theta_{j,i})$;
  end
end

4.2 Goodness-of-fit test

The methods described in Section 4.1 output the best parametrization only relative to a chosen model and its likelihood[7]. We now target the estimates' absolute goodness-of-fit to the sample data on its own and ways to test it.

Suppose that a copula model of the form (4.1) has been calibrated to a sample data matrix $\aleph = \{x_{1,t}, ..., x_{d,t}\}_{t=1}^{T}$.

Remark 4.9. Proposition 3.5 defined the d transformations $T_i(x_i), i = 1, ..., d$, whereby we may obtain uniformly and independently distributed random variables $Z^i, i = 1, ..., d$ from a multivariate (dependent) random variable $X = (X^1, ..., X^d)$ on $\mathbb{R}^d$.

It follows from Theorem 2.2 [cf. 22, Section 1.6] that the Z^i's in Remark 4.9 can be represented as

$$Z^i = C_{i|1,...,i-1}(F_i(X^i)|F_1(X^1), ..., F_{i-1}(X^{i-1})). \tag{4.9}$$

[7] Other than by pure likelihood values, Dias [22] and Embrecht et al. [31] suggest a ranking of dependence models according to the Akaike information criterion [cf. 3]

$$AIC(C(\cdot;\theta_0)) = -2\exp\sum_{t=1}^{T}\ln c(\hat{u}_{1,t}, ..., \hat{u}_{d,t}; \theta_0) + 2|\theta_0|,$$

that imposes a penalty equal to the number of model parameters $|\theta_0|$.

Then the quantile transformed random variables $\Phi^{-1}(Z^i), i = 1, ..., d$ are independently and standard normally distributed, and the sum of squares $S = \sum_{i=1}^{d}(\Phi^{-1}(Z^i))^2$ has a χ^2 distribution with d degrees of freedom. The variable S may then be used to test whether (or not) the basic model (4.9) fits the data $\aleph$. This requires to compute

$$s_t = \sum_{i=1}^{d}(\Phi^{-1}(z_{i,t}))^2, \quad t = 1, ..., T, \tag{4.10}$$

as test statistics, where $z_{i,t} = C_{i|1,...,i-1}(F_i(x_{i,t})|F_1(x_{1,t}), ..., F_{i-1}(x_{i-1,t})), t = 1, ..., T, i = 1, ..., d$. Then it is to be checked whether the s_t's can be understood as samples from a χ^2 distribution.[8] Analogous to the canonical maximum likelihood method, we use the pseudo-versions

$$\hat{z}_{i,t} = C_{i|1,...,i-1}(\hat{u}_{i,t}|\hat{u}_{1,t}, ..., \hat{u}_{i-1,t})) \tag{4.11}$$

where the $\hat{u}_{i,t}$ are defined as in (4.3). Then the test requires the following steps:

(1) compute the pseudo-samples $\hat{u}_{i,t} = \hat{F}_i(x_{i,t}), i = 1, ..., d, t = 1, ..., T$

(2) evaluate the probability integral transforms (4.11)

(3) calculate the summary statistics $s_t = \sum_{i=1}^{d}(\Phi^{-1}(z_{i,t}))^2, t = 1, ..., T$

(4) test whether the $s_t, t = 1, ..., T$ have a χ^2 distribution.

Step (2) is the crucial problem here. After we achieved to avoid the explicit computation of the conditional cdf $C_{i|1,...,i-1}$ of an elliptical copula C in Section 3.1, this is our designated target in the following.

Gauss copula fit To our knowledge, the goodness-of-fit test for a Gauss copula has not yet been described in a detailed manner.[9] Hence the development of an incremental evaluation method for the conditional Gauss copula functions is fairly new.

The multivariate normal distribution is closed under the taking of margins:

Theorem 4.10 (cf. [49], Example 5.1)**.** *Let $X \sim MVN(\mu, \Sigma)$ be a normally distributed random variable on $\mathbb{R}^d$ with mean μ and covariance matrix Σ. Partition X, μ and Σ into*

$$X = \begin{pmatrix} X^1 \\ X^2 \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

with $X^1, \mu_1 \in \mathbb{R}^{m\times 1}, \Sigma_{11} \in \mathbb{R}^{m\times m}, 0 < m < d$. Then $X^i \sim MVN(\mu_i, \Sigma_{ii}), i = 1, 2$.

[8] We employ the Kolmogorov-Smirnov goodness-of-fit hypothesis test [cf. 54] of the distribution of a single sample (KS-test). The test returns whether (or not) to reject the hypothesis that the s_t's are χ^2-distributed, the asymptotic p-value and the KS-statistic $\max_{x>0}(\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}_{s_t\le x} - \chi^2(x))$.

[9] Embrecht et al. [33, 31] give the very basic properties of conditional elliptical distributions and refer the interested reader to the groundwork by Fang et al. [34].

The partition of X, μ and Σ in Corollary 4.11 is hereafter called the *usual partition* (or d by m partition, if stressing the dimensions) of X (or X, μ and Σ). In the same sense, we denote by

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

with $X_1^I, \mu_1^I \in \mathbb{R}^{m\times 1}, \Sigma_{11}^I \in \mathbb{R}^{m\times m}, 0 < m < |I|$, the $|I|$ *by* m *partition of the* I*-margin* X^I *of* X, where $X^I = X^1$, $\mu^I = \mu_1$ and $\Sigma^I = \Sigma_{11}$ from the d by $|I|$ partition of X, μ and Σ. For $I = \{1, ..., i\}, i \in \{1, ..., d\}$ in particular, we write X^i, μ^i and Σ^i for X^I, μ^I and Σ^I, respectively.

The following result shows that the multivariate normal distribution is also closed under conditioning:

Corollary 4.11 (cf. [34], Theorem 2.16 ff.)**.** *Let $X \sim MVN(\mu, \Sigma)$ be a normally distributed random variable on $\mathbb{R}^d$ with the usual partition of μ and Σ. Then the conditional distribution of X^2 given X^1 is $MVN(\mu_{2|1}, \Sigma_{2|1})$, where $\mu_{2|1}$ and $\Sigma_{2|1}$ are defined as follows:*

$$\mu_{2|1} = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X^1 - \mu_1) \tag{4.12}$$

$$\Sigma_{2|1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}. \tag{4.13}$$

Proof. The corollary follows directly from Theorem 2.16 and the following corollary in [34] through permutation of the components of $(X^1, ..., X^d)$. □

It follows from Corollary 4.11 that the Gauss copula is closed under conditioning except for a shift in the mean:

Proposition 4.12. *Let C be the Gauss d-copula with correlation matrix Σ. For $I = \{1, ..., i\}$ and the i by $i - 1$ partition of Σ^I it holds*

$$C_{i|1,...,i-1}(u_i|u_1, ..., u_{i-1}) = \Phi_{\mu_{2|1}^I, \Sigma_{2|1}^I}(\Phi^{-1}(u_i)), \tag{4.14}$$

where

$$\mu_{2|1}^I = \Sigma_{21}^I(\Sigma_{11}^I)^{-1}(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_{i-1}))' \tag{4.15}$$

$$\Sigma_{2|1}^I = \Sigma_{22}^I - \Sigma_{21}^I(\Sigma_{11}^I)^{-1}\Sigma_{12}^I. \tag{4.16}$$

Proof. Let $(U^1, ..., U^d)$ be a random variable with normal copula distribution C. For $I = \{1, ..., i\}$, it is clear [cf. 34] that the cdf of the I-margin $(U^1, ..., U^i)$ is the Gauss i-copula function with correlation matrix Σ^I. Then the transformed random variable $(X^1, ..., X^i)$ with $X^i = \Phi^{-1}(U^i), i = 1, ..., i$ has a joint normal distribution with correlation matrix Σ^I. By Corollary 4.11, the conditional distribution of X_i given $X^1, ..., X^{i-1}$ is a (univariate) normal distribution with mean (4.15) and variance (4.16). Then (4.14) holds by reversing the margin transformation. □

Algorithm 9: PIT for Gauss copula function

Computes, for $t = 1, ..., T$, *the probability integral transforms* $(z_{1,t}, ..., z_{d,t})$ *of* $(u_{1,t}, ..., u_{d,t})$ *by the Gauss copula with correlation matrix* Σ.;

for $t = 1$ *to* T **do**
  $z_{1,t} = u_{1,t}$;
  for $i = 2$ *to* d **do**
    $\mu_i = \Sigma_{21}^i (\Sigma_{11}^i)^{-1} (\Phi^{-1}(u_{1,t}), \ldots, \Phi^{-1}(u_{i-1,t}))'$;
    $\Sigma_i = \Sigma_{22}^i - \Sigma_{21}^i (\Sigma_{11}^i)^{-1} \Sigma_{12}^i$;
    $z_{i,t} = \Phi_{\mu_{2|1}^i, \Sigma_{2|1}^i}(\Phi^{-1}(u_{i,t}))$;
  end
end

Proposition 4.12 leads to Algorithm 9, that can be used to successively transform some sample data.

We implement Algorithm 9 in the context of the general procedure, using all the random draws simulated in Chapter 3. This requires to estimate beforehand the correlation matrix of the Gauss copula by applying Algorithm 5 to the Gauss, the t and the pair 3-copula samples. Figure 4.1 shows the quantile-quantile plots of the test distribution (of sums s_t) and a χ^2 distribution.

- The quantiles of the test statistics s_t computed from the Gauss copula samples align with those of the χ^2-distribution; the fit to the Gauss copula sample is good.
- The quantiles of the test statistics s_t computed from the t copula samples diverge from the χ^2-quantiles; the fit to the t copula sample is rejectable.
- The quantiles of the test statistics s_t computed from the pair copula samples deviate from the χ^2-distribution only in the very upper tail; the fit to the pair copula sample is moderate.

This shows that the Gauss copula function captures the moderate dependence structures, even if those are modelled conditionally in original. But heavy clustering is poorly fitted.

t copula fit The probability integral transformation for the t copula function has barely been touched in the literature. Then the explicit representation of the conditional t copula function is innovative.

Multivariate t distributions belong to a more general subclass of elliptical distributions:

Definition 4.13. A random variable $X \sim ELL(\mu, \Sigma, g)$ on $\mathbb{R}^d$ is said to have a *symmetric multivariate Pearson Type VII distribution*, if it has a density generator g, where

$$g(t) = \frac{\Gamma(N)}{(\pi\nu)^{d/2}\Gamma(N - d/2)}(1 + t/\nu)^{-N}, \quad N > d/2, \nu > 0.$$

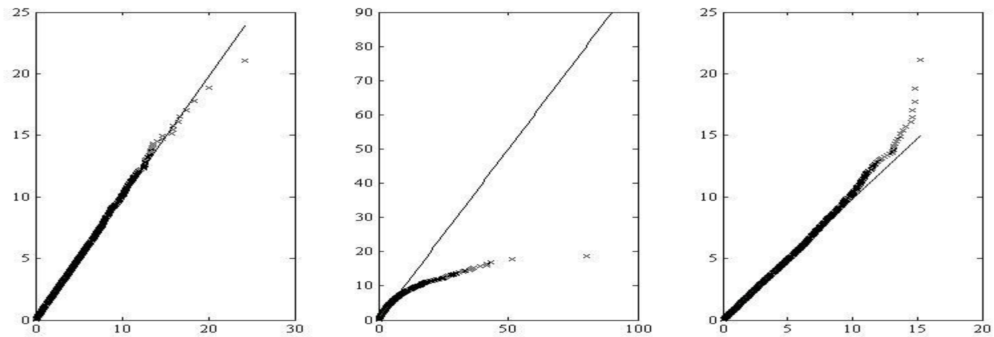

Figure 4.1: Quantile-quantile plot of χ^2-distribution and test statistics computed from Gauss copula fitted to Gauss copula samples (left), t copula samples (center), and pair copula samples (right).

For X of this type we write $X \sim MPVII(\mu, \Sigma, g_{N,\nu})$.

Symmetric multivariate Pearson Type VII distributions include a number of important distributions such as the multivariate t-distribution (for $N = \frac{1}{2}(d+\nu)$ and the multivariate Cauchy distribution (for $\nu = 1, N = \frac{1}{2}(d+1)$).

Theorem 4.14 ([34], Theorem 3.7)**.** *Let $X \sim MPVII(\mu, \Sigma, g_{N,\nu})$ with the usual partition of X, μ and Σ. Then we have*

$$X^1 \sim MPVII(\mu_1, \Sigma_{11}, g_{N-(d-m)/2,\nu}).$$

Theorem 4.14 shows us that a symmetric multivariate Pearson type VII distribution is closed under the taking of margins. In particular, if X has a d-dimensional t distribution with density generator $g_{(d+\nu)/2,\nu}$, then X^1 has density generator $g_{(m+\nu)/2,\nu}$, hence has a m-dimensional t distribution. Similarly, X^2 has a $(d-m)$-dimensional t distribution.

There exists a closed form representation of conditional densities for this class of distributions, too:

Proposition 4.15 (cf. [34], Theorem 3.7)**.** *Let $X \sim MPVII(\mathbf{0}, \mathbf{1}, g_{N,\nu})$ with the usual partition of X, $\mathbf{0}$ and $\mathbf{1}$. Then the conditional pdf f of X^2 given X^1 is*

$$f(x_2|x_1) = \frac{\Gamma(N)(1+\frac{s}{\nu})^{-(d-m)/2}}{(\pi\nu)^{(d-m)/2}\Gamma(N-(d-m)/2)} \left(1 + \frac{t}{\nu+u}\right)^{-N}, \tag{4.17}$$

where $s = x_1.x_1$ and $t = x_2.x_2$.

Proof. As to Theorem 4.14, the density of X^1 is

$$f(x_1) = \frac{\Gamma(N-(d-m)/2)}{(\pi\nu)^{m/2}\Gamma(N-d/2)} \left(1 + \frac{s}{\nu}\right)^{-N+(d-m)/2},$$

where $s = x_1.x_1$. Then, for $t = x_2.x_2$, we have

$$\begin{aligned} f(x_2|x_1) &= \frac{\Gamma(N)}{(\pi\nu)^{d/2}\Gamma(N-d/2)}\left(1+\frac{s+t}{\nu}\right)^{-N}\frac{(\pi\nu)^{m/2}\Gamma(N-d/2)}{\Gamma(N-(d-m)/2)}\left(1+\frac{s}{\nu}\right)^{N-(d-m)/2} \\ &= \frac{\Gamma(N)}{(\pi\nu)^{(d-m)/2}\Gamma(N-(d-m)/2)}\left(1+\frac{s}{\nu}\right)^{-(d-m)/2}\left(\frac{1+(s+t)/\nu}{1+s/\nu}\right)^{-N} \\ &= \frac{\Gamma(N)}{(\pi\nu)^{(d-m)/2}\Gamma(N-(d-m)/2)}\left(1+\frac{s}{\nu}\right)^{-(d-m)/2}\left(1+\frac{t}{\nu+s}\right)^{-N}. \end{aligned}$$

□

The conditional pdf (4.17) is not of type MPVII (even though it is spherical). Hence MPVII distributions are not closed under conditioning. We would very much like to generalize this result to non-standard means $\mu \neq \mathbf{0}$ and $\Sigma \neq \mathbf{1}$. This requires the following auxiliary results:

Lemma 4.16 (cf. [34], Theorem 3.7 ff.). *Let $X \sim MPVII(\mathbf{0}, \mathbf{1}, g_{N,m})$ with the usual partition of $X, \mathbf{0}$ and $\mathbf{1}$. Then the conditional distribution of X^2 given X^1 has the following stochastic representation:*

$$X^2|X^1 \stackrel{d}{=} R_u U,$$

where U a $(d-m)$-dimensional uniform random vector and R_s a generating variable such that $w = R_s^2/(\nu+s)$ has a Bessel II distribution with parameters $(d-m)/2$ and $N-(d-m)/2$).[10]

Lemma 4.17 ([34], Theorem 3.8). *Let $X \sim MPVII(\mu, \Sigma, g_{N,\nu})$ with the usual partition of X, μ and Σ. Then the distribution of X^2 given X^1 has the following stochastic representation:*

$$X^2|X^1 \stackrel{d}{=} \mu_{2|1} + R_s \Sigma_{2|1}^{1/2} U,$$

where $s = X^1.X^1$, U a $(d-m)$-dimensional uniform random vector and R_s a generating variable such that $w = R_s^2/(\nu+s)$ has a Bessel II distribution with parameters $(d-m)/2$ and $N-(d-m)/2)$.

Lemma 4.16 and Lemma 4.17 show that the conditional pdf of a general distribution of MPVII type is given by the common affine transformation of the conditional pdf of a standard distribution of MPVII type. This enables us to formulate Proposition 4.15 in general terms:

Proposition 4.18. *Let $X \sim MPVII(\mu, \Sigma, g_{N,\nu})$ with the usual partition of X, μ and Σ. Then the conditional pdf of X^2 given X^1 is*

$$f(x_2|x_1) = \frac{\Gamma(N)|\Sigma_{2|1}|^{-1/2}}{(\pi\nu)^{(d-m)/2}\Gamma(N-(d-m)/2)}\left(1+\frac{s}{\nu}\right)^{-(d-m)/2}\left(1+\frac{t}{\nu+s}\right)^{-N},$$

[10] A $\mathbb{R}$-valued random variable X has a Bessel II distribution with parameters a and b if $f^X(x) = \Gamma(a+b)/[\Gamma(a)\Gamma(b)]x^{a-1}(1+x)^{-(a+b)}$.

with $s = x_1.x_1$ *and* $t = (x_2 - \mu_{2|1}).\Sigma_{2|1}^{-1}(x_2 - \mu_{2|1})$*, where* $\mu_{2|1}$ *and* $\Sigma_{2|1}$ *are given by* (4.12) *and* (4.13)*, respectively.*

Proof. Proposition 4.18 follows from Proposition 4.15, Lemma 4.16 and Lemma 4.17. □

Given the conditional pdf of a MPVII type distribution, the conditional cdf is obtained by integration:

Proposition 4.19. *Let* $I = \{1, ..., i\}$ *and* $X \sim MPVII(\mu, \Sigma, g_{N,\nu})$ *with the* i *by* $i-1$ *partition of* X^I, μ^I *and* Σ^I*. Then the conditional cdf* F *of* X^i *given* $X^1, ..., X^{i-1}$ *is*

$$F(x_i|x_1, ..., x_{i-1}) = \frac{1}{2} + \frac{\Gamma(N - \frac{d-i}{2})\left(1 + \frac{s}{\nu}\right)^{-1/2} t}{(\pi\nu)^{1/2}\Gamma(N - \frac{d-i}{2} - 1/2)} h\left(\frac{1}{2}, N - \frac{d-i}{2}, \frac{3}{2}, \frac{-t^2}{\nu + s}\right),$$

where $t = (x_i - \mu_{2|1}^I)/\Sigma_{2|1}^I$, $s = \sum_{j=1}^{i} x_j^2$, h *the Gaussian hypergeometric function*[11] *and*

$$\begin{aligned} \mu_{2|1}^I &= \Sigma_{21}^I(\Sigma_{11}^I)^{-1}(x_1, ..., x_{i-1})' \\ \Sigma_{2|1}^I &= \Sigma_{22}^I - \Sigma_{21}^I(\Sigma_{11}^I)^{-1}\Sigma_{12}^I. \end{aligned}$$

Proof. By Theorem 4.14 and Proposition 4.18, we have

$$f(x_i|x_1, ..., x_{i-1}) = \frac{\Gamma(N - \frac{d-i}{2})}{\Sigma_{2|1}^I(\pi\nu)^{1/2}\Gamma(N - \frac{d-(i-1)}{2})}\left(1 + \frac{s}{\nu}\right)^{-1/2}\left(1 + \frac{t}{\nu + s}\right)^{-(N - \frac{d-i}{2})}, \tag{4.18}$$

where $s = \sum_{j=1}^{i} x_j^2$ and $t = (x_i - \mu_{2|1}^I/\Sigma_{2|1}^I)^2$. Then integration with respect to variable x_i leads to

$$\begin{aligned} F(x_i|x_1, ..., x_{i-1}) &= \int_{-\infty}^{x_i} f(x_i|x_1, ..., x_{i-1})dx_i \\ &= \frac{\Gamma(N - \frac{d-i}{2})\left(1 + \frac{s}{\nu}\right)^{-1/2}}{\Sigma_{2|1}^I(\pi\nu)^{1/2}\Gamma(N - \frac{d-(i-1)}{2})}\int_{-\infty}^{x_i}\left(1 + \frac{(\frac{x-\mu_{2|1}^I}{\Sigma_{2|1}^I})^2}{s + \nu}\right)^{-(N - \frac{d-i}{2})} dx \\ &= \frac{\Gamma(N - \frac{d-i}{2})\left(1 + \frac{s}{\nu}\right)^{-1/2}}{\Sigma_{2|1}^I(\pi\nu)^{1/2}\Gamma(N - \frac{d-(i-1)}{2})}\int_{-\infty}^{\frac{x_i-\mu_{2|1}^I}{\Sigma_{2|1}^I}}\left(1 + \frac{y^2}{s + \nu}\right)^{-(N - \frac{d-i}{2})} \Sigma_{2|1}^I dy \\ &= \frac{\Gamma(N - \frac{d-i}{2})\left(1 + \frac{s}{\nu}\right)^{-1/2}}{(\pi\nu)^{1/2}\Gamma(N - \frac{d-(i-1)}{2})}\left[yh\left(\frac{1}{2}, N - \frac{d-i}{2}, \frac{3}{2}, -\frac{y^2}{\nu + s}\right)\right]_{-\infty}^{\frac{x_i-\mu_{2|1}^I}{\Sigma_{2|1}^I}} \\ &= \frac{1}{2} + \frac{\Gamma(N - \frac{d-i}{2})\left(1 + \frac{s}{\nu}\right)^{-1/2}}{(\pi\nu)^{1/2}\Gamma(N - \frac{d-(i-1)}{2})}\frac{x_i - \mu_{2|1}^I}{\Sigma_{2|1}^I} h\left(N - \frac{d-i}{2}, \frac{-(\frac{x_i-\mu_{2|1}^I}{\Sigma_{2|1}^I})^2}{\nu + s}\right). \end{aligned}$$

□

[11] Abramowitz and Stegun [2] gives the definition and some basic examples of the Gaussian hypergeometric function (or series). Section A.4 appends the particular case on hand.

Proposition 4.19 facilitates to formulate the conditional cdf of the t copula analogously to Proposition 4.12:

Proposition 4.20. *Let C be the t d-copula with correlation matrix Σ and ν degrees of freedom. For $I = \{1, ..., i\}$ and the i by $i-1$ partition of Σ^I, it holds*
(4.19)

$$C_{i|1,...,i-1}(u_i|u_1,...,u_{i-1}) = \frac{1}{2} + \frac{\Gamma(\frac{i+\nu}{2})\left(1+\frac{s}{\nu}\right)^{-1/2}}{(\pi\nu)^{1/2}\Gamma(\frac{(i-1)+\nu}{2})}\left(\frac{t-\mu^I_{2|1}}{\Sigma^I_{2|1}}\right) h\left(\frac{1}{2}, \frac{i+\nu}{2}, \frac{3}{2}, \frac{-(\frac{t-\mu^I_{2|1}}{\Sigma^I_{2|1}})^2}{\nu+s}\right),$$

where $s = \sum_{j=1}^{i-1} t_\nu^{-1}(u_j)$, $t = t_\nu^{-1}(u_i)$ and

$$\begin{aligned}\mu^I_{2|1} &= \Sigma^I_{21}(\Sigma^I_{11})^{-1}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_{i-1}))' \\ \Sigma^I_{2|1} &= \Sigma^I_{22} - \Sigma^I_{21}(\Sigma^I_{11})^{-1}\Sigma^I_{12}.\end{aligned}$$

Proof. Let $(U^1, ..., U^d)$ be a random variable with a t copula distribution C. For $I = \{1, ..., i\}$, it is clear that the cdf of the I-margin $(U^1, ..., U^i)$ is the t i-copula function with correlation matrix Σ^I and ν degrees freedom. Then the transformed random variable $(X_1, ..., X_i)$ with $X_i = t_\nu^{-1}(U^i), i = 1, ..., d$ has a joint Student t distribution with correlation matrix Σ^I and ν degrees of freedom. By Proposition 4.19, the conditional cdf of X^i given $X^1, ..., X^{i-1}$ is (4.18). Then (4.19) holds by reversing the margin transformation. □

Proposition 4.20 leads to Algorithm 10, that can be used to successively transform some sample data.

Algorithm 10: PIT for t copula function

Computes, for $t = 1, ..., T$, the probability integral transforms $(z_{1,t}, ..., z_{d,t})$ of $(u_{1,t}, ..., u_{d,t})$ by the t copula with correlation matrix Σ and ν degrees of freedom.;

for t = *1 to* T **do**
  $z_{1,t} = u_{1,t}$;
  for i = *2 to* d **do**
    $\mu^i_{2|1} = \Sigma^i_{21}(\Sigma^i_{11})^{-1}(t_\nu^{-1}(u_{1,t}), \ldots, t_\nu^{-1}(u_{i-1,t}))'$;
    $\Sigma^i_{2|1} = \Sigma^i_{22} - \Sigma^i_{21}(\Sigma^i_{11})^{-1}\Sigma^i_{12}$;
    $t = t_\nu^{-1}(u_{i,t})$;
    $s = \sum_{j=1}^{i-1} t_\nu^{-1}(u_{j,t})$;
    $z_{i,t} = \frac{1}{2} + \frac{\Gamma(\frac{i+\nu}{2})\left(1+\frac{s}{\nu}\right)^{-1/2}}{(\pi\nu)^{1/2}\Gamma(\frac{\nu}{2})}(\frac{t-\mu^i_{2|1}}{\Sigma^i_{2|1}})h\left(\frac{1}{2}, \frac{i+\nu}{2}, \frac{3}{2}, \frac{-(\frac{t-\mu^i_{2|1}}{\Sigma^i_{2|1}})^2}{\nu+s}\right)$;
  end
end

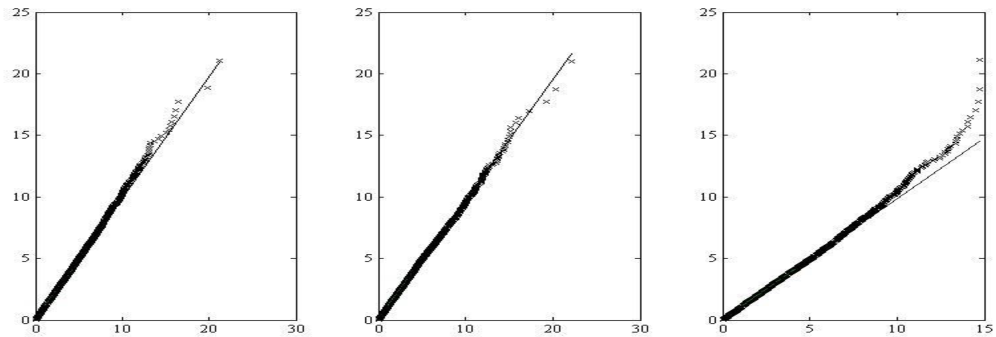

Figure 4.2: Quantile-quantile plot of χ^2-distribution and test statistics computed from t copula fitted to Gauss copula samples (left), t copula samples (center), and pair copula samples (right).

We implement Algorithm 10 in the context of the general procedure, using all the random draws simulated in Chapter 3. This requires again to estimate beforehand the correlation matrix and the degrees of freedom of the t copula by applying Algorithm 6 to the Gauss, the t and the pair 3-copula samples. Figure 4.2 shows the quantile-quantile plots of the test distribution (of sums s_t) and a χ^2 distribution.

- The quantiles of the test statistics s_t computed from the Gauss copula samples align with those of the χ^2-distribution; the fit to the Gauss copula sample is good.
- The quantiles of the test statistics s_t computed from the Gaussian copula samples align with those of the χ^2-distribution; the fit to the t copula sample is good.
- The quantiles of the test statistics s_t computed from the pair copula samples deviate from the χ^2-distribution only in the very upper tail; the fit to the pair copula sample is moderate.

These results signify that the t copula function contains a fairly wide range of dependence structures.

Pair copula fit Chapter 3 showed already how to evaluate the conditional cdf corresponding to a pair copula decomposition. Hence the probability integral transform can easily be specialized to the pair copula approach.

For the D-vine decomposition as described in Definition 2.12, the complex conditional cdf (3.1) can be stripped down to cascades of bivariate conditional cdfs (2.20) with starting point (3.2). This is exploited in Algorithm 11, which can be used to compute the conditional probability transform at some sample data.

In the case of the pair copula decomposition corresponding to a canonical vine as to Definition 2.13, we start iterating at (3.3) and recur to cascades of bivariate conditional

Algorithm 11: PIT for pair copula function based on D-vine decomposition

Computes, for $t = 1, ..., T$*, the probability integral transforms* $(z_{1,t}, ..., z_{d,t})$ *of* $(u_{1,t}, ..., u_{d,t})$ *by the pair copula based on D-vine density decomposition with (conditional) correlation matrix* Σ.;

```
for t = 1 to T do
    z_{1,t} = x_{1,t};
    z_{2,t} = h(x_{2,t}, x_{1,t}; θ_{1,1});
    w_{2,1} = x_{2,t};
    w_{2,2} = h(x_{1,t}, x_{2,t}; θ_{1,1});
    for i = 3 to d do
        z_{i,t} = h(x_{i,t}, x_{i−1,t}; θ_{1,i−1});
        for j = 2 to i-1 do
            z_{i,t} = h(z_{i,t}, w_{i−1,2(j−1)}; θ_{j,i−j});
        end
        if j=d-1 then
            Stop;
        end
        w_{i,1} = x_{i,t};
        w_{i,2} = h(w_{i−1,1}, w_{i,1}; θ_{1,i−1});
        w_{i,3} = h(w_{i,1}, w_{i−1,1}; θ_{1,i−1});
        for j=1 to i − 3 do
            w_{i,2j+2} = h(w_{i−1,2j}, w_{i,2j+1}; θ_{j+1,i−j−1});
            w_{i,2j+3} = h(w_{i,2j+1}, w_{i−1,2j}; θ_{1,i−j−1});
        end
        w_{i,2i−2} = h(w_{i−1,2i−4}, w_{i,2i−3}; θ_{i−1,1});
    end
end
```

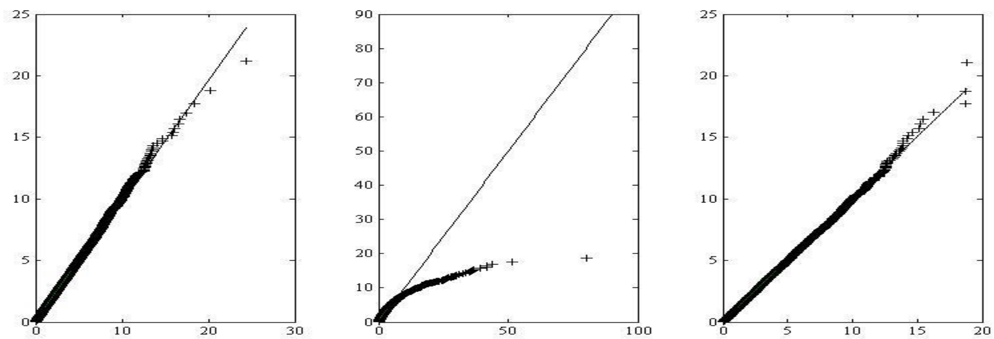

Figure 4.3: Quantile-quantile plot of χ^2-distribution and test statistics computed from pair copula fitted to Gauss copula samples (left), t copula samples (center), and pair copula samples (right).

copulas of the form (2.20) in the end. This is written out in Algorithm 12, that can be used to perform the conditional probability transformations on some sample data.

Algorithm 12: PIT for pair copula function based on canonical vine decomposition

Computes, for $t = 1, ..., T$, *the probability integral transforms* $(z_{1,t}, ..., z_{d,t})$ *of* $(u_{1,t}, ..., u_{d,t})$ *by the pair copula based on a canonical vine density decomposition with (conditional) correlation matrix* Σ.;

```
for t = 1 to T do
    z_{1,t} = x_{1,t};
    for i = 2 to d do
        z_{i,t} = x_{i,t};
        for j = 1 to i-1 do
            z_{i,t} = h(z_{i,t}, z_{j,t}; θ_{j,i-j});
        end
    end
end
```

We implement Algorithm 12 in the context of the general procedure to back transform and test all the copula samples simulated in Chapter 3 (almost identical results would apply to using Algorithm 11 instead). This requires again to estimate beforehand the (conditional) correlation matrix of the pair copula by applying Algorithm 8 (or 7) to the Gauss, the t and the pair 3-copula samples. Figure 4.3 presents the results in form of quantile-quantile plots of the resulting distributions (of sums s_t) and a χ^2 distribution.

- The quantiles of the test statistics s_t computed from the Gauss copula samples align with those of the χ^2-distribution; the fit to the Gauss copula sample is good.

- The quantiles of the test statistics s_t computed from the t copula samples diverge from the χ^2-quantiles; the fit to the t copula sample is poor.

- The quantiles of the test statistics s_t computed from the pair copula samples align with those of the χ^2-distribution; the fit to the pair copula sample is good.

Hence the pair copula function (with Gauss building blocks) can indeed reproduce a normal dependence structure, but it fails to get the extremes of the t copula function.

4.3 Summary

In this chapter we estimated the parameters of the Gauss, the t and the pair copula function on the basis of a given realization by the canonical maximum likelihood method. Then we developed testing methods to assess how good the estimates fit the data.

Having explained the canonical maximum likelihood method, we derived (almost) explicit estimation procedures for the Gauss and the t copula. Here the closed form pdf representations facilitated the formulation of the estimators, although the additional degrees of freedom prevented an analytical solution of the estimation problem in the case of the t copula.

Similarly, the pair copula construction turned out to be well suited to the likelihood method. The conditional construction via reallocated cascades of bivariate building block copulas enabled us to evaluate the likelihood function by a recursive pattern.

Then the derived methods were tested. All estimation methods stood up to the quality check on the sampled data of Chapter 3 in that the initial configurations were adequately reproduced.

Continuing, we showed a keen interest in testing procedures for the goodness-of-fit of the estimated models. This involved the multivariate probability integral transform. One of our major contributions here was the development of analytical transformations in the case of the Gauss and the t copula. Regarding pair copulas, we once more used the specific construction by building block copula functions for a sophisticated but direct method.

Then we successfully applied the goodness-of-fit testing schemes to the sample data, which we estimated before. The findings had us to see that only the t copula captures a wide range of dependence structures, while the Gauss and the pair copula functions fail to handle extreme scenarios.

Chapter 5

Risk management applications

In this chapter we describe the quantitative approaches to risk management in the context of a new regulatory framework. These are subsequently implemented in the savings bank sector, applying the whole copula machinery to business.

Section 5.1 clarifies the risk perspective in up-to-date banking competence and catalogues the main risk types. Section 5.2 puts the current practices onto a firmer mathematical footing. Formal definition to some common risk notions is given. The use of copula methods is important here. Then Section 5.3 applies the risk measurement techniques to business at Sparkasse Leverkusen. Here the risk management tool CopRisk, which is currently being installed for risk profiling and controlling, is described in detail.

The presentation in this chapter is based on the consultative documents by Hendricks et al. [41] and Hendricks and Cole [40]. These contain the status-quo of the regulatory framework. The standard reference for quantitative risk management practice is Embrecht et al. [31]. See Beck and Lesko [8], Schumacher et al. [71], Friedberg and Schumacher [36] for insights into business implementations in the savings bank sector. The software solution is self-elaborated.

5.1 Integrated risk management

The integrated risk management has received a lot of recent attention[1] in contemporary banking business. The managing and monitoring of capital adequacy in relation to the risk profile arrived[2] at the core competence (and duty) of institutional practice.

The goal of *integrated risk management* is to ensure an awareness of the financial risks across the range of diverse business activities.

> An integrated risk management system seeks to have in place management policies and procedures that are designed to help ensure an awareness of, and

[1] Hendricks and Cole [40] survey the latest trends.

[2] Embrecht et al. [31, Section 1.2] present the historical ruling of the financial sector and describe the conceptual changes over the years of supervision.

> accountability for, the risks taken throughout the financial firm, and also to develop the tools needed to address those risks. A key objective is to ensure that the firm does not ignore any material source of risk.[3]

The recent approaches [cf. 41, 40, 31] to the management of the major individual risks in the banking sector allocate three main types of uncertainty : *market risk*, *credit risk* and *operational risk*. Surely the most observable source is market risk, that is the risk associated with the change in value of a financial position due to changes in the value of the underlying components, on which this position depends. Credit risk is the risk of counterparty default and lost repayments on outstanding investments. Operational risk, which draws increasing attention in banking, is the risk of losses resulting from inadequate or failed internal processes, people and systems, or from external events [cf. 31, Section 1.1].[4] The actual contour of each type is very much subjective to the bank's specific businesses. The same holds for the quantitative impact on the overall exposure.[5] Moreover, the boundaries of the risk categories are somewhat floating.

The new challenge regarding capital adequacy is comprised in the multivariate nature of risk. Not least by legal requirements, the financial institutions are required to internalize holistic capital standards, that dwell on the bank's entire economic loss potential and overall risk exposure. For market risk, credit risk or even portfolio risk on enterprize level, the driving risk factors such as individual asset prices, risky loans or sectorial loss potentials assemble to multidimensional vectors of risky positions. Then *aggregation* of sectorial loss potentials becomes an inevitable step towards total risk assessment and a discipline of integrated risk management sine qua non.

> Broadly, risk aggregation refers to efforts by firms to develop quantitative risk measures that incorporate multiple types or sources of risk. The most common approach is to estimate the amount of economic capital that a firm believes is necessary to absorb potential losses associated with each of the included risks. This is typically accomplished via mathematical or statistical techniques designed to assess the likelihood of potential adverse outcomes (...)[6]

Up to date practitioners in the savings bank sector monitor and manage different types of risk in a more *silo-based* manner [cf. 8, 39]. That is, the risk categories are profiled separately from each other. Commercial software is set in place to assess most of the individual risk exposures, the others being profiled by sophisticated in-house solutions. The resulting numbers on individual exposures are then simply added up to get an aggregate exposure. It is obvious from the preceding chapters that simple methods like this only

[3] Quote taken from the opening paragraphs in Hendricks and Cole [40, Paragraph 2]. Kloman [48] writes integrated risk management down very concisely as a discipline for living with the possibility that future events may cause adverse effects.

[4] Some authors [e.g. 9] extend the catalogue of major risks by liquidity risk

[5] Beck et al. [9] sketch qualitative properties of the exposure to market and credit risks, and Friedberg and Schumacher [36] give a quantitative example of risk numbers in the Sparkasse sector.

[6] Quote taken from the opening paragraphs in Hendricks and Cole [40, Paragraph 2].

mark a first step towards true integrated risk management. We take on this fallacy and set up the mathematical framework for the aggregation of dependent risk lines in the following.

5.2 Quantitative risk profile

We consider the quantitative assessment of capital charges in relation to a bank's overall exposure in the style of Embrecht et al. [31, Chapter 2].

Consider a portfolio of risky positions such as a book of risky loans, a collection of stocks or bonds or even an overall position of risky assets. Introduce a calender time $t \in [0, T]$, measured in years, and model the risky factors by a stochastic process $\{Z_t\}$ on $\mathbb{R}^d$. It is convenient to consider process movements over a fixed horizon Δ. This leads us to the discrete process $\{Z_\tau^\Delta\}, \tau = 0, 1, ..., T/\Delta$, which follows from the generic process $\{Z_t\}$ by $Z_\tau^\Delta = Z_{\tau\Delta}$. For daily movements, for example, set $\Delta = 1/365$ (or $\Delta = 1/250$, considering working days).

The value of the portfolio is modelled as a function of time and the risky factors. It is a stochastic process $\{V_t\}$ on $\mathbb{R}$ with

$$V_t = f(t, Z_t), \quad t \in [0, T]$$

for some measurable function $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$. The choice of the risk factors and of f is of course a modelling issue and depends on the portfolio at hand. Frequently used risk factors are logarithmic prices of financial assets, yields and logarithmic exchange rates. The discrete portfolio process, denoted by $\{V_\tau^\Delta\}$, is defined by $V_\tau^\Delta = f(\tau\Delta, Z_{\tau\Delta})$ for $\tau = 0, 1, ..., T/\Delta$. For a given time horizon Δ, the *loss process* $\{L_{\tau+1}^\Delta\}$ of the portfolio over the period $[\tau, \tau + \Delta]$ is defined as

$$\begin{aligned} L_{\tau+1}^\Delta &= -(V_{\tau+1}^\Delta - V_\tau^\Delta) \\ &= -(f((\tau+1)\Delta, Z_{(\tau+1)\Delta}) - f(\tau\Delta, Z_{\tau\Delta})). \end{aligned}$$

The portfolio losses are defined retrospectively. Hence the loss over the period $[\tau\Delta, (\tau+1)\Delta]$ is known at time $(\tau + 1)\Delta$ but random as from time $\tau\Delta$.

The series of risk factor changes $\{X_\tau^\Delta\}$ is defined by $X_\tau^\Delta = Z_\tau^\Delta - Z_{\tau-1}^\Delta$ for $\tau = 1, ..., T/\Delta$. Then

$$L_{\tau+1}^\Delta = -(f((\tau+1)\Delta, Z_\tau^\Delta + X_{\tau+1}^\Delta) - f(\tau\Delta, Z_{\tau\Delta})).$$

At time $\tau\Delta$, Z_τ^Δ is known and the portfolio loss over the next period Δ is a function of the random variable $X_{\tau+1}^\Delta$. Hence, we may introduce the so called *loss operator* $l_{\tau+1}^\Delta : \mathbb{R}^d \to \mathbb{R}$, that maps risk factor changes into period ahead losses. It is defined as

$$l_\tau^\Delta(x) = -(f((\tau+1)\Delta, Z_\tau^\Delta + x) - f(\tau\Delta, Z_{\tau\Delta})), \quad x \in \mathbb{R}.$$

Example 5.1 (cf. [31], Example 2.4)**.** Consider a fixed portfolio of d stocks. Denote by ω_i the number of shares of stock i in the portfolio at all times $t \in [0, T]$ and the price process of stock i by $\{S_{t,i}\}$. Using the logarithmic prices $Z_{t,i} = \ln(S_{t,i}), 1 \leq i \leq d$ as risk factors and fixing time horizon $\Delta > 0$, the risk factor changes $X^{\Delta}_{\tau+1,i} = \ln(S^{\Delta}_{\tau+1,i}) - \ln(S^{\Delta}_{\tau,i})$ then correspond to the log-returns of the stocks in the portfolio. Then $V_t = \sum_{i=1}^{d} \omega_i e^{Z_{t,i}}$ and

$$L^{\Delta}_{\tau+1} = -\sum_{i=1}^{d} \omega_i S^{\Delta}_{\tau,i}(e^{X^{\Delta}_{\tau+1,i}} - 1).$$

Moreover, the loss operator is given by

$$l^{\Delta}_{\tau}(x_1, ..., x_d) = -\sum_{i=1}^{d} \omega_i S^{\Delta}_{\tau,i}(e^{x_i} - 1).$$

Loss distributions using copulas Copula functions apply for the distributional description of the period ahead loss.

Fix a current time τ and a time horizon Δ. Consider the distribution $P^{L^{\Delta}_{\tau+1}}$ of $L^{\Delta}_{\tau+1}$, termed the *loss distribution*. Using the loss operator notation, the loss distribution is determined by the distribution $P^{X^{\Delta}_{\tau+1}}$ of the risk factor changes $X^{\Delta}_{\tau+1}$:

$$P[L^{\Delta}_{\tau+1} \in B] = P[l^{\Delta}_{\tau}(X^{\Delta}_{\tau+1}) \in B], \quad B \in \mathcal{B}(\mathbb{R}).$$

It is relevant in this respect to separate conditional and unconditional viewpoints according to the time series properties of the risk factor changes $\{X^{\Delta}_{\tau+1}\}$. Suppose that the risk factor changes $\{X^{\Delta}_{\tau+1}\}$ are invariant under shifts of time. That is equivalent to saying that the random variables $X^{\Delta}_{\tau+1}, \tau = 1, ..., T/\Delta$ are equally distributed:

$$X^{\Delta}_{\tau+1} \stackrel{d}{=} X, \quad \tau = 1, ..., T/\Delta,$$

where X is some generic risk factor change on $\mathbb{R}^d$ with stationary distribution P^X. The unconditional loss distribution $P^{L^{\Delta}_{\tau+1}}$ is then defined as the distribution of $l^{\Delta}_{\tau}(\cdot)$ under the stationary distribution P^X of the generic risk factor change X:

$$P[L^{\Delta}_{\tau+1} \in B] = P[l^{\Delta}_{\tau}(X) \in B], \quad B \in \mathcal{B}(\mathbb{R}). \tag{5.1}$$

Then

$$F^{L^{\Delta}_{\tau+1}}(l) = P[l^{\Delta}_{\tau}(X) \leq l], \quad l \in \mathbb{R}_{\infty}.$$

Capital adequacy decisions based on the unconditional loss distribution are referred to as *static* risk management.

