Zentrum Mathematik
Lehrstuhl für Mathematische Statistik
der Technischen Universität München

Quantifying Risk: Modelling and Estimation

Klaus Böcker

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Herbert Spohn
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Claudia Klüppelberg
	2.	Univ.-Prof. Dr. Ralf Korn, Technische Universität Kaiserslautern
	3.	Prof. Andrea Resti, Università Commerciale Luigi Bocconi, Milan / Italien (schriftliche Beurteilung)

Die Dissertation wurde am 23.10.2008 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 19.11.2008 angenommen.

Zusammenfassung

Der Schwerpunkt dieser Arbeit dient der Erforschung von Risikomodellen, die über die zur Messung des Markt- und Kreditrisikos angewandten Methoden hinausgehen. Deshalb betrachten wir die Quantifizierung des multivariaten operationellen Risikos, des multivariaten Geschäftsrisikos und die Aggregation von verschiedenen Risikoarten.

Wir beginnen mit dem operationellen Risiko, wobei wir zunächst den eindimensionalen Fall betrachten und zeigen, dass bei langschwänzigen Verlustdaten (wie sie in der Praxis vorkommen) eine geschlossene Näherungsformel für den operationellen Value-at-Risk existiert. Danach erweitern wir unser Modell auf mehrere Risikozellen, wobei wir annehmen, dass das gesamte operationelle Risiko einer Bank durch einen zusammengesetzten Poisson-Prozess beschrieben werden kann. Die Abhängigkeitsstruktur zwischen den verschiedenen Randprozessen wird dabei durch eine Lévy-Copula modelliert und wir stellen analytisch dar, welche Konsequenzen sich daraus für gleichzeitige Verlustereignisse in verschiedenen Risikozellen ergeben. Ähnlich wie im eindimensionalen Fall erlaubt uns diese Technik, geschlossene Formeln für den operationellen Value-at-Risk für verschiedene Abhängigkeitsstrukturen zu berechnen, z.B. vollständige Abhängigkeit, Unabhängigkeit und multivariate reguläre Variation.

Für die Quantifizierung des Geschäftsrisikos nehmen wir an, dass die zukünftige Entwicklung der Geschäftsergebnisse einer Bank durch eine multivariate Brownsche Bewegung modelliert wird. Mit Hilfe dieses Modells wird dann die Auswirkung zukünftiger Ergebnisschwankungen auf den sogenannten Capital-at-Risk untersucht, wobei unterschiedliche Gauß-Prozesse verwendet und numerisch analysiert werden.

Hinsichtlich der Risikoaggregation führen wir eine vergleichende Analyse verschiedener Aggregationsverfahren durch. Besonderen Wert legen wir auf die Darstellung, wie Expertenwissen bei der Bestimmung der Korrelation zwischen verschiedenen Risikoarten berücksichtigt werden kann. Schließlich wird die Abhängigkeit zwischen einem Kreditportfolio und dem aggregierten Marktrisiko ausführlich untersucht. Dabei berechnen wir einen analytischen Ausdruck für die lineare Korrelation beider Teilportfolien und entwickeln einen neuen Schätzer für den Gauß-Copula-Parameter, der im Fall nahezu beliebiger Kreditportfolien angewendet werden kann.

Abtract

This thesis is devoted to the investigation of risk models that go beyond those techniques used for measuring the standard risk types market risk and credit risk. Therefore, we consider the quantification of multivariate operational risk, multivariate business risk, and the aggregation of different risk types.

Starting with operational risk, we show that in the one-dimensional case, when loss data are heavy-tailed (which in practice they are), a simple closed-form approximation for the operational Value-at-Risk can be obtained. We then extend our model to several operational risk cells assuming that a bank's total operational Value-at-Risk can be described by a multivariate compound Poisson process. The dependence structure between different marginal processes is described by a Lévy copula, and we analytically discuss the consequences of this model for simultaneous loss events in different risk cells. Similarly to the one-dimensional case, this technique allows us to calculate closed-form expressions for total operational Value-at-Risk for different dependence structures, including complete dependence, independence, and multivariate regular variation.

For the assessment of business risk, we assume that the evolution of a bank's earnings path is modelled by a multivariate Brownian motion. This model is then used for investigating the impact of future earnings uncertainties on the so-called Capital-at-Risk. In doing so, different Gauss processes are considered and analysed numerically.

With respect to risk aggregation, we perform a comparative analysis of different aggregation methodologies. Particular emphasis is given to the description of how expert knowledge can be included when calculating the correlation between different risk types. Finally, we intensively investigate the dependence between a credit portfolio and aggregated market risk. In doing so, we analytically calculate their linear correlation, and, moreover, we suggest a new estimator for their Gaussian copula parameter, which can be applied to almost arbitrary credit portfolios.

Acknowledgment

I am greatly indebted to my supervisor, Prof. Claudia Klppelberg, for her very helpful support, patience, and ongoing advice. I am especially grateful to her that she enabled me to do this thesis while I have been working as a full-time employee in a large international bank.

I also want to express my gratitude to Prof. Ralf Korn and Prof. Andrea Resti for acting as referees of this thesis.

Finally, my wife Katharina and my children deserve special thanks for their encouragement, support and love.

Contents

Chapter 1

Introduction

1.1 Motivation and Background

> *It would be a mistake to conclude that the only way to succeed in banking is through ever-greater size and diversity. Indeed, better risk management may be the only truly necessary element of success in banking.*
> Alan Greenspan, Speach to the American Bankers Association, October 5, 2004.

> *This crisis is the result of regulatory failure to guard against excessive risk-taking in the financial system, ...*
> Dominique Strauss-Kahn, Financial Times, September 22, 2008.

Risk is an inevitable part of every financial institution. Risks are implicitly accepted when banks or insurance companies provide their financial services to customers and explicitly when they take risk positions that offer profitable, above-average returns. However, the rush to ever growing profits combined with excessive risk-taking and erroneous risk management almost inevitably result in a financial crisis such as we are witnessing today. Then, the growing awareness of risk inherent in the banking industry often leads to an increasing demand for banking supervision at international level. So, spectacular crunches like the Saving & Loans crisis in the 1970s or the Japanese banking crisis in the 1990s were catalysts for the creation of the Basel Committee of Banking Supervision under the auspices of the Bank for International Settlement (BIS) in Basel. Needless to say, also the current financial turmoil will leave its mark in tomorrow's bank capital regulation and risk management practices, as it is for instance indicated by the above quote of Dominique Strauss-Kahn, the Managing Director of the International Monetary Fund (IMF).

There is no unique view on risk and often it is divided in certain sub-classes such as market risk, credit risk and operational risk. Market risk is associated with trading

activities; it is defined as the potential loss arising from adverse price changes of a bank's positions in financial markets and usually encompasses interest rate, foreign exchange, equity, and credit-spread risk. Credit risk is owing to potential losses arising from a customer's default or a deteriorating credit rating. Such risks usually include loan-default risk, counterparty risk, issuer risk and country risk. Finally, operational risk is the risk of losses resulting from inadequate or failed processes, people and systems, or external events.

The basic idea underlying modern banking regulation is that every bank must keep a certain amount of equity capital (the so-called "capital charge") as a buffer against their risks. In doing so, a bank has to quantify its overall risk exposure the best of its knowledge, and, of course, to the satisfaction of its supervising authority. Therefore, sound risk measurement techniques are a prerequisite for a successful banking supervision. With this respect, the bulk of research that had been published over the past two decades, both by academics and practitioners, was within the huge fields of market risk and credit risk. However, with the new Basel Accord "Basel II" [5], which had to be fully implemented by year-end 2007, the focus was moving also to operational risk, for which banks now—similar as it is already the case for market risk—can use their own internal modelling techniques (commonly referred to as advanced measurement approaches (AMA)) to determine operational risk capital charges.

Moreover, Basel II requires banks to model and measure basically all of the "other" material risks they are exposed to, encompassing, e.g., interest rate risk in the banking book, credit concentration risk, and business risk. Besides this sheer regulatory perspective, the increased materiality of business risk, in particular, was recently also confirmed by an industry survey of the IFRI/CRO Forum about economic capital practices [36], according to which 85 % of the participating banks try to include business risk in their overall risk assessment. However, this survey has also revealed that the quantification of business risk is still in its infancy and measurement techniques are often judgemental.

This is exactly where our research comes in. In the first two Chapters we present a thorough investigation of the two "new", Basel-II relevant risk types operational risk and business risk.
For operational risk, our main result is a new method for modelling multivariate operational risk (measured by means of Value-at-risk), which is based on the concept of Lévy copulas. From a methodological perspective, we are not so much interested in numerical simulations; instead, we present a completely analytic treatment of this intriguing subject. As a consequence thereof, we deliver new insight into the general behaviour of multivariate operational risk, which is of great practical relevance for operational

risk management. For instance, for important examples of the Lévy copula (and thus of the dependence structure) and heavy-tailed loss data of regular variation, we derive first-order approximations for multivariate operational Value-at-Risk (OpVAR).

With respect to business risk, we develop a new continuous-time model where future earnings' uncertainty is driven by a multivariate Brownian motion. As a matter of fact, this model can actually be considered as a mark-to-model approach for assessing the economic capital associated with this important risk type. Special attention is paid to the examination of two popular measures used in business risk context, namely the Earnings-at-Risk (EAR) and the Capital-at-Risk (CAR). We perform a sensitivity analysis concerning the relationship between EAR and CAR with respect to different Gauss processes for the future earnings.

Finally we consider the aggregation of different risk types into a single number. This is of utmost relevance for risk control because, e.g., according to the Committee of European Banking Supervisors [25], banks have to calculate an "overall capital number" as an integral part of their capital adequacy assessment process. This single-number metric should encompass all risks associated with different businesses and risk types. Above all, regulators want to understand the extent to which the institution has introduced diversification and correlation effects when aggregating different risk types. Our investigation of risk aggregation mainly consists of three parts. First, a comparative analysis of different aggregation techniques encompassing four different risk types; second, a closer inspection of how expert knowledge can be elicited and used for the estimation of the parameters of the inter-risk-type-dependence copula; and, finally, an extended analysis of the inter-risk dependence structure between a credit risk portfolio and market risk. We present new closed-form results for the inter-risk correlation between market and credit risk and suggest a novel way of how the parameter of a Gaussian copula can easily be approximated for almost arbitrary credit portfolios.

1.2 Outline of the Thesis

Every Chapter of the following thesis is based on one or more research papers cited below. Hence, all Chapters are basically self-contained and the notation is only unified within the individual Chapters. In detail, this thesis is organised as follows.

Chapter 2 is based on Böcker [9], Böcker & Klüppelberg [13, 14, 15, 17, 18], and Böcker & Sprittulla [20]. Section 2.1 is devoted to the standard univariate loss distribution approach where the cumulated operational loss $S(t)$ of a single operational risk cell up

to time t is described by an aggregate loss process

$$S(t) = \sum_{k=1}^{N(t)} X_k \,, \quad t \geq 0 \,,$$

with iid $(X_k)_{k\in\mathbb{N}}$ (positive) loss variables and frequency process $(N(t))_{t\geq 0}$, independent of $(X_k)_{k\in\mathbb{N}}$. For such a model, OpVAR can usually not be calculated analytically. We show, however, that when loss data are heavy-tailed (which in practice they are), a simple first-order approximation in closed form for OpVAR can be obtained. We then apply this approximation to Pareto-like severity models.

Section 2.2 is devoted to multivariate operational risk, i.e., we consider several operational risk cells $i =, 1, \ldots, d$. Assuming that the loss frequency process within each cell follows a homogeneous Poisson process with rate $\lambda_i > 0$, the aggregate loss $S_i(\cdot), i = 1, \ldots, d$, is compound Poisson and thus a special kind of a Lévy process. This allows us to use the new concept of Lévy copulas (see e.g. Cont & Tankov [24]) for calculating operational risk at a high confidence level. Because we model the dependence in frequency and severity between different cells simultaneously, the model has only a few parameters; an important advantage when practical calculations have to be employed.

Similarly to the univariate case, we present closed-form approximations for multivariate OpVAR in Section 2.3, shedding light upon extremal dependence modelling in general. Particular emphasis is given to multivariate regularly dependence (see e.g. Resnick [60, 61]), a useful concept to describe the dependence structure between heavy-tailed data such as observed in operational risk.

Chapter 3 is going to appear in Böcker [11]. We propose a continuous-time model for business risk where the bank's future earnings are modelled by means of multivariate Brownian motion. In particular, we distinguish between the two measures EAR and CAR, see also the discussion in Saita [64, 65]. After discussing some preliminaries in Section 3.2.1, such as a formulation of the discounted-cash-flow method in continuous time, we introduce a first stochastic model for quantifying business risk in Section 3.2.2, which is used in Section 3.2.3 to investigate the relation between EAR and CAR in greater detail. Specifically, we derive an analytic relationship between both measures. Finally, in Section 3.3, we consider a possible extension of the business risk model based on an alternative Gauss process and compare it to the simple model.

Chapter 4 is based on the paper Böcker & Hillebrand [12] and the book contributions Böcker [11] and Böcker & Klüppelberg [16] where we elaborate some specific aspects in the context of the aggregation-by-risk-type approach, a popular way for estimating the aggregated risk profile of a bank. In Section 4.2 we first review the aggregation-by-risk-

type approach and recall some important properties of the (pairwise) linear correlation coefficient, followed by a discussion of alternative dependence measures such as copulas, Kendall's tau, and tail dependence; and finally of the inter-risk-correlation matrix as a whole. Section 4.2.4 investigates how the estimation of inter-risk correlation can be supported by means of expert judgements, a field, which so far has only received little attention in the context of risk aggregation. In Section 4.2.5 we illustrate our results with an extended numerical example, which describes a typical aggregation-by-risk-type problem as it may also occur in practice.

In Section 4.3, we then concentrate on the interaction between a credit risk portfolio and another, normally distributed risk type, say market risk. Combining Merton-like factor models for credit risk with linear factor models for market risk, we calculate closed-form expressions for their inter-risk correlation and, moreover, suggest how their Gaussian copula can be estimated.

Chapter 2

Modelling and Measuring Operational Risk

This Chapter is devoted to the quantification of operational risk. In Section 3.1 we first present a simple closed-form approximation for the OpVAR of a single event type/business line cell. In Section 2.2 we then introduce the new concept of Lévy copulas to model the dependence structure between operational loss events occurring in different cells. Finally, in Section 2.3, we derive first order approximations for multivariate OpVAR for important dependence structures, including complete dependence, independence, and multivariate regular variation.

2.1 Operational Risk of a Single Cell

2.1.1 The Loss Distribution Approach

The new Basel II accord [5] imposes new methods of calculating regulatory capital that apply to the banking industry. Besides credit risk, the new accord focuses on operational risk, defined as the risk of losses resulting from inadequate or failed internal processes, people and systems, or from external events.

According to [5], every bank has to calculate explicit capital charges to cover potential operational risk losses by means of one of three approaches: the basic indicator approach, the standardised approach, and the advanced measurement approach (AMA). This "continuum of approaches" reflects different levels of sophistication and risk sensitivity. AMA as the most flexible approach for operational risk quantification allows the bank to build its own internal operational risk model and measurement system, comparable to market risk standards. Instead of prescribing a particular type of VAR approach, however, the committee focuses only on a set of quantitative and

qualitative standards, which have to be fulfilled by the institution's AMA model. The following two requirements are directly cited from [5] and are most relevant for the issues discussed in this Chapter:

(1) The operational-risk measure is a VAR at confidence level 99.9 % with a holding period of one year (cf. § 667)

(2) The measurement approach must capture potentially severe tail loss events (cf. § 667).

The most popular method in the industry to satisfy the AMA standards is the loss distribution approach (LDA), which is based on modelling the probability distribution of operational losses using bank-internal and external data. Such type of model indeed originates in insurance and has its root in *insurance risk theory*, which goes back to the early work by Filip Lundberg in 1903. A key feature of every LDA model is to split up the total loss amount over a certain time period into a frequency component, i.e. the number of losses, and a severity component, i.e. the individual loss amounts. Then, the total loss is obtained by compounding the frequency and the severity information. A prototypical model of this kind, which is currently best practice and implemented in various commercial software packages, is the following.

Definition 2.1.1. [Univariate Standard LDA model]

(1) The severity process.
The severities $(X_k)_{k\in\mathbb{N}}$ *are positive iid random variables with distribution function* $F \in \mathcal{S}$ *describing the magnitude of each loss event.*

(2) The frequency process.
The number $N(t)$ *of loss events in the time interval* $[0,t]$ *for* $t \geq 0$ *is random. The resulting* counting process $(N(t))_{t\geq 0}$*, is generated by a sequence of points* $(T_n)_{n\geq 1}$ *of non-negative random variables satisfying*

$$0 \leq T_1 \leq T_2 \leq \cdots \ \ a.s.$$

and

$$N(t) = \sup\{n \geq 1 : T_n \leq t\}\,, \quad t \geq 0\,.$$

(3) *The severity process and the frequency process are assumed to be independent.*

(4) The aggregate loss process.
The aggregate loss $S(t)$ *in* $[0,t]$ *constitutes a process*

$$S(t) = \sum_{k=1}^{N(t)} X_k\,, \quad t \geq 0\,. \tag{2.1}$$

With respect to the frequency process, there are two typical specifications which are widely used for LDA models in practice:

Example 2.1.2. (a) The *Poisson-LDA* is a Standard LDA, where $(N(t))_{t\geq 0}$ is a homogenous Poisson process with intensity $\lambda > 0$, in particular,

$$P(N(t) = n) = p_t(n) = e^{-\lambda t}\,\frac{(\lambda t)^n}{n!}, \quad n \in \mathbb{N}_0\,.$$

(b) The *negative-binomial-LDA* is a Standard LDA, where $(N(t))_{t\geq 0}$ is given by a negative binomial process satisfying for $\beta, \gamma > 0$

$$P(N(t) = n) = p_t(n) = \binom{\gamma + n - 1}{n} \left(\frac{\beta}{\beta + t}\right)^{\gamma} \left(\frac{t}{\beta + t}\right)^{n}, \quad n \in \mathbb{N}_0\,.$$

□

The negative binomial distribution is a gamma mixture of a Poisson distribution, i.e. it can be viewed as a Poisson distribution whose parameter λ is a gamma distributed random variable. This allows for modelling *over-dispersion*, which means that for all $t > 0$ the variance of $N(t)$ is greater than its mean, whereas for the Poisson-LDA $\mathrm{var}(N(t)) = EN(t)$ holds. However, we will see later that as far as the OpVAR approximation is concerned, over-dispersion is of minor importance.

Regarding the severity of losses, note that we do not require X_k to have finite mean and/or variance. This is in accordance with empirical research: Moscadelli [54] showed very convincingly that typical severity distributions for operational risk are very heavy-tailed such that even moments of low order may not exist. This fact is also acknowledged by the regulators, who want banks to take the heavy-tail property of operational losses explicitly into account.

Finally, the goal of every LDA model is mainly to determine the *aggregate loss distribution*, which for the Standard LDA can be written as

$$\begin{aligned} G_t(x) &= P(S(t) \leq x) \\ &= \sum_{n=0}^{\infty} p_t(n)\, P(S(t) \leq x | N(t) = n) \\ &= \sum_{n=0}^{\infty} p_t(n)\, F^{n*}(x), \quad x \geq 0, \quad t \geq 0\,, \end{aligned} \tag{2.2}$$

where $F(\cdot) = P(X_k \leq \cdot)$ is the distribution function of X_k, and $F^{n*}(\cdot) = P(\sum_{i=1}^{n} X_i \leq \cdot)$ is the n-fold convolution of F with $F^{1*} = F$ and $F^{0*} = I_{[0,\infty)}$.

For most choices of severity and frequency distributions, G_t cannot be calculated analytically. Approximation methods to overcome this problem include Panjer recursion, Monte Carlo simulation, and FFT (fast Fourier transform) methods, see Klugman,

Name	Distribution function	Parameters
Lognormal	$F(x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right)$	$\mu \in \mathbb{R},\ \sigma > 0$
Weibull	$F(x) = 1 - e^{-(x/\theta)^\tau}$	$\theta > 0,\ 0 < \tau < 1$
Pareto	$F(x) = 1 - \left(1 + \frac{x}{\theta}\right)^{-\alpha}$	$\alpha, \theta > 0$

Table 2.1.3. *Popular severity distributions with support* $(0, \infty)$. *(Φ is the standard normal distribution function).*

Panjer and Willmot [41] for an overview. The drawback of these methods is, however, that their results remain a "black box", and the interaction between different model parameters and their impact on the final result, i.e. OpVAR is only interpretable through extensive sensitivity analyses. However, as both regulatory capital and economic capital for operational risk are based on a very high quantile of the aggregate loss distribution G_t, a natural estimation method for OpVAR is via asymptotic tail and quantile estimation. Before we consider this is more detail, we have to specify more precisely what heavy-tailness of the loss severities actually means.

2.1.2 Subexponentiality and Univariate Regular Variation

Some popular distributions used for modelling the severity of single loss events are given in Table 2.1.3. All of them are heavy-tailed, more precisely, they belong to the class of so-called *subexponential distributions*, meaning that their tails decay slower than any exponential tail.

The defining property of subexponential distributions is that the tail of the sum of n subexponential random variables has the same order of magnitude as the tail of the maximum variable among them, or, more precisely, the sum of n iid severities is most likely to be large because of one of the terms being large. This can be written as follows:

Definition 2.1.4. [Subexponential distributions] *Let $(X_k)_{k\in\mathbb{N}}$ be iid random variables with distribution function F. Then F (or sometimes $\overline{F}$) is said to be a* subexponential distribution function ($F \in \mathcal{S}$) *if*

$$\lim_{x\to\infty} \frac{P(X_1 + \cdots + X_n > x)}{P(\max(X_1, \ldots, X_n) > x)} = 1 \quad \textit{for some (all) } n \geq 2.$$

Focusing on operational risk, this implies that severe overall losses are mainly due to a single big loss rather than the consequence of accumulated small independent losses. Of course, this insight should have implications for operational risk management. For more details on subexponential distributions and related classes see Embrechts et al. [29], Appendix A3.

As a useful semiparametric class of subexponential distributions, we will in the sequel heavily rely on distributions, whose far out right tails behave like a power function.

Definition 2.1.5. [Regularly varying functions] *Let f be a positive measureable function. If for some $\alpha \in \mathbb{R}$*

$$\lim_{t\to\infty} \frac{f(xt)}{f(t)} = x^{-\alpha}\,, \quad x > 0\,, \tag{2.3}$$

then f is called regularly varying with index α.

Here we consider loss variables X whose distribution tails are regularly varying.

Definition 2.1.6. [Regularly varying distribution tails] *Let X be a positive random variable with distribution tail $\overline{F}(x) := 1 - F(x) = P(X > x)$ for $x > 0$. If for $\overline{F}$ relation* (2.3) *holds for some $\alpha \geq 0$, then X is called* regularly varying with index $-\alpha$ *and denoted by $\overline{F} \in \mathcal{R}_{-\alpha}$. The quantity α is also called the* tail index *of F.*
Finally we define $\mathcal{R} := \cup_{\alpha\geq 0}\mathcal{R}_{-\alpha}$.

Regular variation is a powerful concept when working with heavy-tailed data and it is worthwhile to recall some of its properties.

Remark 2.1.7. (a) Subexponentiality implies regular variation, $\mathcal{R} \subset \mathcal{S}$.
(b) Regularly varying distribution functions have representation $\overline{F}(x) = x^{-\alpha}L(x)$ for $x \geq 0$, where L is a *slowly varying function* ($L \in \mathcal{R}_0$) satisfying $\lim_{t\to\infty} L(xt)/L(t) = 1$ for all $x > 0$. Typical examples are functions, which converge to a positive constant or are logarithmic as e.g. $L(\cdot) = \ln(\cdot)$.
(c) The classes $\mathcal{S}$ and $\mathcal{R}_{-\alpha}$, $\alpha \geq 0$, are closed with respect to *tail-equivalence*, which for two distribution functions (or also tail integrals) is defined as $\lim_{x\to\infty} \overline{F}(x)/\overline{G}(x) = c$ for $c \in (0, \infty)$.
(d) We introduce the notation $\overline{F}(x) \sim \overline{G}(x)$ as $x \to \infty$, meaning that the quotient of right-hand and left-hand side tends to 1; i.e. $\lim_{x\to\infty} \overline{G}(x)/\overline{F}(x) = 1$.
(e) In Definition 2.1.6 we have used a functional approach to regular variation. Alternatively, regular variation can be reformulated in terms of vague convergence of the underlying probability measures, and this turns out to be very useful when we consider multivariate operational risk in Section 2.3.5. □

Distributions in $\mathcal{S}$ but not in $\mathcal{R}$ include the heavy-tailed Weibull distribution and the lognormal distribution. Their tail decreases faster than tails in $\mathcal{R}$, but less fast than an exponential tail. The following definition will be useful.

Definition 2.1.8. [Rapidly varying distribution tails] *Let X be a positive random variable with distribution tail $\overline{F}(x) := 1 - F(x) = P(X > x)$ for $x > 0$. If*

$$\lim_{t\to\infty} \frac{\overline{F}(xt)}{\overline{F}(t)} = \begin{cases} 0\,, & \text{if } x > 1\,, \\ \infty & \text{if } 0 < x < 1\,. \end{cases}$$

then $\overline{F}$ is called rapidly varying*, denoted by $\overline{F} \in \mathcal{R}_\infty$.*

Other important results related to the generalised inverse of regularly varying and rapidly varying distributions are gathered by the following Proposition.

Proposition 2.1.9. *(1)* [Regular variation] *Let $\alpha > 0$. Then*
(i) $\overline{F} \in \mathcal{R}_{-\alpha} \quad \Leftrightarrow \quad (1/\overline{F})^{\leftarrow} \in \mathcal{R}_{1/\alpha}$,
(ii) $\overline{F}(x) = x^{-\alpha}L(x)$ *for* $x \geq 0 \quad \Leftrightarrow \quad (1/\overline{F})^{\leftarrow}(z) = z^{1/\alpha}\widetilde{L}(z)$ *for* $z \geq 0$,
where L and $\widetilde{L}$ are slowly varying functions,
(iii) $\overline{F}(x) \sim \overline{G}(x)$ *as* $x \to \infty \quad \Leftrightarrow \quad (1/\overline{F})^{\leftarrow}(z) \sim (1/\overline{G})^{\leftarrow}(z)$ *as* $z \to \infty$.

(2) [Rapid variation] *If $\overline{F}, \overline{G} \in \mathcal{R}_\infty$ such that $\overline{F}(x) \sim \overline{G}(x)$ as $x \to \infty$, then $(1/\overline{F})^{\leftarrow}(z) \sim (1/\overline{G})^{\leftarrow}(z)$ as $z \to \infty$.*

Proof. (1) Proposition 1.5.15 of Bingham, Goldie and Teugels [7] ensures that regular variation of $1/\overline{F}$ is equivalent to regular variation of its (generalised) inverse and provides the representation. Proposition 0.8(vi) of Resnick [60] gives the asymptotic equivalence.
(2) Theorem 2.4.7 of [7](ii) applied to $1/\overline{F}$ ensures that $(1/\overline{F})^{\leftarrow} \in \mathcal{R}_0$. Furthermore, tail equivalence of F and G implies that $(1/\overline{F})^{\leftarrow}(z) = (1/\overline{G})^{\leftarrow}(z(1 + o(1))) = (1/\overline{G})^{\leftarrow}(z)(1 + o(1))$ as $z \to \infty$, where we have used that the convergence in Definition 2.1.6 is locally uniformly. □

2.1.3 A Closed-Form Approximation for OpVAR

OpVAR is simply a quantile of the aggregate loss distribution G_t.

Definition 2.1.10. [OpVAR] *Suppose G_t is a loss distribution function according to eq. (2.2). Then, OpVAR up to time t at confidence level κ,* $\mathrm{VAR}_t(\kappa)$*, is defined as its κ-quantile*

$$\mathrm{VAR}_t(\kappa) = G_t^{\leftarrow}(\kappa)\,, \quad \kappa \in (0,1)\,,$$

where $G_t^{\leftarrow}(\kappa) = \inf\{x \in \mathbb{R} : \ G_t(x) \geq \kappa\}$, $0 < \kappa < 1$, *is the (left continuous) generalized inverse of* G_t. *If* G_t *is strictly increasing and continuous, we may write* $\mathrm{VAR}_t(\kappa) = G_t^{-1}(\kappa)$.

As we already said, $G_t(\kappa)$—and thus also OpVAR—can in general not be calculated analytically. Recall, however, that κ is usually near 1, e.g. 0.999 for regulatory purposes or, in the context of a bank's internal economic capital, even higher such as 0.9995. Hence, instead of considering the entire distribution, it is sufficient to concentrate on the right tail $P(S(t) > x)$ for very large x. Now, in actuarial science, the tail behavior of G_t has been extensively studied both for small claims and large claims models. For the latter, the following key result basically stays that (under weak regularity conditions for the frequency process) for every fixed $t > 0$, G_t can be written in terms of the severity distribution function F.

Theorem 2.1.11. [Embrechts, Klüppelberg and Mikosch [29], Theorem 1.3.9] *Consider the standard LDA* $S(t) = \sum_{n=0}^{N(t)} X_i$, $t \geq 0$, *from Definition 2.1.1. Assume that the severities* X_i *are subexponential with distribution function* F. *Fix* $t > 0$ *and define the frequency distribution by* $P(N(t) = n) = p_t(n)$ *for* $n \in \mathbb{N}_0$. *Then, the aggregate loss distribution is given by*

$$G_t(x) = \sum_{n=0}^{\infty} p_t(n)\, F^{n*}(x)\,, \quad x \geq 0, \quad t \geq 0\,.$$

Assume that for some $\varepsilon > 0$,

$$\sum_{n=0}^{\infty} (1+\varepsilon)^n p_t(n) < \infty\,. \tag{2.4}$$

Then, G_t *is subexponential with tail behaviour given by*

$$\overline{G}_t(x) \sim EN(t)\overline{F}(x)\,, \qquad x \to \infty\,. \tag{2.5}$$

Specifically, it has been shown in Examples 1.3.10 and 1.3.11 of Embrechts, Klüppelberg and Mikosch [29] that the tail estimate (2.5) holds for the Poisson-LDA and the negative-binomial-LDA of Example 2.1.2.

Given relation The result of Theorem 2.1.11, it is straightforward to obtain an first-order approximation for OpVAR, valid at very high confidence levels.

Theorem 2.1.12. [OpVAR for subexponential distributed severities, Part I] *Consider the Standard LDA model for fixed* $t > 0$ *and a subexponential severity with distribution function* F. *Assume, moreover, that the tail estimate (2.5) holds. Then,* $\mathrm{VAR}_t(\kappa)$ *satisfies the approximation*

$$\mathrm{VAR}_t(\kappa) = F^{\leftarrow}\left(1 - \frac{1-\kappa}{EN(t)}(1+o(1))\right), \quad \kappa \to 1\,. \tag{2.6}$$

Proof. Note first that $\kappa \to 1$ is equivalent to $x \to \infty$. Then recall that $o(1)$ always stands for a function, which tends to 0, if its argument tends to a boundary, in our case if $\kappa \to 1$ or $x \to \infty$. With this notation relation (2.5) can be rewritten as

$$G_t(x) = 1 - EN(t)\,\overline{F}(x)(1 + o(1))\,, \quad x \to \infty\,.$$

Setting the right hand side equal to κ gives an asymptotic solution

$$F(x) = 1 - \frac{1-\kappa}{EN(t)}(1 + o(1))\,, \quad x \to \infty\,,$$

and, finally,

$$x = G_t^{\leftarrow}(\kappa) = F^{\leftarrow}\left(1 - \frac{1-\kappa}{EN(t)}(1 + o(1))\right), \quad \kappa \to 1\,. \qquad \square$$

An equivalent but sometimes more convenient and more natural way to express $\mathrm{VAR}_t(\kappa)$ in the context of high quantiles is by means of the tail of the severity distribution $\overline{F}(\cdot)$ instead of $F(\cdot)$ directly. This can easily be achieved by noting that $1/\overline{F}$ is increasing, hence,

$$\begin{aligned} F^{\leftarrow}(\kappa) &= \inf\{x \in \mathbb{R} :\ F(x) \geq \kappa\} \\ &= \inf\{x \in \mathbb{R} :\ 1/\overline{F}(x) \geq 1/(1-\kappa)\} \\ &=: \left(\frac{1}{\overline{F}}\right)^{\leftarrow}\left(\frac{1}{1-\kappa}\right), \quad 0 < \kappa < 1\,. \end{aligned} \tag{2.7}$$

This implies the following Corollary.