If dynamic structures are to be incorporated, introduce the information flow $\{\mathcal{F}^{\Delta}_{\tau}\}$ given by $\mathcal{F}^{\Delta}_{\tau} = \mathcal{F}_{\Delta\tau}$ and consider the conditional risk factor change $X^{\Delta}_{\tau+1}$ over period $[\tau, \tau+1]$ given the current information $\mathcal{F}^{\Delta}_{\tau}$. In this case, the conditional loss distribution $P^{L^{\Delta}_{\tau+1}}$ is determined by the distribution of $l^{\Delta}_{\tau}(\cdot)$ under the conditional distribution of $X^{\Delta}_{\tau+1}$ given $\mathcal{F}^{\Delta}_{\tau}$:

$$P[L^{\Delta}_{\tau+1} \in B|\mathcal{F}^{\Delta}_{\tau}] = P[l^{\Delta}_{\tau}(X^{\Delta}_{\tau+1}) \in B|\mathcal{F}^{\Delta}_{\tau}], \quad B \in \mathcal{B}(\mathbb{R}).$$

Then

$$F^{L^{\Delta}_{\tau+1}|\mathcal{F}^{\Delta}_{\tau}}(l) = P[l^{\Delta}_{\tau}(X^{\Delta}_{\tau+1}) \leq l|\mathcal{F}^{\Delta}_{\tau}], \quad l \in \mathbb{R}_{\infty}.$$

Techniques based on the conditional loss distribution are referred to as *dynamic* risk management.

In either case, the loss operator in conjunction with the distribution of the risk factor changes gives the distribution of the one period ahead loss of our portfolio. Then it is a prime concern of risk managers to describe the distributional aspects of risk factor changes $X^{\Delta}_{\tau+1}$ as from time τ. We follow the static approach in this work.

As to Theorem 2.2, the cdf F^X of the stationary distribution of the generic risk factor changes $X = (X^1, ..., X^d)$ may then be modelled using copulas:

$$F^X(x_1, ..., x_d) = C(F^X_1(x_1), ..., F^X_d(x_d)),$$

where the marginal cdf's $F^X_i, i = 1, ..., d$ and the copula C may certainly depend on parameters $\theta \in \Theta$ as discussed in Section 4.1.[7] In consequence of our designated target, here we assume a predefined model for the marginal cdf's F^X_i for $i = 1, ..., d$, which comes either non-parametric in form of an empirical distribution or in parametric form depending on some set $\theta_i \in \Theta_i$ of admissible parameters.

Remark 5.2. This assumption corresponds to the silo-based risk management, which is already practised in business to handle individual exposures.

We denote the fitted i-margin by $\hat{F}_i(\cdot)$, dropping the eventual parametrization. Then

$$F^X(x_1, ..., x_d) = C(\hat{F}_1(x_1), ..., \hat{F}_d(x_d)). \tag{5.2}$$

This prepares the distributional description of the random losses $L^{\Delta}_{\tau+1}$ as to (5.1).

Risk measures based on loss distributions One of the principle functions of risk management in the financial sector is to determine the amount of capital needed as a buffer against potential losses. This requires the definition of measures, that quantify the firm's belief in its overall risk. It is common in this respect to use statistical quantities of the loss distribution as a summary of the portfolio risk over some time period Δ. We restrict ourselves to the value-at-risk measure and the mean value-at-risk measure, which are most widely used among practitioners in the savings bank sector.[8]

Definition 5.3. Given some confidence level $\alpha \in (0, 1)$, the *Value-at-Risk* of our portfolio at the confidence level α is defined by the smallest number l such that the probability that the loss L exceeds l is no larger than $(1 - \alpha)$:

$$\mathrm{VaR}_{\alpha}(L) = \inf\{l \in \mathbb{R} : P[L > l] \leq 1 - \alpha\}.$$

[7] A dynamic perspective on portfolio risk management requires the notion of a conditional copula functions $C(\cdot|\mathcal{F})$ as to Patton [63] who proved Sklar's Theorem in the conditional setting. His conditional approach is in turn taken up by numerous authors [e.g. 22, 65] for time series modelling. The conditional copula approach is briefly described in Section A.2.

[8] Due to its prevalence in financial institutions around the world, the value-at-risk measure is incorporated in the Basel II capital-adequacy framework [cf. 41, Part 2].

Thus the VaR_α at confidence level α is the $(1-\alpha)$-quantile of the loss distribution P^L:

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

Typical values for α are $\alpha = 0.95$ or $\alpha = 0.99$. Sometimes the statistic $\text{VaR}_\alpha^{\text{mean}} = \text{VaR}_\alpha - E^P[L]$, called the *mean Value at Risk*, is used for capital-adequacy purposes instead of ordinary VaR. The mean-VaR statistic can be interpreted as the economic capital needed as a buffer against unexpected losses. It is clear from Definition 5.3 that the VaR_α (or $\text{VaR}_\alpha^{\text{mean}}$) at confidence level α disregards all those potential losses, which occur with a probability less than $1-\alpha$. Hence no information about the severity of the worst cases is included. This is an acknowledged drawback of VaR (or $\text{VaR}_\alpha^{\text{mean}}$) as a risk measure.[9]

Regarding (5.1), both quantities, VaR_α and $\text{VaR}_\alpha^{\text{mean}}$, are distributional statistics of the loss operator l_t, which are generally not given in closed form.[10] Using (5.2), Monte Carlo methods are employed here to estimate the distributional statistics from generated loss scenarios. This involves the following steps [cf. 31, Section 2.3.3]:

(1) calibrate F^X (or copula C of F^X, in fact) to historical risk factor change data $\aleph = \{x_{1,s}, ..., x_{d,s}\}_{s=\tau-n+1}^{\tau}$

(2) generate independent realizations $x^{(1)}, ..., x^{(M)}$ of ($\mathbb{R}^d$-valued) risk factor changes

(3) apply the loss operator l_t to the simulated risk factor changes to get the sample losses $L_{\tau+1}^{\Delta\,(k)} = l_\tau^\Delta(x_{\tau+1}^{\Delta\,(k)}), \quad k = 1, ..., M.$

(4) approximate the mean loss by $E^P[L_{\tau+1}^\Delta] = \frac{1}{M}\sum_{k=1}^M L_{\tau+1}^{\Delta\,(k)}$ and compute the empirical risk measures

Chapter 4 explained how to estimate a copula based model from some sample data. Inference methods for given margins have been offered there. These apply readily to the calibration of the risk factor change distribution (step 1). Chapter 3 formulated the simulation methods for ordinary copulas. These may now be readily applied to produce sample vectors of risk factor changes according to the chosen model (step 2). Steps 3 and 4 involve only simple calculations.

Concluding, the machinery developed in the previous chapters enables us to assess the overall loss potential of portfolio, that consists of several risk lines.

[9] Many practitioners prefer to implement the related expected shortfall measure

$$\text{ES}_\alpha(L) = \frac{1}{1-\alpha}\int_\alpha^1 (F^L)^{-1}(u)du$$

which is (other than VaR) a coherent measure in the sense of Artzner et al. [4], too.

[10] Embrecht et al. [31] discusses some parametric models that come up with closed value-at-risk measures.

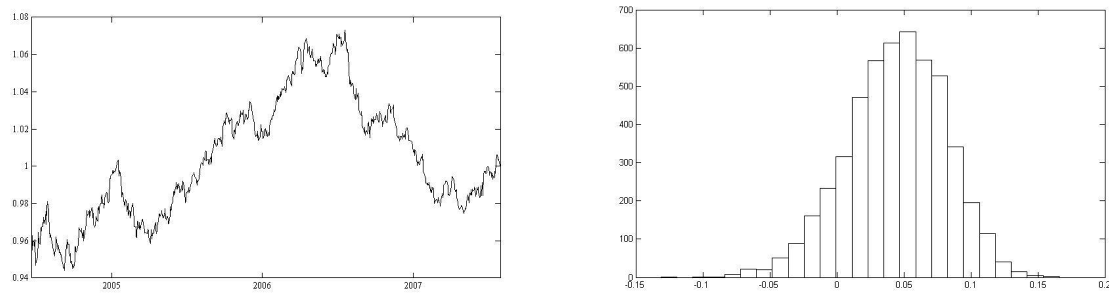

Figure 5.1: Time series and profit-loss distribution of the treasury book (normalized to July 31st, 2008).

5.3 Sparkasse Leverkusen: Business in a savings bank

We document the business application of copula based methods to the risk landscape at Sparkasse Leverkusen[11]. The aim is to quantify the one year overall loss distribution $P^{L^{\Delta}_{\tau+1}}$ as seen from the present time τ. Hence the time horizon is $\Delta = 1$ (we omit the horizon Δ in notations). Currently, the in-house risk team allocates three sectorial risk exposures on enterprize level. Then, as to (5.2), the multivariate risk factor under discussion is a $\mathbb{R}^3$-valued random variable $X_{\tau+1} = (X^1_{\tau+1}, X^2_{\tau+1}, X^3_{\tau+1})$. We touch the individual risk accounting just barely, following which we detail the in-house software solution.

The bulkiest exposure is *treasury risk*, that is the possibly adverse effects on the cash value of a book of loans subjective to yield curve movements (excluding credit events). A savings bank profits mainly from the transformation of loan maturities so that a shift of the curve and a change of its slope may cause a severe slump in the yields. Then $X^1_{\tau+1}$ is the random variable describing the absolute differences in cash values of the loan book over the period $[\tau, \tau + 1]$. In order to cope with potential losses from interest rate risk, the celebrated software solution S-Treasury[12] has been installed. The tool returns historical cash equivalents $x_{1,\tau-n}, ..., x_{1,\tau}$ of the present book as well as the year ahead distribution $P^{X^1_{\tau+1}}$ over changes in cash values of the loan portfolio. Figure 5.1 illustrates the time series and the profit-loss distribution.

Another major source of uncertainty is credit risk, that is the risk of drops in the fair market value of the loan portfolio due to credit events. Typical cases are the actual default of debtors, the degree migration of creditworthiness and the changes in the market's risk aversion. In this respect, $X^2_{\tau+1}$ is the random variable describing the absolute change in the

[11] Sparkasse Leverkusen is a local savings bank and group member of the Rheinische Sparkessen- und Giroverband. Visit [75] for more information.

[12] S-Treasury is a cash-flow based tool for balancing and financial controlling. Visit [38] for more information.

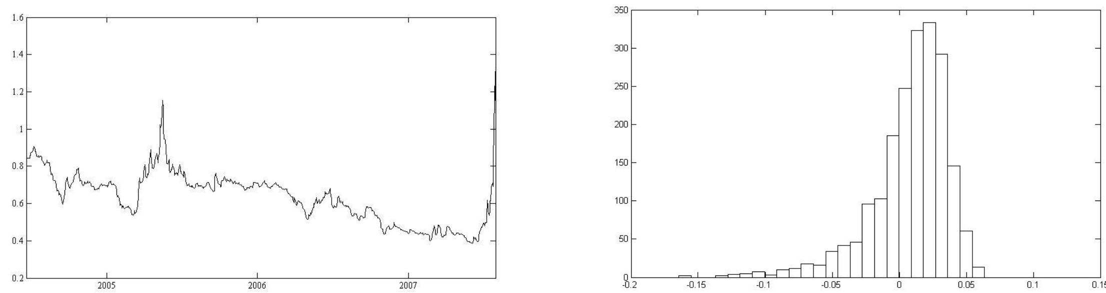

Figure 5.2: Time series of the itraxx index and profit-loss distribution of the credit book (normalized to July 31st, 2008).

value of the credit book in terms of cash equivalents. The individual profiling of these risks is accomplished by the commercial tool CreditPortfolioView[13], which is as well common in the public banking sector. CPV copes with the afore mentioned risk sources and returns the one year ahead probability distribution $P^{X^2_{\tau+1}}$ over possible changes in market values of the credit book. But it fails to offer historic cash values $x_{2,\tau-n}, ..., x_{2,\tau}$ of the present book. The insufficient availability of time series of credit risk data is an acknowledged (and so far unsolved) problem. The following assumption is a first workaround:

Assumption 5.4. There exists an increasing transformation $T : \mathbb{R} \to \mathbb{R}$ such that $X^2_{\tau+1} \overset{d}{=} T(\widetilde{X}^2_{\tau+1})$, where $\widetilde{X}^2_{\tau+1}$ the random variable describing the negative absolute returns of the itraxx crossover index.[14]

This allows us to find remedy in the use of historic index changes, which are available for the itraxx family. In this case, it follows from Theorem 2.4 that the copula C of $(X^1_{\tau+1}, X^2_{\tau+1}, X^3_{\tau+1})$ is equal to the copula C of $(X^1_{\tau+1}, \widetilde{X}^2_{\tau+1}, X^3_{\tau+1})$. Hence the itraxx time series $\widetilde{x}^2_{\tau-n}, ..., \widetilde{x}^2_{\tau}$ may be used to estimate the model (5.2), while it is still appropriate to run CPV for the purpose of the (true) marginal risk distribution $P^{X^2_{\tau+1}}$. Figure 5.2 shows the index time series and the profit-loss distribution.

The third exposure to risk steams from uncertain investments in stock and fund shares, termed *Depot A*. Hence, $X^3_{\tau+1}$ is the random variable describing the change in the value of equity shares due to adverse market price movements. It is acknowledged that capital market assets have long reported time series so that historic portfolio price changes $x_{3,\tau-n}, ..., x_{3,\tau}$ are certainly available. The year ahead distribution $P^{X^3_{\tau+1}}$ of portfolio price changes is assessed in-house by a historical simulation[15] approach. Figure 5.3 shows the historic portfolio prices and charts the profit-loss distribution.

[13] Credit Portfolio View (CPV) is a software tool for the assessing, measuring and illustration of risks in the credit book. Visit Sparkassen Rating und Risikosysteme GmbH-site [76] for more information.

[14] The itraxx index familiy is published by markit [52]. The itraxx crossover index is composed of the premia of credit default swaps written on companies in the subinvestment class. Then Assumption 5.4 is reasonable due to the similarity to the credit book names.

[15] Embrecht et al. [31, Section 2.3] discusses hisorical simulation as a standard method for market risks.

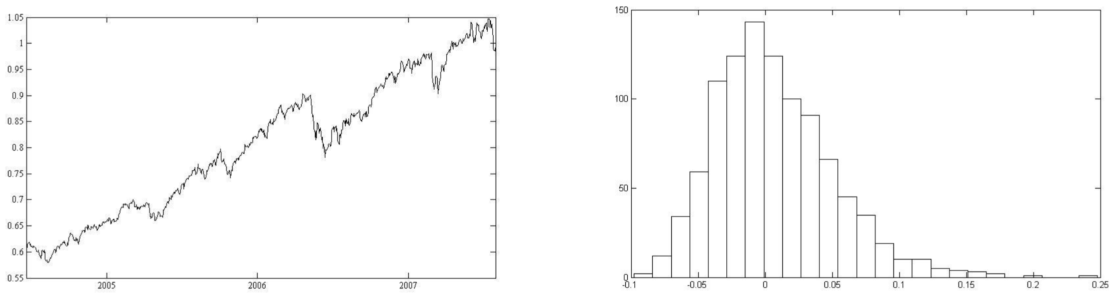

Figure 5.3: Time series and profit-loss distribution of the Depot-A book (normalized to July 31st, 2008).

Business implementation Only a handful of institutions in the Sparkasse group have sophisticated in-house solutions to the aggregation problem in use. Hence Sparkasse Leverkusen somehow pioneers the implementation of a copula based model in this sector.

Regarding the (short) catalogue of risky business lines, the objective cdf of risk factor changes is

$$F^X(x_1, x_2, x_3) = C(\hat{F}_1^X(x_1), \hat{F}_2^X(x_2), \hat{F}_3^X(x_3); \theta),$$

where $X = (X^1, X^2, X^3)$ is characterized as above, $\hat{F}_i^X$ is derived from P^{X_i} for $i = 1, ..., 3$, and $C(\cdot; \theta)$ is a parametric ordinary 3-copula of choice. Which copula to use has more than once been resumed during discussion rounds. The risk team at Sparkasse Leverkusen finalized the Gauss, the t and the pair copula family, because these were the ones most manageable. Then the final management solution requires the following steps:

(1) Choose a parametric copula $C(\cdot; \theta)$ from the Gaussian, the t and the pair copula family.

(2) Estimate the optimal set of copula parameters $\hat{\theta}$ from the historic time series $\aleph = \{x_{1,s}, x_{2,s}, x_{3,s}\}_{s=\tau-n+1}^{\tau}$.

(3) Test the hypothesis that $C(\cdot; \hat{\theta})$ is the true copula of X by means of a goodness-of-fit test.

(4) Simulate M independent realizations $x_{\tau+1}^{(1)}, ..., x_{\tau+1}^{M)}$ of risk factor changes for the period ahead.

(5) Evaluate the loss operator l_τ for each simulated risk factor change to provide the loss distribution $L_{\tau+1}^{(k)} = l_\tau(x_{\tau+1}^{(k)}), \quad k = 1, ..., M$.

(6) Compute the mean, the VaR_α and the VaR_α^{mean} from the sample $L_{\tau+1}^{(k)}, k = 1, ..., M$.

This solution can in principle be implemented in Microsoft-Excel alone, but we decided to use the MatLab Toolbox ExcelLink. Then all the functionality of MatLab is available

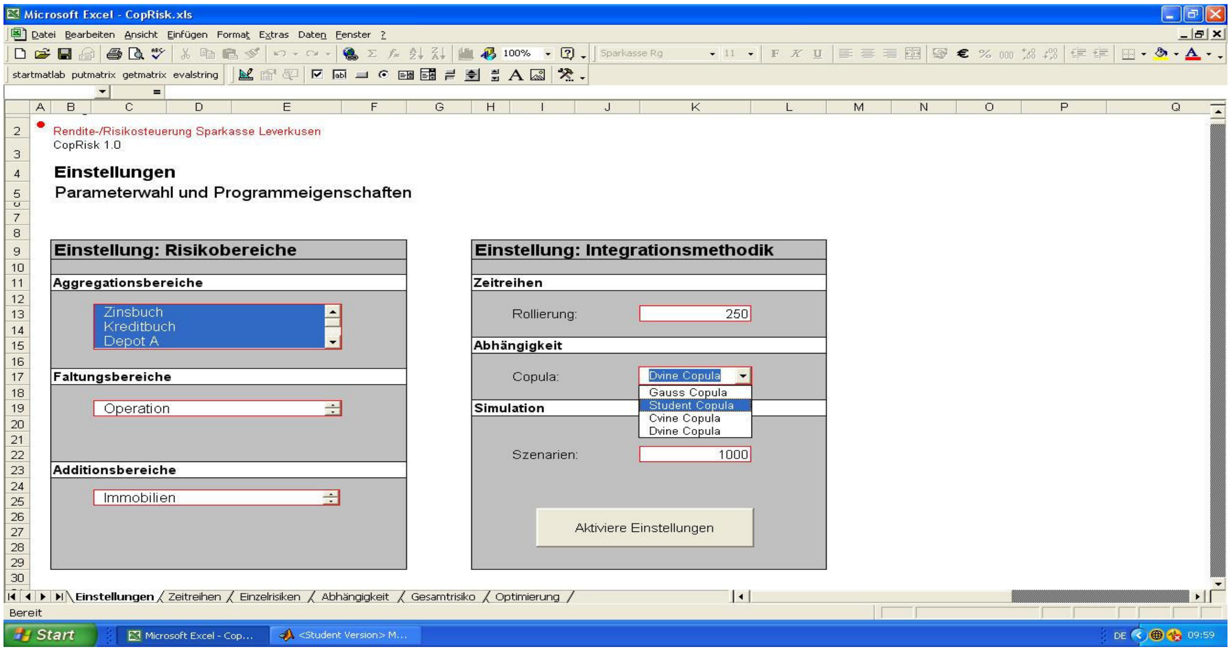

Figure 5.4: Screen-shot of MS-Excel worksheet `Einstellungen` showing the default setup.

from the Excel environment. This allows us to edit and analyze the data with the familiar Excel user interface, while the computational kernel is outsourced to MatLab.

The file `CopRisk.xls`[16] shows a possible realization. It contains the five worksheets `Einstellungen`, `Zeitreihen`, `Einzelrisiken`, `Abhängigkeit` and `Gesamtrisiko`.

On the `Einstellungen`-sheet, screen-shot in Figure 5.4, the user makes the program settings. Here the user selects which copula function to use. The options are the Gauss Copula, the t Copula, the canonical vine pair copula and the D-vine pair copula (default is the Gauss copula). The user also assigns the roll over period Δ, which is measured in days (default is a one year horizon $\Delta = 250$), and the number M of simulated loss scenarios (default is $M = 1000$). Additionally, the user chooses what risk types to incorporate. Here we forego all the alternatives to selecting the risk categories Zinsbuch, Kreditbuch and Depot-A.

Figure 5.5 is a screen-shot of the `Zeitreihen`-sheet, that shows the time series data $\{Z_{1,t}, Z_{2,t}, Z_{3,t}\}$ of historic risk factor values. These are entered on a daily scale in a previous step. On this sheet, the user sets the periods of the historic time series. It is essential here to mark synchronous series $\{Z_{1,t}, Z_{2,t}, Z_{3,t}\}_{t=\tau-m+1}^{\tau}$, albeit the time lag m is so long arbitrary, as it exceeds the roll over period Δ. The risk factor change data $\aleph = \{x_{1,s}, x_{2,s}, x_{3,s}\}_{s=\tau-n+1}^{\tau}$, where $n = \Delta$ is the time stretch, is then worked out internally as the absolute return with a Δ-day roll over:

$$x_{i,s} = Z_{i,s} - Z_{i,s-\Delta}, \quad s = \tau - n + 1, ..., \tau$$

[16]The file CopRisk.xls features some additional functionality not discussed in the course of this work.

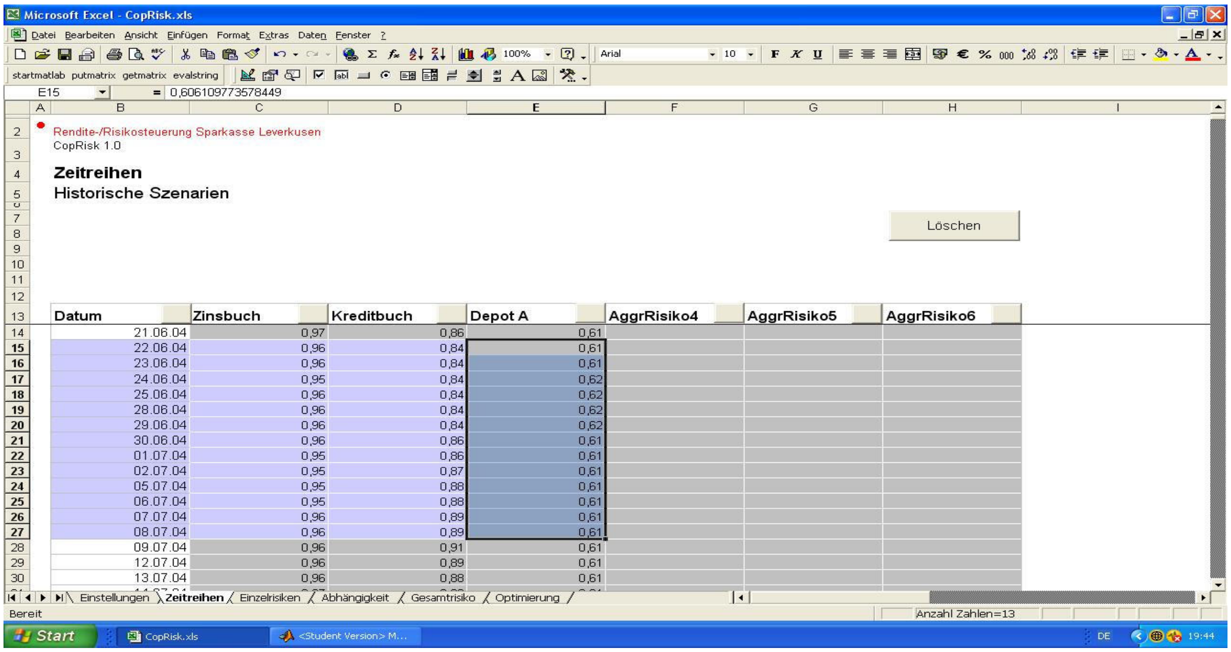

Figure 5.5: Screen-shot of MS-Excel worksheet `Zeitreihen` showing the exemplary selection of risk series of factors `Zinsbuch, Itraxx, Depot A` (normalized to July 31st, 2008).

On the worksheet `Einzelrisiken`, screen-shot in Figure 5.6, the user must in the first instance fill the individual (non-parametric) risk distributions. Then the user selects the data to be used in scenario generation. The discrete margin draws $x_i, i = 1, 2, 3$ need not necessarily have equal lengths, because the kernel uses them separately to reconstruct the marginal cdf's $\hat{F}_i^X, i = 1, 2, 3$.

The worksheet `Abhängigkeit` offers fitting and testing functionality. Here the user initiates the respective procedure, which estimates the selected copula model from the historic data. This is done by pushing the Kalibriere Abhängigkeit-button, which runs one of the Algorithms 5, 6, 8 or 7. The estimated set of parameters is immediately displayed on the worksheet. The true (historic) risk factor change series are further compared graphically to some stylized copula samples. Then the user tests the found estimators for their goodness of historical fit. Depending on the copula function of choice, the user runs one of the procedures 9, 10, 11 or 12 by clicking the Teste Abhängigkeit-button. The test statistics are printed on the worksheet. The numerical results are enforced by a quantile-quantile diagram, that shows significant divergence (if ever existent) of the test statistics (4.10) from a χ^2-distribution. Figures 5.7, 5.8, 5.9 and 5.10 give screen-shots of the calibration and testing interface.

The worksheet `Gesamtrisiko` features some conclusive summary statistics of the individual risk types. In detail, the mean return, the value-at-risk measure at the confidence levels $\alpha = 95\%$ and $\alpha = 99\%$, and the corresponding mean value-at-risk measures are displayed for all selected margins. By clicking the Aggregiere Risiken-button, the user

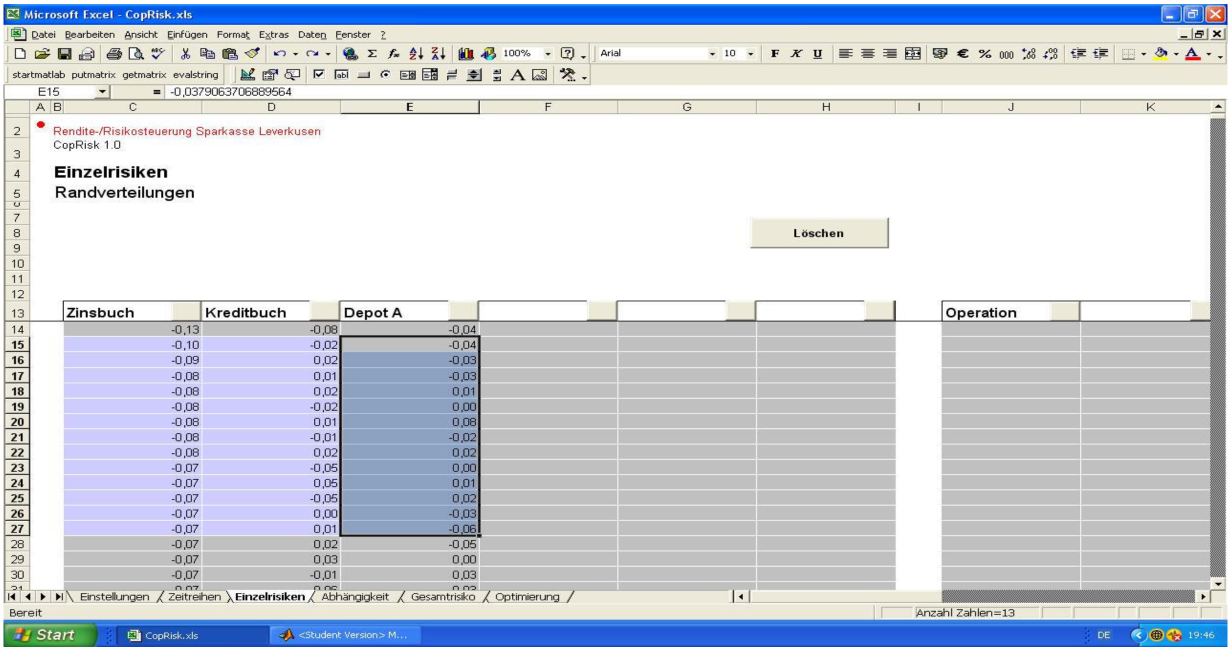

Figure 5.6: Screen-shot of MS-Excel worksheet `Einzelrisiken` showing the exemplary selection of sample losses in lines `Zinsbuch, Kreditbuch, Depot A` (normalized to July 31st, 2008).

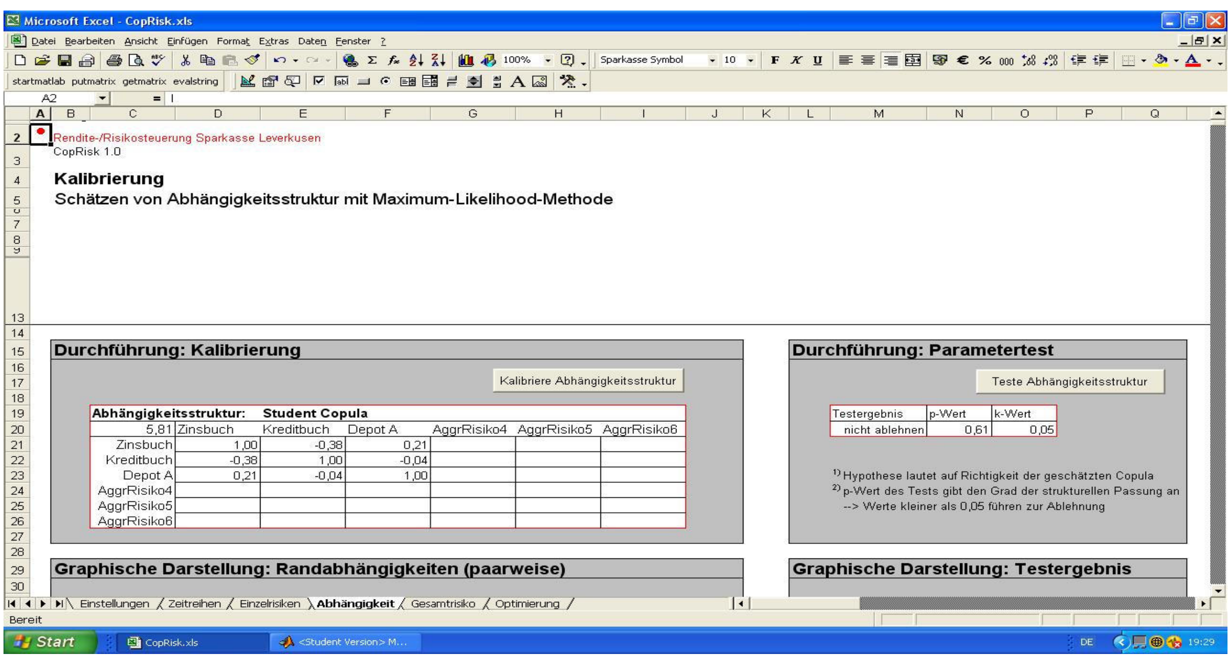

Figure 5.7: Screen-shot of MS-Excel worksheet `Abhängigkeit` showing the fitted parameters of a t copula and results of the ks-test.

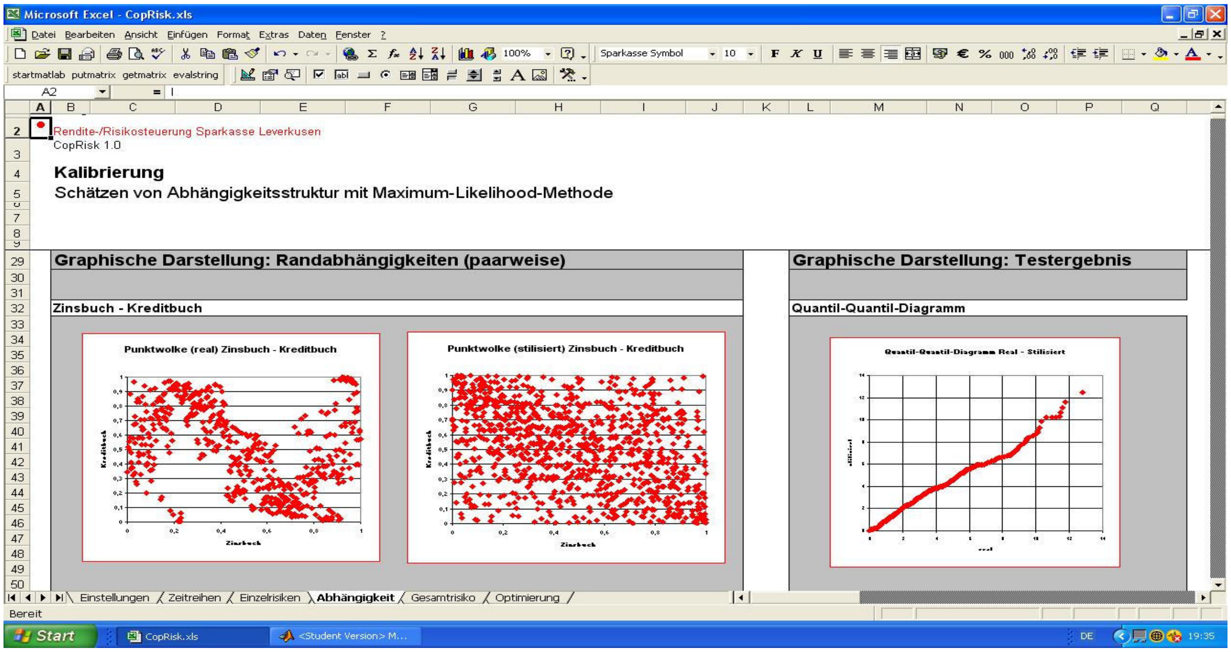

Figure 5.8: Screen-shot of MS-Excel worksheet `Abhängigkeit` showing real (1,2)-marginal pseudo-observations and fitted t copula draws, and t copula ks-test plot.

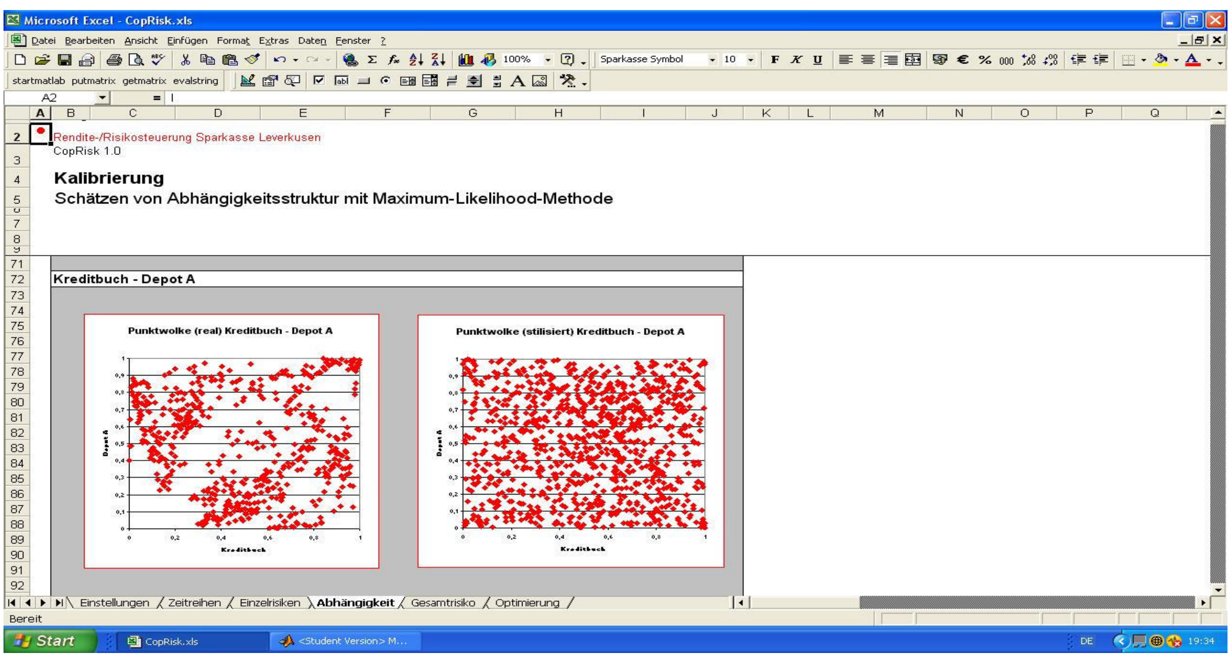

Figure 5.9: Screen-shot of MS-Excel worksheet `Abhängigkeit` showing real (2,3)-marginal pseudo-observations and fitted t copula draws.

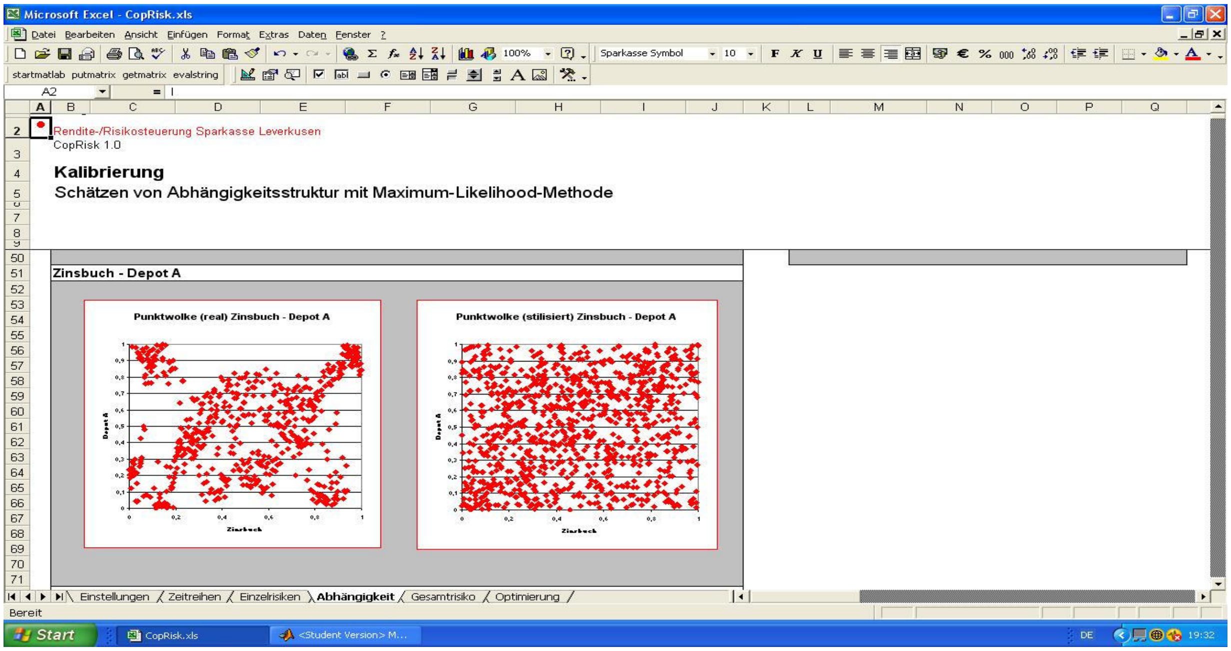

Figure 5.10: Screen-shot of MS-Excel worksheet `Abhängigkeit` showing real (1,3)-marginal pseudo-observations and fitted t copula draws.

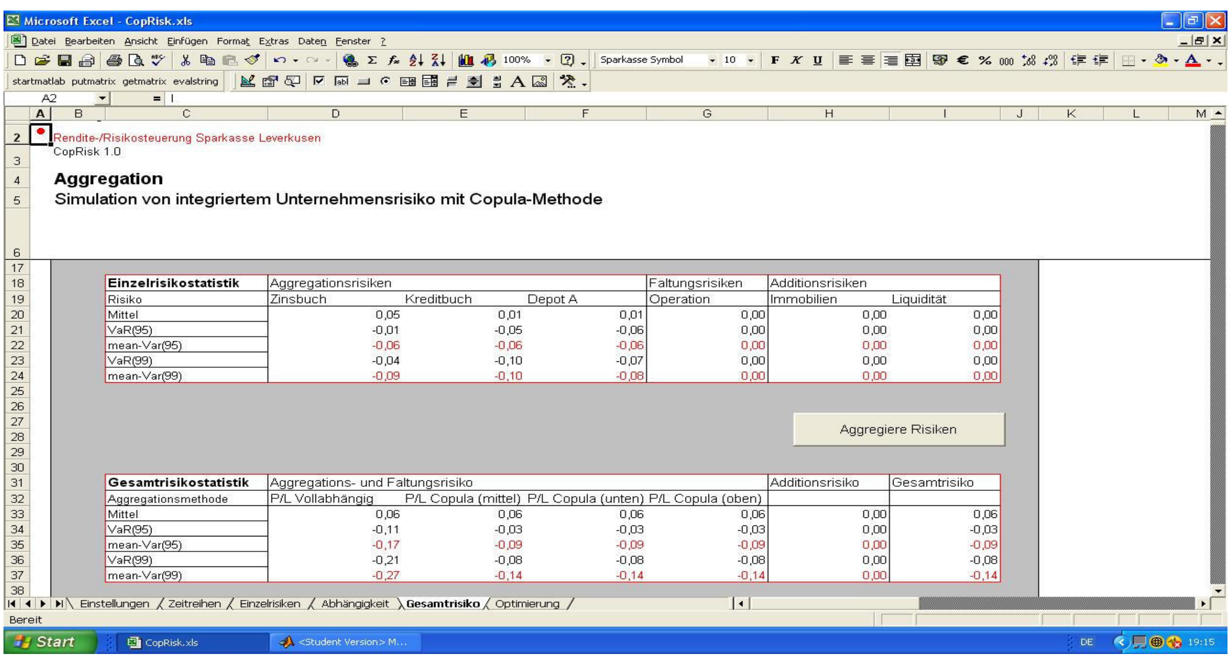

Figure 5.11: Screen-shot of MS-Excel worksheet `Gesamtrisiko` showing the risk measures on individual and enterprize level.

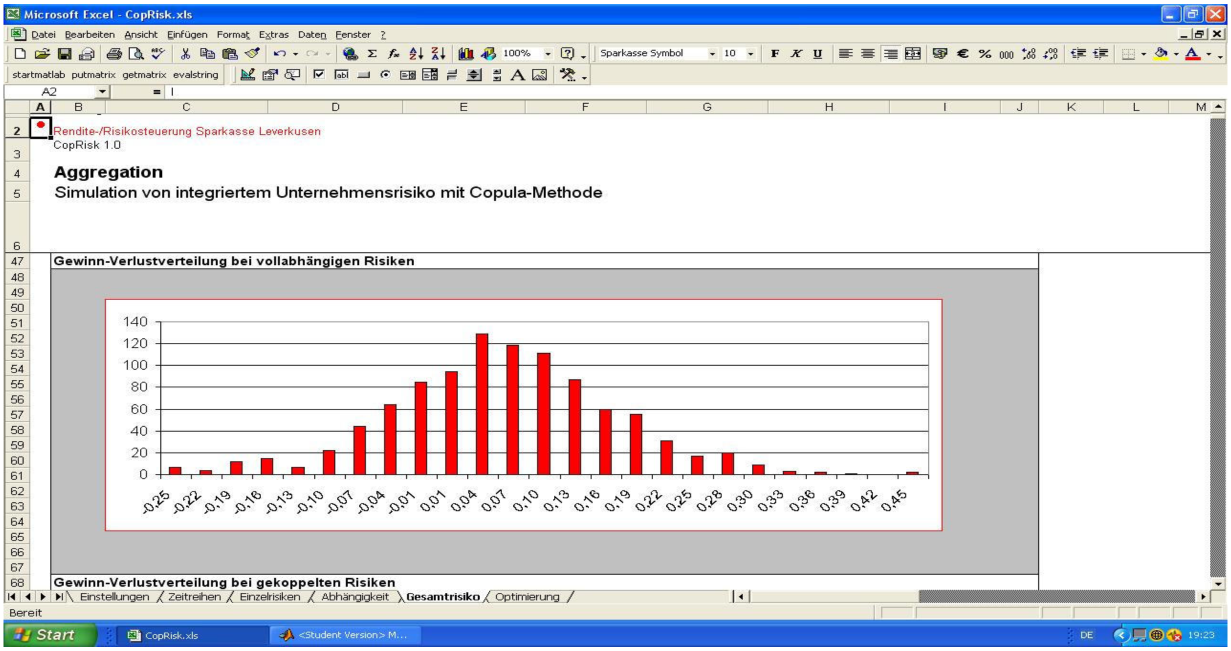

Figure 5.12: Screen-shot of MS-Excel worksheet `Gesamtrisiko` showing the overall loss distribution, if individual factors are simply added.