Corollary 2.1.13. *Consider the situation as in Theorem 2.1.12. Then, OpVAR is given by*

$$\mathrm{VAR}_t(\kappa) = \left(\frac{1}{\overline{F}}\right)^{\leftarrow}\left(\frac{EN(t)}{1-\kappa}(1 + o(1))\right), \quad \kappa \to 1\,. \tag{2.8}$$

For some general subexponential distribution function $F \in \mathcal{S}$, the right-hand side of (2.6) is not always asymptotically equivalent to $F^{\leftarrow}\big(1 - \frac{1-\kappa}{EN(t)}\big)$ as $\kappa \uparrow 1$. Consider the following example.

Example 2.1.14. Consider $(1/\overline{F})^{\leftarrow}(y) = \exp(y + y^{1-\varepsilon})$ for some $0 < \varepsilon < 1$ with $y = 1/(1-\kappa)$, i.e. $\kappa \uparrow 1$ equivalent to $y \to \infty$. Then $(1/\overline{F})^{\leftarrow}(y) = \exp(y(1 + o(1))$, but $(1/\overline{F})^{\leftarrow}(y)/e^y = \exp(y^{1-\varepsilon}) \to \infty$ as $y \to \infty$. This situation typically occurs, when $\overline{F} \in \mathcal{R}_0$, i.e. for extremely heavy-tailed models. $\square$

However, if we restrict the class subexponential distributions, we can formulate the analytical OpVAR result as follows.

Theorem 2.1.15. [OpVAR for subexponential distributed severities, Part II] *Consider the Standard LDA model for fixed $t > 0$ and a subexponential severity with distribution function F. Assume, moreover, that the tail estimate (2.5) holds.*
(i) *If $\overline{F} \in \mathcal{S} \cap (\mathcal{R} \cup \mathcal{R}_\infty)$, then $\mathrm{VAR}_t(\kappa)$ is asymptotically given by*

$$\mathrm{VAR}_t(\kappa) = \left(\frac{1}{\overline{G}_t}\right)^{\leftarrow}\left(\frac{1}{1-\kappa}\right) \sim F^{\leftarrow}\left(1 - \frac{1-\kappa}{EN(t)}\right), \quad \kappa \uparrow 1\,. \tag{2.9}$$

(ii) *The severity distribution tail belongs to $\mathcal{R}_{-\alpha}$ for $\alpha > 0$, i.e. $\overline{F}(x) = x^{-\alpha}L(x)$ for $x \geq 0$ and some slowly varying function L if and only if*

$$\mathrm{VAR}_t(\kappa) \sim \left(\frac{EN(t)}{1-\kappa}\right)^{1/\alpha} \widetilde{L}\left(\frac{1}{1-\kappa}\right), \quad \kappa \uparrow 1\,, \tag{2.10}$$

where $\widetilde{L}\left(\frac{1}{1-\cdot}\right) \in \mathcal{R}_0$.

Proof. (i) This is a consequene of Theorem 2.1.12 in combination with Proposition 2.1.9.
(ii) By Definition 2.1.10, $\mathrm{VAR}_t(\kappa) = G^{\leftarrow}(\kappa)$. By (2.5) we have $\overline{G}_t(x) \sim EN(t)\,\overline{F}(x)$ as $x \to \infty$. From Proposition 2.1.9 it follows that

$$\left(\frac{1}{\overline{G}_t}\right)^{\leftarrow}\left(\frac{1}{1-\kappa}\right) \sim \left(\frac{1}{\overline{F}}\right)^{\leftarrow}\left(\frac{EN(t)}{1-\kappa}\right) = \left(\frac{EN(t)}{1-\kappa}\right)^{1/\alpha} \widetilde{L}\left(\frac{EN(t)}{1-\kappa}\right), \quad \kappa \uparrow 1\,,$$

and the result follows. □

We refrain from giving more information on the relationship between L and $\widetilde{L}$ (which can be found in [7]) as it is rather involved and plays no role in this Chapter. When such a model is fitted statistically, then L and $\widetilde{L}$ are usually replaced by constants; see Embrechts et al. [29], Chapter 6. In that case $L \equiv \theta^\alpha$ results in $\widetilde{L} \equiv \theta$ as in the following example. To indicate that the equivalence of Theorem 2.1.15(ii) does not extend to subexponential distribution tails in $\mathcal{R}_\infty$ we refer to Example 2.3.10.

Theorems 2.1.12 and 2.1.15 hold for a quite general class of LDA models and are remarkable for two reasons. First, they allow for closed-form solutions of asymptotic OpVAR for many popular specifications of the severity distribution, see Table 2.1.16. Moreover, (2.9) says that OpVAR at high confidence levels only depends on the tail and not on the body of the severity distribution. Therefore, if one is only interested in VAR calculations, modelling the whole distribution function F is superfluous. Instead, to obtain a first order approximation for the OpVAR for a specific LDA model, it suffices to combine (2.6) with the subexponential tail of the severity distribution F, see also Section 2.1.5.

Name	$\mathrm{VAR}_t(\kappa)$
Lognormal	$\exp\left[\mu - \sigma\,\Phi^{-1}\left(\dfrac{1-\kappa}{EN(t)}\right)\right]$
Weibull	$\theta\left[\ln\left(\dfrac{EN(t)}{1-\kappa}\right)\right]^{\frac{1}{\tau}}$
Pareto	$\theta\left[\left(\dfrac{EN(t)}{1-\kappa}\right)^{1/\alpha} - 1\right]$

Table 2.1.16. *First order approximations of* $\mathrm{VAR}_t(\kappa)$ *as* $\kappa \to 1$ *for the aggregate loss distribution for popular severity distributions. Set* $EN(t) = \lambda\, t$ *for a Poisson distributed and* $EN(t) = \gamma\, t/\beta$ *for a negative binomially distributed frequency.*

Second, because the frequency enters in expression (2.6) only with its expectation $EN(t)$, it is also not necessary to calibrate a specific counting process; estimating the sample mean of the frequency suffices. As a consequence thereof, over-dispersion as modelled by the negative binomial distribution, has asymptotically no impact on the OpVAR.

2.1.4 A Refinement of the Analytic OpVAR Formula

In this Section we assume that the severity distribution $F \in \mathcal{S}$ has finite expectation $\mu := EX_k$, assuring that the aggregate operational loss has finite expectation given by $ES(t) = EN(t)\,\mu$.

Let us again consider equation (2.5), which has an interesting interpretation. Recall that the distribution tail of the aggregate loss can be written as

$$\begin{aligned}\overline{G}_t(x) &= \sum_{n=0}^{\infty} P(S(t) > x | N(t) = n)\, P(N(t) = n) \\ &= \sum_{n=0}^{\infty} \overline{F^{n*}}(x)\, P(N(t) = n)\,, \quad x \geq 0, \quad t \geq 0\,.\end{aligned}$$

Since F is subexponentially distributed, we have (see Embrechts et al. [29], Definition 1.3.3),

$$\overline{F^{n*}}(x) \sim n\,\overline{F}(x), \qquad x \to \infty\,,$$

and we can write for a fixed $t \geq 0$,

$$P(S(t) > x | N(t) = n) \sim n\overline{F}(x) \sim n\,\overline{F}(x) F^{n-1}(x) =: np(1-p)^{n-1}\,, \quad x \to \infty\,. \tag{2.11}$$

This can be interpreted as follows: Consider a Bernoulli random variable for which "success", defined as a very large loss that impacts the OpVAR figure significantly, occurs only with a very small probability $p = P(X_i > x) = \overline{F}(x)$, and "failure', defined as a small or negligible loss, occurs with probability $1 - p = F(x)$. Hence, from (2.11) we infer that for very large x the conditional aggregate loss $P(S(t) > x|N(t) = n)$ can be approximated by a binomial distribution where the number of successes that occur in n trials equals 1. This is sometimes referred to as the *single-loss approximation* of OpVAR.

A consequence of the single-loss interpretation is that approximation (2.9) usually underestimates OpVAR because in reality all (and not only one) loss events $X_k, k = 1, \ldots, n$, contribute to the aggregate loss $S(t)$ and therefore to OpVAR. Clearly, this effect is expected to be stronger for large frequency expectations and when severities are not extremely heavy-tailed.

These considerations suggest a natural adjustment of the single-loss approximation by taking the expected losses of the individual claims into account. This approach can also be motivated by the so-called large-deviation theory where one investigates the asymptotic behaviour of random sums like (2.1) when both x and n are varying together. Then, uniform convergence can be achieved when S is replaced by the centered random variable $\hat{S}(t) = S(t) - ES(t)$, see for example Proposition 8.6.4. of Embrechts et al. [29] or Klüppelberg and Mikosch [43].

Now, one property of subexponential distributions F is that (see Embrechts et al. [29], Lemma 1.3.5)

$$\lim_{x\to\infty} \frac{\overline{F}(x-y)}{\overline{F}(x)} = 1, \quad y \in \mathbb{R}\,,$$

which implies the following relationships for $\mu > 0$, $EN(t) > 0$, and a fixed $t \geq 0$

$$\begin{aligned} \overline{G}_t(x) &\sim \overline{G}_t(x + EN(t)\mu)\,, \qquad x \to \infty\,, \\ \overline{F}(x) &\sim \overline{F}(x + \mu)\,, \qquad x \to \infty\,. \end{aligned}$$

Together with (2.5) we then obtain

$$G_t(x + EN(t)\mu) \sim EN(t)F(x + \mu)\,, \qquad x \to \infty\,,$$

which finally yields the following refined approximation for OpVAR

$$\mathrm{VAR}_t(\kappa) = F^{\leftarrow}\left(1 - \frac{1-\kappa}{EN(t)}(1 + o(1))\right) + (EN(t) - 1)\mu, \quad \kappa \to 1\,. \tag{2.12}$$

In the light of equation (2.12) and in contrast to the single-loss interpretation (2.6), OpVAR can now be thought of as the result of two different components: first, exactly

one single extreme loss at very high confidence level and, second, $(EN(t)-1)$ expected losses of expected loss size μ, which we refer to as *mean correction*.

Note that the mean correction term $(EN(t)-1)\mu$ of equation (2.12) does not depend on the confidence level κ, it is just a constant. Hence, for $\kappa \to 1$ approximation (2.12) asymptotically equals the standard single loss approximation of OpVAR.

2.1.5 Models with Pareto-Like Severity Tails

We now formulate explicit results for the single loss approximation and discuss its consequences for OpVAR in the important case of Pareto-like distributed severities. With respect to the loss frequency, we confine ourself to the Poisson-LDA model as described in Example 2.1.2, and therefore we will consequently replace $EN(t)$ by $\lambda\, t$.

Pareto's law is the prototypical parametric example for a heavy tailed distribution and suitable for modelling loss severities in operational risk. This was shown by Moscadelli [54] who investigated empirical loss data collected by the Basel Committee during 2001.

The Poisson-Pareto LDA is a Standard LDA as given in Definition 2.1.1, where the loss severities $(X_k)_{k\in\mathbb{N}}$ are Pareto distributed; i.e. for parameters $\alpha, \theta > 0$

$$\overline{F}(x) = \left(1 + \frac{x}{\theta}\right)^{-\alpha}, \quad x > 0\,. \tag{2.13}$$

Motivated by extreme value methods, [54] reports results for $1/\alpha$ in a range between approximately 0.6 and 1.5, corresponding to α roughly between 0.7 and 1.7. For all such α the severity distribution has infinite variance and for $\alpha \leq 1$ even the mean value does not exist.

As a result of Theorem 2.1.12, we obtain for the OpVAR

$$\mathrm{VAR}_t(\kappa) \sim \theta \left(\frac{\lambda\, t}{1-\kappa}\right)^{1/\alpha}, \quad \kappa \to 1\,. \tag{2.14}$$

Figure 2.1.17 compares the analytical OpVAR estimate (2.14) with the results of a Monte Carlo simulation for the Pareto-LDA with different shape parameters α and $\theta = 1$. We see that the best approximation is obtained for extremely heavy-tailed data, i.e. for small values of α. Consequently, for operational loss data, our approximation should be very good.

Time Scaling in the Pareto Severity Model

A well-known formula in risk management is the square-root-of-time rule for deriving multi-period VAR values from 1-period values. This scaling law is based on the well-known property of the normal distribution, which says that the sum of n iid centered

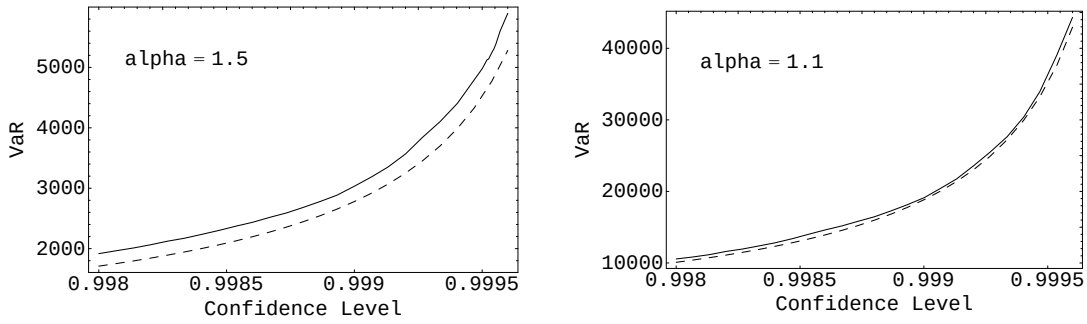

Figure 2.1.17. *Comparison of approximated OpVAR given by* (2.14) *(dashed line) and the simulated OpVAR (solid line) for the Pareto-Poisson LDA with* $\theta = 1$.

normal random variables, when scaled by $\sqrt{n}$ is again normally distributed. As a generalisation, the central limit theorem guarantees that the sum of n iid random variables with finite variance (with arbitrary distribution and centred by its mean) converges for $n \to \infty$, when scaled by $\sqrt{n}$, to a standard normal distribution. It can be shown that the central limit theorem holds also for Pareto-LDA models, when proper adjustments have been made for the random number $N(t)$ of summands; see Embrechts, Klüppelberg and Mikosch [29], Theorems 2.5.7 and 2.5.9. Note that for $\alpha < 2$ neither is scaling by $\sqrt{n}$ correct nor does the normal distribution appear as a limit for $n \to \infty$. Instead scaling has to follow a $1/\alpha$-root and the limit is a so-called stable distribution, which is much heavier-tailed than the normal law.

We are, however, not aiming at a limit law for $n \to \infty$, respectively $N(t) \to \infty$ (which means $t \to \infty$), but for a simple multi-period OpVAR based on 1-period values. Moreover, we consider approximations in the very far out tail of a heavy-tailed distribution. Consequently, a central limit argument may be misleading, and scaling with the square-root factor is even for a finite variance model not justified.

We may, however, infer from (2.14) that for all fixed $t > 0$,

$$\mathrm{VAR}_t(\kappa) \sim t^{1/\alpha}\, \mathrm{VAR}_1(\kappa)\,, \quad \kappa \to 1\,. \tag{2.15}$$

Consequently, in the case of a Pareto-LDA model, we have an α-root-of-time rule for the OpVAR. Inserting typical values for α, (2.15) implies that the threat of losses due to operational risk increases rapidly (and much faster than the outcome of the square-root-rule) when considering future time horizons. To put it simply, operational risk can be a long-term killer!