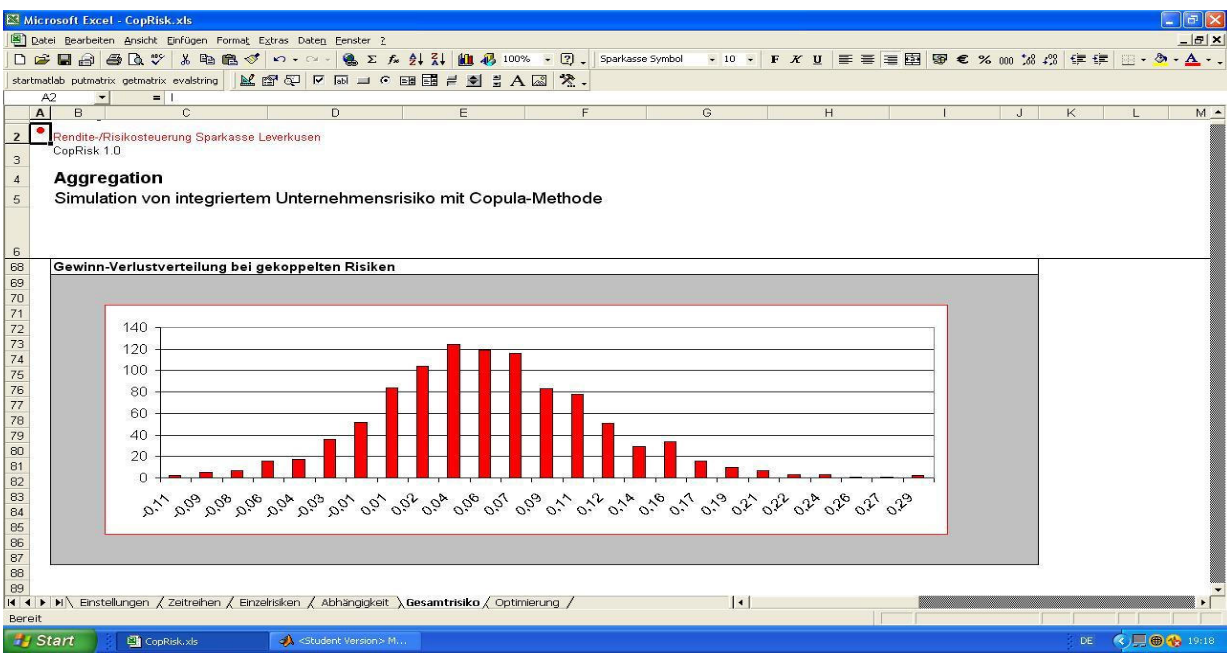

Figure 5.13: Screen-shot of MS-Excel worksheet `Gesamtrisiko` showing the overall loss distribution, if individual factors are aggregated by fitted t copula.

copula	$\rho_{1,2}$	$\rho_{2,3}$	$\rho_{1,3(\|2)}$	ν	hyp.	l.	p	ks
Gauss	-0.20	0.03	0.19	n.a	n.r.	22.429	0.12	0.07
t	-0.38	-0.04	0.21	5.81	n.r.	38.423	0.46	0.05
d-vine	-0.20	0.03	0.20	n.a	n.r.	22.431	0.08	0.08
c-vine	-0.20	0.07	0.19	n.a	n.r.	22.431	0.09	0.07

Table 5.1: Calibration and testing results (n.a. = not available, n.r. = not rejected).

commands to do the aggregation of risks. Here the kernel generates M independent samples $x^{(k)}_{\tau+1}, k = 1, ..., M$, using one of the Algorithms 1, 2, 3 or 4. These are subsequently transformed into samples $L^{(k)}_{\tau+1}, k = 1, ..., M$ of the integrated loss distribution by the specific loss operator

$$l_\tau(x) = \sum_{i=1}^{3} x_i.$$

Then the above risk measures are computed and printed on the worksheet. The distributional statistics, that correspond to complete dependence between the margins, are also printed out for comparison. Figure 5.11 gives a screen-shot of the aggregation interface. The respective frequency bar charts are given for graphical user support. This is screen-shot in Figures 5.12 and 5.13.

Empirical Results We run the over-all machinery on real business data and examine the risk statistics at Sparkasse Leverkusen with respect to the dependence structure at a one year horizon.

Table 5.1 shows the numerical results of the parameter estimation and the goodness-of-fit test for the cases of the Gauss copula, the t copula, the canonical pair copula and the Dvine pair copula. The estimations are based on risk factor series, which date from the 21st of June 2004 through the 31st of July 2007, and a one year roll over $\Delta = 250$.

- The business lines Treasury and Credit have a significant negative dependence, the risk factors Credit and Depot-A are merely independent (in the sense of linear correlations), and the factors Treasury and Depot-A show a moderate positive linear dependence.

- The degrees of freedom ν of the t copula is quite low; there is some extreme dependence present in the data.

- All of the p-values are greater than 0.05; the tests result in not to disclaim the hypothesis that the estimated copula function is the true dependence concept between the empirical margins.

- The likelihood of the estimated t copula function is well ahead those of the Gauss, the Dvine or the canonical pair copula, which are about at the same level.

- The p-statistic (or the ks-statistic) is optimal for the t copula; the t copula is of best absolute fit, followed by the Cvine copula, the Dvine copula and the Gauss copula.

Thus the t copula function fits best to business at Sparkasse Leverkusen. Then we forego the aggregation of individual risk distributions on the basis of the Gauss copula, the canonical and the D-vine pair copula.

Figures 5.8, 5.9, 5.10 (which we already used for documentation) give further graphical support of the goodness-of-fit of the t copula with respect to the empirical data set.

- The stylized scatters are comparable to the real data pairs; the t copula captures the major characteristics of the empirical (margin) copulas such as tendency and extreme value dependence.

- The quantiles of the ks-test statistics do not diverge from the quantiles of the χ^2-distribution; the goodness-of-fit of the t copular is satisfactory.

Altogether, the t copula sustains a graphical comparison with the historical values.

Figures 5.11, 5.12 and 5.13 (which we already used for documentation) screen-shot the conclusions of the enterprize risk distribution. These are found by Monte-Carlo simulation with 1000 scenarios of coupled risk factor changes using the fitted t copula model. The statistics of the individual risk distributions and the rudimentary addition of these serve as a benchmark for the copula based integration.

- The institute's loss distribution, that results from a pragmatic summation of sectors, has heavier tails compared to the aggregated losses.

- The value-at-risk measure at the confidence level $\alpha = 95\%$ is reduced by more than 50% when including a (non-trivial) dependence structure.

- The value-at-risk measure at the confidence level $\alpha = 99\%$ is reduced by more than 50% on account of the interrelations between the various business activities.

- The mean value-at-risk measure at both confidence levels $\alpha = 95\%$ and $\alpha = 99\%$ is cut down almost by half, if we use a copula based risk management approach.

5.4 Summary

In this chapter we applied the copula methodolgy, which was derived in the previous Chapters 2, 3 and 4, to real business at Sparkasse Leverkusen. Our major achievement here is the implementation of a sophisticated risk management solution.

With a view to the supervisory standards, we came across modern risk management as a quantitative discipline, that involves capital adequacy decisions based on multidimensional risk change distributions. We deemed copula models to be appropriate here to assess the overall loss distribution by separately coupling some individually profiled risk types. This was related to the traditional accountancy of one-factor risk silos.

After formulating the multi-source risk distribution on enterprize level, we discussed how the value-at-risk statistic may be used to quantify the charges on the firm specific risk exposure and to meet the regulatory capital requirements.

Using the MS-Excel interface together with the MatLab functionality, we built the user friendly implementation CopRisk 1.0, which makes use of the entire machinery developed before. The dedicated task therein was to design extensible modules for the estimation, simulation and evaluation of the overall risk situation. We subdivided the programm into several worksheets to account for

- the update of historical and distributional properties of individual risks
- the estimation and testing of a copula based model on the basis of historical risk series
- the simulation of the estimated risk model and
- the compiling of distributional facts of the overall year ahead risk exposure.

We documented each of these modules by way of screen-shots.

Finally, we ran CopRisk 1.0 on some actual (but anonymized) risk data. This revealed that the t copula suits best the risk landscape at Sparkasse Leverkusen. Moreover, the subsequent Monte Carlo simulation of risk scenarios showed that the copula based method cuts down the overall risk exposure massively.

Part II

Lévy copula methods

Chapter 6

Introduction

The modern theory of stochastic processes is undergoing an evident renaissance of interest in Lévy processes. In view of non-Gaussian phenomena in fields as diverse as meteorology, finance and insurance, researchers and practitioners have rediscovered the convenience of Lévy processes to cope with the rapidly evolving ramifications of Brownian motion in a simple and flexible way. Multidimensional stochastic modelling still continues to be dominated by Gaussian processes. This produces an undesirable imbalance between available models and possible applications.

Brownian motion has long played the dominant role as the driving process for modelling price movements. This was mainly due to the known technology for handling diffusion processes. During the past several years, Lévy processes have become increasingly popular in the theory of continuous-time processes. The keen interest in discontinuous models is mainly drawn by the eminent presence of jumps in observed prices movements. These diffusions can not handle. Another peace of evidence for discontinuous price behavior is seen from option quotes. The very existence of a market for (out-of-the money) short-term options proves that the market participants recognize large price movements over short periods of time. Diffusion based models do not get the implied volatility surfaces of these options unless by unrealistic high values of volatility of volatility. This is avoided in the wider class of Lévy processes, which cope with structural breaks and extremely irregular behavior including jumps, bursts and spikes; Lévy processes prepare to provide a better fit to real life data. In this regards, the literature became very innovative as for univariate phenomena. The scene of multidimensional stochastic processes still resorts in large parts to diffusion based models and a Gaussian concept of dependence. But the normality assumption becomes more and more obsolete in a joint sense, too.

Here, Lévy copulas offer a versatile modelling approach insofar as they enable to separate the dependence structure and the marginal aspects of a multidimensional stochastic process. Thus Lévy copulas handle non-normality of the dependence structure apart from the possibly extreme marginal behavior. This responds to the downright need for a flexible coupling of one-dimensional Lévy models, that cope with non-Gaussian margins already. Hence Lévy copula functions allow to model dependence on its own. Then our designated target is to develop practicable Lévy copula methods for multidimensional processes. Important issues in the part include:

- the use of Lévy copula functions as an appropriate summary of the dynamic dependence structure in multidimensional Lévy processes separate from marginal aspects,
- the up-to-date construction patterns of parametric dependence models including the Archimedean, the modular and the canonical type Lévy copula,
- the emphasis on structural properties of parametric Lévy copulas with regard to numerical simulation,
- the development of a multidimensional modelling framework for option pricing with Lévy processes,
- the application of Lévy copula methods to the pricing of foreign exchange options and the tracking of market implied dependence structures.

With this agenda, we fit into the fairly novel discussion on modelling complex jump dependence with Lévy copulas, which was pioneered by Tankov [77]. The issue of derivative pricing in Lévy models is dealt with in a multitude of publications.

The basic properties of Lévy processes have been well understood for a long time from Paul Lévy's characterization in the 1930's. Recent discourses on the basic knowledge of Lévy processes include the standard textbooks by Bertoin [12] and Sato [69], of which the latter gives special emphasis to the correspondence between Lévy processes and infinitely divisible distributions. Another adequate reference is Raible [64]'s celebrated dissertation, that touches upon theoretical, numerical and empirical issues. The textbook edited by Barndorff-Nielsen et al. [5] offers a state-of-the art survey on the theory, generalization and application of Lévy processes. A cross-referenced collection of various articles devoted to distributional and path properties as well as to the problems of simulation and statistical estimation is given there. In this collection, a large interest is taken in the use of Lévy processes in finance. Here Barndorff-Nielson and Shephard [7] deal with the transfer of the main stylized features of financial series into Lévy model logic. Eberlein [23] takes up the issue of reproducing empirical finance data, where he focuses on generalized hyperbolic Lévy motions for asset price models. The estimation of parameters in Lévy models is discussed in Nolan [58]. The methods derived therein are applied to exchange rate and stock price data. Series representations of Lévy processes, which are explained in Rosinski [67], prepare the simulation of sample paths. Then Cont and Tankov [20] have taken on the writing of an all-embracing discourse of financial modelling with jump processes in textbook format. They present an understandable and convenient overview of the use of Lévy processes in financial modelling. Their analysis specifically expands to the numerical calibration of a Lévy motion to market data and to the separate modelling of dependence structures via copula functions. In these points the book reviews the main contents of Tankov [77]'s pathbreaking dissertation. A detailed characterization of dependence of multidimensional Lévy processes using Lévy copulas can also be found in Kallsen and Tankov [45]. Tankov [78] applies the concept of Lévy copulas to the valuation of multi-asset options for the first time. Luciano and Schoutens [51] gives an alternative approach to multi-asset jump modelling on the basis of subordination and ordinary copulas with application to equity and credit. In contrast, the articles by Eberlein and Raible [30], Eberlein and Özkan [28] and Eberlein and Koval [27] are the basis of another collection of publications. These involve many discourses on the theory and valuation of multidimensional Lévy processes in term structure modelling. The main result is an extension of the Libor market model to the Lévy world. Change of measure techniques and pricing methods in multidimensional Lévy models are treated therein on the way. Building on these works, Eberlein and Papapantoleon [29] and Eberlein et al. [26] discover symmetries in the pricing of non-standard derivatives. This reference includes, for example, the valuation of multi-asset options in terms of plain vanilla contracts, that are written on some (artificial) one-dimensional underlying. In the present part we choose a combination of these subjects so as to develop a deep methodology for using Lévy copula functions in finance.

Our declared intention is to merge, on a firm theoretical groundwork, the conceptual properties of a Lévy copula with numerical issues. The application of dynamic dependence structures to the pricing of complex options is a secondary objective. One-dimensional models are already available from an abundance of publications (partly mentioned above). So we address the integration of dependence issues and univariate pricing techniques. In this part we distance ourselves from:

- the parametric modelling of univariate Lévy processes as for the generalization of Brownian motion,
- the compound Poisson processes and the related non-copula based dependence structures,
- a guide to dependence measures and orderings.

In consequence, we mostly forego the debate on one-dimensional Lévy processes and target Lévy copula methods on their own. Our main contribution is twofold. On the one hand, we develop complex Lévy copula models (further) with emphasis on probabilistic interpretations and numerical convenience. We make good use of the model features in order to help on the sampling of multidimensional Lévy processes. On the other hand, we formulate the calibration of a Lévy copula on the basis of pricing symmetries in exponential Lévy models. Using a non-parametric approach, we accomplish to back out the dependence structures from option quotes in the foreign exchange market.

The part is organized as follows. On the basis of a concise guide to multidimensional Lévy processes, in Chapter 7 we introduce Lévy copula functions as the conceptual analogue to ordinary copulas for modelling dependence between dynamic Lévy motions. We overlook basic properties and echo how Lévy copulas facilitate the formulation of dependence. We survey the (as yet) short list of existing parametric models. Then we go into details about the sophisticated modular design and the self-issued canonical approach.

In Chapter 8 we treat the subject matter of path generation with Lévy copula functions. We focus on series representations of multidimensional Lévy processes here and give both reasons for and instances of their use in simulation. We argue in particular that the modular and the canonical construction is very much suited to simulation by series representations.

In Chapter 9 we set up the financial market environment and establish exponential Lévy processes for asset price modelling. Then we go into details about measure transformations. Here we describe the relation between the martingale property and the characteristics of a Lévy process. Then we concern the option pricing problem and discuss symmetry techniques for the valuation of multi-name derivatives. Semi-analytical pricing methods, which are based on Fourier inversions, apply here.

In Chapter 10 the findings of Chapters 7 and 9 are applied to the calibration of a two-rate foreign exchange market model. We examine how to extract the model parameters from market quotes via a least-squares approximation. This involves the regularization of the eventual instabilities in an ill-posed problem. In closing this part, we then present the numerical results of fitting a dependence structure between the USD/JPY rate and the EUR/JPY rate. Here we search for the market implied Lévy copula, that reproduces liquid prices of options written on the USD/EUR cross rate.

Chapter 7

Lévy copula functions

In Chapter 2 we introduced the notion of a copula function to model the dependence in a multivariate distribution. This chapter transfers the main results on the coupling of static random variables to the case of multidimensional Lévy processes. We give a precise definition of Lévy copula functions and discuss the basic properties. Then we examine fundamental and advanced approaches to dynamic dependence modelling.

Section 7.1 defines Lévy copula functions of general Lévy processes and gives the basic properties and interpretations. Section 7.2 introduces the Archimedean type construction of Lévy copulas modelled after A.1. The bivariate Clayton Lévy copula is analyzed here. Then Section 7.3 is devoted to high dimensional modular Lévy copulas. It is shown that a building block design, which is due to Tankov [77], enables to arrange general copula functions by way of simple ones. Then Section 7.4 develops the new family of canonical Lévy copulas, which are made from ordinary and bivariate Lévy copula functions.

The derivations in this chapter follow the pioneering thesis by Tankov [77] and the sequel textbook by Cont and Tankov [20]. We remain also close to Kallsen and Tankov [45]'s publication on Lévy copulas for multidimensional jump processes. The model findings of Sections 7.3 and 7.4 present new material in large parts.

The following guide to Lévy processes prepares for the modelling of dynamic dependence structures. We explain general definitions and basic properties enough for what is to come [see 12, 69, for ample discourses].

Definition 7.1. A stochastic process $\{X_t\}$ on $\mathbb{R}^d$ with $X_0 = \mathbf{0}$ is a *Lévy process*, if the following conditions are satisfied:

(1) For any choice of $n \geq 1$ and $0 \leq t_0 < t_1 < ... < t_n \leq T$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, ..., X_{t_n} - X_{t_{n-1}}$ are independent (independent increments property).

(2) The distribution of $X_{t+s} - X_t$ does not depend on t (stationary increment property).

(3) For every $t \in [0, T]$ and $\varepsilon > 0$ it holds $\lim_{s \to t} P[|X_s - X_t| > \varepsilon] = 0$ (stochastic continuity property).

(4) There is $\Omega_0 \in \mathcal{F}$ with $P[\Omega_0] = 1$ such that, for every $\omega \in \Omega_0$, $X_t(\omega)$ is right-continuous in $t \in [0, T]$ and has left limits in $t \in [0, T]$.[1]

The most fundamental example of a Lévy process is the Brownian motion. A stochastic process $\{X_t\}$ on $\mathbb{R}^d$ is a *Brownian motion*, if it is a Lévy process, and if, for $t \in (0, T]$, X_t has a Gaussian distribution with mean 0 and covariance matrix $t\mathbf{1}$ and $\{X_t\}$ is almost surely continuous in $t \in [0, T]$. We reserve the notation $\{B_t\}$ for the Brownian motion. As shown in Sato [69, Proposition 5.2], for every $i \in \{1, ..., d\}$, the component $\{B_t^i\}$ of $\{B_t\}$ is a Brownian motion on $\mathbb{R}$ and $\{B_t^1\}, ..., \{B_t^d\}$ are independent. For $\gamma \in \mathbb{R}^d$ and a symmetric positive definite matrix $A \in \mathbb{R}^{d\times d}$, we call the process $\gamma t + A^{1/2} B_t$ a *Brownian motion with drift γ and covariance matrix A.*

Other than Brownian motion itself, condition (4) in Definition 7.1 allows a Lévy process $\{X_t\}$ to have discontinuities $\Delta X_t := X_t - X_{t-}$, that are seen as sudden events. The jump behavior is described by the *jump measure* μ, which is defined on $[0, T] \times \mathbb{R}^d$ by

$$\mu(A) = \sharp\{(t, \Delta X_t) \in A\}.$$

For every measurable set $A \subset \mathbb{R}^d$, $\mu([t_1, t_2] \times A)$ counts the number of jump times of $\{X_t\}$ between t_1 and t_2 such that their jump sizes are in A. Averaging this number over the unit time interval leads us to the celebrated *Lévy measure* ν, which is defined on $\mathbb{R}^d$ by

$$\nu(A) = E[\sharp\{t \in [0, 1] : \Delta X_t \neq 0, \Delta X_t \in A\}], A \in \mathcal{B}(\mathbb{R}^d) \tag{7.1}$$

For every measurable set $A \subset \mathbb{R}^d$, $\nu(A)$ is the expected number, per unit time, of jumps whose size belongs to A. The Lévy measure is a not necessarily finite Radon measure on $\mathbb{R}^d$. Hence $\{X_t\}$ may have an infinite number of small jumps on $[0, T]$ while $\nu(A)$ is finite for any compact set A such that $\mathbf{0} \notin A$.

The continuous and the discontinuous part of a sample path are distinct and can be decomposed:

Theorem 7.2 ([20], Proposition 3.7)**.** *Let $\{X_t\}$ be a Lévy process on $\mathbb{R}^d$ and ν its Lévy measure. Then*

(1) ν is a Radon measure on $\mathbb{R}^d$ and verifies:

$$\nu(\{\mathbf{0}\}) = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} (1 \wedge |x|^2)\nu(dx) < \infty$$

(2) The jump measure μ of $\{X_t\}$ is a Poisson random measure[2] on $[0, T] \times \mathbb{R}^d$ with intensity measure $\nu(dx)dt$.

[1] This is equivalent to saying that $\{X_t\}$ (or at least a unique modification of it) is a cadlag stochastic process [cf. 20, Definition 3.1].

[2] Poisson random measures are uniquely determined by their intensity measure as described in Cont and Tankov [20].

(3) There exist a Brownian motion $\{B_t\}$ on $\mathbb{R}^d$ with drift $\gamma \in \mathbb{R}^d$ and covariance matrix $A \in \mathbb{R}^{d\times d}$ such that

$$
\begin{aligned}
X_t &= \gamma t + A^{1/2}B_t + X_t^l + \lim_{\varepsilon\downarrow 0} \widetilde{X}_t^\varepsilon, \text{ where} \\
X_t^l &= \int_{|x|\geq 1, s\in[0,t]} x\mu(ds\times dx) \text{ and} \\
\widetilde{X}_t^\varepsilon &= \int_{\varepsilon\leq|x|<1, s\in[0,t]} x(\mu(ds\times dx) - \nu(dx)ds).
\end{aligned}
\tag{7.2}
$$

The terms in (7.2) are independent, and the convergence in the last term is almost sure and uniform in t on $[0,T]$.

Property (3) in Proposition 7.2, called the *Lévy-Ito decomposition* of $\{X_t\}$, allows us to distinguish between the continuous and the discontinuous part of a Lévy process. Let us therefore denote by $\{X_t^C\}$ the continuous part and by $\{X_t^J\}$ the jump part of a Lévy process. Proposition 7.2 also states that every Lévy process $\{X_t\}$ is uniquely defined by a drift γ, a covariance matrix A and a Lévy measure ν. The triplet (γ, A, ν) is hence called the *characteristic triplet*[3] of $\{X_t\}$.

The characteristic triplet is also a determinant of the characteristic function of a Lévy process:

Theorem 7.3 ([20], Theorem 3.1)**.** *Let $\{X_t\}$ be a Lévy process on $\mathbb{R}^d$ with characteristic triplet (γ, A, ν). Then*

$$
\begin{aligned}
E[e^{iz.X_t}] &= e^{t\psi(z)}, \quad z\in\mathbb{R}^d \\
\text{with } \psi(z) &= -\frac{1}{2}z.Az + i\gamma.z + \int_{\mathbb{R}^d}(e^{iz.x} - 1 - iz.x\mathbf{1}_{|x|\leq 1})\nu(dx).
\end{aligned}
\tag{7.3}
$$

The function ψ in (7.3) is continuous [cf. 20, Proposition 3.2] and called the *characteristic exponent* of $\{X_t\}$. The exponential structure shows in particular that the characteristic function of a Lévy processes is multiplicative in t. That is to say a Lévy process has an infinitely divisible distribution[4] for every t. We denote by ψ^C and ψ^J those parts of the characteristic exponent, that are associated to the continuous and the jump part of the Lévy process, respectively.

(7.2) and (7.3) require a compensation of small jumps in general. If the Lévy measure satisfies

$$
\int_{|x|\leq 1} |x|\nu(dx) < \infty, \tag{7.4}
$$

[3] Sato [69] uses the term generating triplet, instead.

[4] Sato [69] gives a comprehensive discourse of the interrelation of Lévy processes and infinitely divisible distributions.

then the discontinuities of the path of $\{X_t\}$ may be expressed as the sum of jumps between 0 and t [cf. 20, Corollary 3.1]:

$$X_t^J = \int_{[0,t]\times\mathbb{R}^d} x\mu(ds \times dx) = \sum_{s\in[0,t]}^{\Delta X_s \neq 0} \Delta X_s. \tag{7.5}$$

As a consequence, the characteristic exponent of $\{X_t\}$ may be written as [cf 20, Corollary 3.1]

$$\psi(z) \;=\; -\frac{1}{2}z.Az + i\gamma^b.z + \int_{\mathbb{R}^d}(e^{iz.x}-1)\nu(dx),$$

where $\gamma^b = \gamma - \int_{|x|\leq 1} x\nu(dx)$ is called the *drift* of $\{X_t\}$. Condition (7.4) relates to the behavior of ν around the origin.[5] If the tails of the Lévy measure ν satisfy

$$\int_{|x|>1} |x|\nu(dx) < \infty$$

instead, then the characteristic exponent of $\{X_t\}$ may be written as [cf 20, Section 3.4]

$$\psi(z) \;=\; -\frac{1}{2}z.Az + i\gamma^c.z + \int_{\mathbb{R}^d}(e^{iz.x}-1-iz.x)\nu(dx),$$

where $\gamma^c = \gamma + \int_{|x|>1} x\nu(dx)$ is called the *center* of $\{X_t\}$. By convention, the characteristic triplet is always expressed in terms of γ rather than γ^b or γ^c. Next, we give some fundamental Lévy processes, stressing the various representations.

Example 7.4. Let $\{X_t\}$ be a Brownian motion on $\mathbb{R}^d$ with drift $\gamma \in \mathbb{R}^d$ and variance $A \in \mathbb{R}^{d\times d}$. Then the characteristic triplet of $\{X_t\}$ is $(\gamma, A, 0)$ and the characteristic exponent can be represented as

$$\psi(z) = -\frac{1}{2}z.Az + i\gamma.z.$$

Example 7.5. Let $\{X_t\}$ be a Lévy process on $\mathbb{R}^d$ with triplet (γ, A, ν) and discrete Lévy measure of the form

$$\nu = \sum_{k=1}^{N} \nu_k \delta_{x^{(k)}}, \tag{7.6}$$

where $\delta_{x^{(k)}}$ the Dirac-measure at $x^{(k)} \in \mathbb{R}^d$. With a view to (7.1), measure (7.6) certainly is a proper Lévy measure, and it satisfies condition (7.4). Hence, its characteristic exponent has the following representation:

$$\psi(z) = -\frac{1}{2}z.Az + i\gamma^b.z + \sum_{k=1}^{N}(e^{iz.x^{(k)}}-1)\nu_k,$$ [6]

[5] Cont and Tankov [20] show that a Lévy process $\{X_t\}$ on $\mathbb{R}^d$ with triplet (γ, A, ν) is of finite variation if and only if $A = 0$ and (7.4) is satisfied.

[6] Cont and Tankov [20] argue that Lévy processes of this kind are superpositions of independent Poisson processes with different jump sizes.

where γ^b is as in Corollary (7.6). This $\{X_t\}$ surely satisfies condition (7.6) on the tail integrability of ν. Consequently, we may just as well express the characteristic exponent as

$$\psi(z) = -\frac{1}{2}z.Az + i\gamma^c.z + \sum_{k=1}^{N}(e^{iz.x^{(k)}} - 1 - iz.x^{(k)})\nu_k.$$

Example 7.6. Let $\{X_t\}$ be a $\mathbb{R}$-valued Lévy process with triplet (γ, A, ν) and discrete Lévy measure of the form

$$\nu(x) = \frac{\lambda}{\delta\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\delta^2}}, \tag{7.7}$$

where $\lambda, \delta > 0$ and $\mu \in \mathbb{R}$. With a view to (7.1), measure (7.6) certainly is a proper Lévy measure, and it satisfies condition (7.4). Hence, its characteristic exponent has the following representation:

$$\psi(z) = -\frac{1}{2}\sigma^2 z^2 + i\gamma^b z + \lambda(e^{-\delta^2 z^2/2 + i\mu z - 1}),$$

where γ^b is as in Corollary (7.6). The process $\{X_t\}$ is known as the *Merton process.*

Example 7.7. Let $\{X_t\}$ be a $\mathbb{R}$-valued Lévy process with triplet (γ, A, ν) and Lévy measure of the form

$$\nu(x) = \frac{\lambda_+}{x^{1+\alpha}}\mathbf{1}_{x>0} + \frac{\lambda_-}{x^{1+\alpha}}\mathbf{1}_{x<0}, \tag{7.8}$$

where λ_+, λ_- are positive constants and $0 < \alpha < 2$. In this case, one can show [cf. 69, Theorem 14.15] that there exists a triplet (β, τ, c) with $c > 0$, $\beta \in [-1, 1]$, and $\tau \in \mathbb{R}$ such that

$$\begin{aligned}
\psi(z) &= -\frac{1}{2}A^2 z^2 + i\tau z - cz^\alpha(1 - i\beta\tan(\frac{\pi\alpha}{2})\operatorname{sgn}(z) \text{ for } \alpha \neq 1,\\
\psi(z) &= -\frac{1}{2}A^2 z^2 + i\tau z - cz(1 + i\beta\frac{2}{\pi}\operatorname{sgn}(z)\log(|z|)) \text{ for } \alpha = 1.
\end{aligned}$$

The parameter τ is identical with the drift γ^b if $0 < \alpha < 1$, and with the center γ^c if $1 < \alpha < 2$. The parameter β represents non-symmetry of the Lévy measure. In this sense, ν is symmetric, if and only if $\beta = 0$, and σ is the scaling parameter. The process $\{X_t\}$ is called an *α-stable process* and α is called the *stability index.*

The findings to come will involve linear transformations of Lévy processes, which are again Lévy processes:

Proposition 7.8 ([69], Proposition 11.10)**.** *Let $\{X_t\}$ be a Lévy process on $\mathbb{R}^d$ with generating triplet (γ, A, ν) and let $U \in \mathbb{R}^{r\times d}$ be a matrix. Then $\{UX_t\}$ is a Lévy process on $\mathbb{R}^r$*

with generating triplet (γ_U, A_U, ν_U) *given by*

$$\begin{aligned}
\gamma_U &= U\gamma + \int_{\mathbb{R}^d} Ux\,(1_E(Ux) - 1_D(x))\,\nu(dx) \\
A_U &= UAU' \\
\nu_U &= [\nu U^{-1}]_{|\mathbb{R}\setminus\{0\}},
\end{aligned}$$

where $\nu U^{-1}(B) = \nu(\{x : Ux \in B\}), D = \{x \in \mathbb{R}^d : |x| \leq 1\}, E = \{y \in \mathbb{R}^r : |y| \leq 1\}$ *and* $[\nu U^{-1}]_{|\mathbb{R}\setminus\{0\}}$ *is the restriction of* νU^{-1} *to* $\mathbb{R} \setminus \{0\}$.

As a special case, projections of Lévy processes are again Lévy processes.

7.1 Definition and basic properties

Lévy copulas have proven advantageous to model complex dependence structures of multivariate jump processes. We give definitions, properties and fundamental representatives of dependence functions of this kind.

The tail integrals of a Lévy measure are defined similar to the cdf of a probability distribution:

Definition 7.9. Let $\{X_t\}$ be a Lévy process on $\mathbb{R}^d$ with Lévy measure ν. The *tail integral* of ν is the function $U : (\mathbb{R}\setminus 0)^d \to \mathbb{R}$ defined by

$$U(x_1, ..., x_d) = \nu\left(\prod_{i=1}^{d} \mathcal{I}(x_i)\right) \prod_{i=1}^{d} \operatorname{sgn}(x_i),$$

where

$$\mathcal{I}(x) = \begin{cases} [x, \infty), & x > 0; \\ (-\infty, x], & x < 0. \end{cases}$$

Although the definition of the tail integral spares the axes, it can be shown [cf. 45, Lemma 3.5] that the Lévy measure is uniquely determined by the set $\{U_I : I \subset \{1, ..., d\}\}$ of its marginal tail integrals and vice versa.

Lévy copulas are defined on the analogy of copulas, which we refer to as *ordinary copulas* for distinction in the following:

Definition 7.10. A function $F : \mathbb{R}_\infty^d \to \mathbb{R}_\infty$ is called a *Lévy d-copula function* (or Lévy d-copula or Lévy copula), if

(1) $F(u_1, ..., u_d) \neq \infty$ for $(u_1, ..., u_d) \neq (\infty, .., \infty)$

(2) $F(u_1, ..., u_d) = 0$ if $u_i = 0$ for at least one $i \in \{1, ..., d\}$

(3) F is d-increasing

(4) $F_i(u) = u$ for any $i \in \{1, ..., d\}$, $u \in \mathbb{R}$.

Sometimes it is useful to limit the notion of a Lévy copula to Lévy processes whose Lévy measures ν are supported by $\mathbb{R}^d_+$:

Definition 7.11. A function $F : \mathbb{R}^d_{\infty,+} \to \mathbb{R}_{\infty,+}$ is called a Lévy copula, if

(1) $F(u_1, ..., u_d) \neq \infty$ for $(u_1, ..., u_d) \neq (\infty, .., \infty)$

(2) $F(u_1, ..., u_d) = 0$ if $u_i = 0$ for at least one $i \in \{1, ..., d\}$

(3) F is d-increasing

(4) $F_i(u) = u$ for any $i \in \{1, ..., d\}$, $u \in \mathbb{R}_{\infty,+}$.

We call a Lévy copula of this kind *positive* and denote it by F^+ for distinction. It can be shown [cf. 77, Section 4.5] that, if F^+ is a Lévy copula on $\mathbb{R}^d_{\infty,+}$, then it can be extended to a Lévy copula F on $\mathbb{R}^d_\infty$ by

$$F(u_1, ..., u_d) = \begin{cases} F^+(u_1, ..., u_d), & (u_1, ..., u_d) \in \mathbb{R}^d_{\infty,+}; \\ 0, & \text{otherwise.} \end{cases}$$

As a consequence, the results on general Lévy copulas also hold for the positive ones.

The next theorem parallels Sklar's theorem in the context of tail integrals instead of probability distributions:

Theorem 7.12 ([77], Theorem 4.8)**.** *Let ν be a Lévy measure on $\mathbb{R}^d$. Then there exists a Lévy copula F such that the tail integrals of ν satisfy*

$$U_I((x_i)_{i \in I}) = F_I((U_i(x_i))_{i \in I})$$

for any non-empty $I \subset \{1, ..., d\}$ and any $(x_i)_{i \in I} \in \mathbb{R}^{|I|}$. Conversely, if F is a d-dimensional Lévy copula and $\nu_1, ..., \nu_d$ are Lévy measures on $\mathbb{R}$ with tail integrals integrals $U_1, ..., U_d$, then there exists a unique Lévy measure $\mathbb{R}^d$ with one-dimensional marginal tail integrals $U_1, ..., U_d$.

Theorem 7.12 shows that Lévy copula functions characterize the dependence structure between marginal tail integrals. Conversely, Lévy copula functions are themselves determined by the joint and marginal tail integrals.[7]

Lévy copulas allow of a probabilistic interpretation, which is explained in the following. Let therefore F be a Lévy copula on $\mathbb{R}^d_\infty$, that satisfies the following continuity condition at infinity:

$$\lim_{(u_i)_{i \in I} \to \infty} F(u_1, ..., u_d) = F(u_1, ..., u_d)|_{(u_i)_{i \in I} = \infty} \tag{7.9}$$

[7] Barndorff-Nielsen and Lindner [6] show that any Lévy copula function itself defines a proper Lévy measure.

for all $I \subset \{1, ..., d\}$. This Lévy copula defines a positive measure μ on $\mathbb{R}^d$ with Lebesgue margins such that for each $a, b \in \mathbb{R}^d$ with $a \leq b$,

$$V_F((a, b]) = \mu((a, b]),$$

where V_F the volume function of F. The Lévy measure ν is described by the measure μ via

$$\nu(A) = \mu(\{u \in \mathbb{R}^d : f(u) \in A\}), \tag{7.10}$$

where $f : (u_1, ..., u_d) \mapsto (U_1^{-1}(u_1), ..., U_d^{-1}(u_d))$. By Tankov [cf. 77, and references therein], there exists a family, indexed by $\xi \in \mathbb{R}$, of positive Radon measures $K(\xi, \cdot)$ on $\mathbb{R}^{d-1}$, such that $\xi \mapsto K(\xi, du_2, ..., du_d)$ is Borel measurable and

$$\mu(du_1, du_2, ..., du_d) = \lambda(du_1) \otimes K(u_1, du_2, ..., du_d). \tag{7.11}$$

$\{K(\xi, \cdot)\}_{\xi \in \mathbb{R}}$ is called the *family of conditional probability distributions associated to the Lévy copula F*. $K(\xi, \cdot)$ is a probability distribution for almost all $\xi \in \mathbb{R}$. Its (conditional) cdf, denoted by K_ξ, has an explicit representation:

Theorem 7.13 ([77], Lemma 5.5)**.** *Let F be a Lévy copula on $\mathbb{R}_\infty$, satisfying* (7.9)*, and $\{K(\xi, \cdot)\}_{\xi \in \mathbb{R}}$ be the family of conditional probability distributions associated to F. Then, there exists a nullset N such that for every $\xi \in \mathbb{R} \setminus N$, K_ξ is a cdf represented by*

$$K_\xi(u_2, ..., u_d) = \operatorname{sgn}(\xi) \frac{\partial}{\partial \xi} V_F((\xi \wedge 0, \xi \vee 0] \times (-\infty, u_2] \times ... \times (-\infty, u_d]) \tag{7.12}$$

in every point $(u_2, ..., u_d)$ where K_ξ is continuous.

Theorem 7.13 associates a family of cdfs K_ξ to every general Lévy copula function F. In particular, the conditional cdf associated to a positive Lévy copula F^+ on $\mathbb{R}^d_{\infty,+}$, denoted by K_ξ^+, is of the form

$$K_\xi^+(u_2, ..., u_d) = \frac{\partial}{\partial \xi} K^+(\xi, u_2, ..., u_d). \tag{7.13}$$

The conditioning in (7.12) (and in (7.13)) is usually on the first variable, albeit the choice is arbitrary.

On account of Theorem 7.12, there exists a *canonical representation* of the density of a Lévy measure ν. We extend (omitting the proof in good faith) the results in Cont and Tankov [20, Proposition 5.8], that target the 2-dimensional case only:

Proposition 7.14. *Let F be a Lévy d-copula, continuous on $\mathbb{R}^d_\infty$, such that the density $\partial^d F(u_1, ..., u_d)/\partial u_1 ... \partial u_d$ exists on $\mathbb{R}^d$, and let $U_1, ... U_d$ be one-dimensional tail integrals with densities $\nu_1, ..., \nu_2$. Then*

$$\nu(dx_1, ..., dx_2) = \frac{\partial^d F}{\partial u_1 ... \partial u_d}(U_1(x_1), ..., U_d(x_d)) \nu_1(dx_1) \ldots \nu_d(dx_d)$$

is the Lévy density of a Lévy measure with marginal Lévy measures $\nu_1, ..., \nu_d$ and Lévy copula F.

Proposition 7.14 enables us to construct d-dimensional Lévy densities from 1-dimensional Lévy densities in conjunction with a smooth Lévy copula. If the Lévy copula and the Lévy measures are not sufficiently smooth, then the tail integrals are used instead of the densities [cf. 20, Section 5.6]. This is discussed only by way of example in the following:

Example 7.15. Let ν_1, ν_2 be discrete Lévy measures on $\mathbb{R}$ in the sense of Example 7.5 supported by a centered equidistant grid $\{x_k = (k - (N+1)/2)\Delta, k = 1, ..., N\}$ with grid size $\Delta > 0$, hence

$$\nu_i = \sum_{k=1}^{N} \nu_i^k \delta_{x_k}, \tag{7.14}$$

with weights $\nu_i^k \in \mathbb{R}, k = 1, ..., N$ for $i = 1, 2$. Then, for $i = 1, 2$, the tail integral U_i of ν_i is

$$U_i(x) = \operatorname{sgn}(x) \sum_{k=1}^{N} \nu_i^k 1_{\mathcal{I}(x_k)}(x). \tag{7.15}$$

Further let F be a Lévy 2-copula function, which is continuous on $\mathbb{R}^2_\infty$. Similar to Proposition 7.14, ν_1, ν_2 and F define a Lévy measure ν_{12} on $\mathbb{R}^2$:

$$\nu_{12} = \sum_{i,j=1}^{N} \nu_{12}^{ij} \delta_{(x_i, y_j)},$$

where the weights ν_{12}^{ij}, for $x_i, y_j > 0$ say, are given by

$$\begin{aligned} \nu_{12}^{ij} &= F(U_1(x_i), U_2(y_j)) - F(U_1(x_i), U_2(y_j + \Delta)) \\ &- F(U_1(x_i + \Delta), U_2(y_j)) + F(U_1(x_i + \Delta), U_2(y_j + \Delta)). \end{aligned} \tag{7.16}$$

If $f(u, v) = \partial F(u, v)/\partial u \partial v$ exists on $\mathbb{R}^2$, then, for small step sizes Δ,

$$\nu_{12}^{ij} \approx f(U_1(x_i), U_2(y_j)) \nu_1^i \nu_2^j, \quad i, j = 1, ..., N$$

is an approximation of (7.16).

Fundamental Lévy copula functions Similar to the ordinary case, there exist Lévy copulas, that correspond to complete negative dependence, independence and complete positive dependence of the jump behavior.

Independence among the components of a Lévy process is very much a Lévy measure property:

Lemma 7.16 ([77], Lemma 4.2). *Let $\{X_t\}$ be a Lévy process. Then $\{X_t^1\}, ..., \{X_t^d\}$ are* independent*, if and only if their continuous martingale parts are independent and the Lévy measure ν is supported by the axes*

$$\nu(B) = \sum_{i=1}^{d} \nu_i(B_i), \quad \textit{for all } B \in \mathcal{B}(\mathbb{R}^d \setminus \{0\}), \tag{7.17}$$

where for every i, ν_i denotes the i-th margin of ν and $B_i = \{x \in \mathbb{R} : (0, ..., 0, x, 0, ..., 0) \in B\}$.

This says that the marginal tail integrals $U_I((x_i)_{i\in I})$ vanish for all $I \subset \{1, ..., d\}$ and all $(x_i)_{i\in I} \in (\mathbb{R} \setminus \{0\})^{|I|}$:

Theorem 7.17 ([77], Theorem 4.10)**.** *Let $\{X_t\}$ be a Lévy process on $\mathbb{R}^d$. Then the* independence copula *is given by*

$$F_\perp(u_1, ..., u_d) = \sum_{i=1}^{d} u_i \prod_{j\neq i} 1_\infty(u_j). \tag{7.18}$$

Note that Lévy copula (7.18) fulfills the conditions of Definition 7.10 but is discontinuous at positive infinity.

Definition 7.18. Define $S_+ = \{x \in \mathbb{R}^d : \mathrm{sgn}(x_1) = ... = \mathrm{sgn}(x_d)\}$ and $S_- = \{x \in \mathbb{R}^2 : \mathrm{sgn}(x_1) \neq \mathrm{sgn}(x_2)\}$. Let $\{X_t\}$ be a Lévy process on $\mathbb{R}^d$. Its jumps are considered *completely positive dependent*, if there exists an increasing set D of S_+ such that $\Delta X_t \subset D, t \geq 0$. For $d = 2$, the jumps of $\{X_t\}$ are completely negative dependent, if there exists a decreasing[8] set D of S_- such that $\Delta X_t \subset D, t \geq 0$.

Complete dependence of jumps is not a Lévy measure property in general, but it can be described by the Lévy copula:

Theorem 7.19 (cf. [77], Theorem 4.11)**.** *Let $\{X_t\}$ be a Lévy process on $\mathbb{R}^d$, whose Lévy measure is supported by an ordered set $D \subset S$. Then the* complete positive dependence Lévy copula *is given by*

$$F_\|(u_1, ..., u_d) = \min(|u_1|, ..., |u_d|) 1_{S^+}(u_1, ..., u_d) \prod_{i=1}^{d} \mathrm{sgn}(u_i).$$

If $d = 2$, then the complete negative dependence Lévy copula *is given by*

$$F_|(u_1, u_2) = -\min(|u_1|, |u_2|) 1_{S^-}(u_1, u_2).$$

Conversely, if $F_\|$ or $F_|$ is a Lévy copula of $\{X_t\}$, then the Lévy measure is supported by a strictly ordered subset $D \subset S$. If, in addition, the tail integrals U_i of X^i are continuous and satisfy $\lim_{x\to 0} U_i(x) = \infty, i = 1, ..., d$, then the jumps of X_t are completely dependent.

Theorem 7.19 shows that the complete positive dependence and the complete negative dependence Lévy copula both resemble the complete positive dependence ordinary copula (2.5).

Remark 7.20. There exists a positive complete positive dependence Lévy copula in a natural way. The complete negative dependence Lévy copula does not possess a positive companion Lévy copula.

[8] Increasing and decreasing sets (or strictly ordered sets) are defined in [77]

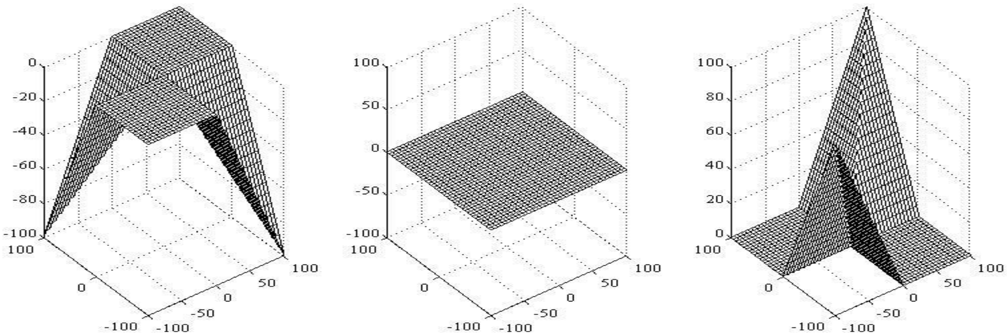

Figure 7.1: Surface of the complete negative dependence Lévy copula (left), the independence Lévy copula (center), and the complete positive dependence Lévy copula (right).

Figure 7.1 illustrates the complete negative dependence, the independence and the complete positive dependence Lévy 2-copula over a finite square.

- The complete dependence Lévy copulas are supported by two opposing quadrants each; complete dependence characterizes jumps of either unequal or equal signs.
- The graphs of the complete dependence Lévy copulas consist of pyramidal bricks; the complete dependence Lévy copulas extend the pattern of the maximum (ordinary) copula.
- The complete dependence Lévy copulas have a kink; they do not possess a continuous density.
- The independence Lévy copula is zero everywhere (except at positive infinity).

7.2 Archimedean Lévy copula functions

We have introduced the Archimedean construction to ordinary copulas just barely in Section A.1. This pattern applies to Lévy copulas, as well.

Definition 7.21 (cf. [77], Theorem 5.2)**.** Let $\phi : [-1, 1] \to [-\infty, \infty]$ be a strictly increasing continuous function with $\phi(1) = \infty, \phi(0) = 0, \phi(-1) = -\infty$, having nonnegative derivatives of order up to d on $(-1, 0)$ and $(0, 1)$, and satisfying

$$\frac{\partial^d \phi(e^t)}{\partial t^d} \geq 0, \ \frac{\partial^d \phi(-e^t)}{\partial t^d} \leq 0.$$

Let $\widetilde{\phi}(t) = 2^{d-2}\{\phi(t) - \phi(-t)\}$ for $t \in [-1, 1]$. Then

$$F(u_1, ..., u_d) = \phi\left(\prod_{i=1}^{d} \widetilde{\phi}^{-1}(u_i)\right) \tag{7.19}$$

is called a *general Archimedean Lévy copula.*[9]

The function ϕ is referred to as *(general) Archimedean copula generator* (in the sense of Lévy copulas). The Archimedean construction of positive Lévy copulas is simpler:

Definition 7.22 (cf. [77], Theorem 5.1)**.** Let $\psi : [0,1] \to [0,\infty]$ be a strictly decreasing continuous function with $\phi(0) = \infty, \phi(\infty) = 0$, having derivatives of order up to d on $(0,1)$, and satisfying

$$(-1)^d \frac{\partial^d \psi(t)}{\partial t^d} \geq 0.$$

Then

$$F^+(u_1, ..., u_d) = \psi\left(\sum_{i=1}^{d} \psi^{-1}(u_i)\right) \tag{7.20}$$

is called an *Archimedean Lévy copula.*[10]

The function ψ is referred to as *Archimedean copula generator* (in the sense of Lévy copulas). Lévy copulas of the form (7.19) or (7.20) are just about the only parametric Lévy copulas, which are discussed in the literature.

Clayton Lévy copula functions The Lévy Clayton copula resembles the ordinary Clayton copula in terms of constructions, hence the naming:

Definition 7.23. Let $\psi(t) = t^{-1/\theta}$. Then ψ is an Archimedean generator producing the one-parameter Archimedean Lévy d-copula

$$F^+(u_1, ..., u_d) = \left(\sum_{i=1}^{d} u_i^{-\theta}\right)^{-1/\theta}, \tag{7.21}$$

which is called the *Clayton Lévy copula function.* Let

$$\phi(t) = \eta(-\log(|t|))^{-1/\theta} 1_{t\geq 0} - (1-\eta)(-\log(|t|))^{-1/\theta} 1_{t<0}, \quad \theta > 0, \eta \in (0,1).$$

Then ϕ is a general Archimedean generator with

$$\begin{aligned} \widetilde{\phi}(t) &= 2^{d-2}\{-\log(|t|)\}^{-1/\theta} \operatorname{sgn}(t), \\ \text{and} \quad \widetilde{\phi}^{-1}(t) &= e^{-|2^{2-d}t|^{-\theta}} \operatorname{sgn}(t). \end{aligned}$$

ϕ produces the two parameter general Archimedean Lévy d-copula

$$F(u_1, ..., u_d) = 2^{2-d} \left(\sum_{i=1}^{d} |u_i|^{-\theta}\right)^{-1/\theta} (\eta 1_{u_1 \dots u_d \geq 0} - (1-\eta) 1_{u_1 \dots u_d < 0}),$$

which is referred to as the *general Clayton Lévy copula.*

[9] The function (7.19) is a Lévy copula and $\widetilde{\phi}$ is used to preserve uniformity at the margins [cf. 77]. Moreover, Tankov [77, Remark 5.1] gives sufficient conditions to satisfy (7.19). These are comparable to the notion of complete monotonicity as defined in Appendix A.1.

[10] The function (7.20) really is a Lévy copula [cf. 77].

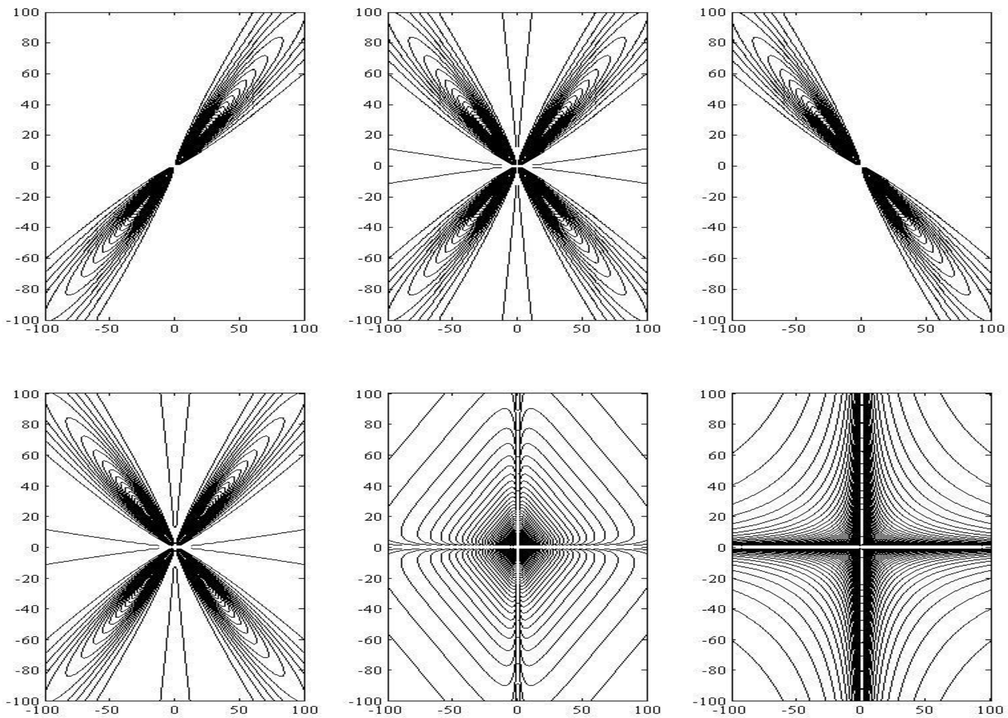

Figure 7.2: Contours of the Gauss copula density using parameters $\eta = 1, \theta = 6$ (upper left), $\eta = 0.5, \theta = 6$ (upper center), $\eta = 0, \theta = 6$ (upper right), $\eta = 0.5, \theta = 6$ (lower left), $\eta = 0.5, \theta = 0.5$ (lower center), $\eta = 0.5, \theta = 0.05$ (lower right).

For $d = 2$, the general Lévy Clayton copula is

$$F(u, v) = \left(|u|^{-\theta} + |v|^{-\theta}\right)^{-1/\theta} (\eta 1_{uv\geq 0} - (1-\eta) 1_{uv<0}), \tag{7.22}$$

and its density

$$f(u, v) = \operatorname{sgn}(u) \operatorname{sgn}(v) \left(|u|^{-\theta} + |v|^{-\theta}\right)^{-\frac{2\theta+1}{\theta}} |u|^{-\theta-1} |v|^{-\theta-1} (\eta 1_{uv\geq 0} - (1-\eta) 1_{uv<0})$$

exists on $\mathbb{R}^2$.

We analyze the role of the parameters on the basis of Figure 7.2, that shows the contours of the Lévy Clayton 2-copula density over various configurations.

- The jumps tend to have, for $\eta \gg 0$, the same directions and, for $\eta \ll 1$ opposite directions; η is responsible for the sign dependence of jumps.
- The higher θ the more associated are the absolute values of the jumps; θ determines the size dependence of of jumps.
- The contours are balanced; the dependence structure is symmetric.

Hence the Lévy Clayton copula models a wide range of ample dependence, albeit this is symmetric whatsoever.

7.3 Modular Lévy copula functions

In view of positive Lévy copulas, the idea to model the jump dependence structure of a $\mathbb{R}^d$-valued Lévy process separately in each of the 2^d corners is obvious. We analyze a modular Lévy copula design with emphasis on the probabilistic interpretation of its component parts [cf. 80].

Definition 7.24 (cf. [77], Theorem 5.3)**.** For $\alpha = \{\alpha_1, ..., \alpha_d\} \in \{-1, 1\}^d$ let $g^\alpha(u) : [0, \infty] \to [0, 1]$ be a nonnegative, increasing function satisfying

$$\sum_{\alpha \in \{-1,1\}^d \text{ with } \alpha_k = -1} g^\alpha(u) = 1 \text{ and } \sum_{\alpha \in \{-1,1\}^d \text{ with } \alpha_k = 1} g^\alpha(u) = 1 \tag{7.23}$$

for all $u \in [0, \infty]$ and all $k \in \{1, ..., d\}$. Moreover, let F^α be positive Lévy copulas that satisfy the following continuity property at infinity: for all $I \subset \{1, ..., d\}$, $(u_i)_{i \in I^c} \in [0, \infty]^{|I^c|}$ we have

$$\lim_{(u_i)_{i \in I} \to (\infty, ..., \infty)} F^\alpha(u_1, ..., u_d) = F^\alpha(v_1, ..., v_d),$$

where $v_i = u_i$ for $i \in I$ and $v_i = \infty$ otherwise. Then we call

$$F(u_1, ..., u_d) = F^\alpha(|u_1| g^\alpha(|u_1|), ..., |u_d| g^\alpha(|u_d|)) \prod_{i=1}^{d} \operatorname{sgn}(u_i)$$

a *modular Lévy copula*.[11]

We refer to the Lévy copulas $F^\alpha, \alpha \in \{-1, 1\}^d$ as the *modules* and to the set of functions $\{g^\alpha, \alpha \in \{-1, 1\}^d\}$ as the *joinder*.

Example 7.25. Consider the case $d = 3$. Let the module copulas F^α be given by (7.20) for all $\alpha \in \{-1, 1\}^3$ and

$$g^\alpha(u) = \begin{cases} 1, & \text{for } \alpha_1 = \alpha_2 = \alpha_3; \\ 0, & \text{otherwise.} \end{cases} \tag{7.24}$$

for all $u \in [0, \infty]$. This produces the modular Lévy copula F on $\mathbb{R}^3_\infty$ with

$$F(u_1, u_2, u_3) = \left(|u_1|^{-\theta} + |u_2|^{-\theta} + |u_3|^{-\theta}\right)^{-1/\theta} \left(1_{u_1>0, u_2>0, u_3>0} - 1_{u_1<0, u_2<0, u_3<0}\right),$$

a co-moving Clayton Lévy copula so to say.

[11] Tankov [77] shows that the condition (7.23) ensures uniformity of the margins.

Example 7.25 uses a piecewise constant joinder in combination with equal modules. Regarding parsimony of the copula model (and notational ease), we resort to equal modules $F^\alpha \equiv F^+$ in the following. As far as the joinder is concerned, we restrict ourselves to piecewise constant functions $g^\alpha(u) \equiv g^\alpha \in \mathbb{R}^+$. We formulated a constructive approach to g^α in [80].

Remark 7.26. In the context of Example 7.25 the joinder $g^\alpha = g^{\{\alpha_1,\alpha_2,\alpha_3\}}$ is a probability distribution, only if we condition on one of the values $\alpha_i, i \in \{1,2,3\}$. For instance, $(\alpha_2,\alpha_3) \mapsto g^{\{1,\alpha_2,\alpha_3\}}$ is a probability distribution function on $\{-1,1\}^2$.

Probabilistic interpretation The modules and the joinder of a general modular Lévy copula allow of probabilistic interpretations in answer to Note 7.26.

Example 7.27. Take a look at the probability of, say, opposing jumps $U_2 < 0, U_3 \geq 0$ given $U_1 = \xi$ under the conditional probability distribution $P_\xi[\cdot] := P[\cdot|U_1 = \xi]$, that is determined by (7.12). Identifying $g^{(u_1,u_2,u_3)}$ with $g^{(\alpha_1,\alpha_2,\alpha_3)}$, it follows from

$$\begin{aligned}
K_\xi(0,\infty) &= \frac{\partial F^+}{\partial u_1}(|\xi|g^{(\xi,1,1)},0,\infty)g^{(\xi,1,1)} + \frac{\partial F^+}{\partial u_1}(|\xi|g^{(\xi,1,-1)},0,\infty)g^{(\xi,1,-1)} \\
&+ \frac{\partial F^+}{\partial u_1}(|\xi|g^{(\xi,-1,1)},\infty,\infty)g^{(\xi,-1,1)} + \frac{\partial F^+}{\partial u_1}(|\xi|g^{(\xi,-1,-1)},\infty,\infty)g^{(\xi,-1,-1)} \\
K_\xi(0,0) &= \frac{\partial F^+}{\partial u_1}(|\xi|g^{(\xi,1,1)},0,\infty)g^{(\xi,1,1)} + \frac{\partial F^+}{\partial u_1}(|\xi|g^{(\xi,1,-1)},0,\infty)g^{(\xi,1,-1)} \\
&+ \frac{\partial F^+}{\partial u_1}(|\xi|g^{(\xi,-1,1)},\infty,0)g^{(\xi,-1,1)} + \frac{\partial F^+}{\partial u_1}(|\xi|g^{(\xi,-1,-1)},\infty,\infty)g^{(\xi,-1,-1)} \\
\lim_{c\to\infty} K_\xi(-c,a) &= 0 \quad \forall a \in \mathbb{R}_\infty
\end{aligned}$$

that

$$\begin{aligned}
P_\xi[U_2 < 0, U_3 \geq 0] &= K_\xi(0,\infty) - \lim_{c\to\infty} K_\xi(-c,\infty) - K_\xi(0,0) + \lim_{c\to\infty} K_\xi(-c,0) \\
&= \frac{\partial F^+}{\partial u_1}(|\xi|g^{(\xi,-1,1)},\infty,\infty)g^{(\xi,-1,1)}.
\end{aligned}$$

Further we have

$$\begin{aligned}
\frac{\partial F^+}{\partial u_1}(|\xi|g^{(\xi,-1,1)},\infty,\infty)g^{(\xi,-1,1)} &= \frac{\partial F_1^+}{\partial u_1}(|\xi|g^{(\xi,-1,1)})g^{(\xi,-1,1)} \\
&= g^{(\xi,-1,1)},
\end{aligned}$$

because of the uniformity of F^+ at the margins (as part of Definition 7.11). The argumentation used here does not depend on a specific sign vector, and so the result can be generalized to

$$P_\xi[\mathrm{sgn}(U_2) = \alpha_2, \mathrm{sgn}(U_3) = \alpha_3] = g^{(\xi,\alpha_2,\alpha_3)}.$$

In the multidimensional case, (7.23) allows likewise of the interpretation of g^α as a conditional distribution function given $\alpha_k \in \{-1, 1\}$ for one $k \in \{1, ..., d\}$:

Theorem 7.28. *Let F be a modular Lévy copula on $\mathbb{R}^d_\infty$ with joinder $g^\alpha \in \mathbb{R}^+$ and modules $F^\alpha = F^+$. Further let $U_1 = \xi$ be a given realization. Then it holds*

$$P_\xi[\mathrm{sgn}(U_2) = \alpha_2, ..., \mathrm{sgn}(U_d) = \alpha_d] = g^{(\xi,\alpha_2,...,\alpha_d)}. \tag{7.25}$$

Proof. The proof is essentially the same as in the three-dimensional case. Suppose $\alpha_i = 1, i = 2, ..., d$, the other cases being derived analogously. It holds

$$\begin{aligned} P_\xi[\mathrm{sgn}(U_2) = \alpha_2, ..., \mathrm{sgn}(U_d) = \alpha_d] &= \sum_{(u_2,...,u_d)\in\{0,\infty\}^{d-1}} (-1)^{N(u_2,...,u_d)} K_\xi(u_2, ..., u_d) \\ &= \frac{\partial F^+}{\partial u_1}(|\xi| g^{(\xi,\alpha_2,...,\alpha_d)}, \infty, ..., \infty) g^{(\xi,\alpha_2,...,\alpha_d)} \end{aligned}$$

by cancellation of terms, where we denote $N(u_2, ..., x_d) = \sharp\{k : u_k = 0\}$. Then uniformity of F^+ at the margins gives the desired result. □

Theorem 7.28 distinguishes the piecewise constant joinder g^α of a modular Lévy copula as the joint jump sign probabilities. This suggests to analyze the conditional absolute jumps size distribution of $(U_2, ..., U_d)$ given their signs $(\alpha_2, ..., \alpha_d)$ and U_1.

Example 7.29. Assume $(\alpha_2, \alpha_3) = (-1, 1)$. Then

$$P_\xi[U_2 \in \mathcal{I}^c(u_2), U_3 \in \mathcal{I}^c(u_3)|\alpha_2, \alpha_3] = \frac{P_\xi([0 \geq U_2 \geq u_2, 0 \leq U_3 \leq u_3]}{g^{(\xi,-1,1)}},$$

where

$$\mathcal{I}^c(x) = \begin{cases} [0, x), & x > 0; \\ (-x, 0], & x < 0. \end{cases}$$

By (7.12), it holds

$$\begin{aligned} P_\xi[0 \geq U_2 \geq u_2, 0 \leq U_3 \leq u_3] &= K_\xi(0, u_3) - K_\xi(u_2, u_3) - K_\xi(0, 0) + K_\xi(u_2, 0) \\ &= \frac{\partial F^+}{\partial u_1}(|\xi| g^{(\xi,-1,1)}, |u_2| g^{(\xi,-1,1)}, |u_3| g^{(\xi,-1,1)}) g^{(\xi,-1,1)}. \end{aligned}$$

Then the absolute jump size probability conditional on the jump signs is

$$P_\xi[0 \geq U_2 \geq u_2, 0 \leq U_3 \leq u_3|\alpha_2, \alpha_3] = \frac{\partial F^+}{\partial u_1}(|\xi| g^{(\xi,-1,1)}, |u_2| g^{(\xi,-1,1)}, |u_3| g^{(\xi,-1,1)}).$$

This result is not unexpected due to the modular design.

The arguments used in the three-dimensional case apply just as well for arbitrary dimensions:

Theorem 7.30. *Let F be a modular Lévy copula on $\mathbb{R}^d$ with joinder $g^\alpha \in \mathbb{R}^+$ and modules $F^\alpha = F^+$. Further let $\xi = U_1$ be a given realization. Then it holds*

(7.26)

$$P_\xi[U_i \in \mathcal{I}^c(u_i), i = 2, ..., d | \alpha_2, ..., \alpha_d] = \frac{\partial F^+}{\partial u_1}(|\xi| g^{\xi,\alpha_2,...,\alpha_d}, |u_2| g^{(\xi,\alpha_2,...,\alpha_d)}, ..., |u_d| g^{(\xi,\alpha_2,...,\alpha_d)}).$$

Proof. The proof is essentially the same as in the three-dimensional case. Suppose $\alpha_i = 1, i = 2, ..., d$, the other cases being derived analogously. By formula (7.12), we have

$$\begin{aligned} P_\xi[U_i \in \mathcal{I}^c(u_i), i = 2, ..., d | \alpha_2, ..., \alpha_d] &= \sum_{v_i \in \{0, u_i\}} (-1)^{N(v_2,...,v_d)} K_\xi(v_2, ..., v_d) \\ &= \frac{\partial F^+}{\partial u_1}(|\xi| g^{(\xi,\alpha_2,...,\alpha_d)}, u_2, ..., u_d) g^{(\xi,\alpha_2,...,\alpha_d)} \end{aligned}$$

due to cancellation of terms. Together with formula (7.25), we conclude

$$P_\xi[U_i \in \mathcal{I}^c(u_i), i = 2, ..., d | \alpha_2, ..., \alpha_d] = \frac{\partial F^+}{\partial u_1}(|\xi| g^{(\xi,\alpha_2,...,\alpha_d)}, u_2, ..., u_d)$$

for the stand-alone probability of the conditional size vector. □

Theorem 7.30 shows that the distribution of the transformed absolute values of jumps $(|U_2| g^\alpha, ..., |U_d| g^\alpha)$ conditional on the realized jump $U_1 = \xi$ and jump signs α is that associated to the module F^+ in the sense of (7.13) given $|\xi| g^\alpha$. This reinforces the idea of a modular design in a way that jump sizes are close upon separated from jump signs.

Modular Clayton Lévy copula The modular Clayton Lévy copula is obtained from using Clayton modules:

Example 7.31. For all $\alpha \in \{-1, 1\}^3$, let $F^\alpha = F^+$ be the positive Clayton Lévy 3-copula given by (7.21),

$$F^+(u_1, u_2, u_3) = \left(u_1^{-\theta} + u_2^{-\theta} + u_3^{-\theta}\right)^{-1/\theta},$$

and $g^\alpha(u_1, u_2, u_3) \equiv g^\alpha$ the constant joinder defined by

$$g^\alpha = \begin{cases} \frac{27}{36}, & \alpha_1 = \alpha_2, \alpha_1 \neq \alpha_3; \\ \frac{3}{36}, & \text{otherwise}. \end{cases} \tag{7.27}$$

According to Definition 7.24, F^+ and g^α give the following modular Lévy copula F on $\mathbb{R}^3_\infty$:

$$F(u_1, u_2, u_3) = \begin{cases} \left(\sum_{i=1}^3 (|u_i| \frac{27}{36})^{-\theta}\right)^{-1/\theta} \prod_{i=1}^3 \operatorname{sgn}(u_i), & \alpha_1 = \alpha_2, \alpha_1 \neq \alpha_3; \\ \left(\sum_{i=1}^3 (|u_i| \frac{3}{36})^{-\theta}\right)^{-1/\theta} \prod_{i=1}^3 \operatorname{sgn}(u_i), & \text{otherwise}. \end{cases} \tag{7.28}$$

As to Theorem 7.28, the conditional probability of sign vector $\{\alpha_2, \alpha_3\} \in \{-1, 1\}^2$ given $U_1 = \xi$ is

$$P_\xi[\operatorname{sgn}(U_i) = \alpha_i, i = 2, 3] = \begin{cases} \frac{27}{36}, & \alpha_1 = \alpha_2, \alpha_1 \neq \alpha_3; \\ \frac{3}{36}, & \text{otherwise}. \end{cases} \tag{7.29}$$

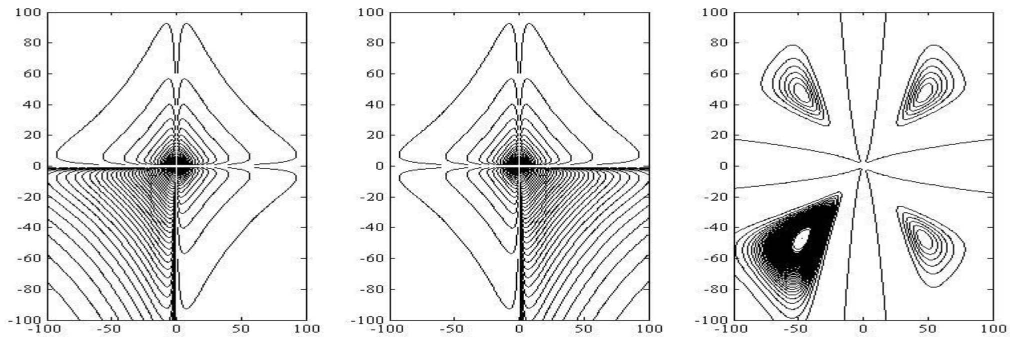

Figure 7.3: Contours of the (conditional) modular Lévy copula density in (1,2)-marginal cross section at $u_3 = 50$ using $\theta = 0.5$ (left), (1,3)-marginal cross section at $u_2 = 50$ using $\theta = 0.5$ (center), and (1,2)-marginal cross section at $u_3 = 50$ using $\theta = 6$ (right).

Regarding Theorem 7.30, the conditional probability distribution of jump sizes is determined by

$$P_\xi[U_i \in \mathcal{I}^c(u_i), i = 2, 3|\alpha_2, \alpha_3] = \begin{cases} \left(\sum_{i=1}^3 (|u_i|\frac{27}{36})^{-\theta}\right)^{-\frac{1+\theta}{\theta}} (|u_1|\frac{27}{36})^{-1-\theta}, & \alpha_1 = \alpha_2, \alpha_1 \neq \alpha_3; \\ \left(\sum_{i=1}^3 (|u_i|\frac{3}{36})^{-\theta}\right)^{-\frac{1+\theta}{\theta}} (|u_1|\frac{3}{36})^{-1-\theta}, & \text{otherwise.} \end{cases}$$

We have plotted the contours for the bivariate marginal densities of copula (7.28) over different module parameters in Figure 7.3. Each plot shows a 2-dimensional cross section of the trivariate copula density from a different perspective but at the same level $u_i = 50$ of the hidden ith variable.

- Given $u_3 = 50$, the (1,2)-margin has greater mass in the 3rd quadrant than in the others; a co-movement of the 1st and 2nd component in negative direction is more likely than jumps in other directions.
- Given $u_2 = 50$, the (1,3)-margin (or the (2,3)-margin, which we omit in good faith) has greater mass in the 4th quadrant than in the others; negative jumps in the 3rd component are most likely to go together with positive jumps in the 1rd (and 2nd) component.
- Given $u_3 = 50$, the level curves are the more clustered around 50 the higher the θ; the probability masses into the absolute value of the hidden variable, if θ increases.

These findings are feasible with respect to Example 7.31 and the aforementioned notices of the Lévy Clayton copula. In conclusion, the modularly designed Lévy copula generalizes the Archimedean construction insofar as it allows for a more complex yet separate modelling of jump sign and jump size dependence.

7.4 Canonical Lévy copula functions

Using the family of conditional probability distributions associated to a Lévy copula, we develop a new Lévy copula model and discuss some relevant properties [cf. 79].

Theorem 7.13 related conditional probability measures and Lévy copulas. This suggests an alternative modelling approach: instead of designing the generic Lévy copula F, from which K_ξ can be derived, we propose to define an implicit dependence structure by modelling the conditional probability distribution K_ξ in the first place. From there, the multivariate Lévy measure obtains via the interrelations (7.10) and (7.11). Here the crux is to keep the jump dependence structure separate from the marginal process evolution. This is certainly not fulfilled per se. The following result establishes a sufficient (and necessary) design of a qualified conditional measure:

Theorem 7.32. *Let $\nu_i, i = 1, ..., d$ be marginal Lévy measures on $\mathbb{R}$ with corresponding tail integrals U_i and $f : (u_1, ..., u_d) \mapsto (U_1^{-1}(u_1), ..., U_d^{-1}(u_d))$. Further let $K(\xi, \cdot)$ be a conditional measure on $\mathbb{R}^{d-1}$ such that $\nu = \mu(f)$ in the sense of* (7.10) *and* (7.11) *is a Lévy measure on $\mathbb{R}^d$ with margins ν_i. Then there exist Lévy copulas $F^i : \mathbb{R}^2_\infty \to \mathbb{R}_\infty, i = 2, ..., d$ and a family, indexed by $\xi \in \mathbb{R}$, of ordinary copula functions $C_\xi : [0,1]^{d-1} \to [0,1]$ such that*

$$K_\xi(u_2, ..., u_d) = C_\xi\left(G^2_\xi(u_2), ..., G^d_\xi(u_d)\right), \tag{7.30}$$

where

$$G^i_\xi(u) := \operatorname{sgn}(\xi)\frac{\partial}{\partial \xi} V_{F^i}((0 \wedge \xi, 0 \vee \xi] \times (-\infty, u]). \tag{7.31}$$

Conversely, if $F^i : \mathbb{R}^2_\infty \to \mathbb{R}_\infty, i = 2, ..., d$ Lévy copulas, $C_\xi : [0,1]^{d-1} \to [0,1], \xi \in \mathbb{R}$ ordinary copula functions and conditional measure K on $\mathbb{R}^{d-1}$ defined as in (7.30)*, then $\nu = \mu(f)$ in the sense of* (7.10) *and* (7.11) *is a Lévy measure on $\mathbb{R}^d$ with margins ν_i.*

Proof. First part. For all $\xi \in \mathbb{R} \setminus N$, there exists by Sklar's theorem a $(d-1)$-dimensional ordinary copula C_ξ and univariate marginal distribution functions $G^i_\xi, i = 2, ..., d$ such that

$$K_\xi(u_2, ..., u_d) = C_\xi(G^2_\xi(u_2), ..., G^d_\xi(u_d)).$$

The G^i_ξ's are distribution functions by (7.12) and the arguments used therefore. Our goal is to represent the G^i_ξ's by (7.31). For this purpose, let $i \in \{2, ..., d\}$ and consider the bivariate tail integral $U^{1,i}_\nu$. By Theorem 7.12, there exists a Lévy copula F^i on $\mathbb{R}^2_\infty$ so that

$$U^{1,i}_\nu(x_1, x_i) = F^i(U_1(x_1), U_i(x_i)). \tag{7.32}$$

It follows from the construction of the Lévy measure ν that, for $x_1 < 0, x_i \geq 0$ say,[12]

$$\begin{aligned} U_\nu^{1,i}(x_1, x_i) &= -\mu(\{u \in \mathbb{R}^d : u_1 \in (U_1(x_1), 0], u_i \in (0, U_i(x_i)]\}) \\ &= -\int_{U_1(x_1)}^0 \int_0^{U_i(x_i)} \int_{\mathbb{R}^{d-2}} K(\xi, dx_2, ...dx_d) d\xi \\ &= \int_0^{U_1(x_1)} K_\xi(\infty, .., U_i(x_i).., \infty) - K_\xi(\infty, .., 0, .., \infty) d\xi. \end{aligned}$$

By (7.12) and uniformity at the margins of an ordinary copula, the bivariate tail integral can be written as

$$U_\nu^{1,i}(x_1, x_i) = \int_0^{U_1(x_1)} G_\xi^i(U_i(x_i)) - G_\xi^i(0) d\xi. \tag{7.33}$$

Equating (7.32) and (7.33) leads to

$$\int_0^{U_1(x_1)} G_\xi^i(U_i(x_i)) - G_\xi^i(0) d\xi \quad = \quad F^i(U_1(x_1), U_i(x_i)),$$

where we express the right hand side in terms of volume functions as follows

$$\begin{aligned} F^i(U_1(x_1), U_i(x_i)) &= -V_{F^i}((U_1(x_1), 0] \times (-\infty, U_i(x_i)]) \\ &\quad +V_{F^i}((U_1(x_1), 0] \times (-\infty, 0]). \end{aligned}$$

Differentiation then yields $G_\xi^i(u) := -\frac{\partial}{\partial \xi} V_{F^i}((\xi, 0] \times (-\infty, u])$. The general case $x_1, x_i \in \mathbb{R}$ can be derived analogously.

Second part. In order to show that $\nu = \mu(f)$ has margins $\nu_i, i = 1, ..., d$, it suffices to consider its marginal tail integrals $U_\nu^i, i = 1, ..., d$. The goal is to prove equality between the input tail integrals U_i and the implicit tail integrals U_ν^i. Similar to the first part, it holds, for $x_i < 0$ say,

$$U_\nu^i(x_i) = -\mu(\{u \in \mathbb{R}^d : u_i \in (U_i(x_i), 0]\}).$$

Since $K(\xi, \cdot)$ is a probability measure on $\mathbb{R}^{d-1}$, the first tail integral turns out to be

$$\begin{aligned} U_\nu^1(x_1) &= -\mu(\{u \in \mathbb{R}^d : u_1 \in (U_1(x_1), 0)]\}) \\ &= \int_0^{U_1(x_1)} \int_{\mathbb{R}^{d-1}} K(\xi, dx_2, ...dx_d) d\xi \\ &= U_1(x_1). \end{aligned}$$

We further have, for the i-th tail, $i \in \{2, ..., d\}$,

$$\begin{aligned} U_\nu^i(x_i) &= -\mu(\{u \in \mathbb{R}^d : u_i \in (U_i(x_i), 0]\}) \\ &= -\int_{\mathbb{R}} \int_{U_i(x_i)}^0 \int_{\mathbb{R}^{d-2}} K(\xi, dx_2, ..., dx_d) d\xi. \\ &= \int_{\mathbb{R}} \left(K_\xi(\infty, .., U_i(x_i), .., \infty) - K_\xi(\infty, .., 0, .., \infty)\right) d\xi, \end{aligned}$$

[12]The first identity is also used for the proof of Theorem 7.12 [cf. 45, Theorem 3.6].

by means of the same arguments as used in the previous case. By (7.12) and uniformity at the margins of an ordinary copula, it follows that

$$U_\nu^i(x_i) \;=\; \int_{\mathbb{R}} G_\xi^i(U_i(x_i)) - G_\xi^i(0)d\xi.$$

Since $G_\xi^i(y_i)$ is the derivative with respect to the integration variable, the desired result follows from uniformity at the margins of a Lévy copula:

$$\begin{aligned}\int_{\mathbb{R}}(G_\xi^i(U_i(x_i)) - G_\xi^i(0))d\xi &= -V_{F^i}(\mathbb{R}_\infty \times (U_i(x_i),0]) \\ &= F^i(\infty, U_i(x_i)) - F^i(-\infty, U_i(x_i)) \\ &= F_2^i(U_i(x_i)) \\ &= U_i(x_i).\end{aligned}$$

The general case $x_1, x_i \in \mathbb{R}$ can be derived analogously. It is not yet concluded that ν really is a Lévy measure satisfying the integrability condition $\int_{\mathbb{R}^d}(|x|^2 \wedge 1)\nu(dx) < \infty$. But this is automatically [cf. 77, Proof of Theorem 4.8] fulfilled, if its one-dimensional margins are Lévy measures:

$$\int_{\mathbb{R}^d}(|x|^2 \wedge 1)\nu(dx) \leq \int_{\mathbb{R}^d}\sum_{i=1}^d (x_i^2 \wedge 1)\nu(dx) \leq \sum_{i=1}^d \int_{\mathbb{R}}(x_i^2 \wedge 1)\nu_i(dx_i) < \infty.$$

□

Remark 7.33. It is worth noting that Theorem 7.32 imposes a dependence structure between the marginal Lévy measures without mention of the Lévy copula. There still exists a corresponding Lévy copula by Theorem 7.12, but it is only given implicitly through the designed measure ν.

In the sense of Remark 7.33, we call the implicit Lévy copula the *canonical Lévy copula* and the implicit pattern the *canonization*.[13] Moreover, Theorem 7.32 distinguishes the first margin, which we hereafter call the *canon*.

The proof (first part) of Theorem 7.32 also provides the groundwork for the following result on the side:

Corollary 7.34. *Let $C_\xi, \xi \in \mathbb{R}$ be arbitrary ordinary copulas, $F^i, i = 2, ..., d$ Lévy copulas and ν as in the previous theorem. Then the marginal Lévy copulas satisfy $F^{1,i} = F^i$, $\forall i \in \{2, ..., d\}$.*

Corollary 7.34 indicates that the building block Lévy copulas $F^i, i = 1, ..., d$ coincide with the bivariate marginal Lévy copulas $F^{1,i}, i = 2, ..., d$, called the *canon dependence structures*. The following lemma targets the bivariate Lévy copulas $F^{i,j}, i, j = 2, ..., d$, called the *non-canon dependence structures*, in a particular case. First reasons for our naming is given thereby:

[13]Luciano and Schoutens [51] and Tankov [77] give two alternative approaches to bringing together ordinary and Lévy copula models.

Lemma 7.35. *Let $C_\xi, \xi \in \mathbb{R}$ be arbitrary ordinary copulas, $F^i = F_\|, i = 2, ..., d$ and ν as in the previous theorem. Then the marginal Lévy copulas satisfy $F^{i,j} = F_\|,\ \forall i, j \in \{1, ..., d\}$.*

Proof. Let $F^i = F_\| = \min\{|x_1|, |x_i|\} 1_{S^+}(x_1, x_i) \operatorname{sgn}(x_1) \operatorname{sgn}(x_i)$. Then

$$G^i_\xi(x_i) = 1_{x_i \geq \xi \geq 0} + 1_{\xi < 0} - 1_{0 > \xi \geq x_i}. \tag{7.34}$$

It suffices again to consider the tail integrals. We want to show that the implicit bivariate tail integral $U^{i,j}_\nu$ can be represented as the univariate tails U_i, U_j, that are coupled by the complete dependence copula $F_\|$. For this purpose, we renew the argumentation from the proof of the second part of Theorem 7.32 and obtain

$$\begin{aligned} U^{i,j}_\nu(x_i, x_j) &= \int_{\mathbb{R}} [K_\xi(\infty, U_i(x_i), U_j(x_j), \infty) - K_\xi(\infty, ., U_i(x_i), 0, ., \infty) \\ &\quad - K_\xi(\infty, ., 0, U_j(x_j), ., \infty) + K_\xi(\infty, ., 0, 0, ., \infty)] d\xi. \end{aligned}$$

Assuming $\xi < 0$, we have

$$\begin{aligned} K_\xi(\infty, .., U_i(x_i), .., U_j(x_j), .., \infty) &= C^{i,j}_\xi(\{1_{U_k(x_k) \geq \xi \geq 0} + 1_{\xi < 0} - 1_{0 > \xi \geq U_k(x_k)}\}_{k=i,j}) \\ &= (1_{\xi<0} - 1_{0>\xi \geq U_i(x_i)}) \cdot (1_{\xi<0} - 1_{0>\xi \geq U_j(x_j)}) \\ &= 1_{\xi \leq U_i(x_i), \xi \leq U_j(x_j)} \\ K_\xi(\infty, .., U_i(x_i), .., 0, .., \infty) &= C^{i,j}_\xi(1_{U_i(x_i) \geq \xi \geq 0} + 1_{\xi<0} - 1_{0>\xi\geq U_i(x_i)}, 1_{\xi<0}) \\ &= (1_{\xi<0} - 1_{0>\xi \geq U_i(x_i)}) \cdot 1_{\xi<0} \\ &= 1_{\xi \leq U_i(x_i)} \\ K_\xi(\infty, .., 0, .., U_j(x_j), .., \infty) &= C^{i,j}_\xi(1_{\xi<0}, 1_{U_i(x_i) \geq \xi \geq 0} + 1_{\xi<0} - 1_{0>\xi\geq U_i(x_i)}) \\ &= 1_{\xi<0} \cdot (1_{\xi<0} - 1_{0>\xi \geq U_j(x_j)}) \\ &= 1_{\xi \leq U_j(x_j)} \\ K_\xi(\infty, .., 0, .., 0, .., \infty) &= C^{i,j}_\xi(1_{\xi<0}, 1_{\xi<0}) \\ &= 1, \end{aligned}$$

by (7.30), (7.34) and the properties of ordinary copula functions. This gives us immediately the characteristic integrand

$$1_{\xi \leq U_i(x_i), \xi \leq U_j(x_j)} - 1_{\xi \leq U_i(x_i)} - 1_{\xi \leq U_j(x_j)} + 1 \quad = \quad 1_{\xi \geq U_i(x_i), \xi \geq U_j(x_j)}.$$

In a similar way, we obtain $1_{\xi \leq U_i(x_i), \xi \leq U_j(x_j)}$ for $\xi > 0$. Then

$$\begin{aligned} U^{i,j}_\nu(x_i, x_j) &= \int_{-\infty}^0 1_{\xi \geq U_i(x_i), \xi \geq U_j(x_j)} d\xi + \int_0^\infty 1_{\xi \leq U_i(x_i), \xi \leq U_j(x_j)} d\xi \\ &= \min\{|U_i(x_i)|, |U_j(x_j)|\} 1_{S^+}(x_i, x_j). \end{aligned}$$

□

Lemma 7.35 demonstrates that canonization carries complete dependence to any bivariate margin, no matter what the association between the non-canon variables may be. This is a very strong result and enforces the interpretation of the projected variable as the system's rule, the canon.

Canonical Clayton Lévy copula The canonical Clayton Lévy copula is the combination of the Clayton ordinary and the general Clayton Lévy copula function.

Example 7.36. For $\xi \in \mathbb{R}$, let C_ξ be the ordinary Clayton 2-copula given by

$$C_\xi(u_1, u_2; \kappa) = (u_1^{-\kappa} + u_2^{-\kappa} - 1)^{-1/\kappa}, \tag{7.35}$$

and $F^i, i = 2, 3$ be general Clayton Lévy 2-copula functions given by

$$F^i(u_1, u_i) = \left(|u_1|^{-\theta} + |u_i|^{-\theta}\right)^{-1/\theta} (\eta 1_{u_1 u_i \geq 0} - (1-\eta) 1_{u_1 u_i < 0}). \tag{7.36}$$

As to Tankov [cf. 77, p.180], a straight forward computation leads to

$$\begin{aligned} G_\xi^i(u_i) &:= \operatorname{sgn}(\xi) \frac{\partial}{\partial \xi} V_F((0 \wedge \xi, 0 \vee \xi] \times (-\infty, u_i]) \\ &= \left\{ (1-\eta) + \left(1 + \frac{|\xi|}{|u_i|}^{\theta}\right)^{-1-1/\theta} (\eta - 1_{u_i < 0}) \right\} 1_{\xi \geq 0} \\ &+ \left\{ \eta + \left(1 + \frac{|\xi|}{|u_i|}^{\theta}\right)^{-1-1/\theta} (1_{u_i \geq 0} - \eta) \right\} 1_{\xi < 0}. \end{aligned}$$

Hence, the conditional cdf K_ξ associate to the canonical Clayton Lévy copula is given in closed form by

$$K_\xi(u_2, u_3) = (G_\xi^2(u_2)^{-\kappa} + G_\xi^3(u_3)^{-\kappa} - 1)^{-1/\kappa}, \tag{7.37}$$

where G_ξ^i as in (7.37).

Figure 7.4 illustrates the implicit Lévy copula associate to the conditional cdf (7.37). It shows the level curves of its (2,3)-marginal density (the other margins being clear from Proposition 7.34 and our findings on the general Clayton Lévy copula) over various parameter choices. Each plot represents a bivariate cross section of the trivariate density as viewed from the $(2, 3)$-perspective at the angle $u_1 = 50$. We observe that

- Given $u_1 = 50$, the (2,3)-margin has greater mass in the 1st quadrant than in the others; a co-movement of the 2nd and 3rd component in positive direction is more likely than jumps in other directions.

- Given $u_1 = 50$, the level curves are the more clustered around 50 the higher the θ; the probability masses into the absolute value of the hidden variable, if θ increases.

- If κ increases, then the contours become arrow shaped; the dependence structure is asymmetric.

As to Lemma 7.35, these observations enforce the typical jump sign and jump size dependence structures of the Lévy Clayton copula function. The asymmetric dependence structure of the ordinary Clayton copula [cf. Section A.1 and 31] are obvious, too. Altogether, the canonically designed Lévy copula offers a radically new way to non-standard dependence patterns and models asymmetries as well as driving jump forces.

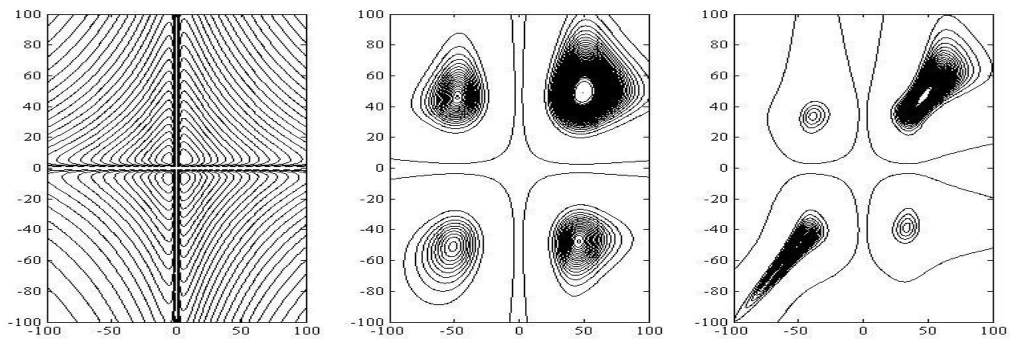

Figure 7.4: Contours of the (conditional) canonical Lévy copula density in (2,3)-marginal cross section at $u_1 = 50$ using $\eta = 0.75, \theta = 0.5, \kappa = 0.5$ (left), (2,3)-marginal cross section at $u_1 = 50$ using $\eta = 0.75, \theta = 6, \kappa = 0.5$ (center), and (2,3)-marginal cross section at $u_1 = 50$ using $\eta = 0.75, \theta = 6, \kappa = 6$ (right).

7.5 Summary

In this chapter we isolated the jump dependence structure of multidimensional Lévy processes using Lévy copula functions. With only a few parametric models available, we developed new Lévy copulas and gave innovative probabilistic interpretations for existent ones.

Building on a short and precise guide to Lévy processes, we decoupled the dependence structure of a multidimensional jump processes from its marginal dynamics using Lévy copula functions (much in the same way as in the case of ordinary copulas). We gave basic properties and made ourselves familiar with the fundamental and the Archimedean Lévy copulas.

Then we came to appreciate a clear distinction of jump sign and jump size dependence by instance of the general Lévy Clayton copula. We elaborated on this idea in the context of modularly designed Lévy copulas, that universalize the separate modelling of jump sign and jump size dependence. Our contribution here is the interpretation of the model's two component parts as jump sign and jump size probabilities. We recognised this interpretation by the contour lines of the density of the modular Lévy copula.