Maxima of Operational Losses

Consider OpVAR at confidence level κ and time horizon $t = 1$ year, i.e. the potential 1-year loss that is exceeded only with small probability $1 - \kappa$. From the law of large numbers we know that for large N an event with probability p occurs on average Np times in a series of N observations. Therefore, in case of yearly data, for $\kappa = 0.1\%$, OpVAR can be heuristically interpreted as the once-in-a-thousand-year event. There is, however, a different interpretation of OpVAR that is closely related to the *sample maxima* among a sequence of $N(t)$ iid loss variables X_i within a given time period $[0, t]$,

$$M(t) = \max(X_1, \ldots, X_{N(t)}), \qquad t \geq 0.$$

For the Standard LDA from Definition 2.1.1, setting $P(N(t) = n) = p_t(n)$ and defining $M_n = \max(X_1, \ldots, X_n)$ for $n \in \mathbb{N}$, we can immediately calculate the distribution function G_M of $M(t)$ for any fixed $t > 0$.

$$G_M(x) = P(M(t) \leq x) = \sum_{n=0}^{\infty} p_t(n)\, P(M_n \leq x) = \sum_{n=0}^{\infty} p_t(n)\, F^n(x)\,, \quad x \geq 0\,.$$

If the frequency follows a Poisson process with intensity $\lambda > 0$, we obtain

$$G_M(x) = \sum_{n=0}^{\infty} e^{-\lambda t}\, \frac{(\lambda t)^n}{n!}\, F^n(x) \;=\; e^{-\lambda\, t\, \overline{F}(x)}\,, \quad x \geq 0\,. \tag{2.16}$$

We now ask for the most probable value x_{mp} of the maximum, the *mode* of G_M. If F has a differentiable density f with derivative f', then also G_M has a differentiable density g_M with derivative g'_M. In this case, the mode of G_M is determined as the solution x_{mp} to

$$g'_M(x) = e^{-\lambda t\, \overline{F}(x)}\, \lambda\, t\, \left[\lambda\, t\, f^2(x) + f'(x)\right] = 0$$

and, thus, x_{mp} is the solution to

$$\lambda\, t\, f^2(x) + f'(x) = 0\,.$$

For most realistic severity distributions x_{mp} will be unique. In the important example of a Pareto distribution we have

$$x_{mp} = \theta \left[\left(\frac{\alpha\, \lambda\, t}{1 + \alpha} \right)^{1/\alpha} - 1 \right] \approx \theta \left(\frac{\alpha\, \lambda\, t}{1 + \alpha} \right)^{1/\alpha}\,. \tag{2.17}$$

Note the similarity between the VAR formula (2.14) and the right hand side of (2.17). We finally arrive at the following approximate relationship between the OpVAR at

α	κ		
	99.0 %	99.9 %	99.95 %
1.2	77	524	934
1.0	200	2 000	4 000
0.8	871	15 496	36 857

Table 2.1.18. *The factor* $\left(\frac{1+1/\alpha}{1-\kappa}\right)^{1/\alpha}$ *of equation (2.18) for* α *and* κ *in a realistic range.*

time horizon t and the most probable value of the maximum loss event during that time period for κ near 1,

$$\mathrm{VAR}_t(\kappa) \approx \left(\frac{1+1/\alpha}{1-\kappa}\right)^{1/\alpha} x_{mp}\,. \tag{2.18}$$

It is worth mentioning that this result does not depend on the frequency process, but only on the shape parameter α and the confidence level κ. For any given x_{mp}, Table 2.1.18 clearly shows the sensitivity of the corresponding OpVar of the shape parameter and the confidence level.

The question arises, whether (2.18) can be used as an alternative approximation for OpVar. Unfortunately, estimating x_{mp} by a reliable empirical method would require a vast amount of loss data, which are currently not available. The underlying data should consist of annual maximal losses for the last years, which define a histogram, from which x_{mp} can be read off. Therefore, a large amount of annual maxima would have to be collected before x_{mp} could be estimated, where presumably the iid property would be violated simply by non-stationarity in a long time series.

However, the right hand side of (2.18) can, for instance, be estimated by scenario analyses and expert-based risk assessment. An experienced risk manager may guesstimate the maximum-one-year loss caused by a single event within the next year. Annual maximal losses of previous years may guide the way. Such estimates, interpreted as the most probable value x_{mp}, then yield an expert-approximation of the OpVAR as it is required by the Basel Committee.

The Generalised Pareto Distribution Model

Needless to say, (2.14) asymptotically holds for all Pareto-like severity distributions satisfying $\overline{F}(x) \sim (x/\theta)^{-\alpha}$ as $x \to \infty$. Nevertheless, instead of directly using the general concept of regular variation for modelling the entire severity distribution, we now invoke here the so-called POT method (an achronym for "peaks over threshold"), which is a classical technique of extreme value theory. The POT method is based on the fact

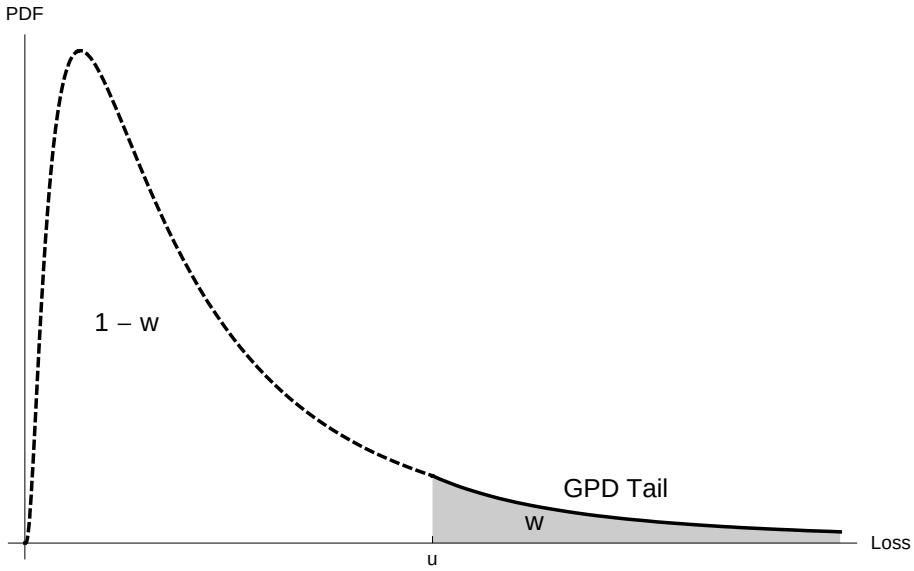

Figure 2.1.19. Probability density function of a model where high severity losses above a threshold $u > 0$ follow a GPD (solid line). The distribution body (dotted line) is differently modelled by e.g. a lognormal or Weibull distribution. The tail weight w corresponds to the shaded area under the curve.

that (under weak regularity conditions, see Embrechts, Klüppelberg & Mikosch [29], Section 3.4 for details) loss data above a high threshold u follow a generalised Pareto distribution (GPD). The body of the severity distribution is estimated by the empirical distribution function, i.e. for losses with moderate size below the threshold u any arbitrary distribution is possible. This situation is depicted in Figure 2.1.19, which schematically shows the probability density function of such a mixed severity distribution. The appropriateness of heavy-tailed GPD models in the context of operational risk has been justified very convincingly e.g. in Moscadelli [54]; an example for its practical implementation can be found in Nguyen & Ottmann [55].

In such a model, the heavy-tailed severity distribution above a high threshold $u > 0$ (i.e. high severity loss) is parameterised for $\xi, \beta, > 0$ by

$$\overline{F}(x) = w \left(1 + \xi \, \frac{x-u}{\beta}\right)^{-1/\xi}, \quad x > u > 0\,, \tag{2.19}$$

where $w = w(u) \in (0,1)$ describes the relative number of losses above u, sometimes referred to as the tail weight of F, and ξ is called the tail parameter. Note that (as $\xi > 0$ is required) the GPD is a Pareto distribution including besides the shape parameter ξ also a location and scale parameter. This proves particularly useful for statistical analyses. Moreover, in contrast to the general concept of regular variation, it makes the slowly varying function precise, namely a constant: $\overline{F}(x) \sim w(\xi/\beta)^{-1/\xi} x^{-1/\xi}$ as $x \to \infty$. As a consequence of (2.9), analytic stand-alone OpVAR is then for fixed $t > 0$

given explicitly by

$$\mathrm{VAR}_t(\kappa) \sim u + \frac{\beta}{\xi}\left[\left(\frac{w\,\lambda\,t}{1-\kappa}\right)^{\xi} - 1\right], \quad \kappa \uparrow 1 \tag{2.20}$$

where $w\,\lambda$ denotes the expected number of losses per unit time above a threshold u. Note that the threshold u enters the OpVAR approximation in two different ways. First, explicitly by the linear term on the right-hand side of (2.20). However, in practical calculations, for typical parameterisations of LDA models OpVAR at high confidence level κ will be approximated sufficiently well by neglecting u and only evaluating the second term on the right-hand side of (2.20). Second, and even more important, the threshold u implicitly enters OpVAR by influencing w and thus the effective loss frequency used in formula (2.20). Hence our result again shows how important a careful, sound and proper calibration of an operational risk models is in order not to fall into the *model and calibration risk's trap.*

2.2 Multivariate Models for Operational Risk

2.2.1 Introduction

Usually, total aggregate loss of a financial institution is not modelled directly by means of (2.1). Instead, operational losses (and thus operational risk) is attributed to different loss types and business lines. For instance, Basel II distinguishes 7 loss types and 8 business lines, yielding a matrix of 56 operational risk cells. Then, in the prototypical multivariate LDA model, for each cell $i = 1, \ldots, d$, the cumulated operational loss $S_i(t)$ up to time t is described by an aggregate loss process

$$S_i(t) = \sum_{k=1}^{N_i(t)} X_k^i\,, \quad t \geq 0\,, \tag{2.21}$$

where for each i the sequence $(X_k^i)_{k\in\mathbb{N}}$ are iid positive random variables describing the severity of each loss event, and $(N_i(t))_{t\geq 0}$ counts the number of losses in the time interval $[0, t]$, the frequency process, independent of $(X_k^i)_{k\in\mathbb{N}}$, see also Definition 2.1.1.

A required feature of any AMA model of Basel II [5] is that it allows for explicit correlations between different operational risk events. More precisely, in the context of multivariate LDA models, banks can explicitly consider the dependency between losses occurring in different event type/business line cells, and thus the question is how such a dependence structure affects a bank's total operational risk.

The bank's total operational risk is simply given by the stochastic process

$$S^+(t) := S_1(t) + S_2(t) + \cdots + S_d(t)\,, \quad t \geq 0\,. \tag{2.22}$$

The present literature suggests to model dependence between different operational risk cells by means of different concepts, which basically split into models for frequency dependence on the one hand and for severity dependence on the other hand. Some important techniques are

- modelling dependence between the number of losses $N_1(t), \ldots, N_d(t)$ occurring within $t_1 = 1$ year via correlation or copulas, see e.g. Aue & Kalkbrenner [1], Bee [3], or Frachot, Roncalli & Salomon [32],
- introducing coincident loss events by a common-shock model, see e.g. Lindskog & McNeil [46], or Powojowski, Reynolds & Tuenter [57],
- modelling dependence between the severities of those losses that occur at the same points in time, see e.g. Chavez-Demoulin, Embrechts & Nešlehová [22],
- for $t_1 = 1$ year, modelling dependence between the distribution functions of the aggregate marginal processes S_i for $i = 1, \ldots, d$ by means of distributional copulas.

For the practical implementation of LDA models, the first of the approaches above is probably most popular in the banking industry. Its main advantage is that correlation estimates between the yearly number of loss events within each operational risk cell can be quite easily calculated from empirical loss data. However, it has been reported by e.g. Aue & Kalkbrenner [1] or Bee [3] that the impact of a specific copula or the level of loss-number correlation (sometimes referred to as frequency correlation) has only little impact on the economic capital for operational risk. We will give some mathematical reasoning to support this observation later in this Chapter.

Here we suggest a different model. From a mathematical point of view, in contrast to the models proposed in Chavez-Demoulin et al. [22], we stay within the class of multivariate Lévy processes, a class of stochastic processes, which has been well studied also in the context of derivatives pricing; see e.g. Cont and Tankov [24]. In contrast to the approaches above, dependence in frequency and severity between different cells is modelled at the same time using one and the same concept, namely that of the *Lévy copula* (see e.g. Cont & Tankov [24], Kallsen and Tankov [39], Barndorff-Nielsen and Lindner [4], and also Klüppelberg & Resnick [40] for a related concept). This yields a model with comparably few parameters, which particularly in the light of spare data at hand may be a viable alternative to other, more complex models.

Our model has the same advantage as distributional copulas: the dependence structure between different cells can be separated from the univariate problem, i.e. the marginal loss processes S_i for $i = 1, \ldots, d$. Consequently, with a rather transparent dependence model, we are able to model the possibility of coincident losses occurring in different cells.

2.2.2 Lévy Processes, Tail Integrals, and Lévy Copulas

Since operational risk is only concerned with losses, we restrict ourselves to Lévy processes admitting only positive jumps in every component, hereafter called *spectrally positive Lévy processes*. As a consequence of their independent and stationary increments, Lévy processes can be represented by the *Lévy-Khintchine formula*, which for a d-dimensional spectrally positive Lévy processes S without drift and Gaussian component simplifies to

$$E(e^{i\,(z,S_t)}) = \exp\left\{t \int_{\mathbb{R}_+^d} (e^{i(z,x)} - 1)\,\Pi(dx)\right\}, \quad z \in \mathbb{R}^d\,,$$

where Π is a measure on $\mathbb{R}_+^d = [0,\infty)^d$, called the *Lévy measure* of S and $(x,y) := \sum_{i=1}^d x_i y_i$ for $x, y \in \mathbb{R}^d$ denotes the inner product.

Whereas the dependence structure in a Gaussian model is well-understood, dependence in the Lévy measure Π is much less obvious. Nevertheless, as Π is independent of t, it suggests itself for modelling the dependence structure between the components of S, leading to the concept of Lévy copulas.

Distributional copulas are multivariate distribution functions with uniform marginals. They are used for dependence modelling within the context of Sklar's theorem, which states that any multivariate distribution with continuous marginals can be transformed into a multivariate distribution with uniform marginals. This concept exploits the fact that distribution functions have values only in $[0,1]$. In contrast, Lévy measures are in general unbounded on $\mathbb{R}^d$ and may have a non-integrable singularity at 0, which causes problems for the copula idea. Within the class of spectrally positive compound Poisson models, the Lévy measure of the cell process S_i is given by $\Pi_i([0,x)) = \lambda_i P(X^i \le x)$ for $x \in [0,\infty)$. It follows that the Lévy measure is a finite measure with total mass $\Pi_i([0,\infty)) = \lambda_i$ and, therefore, is in general not a probability measure. Since we are interested in extreme operational losses, we prefer (as is usual in the context of general Lévy process theory) to define a copula for the tail integral. Although we shall mainly work with compound Poisson processes, we formulate definitions and some results and examples for the slightly more general case of spectrally positive Lévy processes.

Definition 2.2.1. [Tail integral] *Let X be a spectrally positive Lévy process in $\mathbb{R}^d$ with Lévy measure Π. Its* tail integral *is the function $\overline{\Pi} : [0,\infty]^d \to [0,\infty]$ satisfying for $x = (x_1,\ldots,x_d)$,*

(1) $\overline{\Pi}(x) = \Pi([x_1,\infty) \times \cdots \times [x_d,\infty))\,, \quad x \in [0,\infty)^d\,,$
where $\overline{\Pi}(0) = \lim_{x_1 \downarrow 0,\ldots,x_d \downarrow 0} \Pi([x_1,\infty) \times \cdots \times [x_d,\infty))$
(this limit is finite if and only if X is compound Poisson);

(2) $\overline{\Pi}$ *is equal to 0, if one of its arguments is* ∞*;*

(3) $\overline{\Pi}(0,\dots,x_i,0,\dots,0) = \overline{\Pi}_i(x_i)$ *for* $(x_1,\dots,x_d) \in \mathbb{R}^d_+$*, where* $\overline{\Pi}_i(x_i) = \Pi_i([x_i,\infty))$ *is the tail integral of component* i*.*

Definition 2.2.2. [Lévy copula] *A d-dimensional* Lévy copula *of a spectrally positive Lévy process is a measure defining function* $\widehat{C} : [0,\infty]^d \to [0,\infty]$ *with marginals, which are the identity functions on* $[0,\infty]$*.*

The following is Sklar's theorem for spectrally positive Lévy processes.