Then we developed the new canonical design of a Lévy copula. Our main contribution here was a representation result, that gives sufficient and necessary model conditions for the canonical Lévy copula function. Continuing, we gave proof to the built-in projection of a driving jump force component. This model feature was then reinforced by the graphical illustration of the contours of the canonical Lévy copula density.

Chapter 8

Simulation of Lévy processes

In this chapter we simulate Lévy processes whose dependence structure is given by a Lévy copula function. Here series representations prove very useful in sampling from copula based models. We reveal, in particular, that the modular and the canonical design are very well suited to this approach.

Section 8.1 expands a general Lévy process into a series of random variables. This suggests the simulation approach. Full detail is given to the representation of a $\mathbb{R}^d$-valued symmetric α-stable process. Then multivariate Lévy processes are represented in the specific case of the Clayton, the modular and the canonical Lévy copula. Section 8.2 develops the explicit sampling procedure for general Lévy processes. This is applied to the $\mathbb{R}^d$-valued symmetric α-stable process first. Then the trajectories from coupled Lévy processes are generated. Here the modular and the canonical construction of Lévy copulas prepare efficient sampling approaches.

The use of series representations in conjunction with Lévy copulas originates mostly in Cont and Tankov [20] and Tankov [77]. The implementation of modularly and canonically coupled Lévy processes presents merely new material.

8.1 Series representations

For general Lévy processes, infinite variation processes in particular, simulation is a delicate problem, because jumps may arrive infinitely often in every time interval. This excludes a simple adding of jumps to the Brownian component as suggested by (7.5). Here series representations provide a natural sampling approach.

A general representation theorem The original representation theorem is due to Rosinski [67]:

Theorem 8.1 ([20], Theorem 6.2)**.** *Let $\{V_k : k \geq 1\}$ be i.i.d. sequence of random variables on a measurable space S. Assume that $\{V_k : k \geq 1\}$ is independent of the sequence $\{\Gamma_k : k \geq 1\}$ of jumping times of a standard Poisson process [cf. 20, Section 2.5.3]. Let*

$\{U_k : k \geq 1\}$ be a sequence of independent random variables, uniformly distributed on $[0,1]$ and independent from $\{V_k : k \geq 1\}$ and $\{\Gamma_k : k \geq 1\}$. Let

$$H : (0,\infty) \times S \to \mathbb{R}^d$$

be a measurable function. We define measures on $\mathbb{R}^d$ by

$$\begin{aligned} \sigma(r,B) &= P[H(r,V_k) \in B], \quad r > 0, B \in \mathcal{B}(\mathbb{R}^d) \\ \nu(B) &= \int_0^\infty \sigma(r,B)dr \end{aligned}$$

and denote

$$A(s) = \int_0^s \int_{|x|\leq 1} x\sigma(r,dx)dr, \quad s > 0.$$

(1) *If ν is a Lévy measure on $\mathbb{R}^d$, that is,*

$$\nu(\{\mathbf{0}\}) = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} (|x|^2 \wedge 1)\nu(dx) < \infty$$

and the limit $\gamma = \lim_{s\to\infty} A(s)$ exists in $\mathbb{R}^d$ then the series

$$\sum_{k=1}^\infty H(\Gamma_k, V_k)\mathbf{1}_{U_k \leq t}$$

converges almost surely and uniformly on $t \in [0,1]$ to a Lévy process with characteristic triplet $(\gamma, 0, \nu)$, that is, with characteristic exponent

$$\psi(z) = iz.\gamma + \int_{\mathbb{R}^d} (e^{iz.x} - 1 - iz.x\mathbf{1}_{|x|\leq 1})\nu(dx).$$

(2) *If ν is a Lévy measure on $\mathbb{R}^d$ and for each $v \in S$ the function*

$$r \to |H(r,v)| \text{ is nondecreasing}$$

then

$$\sum_{k=1}^\infty H(\Gamma_k, V_k)\mathbf{1}_{U_k \leq t} - t[A(k) - A(k-1)]$$

converges almost surely and uniformly on $t \in [0,1]$ to a Lévy process with characteristic triplet $(0,0,\nu)$.

Theorem 8.1 shows us how to expand a multivariate Lévy process into an infinite series of random variables.[1] A closed form series representation is available for an $\mathbb{R}$-valued α-stable process. This requires to reformulate the Lévy density of an α-stable process:

[1] Cont and Tankov [cf. 20, Remark 6.6] show that the truncated series

$$X_t^\tau = \sum_{k:\Gamma_k \leq \tau} H(\Gamma_k, V_k)\mathbf{1}_{U_k \leq t} - t[A(k) - A(k-1)]$$

is a compound Poisson process with characteristic triplet $(0,0,\nu_\tau)$, where $\nu_\tau(A) = \int_0^\tau \sigma(r,A)dr$. An alternative approximation of an infinite activity Lévy processes by a compound Poisson process is discussed there [cf. 20, Section 6.3], too.

Lemma 8.2 (cf. [20], Example 6.15)**.** *Consider a symmetric random variable V such that $E[|V|^\alpha] < \infty$. Denote the distribution of V by P^V. Then we can write for any measurable set B*

$$\int_0^\infty r^{\frac{1}{\alpha}} P^V(r^{\frac{1}{\alpha}} B) dr = \frac{\alpha}{2} E[|V|^\alpha] \int_B \frac{dx}{|x|^{1+\alpha}},$$

where the set $r^{\frac{1}{\alpha}} B$ contains all points such that $r^{-\frac{1}{\alpha}} x \in B$.

Using Lemma 8.2, the series representation of an α-stable Lévy process follows direct from Theorem 8.1:

Example 8.3. Let ν be a symmetric α-stable process on $\mathbb{R}$ with density (7.8), choosing $\lambda_+ = \lambda_- = 1$, and tail integral U. Define $\sigma(r, B) = \frac{1}{\alpha} r^{\frac{1}{\alpha}} P^V(r^{\frac{1}{\alpha}} B)$ and V as a random variable taking values 1 or -1 with equal probability. Then Lemma 8.2 applies:

$$\begin{aligned} \int_0^\infty \sigma(r, \mathcal{I}(x)) &= \frac{1}{2} \int_{\mathcal{I}(x)} \nu(dx) \\ &= \frac{1}{2} U(x). \end{aligned}$$

On the other hand, for $H(r, v) = U^{(-1)}(r) v$, we have

$$\begin{aligned} \int_0^\infty \sigma(r, \mathcal{I}(x)) &= \int_0^\infty P[U^{(-1)}(r) V_i \in \mathcal{I}(x)] dr \\ &= \int_0^{U(x)} P[\operatorname{sgn}(V_k) = \operatorname{sgn}(x)] dr \\ &= \frac{1}{2} U(x) \end{aligned}$$

by Theorem 8.1. Because P^V is symmetric, $A(s) = 0$ and we find by application of Theorem 8.1 that

$$\sum_{k=1}^\infty (\alpha \Gamma_k)^{-\frac{1}{\alpha}} V_k \mathbf{1}_{U_k \le t}$$

is an α-stable process on the interval $[0, 1]$, where $V_k, k \ge 1$ are independent and distributed with the same law as V, and $U_k, k \ge 1$ are independent uniform on $[0, 1]$.

There are other $\mathbb{R}$-valued Lévy processes, which admit explicit series representations [cf. 20, Example 6.4 and Example 6.16]. However, it is cumbersome to find closed form expansions of general (dependent) Lévy processes, although Theorem 8.1 is formulated for the multidimensional case already.

A representation using Lévy copulas Tankov [77] has reformulated the representation theorem to incorporate Lévy copulas:

Theorem 8.4 ([77], Theorem 5.6)**.** *Let ν be a finite variational Lévy measure*[2] *on $\mathbb{R}^d$ with marginal tail integrals $U_i, i = 1, ..., d$, Lévy copula F and conditional probability distribution $K(\xi, \cdot)$. Let $\{V_k : k \geq 1\}$ be a sequence of independent random variables, uniformly distributed on $[0, 1]$. Introduce d random sequences $\{\Gamma_k^1 : k \geq 1\}, ..., \{\Gamma_k^d : k \geq 1\}$, independent from $\{V_k : k \geq 1\}$ such that*

(1) $N = \sum_{k=1}^{\infty} \delta_{\Gamma_k^1}$ is a Poisson random measure on $\mathbb{R}$ with intensity measure λ.

(2) Conditionally on Γ_k^1, the random variable $(\Gamma_k^2, ..., \Gamma_k^d)$ is independent from Γ_l^i with $l \neq k$ and all i and is distributed on $\mathbb{R}^{d-1}$ with law $K(\Gamma_k^1, dx_2, ..., dx_d)$.

Then $\{X_t : 0 \leq t \leq 1\}$ with

$$X_t^i = \sum_{k=1}^{\infty} U_i^{-1}(\Gamma_k^i) 1_{[0,t]}(V_k), \ i = 1, ..., d \tag{8.1}$$

is a Lévy process on the time interval $[0, 1]$ with characteristic exponent

$$\psi^{X_t}(z) = \int_{\mathbb{R}^d} (e^{iz.x} - 1)\nu(dx).$$

Theorem 8.4 involves the conditional probability distribution associated to the Lévy copula inasmuch as it supports the separation of the dependence structure and the margins.

Remark 8.5. The conditional probability distribution K_ξ is known from Theorem 7.13.

The following lemma constructs an initial series $\{\Gamma_k^1 : k \geq 1\}$ so that condition (1) in Theorem 8.4 is satisfied:

Lemma 8.6 (cf. [77], Remark 5.4)**.** *Let $\{\Gamma_k^1, k \geq 1\}$ be an alternating series of jump times $\{T_k : k \geq 1\}$ of a Poisson processes with intensity equal to 2, hence $\Gamma_k^1 = T_k(-1)^k, k \geq 1$. Then $N = \sum_{k=1}^{\infty} \delta_{\Gamma_k^1}$ is a Poisson random measure on $\mathbb{R}$ with intensity measure λ.*

The series representation of a symmetric α-stable process in Example 8.3 is slightly different from using (8.1):

Example 8.7. Let ν be a univariate α-stable Lévy measure with density (7.8) using $\lambda_+ = \lambda_- = 1$ and tail integral U. It follows from Theorem 8.4 that

$$\sum_{k=1}^{\infty} (\alpha \Gamma_k)^{-\frac{1}{\alpha}} \mathbf{1}_{[0,t]} V_k$$

is an α-stable process on the interval $[0, 1]$, where $V_k, k \geq 1$ are independent and uniformly distributed on $[0, 1]$, and $\Gamma_k, k \geq 1$ are Poisson arrivals as described in Lemma 8.6.

[2] Tankov [77, Theorem 5.7] generalizes this result to the case of an infinite variational Lévy measure

Of course, there is no need of a $K(\Gamma_k^1, \cdot)$-distributed random variable $(\Gamma_k^2, ..., \Gamma_k^d)$ for all $k \geq 1$ in the 1-dimensional case. Then let us expand a Lévy process $\{X_t\}$ on $\mathbb{R}^2$ with symmetric α-stable margins and an Archimedean dependence structure:

Example 8.8. Let ν_1, ν_2 be univariate α-stable Lévy measures with density (7.8) using $\lambda_+ = \lambda_- = 1$ and tail integral U_1, U_2. Further let $K(\xi, \cdot)$ be the conditional probability measure (7.37) associated to the general Clayton Lévy 2-copula (7.22). By Theorem 8.4, we have that

$$\sum_{k=1}^{\infty} (\alpha\Gamma_k^i)^{-\frac{1}{\alpha}} \mathbf{1}_{[0,t]} V_k, \quad i = 1, 2$$

is an α-stable process on $\mathbb{R}^2$ with general Clayton Lévy copula, where $V_k, k \geq 1$ are independent and uniformly distributed on $[0, 1]$, $\Gamma_k^1, k \geq 1$ are Poisson arrivals as described in Lemma 8.6, and the cdf of $\Gamma_k^2, k \geq 1$ is K_ξ.

The following examples give in the same manner the series representations of Lévy processes $\{X_t\}$ on $\mathbb{R}^3$ with modular and canonical type dependence structures as defined, respectively, in Section 7.3 and Section 7.4:

Example 8.9. Let ν_1, ν_2, ν_3 be univariate α-stable Lévy measures with densities (7.8) using $\lambda_+ = \lambda_- = 1$, and $K(\xi, \cdot)$ be the conditional measure induced by the probabilities (7.29) and (7.30). Let $\Gamma_k^1, k \geq 1$ be as in the previous Example, and for all $k \geq 1$, let the conditional cdf of (Γ_k^2, Γ_k^3) given Γ_k^1 be $K_{\Gamma_k^1}$. By Theorem 8.4, the process $\{X_t : 0 \leq t \leq 1\}$ on $\mathbb{R}^3$ with

$$X_t^i = \sum_{k=1}^{\infty} (\alpha\Gamma_k^i)^{-\frac{1}{\alpha}} 1_{[0,t]}(V_k), \ i = 1, ..., 3$$

is a Lévy process on the time interval $[0, 1]$ with α-stable margins and modular Lévy copula (7.28).

Example 8.10. Let ν_1, ν_2, ν_3 be univariate α-stable Lévy measures with densities (7.8) using $\lambda_+ = \lambda_- = 1$, and $K(\xi, \cdot)$ be the conditional distribution given explicitly by its cdf (7.37). Define the series $\{\Gamma_k^i\}, i = 1, ..., 3$ as in the previous example. Then the process $\{X_t : 0 \leq t \leq 1\}$ on $\mathbb{R}^3$ with

$$X_t^i = \sum_{k=1}^{\infty} (\alpha\Gamma_k^i)^{-\frac{1}{\alpha}} 1_{[0,t]}(V_k), \ i = 1, ..., 3$$

is a Lévy process on the time interval $[0, 1]$ with α-stable margins and a canonical Clayton Lévy copula induced by (7.35) and (7.36).

Examples 8.7 through 8.10 show that series representations of dependent Lévy processes with equal margins only differ in terms of the conditional probability distribution associated to the Lévy copula.

8.2 Sampling dependent jumps

With regard to Examples 8.7, 8.8, 8.9 and 8.10, we simulate (dependent) Lévy processes by series representations. Efficiency of the sampling methods is concerned, too.

Together, Theorem 8.4 and Lemma 8.6 form a constructive approach to the generation of dependent Lévy paths. What is left is truncation of the infinite series (8.1). Here Tankov [cf. 77, Section 5.3] suggests to randomly truncate the series in the following way:

(8.2)
$$X_t^i = \sum_{k:\Gamma_k^1<\tau} U_i^{-1}(\Gamma_k^i)1_{[0,t]}(V_k), \quad i = 1, ..., d,$$

where $\tau > 0$ is some truncation level. This is employed in Algorithm 13, which contains the general procedure to sample paths of a dependent Lévy process $\{X_t\}$ on $\mathbb{R}^d$ with dependence specified by some Lévy copula F (which is given either explicitly or implicitly).
Starting from the scratch, we instance the sampling of a $\mathbb{R}$-valued Lévy process $\{X_t\}$ in

Algorithm 13: Simulation of dependent Lévy processes

Generates trajectory of a Lévy process $\{X_t\}$ on $\mathbb{R}^d$ via series representation. The dependence is given by a Lévy copula F, endowed with its associated conditional cdf $K(\xi,\cdot)$. The marginal Lévy measures have tails $U_1, ..., U_d$. Let a number τ be fixed depending on the required precision and computational capacity.;

Initialize $k = 0, \Gamma_0^1 = 0$;
while $|\Gamma_k^1| < \tau$ **do**
  Set $k = k + 1$;
  Simulate exponential(2) T_k and set $\Gamma_k^1 = (-1)^k(|\Gamma_{k-1}^1| + T_k)$;
  Simulate $(\Gamma_k^2, ..., \Gamma_k^d)$ from distribution $K(\Gamma_k^1, \cdot)$;
  Simulate V_k uniform on $[0, 1]$;
end
The trajectory is then given by $X_t^i = \sum_{k:\Gamma_k^1<\tau} U_i^{-1}(\Gamma_k^i)1_{[0,t]}(V_k),\ i = 1, ..., d.$;

the context of Example 8.7.[3] The resulting is plotted in Figure 8.4.

- The higher the α the more frequent occur the jumps; α determines the jump activity.
- The lower the τ the more jumps are discarded; τ is the truncation level of small jumps.

Hence α is responsible for the jump intensity and τ determines the jump threshold.

[3] The Lévy measure ν of an α-stable process is of finite variation, if and only if $\alpha \in [0, 1]$. Hence, for $\alpha > 1$, Theorem 8.4 does not apply. Even so we stick with Algorithm 13 by thinking of the truncated Lévy measure $\nu_\varepsilon(x) = \nu(x)1_{|x|>\varepsilon}$ where $\varepsilon = U^{-1}(\tau)$., which produces finite variational processes in any case.

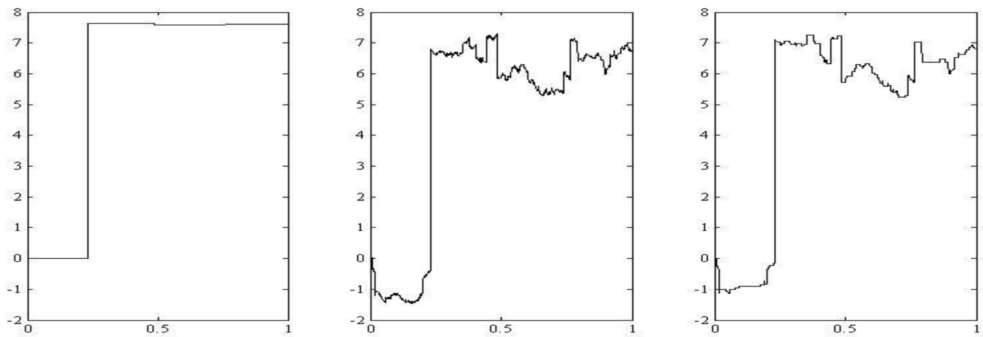

Figure 8.1: Sample paths of α-stable Lévy process on $\mathbb{R}$ using $\alpha = 0.5, \tau = 10000$ (left), $\alpha = 1.5, \tau = 10000$ (center), and $\alpha = 1.5, \tau = 100$ (right).

Having discussed the simulation of an α-stable process, let us now turn to the generation of sample paths from dependent Lévy processes. Here the simulation from the conditional measure associate to a Lévy copula was found to be crucial. That is why we forego the implementation details on the sampling of the marginal process in the following.

Lévy Clayton copula sampling After the conditional cdf K_ξ associated to the general Lévy Clayton 2-copula was given by (7.37), the quantile function K_ξ^{-1} can also be evaluated in closed form [cf. 77, Section 5.3]:

$$\begin{aligned}
K_\xi^{-1}(u) &= B(\xi,u)|\xi| \left\{ C(\xi,u)^{-\frac{\theta}{\theta+1}} - 1 \right\}^{-1/\theta} \\
\text{with } B(\xi,u) &= \operatorname{sgn}(u-1+\eta)1_{\xi\ge 0} + \operatorname{sgn}(u-\eta)1_{\xi<0} \\
\text{and } C(\xi,u) &= \left\{ \frac{u-1+\eta}{\eta} 1_{u\ge 1-\eta} + \frac{1-\eta-u}{1-\eta} 1_{u<1-\eta} \right\} 1_{\xi\ge 0} \\
&+ \left\{ \frac{u-\eta}{1-\eta} 1_{u\ge\eta} + \frac{\eta-u}{\eta} 1_{u<\eta} \right\} 1_{\xi<0}.
\end{aligned} \tag{8.3}$$

As a consequence, for every conditioning realization Γ_k^1 in Algorithm 13, a $K(\Gamma_k^1, \cdot)$-distributed random variable on $\mathbb{R}$ can be obtained by the quantile transformation of a uniform random variable $U \sim UNF(0,1)$.

We implement Algorithm 13 in the context of Example 8.8, using (8.3). Figure 8.2 shows the trajectories of two dependent α-stable processes over various configurations, where we keep the stability index $\alpha = 1.5$ and the truncation level $\tau = 10000$ fixed.[4]

- If $\eta \gg 0$ and $\theta \gg 0$, then the trajectories are merely congruent; together, strong jump sign and jump size dependence leads to almost identical jump behavior.

[4] Figure 8.2 can be compared to the dependent variance gamma processes pictured in Luciano and Schoutens [cf. 51, Figure 1] and Tankov [cf. 78, Figure 1]

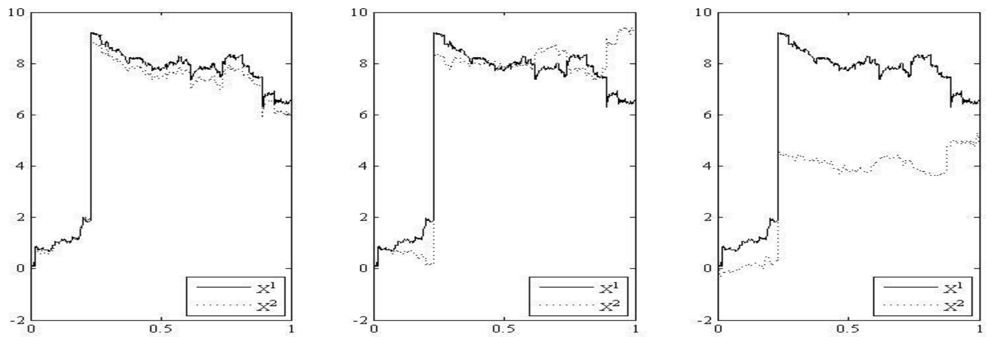

Figure 8.2: Sample paths of 1.5-stable Lévy process on $\mathbb{R}^2$ with Clayton Lévy copula using $\eta = 1, \theta = 6$ (left), $\eta = 0.5, \theta = 6$ (center), and $\eta = 0.5, \theta = 0.5$ (right).

- If $\eta = 0.5$ and $\theta \gg 0$, then there are congruent as well as mirror-inverted partial trajectories; weak jump sign dependence and strong jump size dependence produces divergent jump processes.
- If $\eta \ll 1$ and θ the trajectories are close to being decoupled from each other; weak jump sign and jump size dependence results in an almost independent jump behavior.

Reconsidering the roles of the Lévy Clayton copula parameters from Section 7.2, these findings are not unexpected. The sample paths indicate the same distinction between jump sign and jump size dependence, which we analyzed before.

Quantile transforms of the conditional measures associated with Lévy copulas are generally not available in closed form. This makes numerical sampling procedures necessary. It is true that we know the conditional distribution $K(\Gamma_k^1, \cdot)$ associated with any Lévy copula F by formula (7.12), but its complexity interferes in most cases with a simple random number generation.

Modular Lévy copula sampling The probabilistic interpretations of the modular Lévy copula are derived on the assumption of a given realization. This goes obviously along with the conditional perspective of point (2) in Theorem 8.4.

Regardless of the underlying Lévy copula, the conditional measure $K(\xi, \cdot)$ on $\mathbb{R}^{d-1}$ is totally described by the probabilities $P_\xi[X_i \in \mathcal{I}^c(x_i), 2 \le i \le d], (x_2, ..., x_d) \in \mathbb{R}^{d-1}$. By a simple conditioning argument, we have

$$\begin{aligned} P_\xi[X_i \in \mathcal{I}^c(x_i), i = 2, ..., d] &= P_\xi[|X_i| \le |x_i|, i = 2, ..., d | \alpha_2, ..., \alpha_d] \\ &\cdot \; P_\xi[\operatorname{sgn}(X_i) = \alpha_i, i = 2, ..., d]. \end{aligned}$$

In the case of the modular Lévy copula, the terms on the right hand side are given explicitly by (7.25) and (7.26). Then the general sampling procedure is as follows:

(1) pick a corner α according to g^α given $\alpha_1 = \operatorname{sgn}(\xi)$,

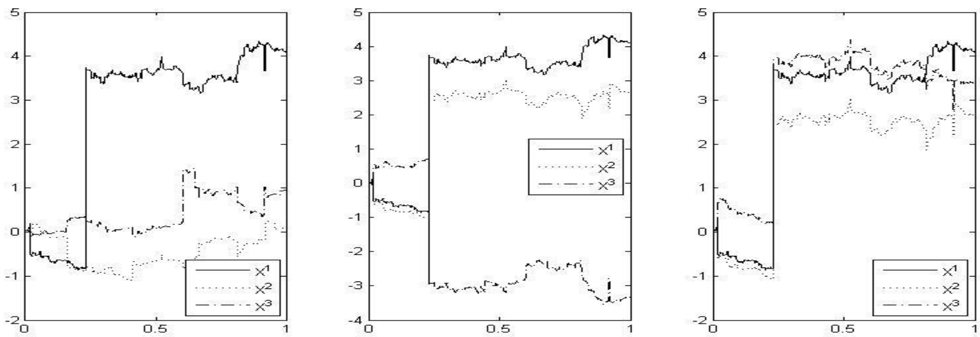

Figure 8.3: Sample paths of 1.5-stable Lévy process on $\mathbb{R}^3$ with modular Lévy copula using g^α as in (7.27), $\theta = 0.5$ (left), g^α as in (7.27), $\theta = 6$ (center), g^α as in (7.24), $\theta = 6$ (right).

(2) and simulate absolute jump sizes $(|x_2|, ..., |x_d|)$ according to F^α given $|x_1| = \xi$.

This is formulated in Algorithm 14, which can be used to sample from the conditional measure associate to a modular Lévy copula.

Algorithm 14: Simulation of conditional measure associated to modular Lévy copula

Samples $(x_2, ..., x_d)$ *from conditional measure* $K(x_1, \cdot)$ *given a realization* x_1*, a Lévy copula* F^+ *on* $[0, \infty]^d$ *and constant joinder* g^α*.*;

Pick corner α according to the conditional probability function associated to g^α;
Simulate absolute jump sizes $(y_2, ..., y_d)$ from the conditional distribution associated to F^+ with conditioning argument $|x_1|g^\alpha$;
Set $x_i = \alpha_i y_i / g^\alpha, i = 2, ..., d$;
The conditional sample vector is given by $(x_2, ..., x_d)$;

Of course, we can not free ourselves from the conditional distribution associated to a Lévy copula by conditioning arguments. However we do produce relief in a way that the distribution is now associated to a positive Lévy copula, which is more manageable. The corner picking should not be harmful either, because, in most cases, one is dealing with simple discrete probability distributions.

Then Algorithm 14 fits well into Algorithm 13. We implement Algorithm 14 to sample a $\mathbb{R}^3$-valued Lévy process with modular Lévy copula in the context of Examples 8.9 and 7.31 over various dependence configurations, where we fix $\alpha = 1.5$ and $\tau = 1000$. The results are shown in Figure 8.3.

- If $\theta < 1$ and g^α as defined in (7.27), then the trajectories are divergent; a moderate sign coupling and a weak jump size dependence induce a loose overall coupling.

- If $\theta \gg 1$ and g^α as defined in (7.27), then the 1st and the 2nd trajectories are nearly aligned while either of these and the 3rd trajectory are rather mirror-inverted; for deeply dependent jump sizes, g^α determines the jump behavior.
- If $\theta \gg 1$ and g^α as defined in (7.24), then the trajectories are almost congruent; strong jump sign and strong jump size dependence yields almost completely dependent jumps.

These observations are comparable to the analysis of the modularly designed Lévy copula in Section 7.3. Here the separate modelling of jump sign and jump size dependence is reinforced. We plotted the scatters of the respective random jump draws in [80].

Canonical Lévy copula sampling Theorem 8.4 requires a slight modification in order to match Tankov's series representations of generally dependent Lévy processes with a canonically designed dependence structure:

Theorem 8.11. *Let v_i be marginal Lévy measures on $\mathbb{R}$ with tail integrals $U_i, i = 1, ..., d$ and $K(\xi, \cdot)$ be a conditional probability measure on $\mathbb{R}^{d-1}$, such that $\nu = f(\mu)$ with μ and f as before is a Lévy measure preserving the margins. Further let $\{V_k : k \geq 1\}$ be a sequence of independent random variables, uniformly distributed on $[0, 1]$. Introduce d random sequences $\{\Gamma_k^1 : k \geq 1\}, ..., \{\Gamma_k^d : k \geq 1\}$, independent from $\{V_k : k \geq 1\}$ such that*

(1) $N = \sum_{k=1}^\infty \delta_{\Gamma_k^1}$ is a Poisson random measure on $\mathbb{R}$ with intensity measure λ

(2) Conditionally on Γ_k^1, the random vector $(\Gamma_k^2, ..., \Gamma_k^d)$ is independent from $\{\Gamma_l^i\}$ with $l \neq k$ and all i and is distributed on $\mathbb{R}^{d-1}$ with law $K(\Gamma_k^1, dx_2, ..., dx_d)$.

Then $\{X_t\}$ defined by

$$X_t^i = \sum_{k=1}^\infty U_i^{-1}(\Gamma_k^i) 1_{[0,t]}(V_k), \ i = 1, ..., d$$

is a Lévy process on the time interval $[0, 1]$ with characteristic exponent

$$\psi^{X_t}(z) = \int_{\mathbb{R}^d} (e^{iz.x} - 1)\nu(dx).$$

Proof. The proof is essentially the same as given in Tankov [77] for the case of a generic Lévy copula. The only difference is that the copula is given implicitly, although it exists by Theorems 7.12 and 7.32, and so the proof applies. □

Hence Algorithm 13 applies, where simulation of the conditional probability measure $K(\Gamma_k^1, \cdot)$ remains crucial. Here the canonical approach (7.30) is advantageous insofar as the sampling from a K_ξ only requires the following two steps:

(1) simulate $(v_2, ..., v_d)$ from the ordinary copula function C_ξ,

Algorithm 15: Simulation of conditional measure associate to canonical Lévy copula

Samples $(x_2, ..., x_d)$ *from conditional measure* $K(x_1, \cdot)$ *given a canon realization* x_1, *an ordinary copula* C *and Lévy copulas* F^i, $i = 2, ..., d$.;

Generate sample $(u_2, ..., u_d)$ from ordinary copula C_{x_1};
Set $x_i = (G^i_{x_1})^{-1}(u_i), i = 2, ..., d$ with
$G^i_{x_1}(x) = \mathrm{sgn}(\xi)\frac{\partial}{\partial\xi}V_{F^i}((0 \wedge \xi, 0 \vee \xi] \times (-\infty, x])$;
The conditional sample vector is given by $(x_2, ..., x_d)$;

(2) and compute the quantile transforms $u_i = (G^i_\xi)^{-1}(v_i)$ for $i = 2, ..., d$.

This idea is formulated in Algorithm 15, that may be used to sample from a conditional measure $K(\xi, \cdot)$ associated to a canonical Lévy copula.

It is worth mentioning that simulation from the conditional distribution associated to a canonical Lévy copula can be managed without high dimensional Lévy copulas. All we need to worry about is Lévy 2-copulas and ordinary $(d-1)$-copulas. This produces relief to the simulation problem, which is complex otherwise [cf. 79, Section 3.2].

Remark 8.12. We discussed the sampling from ordinary copula functions in Chapter 3.

Remark 8.13. We gave the inverse of the conditional distribution function $G^i_{x_1}$ associated to the bivariate Lévy Clayton copula analytically in (8.3).

Then Algorithm 15 fits well into Tankov's Algorithm 13. We implement Algorithm 13 to sample a $\mathbb{R}^3$-valued Lévy process with canonical Lévy copula in the context of Example 8.10 over various dependence configurations, where we keep $\alpha = 1.5$ and $\tau = 1000$ fixed. Figure 8.3 shows the resulting sample paths, where we restrict ourselves to perfect jump sign dependence $\eta = 1$ for the better interpretation.

- If $\theta < 1$ and $\kappa < 1$, then the trajectories are merely decoupled from one another, albeit the 2nd and the 3rd margin signify a minimum of interdependence; neither the canon nor the non-canon dependence is strong and so the margins are nearly independent unless the nested structures add up.

- When increasing the κ, the 2nd and 3rd trajectory converge, while the 1st marginal trajectory is still divergent; the ordinary copula parameter κ is responsible for the stand alone (conditional) dependence of the non-canon variables.

- When increasing the θ instead, the 2nd and the 3rd component converge to the 1st; the dependence on the canon overcomes the conditional association between the non-canon margins.

These results correspond to our analytical findings on the canonical Lévy copula of Section 7.4. Moreover, we detected graphically that the conditional dependence, which is induced by the ordinary copula parameter, is carried forward to the association between non-canon

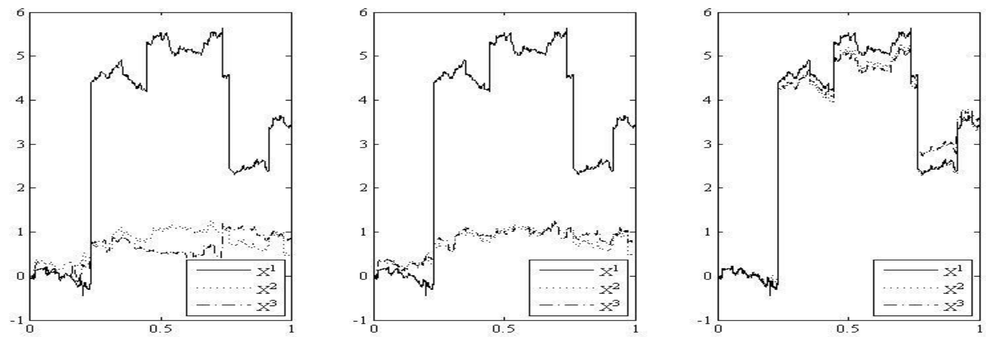

Figure 8.4: Sample paths of 1.5-stable Lévy process on $\mathbb{R}^3$ with canonical Lévy copula using $\theta = 0.5, \kappa = 0.5$ (left), $\theta = 0.5, \kappa = 6$ (center), $\theta = 6, \kappa = 0.5$ (right).

margins only. The canon dependence, which is characterized by the Lévy copulas, proves relevant for all pairwise associations. This supports the idea of a driving force component. We plotted the scatters of the respective random jump draws in [79].

8.3 Summary

In this chapter we developed simulation procedures for multidimensional Lévy processes on the basis of truncated series representations. Having discussed the general procedure, we showed a large interest in sampling modularly and canonically dependent Lévy processes.

First, we thoroughly expanded a 1-dimensional α-stable process to an infinite series of random variables in the light of the general representation theorem.

Then we came to appreciate that Tankov's modified series expansion offers a very elegant way to include dependence issues. This was explained using the example of α-stable margins in combination with the general Clayton, the modular and the canonical Lévy copula.

We formulated the respective sampling approaches. Our contribution here is that we worked out tailor-made solutions for the modular and the canonical Lévy copula. Specifically, we exploited the probabilistic interpretation of the modular Lévy copula in so far as it allows us to simulate the jump signs and the jump sizes separately. The canonical design, instead, proved advantageous by the use of ordinary copula methods.

Finally, we implemented the procedures so-devised. Regarding the dependent sample paths, we rediscovered the analytical features of both Lévy copula models, such as the sharp distinction of jump sign and jump size or the implication of a driving jump component.

Chapter 9

Lévy processes in finance

In this chapter we introduce exponential Lévy processes into derivative pricing models. This entails measure transformations in Lévy models and semianalytical pricing techniques.

Having prepared financial markets with exponential Lévy assets, Section 9.1 describes the equivalent change of measure in full detail with regard to the martingale property of an exponential Lévy process. This is subsequently matched with price relations between plain vanilla and exotic options. Section 9.2 then targets semianalytical pricing techniques. Here Fourier inversion is the approach of choice.

The presentation of exponential Lévy models in this chapter follows closely along Cont and Tankov [20] and Eberlein and Papapantoleon [29]. The pricing techniques originate mainly from Raible [64] and Carr and Madan [14].

Keeping to the standard assumptions in the literature [cf. 64, 29], we restrict ourselves to Lévy processes, which satisfy the following integrability conditions:

Assumption 9.1 (cf. [64], Section 3.2)**.** For a Lévy process $\{X_t\}$ on $\mathbb{R}^d$ with characteristic triplet (γ, A, ν), it holds

$$\int_{|x|\leq 1} |x|\nu(dx) < \infty,$$

and there exists $M > 1$ such that

$$\int_{|x|>1} e^{u.x}\nu(dx) < \infty \quad \forall u \in [-M, M]^d.$$

Assumption 9.1 characterizes finite variational Lévy processes $\{X_t\}$ on $\mathbb{R}^d$ with exponential moments up to an order $M > 1$.

Remark 9.2. α-stable processes are of finite variation, only if $\alpha < 1$ but fail to have an exponential moment of any order. In contrast, Lévy processes whose Lévy measure is either discrete or of Merton type are of finite variation in all cases.

9.1 Exponential Lévy models

Exponential models have long been considered in the theory of stochastic finance, the most prominent of which is the geometric Brownian motion. We go into details about the analytical properties relating to measure transformations and martingales, following which pricing relations are given for plain vanilla and Margrabe options.

The characterization of exponential models due to Karatzas and Shreve [46] will be useful for us:

Definition 9.3. An *exponential Lévy model* on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ consists of

(1) a d-dimensional Lévy process $(\{X_t\}, P)$ with triplet (γ, A, ν)

(2) a constant *risk-free rate* $r \geq 0$

(3) a constant *dividend rate* $\delta \in \mathbb{R}^d$

(4) a vector of positive, constant *initial asset prices* $S_0 \in \mathbb{R}^d$.

We refer to this financial market as $\mathcal{M} = \{S_0, \gamma, A, \nu, r, \delta\}$. $\mathcal{M}$ involves the 1-dimensional price process $(\{S_t^0\}, P)$, which accounts for the *money market* process given by

$$S_t^0 = e^{rt}, \quad t \geq 0, \tag{9.1}$$

and d *asset price* processes $(\{S_t^i\}, P)$ given by

$$S_t^i = S_0^i e^{(r-\delta_i)t + X_t^i}, \quad t \geq 0, i = 1, ..., d. \tag{9.2}$$

We will mainly use 2-asset models in the context of stock and foreign exchange markets:

Example 9.4. Let $\mathcal{M} = ((S_0^1, S_0^2), A, \nu, r, (\delta_1, \delta_2))$ be a 2-dimensional market model on $(\Omega, \mathcal{F}, P)$ with money account price $S_t^0 = e^{rt}, t \in [0, T]$ and stock prices

$$S_t^1 = e^{(r-\delta_1)+X_t^1}, S_t^2 = e^{(r-\delta_2)+X_t^2}, \quad t \in [0, T],$$

where $X_t = (X_t^1, X_t^2)$ a Lévy process on $\mathbb{R}^2$ with triplet (γ, A, ν) satisfying (9.4), $r > 0$ the risk-free interest rate and $\delta \in \mathbb{R}^2$ the dividend vector. This market is referred to as the *(2-dimensional) stock market model* in the following. The covariance matrix consists of σ_1, σ_2 and ρ in the usual way:

$$A = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}.$$

σ_1, σ_2 are the volatility parameters of S_t^1, S_t^2 and ρ is the linear correlation coefficient.

Example 9.5. Let $\mathcal{M} = \big((R_0^1, R_0^2), A, \nu, r_h, (r_f^1, r_f^2)\big)$ be a 2-dimensional market model on $(\Omega, \mathcal{F}, Q)$ with money account price $S_t^0 = e^{rt}, t \in [0,T]$ and foreign exchange rates

$$R_t^1 = e^{(r_h - r_f^1) + X_t^1}, R_t^2 = e^{(r_h - r_f^2) + X_t^2}, \quad t \in [0,T],$$

where $X_t = (X_t^1, X_t^2)$ a Lévy process on $\mathbb{R}^2$ with triplet (γ, A, ν) satisfying (9.4), $r_h > 0$ the domestic risk-free rate and $r_f^1, r_f^2 > 0$ the foreign risk-free rates. This market is referred to as the *(2-dimensional) foreign exchange market model* in the following. The covariance matrix consists of σ_1, σ_2 and ρ in the usual way:

$$A = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}.$$

σ_1, σ_2 are the volatility parameters of R_t^1, R_t^2 and ρ is the linear correlation coefficient. The quotient process $\{R_t^c\}$ defined by

$$R_t^c = R_t^2 / R_t^1 = R_0^2 / R_0^1 e^{(r_f^1 - r_f^2)t + X_t^2 - X_t^1}$$

is referred to as the *cross rate*.

The notion of the martingale property of the market $\mathcal{M}$ will be important:

Definition 9.6 (cf. [20], Proposition 9.1)**.** Let $\mathcal{M}$ be a financial market on $(\Omega, \mathcal{F}, P)$. A measure $Q \approx P$ on $(\Omega, \mathcal{F})$ such that

$$E^Q\left[e^{\delta_i(T-t)}\frac{S_T^i}{S_T^0}|\mathfrak{F}_t\right] = \frac{S_t^i}{S_t^0}, \quad \forall i = 1, ..., d, \tag{9.3}$$

is called an *equivalent martingale measure*. Moreover, for $i = 1, ..., d$, the process $(\{\bar{S}_t\}, Q)$ with $\bar{S}_t = e^{\delta_i t} S_t^i / S_t^0$ is said to be *martingale* (or a martingale).

It follows from the deterministic money market account (9.1) that the martingale condition (9.3) can be written as

$$e^{-(r-\delta_i)(T-t)} E^Q\left[S_T^i | \mathcal{F}_t\right] = S_t^i.$$

Hence the asset prices $(\{S_t^i\}, Q)$, discounted at rate r and reinvested at rate δ_i, have mean rate of return equal to zero. This is equal to saying that $\{e^{X_t^i}\}$ fulfills the martingale condition

$$E^Q\left[e^{X_T^i} | \mathcal{F}_t\right] = e^{X_t^i}, \quad i = 1, ..., d.$$

Existence and uniqueness of an equivalent martingale measure correspond to the notions of an arbitrage-free and an complete market. We use these relations to define arbitrage-free and complete markets:

Definition 9.7. A market model $\mathcal{M}$ on $(\Omega, \mathcal{F}, P)$ is *arbitrage-free* if and only if there exists a martingale measure $Q \approx P$. A market model $\mathcal{M}$ on $(\Omega, \mathcal{F}, P)$ is *complete* if and only if there is a unique martingale measure $Q \approx P$.

The martingale property of (exponential) Lévy processes is related to the characteristic triplet of $\{X_t\}$:

Proposition 9.8 (cf. [29], Section 2)**.** *Let $\{X_t\}$ be a Lévy process on $\mathbb{R}^d$ with characteristic triplet (γ, A, ν).*

(1) $\{X_t^i\}$ is a martingale if and only if $\int_{|x|\geq 1} |x|\nu(dx) < \infty$ and

$$\gamma_i + \int_{|x|\geq 1} x_i \nu(dx) = 0.$$

(2) $\{e^{X_t^i}\}$ is a martingale if and only if $\int_{|x|\geq 1} e^{x_i}\nu(dx) < \infty$ and

$$\frac{1}{2}A_{ii} + \gamma_i + \int_{\mathbb{R}^d} (e^{x_i} - 1 - x_i \mathbf{1}_{|x|\leq 1})\nu(dx) = 0. \tag{9.4}$$

Proof. The proof follows directly from $e^{X_t^i} = \varphi^{X_t}((0, ..., -i, ..., 0))$ and Theorem 7.3. Note that the requirements of the conditions (1) and (2) are met by Assumption 9.1. □

Measure transformations Following Sato [69], the equivalence of two probability measures is related to the characteristic triplet of the Lévy process:

Theorem 9.9 ([69], Theorem 33.1)**.** *Let $(\{X_t\}, P)$ and $(\{X_t\}, \widehat{P})$ be two Lévy processes on $\mathbb{R}^d$ with generating triplets (γ, A, ν) and $(\widehat{\gamma}, \widehat{A}, \widehat{\nu})$, respectively. Then the following two statements (1) and (2) are equivalent:*

(1) $P|_{\mathcal{F}_t} \approx \widehat{P}|_{\mathcal{F}_t}$ for every $t \in (0, T]$.

(2) The generating triplets satisfy

$$\begin{aligned} A &= \widehat{A}, \\ \nu &\approx \widehat{\nu} \\ \text{and} \quad \widehat{\gamma} - \gamma - \int_{|x|\leq 1} x(\widehat{\nu} - \nu)(dx) &\in \{Ax : x \in \mathbb{R}^d\} \end{aligned}$$

with the function $\phi(x)$ defined by $\phi(x) = \ln \frac{d\widehat{\nu}}{d\nu}$ satisfying

$$\int_{\mathbb{R}^d} (e^{\phi(x)/2} - 1)^2 \nu(dx) < \infty.$$

Theorem 9.9 describes the equivalence of the two probability measures P and $\widehat{P}$ in terms of existing triplets (γ, A, ν) and $(\widehat{\gamma}, \widehat{A}, \widehat{\nu})$. The following result prepares the ground for a constructive approach to equivalent measures:

Theorem 9.10 ([69], Theorem 33.2)**.** *Let* $(\{X_t\}, P)$ *and* $(\{X_t\}, \widehat{P})$ *be two Lévy processes on* $\mathbb{R}^d$ *with generating triplets* (γ, A, ν) *and* $(\widehat{\gamma}, \widehat{A}, \widehat{\nu})$*, respectively. Suppose that the equivalent conditions (1) and (2) in the previous theorem are satisfied. Choose* $\eta \in \mathbb{R}^d$ *such that*

$$\widehat{\gamma} - \gamma - \int_{|x|\leq 1} x(\widehat{\nu} - \nu)(dx) = A\eta. \tag{9.5}$$

Then we define, P-a.s.,

$$\begin{aligned} U_t &= \eta.(X_t^c) - \frac{t}{2}\eta.A\eta - t\gamma.\eta \\ &\quad + \lim_{\varepsilon \downarrow 0}\left(\sum_{(s,\Delta X_s)\in(0,t]\times\{|x|>\varepsilon\}} \phi(\Delta X_s) - t\int_{|x|>\varepsilon} (e^{\phi(x)/2} - 1)^2 \nu(dx) \right), \end{aligned} \tag{9.6}$$

where ϕ *is the function in (2),* $(\{X_t^C\}, P)$ *the continuous part of* $(\{X_t\}, P)$ *and the convergence in the right-hand side of* (9.6) *is uniform in* t *on any bounded interval, P-a.s. Moreover, for every* $t \in [0, T]$*,*

$$\begin{aligned} E^P[e^{U_t}] &= E^{\widehat{P}}[e^{-U_t}] = 1 \\ \text{and} \quad \frac{dP}{d\widehat{P}}|_{\mathcal{F}_t} &= e^{U_t}, \quad P\text{-a.s.} \end{aligned}$$

The process $(\{U_t\}, P)$ *is a Lévy process on* $\mathbb{R}$ *with generating triplet* (γ_U, A_U, ν_U) *expressed by*

$$\begin{aligned} \gamma_U &= -\frac{1}{2}\eta.A\eta - \int_{\mathbb{R}} (e^y - 1 - y\mathbf{1}_{0<|y|\leq 1}(y)(\nu\phi^{-1})(dy).a \\ A_U &= \eta.A\eta \\ \nu_U &= [\nu\phi^{-1}]_{|\mathbb{R}\setminus\{0\}} \end{aligned}$$

Theorem 9.10 may be used to build a new Lévy process $(\{X_t\}, \widehat{P})$ from $(\{X_t\}, P)$ by an equivalent measure transformation from P to $\widehat{P}$ [cf. 69, Definition 33.4]:

Definition 9.11. Let $(\{X_t\}, P)$ be a Lévy process on $\mathbb{R}^d$ with generating triplet (γ, A, ν). Given $\phi(x)$ satisfying (9.5) and given $\eta \in \mathbb{R}^d$, define $\widehat{A} = A$, $\widehat{\nu}(dx) = e^{\phi(x)}\nu(dx)$ and $\widehat{\gamma}$ by (9.5). This gives us U_t as in Theorem 9.10 and the probability measure $\widehat{P}$ by

$$\widehat{P}|_{\mathcal{F}_t}[B] = E^P[e^{U_t}\mathbf{1}_B] \text{ for } B \in \mathcal{F}_t.$$

Then $(\{X_t\}, \widehat{P})$ is a Lévy process, which has generating triplet $(\widehat{\gamma}, \widehat{A}, \widehat{\nu})$.

We call this construction of $(\{X_t\}, \widehat{P})$ by $(\{X_t\}, P)$ *density transformation with* $\phi(x)$ *and* η, and $\phi(x)$ and η the *Girsanov quantities* [cf. 82, Theorem 1.20]. The following result relates the existence of an equivalent martingale measure to the Girsanov quantities:

Theorem 9.12 (cf. [82], Theorem 1.22)**.** *Let $(\{X_t\}, P)$ be a Lévy process on $\mathbb{R}^d$ with triplet (γ, A, ν), and let $(\{X_t\}, \widehat{P})$ be constructed from $(\{X_t\}, P)$ via the density transformation with $\phi(x)$ and η. Then $(\{e^{X_t}\}, \widehat{P})$ is a martingale if and only if*

$$\gamma + A\eta + \frac{1}{2}diag(A) + \int_{\mathbb{R}^d} (e^x - \mathbf{1})e^{\phi(x)} - x1_{|x|\leq 1}\nu(dx) = 0. \tag{9.7}$$

From Theorem 9.12, we can easily show existence of an equivalent martingale measure in the case of a regular Brownian part.

Corollary 9.13. *Let $(\{X_t\}, P)$ be a Lévy process on $\mathbb{R}^d$ with triplet (γ, A, ν), where A a regular matrix. Then for any $\xi \in [-(M-1), M-1]^d$ there exists an $\eta \in \mathbb{R}^d$ such that $(\{X_t\}, \widehat{P})$ constructed from $(\{X_t\}, P)$ via density transformation by $\phi(x) = \xi.x$ and η is a martingale.*

Proof. By Assumption 9.1, for $\phi(x) = \xi.x, \xi \in [-(M-1), M-1]^d$, the integral

$$\int_{\mathbb{R}^d} (e^x - \mathbf{1})e^{\xi.x} - x1_{|x|\leq 1}\nu(dx) = 0$$

is finite. Hence, condition (9.7) turns into a linear equation for η, which has a unique solution if A is regular. □

Corollary 9.13 in conjunction with (9.5) shows that, for Lévy processes with regular Brownian part, an equivalent martingale process can be constructed by a simple change of the drift, whatever the transformation of the Lévy measure.[1] This is equivalent to saying that there exists a generally non-unique equivalent martingale measure in merely all cases of interest. Then exponential Lévy models are arbitrage-free but incomplete in general. Moreover, the findings reveal a clear connection of a probability measure P to the characteristic triplet (γ, A, ν). Hence we denote $P = P(\gamma, A, \nu)$ whenever indexing is advantageous.

The density transformation discussed most frequently in finance literature is the Esscher transform of a probability measure:

Example 9.14 (Sato [69], Example 33.14)**.** Given a Lévy process $(\{X_t\}, P)$ with (γ, A, ν), let $\eta \neq \mathbf{0}$ and $\phi(x) = \eta.x$. If further $E^P[e^{\eta.X_t}] < \infty$ is satisfied, we can determine a new Lévy process $(\{X_t\}, \widehat{P})$ by the density transformation by our η and ϕ. It follows from the definition of U_t and from the Levy-Ito decomposition that

$$\begin{aligned} U_t &= \eta.X_t - t\int_{\mathbb{R}^d} (e^{\eta.x} - 1 - \eta.x\mathbf{1}_{|x|\leq 1}(x))\nu(dx) + \frac{1}{2}\eta.A\eta + \gamma.\eta \\ &= \eta.X_t - t\psi(-i\eta), \end{aligned}$$

where ψ the characteristic exponent of $(\{X_t\}, P)$. This density transformation is known as the *Esscher transform*. By (9.4), if the $e^{X_t^i}$ are martingales for $i = 1, ..., d$, then the density process $\{U_t\}$ reduces to $U_t = \eta.X_t$.

[1] The existence result can be generalized onto the case of non-regular (or vanishing) Brownian part [cf. 20, Section 9.5, for the one-dimensional case].

Next, linear transformations of a Lévy process $(\{X_t\}, P)$ on $\mathbb{R}^d$ under an equivalent change of measure are characterized:

Proposition 9.15 ([29], Proposition 6.1)**.** *Let $(\{X_t\}, P)$ be a Lévy process on $\mathbb{R}^d$ with triplet (γ_c, A, ν), let u, v be vectors in $\mathbb{R}^d$ such that $v \in [-M, M]^d$. Moreover, let $\widehat{P} \approx P$, with density*

$$\frac{d\widehat{P}|_{\mathcal{F}_t}}{dP|_{\mathcal{F}_t}} = \frac{e^{v.X_t}}{E^P[e^{v.X_t}]}.$$

Then, the 1-dimensional process $(\{\widehat{X}_t\}, \widehat{P})$ defined by $\widehat{X}_t := u.X_t$ is a Lévy process on $\mathbb{R}^d$ with the characteristic triplet $(\widehat{\gamma}_c, \widehat{A}, \widehat{\nu})$ given by

$$\begin{aligned}
\widehat{\gamma}_c &= u.\gamma_c + \frac{1}{2}(u.Av + v.Au) + \int_{\mathbb{R}^d} u.x(e^{v.x} - 1)\nu(dx) \\
\widehat{A} &= u.Au \\
\widehat{\nu} &= \mathcal{T}(\mu),
\end{aligned}$$

where $\mathcal{T}$ is a mapping $\mathcal{T} : \mathbb{R}^d \to \mathbb{R}$ such that $x \mapsto \mathcal{T}(x) = u.x$, and μ is a measure defined by

$$\mu(B) = \int_B e^{v.x}\nu(dx)$$

Proposition 9.15 allows us to compute the characteristic triplet of a linearly transformed Lévy process under some equivalent probability measure explicitly.[2] This is now matched with the martingale property in the 2-dimensional case:

Corollary 9.16. *Let $(\{X_t\}, P)$ be a Lévy process on $\mathbb{R}^2$ and such that e^{X_t} is a P-martingale. Define $\widehat{P} \approx P$ by the density $d\widehat{P}|_{\mathcal{F}_t}/dP|_{\mathcal{F}_t} = e^{X_t^1}$. Then $(\{\widehat{X}_t\}, \widehat{P})$ defined by $\widehat{X}_t := X_t^2 - X_t^1$ is a Lévy process and $e^{\widehat{X}_t}$ is a $\widehat{P}$-martingale.*

Proof. Suppose $(\{X_t\}, P)$ has characteristic triplet (γ_c, A, ν). By the martingale condition, it holds

$$\gamma_c^i = -\frac{1}{2}A_{ii} - \int_{\mathbb{R}^2} (e^{x_i} - 1 - x_i)\nu(dx), \ i = 1, 2.$$

For $u = (-1, 1)$ and $v = (1, 0)$, Proposition 9.15 then gives us the characteristic triplet $(\widehat{\gamma}_c, \widehat{A}, \widehat{\nu})$ of $\widehat{X}_t$ under $\widehat{P}$. We show that $(\widehat{\gamma}_c, \widehat{A}, \widehat{\nu})$ fulfills the martingale condition by

[2] Papapantoleon [61] discusses linearly transformed time-inhomogeneous Lévy process under equivalent measure transformations.

comparing the Brownian motion and the jump characteristics separately:

$$\begin{aligned}
\widehat{\gamma}_c^C &= u.\gamma_c^C + \frac{1}{2}(u.Av + v.Au) \\
&= -\frac{1}{2}u.diag(A) + \frac{1}{2}(u.Av + v.Au) \\
&= \frac{1}{2}\{(2u_1v_1 - u_1)A_{11} + (2u_2v_1 + 2u_1v_2)A_{12} + (2u_2v_2 - u_2)A_{22}\} \\
&= \frac{1}{2}\{-A_{11} + 2A_{12} - A_{22}\} \\
&= -\frac{1}{2}\{u_1u_1A_{11} + 2u_1u_2A_{12} + u_2u_2A_{22}\} \\
&= -\frac{1}{2}u.Au \\
&= -\frac{1}{2}\widehat{A}
\end{aligned}$$

and

$$\begin{aligned}
\widehat{\gamma}_c^J &= u.\gamma_c^J + \int_{\mathbb{R}^d} u.x(e^{v.x} - 1)\nu(dx) \\
&= \int_{\mathbb{R}^2} (e^{x_1} - 1 - x_1) - (e^{x_2} - 1 - x_2)\nu(dx) + \int_{\mathbb{R}^2} (x_2 - x_1)(e^{x_1} - 1)\nu(dx) \\
&= \int_{\mathbb{R}^2} (e^{x_1} - e^{x_2} + (x_2 - x_1)e^{x_1})\nu(dx) \\
&= -\int_{\mathbb{R}^2} (e^{x_2 - x_1} - 1 - (x_2 - x_1))e^{x_1}\nu(dx) \\
&= -\int_{\mathbb{R}} (e^y - 1 - y)\widehat{\nu}(dy).
\end{aligned}$$

Altogether, we have

$$\widehat{\gamma}_c = -\frac{1}{2}\widehat{A} - \int_{\mathbb{R}} (e^y - 1 - y)\widehat{\nu}(dy).$$

This characterizes $e^{\widehat{X}_t}$ as a $\widehat{P}$-martingale. □

Corollary 9.16 shows that, for an exponential martingale process $(\{X_t\}, P)$ on $\mathbb{R}^2$, the ratio process $e^{X_t^2}/e^{X_t^1}$ is a martingale under the measure $\widehat{P}$, which results from the Esscher transform by $\eta = (1, 0)$.[3]

Remark 9.17. It is worth stressing that Proposition 9.15 produces 1-dimensional Lévy processes. The characteristic triplet (γ, A, ν) of an $\mathbb{R}^d$-valued Lévy process $(\{X_t\}, P)$ generates the characteristic triplet $(\widehat{\gamma}, \widehat{A}, \widehat{\nu})$ of the $\mathbb{R}$-valued Lévy process $(\{\widehat{X}_t\}, \widehat{P})$.

[3] Corollary 9.16 can alternatively be proved using change of numeraire techniques [cf. 11].

As to Remark 9.17, what comes to mind immediately in the context of Lévy copulas is how the dependence structure of measure ν is involved in the construction of measure $\widehat{\nu}$. This is explain in terms of a discrete Lévy measure ν on $\mathbb{R}^2$ in the sense of Example 7.15:

Corollary 9.18. *Let ν_{12} be the discrete Lévy measure on $\mathbb{R}^2$ constructed in Example 7.15. Define $\widehat{P} \approx P$ by the density $d\widehat{P}/dP = e^{X_T^1}/E[e^{X_T^1}]$. Then, the Lévy measure $\widehat{\nu}$ of $(\{\widehat{X}_t\}, \widehat{P})$ with $\widehat{X}_t = X_t^2$ has discrete support*

$$\{z_k, k = 1, ..., N\}, \text{ where } z_k = (k - \frac{N+1}{2})\Delta$$

and weights

$$\widehat{\nu}^k = \sum_{i=1}^{N} e^{x_i} \nu_{12}^{ik} \tag{9.8}$$

Proof. Proposition 9.15 applied to Lévy measure ν_{12} gives, for $u = (0, 1)$ and $v = (1, 0)$,

$$\begin{aligned} \widehat{\nu}(B) &= \int_{\mathbb{R}} e^x \int_{y \in B} \nu_{12}(dx, dy) \\ &= \sum_{i=1}^{N} e^{x_i} \sum_{j: y_j \in B} \nu_{12}^{ij}, \quad B \in \mathcal{B}(\mathbb{R}). \end{aligned}$$

Due to the constraint $y_j \in B$, the measure $\widehat{\nu}$ has again discrete support 9.8. The weights 9.8 then result from the inner product of the discrete integrant $(e^{x_1}, ..., e^{x_N})$ and the k-th row of the weight matrix $\nu_{12}^{ij}, i, j = 1, ..., N$. □

Corollary 9.19. *Let ν_{12} be the discrete Lévy on $\mathbb{R}^2$ constructed in Example 7.15. Define $\widehat{P} \approx P$ by the density $d\widehat{P}/dP = e^{X_T^1}/E[e^{X_T^1}]$. Then, the Lévy measure $\widehat{\nu}$ of $(\{\widehat{X}_t\}, \widehat{P})$ with $\widehat{X}_t = X_t^2 - X_t^1$ has (extended) discrete support*

$$\{z_k, k = 1, ..., 2 * N - 1\}, \text{ where } z_k = (k - N) \tag{9.9}$$

and weights

$$\widehat{\nu}^k = \sum_{l} e^{x_l} \nu_{12}^{l, c(l)}, \tag{9.10}$$

where summation is from $\underline{s} := \max\{\frac{N+1}{2} - (k-1), 1\}$ to $\overline{s} := \min\{N - (k - \frac{N+1}{2}), N\}$ and $c(l) = l - \underline{s} + 1 + \max\{k - \frac{N-1}{2}\}$.

Proof. Proposition 9.15 applied to Lévy measure ν_{12} gives, for $u = (-1, 1)$ and $v = (1, 0)$,

$$\begin{aligned} \widehat{\nu}(B) &= \int_{\mathbb{R}} e^x \int_{y \in B + x} \nu_{12}(dx, dy) \\ &= \sum_{i=1}^{N} e^{x_i} \sum_{j: y_j \in B + x_i} \nu_{12}^{ij}, \quad B \in \mathcal{B}(\mathbb{R}^2). \end{aligned}$$

Due to the constraint $y_j \in B+x_i$, the measure $\widehat{\nu}$ has discrete support (9.9). Weights (9.10) then result from the (partial) inner product of the discrete integrant $(e^{x_1}, ..., e^{x_N})$ and the $(k-(N+1)/2)$-th diagonal of weights $\nu_{12}^{ij}, i, j = 1, ..., N$. The assignement of indices is purely technical. □

Symmetries for pricing derivatives There is a one-to-one correspondence between arbitrage-free pricing rules and equivalent martingale measures:

Proposition 9.20 ([20], Proposition 9.1). *In a financial market $\mathcal{M}$ on $(\Omega, \mathcal{F}, P)$ any arbitrage-free linear pricing rule V for a payoff H can be represented as*

$$V_t^H = e^{-r(T-t)} E^Q[H|\mathcal{F}_t],$$

where Q an equivalent martingale measure.

In other words, the fair price of a contract is its discounted payoff as expected under an equivalent martingale measure and the current information.

Remark 9.21. We have shown how to construct equivalent martingale measures in nearly all cases of interest[4] by the Esscher transform. This has characterized exponential Lévy markets as arbitrage-free models in the context of Definition 9.7.

In view of Remark 9.21, we assume existence of an equivalent martingale measure Q and model the financial market $\mathcal{M}$ on $(\Omega, \mathcal{F}, Q)$ arbitrage-free. Then the asset prices have mean rates of return $\mu_i = r - \delta_i, i = 1, ..., d$ such that the price processes $e^{-(r-\delta_i)t} S_t^i$, discounted at rate r and reinvested at rate δ_i, are martingales under Q. We assume $t = 0$ for convenience in the following.

Example 9.22. Let $\mathcal{M} = ((S_0^1, S_0^2), A, \nu, r, (\delta_1, \delta_2))$ be the 2-dimensional stock market on $(\Omega, \mathcal{F}, Q)$ as defined in Example 9.4, where Q a martingale measure. Then the value of a European plain vanilla put option V_0^p on stock price S_t^i with strike K and maturity T is given by

$$V_0^p(K, T; S_0^i, A_{ii}, \nu_i, r, \delta_i) = e^{-rT} E^Q[(K - S_T^i)^+]$$

for $i = 1, 2$. The value of a Margrabe option V_0^m on stock prices S_t^1, S_t^2 with maturity T is given by

$$V_0^m(T; S_0^1, S_t^2, A, \nu, r, \delta_1, \delta_2) = e^{-rT} E^P[(S_T^1 - S_T^2)^+].$$

Eberlein and Papapantoleon [29] have proven a symmetry between the values of a Margrabe and some plain vanilla option in exponential Lévy models. We adopt the arguments used therein[5] for the case of homogeneous Lévy processes:

---

[4]One can show that if the trajectories of $(\{X_t\}, P)$ are neither almost surely increasing nor almost surely decreasing, then there exists an equivalent martingale measure.

[5]Eberlein and Papapantoleon [29] use time-inhomogeneous Lévy processes, which are briefly addressed in Section B.1.

Theorem 9.23 (cf. [29], Theorem 6.2)**.** *The value of a Margrabe and a European plain vanilla option are related via the following symmetry:*

$$V_0^m(T; S_0^1, S_0^2, A, \nu, r, \delta_1, \delta_2) = S_0^1 V_0^p(1, T; S_0^2/S_0^1, \widehat{A}, \widehat{\nu}, \delta_1, \delta_2),$$

where $\widehat{A}$ and $\widehat{\nu}$ are defined as in 9.15 for $v = (1,0)$ and $u = (-1,1)$.

Proof. Expressing the value of the Margrabe option in terms of asset S^1, we get:

$$\begin{aligned} V_0^m(T; S_0^1, S_0^2, A, \nu, r, \delta_1, \delta_2) &= e^{-rT} E^Q[(S_T^1 - S_T^2)^+] \\ &= e^{-rT} E^Q[S_T^1(1 - \frac{S_T^2}{S_T^1})^+] \\ &= S_0^1 e^{-\delta_1 T} E^Q[e^{X_T^1}(1 - \widehat{S}_T)^+], \end{aligned} \tag{9.11}$$

where $\widehat{S}_T = e^{(\delta_1 - \delta_2)T + X_T^2 - X_T^1}$. As to Definition 9.11 and Example 9.14, we can construct a new Lévy process $(\{X_t\}, \widehat{Q})$ by the linear density transformation

$$\frac{d\widehat{Q}|_{\mathcal{F}_t}}{dQ|_{\mathcal{F}_t}} = e^{X_t^1}.$$

Then the right-hand side in (9.11) may be rewritten as

$$S_0^1 e^{-\delta_1 T} E^Q[e^{X_T^1}(1 - \widehat{S}_T)^+] = S_0^1 e^{-\delta_1 T} E^{\widehat{Q}}[(1 - \widehat{S}_T)^+].$$

By Corollary 9.16, the exponential Lévy process $e^{-(\delta_1 - \delta_2)t}\widehat{S}_t$ is a martingale under $\widehat{Q}$. By 9.20, we then have that $e^{-\delta_1 T} E^{\widehat{Q}}[(1 - \widehat{S}_T)^+]$ is a pricing formula for a European plain vanilla put option on asset $\widehat{S}_t$ with strike $K = 1$, risk-free rate δ_1 and dividend rate δ_2. □

The line of argumentation used in the proof of Theorem 9.23 provides an elegant interpretation of the pricing symmetry as an equivalent change between the arbitrage-free markets $\mathcal{M}$ and $\widehat{\mathcal{M}} = (\widehat{S}_0, \widehat{A}, \widehat{\nu}, \delta_1, \delta_2)$, where $\widehat{\mathcal{M}}$ is 1-dimensional.[6]

Example 9.24. Let $\mathcal{M} = \big((R_0^1, R_0^2), A, \nu, r_h, (r_f^1, r_f^2)\big)$ be the 2-dimensional foreign exchange market on $(\Omega, \mathcal{F}, Q)$ as defined in Example 9.5, where Q a martingale measure. Then the value of a European plain vanilla call option V_0^c on foreign exchange rate R_t^i with strike K and maturity T is given by

$$V_0^c(K, T; R_0^i, A_{ii}, \nu_i, r_h, r_f^i) = e^{-r_h T} E^Q[(R_T^i - K)^+]$$

for $i = 1, 2$. Let us in this context denote the value of an European plain vanilla call option on the cross exchange rate R_t^c associated to R_t^1 and R_t^2 with strike K and maturity T by $V_0^c(K, T; R_0^c, A, \nu, r_h, r_f^1, r_f^2)$.

[6] In fact, we have used a change-of-numeraire technique as described in [11]. Papapantoleon [61] refers to this technique as duality method. One can show similar symmetries for asset-or-nothing options and quanto-options [cf. 29, Theorem 6.3 and Theorem 6.4].

An arbitrage-free pricing rule for the call option on the cross exchange rate is obtained similarly to Theorem 9.23:

Theorem 9.25. *The value of a European call option on the cross exchange rate associated to R^1 and R^2 is given by*

$$V_0^c(K,T;R_0^c,A,\nu,r_h,r_f^1,r_f^2) = V_0^c(K,T;R_0^2/R_0^1,\widehat{A},\widehat{\nu},r_f^1,r_f^2), \tag{9.12}$$

where $\widehat{A}$ and $\widehat{\nu}$ are defined as in 9.15 for $v=(1,0)$ and $u=(-1,1)$.

Proof. The proof of Theorem 9.25 follows the argumentation used in the previous proof. As to Definition 9.11 and Example 9.14, choosing $\eta=(1,0)$, we can construct a new Lévy process $(\{X_t\},\widehat{Q})$ by the linear density transformation

$$\frac{d\widehat{Q}|_{\mathcal{F}_t}}{dQ|_{\mathcal{F}_t}} = e^{X_t^1}.$$

By Corollary 9.16, the exponential Lévy process $e^{-(r_f^1-r_f^2)t}R_t^c$ is a martingale under $\widehat{Q}$. By 9.20, we then have that $e^{-r_f^1T}E^{\widehat{Q}}[(R_T^c-K)^+]$ is a pricing formula for a European plain vanilla call option on the cross rate R_t^c with strike K, domestic risk-free rate r_f^1 and foreign risk-free rate r_f^2. □

The equivalent martingale transformation in the proof of Theorem 9.25 may again be interpreted as an equivalent change between the arbitrage-free markets $\mathcal{M}$ and $\widehat{\mathcal{M}} = (R_t^c,\widehat{A},\widehat{\nu},r_f^1,r_f^2)$, where $\widehat{\mathcal{M}}$ is now 1-dimensional.

Remark 9.26. Theorems 9.23 and 9.25 indicate, in particular, that both the Margrabe option in the context of a stock market and the call option on the cross rate in a foreign exchange market are redundant assets. Hence option prices should be fixed by option prices of the primary assets.

9.2 Semi-analytical option pricing

Different from the Black-Scholes model, closed form solutions of option prices are generally not available in exponential Lévy models. This is mainly due to the fact that the pdf of Lévy processes are not known in most cases. Instead, the characteristic function of a Lévy process is given explicitly by the Levy-Khintchin representation. This has lead to the development of Fourier based option pricing methods.

Definition 9.27. Let h be a complex-valued integrable function on $\mathbb{R}^d$. The *Fourier transform Fh of h* is a complex-valued function on $\mathbb{R}^d$ defined by

$$(Fh)(v) = \int_{\mathbb{R}^d} e^{iv.x}h(x)dx, \quad v\in\mathbb{R}^d.$$

Note that for a random variable X on $\mathbb{R}^d$ with pdf f^X the Fourier transform (Ff^X) is the characteristic function φ^X of X. We will need to invert the Fourier transform for pricing purposes:

Proposition 9.28 ([69], Proposition 37.2)**.** *Let h be a complex-valued integrable function on $\mathbb{R}^d$. Then, Fh is continuous and bounded. If Fh is integrable, then*

$$h(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-ix.v}(Fh)(v)dv$$

for almost every $x \in \mathbb{R}^d$ and the function on the right-hand side is continuous and bounded.

In other words, under certain continuity and integrability conditions we may transform a function h back and forth between the Fourier space and the original space. This is the key to semi-analytical pricing methods.

Convolution representation The value of an option can be represented by the convolution of the Laplace transform of the option payoff and the characteristic function of the underlying price process.

Definition 9.29. Let $h(x)$ be a real-valued function. The *bilateral (or two-sided) Laplace transform* of h is defined as

$$\mathcal{L}_h(z) := \int_{\mathbb{R}} e^{-zx}h(x)dx, \quad z \in \mathbb{C}.$$

The bilateral Laplace transform of plain vanilla put and call options can be written out explicitly:

Example 9.30 (cf. [62], Example 15.3)**.** Consider the modified payoff function $h(x) = (e^{-x} - K)^+$ of a plain vanilla call option. The bilateral Laplace transform of h at $z \in \mathbb{C}$

$$\begin{aligned} \mathcal{L}_h(z) &= \int_{\mathbb{R}} e^{-zx}(e^{-x} - K)^+ dx \\ &= \int_{\mathbb{R}} e^{-(z+1)x} - Ke^{-zx} dx \end{aligned}$$

is finite, only if $\Re z \in (-\infty, -1)$. In this case

$$\mathcal{L}_h(z) = K^{z+1} \frac{1}{z(z+1)}. \tag{9.13}$$

In an analogous way, for $z \in \mathbb{C}$ with $\Re z \in (0, \infty)$, 9.13 is the bilateral Laplace transform of the modified payoff function $h(x) = (K - e^{-x})^+$ of a plain vanilla put option.

Proposition 9.20 gives the value of a payoff in terms of the expectation under a martingale measure. The following result, which is due to Raible [64], shows how this is written in terms of convolutions. Here we extend the argumentation to incorporate for dividend yields:

Theorem 9.31 (cf. [64], Theorem 3.2)**.** *Consider a European option with payoff $H(S_T)$ at time T. Let $h(x) := H(e^{-x})$ denote the modified payoff function. Assume that $x \mapsto e^{-Rx}|h(x)|$ is bounded and integrable for some $R \in \mathbb{R}$ such that $|R| < M$. Let $V_0^h(\xi)$ denote the time-0 price of this option taken as a function of the negative log forward price $\xi := -ln\{e^{(r-\delta)T}S_0\}$. Then we have*

$$V_0^h(\xi) = \frac{e^{\xi R - rT}}{2\pi}\int_{-\infty}^{\infty} e^{iu\xi}\mathcal{L}_h(R+iu)\varphi^{X_T}(iR-u)du,$$

whenever the integral on the r. h. s. exists (at least as a Cauchy principal value).

Proof. Raible's proof can easily be adopted to the case of dividend payments. □

Remark 9.32. Theorem 9.31 allows for different representations of the inverse Fourier integral. The choice of R corresponds to shifting the contour of integration along the complex plane.[7]

Theorem 9.31 then allows to represent, for example, the value of a Margrabe option in the context of a stock market model by convolutions[8], using the symmetry proved in Theorem 9.23:

Corollary 9.33. *Let $\mathcal{M} = ((S_0^1, S_0^2), A, \nu, r, (\delta_1, \delta_2))$ be the 2-dimensional stock market on $(\Omega, \mathcal{F}, Q)$ as defined in Example 9.4, where Q a martingale measure. Then the value of a Margrabe option V_0^m on stock prices S_t^1, S_t^2 with maturity T is given by*

$$V_0^m(T; S_0^1, S_t^2, A, \nu, r, \delta_1, \delta_2) = \frac{S_0^1 e^{\xi R - \delta_1 T}}{2\pi}\int_{-\infty}^{\infty} \frac{e^{iu\xi}}{(R+iu)(R+iu+1)}\varphi^{\widehat{X}_T}(iR-u)du, \tag{9.14}$$

where $\xi = -\ln(e^{(\delta_1-\delta_2)T}S_0^2/S_0^1)$, $R \in (0, M)$ and $\phi^{\widehat{X}_T}$ the characteristic function for time T of the martingale process $(\{\widehat{X}_t\}, \widehat{Q})$ with $\widehat{A}$ and $\widehat{\nu}$ as defined in Theorem 9.23.

Proof. Corollary 9.33 follows directly from Theorems 9.23 and 9.31. □

The integral in (9.14) is in fact a Fourier transform. Hence FFT methods[9] apply for its calculation.

Fourier inversion In a similar way, Carr and Madan express the value of an European option on the basis of another Fourier transform.

Let us extend their reasoning, which is described in Cont and Tankov [20, Section 11.1.3], by dividend payments. Hence consider the price of a European call option

$$V_0^c(k) = e^{-rT}E[(e^{(r-\delta)T + X_T} - e^k)^+],$$

[7] Kahl and Lord [44] discuss optimal contours of integration from a numerical perspective.

[8] Eberlein and Özkan [28], Eberlein and Koval [27] apply Raible's convolution representation to the pricing of interest rate caps and cross-currency derivatives.

[9] Fast Fourier transform methods are described in Carr and Madan [14].

in terms of the logarithmic strike $k = \log(K)$, on a dividend paying asset with normalized initial value $S_0 = 1$. We would like to express its Fourier transform in terms of the characteristic function φ^{X_T} of X_T to find the prices by Fourier inversion. This requires to proceed to the modified time value of the option,

$$v(k) = e^{-rT}E[(e^{(r-\delta)T+X_T} - e^k)^+] - (e^{-\delta T} - e^{k-rT})^+, \tag{9.15}$$

due to the integrability constraint in Definition 9.27. Then the Fourier transform of (9.15) may be worked out in the style of Cont and Tankov [20], where we extend their line of argumentation by dividend payments (which is not trivial):

Theorem 9.34 (cf. [20], Section 11.1.3)**.** *The Fourier transform of the modified time value of a European call option on a dividend paying asset is given by*

$$(Fv)(z) = e^{-rT}e^{(iz+1)(r-\delta)T}\frac{\varphi^{X_T}(z-i) - 1}{iz(iz+1)}. \tag{9.16}$$

Proof. Let $S_t = e^{(r-\delta)t+X_t}$ be the price process for a dividend paying asset, modelled under some arbitrage-free pricing measure. That is equivalent to saying that $e^{-(r-\delta)t}S_t = e^{X_t}$ is a martingale. By

$$\int_{\mathbb{R}} \rho_T(x)e^{-\delta T+x}\mathbf{1}_{k\leq(r-\delta)T}dx = e^{-\delta T}\mathbf{1}_{k\leq(r-\delta)T},$$

we then have

$$\begin{aligned}
v(k) &= e^{-rT}E^Q[(e^{(r-\delta)T+X_T} - e^k)^+] - (e^{-\delta T} - e^{k-rT})^+ \\
&= e^{-rT}\int_{\mathbb{R}} \rho_T(x)(e^{(r-\delta)T+x} - e^k)\mathbf{1}_{k\leq(r-\delta)T+x}dx \\
&\quad - \int_{\mathbb{R}} \rho_T(x)(e^{-\delta T+x} - e^{k-rT})\mathbf{1}_{k\leq(r-\delta)T}dx \\
&= e^{-rT}\int_{\mathbb{R}} \rho_T(x)(e^{(r-\delta)T+x} - e^k)(\mathbf{1}_{k\leq(r-\delta)T+x} - \mathbf{1}_{k\leq(r-\delta)T})dx.
\end{aligned}$$

Hence it holds

$$\begin{aligned}
(Fv)(z) &= \int_{\mathbb{R}} e^{izk}e^{-rT}\int_{\mathbb{R}} \rho_T(x)(e^{(r-\delta)T+x} - e^k)(\mathbf{1}_{k\leq(r-\delta)T+x} - \mathbf{1}_{k\leq(r-\delta)T})dxdk \\
&= e^{-rT}\int_{\mathbb{R}} \rho_T(x)\int_{(r-\delta)T}^{(r-\delta)T+x} e^{izk}(e^{(r-\delta)T+x} - e^k)dkdx,
\end{aligned}$$

by interchanging integrals. The inner integral $\int_{(r-\delta)T}^{(r-\delta)T+x} e^{izk}(e^{(r-\delta)T+x} - e^k)dk$ is

$$
\begin{aligned}
\left[\frac{e^{izk+(r-\delta)T+x}}{iz} - \frac{e^{(iz+1)k}}{iz+1}\right]_{(r-\delta)T}^{(r-\delta)T+x} &= \frac{e^{(iz+1)((r-\delta)T+x)}}{iz} - \frac{e^{(iz+1)((r-\delta)T+x)}}{iz+1} \\
&- \frac{e^{(iz+1)(r-\delta)T+x}}{iz} + \frac{e^{(iz+1)(r-\delta)T}}{iz+1} \\
&= \frac{(iz+1)e^{(iz+1)((r-\delta)T+x)} - ize^{(iz+1)((r-\delta)T+x)}}{iz(iz+1)} \\
&- \frac{(iz+1)e^{(iz+1)(r-\delta)T+x} + ize^{(iz+1)(r-\delta)T}}{iz(iz+1)} \\
&= \frac{e^{(iz+1)((r-\delta)T+x)}}{iz(iz+1)} - \frac{e^{(iz+1)(r-\delta)T+x}}{iz(iz+1)} \\
&+ \frac{e^{(iz+1)(r-\delta)T}(1-e^x)}{iz+1}.
\end{aligned}
$$

Then

$$
\begin{aligned}
(Fv)(z) &= e^{-rT}\int_{\mathbb{R}} \rho_T(x)\left\{\frac{e^{(iz+1)((r-\delta)T+x)}}{iz(iz+1)} - \frac{e^{(iz+1)(r-\delta)T+x}}{iz(iz+1)} + \frac{e^{(iz+1)(r-\delta)T}(1-e^x)}{iz+1}\right\}dx \\
&= e^{-rT}\left\{\frac{e^{(iz+1)(r-\delta)T}}{iz(iz+1)}E[e^{(iz+1)X_T}] - \frac{e^{(iz+1)(r-\delta)T}}{iz(iz+1)}E[e^{X_T}] + \frac{e^{(iz+1)(r-\delta)T}}{iz+1}E[1-e^{X_T}]\right\} \\
&= e^{-rT}e^{(iz+1)(r-\delta)T}\frac{\varphi^{X_T}(z-i)-1}{iz(iz+1)}.
\end{aligned}
$$

□

Given the Fourier transform of the modified time value of the call option in closed form by (9.16), the original value can be found by Fourier inversion:

$$
v(k) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-izk}(Fv)(z)dz. \tag{9.17}
$$

As to (9.15), the put option price is obtained by simply adding $(e^{-\delta T} - e^{k-rT})^+$ to (9.17).

Theorem 9.34 may then be used to represent, for example, the value of a call option on the cross rate in the context of a foreign exchange market model by Fourier transforms, taking into account the symmetry established in Theorem 9.25.

Corollary 9.35. *Let $\mathcal{M} = \big((R_0^1, R_0^2), A, \nu, r_h, (r_f^1, r_f^2)\big)$ be the 2-dimensional foreign exchange market on $(\Omega, \mathcal{F}, Q)$ as defined in Example 9.5, where Q a martingale measure. Then the value of a call option on the cross exchange rate R^c associated to R^1 and R^2 is given by*

$$
V_0^c(K,T;R_0^c,A,\nu,r_h,r_f^1,r_f^2) = \frac{e^{-r_f^1 T}R_0^c}{2\pi}\int_{\mathbb{R}} e^{-izk}e^{(iz+1)(r_f^1-r_f^2)T}\frac{\varphi^{\widehat{X}_T}(z-i)-1}{iz(iz+1)}dz, \tag{9.18}
$$

where $k = -\ln(K)$ *and* $\varphi^{\widehat{X}_T}$ *the characteristic function at time* T *of the martingale process* $(\{\widehat{X}_t\}, \widehat{Q})$ *with* $\widehat{A}$ *and* $\widehat{\nu}$ *as defined in Theorem 9.25.*

Proof. Corollary 9.35 follows directly from Theorems 9.25 and 9.34. □

The integral in (9.18) is in fact a Fourier transform. Hence, we will again be able to apply FFT methods for its calculation.

9.3 Summary

In this chapter we described exponential Lévy models for use in finance applications. We called detailed attention to the martingale property of a Lévy process with regard to pricing measure transformations.

First, we introduced exponential Lévy processes into financial asset price modelling. We gave both a stock market and a foreign exchange market as instances of a coupled process environment.

Then we explained equivalent measure transformations in exponential Lévy models, emphasizing the relation between equivalent martingale measures and the characteristic Lévy triplet. It was interesting to see that we may construct infinitely many equivalent measures (or Lévy processes, in fact), that are martingale, in merely all cases of interest. This rendered exponential Lévy models incomplete in general.

Continuing, we detailed the change of measure technique using the example of discrete Lévy characteristics in two dimensions. Specifically, we proved the martingale property for the quotient of the two processes under some equivalent measure resulting from the Esscher transform.

Having mentioned the principles of arbitrage-free option pricing, we denoted the fair value of a plain vanilla and a Margrabe option written on stock prices or foreign exchange rates. Using the change of measure technique, pricing symmetries could be found. Our contribution here is that we related the value of a call option written on the cross rate of two exchange rates to the value of a call option written on their quotient rate. This made the cross rate a redundant asset in the foreign exchange model.

In the end, we showed that the fair value of an option can basically be represented as the inverse Fourier transform of the characteristic function of the underlying Lévy process. Our contribution here is to generalize the existent representations to include dividend payments. This was subsequently applied to the call option on the cross exchange rate using the previous symmetry result.

Chapter 10

Option pricing applications

In this chapter we explain how to back out the parameters of an exponential Lévy model from liquid option prices. Our goal here is to detect the dependence structure implied in a foreign exchange market.

Section 10.1 is devoted to the inverse problem of calibrating a specified model to market data in the least-squares sense. The calibration method is elaborated on a non-parametric approach. Here semi-analytical pricing techniques prove useful for the efficient dealing with implied Lévy copulas. In response to the ill-posedness of the inverse calibration problem, regularization is treated in Section 10.2. Then Section 10.3 analyzes how the options on the cross rate are redundant in the context of a liquid triangular foreign exchange market. Specifically, Lévy copulas cope with nearly reproducing option prices on the cross rate.

The non-parametric calibration approach dealt with in this chapter follows the exhaustive presentation in Cont and Tankov [20]. The handling of market implied dependence structures and the application to the foreign exchange market is new.

10.1 Copula model calibration

We have mentioned that an exponential Lévy model $\mathcal{M}$ as to Definition 9.3 is arbitrage-free but incomplete in general. This prevents the identification of a unique price V^H for a claim H on the asset price process $\{S_t\}$ by arbitrage arguments alone. Following Proposition 9.20, the value V^H of a claim H is given by the discounted expected payoff under an (equivalent) martingale measure Q, however different choices of the martingale measure lead to a range of possible prices.[1] Whenever option prices are traded on the market, the market quotes can be used as a source of information to help selecting the equivalent pricing measure Q. Extracting the model parameters from observed market quotes so as to reproduce the prices of traded options is referred to as model *calibration*. We explain how to calibrate a 2-dimensional financial market in the following.

[1] It is proved by Eberlein and Jacod [25] that, for a European call option with strike K and maturity T, the range of option prices under all possible equivalent martingale measures spans the whole no-arbitrage interval $[(S_0 - Ke^{-rT})^+, S_0]$.

Let therefore $V_j, j \in J$ be prices quoted on the market for a set of benchmark options. Then we search for an arbitrage-free model $\mathcal{M}$ on $(\Omega, \mathcal{F}, Q)$ which prices these options market-consistently:

$$V_j = e^{-r(T-t)} E^Q[H_j|\mathcal{F}_t], \quad \forall j \in J.$$

The calibration problem as such is ill-posed: there may be many pricing models which reproduce the benchmark prices. In other words, there need not be a unique solution. Therefore calibration constraints are added to the effect that all candidate models $\mathcal{M}$ belong to a certain class of models. Then it is no longer guaranteed that the models in the chosen class are consistent with the benchmark options. A more realistic interpretation of the calibration problem is to achieve the best approximation of market prices of options within a given model class.

In our context, we understand the approximation problem in a least-squares sense. Hence we consider a market model $\mathcal{M} = \mathcal{M}(\theta) = \{S_0, \gamma(\theta), A(\theta), \nu(\theta), r, \delta\}$ on $(\Omega, \mathcal{F}, Q(\theta))$, where S_0, r, δ are observed from the market and $\theta \in \Theta$ is some set of triplet parameters such that $Q(\theta) = Q(\gamma(\theta), A(\theta), \nu(\theta))$ is a martingale measure. Then we try to minimize the quadratic pricing error $\varepsilon(\theta)$ over all parameterizations $\theta \in \Theta$ in the following sense:

Definition 10.1. Let $V_j, j \in J$ be market quotes of European call options with strike K_j and maturity T_j. Define a financial market $\mathcal{M}(\theta) = \{S_0, \gamma(\theta), A(\theta), \nu(\theta), r, \delta\}$ as in Definition 9.3, where $\theta \in \Theta$ arbitrage-free model parameters. Then the *quadratic pricing error* $\varepsilon : \Theta \mapsto \mathbb{R}$ is defined by

$$\varepsilon(\theta) = \sum_{j \in J} |V(H_j; \theta) - V_j|^2, \tag{10.1}$$

where $V(H_j; \theta)$ are the model prices corresponding to payoff H_j.[2]

The best model calibration $\theta^* \in \Theta$ is then defined as $\theta^* = \arg\min_{\theta \in \Theta} \varepsilon(\theta)$. This is the set of parameters most consistent with the market quotes. The corresponding triplet $(\gamma(\theta^*), A(\theta^*), \nu(\theta^*))$ so obtained is referred to as the *implied characteristic triplet*. It involves the *implied covariance matrix* $A(\theta^*)$ and the *implied Lévy measure* $\nu(\theta^*)$.

Remark 10.2. Recall that $\gamma(\theta)$ is determined by $A(\theta)$ and $\nu(\theta)$ through Proposition 9.8, if θ corresponds to an arbitrage-free market parametrization.

If the implied characteristic triplet corresponds to a $\mathbb{R}^d$-valued Lévy process, then we call the correlation matrix contained in $A(\theta^*)$ the *implied correlation* matrix and the Lévy copula inherent in $\nu(\theta^*)$ the *implied Lévy copula*.

Let us now be more specific about the actual pricing rule for $V(H_j; \theta)$ by instance of the foreign exchange market of Example 9.5:

Example 10.3. Let $\mathcal{M}(\theta) = \big((R_0^1, R_0^2), A(\theta), \nu(\theta), r_h, (r_f^1, r_f^2)\big)$ be the 2-dimensional foreign exchange market on $(\Omega, \mathcal{F}, Q(\theta))$ as defined in Example 9.5, where $\theta \in \Theta$ arbitrage-free

[2] Cont and Tankov [20] and Cont and Luciano [19] use the weighted pricing error $\varepsilon(\theta) = \sum_{j \in J} \omega_j |V(K_j, T_j; \theta) - V_j|^2$, instead, to take into account the reliability of quotes.

model parameters. If $V_j, j \in J$ correspond to market quotes of European call options on rate $\{R_t^i\}$ with strike K_j and maturity T_j, then the quadratic pricing error $\varepsilon : \Theta \mapsto \mathbb{R}$ is

$$\varepsilon(\theta) = \sum_{j \in J} |V_0^c(K_j, T_j; R_0^i, A_{ii}(\theta), \nu_i(\theta), r_h, r_f^i) - V_j|^2,$$

where $V_0^c(K_j, T_j; R_0^i, A_{ii}(\theta), \nu_i(\theta), r_h, r_f^i)$ is given by (9.12). If $V_j, j \in J$ correspond to market quotes of European call options on the cross rate $\{R_t^c\}$ with strike K_j and maturity T_j, then the quadratic pricing error $\varepsilon : \Theta \mapsto \mathbb{R}$ is

$$\varepsilon(\theta) = \sum_{j \in J} |V_0^c(K_j, T_j; R_0^2/R_0^1, \widehat{A}(\theta), \widehat{\nu}(\theta), r_f^1, r_f^2) - V_j|^2,$$

where $V_0^c(K_j, T_j; R_0^2/R_0^1, \widehat{A}(\theta), \widehat{\nu}(\theta), r_f^1, r_f^2)$ is given by (9.12).