Theorem 2.2.3. [Cont and Tankov [24], Theorem 5.6] *Let* $\overline{\Pi}$ *denote the tail integral of a d-dimensional spectrally positive Lévy process, whose components have Lévy measures* $\Pi_1,\dots,\Pi_d$*. Then there exists a Lévy copula* $\widehat{C} : [0,\infty]^d \to [0,\infty]$ *such that for all* $x_1,\dots,x_d \in [0,\infty]$

$$\overline{\Pi}(x_1,\dots,x_d) = \widehat{C}(\overline{\Pi}_1(x_1),\dots,\overline{\Pi}_d(x_d)). \tag{2.23}$$

If the marginal tail integrals $\overline{\Pi}_1,\dots,\overline{\Pi}_d$ *are continuous, then this Lévy copula is unique. Otherwise, it is unique on* $\mathrm{Ran}\overline{\Pi}_1 \times \cdots \times \mathrm{Ran}\overline{\Pi}_d$*.*
Conversely, if $\widehat{C}$ *is a Lévy copula and* $\overline{\Pi}_1,\dots,\overline{\Pi}_d$ *are marginal tail integrals of spectrally positive Lévy processes, then* (2.23) *defines the tail integral of a d-dimensional spectrally positive Lévy process and* $\overline{\Pi}_1,\dots,\overline{\Pi}_d$ *are tail integrals of its components.*

The following two important Lévy copulas model extreme dependence structures.

Example 2.2.4. [Complete (positive) dependence]
Let $S(t) = (S_1(t),\dots,S_d(t))$, $t \geq 0$, be a spectrally positive Lévy process with marginal tail integrals $\overline{\Pi}_1,\dots,\overline{\Pi}_d$. Since all jumps are positive, the marginal processes can never be negatively dependent. Complete dependence corresponds to a Lévy copula

$$\widehat{C}_{\|}(x) = \min(x_1,\dots,x_d)\,,$$

implying for the tail integral of S

$$\overline{\Pi}(x_1,\dots,x_d) = \min(\overline{\Pi}_1(x_1),\dots,\overline{\Pi}_d(x_d))$$

with all mass concentrated on $\{x \in [0,\infty)^d : \overline{\Pi}_1(x_1) = \cdots = \overline{\Pi}_d(x_d)\}$. □

Example 2.2.5. [Independence]
Let $S(t) = (S_1(t),\dots,S_d(t))$, $t \geq 0$, be a spectrally positive Lévy process with marginal tail integrals $\overline{\Pi}_1,\dots,\overline{\Pi}_d$. The marginal processes are independent if and only if they never jump together, i.e. the Lévy measure Π of S can be decomposed into

$$\Pi(A) = \Pi_1(A_1) + \cdots + \Pi_d(A_d), \qquad A \in [0,\infty)^d \tag{2.24}$$

with $A_1 = \{x_1 \in [0,\infty) : (x_1, 0, \ldots, 0) \in A\}, \ldots, A_d = \{x_d \in [0,\infty) : (0, \ldots, x_d) \in A\}$. Obviously, the support of Π are the coordinate axes. Equation (2.24) implies for the tail integral of S

$$\overline{\Pi}(x_1, \ldots, x_d) = \overline{\Pi}_1(x_1)\, 1_{x_2=\cdots=x_d=0} + \cdots + \overline{\Pi}_d(x_d)\, 1_{x_1=\cdots=x_{d-1}=0}\,.$$

It follows that the independence copula for spectrally positive Lévy processes is given by

$$\widehat{C}_\perp(x) = x_1\, 1_{x_2=\cdots=x_d=\infty} + \cdots + x_d\, 1_{x_1=\cdots=x_{d-1}=\infty}\,.$$

□

2.2.3 The Lévy Copula Model

We now want to motivate our approach for modelling multivariate operational risk by using Lévy copulas. Needless to say, multivariate OpVAR is still in its infancy and so far the question regarding the right model cannot be answered only by statistical analysis because reliable data are often still not available. There exists, however, a model-theoretic rationale for our approach, which we want to briefly explain.

In accordance with the findings of a recent survey of the Basel Committee on Banking Supervision [6] about AMA practices at financial services firms, we assume that within each cell i the loss frequency process N_i in (2.21) follows a homogeneous Poisson process with rate $\lambda_i > 0$, in particular, for every fixed $t > 0$,

$$P(N_i(t) = n)) = e^{-\lambda_i t}\, \frac{(\lambda_i t)^n}{n!}, \quad n \in \mathbb{N}_0\,.$$

All loss severities within cell i are independent and have the same severity distribution function $F_i(x) = P(X^i \leq x)$ for $x \in [0,\infty)$. Then the aggregate loss (2.21) for each cell constitutes a compound Poisson process and, hence, is a Lévy process (the only Lévy process with piecewise constant sample paths).

As a matter of fact, the definition of 56 different cells based on seven loss event types and eight business lines as suggested by the Basel Committee [5] is quite arbitrary. Actually, many banks are using a less dimensional cell matrix, which basically means e.g. that they apply a compound Poisson LDA to a *union* of some of the Basel II cells. Such a procedure, however, is only consistent with the overall framework of a compound Poisson model, if we require that every additive conjunction of different cells constitutes an aggregate loss process that again is a univariate compound Poisson process; in particular, with common severity distribution $F_{i+j}(\cdot)$ and frequency λ_{i+j}, i.e. for such $i \neq j$,

$$S_i(\cdot) + S_j(\cdot) := S_{i+j}(\cdot) \quad \in \text{ compound Poisson processes}\,. \tag{2.25}$$

Or, put another way, a natural requirement of a multivariate LDA model should be that it does not directly depend on the structure of the event type/business line matrix and thus on the business organization. As a direct consequence of the standard LDA model, rigorously applied to several operational risk cells we obtain an "invariance principle" any mathematical OpRisk model has to satisfy.

As we will show below, (2.25) holds true, whenever the vector of all marginal aggregate loss processes $(S_1(t), \ldots, S_d(t))_{t\geq 0}$ constitutes a d-dimensional compound Poisson process. Therefore, the problem is how the different one-dimensional compound Poisson processes $S_i(\cdot) = \sum_{k=1}^{N_i(\cdot)} X_k^i$ can be combined to form a d-dimensional compound Poisson process $S(t) = (S_1(t), \ldots, S_d(t))_{t\geq 0}$ with, in general, dependent components. If we are only interested in one fixed time point, say $t_1 = 1$ year, we can consider $(S_1(t_1), \ldots, S_d(t_1))$ simply as a vector of static random variables with distribution functions $G_{i,t_1}(x_i) = P(S_i(t_1) \leq x_i)$ for $i = 1, \ldots, d$ and $x_i \geq 0$. Now, it is well-known that the dependence structure of a multidimensional random vector can be disentangled from its marginals by introducing a distributional copula. More precisely, Sklar's now famous theorem states that any multivariate distribution with continuous marginals can be transformed into a distribution with uniform marginals. Therefore, choosing an appropriate distributional copula C at t_1 we could write for the aggregate loss distribution function

$$P(S_1(t_1) \leq x_1, \ldots, S_d(t_1) \leq x_d) = G_{t_1}(x_1, \ldots, x_d) = C(G_{1,t_1}(x_1), \ldots, G_{d,t_1}(x_d)) \,.$$

However, switching on time-dependence again in the marginals, the related process $(S(t))_{t\geq 0}$, will in general not be a multivariate compound Poisson process and thus, contradictory to our requirement (2.25), the multivariate model may not be invariant under a re-design of the cell matrix.

Another kind of model frequently used is the following one, or at least a variation of it. Fix a time horizon $t > 0$, and again model the accumulated losses of each operational risk cell $i = 1, \ldots, d$ by a compound Poisson random variable $S_i(t)$. Then, model the dependence between both the loss sizes in different cells and the dependence between the frequency variables $N_i(t)$ by appropriate distributional copulas, where for the latter one has to take the discreteness of these variables into account. Considering this model as a dynamic model in time, it again does not constitute a multivariate compound Poisson model as it is required by (2.25), but instead leads outside the well-studied class of Lévy processes. This can be easily seen as follows: since a Poisson process jumps with probability 0 at any fixed time $s > 0$, we have for any jump time s of $N_j(\cdot)$ that $P(\Delta N_i(s) = 1) = 0$ for $i \neq j$, hence any two of such processes never jump at the same time. However, as described in Section 2.2.2, dependence in multivariate compound Poisson processes—as in every multivariate Lévy process—

means dependence in the jump measure, i.e. the possibility of joint jumps. Finally, from a statistical point of view such a model requires a large number of parameters, which, given the sparsity of data in combination with the task of estimating high quantiles, will be almost impossible to fit.

Hence, the question is how dependence between different risk cells can be established by at the same time also conserving the compound Poisson property of the multivariate process over time. Clearly, the answer is given by Theorem 2.2.3. Since the multivariate tail integral

$$\overline{\Pi}(x_1, \dots, x_d) = \widehat{C}(\overline{\Pi}_1(x_1), \dots, \overline{\Pi}_d(x_d)), \qquad x \in [0, \infty]^d .$$

defines a multivariate compound Poisson process $(S_1, \dots, S_d)$ (cf. Cont & Tankov [24], Theorem 3.1), we conclude that the multivariate compound Poisson model in which the marginal compound Poisson processes S_i are coupled via a Lévy copula satisfies (2.25). Hence, our Lévy copula model can be considered as the most natural and straight-forward extension of the well-known univariate standard compound Poisson LDA model to several dependent operational risk cells.

We now formulate the multivariate compound Poisson model.

Definition 2.2.6. [Multivariate compound Poisson model] *The* multivariate compound Poisson model *consists of:*

(1) Cell processes.
All operational risk cells, indexed by $i = 1, \dots, d$, are described by an univariate compound Poisson model with aggregate loss process S_i, subexponential severity distribution function F_i and Poisson intensity $\lambda_i > 0$, respectively.

(2) Dependence structure.
The dependence between different cells is modelled by a Lévy copula. More precisely, let $\overline{\Pi}_i : [0, \infty) \to [0, \infty)$ be the tail integral associated with S_i, i.e. $\overline{\Pi}_i(\cdot) = \lambda_i \overline{F}_i(\cdot)$ for $i = 1, \dots, d$, and let $\widehat{C} : [0, \infty)^d \to [0, \infty)$ be a Lévy copula. Then

$$\overline{\Pi}(x_1, \dots, x_d) = \widehat{C}(\overline{\Pi}_1(x_1), \dots, \overline{\Pi}_d(x_d))$$

defines the tail integral of the d-dimensional compound Poisson process $S = (S_1, \dots, S_d)$.

(3) Total aggregate loss process.
The bank's total aggregate loss process *is defined as*

$$S^+(t) = S_1(t) + S_2(t) + \cdots + S_d(t), \qquad t \geq 0$$

with tail integral

$$\overline{\Pi}^+(z) = \Pi(\{(x_1, \dots, x_d) \in [0, \infty)^d : \sum_{i=1}^{d} x_i \geq z\}), \quad z \geq 0 . \tag{2.26}$$

The following result states an important property of the multivariate compound Poisson model.

Proposition 2.2.7. *Consider the multivariate compound Poisson model of Definition 2.2.6. Its total aggregate loss process S^+ is compound Poisson with frequency parameter*

$$\lambda^+ = \lim_{z \downarrow 0} \overline{\Pi}^+(z)$$

and severity distribution

$$F^+(z) = 1 - \overline{F}^+(z) = 1 - \frac{\overline{\Pi}^+(z)}{\lambda^+}, \quad z \geq 0\,.$$

Proof. Projections of Lévy processes are Lévy processes. For any compound Poisson process with intensity $\lambda > 0$ and only positive jumps with distribution function F, the tail integral of the Lévy measure is given by $\overline{\Pi}(x) = \lambda \overline{F}(x)$, $x > 0$. Consequently, $\lambda = \overline{\Pi}(0)$ and $\overline{F}(x) = \overline{\Pi}(x)/\lambda$. We apply this relation to the Lévy process S^+ and obtain the total mass λ^+ of S^+, which ensures that S^+ is compound Poisson with the parameters as stated. □

Definition 2.2.8. [Total OpVAR] *Consider the multivariate compound Poisson model of Definition 2.2.6. Then,* total OpVAR *up to time t at confidence level κ is the κ-quantile of the total aggregate loss distribution $G_t^+(\cdot) = P(S^+(t) \leq \cdot)$:*

$$\mathrm{VAR}_t^+(\kappa) = G_t^{+\leftarrow}(\kappa)\,, \quad \kappa \in (0,1)\,,$$

with $G_t^{+\leftarrow}(\kappa) = \inf\{z \in \mathbb{R} :\ G_t^+(z) \geq \kappa\}$ for $0 < \kappa < 1$.

2.2.4 A Bivariate Example

A bivariate model is particularly useful to illustrate how dependence modelling via Lévy copulas works. Therefore, we now focus on two operational risk cells (index $i = 1, 2$) with frequency parameters λ_i and severity distributions F_i so that the marginal tail integrals are given by $\overline{\Pi}_i(\cdot) = \lambda_i \overline{F}_i(\cdot)$ as explained in (2.2.1).

Before we consider the so-called Clayton Lévy copula in greater detail, we briefly mention how in general parametric Lévy copulas can be constructed. The following Proposition shows how Lévy copulas can be derived from distributional copulas and, therefore, ensures that there exists a wide variety of potentially useful Lévy copulas (see Cont & Tankov [24], Proposition 5.5).

Proposition 2.2.9. *Let C be a two-dimensional distributional copula and $f : [0,1] \to [0,\infty]$ an increasing convex function. Then*

$$\widehat{C}(u,v) = f(C(f^{-1}(u), f^{-1}(v)))\,, \quad u,v \in [0,\infty)\,,$$

defines a two-dimensional positive Lévy copula.

Example 2.2.10. [Clayton Lévy copula]
Henceforth, the dependence structure between two operational risk cells shall be modelled by a Clayton Lévy copula, which is similar to the well-known Clayton copula for distribution functions and parameterized by $\delta > 0$ (see Cont & Tankov [24], Example 5.5):

$$\widehat{C}_\delta(u,v) = (u^{-\delta} + v^{-\delta})^{-1/\delta}\,, \quad u,v \geq 0\,.$$

□

We use this copula for mainly two reasons:

- This copula covers the whole range of positive dependence: For $\delta \to 0$ we obtain independence of the marginal processes given by $\widehat{C}_\perp(u,v) = u1_{v=\infty} + v1_{u=\infty}$, and losses in different cells never occur at the same time. For $\delta \to \infty$ we get the complete positive dependence Lévy copula given by $\widehat{C}_\parallel(u,v) = \min(u,v)$, and losses always occur at the same points in time. By varying δ, the cell dependence changes smoothly between these two extremes. However, it should be stressed that $\widehat{C}_\parallel$ only leads to completely dependent processes, if the marginal tail integrals are continuous. This is relevant for dependence of compound Poisson processes, which create by definition a discontinuity of the tail integral in 0, and this has to be discussed in detail below.

- The Clayton copula has a quite simple parametrization (only one parameter δ) and, as we will see later, it allows for precise analytical calculations regarding total aggregated OpVAR. Therefore, we consider the Clayton Lévy copula as particularly useful to investigate different dependence scenarios.

The question what dependence between operational risk cells actually means is not trivial, and with this regard we already mentioned some common modelling techniques in the introduction. Some of them are quite flexible and sophisticated, however, also very complex and often difficult to parameterize. In contrast, Lévy copulas and tail integrals together lead to a quite natural interpretation of dependence in the context of the multivariate compound Poisson model of Definition 2.2.6. To see this, we start with the following decomposition of the marginal tail integral $\overline{\Pi}_1$ for $x_1 \geq 0$,

$$\overline{\Pi}_1(x_1) \;=\; \Pi([x_1,\infty) \times [0,\infty))\,, \qquad x_1 \geq 0\,,$$

which basically measures the number of jumps larger than x_1 in the first component, regardless of the jumps in the second component (i.e. whether jumps with arbitrary size occur or not). This together with Definition (2.2.1) and (2.23) leads to

$$\begin{aligned}
\overline{\Pi}_1(x_1) &= \Pi([x_1,\infty)\times[0,\infty)) \\
&= \Pi([x_1,\infty)\times\{0\}) + \lim_{x_2\downarrow 0}\Pi([x_1,\infty)\times[x_2,\infty)) \\
&= \Pi([x_1,\infty)\times\{0\}) + \lim_{x_2\downarrow 0}\overline{\Pi}(x_1,x_2) \\
&= \Pi([x_1,\infty)\times\{0\}) + \lim_{x_2\downarrow 0}\widehat{C}(\overline{\Pi}_1(x_1),\overline{\Pi}_2(x_2)) \\
&= \Pi([x_1,\infty)\times\{0\}) + \widehat{C}(\overline{\Pi}_1(x_1),\lambda_2) \\
&=: \overline{\Pi}_{\perp 1}(x_1) + \overline{\Pi}_{\parallel 1}(x_1)\,, \qquad x_1\geq 0\,,
\end{aligned} \tag{2.27}$$

where $\overline{\Pi}_{\perp 1}(\cdot)$ describes losses that occur in the first cell only without any simultaneous loss in the second cell. In contrast, $\overline{\Pi}_{\parallel 1}(\cdot)$ describes the expected number of losses per unit time above x_1 in the first cell that coincide with losses of arbitrary size in the second cell (occurring with frequency λ_2). Similarly we may write

$$\overline{\Pi}_2(x_2) =: \overline{\Pi}_{\perp 2}(x_2) + \overline{\Pi}_{\parallel 2}(x_2)\,, \qquad x_2\geq 0\,. \tag{2.28}$$

Connected with these decompositions of the marginal tail integrals, we obtain the following split of the cells' aggregate loss processes (the time parameter t is dropped for simplicity):

$$\begin{aligned}
S_1 &= S_{\perp 1} + S_{\parallel 1} = \sum_{k=1}^{N_{\perp 1}} X^1_{\perp k} + \sum_{l=1}^{N_{\parallel}} X^1_{\parallel l}\,, \\
S_2 &= S_{\perp 2} + S_{\parallel 2} = \sum_{m=1}^{N_{\perp 2}} X^2_{\perp m} + \sum_{l=1}^{N_{\parallel}} X^2_{\parallel l}\,,
\end{aligned} \tag{2.29}$$

where $S_{\parallel 1}$ and $S_{\parallel 2}$ describe the aggregate losses of cell 1 and 2, respectively, that are generated by "common shocks", and $S_{\perp 1}$ and $S_{\perp 2}$ are independent loss processes. Note that apart from $S_{\parallel 1}$ and $S_{\parallel 2}$, all compound Poisson processes on the right-hand side of (2.29) are mutually independent.

So far all these considerations are regardless of a specific Lévy copula. However, it is clear that the relative "weights" of $S_{\parallel 1}$ and $S_{\parallel 2}$ compared to $S_{\perp 1}$ and $S_{\perp 2}$ directly reflect the dependence structure and so all their parameters can be written in terms of the Lévy copula. In the following we disentangle the dependence introduced by a Lévy copula and describe precisely, what it results for loss times and loss severities.

Simultaneous loss times. We begin with the frequency of simultaneous losses, which may in principle be arbitrarily small and, therefore, are given by

$$\lim_{x_1,x_2\downarrow 0}\overline{\Pi}(x_1,x_2) = \widehat{C}(\lambda_1,\lambda_2) = \lim_{x\downarrow 0}\overline{\Pi}_{\parallel 2}(x) = \lim_{x\downarrow 0}\overline{\Pi}_{\parallel 1}(x) =: \lambda_{\parallel}\,.$$

On one hand, in the case of independence, losses never occur at the same points in time; on the other hand, for complete positive dependence we have $\widehat{C}_{\|}(u,v) = \min(u,v)$. Obviously,

$$0 \leq \lambda_{\|} \leq \min(\lambda_1, \lambda_2)\,, \tag{2.30}$$

In particular, maximum dependence is reached if all losses in the cell with the smaller number of expected losses coincide with losses of the other cell.

A widespread concept for modelling dependence in operational risk is that of the frequency correlation between two aggregate loss processes. In the compound Poisson process approach (recall Sklar's theorem for Lévy copulas), the correlation between the number of losses $N_1(t)$ and $N_2(t)$ up to time t associated with S_1 and S_2, respectively, is simply given by

$$\rho(N_1(t), N_2(t)) \;=\; \frac{\mathrm{cov}(N_1(t), N_2(t))}{\sqrt{\mathrm{var}(N_1(t))\,\mathrm{var}(N_2(t))}} = \frac{\lambda_{\|}}{\sqrt{\lambda_1\,\lambda_2}}\,. \tag{2.31}$$

Obviously, for $\lambda_1 > \lambda_2$ the maximum possible frequency correlation is $\rho_{\max} = \sqrt{\lambda_2/\lambda_1}$. So, for two cells with $\lambda_1 \gg \lambda_2$ this frequency correlation is restricted to relatively low values.

Independent loss times. We now turn to the frequencies of the independent loss processes $S_{\perp 1}$ and $S_{\perp 2}$. Using (2.27) and (2.28) we can write their tail integrals for $x_1, x_2 \geq 0$ as

$$\begin{aligned} \overline{\Pi}_{\perp 1}(x_1) &= \overline{\Pi}_1(x_1) - \overline{\Pi}_{\|1}(x_1) = \overline{\Pi}_1(x_1) - \widehat{C}(\overline{\Pi}_1(x_1), \lambda_2)\,, \\ \overline{\Pi}_{\perp 2}(x_2) &= \overline{\Pi}_2(x_2) - \overline{\Pi}_{\|2}(x_2) = \overline{\Pi}_2(x_2) - \widehat{C}(\lambda_1, \overline{\Pi}_2(x_2))\,, \end{aligned} \tag{2.32}$$

so that

$$\lambda_{\perp 1} = \lim_{x \downarrow 0} \overline{\Pi}_{\perp 1}(x) = \lambda_1 - \lambda_{\|}\,, \qquad \lambda_{\perp 2} = \lim_{x \downarrow 0} \overline{\Pi}_{\perp 2}(x) = \lambda_2 - \lambda_{\|}\,. \tag{2.33}$$

Example 2.2.11. [Continuation of Example 2.2.10]
Recall the Clayton Lévy copula

$$\widehat{C}_{\delta}(u,v) = (u^{-\delta} + v^{-\delta})^{-1/\delta}\,, \quad u, v \geq 0$$

for $\delta \in (0, \infty)$. In this case we calculate the frequency of simultaneous jumps as

$$\lambda_{\|} \;=\; (\lambda_1^{-\delta} + \lambda_2^{-\delta})^{-1/\delta}\,, \tag{2.34}$$

and the frequency correlation is given by

$$\rho(N_1(t), N_2(t)) \;=\; \frac{\lambda_{\|}}{\sqrt{\lambda_1\,\lambda_2}} \;=\; \frac{(\lambda_1^{-\delta} + \lambda_2^{-\delta})^{-1/\delta}}{\sqrt{\lambda_1\,\lambda_2}}\,.$$

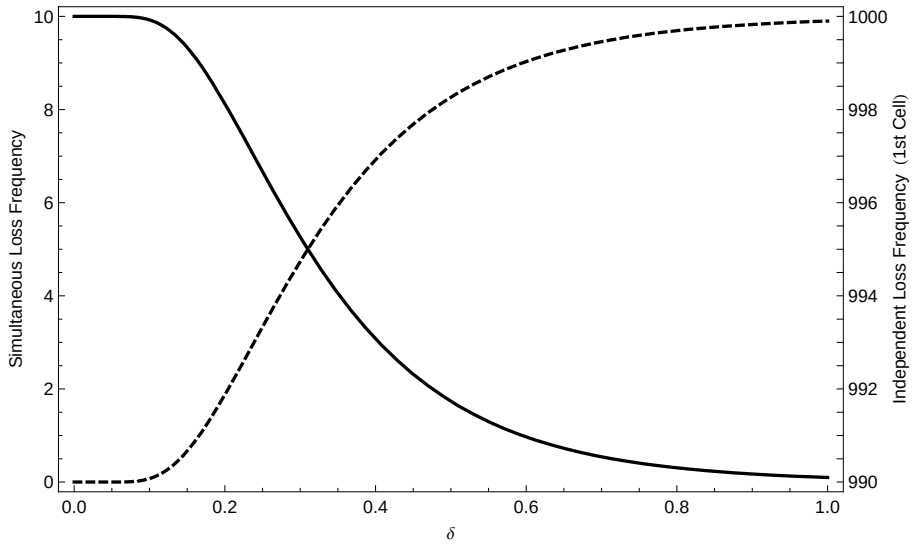

Figure 2.2.12. *Example how the cells' loss frequencies are controlled by the Clayton Lévy copula for* $\lambda_1 = 1000$ *and* $\lambda_2 = 10$*. Left axis: frequency* $\lambda_{\|}$ *of the simultaneous loss processes* $S_{\|1}$ *and* $S_{\|2}$ *as a function of the Lévy Clayton copula parameter* δ *(dashed line). Right axis: frequency* $\lambda_{\perp 1}$ *of the independent loss process* $S_{\perp 1}$ *of the first cell as a function of the Lévy Clayton copula parameter* δ *(solid line).*

We show that, although the Clayton Lévy copula tends to $\widehat{C}_\infty(u,v) = \min(u,v)$ (i.e. the complete dependence copula) as $\delta \to \infty$, the processes S_1 and S_2 are not completely dependent. Take two cells with $\lambda_1 = 1000$ and $\lambda_2 = 10$, then in Figure 2.2.12 both $\lambda_{\|}$ and $\lambda_{\perp 1}$ are plotted as a function of the Lévy Clayton copula parameter δ. One can see that even for $\delta \to \infty$ there are non-simultaneous losses occurring in only the first cell with intensity $\lambda_{\perp 1} = \lambda_1 - \lambda_2 = 990$. Furthermore, the maximal possible correlation in this model is $\rho_{\max} = 10\ \%$. □

Summarising we can say that since dependence of the loss frequency processes only influence the number of expected losses, it follows that frequency correlation for every model has only a very restricted impact on OpVAR.

Simultaneous loss severities and their distributional copula. Also the severity distributions of $X_{\|}^1$ and $X_{\|}^2$ as well as their dependence structure are determined by the Lévy copula. To see this, define the joint survival function as

$$\overline{F}_{\|}(x_1, x_2) := P(X_{\|}^1 > x_1, X_{\|}^2 > x_2) = \frac{1}{\lambda_{\|}} \widehat{C}(\overline{\Pi}_1(x_1), \overline{\Pi}_2(x_2)) \tag{2.35}$$

with marginals

$$\overline{F}_{\|1}(x_1) = \lim_{x_2 \downarrow 0} \overline{F}_{\|}(x_1, x_2) = \frac{1}{\lambda_{\|}} \widehat{C}(\overline{\Pi}_1(x_1), \lambda_2) \tag{2.36}$$

$$\overline{F}_{\|2}(x_2) = \lim_{x_1 \downarrow 0} \overline{F}_{\|}(x_1, x_2) = \frac{1}{\lambda_{\|}} \widehat{C}(\lambda_1, \overline{\Pi}_2(x_2)) \,. \tag{2.37}$$

To explicitly extract the dependence structure between the severities of simultaneous losses $X_{\|}^1$ and $X_{\|}^2$ we use the concept of a distributional survival copula. In general, if $\overline{F}(x_1, x_2)$ is a joint survival function with continuous marginals $\overline{F}_i(x_i)$, there exists a unique survival copula $\widehat{S}$ such that $\overline{F}(x_1, x_2) = \widehat{S}(\overline{F}_1(x_1), \overline{F}_2(x_2))$ giving together with the rhs of (2.35) the relation between Lévy copula and survival copula.

Example 2.2.13. [Continuation of Examples 2.2.10 and 2.2.11]
For the Clayton copula a straight-forward calculation using (2.35)–(2.37) shows that the survival copula $\widehat{S}_\delta$ for the tail severity distributions $\overline{F}_{\|1}(\cdot)$ and $\overline{F}_{\|2}(\cdot)$ is the well-known distributional Clayton copula; i.e. for $\delta > 0$,

$$\widehat{S}_\delta(u, v) = (u^{-\delta} + v^{-\delta} - 1)^{-1/\delta}, \qquad 0 \le u, v \le 1 \,.$$

Consequently, the distribution functions $F_{\|1}$ and $F_{\|2}$ (and thus the simultaneous losses $X_{\|}^1$ and $X_{\|}^2$) are linked by a copula C_δ that is related to $\widehat{S}_\delta$ via

$$\begin{aligned} C_\delta(u, v) &= \widehat{S}_\delta(1-u, 1-v) + u + v - 1 \\ &= ((1-u)^{-\delta} + (1-v)^{-\delta} - 1)^{-1/\delta} + u + v - 1, \ 0 \le u, v \le 1 \,. \end{aligned} \tag{2.38}$$

Specifically, for $\delta \to \infty$ we obtain the complete dependence distributional copula $C_{\|}(u, v) = \min(u, v)$, implying comonotonicity of the simultaneous losses $X_{\|}^1$ and $X_{\|}^2$. □

Independent loss severities and their distributions. We obtain from (2.32) for the severity distributions of non-simultaneous losses

$$\begin{aligned} \overline{F}_{\perp 1}(x_1) &= \frac{\lambda_1}{\lambda_1^{\perp}} \overline{F}_1(x_1) - \frac{1}{\lambda_1^{\perp}} \widehat{C}(\lambda_1 \overline{F}_1(x_1), \lambda_2) \,, \\ \overline{F}_{\perp 2}(x_1) &= \frac{\lambda_2}{\lambda_2^{\perp}} \overline{F}_2(x_2) - \frac{1}{\lambda_2^{\perp}} \widehat{C}(\lambda_1, \lambda_2 \overline{F}_2(x_2)) \,. \end{aligned}$$

Let us summarise the interpretation of multivariate operational risk as it is suggested by our model.

- Dependence between different cells is solely due to the occurrence of simultaneous loss events in different cells.
- There are two types of losses: independent ones, which happen in one single cell

only and dependent ones, which happen simultaneously. The severity distributions of dependent losses are themselves coupled by a distributional copula, which can be derived from the Lévy copula (2.35), e.g. (2.38) in the case of a Clayton Lévy copula. In particular, it follows that in general $F_{\|1}$ and $F_{\|2}$ are different from F_1 and F_2, respectively. Also $F_{\perp 1}$ and $F_{\perp 2}$ are different from F_1 and F_2 as well as from $F_{\|1}$ and $F_{\|2}$, respectively.

- Independence of different cells means that their losses never happen at the same time, whereas complete dependence is equivalent to losses that always occur together.

This pattern is depicted in Figures 2.3.33-2.3.38 at the end of this Chapter where for the Clayton Lévy copula sample paths and occurrence times of the bivariate compound Poisson model are simulated for different copula parameters of $\delta = 0.3, 1$ and 7. For the purpose of a clearer illustration of the dependence structure, both cells are assumed to have identical frequencies of $\lambda_1 = \lambda_2 = 10$. With respect to the severity, we used Pareto distributed severities with tail parameters $\alpha_1 = 1.2$ and $\alpha_2 = 2$ and scale parameters $\theta_1 = \theta_2 = 1$ in Figures 2.3.33-2.3.35, whereas in Figures 2.3.36-2.3.38 we used lognormally distributed severities with parameters $\mu_1 = \mu_2 = 5$ and $\sigma_1 = \sigma_2 = 1.5$. According to (2.34), the percentage average number of common losses related to the different δ used are 10%, 50%, and 90%. The simulation is based on Algorithm 6.15 of Cont & Tankov [24], which can be used for arbitrary severity distributions as well.

2.3 Approximating Multivariate OpVAR

2.3.1 Overview

In this Section we turn to the quantification of total operational loss encompassing all operational risk cells and, therefore, we focus on the total aggregate loss process S^+ defined in Definition 2.2.6. Our goal is to provide some general insight to multivariate operational risk and to find out, how different dependence structures (modelled by Lévy copulas) affect OpVAR.