Gradient based methods A convenient way to search for the optimal model parametrization $\theta^* \in \Theta$ numerically is using a *gradient descent method*[3]. This requires the computation of the quadratic pricing error and its gradient. It can be seen from (10.1) that the evaluation of model prices $V(H_j;\theta)$ and gradients $\partial V(H_j;\theta)/\partial\theta$ is the only crucial problem in this respect. We concentrate on European call options for benchmark products V_j in the following.[4].

By Theorem 9.34, the price V_0^c of a European plain vanilla call option on asset price S_t with strike K and maturity T can be written in terms of the logarithmic strike $k = \ln(K)$. Denoting $V(k) = V_0^c(e^k, T; S_0, \gamma(\theta), \sigma(\theta), \nu(\theta), r, \delta)$, we have

$$V(k) = \frac{e^{-rT} S_0}{2\pi} \int_{\mathbb{R}} e^{-izk} e^{(iz+1)(r-\delta)T} \frac{\varphi^{X_T}(z-i) - 1}{iz(iz+1)} dz, \tag{10.2}$$

where φ^{X_T} the characteristic function of the Lévy process with triplet $(\gamma(\theta), \sigma(\theta), \nu(\theta))$ associated to $\theta \in \Theta$.

Remark 10.4. We have argued before that FFT methods are a very efficient way to compute (10.2) over all benchmark strikes K_i.

This enables us to compute the quadratic pricing error immediately. However the computation of the gradient is cumbersome in general. In this respect, Cont and Tankov [20] argue that a non-parametric Lévy model allows of vast savings in computation of the gradient. Then fix a 1-dimensional financial market $\mathcal{M}(\theta) = (S_0, \gamma(\theta), A(\theta), \nu(\theta), r, \delta)$ on $(\Omega, \mathcal{F}, Q(\theta))$ to begin with, where $Q(\theta)$ is a martingale measure and $\nu(\theta)$ is a discrete Lévy measure defined on an equidistant grid by (7.14) in Example 7.15. Then Θ is the set of no-arbitrage parameters $\theta = (\sigma, \nu_1, ..., \nu_N)$.

There are explicit expressions available for the derivatives of the call prices with respect to the model parameters σ and $\nu_j, j = 1, ..., N$. The following lemma will be used:

[3]The principal algorithm is the BFGS quasi-Newton method offered by the MatLab Optimization Toolbox routines. See the MatLab Help for more insights.

[4]Eberlein and Koval [27] treats cross-currency derivatives in the context of a Lévy Libor market model.

Lemma 10.5. *Let* $w(k) = (e^{-\delta T} - e^{k-x_j-rT})^+ - (e^{-\delta T} - e^{k-rT})^+$*. Then the Fourier transform* Fw *of* w *is given by*

$$(Fw)(z) \quad = \quad e^{-rT}\frac{e^{(iz+1)(r-\delta)T}(e^{izx_j}-1)}{iz(iz+1)}.$$

Proof. Consider $x_j > 0$, the other case being derived analogously. Then

$$\int_{\mathbb{R}} e^{izk} w(k)dk \quad = \quad \int_{-\infty}^{(r-\delta)T} e^{izk}e^k(e^{-rT} - e^{-x_j-rT})dk + \int_{(r-\delta)T}^{(r-\delta)T+x_j} e^{izk}(e^{-\delta T} - e^{k-x_j-rT})dk.$$

Exact computation of the integrals on the right hand side leads to

$$\begin{aligned}
\int_{-\infty}^{(r-\delta)T} e^{izk}e^k(e^{-rT} - e^{-x_j-rT})dk &= (e^{-rT} - e^{-x_j-rT})\left[\frac{e^{(iz+1)k}}{iz+1}\right]_{-\infty}^{(r-\delta)T} \\
&= e^{-rT}\frac{e^{(iz+1)(r-\delta)T}(1-e^{-x_j})}{iz+1}
\end{aligned}$$

and

$$\begin{aligned}
\int_{(r-\delta)T}^{(r-\delta)T+x_j} e^{izk}(e^{-\delta T} - e^{k-x_j-rT})dk &= e^{-\delta T}\left[\frac{e^{izk}}{iz}\right]_{(r-\delta)T}^{(r-\delta)T+x} + e^{-x_j-rT}\left[\frac{e^{(iz+1)k}}{iz+1}\right]_{(r-\delta)T}^{(r-\delta)T+x} \\
&= e^{-\delta T}\frac{e^{iz((r-\delta)T+x_j)} - e^{iz((r-\delta)T+x_j)}}{iz} \\
&- e^{-x_j-rT}\frac{e^{(iz+1)((r-\delta)T+x_j)} - e^{(iz+1)((r-\delta)T+x_j)}}{iz} \\
&= e^{-rT}\frac{e^{(iz+1)(r-\delta)T}(e^{izx_j}-1)}{iz} \\
&- e^{-rT}\frac{e^{(iz+1)(r-\delta)T}(e^{izx_j}-e^{-x_j})}{iz+1} \\
&= e^{-rT}\frac{e^{(iz+1)(r-\delta)T}(e^{izx_j}-1)}{iz(iz+1)} \\
&- e^{-rT}\frac{e^{(iz+1)(r-\delta)T}(1-e^{-x_j})}{iz+1}.
\end{aligned}$$

By cancellation of terms, we have

$$\int_{\mathbb{R}} e^{izk}w(k)dk \quad = \quad e^{-rT}\frac{e^{(iz+1)(r-\delta)T}(e^{izx_j}-1)}{iz(iz+1)}.$$

□

Lemma 10.5 is used to differentiate the option price with respect to the volatility and the Lévy measure weights. We extend the result by Cont and Tankov [20] to incorporate for dividend rates:

Proposition 10.6 (cf. [20], Section 13.4)**.** *The derivative of the option price with respect to the the volatility* σ *is*

$$\frac{\partial V}{\partial \sigma}(k) = T\sigma e^{-rT}\frac{1}{2\pi}\int_{\mathbb{R}} e^{-izk}e^{(iz+1)(r-\delta)T}\varphi^{X_T}(z-i)dz$$

The derivative with respect to the discretized variable ν_j *is*

$$\frac{\partial V}{\partial \nu_j}(k) = T(1-e^{x_j})\frac{1}{2\pi}\int_{\mathbb{R}} e^{-izk}e^{(iz+1)(r-\delta)T}\frac{\varphi^{X_T}(z-i)}{(iz+1)}dz + Te^{x_j}\frac{1}{2\pi}\left\{V(k-x_j)-V(k)\right\}.$$

Proof. First observe that the derivative of the option price with respect to the model parameters σ and ν_j is equal to the derivative of the modified time value v as given in (9.17). Using Theorem 9.34, the derivative with respect to the volatility σ of the Fourier transform Fv of v is

$$\begin{aligned}\frac{\partial Fv}{\partial \sigma}(z) &= Te^{-rT}e^{(iz+1)(r-\delta)T}\frac{\varphi^{X_T}(z-i)}{iz(iz+1)}\left\{\sigma iz(iz+1)\right\}\\ &= T\sigma e^{-rT}e^{(iz+1)(r-\delta)T}\varphi^{X_T}(z-i).\end{aligned}$$

Similarly, for $j = 1, ..., N$, the derivative with respect to the parameter ν_j of the Fourier transform Fv of v is given by

$$\begin{aligned}\frac{\partial Fv}{\partial \nu_j}(z) &= Te^{-rT}e^{(iz+1)(r-\delta)T}\frac{\varphi^{X_T}(z-i)}{iz(iz+1)}\left\{iz(1-e^{x_j})+e^{x_j}(e^{izx_j}-1)\right\}\\ &= Te^{-rT}e^{(iz+1)(r-\delta)T}(1-e^{x_j})\frac{\varphi^{X_T}(z-i)}{(iz+1)}\\ &+ Te^{x_j}e^{izx_j}\left\{(Fz)(z)-\frac{e^{-rT}e^{(iz+1)(r-\delta)T}}{iz(iz+1)}\right\}\\ &- Te^{x_j}\left\{(Fz)(z)-\frac{e^{-rT}e^{(iz+1)(r-\delta)T}}{iz(iz+1)}\right\}\\ &= Te^{-rT}e^{(iz+1)(r-\delta)T}(1-e^{x_j})\frac{\varphi^{X_T}(z-i)}{(iz+1)}\\ &+ Te^{x_j}\left\{e^{izx_j}(Fz)(z)-(Fz)(z)\right\}\\ &- Te^{x_j}\left\{\frac{e^{-rT}e^{izx_j}e^{(iz+1)(r-\delta)T}}{iz(iz+1)}-\frac{e^{-rT}e^{(iz+1)(r-\delta)T}}{iz(iz+1)}\right\}.\end{aligned}$$

We invert the Fourier transform term by term and get

$$\frac{\partial v}{\partial \sigma}(k) = T\sigma e^{-rT}\frac{1}{2\pi}\int_{\mathbb{R}} e^{-izk}e^{(iz+1)(r-\delta)T}\varphi^{X_T}(z-i)dz$$

and

$$\begin{aligned}
\frac{\partial v}{\partial \nu_j}(k) &= Te^{-rT}(1-e^{x_j})\frac{1}{2\pi}\int_{\mathbb{R}} e^{-izk}e^{(iz+1)(r-\delta)T}\frac{\varphi^{X_T}(z-i)}{(iz+1)}dz \\
&+ Te^{x_j}\frac{1}{2\pi}\int_{\mathbb{R}} e^{-izk}\left\{e^{izx_j}(Fz)(z)-(Fz)(z)\right\}dz \\
&- Te^{x_j}\frac{1}{2\pi}\int_{\mathbb{R}} e^{-izk}\left\{e^{-rT}e^{(iz+1)(r-\delta)T}\frac{e^{izx_j}-1}{iz(iz+1)}\right\}dz \\
&= T(1-e^{x_j})\frac{1}{2\pi}\int_{\mathbb{R}} e^{-izk}e^{(iz+1)(r-\delta)T}\frac{\varphi^{X_T}(z-i)}{(iz+1)}dz \qquad (10.3)\\
&+ Te^{x_j}\frac{1}{2\pi}\left\{v(k-x_j)-v(k)\right\} \\
&- Te^{x_j}\frac{1}{2\pi}\left\{(e^{-\delta T}-e^{k-x_j-rT})^+ - (e^{-\delta T}-e^{k-rT})^+\right\} \\
&= T(1-e^{x_j})\frac{1}{2\pi}\int_{\mathbb{R}} e^{-izk}e^{(iz+1)(r-\delta)T}\frac{\varphi^{X_T}(z-i)}{(iz+1)} \\
&+ Te^{x_j}\frac{1}{2\pi}\left\{V(k-x_j)-V(k)\right\},
\end{aligned}$$

respectively, where we used Lemma 10.5 for (10.3). □

Proposition 10.6 allows us to decompose the derivatives of plain vanilla option prices with respect to each of the model parameters into a difference of option prices themselves and another Fourier transform, which involves the characteristic function φ^{X_T} of the Lévy exponent. Hence FFT methods apply again and prepare to compute the gradient in a very efficient manner. Then computation of the gradient requires only twice as much evaluations as the price computation itself.

Having discussed 1-dimensional calibration, consider now the 2-dimensional market $\mathcal{M}(\theta) = (S_0, \gamma(\theta), A(\theta), \nu(\theta), r, \delta)$ on $(\Omega, \mathcal{F}, Q(\theta))$, where $Q(\theta)$ is a martingale measure and $\nu(\theta)$ the discrete Lévy measure defined by (7.16) in Example 7.15 with discrete Lévy measures $\nu_i^1, \nu_j^2, i, j = 1, ..., N$ as before and a parametric Lévy 2-copula $F(\cdot;\kappa)$.

Using Proposition 10.6 in combination with a gradient based optimization routine, we may calibrate each marginal market $\mathcal{M}_i(\theta_i) = (S_0^i, \gamma_i(\theta_i), A_{ii}(\theta_i), \nu_i(\theta_i), r_i, \delta_i), i = 1, 2$, over all $(\theta_i) = (\sigma_i, \nu_i^1, ..., \nu_i^N)$ to a set of benchmark options, whose underlyings are the marginal asset price process $\{S_t^i\}$. As regards Corollary 9.19 and Theorem 9.23 (or 9.25), we may just as well fit the changed market model $\widehat{\mathcal{M}}(\theta) = (\widehat{S}_0, \widehat{\gamma}(\theta), \widehat{A}(\theta), \widehat{\nu}(\theta), \delta_1, \delta_2)$ to a set of benchmark options, whose underlyings are the process $\{\widehat{S}_t\}$.

Remark 10.7. Model prices in the changed market $\widehat{\mathcal{M}}$ are available as to Corollary 9.33 or Corollary 9.35.

Remark 10.8. By Corollary 9.19, $\widehat{\nu}$ is again discrete (on an extended grid). Hence Proposition 10.6 applies to the differentiation of plain vanilla options written on $\{\widehat{S}_t\}$ with respect to the new model parameters $(\widehat{\sigma}, \widehat{\nu}_1, ..., \widehat{\nu}_{2*N-1})$.

The derivative of the model prices with respect to the primary market parameters $(\sigma_1, \sigma_2, \rho, \nu_1^1, ..., \nu_1^N, \nu_2^1, ..., \nu_2^N, \kappa)$ is obtained via the chain rule. It is clear from Proposition 9.15 that σ_1, σ_2 and ρ correspond to $\widehat{\sigma}$, whereas ν_1, ν_2 and κ determine $\widehat{\nu}$ (and $\widehat{\gamma}$ is defined by the martingale condition on the way). Specifically, we have

$$\frac{\partial \widehat{\sigma}}{\partial \sigma_1} = \frac{\sigma_1 - \rho\sigma_2}{\sqrt{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}}, \quad \frac{\partial \widehat{\sigma}}{\partial \sigma_2} = \frac{\sigma_2 - \rho\sigma_1}{\sqrt{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}} \tag{10.4}$$

and

$$\frac{\partial \widehat{\sigma}}{\partial \rho} = -\frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}}. \tag{10.5}$$

The derivatives of $\widehat{\nu}$ with respect to ν_1, ν_2 and κ are more involved. The following lemma will be used:

Lemma 10.9. *Let ν_1, ν_2 and ν_{12} be as in Example 7.15 (using the approximation). Then*

$$\begin{aligned}
\frac{\partial \nu_{12}^{ij}}{\partial \nu_l^1} &= \frac{\partial f(U_1(x_i), U_2(x_j); \kappa)}{\partial U_1(x_i)} \nu_i^1 \nu_j^2 \mathbf{1}_{i \in \mathcal{I}(l)} + f(U_1(x_i), U_2(x_j); \kappa) \nu_j^2 \mathbf{1}_{i=l}, \\
\frac{\partial \nu_{12}^{ij}}{\partial \nu_l^2} &= \frac{\partial f(U_1(x_i), U_2(x_j); \kappa)}{\partial U_2(x_j)} \nu_i^1 \nu_j^2 \mathbf{1}_{j \in \mathcal{I}(l)} + f(U_1(x_i), U_2(x_j); \kappa) \nu_i^1 \mathbf{1}_{j=l}, \\
\frac{\partial \nu_{12}^{ij}}{\partial \kappa} &= \frac{\partial f(U_1(x_i), U_2(y_j); \kappa)}{\partial \kappa} \nu_i^1 \nu_j^2
\end{aligned}$$

for $i, j, l = 1, ..., N$.

Proof. From Example 7.15 we have approximately $\nu_{12}^{ij} = f(U_1(x_i), U_2(y_j); \kappa)\nu_1^i \nu_2^j, i, j = 1, ..., N$. By the chain rule, it follows that

$$\begin{aligned}
\frac{\partial \nu_{12}^{ij}}{\partial \nu_l^1} &= \frac{\partial (f(U_1(x_i), U_2(y_j); \kappa)\nu_i^1 \nu_j^2)}{\partial \nu_l^1} \\
&= \frac{\partial f(U_1(x_i), U_2(y_j); \kappa)}{\partial \nu_l^1} \nu_i^1 \nu_j^2 + f(U_1(x_i), U_2(y_j); \kappa) \frac{\partial \nu_i^1 \nu_j^2}{\partial \nu_l^1} \\
&= \frac{\partial f(U_1(x_i), U_2(y_j); \kappa)}{\partial U_1(x_i)} \nu_i^1 \nu_j^2 \mathbf{1}_{i \in \mathcal{I}(l)} + f(U_1(x_i), U_2(y_j); \kappa) \nu_j^2 \mathbf{1}_{i=l},
\end{aligned}$$

where the tail integrals $U_i, i = 1, 2$ are given by (7.15). The derivatives $\partial \nu_{12}^{ij} / \partial \nu_l^2$ and $\partial \nu_{12}^{ij} / \partial \kappa$ are derived analogously. □

Lemma 10.9 allows us to compute the derivative of ν_{12} with respect to all its building blocks. As to Corollary 9.18 and 9.19, this facilitates the direct differentiation of $\widehat{\nu}$ with respect to ν_1, ν_2 and κ:

Proposition 10.10. *Let ν be a discrete Lévy measure on $\mathbb{R}^2$ and let $\widehat{\nu}$ be constructed from ν via* (9.8) *in Proposition 9.15. Then*

$$\frac{\partial \widehat{\nu}^k}{\partial \nu_1^i} = \sum_l e^{x_l} \frac{\partial \nu_{12}^{l,c(l)}}{\partial \nu_1^i}, \quad \frac{\partial \widehat{\nu}^k}{\partial \nu_2^j} = \sum_l e^{x_l} \frac{\partial \nu_{12}^{l,c(l)}}{\partial \nu_2^j}, \quad \frac{\partial \widehat{\nu}^k}{\partial \kappa} = \sum_l e^{x_l} \frac{\partial \nu_{12}^{l,c(l)}}{\partial \kappa},$$

where $\partial \nu_{12}^{l,c(l)}/\partial \nu_1^i$, $\partial \nu_{12}^{l,c(l)}/\partial \nu_2^j$ and $\partial \nu_{12}^{l,c(l)}/\partial \kappa$ are given by Lemma 10.9, and summation is as in Corollary 9.19.

Proof. Proposition 10.10 follows directly from Corollary 9.19 and the chain rule. □

Together with Proposition 10.6, these latest results enable us to compute the gradient of the model price of a European call written on some transformed asset price with respect to all primary model parameters.

10.2 Regularization using prior views

Cont and Tankov [cf. 20, Section 13.3] warn that reformulating the calibration problem as a nonlinear least-squares problem does not resolve the uniqueness and stability issues. One way to obtain a unique solution in a stable manner is to introduce a regularization method. This is achieved by adding a penalization term to the least-squares criterion (10.1):

$$\varepsilon(\theta) = \sum_{i=1}^{N} |V(K_i, T_i; \theta) - V_i|^2 + \alpha D(Q(\theta), P), \tag{10.6}$$

where P an equivalent probability measure and α is the so-called *regularization parameter*. Here P corresponds to any prior believe in the model configuration, which is obtained from a historical time series analysis for example [cf. 20, Section 13.1]. The regularization parameter α characterizes the importance of the information contained in option prices relative to the prior knowledge of the Lévy measure.

The penalization function D is most commonly taken as the relative entropy of the pricing measure Q with respect to the prior measure P:

Definition 10.11. For two equivalent probability measures P, Q on $(\Omega, \mathcal{F})$ the *relative entropy of Q w.r.t. P* is

$$\mathcal{E}(Q, P) = E^Q[\ln \frac{dQ}{dP}] = E^P[\frac{dQ}{dP} \ln \frac{dQ}{dP}].$$

Relative entropy is often used as a measure of proximity of two equivalent probability measures (hence referred to as the *Kullback-Leibler distance*). Then (10.6) penalizes huge distances between the prior and the pricing model.

In exponential Lévy models, relative entropy is related to the characteristic triplets:

Proposition 10.12 ([20], Proposition 9.10)**.** *Let P and Q be equivalent measures on $(\Omega, \mathcal{F})$ generated by exponential Lévy models with Lévy triplets (γ^P, A^P, ν^P) and $(\gamma^Q, A^Q, \nu^Q,)$, where $A^Q = A^P = A$ with $A > 0$. The relative entropy $\mathcal{E}(Q, P)$ is then given by:*

$$\begin{aligned}\mathcal{E}(Q,P) &= \frac{T}{2A}\left(\gamma^Q - \gamma^P - \int_{|x|\leq 1} x(\nu^Q - \nu^P)(dx)\right)^2 \\ &+ T\int_{\mathbb{R}^d}\left(\frac{d\nu^Q}{d\nu^P}\ln\frac{d\nu^Q}{d\nu^P} + 1 - \frac{d\nu^Q}{d\nu^P}\right)\nu^P(dx).\end{aligned}$$

If P and Q correspond to arbitrage-free Lévy models, the relative entropy reduces to:

$$\begin{aligned}\mathcal{E}(Q,P) &= \frac{T}{2A}\left(\int_{\mathbb{R}^d}(e^x - 1)(\nu^Q - \nu^P)(dx)\right)^2 \\ &+ T\int_{\mathbb{R}^d}\left(\frac{d\nu^Q}{d\nu^P}\ln\frac{d\nu^Q}{d\nu^P} + 1 - \frac{d\nu^Q}{d\nu^P}\right)\nu^P(dx).\end{aligned}$$

Proposition 10.12 shows that the relative entropy between two equivalent (martingale) measures is very much a Lévy measure property. It is worth stressing that penalization by relative entropy reduces the number of free parameters, because equivalence of the prior and the optimal configuration sets the volatility parameters per se. This motivates the following functional interpretation [cf. 20, Section 9.6]:

Definition 10.13. Let P and Q be equivalent martingale measures on $(\Omega, \mathcal{F})$ generated by exponential Lévy models with Lévy triplets (γ^P, A^P, ν^P) and (γ^Q, A^Q, ν^Q), where $A^Q = A^P = A$ with $A > 0$. Then, for a given reference measure ν^P, expression (10.7) viewed as a function of ν^Q defines a positive (possibly infinite) functional on the set of Lévy measures $\mathcal{L}(\mathbb{R}^d)$:

$$H : \mathcal{L}(\mathbb{R}^d) \to [0, \infty], \quad \nu^Q \mapsto H(\nu^Q) = \mathcal{E}(Q(\gamma^Q, A, \nu^Q), P(\gamma^P, A, \nu^P)).$$

H is called the *relative entropy functional* and is a positive convex functional of ν^Q, which is equal to zero, only if $\nu^Q \equiv \nu^P$.

Convexity of the relative entropy functional then makes the difference between (10.1) and (10.6) in view of regularity. If α is large, the convexity properties of the entropy functional stabilize the solution, while for vanishing α we recover the non-regularized least squares problem. The right parameter value cannot be given a priori, because it is very much subjective to the actual data at hand and the level of error present in it.

Example 10.14. Let P and Q be equivalent measures on $(\Omega, \mathcal{F})$ generated by exponential Lévy models with Lévy triplets (γ^P, A^P, ν^P) and $(\gamma^Q, A^Q, \nu^Q,)$, where $A^Q = A^P = A$ with $A > 0$ and ν^Q, ν^P are discrete on $\mathbb{R}$ with

$$\nu^Q = \sum_{j=1}^N \nu_j^Q \delta_{x_j}, \quad \nu^P = \sum_{j=1}^N \nu_j^P \delta_{x_j}.$$

If we assume that P and Q correspond to arbitrage-free markets, then

$$\begin{aligned}
\mathcal{E}(Q,P) &= \frac{T}{2A}\sum_{j=1}^{N}(e^{x_j}-1)(\nu_j^Q-\nu_j^P) \\
&+ T\sum_{j=1}^{N}\left(\frac{\nu_j^Q}{\nu_j^P}\ln\frac{\nu_j^Q}{\nu_j^P}+1-\frac{\nu_j^Q}{\nu_j^P}\right)\nu_j^P.
\end{aligned} \tag{10.7}$$

Then Proposition 10.12 enables us to compute the regularized quadratic pricing error (10.6) explicitly in the context of Example 10.14.[5] Regarding the use of gradient based methods, we would very much like to have closed form representations of the derivatives of the penalization term with respect to the model parameters. These are available for discrete Lévy measures:

Proposition 10.15. *Let P and Q be equivalent measures on $(\Omega,\mathcal{F})$ generated by exponential Lévy models with Lévy triplets (γ^P, A^P, ν^P) and $(\gamma^Q, A^Q, \nu^Q,)$, where $A^Q = A^P = A$ with $A > 0$ and ν^Q, ν^P are discrete on $\mathbb{R}$ with*

$$\nu^Q = \sum_{j=1}^{N}\nu_j^Q\delta_{x_j} \quad \nu^P = \sum_{j=1}^{N}\nu_j^P\delta_{x_j}.$$

Then

$$\frac{\partial H(\nu^Q)}{\partial \nu_k^Q} = \frac{T}{A}(e^{x_k}-1)\sum_{j=1}^{N}(e^{x_j}-1)(\nu_j^Q-\nu_j^P)+T\ln\frac{\nu_k^Q}{\nu_k^P}.$$

Proof. The proof follows directly from differentiating (10.7) with respect to the weights $\nu_k^Q, k = 1,...,N$. □

Proposition 10.15 then rounds out the analytical means for numerical minimization of the penalized quadratic pricing error.

10.3 Foreign Exchange: Redundancy of a liquid market

The foreign exchange (or forex or fx) market is the most actively traded market in the world with more than a trillion worth of transactions each day. The trading is in currencies and

> (...) involves the simultaneous purchase of one currency while selling another currency. Currencies are traded in pairs, such as U.S. dollar/Euro or Japanese

[5] Cont and Luciano [19] exemplify the explicit penalty computation in the case of tempered stable Lévy measures.

> Yen/British Pound. (...) fx traders include governments, corporations and fund managers doing business with foreign countries, that need to exchange one currency for another, and speculators, who seek to profit from price movements in the markets.[6]

The most liquid currencies include the U.S. dollar (USD), the euro (EUR), the Japanese yen (JPY), the British pound, the Swiss franc, the Australian dollar and the Canadian dollar. Traders do business with fx products as diverse as spot and futures transactions or options. Exchange rate transactions take mostly place over-the-counter (OTC) with only a few regularized marketplaces existent.[7]

Our belief is that the fx market is liquid enough to expose not only the implied dynamics of the traded exchange rates but also the implied dependence structure between them. We have explained in the previous sections how the dependence structure (in terms of a Lévy copula function) influences the fair value of an option, that is written on the cross rate.

> An exchange rate transaction is termed a cross rate when the home country currency is not a party in the trade. For example, for a trader in the U.S., a cross rate would be euro/yen, or the euro against the Japanese yen.[8]

This suggests to back out the dependence parameters from cross rate related products in the context of a real connected foreign exchange market.

A real fx model Consider the case of Example 10.3 in a real world foreign exchange market, focusing on the JPY/USD rate and the JPY/EUR rates. Hence we play a trader who operates a Japanese business with international currency exposure in the U.S. Dollar and the Euro. Then the USD/EUR rate acts the counterpart of the cross rate.

Let us fix the 29th of June 2008 as the current date. Figure 10.1 shows the recent historical development [cf. 59] of the delta-connected U.S. dollar-euro-Japanese yen market. The current spot market rates [cf. 75] are 106.13 JPY = 1 USD, 167.61 JPY = 1 EUR and 1.5793 USD = 1 EUR.

For the time being, the (domestic) Japanese money market offers very low interest rates compared to the (foreign) U.S. and European currency accounts. Table 10.1 shows the international short-term interest rate curves [cf. 75].

(in %)	1 day	1 month	3 month	6 month
JPY	3.71	4.44	4.95	5.13
USD	2.25	2.45	3.00	3.15
EUR	0.25	0.63	0.85	0.94

Table 10.1: Interest rates offered in currencies JPY, USD and EUR.

[6] Quote taken from [17, Trading CME fx].

[7] A fraction of the worldwide volume is traded on exchanges like the Philadelphia Stock exchange [56] or the Chicago Mercantile Exchange [17].

[8] Quote taken from [17, Trading CME fx].

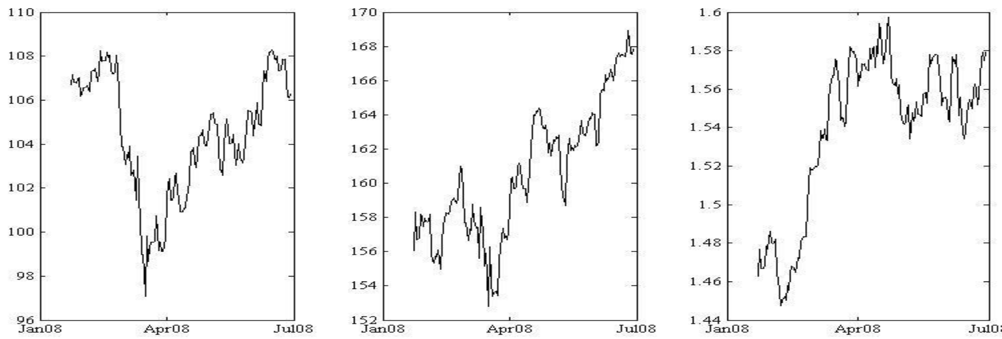

Figure 10.1: Time series of the JPY/USD rate (left), the JPY/EUR rate (center), and the USD/EUR rate (right).

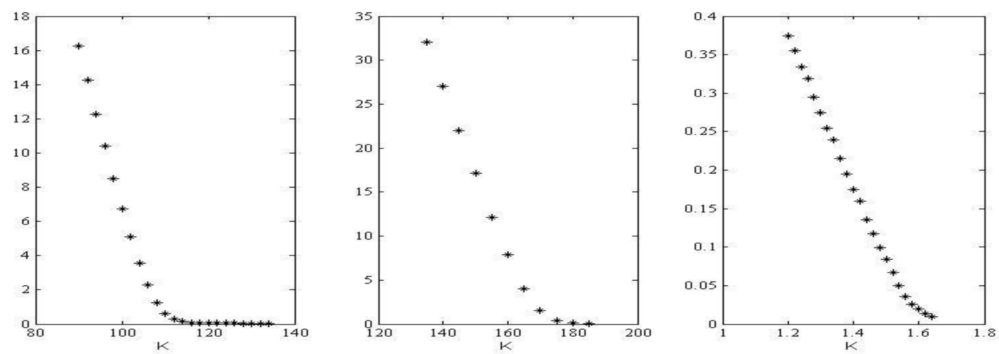

Figure 10.2: Market quotes for European style call options written on the JPY/USD exchange rate (left), the JPY/EUR exchange rate (center), and the USD/EUR cross rate (right).

We refer to market quotes [cf. 59] of OTC-traded European style call options written on the JPY/USD rate and the JPY/EUR rate with a 80 days maturity $T = 80/365$ for benchmark prices. Reference prices of cross rate based products are taken from market quotes [cf. 59] of OTC-trading in European style call options written on the USD/EUR rate with the same lifetime. Figure 10.2 illustrates the available price information.

This set of real data defines a 2-dimensional exponential Lévy model, featuring the exchange rate dynamics

$$\begin{aligned} R_t^1 &= 106.13e^{(0.0084-0.0286)t+X_t^2} \\ \text{and} \quad R_t^2 &= 167.61e^{(0.0084-0.0490)t+X_t^1}, \end{aligned}$$

where $\{X_t\}$ is a Lévy process on $\mathbb{R}^2$ with characteristic triplet (γ, A, ν). Here we interpolated the interest rate curves at the $T = 80/365$ years duration for the risk free rates

$r_h = .0084, r_f^1 = .0286$ and $r_f^2 = .0490$ offered in the domestic and the foreign markets. The corresponding cross rate dynamics are (consistently) given by

$$\widehat{R}_t = 1.5793 e^{(0.0286-0.0490)t+\widehat{X}_t},$$

where $\{\widehat{X}_t\}$, defined by $\widehat{X}_t = X_t^2 - X_t^1$, is a $\mathbb{R}$-valued Lévy process with triplet $(\widehat{\gamma}, \widehat{A}, \widehat{\nu})$. The benchmark price chains V^1, V^2, V^c are given according to the above quotes.

We assume the marginal Lévy measures ν_1, ν_2 to be discrete and coupled by a general Lévy Clayton copula F. Hence $(\gamma, A, \nu) = (\gamma(\theta), A(\theta), \nu(\theta))$ with model parameter vector $\theta = (\sigma_1, \sigma_2, \rho, \nu_1^1, ..., \nu_1^N, \nu_2^1, ..., \nu_2^N, \kappa, \eta)$. Our goal is to calibrate the characteristic triplet $(\gamma(\theta), A(\theta), \nu(\theta))$ (or θ, indeed) so as to keep the pricing error(s) as low as possible.[9]

Numerical implementation Given the available information, the non-regularized objective (error) function in our triangular foreign exchange model is

$$\varepsilon(\theta) = \sum_{j=1}^{23} |V^1(K_j;\theta_1) - V_j^1|^2 + \sum_{j=1}^{11} |V^2(K_j;\theta_2) - V_j^2|^2 + \sum_{j=1}^{23} |V^c(K_j;\theta) - V_j^c|^2, \tag{10.8}$$

where $\theta^1 = (\sigma_1, \nu_1^1, ..., \nu_1^N)$ and $\theta^2 = (\sigma_2, \nu_2^1, ..., \nu_2^N)$. Here (10.4), (10.5) and Propositions 10.6, 10.15 and 10.10 apply to the evaluation of (10.8) and its differentiation with respect to all parameters $\theta = (\sigma_1, \sigma_2, \rho, \nu_1^1, ..., \nu_1^N, \nu_2^1, ..., \nu_2^N, \kappa, \eta)$.

The *exact model calibration* $\theta_{EMC}^* \in \Theta$ is then given by the optimal argument

$$\theta_{EMC}^* = \arg\min_{\theta\in\Theta} \varepsilon(\theta).$$

The estimator is said to be exact, because optimization is over the complex 2-dimensional model. The exact calibration is computationally expensive, because it estimates $2N+2$ parameters simultaneously.

Another strategy is to split optimizations similar to the inference for margins method in maximum likelihood estimations. That is why we refer to the split approach as the *calibration of margins* method. Considering the first and the second term on the right-hand side of (10.8), we recognize the pricing errors

$$\varepsilon_1(\theta_1) = \sum_{j=1}^{23} |V^1(K_j;\theta_1) - V_j^1|^2 \quad \text{and} \quad \varepsilon_2(\theta_2) = \sum_{j=1}^{11} |V^2(K_j;\theta_2) - V_j^2|^2, \tag{10.9}$$

which correspond to the separate problems of calibrating θ_1, θ_2 to the individual benchmark price chains V^1 and V^2. This suggests to calibrate, in a first step, the parameters θ_i associated to the margin $\{R_t^i\}$ to the price chain V^i in the following sense:

$$\hat{\theta}_i = \arg\min_{\theta_i\in\Theta_i} \varepsilon_i(\theta_i).$$

[9] Salmon and Schleicher [68] proceed in a similar way using bivariate ordinary copulas in the style of Cherubini and Luciano [15] to preserve the (as they called it) triangular no-arbitrage relation.

In a second step, the term on the right-hand side of (10.8), which penalizes price differences to the cross rate benchmark, is minimized. This involves the model prices $V^c(K_i;\theta) = V^c(K_i;\theta_c,\theta_1,\theta_2)$, where $\theta_c = (\kappa,\eta)$ the Clayton Lévy copula parameters. Here we employ the previously calibrated (marginal) parameters $\hat{\theta}_i = (\hat{\sigma}_i, \hat{\nu}_i^1, ..., \hat{\nu}_i^N), i = 1,2$ to evaluate the dependence error

$$(10.10) \qquad \varepsilon_c(\theta_c) = \sum_{j=1}^{23} |V^c(K_j;\theta_c,\hat{\theta}_1,\hat{\theta}_2) - V_j^c|^2.$$

From there, the optimal dependence structure is parameterized by

$$\hat{\theta}_c = \arg\min_{\theta_c\in\Theta_c} \varepsilon_c(\theta_c).$$

Then the calibration for margins method gives the best model parametrization

$$\hat{\theta}_{CFM} = (\hat{\theta}_c,\hat{\theta}_1,\hat{\theta}_2).$$

The reduced calibration problems (10.9) remain particularly ill-posed, even though they are moderated to marginal perspectives.

Remark 10.16. In Section 10.2 we have discussed how to regularize the uniqueness and stability issues of the calibration problem by prior views on the pricing measure (or the model parameters, in fact).

Regularization by relative entropy changes the (marginal) objective functions (10.9) into

$$(10.11) \qquad \varepsilon_j(\theta_i) = \sum_j |V(K_j;\theta_i) - V_j^i|^2 + \alpha\mathcal{E}(Q(\theta_i), P(\theta_i^P)),$$

where α the regularization parameter and θ_i^P the prior belief in the model parameters. We do not use historical price data for choosing the prior in this study. Instead we calibrate a simple (few parameter) Merton jump-diffusion model to the benchmark prices to obtain an estimate of the volatility and a candidate ν^P for the Lévy measure [cf. 20, Section 13.4]. With a view to Proposition 10.12, the prior ν^P must also be discretized on the same grid, using, for example, the formula

$$\nu_i^P = \int_{x_i-\Delta x}^{x_i+\Delta x} \nu^P(dx)$$

for the points $x_1, ..., x_{N-1}$ and integrating up to plus infinity or minus infinity for the points x_0, x_N. Then the calibrated Lévy measure will be equivalent to the discretized Lévy measure [cf. 20, Section 13.4] and not to the Merton prior. But if the grid is sufficiently fine, the approximation to the continuous prior is quite good.

The regularization parameter α is chosen so as to achieve an effective trade-off between precision and stability of the calibration. One possible solution [cf. 20, Section 13.3] is to increase the importance of regularization (α, so to say) so long as to stay approximately within model precision. To do so, we first minimize the quadratic pricing error (10.11) with zero regularization $\alpha = 0$.[10] This produces the least squares error ε^0, which is interpreted as the distance of the market to the model and gives an a priori error level that one can not really hope to improve upon while staying in the same class of models. The optimal α is then defined as the maximal value to stay within the same order of magnitude as ε^0:

$$\hat{\alpha} = \sup_{\alpha}\{\varepsilon_i(\hat{\theta}_i(\alpha)) \leq \delta\varepsilon^0\}, \delta > 1, \tag{10.12}$$

where $\theta_i(\alpha)$ is the optimal argument resulting from the minimization of (10.11) with regularization parameter α. Here δ indicates the sacrifice of accuracy for the benefit of stability. Typical values for δ so as to produce satisfactory results are in the interval $1 < \delta < 1.5$.

The following steps summarize the final calibration method as implemented in the numerical examples of the next section:[11]

(1) For $i = 1, 2$, fit a Merton jump-diffusion model to the benchmark price chains V^i and discritize it in order to get a prior parametrization (σ_i, ν_i^P).

(2) For $i = 1, 2$, fix $\hat{\sigma}_i = \sigma_i$ and run the optimization routine on (10.11) with $\alpha_i = 0$ to get an estimate of the model error.

(3) For $i = 1, 2$, choose a trade-off δ between precision and stability, and compute the optimal regularization parameter $\hat{\alpha}_i$ as to (10.12) by running the optimization routine several times.

(4) For $i = 1, 2$, solve the minimization problem (10.11) with optimal $\hat{\alpha}_i$.

(5) Fix $\hat{\sigma}_1, \hat{\nu}_1^1, ..., \hat{\nu}_1^N, \hat{\sigma}_2, \hat{\nu}_2^1, ..., \hat{\nu}_2^N$ and solve the variational problem (10.10).

We have only discussed the penalization of large prior discrepancy in the calibration of the marginal markets, leaving aside uniqueness and stability issues of the calibration of the Lévy copula. This will be resumed with regard to the empirical results in the next section.

Empirical results We apply the regularized calibration procedure to the delta connected fx market and examine the implied Lévy measures and the implied Lévy copulas so-obtained.

We assess the marginal results first. Table 10.2 shows the Merton parameters of the candidate Lévy processes, that were fitted to the references options in a prior calibration (we add the prior calibration of the cross rate products for later reference).

- The marginal Lévy measures are non-trivial; the market anticipates some jumps of the foreign exchange rates.

[10] Using regularization with $\alpha = 0$ is different from omitting the regularization term entirely in that the volatility parameter is kept fixed.

[11] Cont and Luciano [19] enclose a testing of the original algorithm over a range of model examples in the univariate case.

	σ	λ	μ	δ	ε
ν_1^P	0.0759	0.5158	-0.1679	0.1674	1.1738e-6
ν_2^P	0.0732	0.2457	-0.2231	0.2757	2.5419e-6
ν_3^P	0.0826	0.1943	-0.0959	0.3149	6.4952e-6

Table 10.2: Parameters of the Merton-prior as fitted to market implied volatilities.

- The mode is located left to the origin; the market fears a downgrading of the Japanese yen.
- The jump intensity is higher for the JPY/USD rate, while jump diffusion is higher for the JPY/EUR rate; the market anticipates a more frequent but moderate downgrading of Japanese yen compared to U.S. dollar, and a seldom but potentially heavy drop of the JPY/EUR rate.
- The prior implied volatilities are lower than the at-the-money market implied volatilities; the existence of jumps alone produces additional volatility.

Remark 10.17. Regarding Proposition 10.12, the diffusion parameters are from now on kept fixed at the level of the prior because of the equivalence constraints for the regularizing and the regularized measure.

The errors made in reproducing the benchmark prices by Merton's jump diffusion are $\varepsilon_1^P = 1.1738 \cdot 10^{-6}$ and $\varepsilon_2^P = 2.5419 \cdot 10^{-6}$. The model distances to the marginal markets are $\varepsilon_1^0 = 1.0664 \cdot 10^{-6}$ and $\varepsilon_2^0 = 2.0530 \cdot 10^{-6}$, which is found by running the unregularized calibration method on a grid with step size $\Delta x = 2\pi(1024-1)/(600 * 1024) \simeq 0.0105$ and 256 points[12]. The optimal trade-off between pricing accuracy and stability is found by running the optimization repeatedly. We choose a tolerance $\delta = 1.01$ (meaning, that we allow for a 1%-loss of accuracy) and halve the regularization parameter α in each run till the error constraint is met. The outcome is illustrated in Figure 10.3, which gives reason to fixing the relative importance parameters at $\alpha_1 = 6 \cdot 10^{-6}$ and $\alpha_2 = 5 \cdot 10^{-7}$. Figure 10.4 shows the final outcome of marginal calibrations to the JPY/USD rate and the JPY/EUR rate with optimal regularization by a Merton prior.[13]

- The implied market volatilities are adequately reproduced; the fit to the market quotes is better with the non-parametric approach as with the prior fitting.
- The left tails of the Lévy measures are much heavier than the right tails; the jump behavior is strongly asymmetric.
- The Lévy measure calibrated to the JPY/USD rate and its prior are almost coincident; the Merton model is rich enough to capture the rate's market implied characteristic.

[12]The grid size here is due to the constraints in using FFT methods.

[13]Figure 10.4 can be compared to the calibration results in Tankov [cf. 77, Figures 3.8 to 3.11].

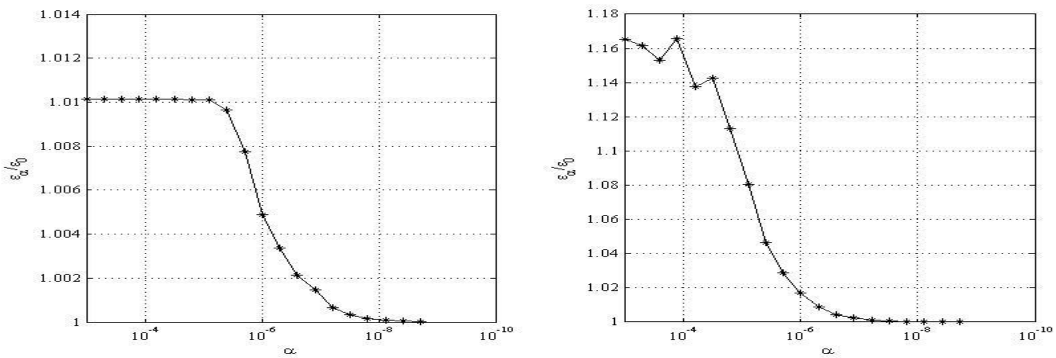

Figure 10.3: Trade off between accuracy and stability of the regularized calibration problem for the JPY/USD rate (left) and JPY/EUR rate (right).