In the context of multivariate operational risk, the estimate of the OpVAR of a single cell (e.g. obtained by means of the single loss approximation od Theorem 2.1.15) is often referred to as the cell's *stand alone* OpVAR. Then, a first approximation to the bank's total OpVAR is obtained by summing up all different stand alone VAR numbers. Indeed, the Basel committee requires banks to sum up all their different operational risk estimates unless sound and robust correlation estimates are available; cf. [5], paragraph 669(d). Moreover, this "simple-sum VAR" is often interpreted as an upper bound for total OpVAR, with the implicit understanding that every other (realistic) cell dependence model necessarily reduces overall operational risk.

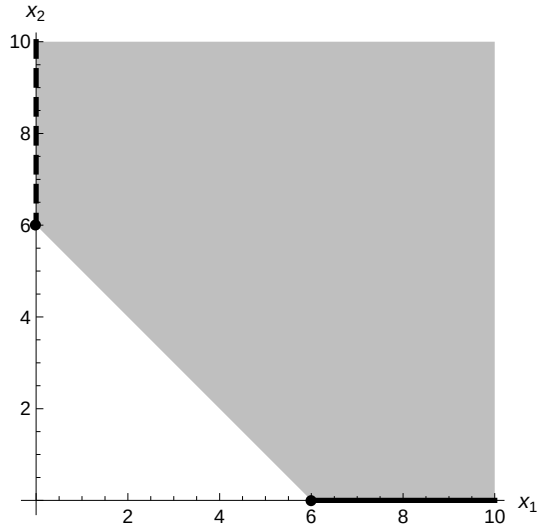

Figure 2.3.1. *Decomposition of the domain of the tail integral* $\overline{\Pi}^+(z)$ *for* $z = 6$ *into a simultaneous loss part* $\overline{\Pi}^+_{\|}(z)$ *(grey area) and independent parts* $\overline{\Pi}_{\perp 1}(z)$ *(solid black line) and* $\overline{\Pi}_{\perp 2}(z)$ *(dashed black line).*

However, as is well recognised, simple-sum VAR may even underestimate total OpVAR when severity data is heavy-tailed, which in practice it is, see e.g. Moscadelli [54]. Therefore, to obtain a more accurate result, one needs more general models for multivariate operational risk, such as the Lévy copula model we discuss here.

Though operational risk is usually modelled by separating several event type/business line cells, we already know from Proposition 2.2.7 that a bank's total OpVAR can be thought of as effectively being a compound Poisson process with risk inter-arrival times being exponentially distributed with finite mean λ^+ and loss severities, which are independent and identically distributed with distribution function $F^+(\cdot)$. The associated tail integral $\overline{\Pi}^+$ of this total aggregate loss process S^+ is given by (2.26) For $d = 2$ we can write

$$\overline{\Pi}^+(z) \;=\; \overline{\Pi}_{\perp 1}(z) + \overline{\Pi}_{\perp 2}(z) + \overline{\Pi}^+_{\|}(z)\,, \qquad z \geq 0\,, \tag{2.39}$$

where $\overline{\Pi}_{\perp 1}(\cdot)$ and $\overline{\Pi}_{\perp 2}(\cdot)$ are the independent parts defined in (2.27)-(2.28) and

$$\overline{\Pi}^+_{\|}(z) \;=\; \Pi(\{(x_1, x_2) \in (0,\infty)^2 : x_1 + x_2 \geq z\})\,, \qquad z \geq 0\,,$$

describes the dependent part due to simultaneous loss events. This is depicted in Figure 2.3.1, where the support of $\overline{\Pi}^+_{\|}(\cdot)$ is shaded in orange, and the support of $\overline{\Pi}_{\|1}(\cdot)$ and $\overline{\Pi}_{\|2}(\cdot)$ are solid black and dashed black lines, respectively.

Finally, if F^+ is subexponentially distributed, we can apply the single loss approximation of Theorem 2.1.12 to estimate total OpVAR. In combination with the fact that

(at first order) the loss frequency processes affects OpVAR only through the number of expected losses, we conclude that total OpVAR is asymptotically only impacted by the expected number of total loss events, $EN^+(t) = EN_1(t) + \cdots + EN_d(t) = \lambda^+$ for $t \geq 0$. It follows that frequency correlation for every model has only a very restricted impact on OpVAR and does not deserve much attention.

It has also been observed that total OpVAR is presumably affected by business volume at time. Actually, this belief is a basic assumption both for the Basic Indicator Approach and the Standardized Approach of Basel II [5], where capital charges for operational risk are scaled by gross income. This idea can be included in the above multivariate compound Poisson model by adapting the frequency. For each $i = 1, \ldots, d$, we leave the severity models X^i untouched, as well as the independence of the severities of the (no longer homogeneous) Poisson process N_i. However, instead of a constant intensity, we model a time-dependent frequency depending on business volume: in each single cell we replace $EN_i(t) = \lambda_i t$ by $EN_i(t) = \int_0^t \lambda_i(s)\, ds$ for $t > 0$. This is then plugged into formula (2.6) for the stand-alone OpVAR of cell i and consequently into $EN^+(t) = \lambda^+ t$ for total OpVAR.

2.3.2 Losses Dominant in One Cell

Before we discuss different kinds of Lévy copula dependence structures, we formulate a very general result for the situation, where the losses in one cell are regularly varying and dominate all others. Indeed the situation of the model is such that it covers arbitrary dependence structures, including also the practitioner's models described above.

Assume for fixed $t > 0$ for each cell model a compound Poisson random variable. Dependence is introduced by an arbitrary correlation or copula for $(N_1(\cdot), \ldots, N_d(\cdot))$ and an arbitrary copula between the severity distributions $F_1(\cdot) = P(X^1 \leq \cdot), \ldots, F_d(\cdot) = P(X^d \leq \cdot)$. Recall that the resulting model $(S_1(t), \ldots, S_d(t))_{t \geq 0}$ does NOT constitute a bivariate compound Poisson process and so is not captured by the multivariate compound Poisson model of Definition 2.2.6. We want to calculate an approximation for the tail $P(S_1(t) + S_2(t) > x)$ for large x and total OpVAR for high levels κ. We formulate the result in arbitrary dimension.

Theorem 2.3.2. [Dominating cell OpVAR] *For fixed $t > 0$ let $S_i(t)$ for $i = 1, \ldots, d$ have compound Poisson distributions. Assume that $\overline{F}_1 \in \mathcal{R}_{-\alpha}$ for $\alpha > 0$. Let $\rho > \alpha$ and suppose that $E[(X^i)^\rho] < \infty$ for $i = 2, \ldots, d$. Then regardless of the dependence*

structure between $(S_1(t),\dots,S_d(t))$,

$$\begin{aligned} P(S_1(t)+\cdots+S_d(t)>x) &\sim EN_1(t)\,P(X^1>x)\,, \quad x\to\infty\,, \\ \mathrm{VAR}_t^+(\kappa) &\sim F_1^{\leftarrow}\Big(1-\frac{1-\kappa}{EN_1(t)}\Big)=\mathrm{VAR}_t^1(\kappa)\,, \quad \kappa\uparrow 1\,. \end{aligned} \tag{2.40}$$

Proof. Consider $d=2$. Note first that

$$\begin{aligned} &\frac{P(S_1(t)+S_2(t)>x)}{P(X^1>x)} \\ &= \sum_{k,m=1}^{\infty} P(N_1(t)=k,N_2(t)=m)\frac{P\Big(\sum_{i=1}^k X_i^1+\sum_{j=1}^m X_j^2>x\Big)}{P\Big(\sum_{i=1}^k X_i^1>x\Big)}\frac{P\Big(\sum_{i=1}^k X_i^1>x\Big)}{P\Big(X^1>x\Big)}\,. \end{aligned} \tag{2.41}$$

We have to find conditions such that we can interchange the limit for $x\to\infty$ and the infinite sum. This means that we need uniform estimates for the two ratios on the right-hand side for $x\to\infty$. We start with an estimate for the second ratio: Lemma 1.3.5 of Embrechts et al. [29] applies giving for arbitrary $\varepsilon>0$ and all $x>0$ a finite positive constant $K(\varepsilon)$ so that

$$\frac{P\Big(\sum_{i=1}^k X_i^1>x\Big)}{P(X^1>x)}\le K(\varepsilon)(1+\varepsilon)^k\,.$$

For the first ratio we proceed as in the proof of Lemma 2 of Klüppelberg, Lindner and Maller [42]. For arbitrary $0<\delta<1$ we have

$$\frac{P\Big(\sum_{i=1}^k X_i^1+\sum_{j=1}^m X_j^2>x\Big)}{P\Big(\sum_{i=1}^k X_i^1>x\Big)} \le \frac{P\Big(\sum_{i=1}^k X_i^1>x(1-\delta)\Big)}{P\Big(\sum_{i=1}^k X_i^1>x\Big)}+\frac{P\Big(\sum_{j=1}^m X_j^2>x\delta\Big)}{P\Big(\sum_{i=1}^k X_i^1>x\Big)} \tag{2.42}$$

Regular variation of the distribution of X^1 implies regular variation of the distribution of $\sum_{i=1}^k X_i^1$ with the same index $-\alpha$. We write for the first term

$$\frac{P\Big(\sum_{i=1}^k X_i^1>x(1-\delta)\Big)}{P\Big(X^1>x(1-\delta)\Big)}\frac{P\Big(X^1>x(1-\delta)\Big)}{P\Big(X^1>x\Big)}\frac{P\Big(X^1>x\Big)}{P\Big(\sum_{i=1}^k X_i^1>x\Big)}\,.$$

For the first ratio we use the same estimate as above and obtain for all $x>0$ the upper bound $K'(\varepsilon)(1+\varepsilon)^k$. For the second ratio, using the so-called Potter bounds (cf. Theorem 1.5.6 (iii) of Bingham, Goldie and Teugels [7]), for every chosen constants

$a > 0, A > 1$, we obtain an upper bound $A(1-\delta)^{-(\alpha+a)}$ uniformly for all $x \geq x_0 \geq 0$. The third ratio is less or equal to 1 for all k and x.
As the denominator of the second term of the rhs of (2.42) is regularly varying, it can be bounded below by $x^{-(\alpha+\rho')}$ for some $0 < \rho' < \rho - \alpha$. By Markov's inequality, we obtain for the numerator

$$P\Big(\sum_{j=1}^{m} X_j^2 > x\delta\Big) \leq (x\delta)^{-\rho}\, E\Big[\big(\sum_{j=1}^{m} X_j^2\big)^{\rho}\Big]\,.$$

The so-called c_ρ-inequality (see e.g. Loéve [48], p. 157) applies giving

$$E\big[(\sum_{j=1}^{m} X_j^2)^{\rho}\big] \leq m\, c_\rho E(X_j^2)^{\rho}$$

for $c_\rho = 1$ or $c_\rho = 2^{\rho-1}$, according as $\rho \leq 1$ or $\rho > 1$. We combine these estimates and obtain in (2.41) for $x \geq x_0 > 0$,

$$\begin{aligned}
&\frac{P(S_1(t) + S_2(t) > x)}{P(X^1 > x)} \\
&\leq \sum_{k,m=1}^{\infty} P(N_1(t) = k, N_2(t) = m) \qquad (2.43)\\
&\quad \Big(K'(\varepsilon)(1+\varepsilon)^k\, A(1-\delta)^{-(\alpha+a)} + x^{\alpha+\rho'}(x\delta)^{-\rho}\, m\, c_\rho\, E[(X_j^2)^{\rho}]\Big)\, K(\varepsilon)(1+\varepsilon)^k\,.
\end{aligned}$$

Now note that $x^{\alpha+\rho'-\rho}$ tends to 0 as $x \to \infty$. Furthermore, we have

$$\begin{aligned}
\sum_{k,m=0}^{\infty} P(N_1(t) = k, N_2(t) = m) &= 1\,,\\
\sum_{k,m=1}^{\infty} P(N_1(t) = k, N_2(t) = m)\, k &= \sum_{k=1}^{\infty} P(N_1(t) = k)\, k = EN_k(t) < \infty\,.
\end{aligned}$$

Consequently, the rhs of 2.43 converges. By Pratt's Lemma (see e.g. Resnick [60], Ex. 5.4.2.4), we can interchange limit and infinite sum on the rhs of (2.41) and obtain

$$\lim_{x\to\infty} \frac{P(S_1(t) + S_2(t) > x)}{P(X^1 > x)} = \sum_{k=1}^{\infty} P(N_1(t) = k)\, k = EN_1(t)\,.$$

The result for $d > 2$ follows by induction.
Approximation (2.40) holds by Theorem 2.1.15(1). □

Within the context of multivariate compound Poisson models, the proof of this result simplifies. Moreover, since a possible singularity of the tail integral in 0 is of no consequence, it even holds for all spectrally positive Lévy processes. We formulate this as a corollary.

Corollary 2.3.3. *Consider a multivariate spectrally positive Lévy process and suppose that* $\overline{\Pi}_1 \in \mathcal{R}_{-\alpha}$. *Furthermore, assume that for all* $i = 2, \dots, d$ *the integrability condition*

$$\int_{x \geq 1} x^{\rho} \, \Pi_i(dx) < \infty \tag{2.44}$$

for some $\rho > \alpha$ *is satisfied. Then*

$$\lim_{z \to \infty} \frac{\overline{\Pi}^{+}(z)}{\overline{\Pi}_1(z)} = 1 \,. \tag{2.45}$$

Moreover,

$$\mathrm{VAR}_t^{+}(\kappa) \sim \mathrm{VAR}_t^{1}(\kappa) \,, \qquad \kappa \uparrow 1 \,, \tag{2.46}$$

i.e. total OpVAR is asymptotically dominated by the stand alone OpVAR of the first cell.

Proof. We first show that (2.45) holds. From equation (2.44) it follows that for $i = 2, \dots, d$

$$\lim_{z \to \infty} z^{\rho} \, \overline{\Pi}_i(z) = 0 \,. \tag{2.47}$$

Since $\alpha < \rho$, we obtain from regular variation for some slowly varying function L, invoking (2.47),

$$\lim_{z \to \infty} \frac{\overline{\Pi}_i(z)}{\overline{\Pi}_1(z)} = \lim_{z \to \infty} \frac{z^{\rho} \overline{\Pi}_i(z)}{z^{\rho - \alpha} L(z)} = 0 \,, \quad i = 2, \dots, d \,,$$

because the numerator tends to 0 and the denominator to ∞. (Recall that $z^{\varepsilon} L(z) \to \infty$ as $z \to \infty$ for all $\varepsilon > 0$ and $L \in \mathcal{R}_0$.)

We proceed by induction. For $d = 2$ we have by the decomposition as in (2.42)

$$\overline{\Pi}_2^{+}(z) := \overline{\Pi}^{+}(z) \leq \overline{\Pi}_1(z(1-\varepsilon)) + \overline{\Pi}_2(z\,\varepsilon), \quad z > 0, \quad 0 < \varepsilon < 1.$$

It then follows that

$$\limsup_{z \to \infty} \frac{\overline{\Pi}_2^{+}(z)}{\overline{\Pi}_1(z)} \leq \lim_{z \to \infty} \frac{\overline{\Pi}_1(z(1-\varepsilon))}{\overline{\Pi}_1(z)} + \lim_{z \to \infty} \frac{\overline{\Pi}_2(z\,\varepsilon)}{\overline{\Pi}_1(z\,\varepsilon)} \frac{\overline{\Pi}_1(z\,\varepsilon)}{\overline{\Pi}_1(z)} = (1-\varepsilon)^{-\alpha} \,. \tag{2.48}$$

Similarly, $\overline{\Pi}_2^{+}(z) \geq \overline{\Pi}_1((1+\varepsilon)z)$ for every $\varepsilon > 0$. Therefore,

$$\liminf_{z \to \infty} \frac{\overline{\Pi}_2^{+}(z)}{\overline{\Pi}_1(z)} \geq \lim_{z \to \infty} \frac{\overline{\Pi}_1((1+\varepsilon)z)}{\overline{\Pi}_1(z)} = (1+\varepsilon)^{-\alpha} \,. \tag{2.49}$$

Assertion (2.45) follows for $\overline{\Pi}_2^+$ from (2.48) and (2.49). This implies that $\overline{\Pi}_2^+ \in \mathcal{R}_\alpha$. Now replace $\overline{\Pi}_1$ by $\overline{\Pi}_2^+$ and $\overline{\Pi}_2^+$ by $\overline{\Pi}_3^+$ and proceed as above to obtain (2.45) for general dimension d. Finally, Theorem 2.1.15(1) applies giving (2.46). □

Hence, for arbitrary dependence structures, when the severity of *one* cell has regularly varying tail dominating those of all other cells, total OpVAR is tail-equivalent to the OpVAR of the dominating cell. This implies that the bank's total loss at high confidence levels is likely to be due to one big loss occurring in one cell rather than an accumulation of losses of different cells regardless of the dependence structure.