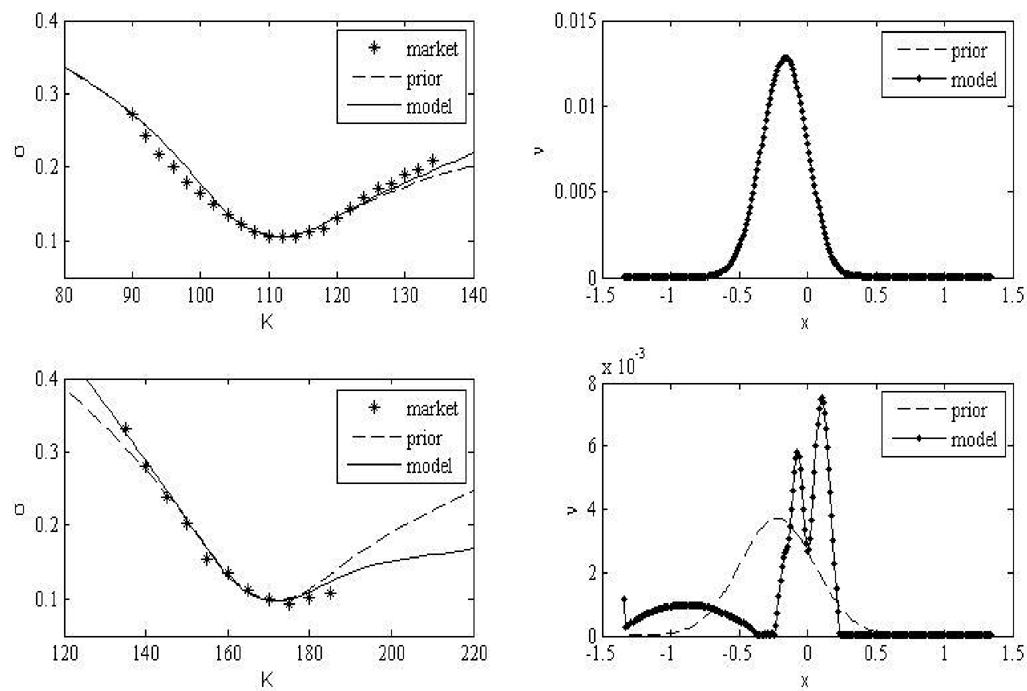

Figure 10.4: Calibrated marginal Lévy measures (right) and corresponding implied volatilities (left) in the case of JPY/USD rate (upper) and JPY/EUR rate (lower).

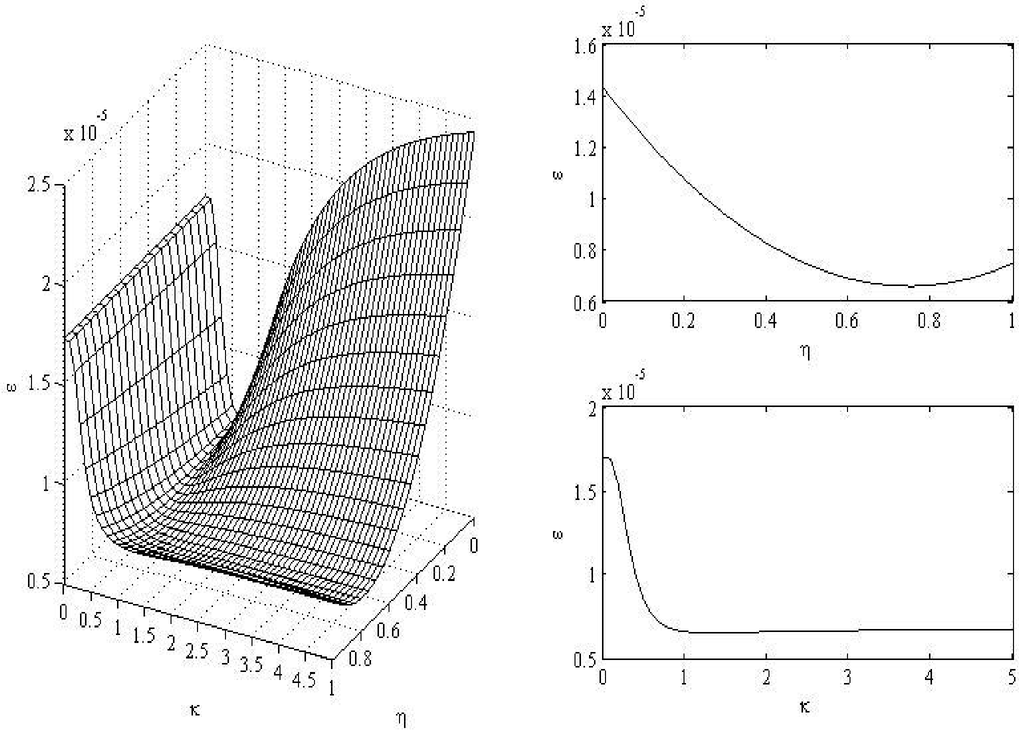

Figure 10.5: Surface of the pricing error at the optimal $\rho = 0.3055$ (left), error function at $\rho = 0.3055, \kappa = 1.4791$ (upper right), and error function at $\rho = 0.3055, \eta = 0.7281$ (lower right).

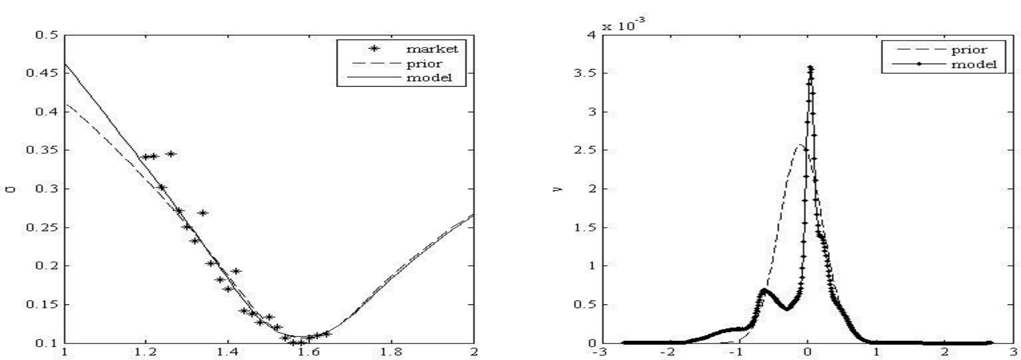

Figure 10.6: Calibrated marginal Lévy measures (right) and corresponding implied volatilities (left) in the case of the USD/EUR cross rate.

- The Lévy measure calibrated to the JPY/EUR rate is trimodal[14] with two alternate modes sharply nearby the origin and a low mode in the left tail; the Merton model is incapable of reproducing the diverging market perceptions.

Then we fit the model's dependence parameters so as to reproduce the quotes corresponding to the EUR/USD cross rate based on the previously calibrated margins. In the least-squares sense, the best fit to market quotes is achieved by $\rho = 0.3055$, $\eta = 0.7281$ and $\kappa = 1.4791$. Hence the market prices some positive normal dependence, a significant sign dependence of sudden jumps in the rates and a moderate size dependence of simultaneous bursts.

While the introduction of a convex prior discrepancy term has worked as a stabilizer for margin calibrations, the calibration of the dependence parameters to option quotes by minimal least squares is found sufficiently stable on its own. Figure 10.5 shows the surface (left graph) of the least square error (10.10) over a range of Lévy copula parameters at $\rho = 0.3055$. The error development over a changing sign (upper right graph) or size dependence (lower right graph) is plotted for additional visualization. Here every other Lévy copula parameter is kept fixed at its optimal value.

- The error is a smooth function of jump sign and jump size dependence in the region at hand.

- Around its minimal value, the error is a strictly convex function of the jump sign dependence η but exhibits a flat yet convex behavior in direction of the jump size dependence κ.

Figure 10.6 shows the fitted conjugate Lévy measure and how accurate option prices are reproduced. For reasons of validity, we have redone a prior fitting of the EUR/USD related option prices to the same effect as before.

- The implied market volatilities are adequately reproduced; the calibration accuracy of the conjugate Lévy measure does not rank behind a direct approach.

- The calibrated Lévy measure is multimodal with one spike precisely at the origin, one clear mode in the left tail and some plateau building on the line.

10.4 Summary

In this chapter we have solved the inverse problem of calibrating the characteristics of an exponential Lévy model to market quotes. Our main contribution here is to extract the dependence structure, that is implied in a liquid two rate foreign exchange market.

First, we formulated the general calibration problem in the style of a least-squares approximation to observable option prices. This was subsequently applied to a discrete

[14] the extremal mode in the very left tail is due to numerical procedure and was already encountered in Cont and Tankov [cf. 20, Section 13.3] for the case of Dax option prices.

2-rate foreign exchange model, that features call options written on the primary exchange rates and the cross rate.

It proved to be crucial for the use of gradient based optimization routines that we are able to determine the derivatives of the approximation error with respect to the model parameters. Our contribution here is that we differentiated the semi-analytical representation of the fair value of a call option written on a primary exchange rate with respect to the diffusion and the Lévy measure parameters in the event of dividend payments. Moreover, we determined the derivatives of the fair model value of a call option written on the cross rate with respect to the Lévy copula parameters.

On account of uniqueness and stability concerns, we included a regularization term in the objective error function, that penalizes large entropy distances of the candidate parameters from some prior belief. We deduced the derivatives of the penalty term with respect to all model parameters involved.

Having described the real delta-connected USD-EUR-JPY exchange rates market in short, we incorporated the previous results into a gradient based optimization scheme. A split strategy to calibrate the marginal parameters to the exchange rates related quotes and the dependence parameters to the cross rate related quotes proved to be useful.

We implemented the margin calibrations first. The volatilities implied in European call options written on the USDJPY and the EURJPY exchange rate were accurately reproduced. Specifically, the calibrated Lévy measures exposed the markets worries about negative jumps of both exchange rates, although the jump behavior of the EURJPY exchange rate seemed more erratic.

Next, we implemented the dependence calibration. The resulting fit to the volatilities implied in European call options written on the EURUSD cross rate turned out to be fairly adequate. Specifically, we recognized that a (in every respect) moderate dependence structure produces the best fit to the cross rate data.

We also argued for the stability of the dependence calibration (as opposed to the margin calibration) graphically. This validated the use of penalty terms in the marginal error functions only.

Appendix A

Ordinary copula related topics

A.1 Archimedean copula functions

We study briefly the pre-eminent class of Archimedean copulas and present the Clayton copula for reference.

Definition A.1. A function f is said to be *completely monotonic* on an interval $[a, b]$ if it satisfies

$$(-1)^k \frac{d^k}{dt^k} f(t) \geq 0, \quad k \in \mathbb{N}, t \in (a, b),$$

i.e. the derivatives alternate in sign.

Let $\phi : [0, 1] \to [0, \infty]$ be a continuous and strictly decreasing function such that $\phi(1) = 0$ and $\phi(0) \leq \infty$. Then the inverse function ϕ^{-1} exists on $[0, \phi(0)]$, and is continuous and strictly decreasing. We extend the notion of the inverse function to the domain $[0, \infty]$.

Definition A.2. Suppose $\phi : [0, 1] \to [0, \infty]$ is continuous and strictly decreasing with $\phi(1) = 0$ and $\phi(0) \leq \infty$. We define a *pseudo-inverse* of ϕ with domain $[0, \infty]$ by

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t), & 0 \leq t \leq \phi(0); \\ 0, & \phi(0) \leq t \leq \infty. \end{cases} \tag{A.1}$$

One can show [cf. 57, Section 4.6] that if the pseudo-inverse $\phi^{[-1]}$ of ϕ is completely monotonic, then $\phi(0) = \infty$, and the usual inverse ϕ^{-1} exists on $[0, \infty]$.

Theorem A.3 ([57], Theorem 4.6.2)**.** *Let $\phi : [0, 1] \to [0, \infty]$ be continuous and strictly decreasing with $\phi(1) = 0$ and $\phi(0) = \infty$, and let ϕ^{-1} denote the inverse of ϕ. Then*

$$C(u_1, ..., u_d) = \phi^{-1}(\phi(u_1) + ... + \phi(u_d)) \tag{A.2}$$

is a copula if and only if ϕ^{-1} is completely monotonic on $[0, \infty)$.

A function ϕ, such that ϕ^{-1} is completely monotonic, is referred to as a *generator* and the copula functions generated as to (A.2) are classified as *Archimedean*[1]. In fact, generators ϕ fulfilling $\phi(0) = \infty$ are termed *strict generators*. We relate strict generators to the Laplace transform of some univariate distribution in the following.

Lemma A.4 (cf. [43], Appendix 1)**.** *A function ψ on $[0, \infty)$ is the Laplace transform of a distribution P, if and only if ψ is completely monotonic and $\psi(0) = 1$.*

As a consequence, any Laplace transform ψ is positive on $[0, \infty)$ and $\psi(\infty) = 0$. Then the inverse ψ^{-1} of Laplace transforms ψ (that exists on $[0, 1]$) determine generators [2] of Archimedean copulas.

Example A.5. Let $X \sim GAM(\frac{1}{\theta}, 1)$ be a gamma distributed random variable on $\mathbb{R}$ with $\theta > 0$. Then the Laplace transform ψ^X of X is

$$\psi(t) = (1 + t)^{-\frac{1}{\theta}}, \quad t > 0.$$

Hence $\phi(t) = \psi^{-1}(t) = t^{-\theta} - 1$ is a strict generator and ϕ produces the Archimedean d-copula

$$C(u_1, ..., u_d; \theta) = (u_1^{-\theta} + ... + u_d^{-\theta} - d + 1)^{-\frac{1}{\theta}}. \tag{A.3}$$

The function ϕ is called the *(strict) Clayton generator* and copula (A.3) is called the *Clayton* [3] *copula.*

In the limit $\theta \to 0$ we arrive at the independence copula. Conversely, in the limit $\theta \to \infty$ the Clayton copula tends to the maximum copula. The Fréchet lower bound is unattainable in our setting for all $d \geq 2$.

In Figure A.1 we plotted the generator (left) and the contour lines of the Clayton copula density over a range of dependence parameters θ.

- Depending on θ, the Clayton generator is a slow or fast decreasing function
- For $\theta = 0.5$ and $\theta = 6$, the density is highest in the lower left corner; the Clayton copula induces asymmetric and positive dependence only.[4]
- The higher the θ the denser is the mass in the corner; θ is responsible for the degree of (tail) dependence.[5]

[1] The term *Archimedean* is due to Ling [50] who proofed that for an Archimedean 2-copula C and any $u_1, u_2 \in [0, 1]$, there exists a positive integer $n \in \mathbb{N}$ such that $u_1^{(n)} < u_2$, where $u_1^{(n)}$ the C-powers defined recursively by $u_1^{(1)} = u_1$ and $u_1^{(n+1)} = C(u_1, u_1^{(n)})$.

[2] Some authors [e.g. 43] use the alternative notation $C(u_1, ..., u_d) = \psi(\psi^{-1}(u_1) + ... + \psi^{-1}(u_d))$ for defining Archimedean copula functions.

[3] Clayton [18] introduced this type of dependence function that was also called the Cook and Johnson copula [37] or the Pareto copula [42].

[4] Archimedean copulas generated by any strict ϕ have always positive dependence structures [see 43, for an extension to negatively dependent margins].

[5] Joe [43] and Nelson [57] give dependence orderings of Archimedean (one-parameter) families.

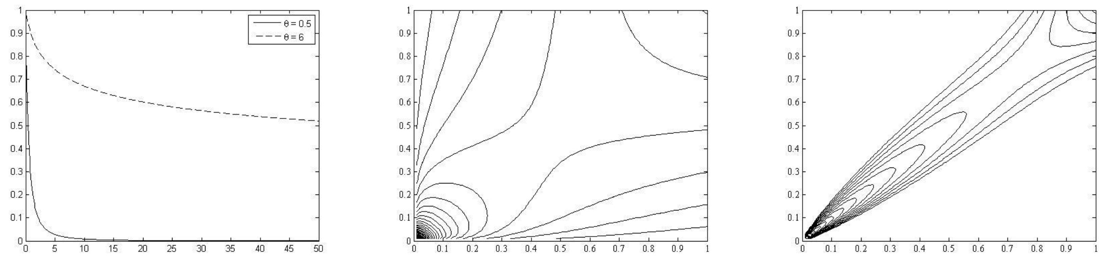

Figure A.1: Generators (left) and contours of the Clayton copula density using $\theta = 0.5$ (center), and $\theta = 6$ (right).

Thus the Clayton copula contrasts the elliptical family regarding symmetry and sign of the attainable dependence structures. It is worth stating that the Clayton copula family is uniparametric. Hence the copula function uses a single parameter θ whatever the dimension.

A.2 Conditional copula functions

Following Patton [63], there exists a natural extension of the concept of an ordinary copula to a conditional copula.

Definition A.6. For any σ-algebra $\mathcal{G} \subset \mathcal{F}$, a conditional d-dimensional copula function C (or conditional d-copula or conditional copula) given $\mathcal{G}$ is a mapping of the form $C : [0,1]^d \to [0,1]$ such that

(1) $C(\cdot, |\mathcal{G})$ is grounded

(2) $C(\cdot, |\mathcal{G})$ is d-increasing and

(3) $C(1, ..., 1, u_i, 1, ..., 1|\mathcal{G}) = u_i$ for all $i \in \{1, ..., d\}, u_i \in [0,1]$

Theorem 2.2 can be extended to incorporate for conditional distributions:

Theorem A.7 (cf. [63], Theorem 3)**.** *Let F be a joint cdf with margins $F_1, ..., F_d$ and $\mathcal{F}$ some conditioning set. Then there exists a conditional d-copula C such that, for all $x_1, ..., x_d \in \mathbb{R}_\infty$,*

$$F(x_1, ..., x_d|\mathcal{F}) = C(F_1(x_1|\mathcal{F}), ..., F_d(x_d|\mathcal{F})|\mathcal{F}). \tag{A.4}$$

If the margins are continuous, then C is unique. Conversely, if C is a conditional d-copula and $F_1, ..., F_d$ are univariate conditional cdf's, then the function F defined by A.4 is a joint conditional cdf with margins $F_1, ..., F_d$.

Referring to Patton [63], the reader may be confirmed that conditional copulas are handled much the same way as ordinary copulas. Then the results on the static case can be readily applied to the dynamic perspective.

A.3 Elliptical distributions

Elliptical distributions are heavily used in this work. We give a very short introduction.

Definition A.8. A random variable X on $\mathbb{R}^d$ is said to have a *spherically symmetric distribution* (or simply spherical distribution) if for every orthogonal matrix $\Gamma \in \mathbb{R}^{d\times d}$

$$\Gamma X \stackrel{d}{=} X.$$

Definition A.8 shows that spherically distributed random variables are distributionally invariant under rotations. That is to say they have uncorrelated (not necessarily independent) components and identical, symmetric marginal distributions. The following result characterizes a spherical distribution by means of its characteristic functions:

Theorem A.9 (cf. [31], Theorem 3.19)**.** *A random variable X on $\mathbb{R}^d$ has a spherical distribution, if and only if its characteristic function $\varphi(t)$ satisfies one of the following equivalent conditions:*

- $\varphi(\Gamma' t) = \varphi(t)$ *for any orthogonal matrix* $\Gamma \in \mathbb{R}^{d\times d}$
- *there exists a function ϕ of a scalar variable such that $\varphi(t) = \phi(t.t)$.*

If existent, the function ϕ is referred to as the *characteristic generator* of the spherical distribution. It is convenient to write $X \sim SPH(\phi)$ for a spherically distributed random variable X on $\mathbb{R}^d$ with characteristic generator ϕ. The pdf of X, if it exists, must be of the form $f^X(x) = g(x.x)$ for some nonnegative function g of a scalar variable. In this case, g is called the *pdf generator*. Then we may write $X \sim SPH(g)$ instead of $X \sim SPH(\phi)$ whenever the distribution is specified by its pdf.

Example A.10 (cf. [31], Example 3.20)**.** Let $X \sim MVN(\mathbf{0}, \mathbf{1})$. Then the characteristic function φ^X of X is

$$\varphi^X(u) = e^{-\frac{1}{2}(u.u)}.$$

From Theorem A.9 it follows that X has a spherical distribution with characteristic generator $\phi(t) = e^{-\frac{1}{2}t}$. The pdf f^X of X is given by

$$f^X(x) = (2\pi)^{d/2} e^{-\frac{1}{2}(x.x)},$$

and can thus be written as $f^X(x) = g(x.x)$ with scalar function $g(t) = (2\pi)^{d/2} e^{-\frac{1}{2}t}$.

Example A.11. Let $X \sim MVT(\mathbf{0}, \mathbf{1}, \nu)$. Then the pdf f^X of X is

$$f^X(x) = \frac{\Gamma(\frac{1}{2}(\nu + d))}{\Gamma(\frac{1}{2}\nu)(\pi\nu)^{d/2}} \left(1 + \frac{x.x}{\nu}\right)^{-(\nu+d)/2}$$

By Theorem A.9, X has a spherical distribution with pdf generator

$$g(t) = \frac{\Gamma(\frac{1}{2}(\nu + d))}{\Gamma(\frac{1}{2}\nu)(\pi\nu)^{d/2}} \left(1 + \frac{t}{\nu}\right)^{-(\nu+d)/2}.$$

The class of elliptical distributions results from affine transformations of spherical distributions:

Definition A.12. A random variable X on $\mathbb{R}^d$ is said to have an *elliptically symmetric distribution* (or simply elliptical distribution) with parameters $\mu \in \mathbb{R}^d$ and $\Sigma \in \mathbb{R}^{d\times d}$ (and ϕ) if

$$X = \mu + A'Y, \quad Y \sim SPH(\phi),$$

where $A \in \mathbb{R}^{k\times d}$ such that $A'A = \Sigma$ with $rank(\Sigma) = k$.

We write $X \sim ELL(\mu, \Sigma, \phi)$ for elliptically distributed random variables X on $\mathbb{R}^d$ with parameters μ and Σ, and characteristic generator ϕ. In the course of this work, we only consider elliptically distributed random variables $X \sim ELL(\mu, \Sigma, \phi)$ whose parameter Σ is of full rank d. Given that Y possesses a pdf generator g, the pdf of X exists and is of the form [cf. 31, Section 3.3.2]

$$f^X(x) = |\Sigma^{-1/2}| g((x-\mu).\Sigma^{-1}(x-\mu)). \tag{A.5}$$

In this case we shall sometimes use the notation $ELL(\mu, \Sigma, g)$ instead of $ELL(\mu, \Sigma, \phi)$.

Example A.13. Let $X \sim MVN(\mu, \Sigma)$. Then we clearly have

$$X = \mu + \Sigma^{1/2}Y, \quad Y \sim MVN(\mathbf{0}, \mathbf{1}).$$

Regarding (A.5), the pdf of X is given by

$$f^X(x) = (2\pi)^{-d/2}|\Sigma|^{-1/2} e^{-\frac{1}{2}((x-\mu).\Sigma^{-1}(x-\mu))}.$$

Example A.14. Let $X \sim MVT(\mu, \Sigma, \nu)$. Then we have

$$X = \mu + \Sigma^{1/2}Y, \quad Y \sim MVT(\mathbf{0}, \mathbf{1}, \nu).$$

Regarding (A.5), the pdf of X is given by

$$f^X(x) = \frac{\Gamma(\frac{1}{2}(\nu + d))}{|\Sigma|^{1/2}\Gamma(\frac{1}{2}\nu)(\pi\nu)^{d/2}} \left(1 + \frac{(x-\mu).\Sigma^{-1}(x-\mu))}{\nu}\right)^{-(\nu+d)/2}.$$

A.4 Gaussian hypergeometric function

Proposition 4.19 involves the Gaussian hypergeometric function [cf. 2, Chapter 15]. This is analyzed (in part graphically) in order to support the proof of Proposition 4.19.

Definition A.15. For $a, b, c \in \mathbb{R}$, the *Gaussian hypergeometric series* $h(a, b, c; z)$ is defined as

$$h(a, b, c; z) = \sum_{k=1}^{\infty} \frac{\Gamma(k+a)}{\Gamma(a)} \frac{\Gamma(k+b)}{\Gamma(b)} \frac{\Gamma(c)}{\Gamma(k+c)} \frac{z^k}{k!}.$$

Its circle of convergence is the unit circle $|z| < 1$.

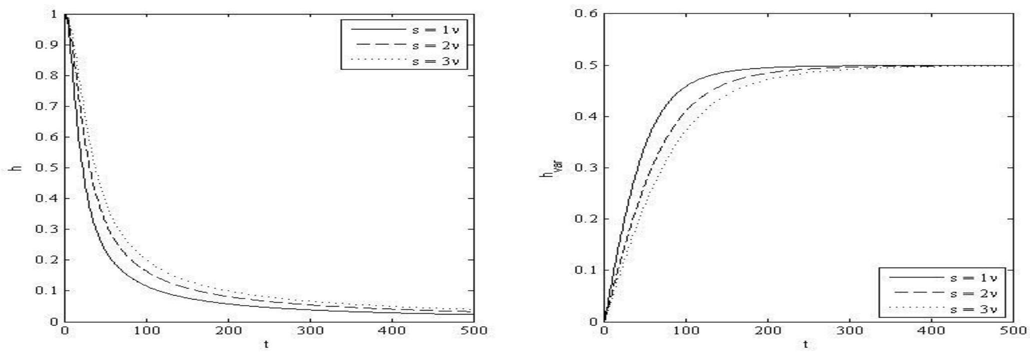

Figure A.2: Line graphs of the Gaussian hypergeometric function and its variant including the Pearson type VII characteristic coefficient.

The Gaussian hypergeometric function $h(a, b, c; z)$ with $a = 1/2, b = N - (d-i)/2, c = 3/2$ may be used to rewrite the integral (we found this representation of the integral using the MatLab Symbolic Toolbox)

$$\int \left(1 + \frac{y^2}{s+\nu}\right)^{-(N-\frac{d-i}{2})} dy = yh\left(\frac{1}{2}, N - \frac{d-i}{2}, \frac{3}{2}, \frac{-y^2}{\nu+s}\right),$$

which comes up in the proof of Proposition 4.19. Another term, that steams from the multivariate Pearsom type VII density is involved there. Then we are interested in the properties of the mapping

$$\text{(A.6)} \qquad t \mapsto \frac{\Gamma(N - \frac{d-i}{2})\left(1 + \frac{s}{\nu}\right)^{-1/2} t}{(\pi\nu)^{1/2}\Gamma(N - \frac{d-i}{2} - 1/2)} h\left(\frac{1}{2}, N - \frac{d-i}{2}, \frac{3}{2}, \frac{-t^2}{\nu+s}\right).$$

For graphical support of the proof of Proposition 4.19, Figure A.2 plots the Gaussian hypergeometric function $h(\frac{1}{2}, N - \frac{d-i}{2}, \frac{3}{2}, \frac{-t^2}{\nu+s})$ and its variant (A.6) over t using $N = \frac{1}{2}(d+\nu)$.

- The Gaussian hypergeometric function $h(\frac{1}{2}, N - \frac{d-i}{2}, \frac{3}{2}, \frac{-t^2}{\nu+s})$, seen as a function of t, is a convex function tending to 0 for $|t| \to \infty$.
- The variant function (A.6) is a concave function with limit value 1/2 for $|t| \to \infty$.

This validates the proof of Proposition 4.19.

Appendix B

Lévy copula related topics

B.1 Time inhomogeneous Lévy process

Time inhomogeneous Lévy processes are obtained from relaxing the condition of stationary increments. They are special examples of additive processes.

Definition B.1. A stochastic process $\{X_t\}$ on $\mathbb{R}^d$ with $X_0 = \mathbf{0}$ is an *additive process* if the following conditions are satisfied:

(1) For any choice of $n \geq 1$ and $0 \leq t_0 < t_1 < ... < t_n \leq T$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, ..., X_{t_n} - X_{t_{n-1}}$ are independent (independent increments property).

(2) For every $t \in [0, T]$ and $\varepsilon > 0$ it holds $\lim_{s \to t} P[|X_s - X_t| > \varepsilon] = 0$ (stochastic continuity property).

(3) There is $\Omega_0 \in \mathcal{F}$ with $P[\Omega_0] = 1$ such that, for every $\omega \in \Omega_0$, $X_t(\omega)$ is right-continuous in $t \in [0, T]$ and has left limits in $t \in [0, T]$.

In an obvious way, Lévy processes are additive. Additive processes admit a generalization of the Levy-Khintchin representation:

Theorem B.2 ([20], Theorem 14.1)**.** *Let $\{X_t\}$ be an additive process on $\mathbb{R}^d$. Then X_t has infinitely divisible distribution for all t. The law of $\{X_t\}$ is uniquely determined by its* spot characteristics $(\gamma_t, A_t, \nu_t)_{t \geq 0}$*:*

$$E[e^{iz.X_t}] = e^{\psi_t(z)}, \quad z \in \mathbb{R}^d$$

$$\psi_t(z) = -\frac{1}{2} z.A_t z + i\gamma_t.z + \int_{\mathbb{R}^d} (e^{iz.x} - 1 - iz.x\mathbf{1}_{|x| \leq 1})\nu_t(dx). \tag{B.1}$$

The spot characteristics satisfy the following conditions:

(1) For all t, A_t is a positive definite $d \times d$-matrix and ν_t is a positive Radon measure on $\mathbb{R}^d$ satisfying $\nu_t(\{0\}) = 0$ and $\int_{\mathbb{R}^d} (|x|^2 \wedge 1)\nu_t(dx) < \infty$.

(2) $\gamma_0 = 0, A_0 = 0, \nu_0 = 0$ *and for all* s, t*, such that* $s \leq t$*,* $A_t - A_s$ *is a positive definite matrix and* $\nu_t(B) \geq \nu_s(B)$ *for all measurable sets* $B \in \mathcal{B}(\mathbb{R}^d)$ *(positiveness property).*

(3) if $s \to t$ *then* $A_s \to A_t, \gamma_s \to \gamma_t$ *and* $\nu_s(B) \to \nu_t(B)$ *for all measurable sets* $B \in \mathcal{B}(\mathbb{R}^d)$ *such that* $B \subset \{x : |x| \geq \varepsilon\}$ *for some* $\varepsilon > 0$ *(continuity property).*

Conversely, for a family of triplets (γ_t, A_t, ν_t) *satisfying the above conditions there exists an additive process* $\{X_t\}$ *with* $(\gamma_t, A_t, \nu_t)_{t\geq 0}$ *as spot characteristics.*

A fundamental additive process encountered for in many finance applications is the Brownian motion with time dependent volatility [cf. 20, Example 14.1.]. For $\{B_t\}$ a standard Brownian motion on $\mathbb{R}$, $\sigma(t) : \mathbb{R}_+ \to \mathbb{R}_+$ a measurable function such that $\int_0^t \sigma^2(s)ds < \infty$ for all $t \in [0, T]$ and $b(t) : \mathbb{R}_+ \to \mathbb{R}$ a continuous function, the process

$$X_t = b(t) + \int_0^t \sigma(s) dB_s$$

is an additive process. The Brownian motion with time dependent drift and volatility is still pathwise continuous. A simple time inhomogeneity can also be incorporated into the jump part of a Lévy process:

Lemma B.3 (cf. [20], Section 14.1)**.** *Let* $(\gamma_t, A_t, \mu_t)_{t\geq 0}$ *be defined by*

(1) $A_t = \int_0^t \sigma^2(s)ds$*, where* $\sigma : [0, T] \to \mathbb{R}^{d\times n}$ *a matrix valued function such that* $\sigma(t)$ *is symmetric and verifies* $\int_0^T \sigma^2(t)dt < \infty$*.*

(2) $\mu_t(B) = \int_0^t \nu_s(B)ds, \quad \forall B \in \mathcal{B}(\mathbb{R}^d)$*, where* $\{\nu_t\}_{t\in[0,T]}$ *a family of Lévy measures verifying* $\int_0^T dt \int_{\mathbb{R}^d} (1 \wedge |x|^2)\nu_t(dx) < \infty$*.*

(3) $\Gamma_t = \int_0^t \gamma(s)ds$*, where* $\gamma : [0, T] \to \mathbb{R}$ *a deterministic function with finite variation.*

Then $(\gamma_t, A_t, \mu_t)_{t\geq 0}$ *are spot characteristics and define a unique additive process* $\{X_t\}$*.* $(\gamma(t), \sigma^2(t), \nu_t)$ *are called* local characteristics *of the additive process.*

An additive process $\{X_t\}$ with local characteristics $(\gamma(t), \sigma^2(t), \nu_t)$ is usually called *time-inhomogeneous Lévy process*[1]. Note that the time inhomogeneous Lévy process is a special case of an additive process.

It is straight forward to imagine Lévy copula functions for time-inhomogeneous Lévy processes:

Definition B.4. Let $\{X_t\}$ be a time-inhomogeneous Lévy process on $\mathbb{R}^d$ with local characteristics $(\gamma(t), \sigma^2(t), \nu_t)$. By 7.12 there exists, for each $t \in [0, T]$, a Lévy copula function F_t of ν_t. We call $\{F_t\}_{t\in[0,T]}$ the *family of local Lévy copulas of* $\{X_t\}$.

Then many of the findings on Lévy copula functions should apply naturally to a family of local Lévy copulas.

[1] Time inhomogeneous Lévy processes arise naturally in the context of the Lévy Libor market model pioneered by Eberlein and Özkan [28]

Bibliography

[1] Aas, K., H. Bakken, C. Czado, A. Frigessi (2007), Pair-copula constructions of multiple dependence, *Insurance: Math. and Econ.* (In Press).

[2] Abramowitz, B., I. Stegun (1965), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, New York.

[3] Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, in *Proc. 2nd Int. Symp. Inf. Theory*, 267–281.

[4] Artzner, P., F. Delbean, J. M. Eber, D. Heath (1999), Coherent measures of risk, *Math. Finance*, 9, 203–208.

[5] Barndorff-Nielsen, O., T. Mikosch, S. Resnick (2001), *Lévy Processes: Theory and Applications*, Birkhäuser, Bosten.

[6] Barndorff-Nielsen, O. E., A. M. Lindner (2007), Lévy copulas: Dynamics and transforms of upsilon type, *Scand. J. of Stat.*, 34(2), 298–316.

[7] Barndorff-Nielson, O. e., N. Shephard (2001), Modelling by Lévy processes for financial econometrics, in *Lévy Processes: Theory and Applications*, eds. O. Barndorff-Nielsen, T. Mikosch, S. Resnick, 283–318, Birkhäuser, Boston.

[8] Beck, A., M. Lesko (2006), Copula-Funktionen zur Ermittlung des Gesamtbankrisikoprofils, *Betriebswirt. Blätter*, 5, 289–293.

[9] Beck, A., M. Lesko, F. Schlottmann, K. Wimmer (2006), Copulas im Risikomanagement, *Kreditwesen*, 14, 727–731.

[10] Bedford, T., R. Cooke (2002), Vines - A new graphical model for dependent random variables, *Ann. of Stat.*, 30(4), 1031–1068.

[11] Benninga, S., T. Björk, Z. Wiener (2001), On the use of numeraires in option pricing. Working paper, Stockholm School of Economics.

[12] Bertoin, J. (1996), *Lévy Processes*, Cambridge University Press, Cambridge.

[13] Bouyé, E., V. Durrlmann, A. Nikeghbali, G. Riboulet, T. Roncalli (2000), Copulas for finance: A reading guide and some applications. Working paper, Crédit Lyonnais, Paris.

[14] Carr, P., D. Madan (1999), Option valuation using the Fast Fourier Transform, *J. of Comp. Finance*, 2, 61–73.

[15] Cherubini, U., E. Luciano (2002), Bivariate option pricing with copulas, *Appl. Math. Finance*, 8, 69–85.

[16] Cherubini, U., E. Luciano, W. Vecchiato (2004), *Copula Methods in Finance*, John Wiley & Sons, Chichester.

[17] Chicago Mercantile Exchange Inc. (2008), CME - Chicago Mercantile Exchange. Available from: `http://www.cme.com`, accessed on June 29th, 2008.

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

[19] Cont, R., E. Luciano (2004), Non-parametric calibration of jump-diffusion option-pricing models, *Comp. Finance*, 7(3), 1–49.

[20] Cont, R., P. Tankov (2004), *Financial Modelling with Jump Processes*, Chapman & Hall/CRC, London.

[21] Deheuvals, P. (1978), Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes, *Publ. Inst. Statst. Univ. Paris*, 23(3), 1–36.

[22] Dias, A. d. C. (2004), *Copula Inference for Finance and Insurance*, Ph.D. thesis, Eidgenössische Technische Hochschule, Zürich.

[23] Eberlein, E. (2001), Application of generalized hyperbolic Lévy motions to finance, in *Lévy Processes: Theory and Applications*, eds. O. Barndorff-Nielsen, T. Mikosch, S. Resnick, 319–336, Birkhäuser, Boston.

[24] Eberlein, E., R. Frey, M. Kalkbrener, L. Overbeck (2007), Mathematics in risk management, *Jahresber. der Deutsch. Math. Verein.*, 109, 165–193.

[25] Eberlein, E., J. Jacod (1997), On the range of options prices, *Finance and Stoch.*, 1(2), 131–140.

[26] Eberlein, E., W. Kluge, A. Papapantoleon (2006), Symmetries in Lévy term structure models, *Int. J. of Theor. and Appl. Finance*, 9, 967–986.

[27] Eberlein, E., N. Koval (2006), A cross-currency Lévy market model, *Quant. Finance*, 6(6), 465–480.

[28] Eberlein, E., F. Özkan (2005), The Lévy Libor Model, *Finance and Stoch.*, 9, 31–53.

[29] Eberlein, E., A. Papapantoleon (2005), Symmetries and pricing of exotic options in Lévy models, in *Exotic Option Pricing and Advanced Lévy Models*, eds. A. Kypriano, W. Schoutens, P. Wilmott, 99–128, John Wiley & Sons, Chichester.

[30] Eberlein, E., S. Raible (1999), Term structure models driven by general Lévy processes, *Math. Finance*, 9, 31–53.

[31] Embrecht, P., R. Frey, A. McNeil (2005), *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton University Press, New Jersey.

[32] Embrecht, P., F. Lindskog, A. McNeil (2003), Modelling dependence with copulas and applications to risk management, in *Handbook of Heavy Tailed Distributions in Finance*, ed. S. Rachev, 329–384, Elsevier/North-Holland, Amsterdam.

[33] Embrecht, P., A. McNeil, D. Straumann (2002), Correlation and dependence in risk management: Properties and pitfalls, in *Risk Management: Value at Risk and Beyond*, ed. M. Dempster, 176–223, Cambridge University Press, Cambridge.

[34] Fang, K.-T., S. Kotz, K.-W. Ng (1990), *Symmetric Multivariate and Related Distributions*, Chapman & Hall/CRC, London.

[35] Fréchet, M. (1951), Sur les tableaux de corrélations dont les marges sont données, *Ann. de l'Univ. de Lyon*, 9, 53–77.

[36] Friedberg, J., M. Schumacher (2006), Risikotragfähigkeitskonzept bietet betriebswirtschaftliche Vorteile, *Betriebswirt. Blätter*, 12, 676–686.

[37] Genest, C., J. MacKay (1986), The joy of copulas: Bivariate distributions with uniform marginals, *Amer. Stat.*, 40, 280–285.

[38] Gillardon AG Financial Software (2006), Gillardon Finanzsoftware - Gesamtbanksteuerung, Produktkalkulation, Kundenberatung. Available from: `http://www.gillardon.de`, accessed on September 12th, 2006.

[39] Grabau, M. (2006), MaRisk-Anforderungen an die Risikotragfähigkeit von Instituten, *Betriebswirt. Blätter*, 12, 687–689.

[40] Hendricks, D., R. Cole (2003), Trends in risk integration and aggregation. Joint Forum publication, Bank for international Settlement, Basel.

[41] Hendricks, D., R. Cole, T. J. Forum (2001), The New Basel Capital Accord. Joint Forum publication, Bank for international Settlement, Basel.

[42] Hutchinson, T., C. Lai (1990), *Continuous Bivariate Distributions, emphasizing Applications*, Rumsby Scientific Publishing, Adelaide.

[43] Joe, H. (1997), *Multivariate Models and Dependence Concepts*, Chapman & Hall/CRC, London.

[44] Kahl, C., R. Lord (2006), Optimal Fourier inversion in semi-analytical option pricing. Working paper, Bergische Universität Wuppertal.

[45] Kallsen, J., P. Tankov (2004), Characterization of dependence of multidimensional Lévy processes using Lévy copulas, *J. of Mult. Analysis*, 97, 1551–1573.

[46] Karatzas, I., S. Shreve (1998), *Methods of Mathematical Finance*, Springer, Berlin.

[47] Kimeldorf, G., A. R. Sampson (1989), A framework for positive dependence, *Ann. Inst. Stat. Math.*, 41, 31–45.

[48] Kloman, H. F. (1990), Risk management agonists, *Risk Analysis*, 10, 201–205.

[49] Lindskog, F. (2000), *Modelling Dependence with Copulas and Applications to Risk Management*, Master's thesis, Eidgenössische Technische Hochschule, Zürich.

[50] Ling, C.-H. (1965), Representation of associative functions, *Publ. Math. Debrecen*, 12, 189–212.

[51] Luciano, E., W. Schoutens (2005), A multivariate jump-driven financial asset model. UCS report, Katholieke Universität Leuven.

[52] Markit Group Limited (2007), Markit. Available from: `http://www.markit.com`, accessed on July 31st, 2007.

[53] Mashal, R., A. Zeevi (2002), Beyond correlation: Extreme co-movements between financial assets. Working paper, Columbia Business School.

[54] Massey, F. J. (1956), The Kolmogorov-Smirnov test for goodness of fit, *J. of the Amer. Stat. Assoc.*, 46, 68–77.

[55] Melchiori, M. (2003), Which Archimedean copula is the right one?, *Yield Curve.com.*

[56] Nasdaq OMX Group Inc. (2008), NASDAQ OMX PHLX. Available from: `http://www.phlx.com`, accessed on June 29th, 2008.

[57] Nelson, R. (2006), *An Introduction to Copulas*, Springer, New York.

[58] Nolan, J. P. (2001), Maximum Likelihood Estimation and diagnostics for stable distributions, in *Lévy Processes: Theory and Applications*, eds. O. Barndorff-Nielsen, T. Mikosch, S. Resnick, 379–400, Birkhäuser, Boston.

[59] OnVista Media GmbH (2008), OnVista: Wertpapieranalyse - Aktien, Optionsscheine, Fonds, Charts, News, Kurse, Kennzahlen, Märkte. Available from: `http://www.onvista.de`, accessed on June 29th, 2008.

[60] Palaro, H. P., L. K. Hotta (2006), Using conditional copula to estimate value at risk, *J. of Data Science*, 4, 93–115.

[61] Papapantoleon, A. (2006), *Application of Semi-martingales and Lévy Processes in Finance: Duality and Valuation*, Ph.D. thesis, Albert Ludwigs Universität Freiburg.

[62] Papapantoleon, A. (2008), An introduction to Lévy processes in finance. Lecture notes, Technische Universität Wien.

[63] Patton, A. J. (2001), Modelling time varying exchange rate dependence using the conditional copula. Working paper, University of California San Diego.

[64] Raible, S. (2000), *Lévy Processes in Finance: Theory, Numerics and Empirical Facts*, Ph.D. thesis, Albert Ludwigs Universität Freiburg.

[65] Rockinger, M., E. Jondeau (2002), Conditional dependency of financial series: an application of copulas. Research Paper, International Center for Financial Asset Management and Engineering.

[66] Rosenblatt, M. (1986), Remarks on a multivariate transformation, *The Ann. of Math. Stat.*, 40, 470–472.

[67] Rosinski, J. (2001), Series representations of Lévy processes from the perspective of point processes, in *Lévy Processes: Theory and Applications*, eds. O. Barndorff-Nielsen, T. Mikosch, S. Resnick, 401–415, Birkhäuser, Boston.

[68] Salmon, M., C. Schleicher (2006), Pricing multivariate currency options with copulas, in *Copulas, from Theory to Application in Finance*, ed. J. Rank, Risk Books, London.

[69] Sato, K.-I. (1999), *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, New York.

[70] Savu, C., M. Trede (2006), Hierarchical archimedean copulas. Working paper, Universät Münster.

[71] Schumacher, M., C. Sievi, O. Wegner (2006), Risikotragfähigkeitskonzept bietet betriebswirtschaftliche Vorteile, *Betriebswirt. Blätter*, 12, 676–686.

[72] Schweizer, B. (1991), Thirty years of copulas, in *Advances in Probability Distributions with given Marginals*, eds. G. Dall'Aglio, S. Kotz, G. Salinetti, 13–50, Kluwer, Dortrecht.

[73] Serfling, R. (1980), *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons, New York.

[74] Sklar, A. (1959), Fonctions de repartition á n dimensions et leurs marges, *Publ. de l'Inst. Stat. de l'Univ. Paris*, 8, 229–231.

[75] Sparkasse Leverkusen (2008), Ihre Sparkasse Leverkusen. Available from: `http://www.sparkasse-lev.de`, accessed on June 29th, 2008.

[76] Sparkassen-Finanzportal GmbH (2008), www.sparkasse.de Homepage. Available from: `http://www.sparkasse.de`, accessed on September 12th, 2008.

[77] Tankov, P. (2004), *Lévy Processes in Finance: Inverse Problems and Dependence Modelling*, Ph.D. thesis, École Polytechnique de Paris.

[78] Tankov, P. (2006), Simulation and option pricing in Lévy copula models. Working paper, Université de Paris.

[79] Tappe, K. (2007), Launching a canon into multivariate Lévy processes. Working paper, Bergische Universität Wuppertal.

[80] Tappe, K. (2008), Modular Lévy copulas join up for series representation. Working paper, Bergische Universität Wuppertal.

[81] Wakefield, C. (2005), Copula methods for pricing multi-asset options. Working paper, University of Johannesburg.

[82] Wannenwetsch, J. (2005), *Lévy Processes in Finance: The Change of Measure and Non-linear Dependence*, Ph.D. thesis, Rheinische Friedrich-Wilhelms-Universität Bonn.

[83] Whelan, N. (2004), Sampling from Archimedean copulas, *Quant. Finance*, 4(3), 339–352.

[84] Xu, J. (1996), *Statistical Modelling and Inference for Multivariate and Longitudinal Discrete Response Data*, Ph.D. thesis, University of British Columbia.