Example 2.3.4. [Dominating cell in the GPD model]
Consider the multivariate compound Poisson model of Definition 2.2.6 with arbitrary Lévy copula and assume that large losses above a high threshold $u > 0$ in the first cell have a GPD tail with tail weight w and parameters $\xi, \beta > 0$, given by (2.19). Assume that $\lambda > 0$ denotes the frequency in the first cell and its severity distribution F_1 as above is tail-dominant to all other cell severities (which apart from that can have arbitrary distribution functions), i.e. $\overline{F}_i(x)/\overline{F}_1(x) \to 0$ as $x \to \infty$ for all $i = 2, \dots, d$. Then, according to Theorem 2.3.2, S^+ is a compound Poisson process with

$$P(S^+(t) > z) \sim \lambda t \, \overline{F}_1(z) = \lambda t \, w \left(1 + \xi \frac{z-u}{\beta}\right)^{-1/\xi}, \quad z \to \infty,$$

and total OpVAR is asymptotically given by

$$\mathrm{VAR}_t^+(\kappa) \sim \mathrm{VAR}_t^1(\kappa) \sim u + \frac{\beta}{\xi}\left[\left(\frac{w\lambda t}{1-\kappa}\right)^{\xi} - 1\right] \sim \frac{\beta}{\xi}\left(\frac{w\lambda t}{1-\kappa}\right)^{\xi}, \qquad \kappa \uparrow 1.$$

□

Note however, that from our equivalence results of Proposition 2.1.9 and Theorem 2.1.15 we know that this is not a general property of any completely dependent compound Poisson model with arbitrary subexponential severity distribution F_i. We shall see in Example 2.3.10 below that the following does NOT hold in general for $x \to \infty$ (equivalently $\kappa \uparrow 1$):

$$\overline{F}_i(x) = o(\overline{F}_1(x)) \quad \Longrightarrow \quad \mathrm{VAR}_t^i(\kappa) = o(\mathrm{VAR}_t^1(\kappa)), \quad i = 2, \dots, d.$$

2.3.3 Multivariate Compound Poisson Model with Completely Dependent Cells

The assumptions of Theorem 2.3.2 may in many cases be quite realistic, however, it is of course possible that two ore more cells' severity distributions are tail equivalent, and

then total OpVAR is expected to depend on the dependence structure, i.e. the Lévy copula. Here and in the next sections we study two very basic multivariate compound Poisson models in more detail, namely the completely dependent and the independent one. Despite their extreme dependence structure, both models provide interesting and valuable insight into multivariate operational risk.

Consider a multivariate compound Poisson model and assume that its cell processes $S_i,\ i = 1, \dots, d$, are completely positively dependent. In the context of Lévy processes this means that they always jump together, implying that also the expected number of jumps per unit time of all cells, i.e. the intensities λ_i, must be equal,

$$\lambda := \lambda_1 = \cdots = \lambda_d \,. \tag{2.50}$$

The severity distributions F_i, however, can be different. Indeed, from Example 2.2.4 we infer that in the case of complete dependence, all Lévy mass is concentrated on

$$\{(x_1, \dots, x_d) \in [0,\infty)^d : \overline{\Pi}_1(x_1) = \cdots = \overline{\Pi}_d(x_d)\} \,,$$

or, equivalently,

$$\{(x_1, \dots, x_d) \in [0,\infty)^d : F_1(x_1) = \cdots = F_d(x_d)\} \,. \tag{2.51}$$

Until further notice, we assume for simplicity that all severity distributions F_i are strictly increasing and continuous so that $F_i^{-1}(q)$ exists for all $q \in [0,1)$. Together with (2.51), we can express the tail integral of S^+ in terms of the marginal $\overline{\Pi}_1$.

$$\begin{aligned} \overline{\Pi}^+(z) &= \Pi(\{(x_1, \dots, x_d) \in [0,\infty)^d : \sum_{i=1}^d x_i \geq z\}) \\ &= \Pi_1(\{x_1 \in [0,\infty) : x_1 + \sum_{i=2}^d F_i^{-1}(F_1(x_1)) \geq z\}) \,, \quad z \geq 0 \,. \end{aligned}$$

Set $H(x_1) := x_1 + \sum_{i=2}^d F_i^{-1}(F_1(x_1))$ for $x_1 \in [0,\infty)$ and note that it is strictly increasing and therefore invertible. Hence,

$$\overline{\Pi}^+(z) = \Pi_1(\{x_1 \in [0,\infty) : x_1 \geq H^{-1}(z)\}) = \overline{\Pi}_1\left(H^{-1}(z)\right) \,, \quad z \geq 0 \,. \tag{2.52}$$

Now we can derive an asymptotic expression for total OpVAR.

Theorem 2.3.5. [OpVAR for the completely dependent compound Poisson model] *Consider a multivariate compound Poisson model with completely dependent cell processes $S_1, \dots, S_d$ and strictly increasing and continuous severity distributions F_i. Then, S^+ is compound Poisson with parameters*

$$\lambda^+ = \lambda \qquad \textit{and} \qquad \overline{F}^+(\cdot) = \overline{F}_1\left(H^{-1}(\cdot)\right) \,. \tag{2.53}$$

If furthermore $\overline{F}^+ \in \mathcal{S} \cap (\mathcal{R} \cup \mathcal{R}_\infty)$*, total OpVAR is asymptotically given by*

$$\mathrm{VAR}_t^+(\kappa) \sim \sum_{i=1}^{d} \mathrm{VAR}_t^i(\kappa)\,, \quad \kappa \uparrow 1\,, \tag{2.54}$$

where $\mathrm{VAR}_t^i(\kappa)$ *denotes the stand alone OpVAR of cell* i*.*

Proof. Expression (2.53) immediately follows from (2.50) and (2.52),

$$\lambda^+ = \lim_{z\to 0} \overline{\Pi}^+(z) = \lim_{z\to 0} \lambda\,\overline{F}_1\left(H^{-1}(z)\right) = \lambda\,\overline{F}_1\left(\lim_{z\to 0} H^{-1}(z)\right) = \lambda\,.$$

If $\overline{F}^+ \in \mathcal{S} \cap (\mathcal{R} \cup \mathcal{R}_\infty)$, we may use (2.9) and the definition of H to obtain

$$\begin{aligned}\mathrm{VAR}_t^+(\kappa) \;&\sim\; H\left[F_1^{-1}\left(1-\frac{1-\kappa}{\lambda\,t}\right)\right] \;=\; F_1^{-1}\left(1-\frac{1-\kappa}{\lambda\,t}\right) + \cdots + F_d^{-1}\left(1-\frac{1-\kappa}{\lambda\,t}\right)\\ &\sim\; \mathrm{VAR}_t^1(\kappa) + \cdots + \mathrm{VAR}_t^d(\kappa), \qquad \kappa \uparrow 1\,.\end{aligned}$$

□

Theorem 2.3.5 states that for the completely dependent compound Poisson model, total asymptotic OpVAR is simply the sum of the asymptotic stand alone cell OpVARs. Recall that this is similar to the new proposals of Basel II, where the standard procedure for calculating capital charges for operational risk is just the simple-sum VAR. Or stated another way, regulators implicitly assume complete dependence between different cells, meaning that losses within different business lines or risk categories always happen at the same instants of time. This is often considered as the worst case scenario, which, however, in the heavy-tailed case can be grossly misleading.

To illustrate the results so far, let us begin with a simple example, which describes another regime for completely dependent cells.

Example 2.3.6. [Identical severity distributions]
Assume that all cells have identical severity distributions, i.e. $F := F_1 = \ldots = F_d$. In this case we have $H(x_1) = d\,x_1$ for $x_1 \geq 0$ and, therefore,

$$\overline{\Pi}^+(z) = \lambda\,\overline{F}\left(\frac{z}{d}\right)\,, \quad z \geq 0\,.$$

If furthermore $\overline{F} \in \mathcal{S} \cap (\mathcal{R} \cup \mathcal{R}_\infty)$, it follows that $\overline{F}^+(\cdot) = \overline{F}(\cdot\,/d)$ is, and we obtain

$$\mathrm{VAR}_t^+(\kappa) \sim d\,\overline{F}\left(1-\frac{1-\kappa}{\lambda\,t}\right)\,, \quad \kappa \uparrow 1\,.$$

□

We can derive very precise asymptotics in the case of dominating regularly varying severities. In this case we can expect from Theorem 2.3.2 that only the dominant cells contribute to total OpVAR.

Proposition 2.3.7. *Assume that the conditions of Theorem 2.3.5 hold. Assume further that $\overline{F}_1 \in \mathcal{R}_{-\alpha}$ with $\alpha > 0$ and that for all $i = 2, \ldots, d$ there exist $c_i \in [0, \infty)$ such that*

$$\lim_{x \to \infty} \frac{\overline{F}_i(x)}{\overline{F}_1(x)} = c_i \,. \tag{2.55}$$

Assume that $c_i \neq 0$ for $2 \leq i \leq b \leq d$ and $c_i = 0$ for $i \leq b+1 \leq d$. For $\overline{F}_1(x) = x^{-\alpha} L(x)$, $x \geq 0$, let $\widetilde{L}$ be the function as in Theorem 2.1.15(ii). Then

$$\mathrm{VAR}_t^+(\kappa) \;\sim\; \sum_{i=1}^{b} c_i^{1/\alpha}\, \mathrm{VAR}_t^1(\kappa) \;\sim\; \sum_{i=1}^{b} c_i^{1/\alpha} \left(\frac{\lambda\, t}{1-\kappa}\right)^{1/\alpha} \widetilde{L}\left(\frac{1}{1-\kappa}\right), \quad \kappa \uparrow 1 \,.$$

Proof. From Theorem 2.1.15(ii) we know that

$$\mathrm{VAR}_t^1(\kappa) \sim \left(\frac{\lambda\, t}{1-\kappa}\right)^{1/\alpha} \widetilde{L}\left(\frac{1}{1-\kappa}\right), \quad \kappa \uparrow 1 \,,$$

where $\widetilde{L}\left(\frac{1}{1-\cdot}\right) \in \mathcal{R}_0$. Note: If all $c_i = 0$ holds for $i = 2, \ldots, d$ then Corollary 2.3.3 applies. So assume that $c_i \neq 0$ for $2 \leq i \leq b$. From (2.55) and Resnick [60], Proposition 0.8(vi), we get $F_i^{\leftarrow}(1 - \frac{1}{z}) \sim c_i^{1/\alpha} F_1^{\leftarrow}(1 - \frac{1}{z})$ as $z \to \infty$ for $i = 1, \ldots, d$. This yields for $x_1 \to \infty$

$$\begin{aligned} H(x_1) &= x_1 + \sum_{i=2}^{d} F_i^{-1}(1 - \overline{F}_1(x_1)) \\ &= x_1 + \sum_{i=2}^{d} c_i^{1/\alpha} F_1^{-1}\left(1 - \overline{F}_1(x_1)\right)(1 + o_i(1)) \\ &= x_1 \sum_{i=1}^{b} c_i^{1/\alpha}\,(1 + o(1)) \,, \end{aligned}$$

where we have $c_1 = 1$. Defining $C := \sum_{i=1}^{b} c_i^{1/\alpha}$, then $H(x_1) \sim C x_1$ as $x_1 \to \infty$, and hence $H^{-1}(z) \sim z/C$ as $z \to \infty$, which implies by (2.52) and regular variation of $\overline{F}_1$

$$\overline{\Pi}^+(z) \,=\, \overline{\Pi}_1(H^{-1}(z)) \,\sim\, \lambda\, \overline{F}_1(z/C) \,\sim\, \lambda\, C^\alpha\, \overline{F}_1(z) \,, \quad z \to \infty \,.$$

Obviously, $\overline{F}^+(z) = C^\alpha\, \overline{F}_1(z) \in \mathcal{R}_{-\alpha}$ and Theorem 2.3.5 applies. By (2.10) together with the fact that all summands from index $b+1$ on are of lower order, (2.54) reduces

to

$$\begin{aligned}\mathrm{VAR}_t^+(\kappa) &\sim F_1^{\leftarrow}\left(1-\frac{1-\kappa}{\lambda\, t}\right)+\cdots+F_b^{\leftarrow}\left(1-\frac{1-\kappa}{\lambda\, t}\right)\\ &\sim F_1^{\leftarrow}\left(1-\frac{1-\kappa}{\lambda\, t\, C^\alpha}\right)\\ &\sim \sum_{i=1}^{b} c_i^{1/\alpha}\left(\frac{\lambda\, t}{1-\kappa}\right)^{1/\alpha}\widetilde{L}\left(\frac{1}{1-\kappa}\right), \quad \kappa\uparrow 1\,.\end{aligned}$$

□

Important examples of Proposition 2.3.7 are the Pareto cases.

Example 2.3.8. [Pareto distributed severities, complete dependence]
Consider a multivariate compound Poisson model of Definition 2.2.6 with completely dependent cells and Pareto distributed severities as in (2.13). Then we obtain for the c_i

$$\lim_{x\to\infty}\frac{\overline{F}_i(x)}{\overline{F}_1(x)}=\left(\frac{\theta_i}{\theta_1}\right)^\alpha, \quad i=1,\ldots,b, \quad \text{and} \quad \lim_{x\to\infty}\frac{\overline{F}_i(x)}{\overline{F}_1(x)}=0, \quad i=b+1,\ldots,d,$$

for some $1\le b\le d$. This, together with Proposition 2.3.7 leads to

$$\overline{F}^+(z)\sim\left(\sum_{i=1}^{b}\frac{\theta_i}{\theta_1}\right)^\alpha\left(1+\frac{z}{\theta_1}\right)^{-\alpha}\sim\left(\sum_{i=1}^{b}\theta_i\right)^\alpha z^{-\alpha}, \quad z\to\infty\,.$$

Finally, from (2.20) and (2.54) we obtain total OpVAR as

$$\mathrm{VAR}_t^+(\kappa)\sim\sum_{i=1}^{b}\mathrm{VAR}_t^i(\kappa)\sim\sum_{i=1}^{b}\theta_i\left(\frac{\lambda\, t}{1-\kappa}\right)^{1/\alpha}, \qquad \kappa\uparrow 1\,.$$

□

Example 2.3.9. [GPD distributed severities, complete dependence]
Consider a multivariate compound Poisson model of Definition 2.2.6 with completely dependent cells and assume that within each cell large losses above some high threshold $u_i>0$ have a GPD tail given by (2.19). Similarly as above we obtain

$$\lim_{x\to\infty}\frac{\overline{F}_i(x)}{\overline{F}_1(x)}=\frac{w_i}{w_1}\left(\frac{\beta_i}{\beta_1}\right)^{1/\xi}, \ i=1,\ldots,b, \qquad \lim_{x\to\infty}\frac{\overline{F}_i(x)}{\overline{F}_1(x)}=0, \ i=b+1,\ldots,d \quad (2.56)$$

and thus $0\le\xi_i<\xi,\ i=b+1,\ldots,d$. The case $\xi_i=0$ corresponds to a tail lighter than any Pareto tail. Using Proposition 2.3.7, we finally arrive at

$$\overline{F}^+(z)\sim w^+\left(1+\xi\frac{z-u_1}{\beta_1}\right)^{-1/\xi}\sim\left(\sum_{i=1}^{b}\frac{\beta_i}{\xi}\, w_i^\xi\right)^{1/\xi} z^{-1/\xi}, \quad z\to\infty\,,$$

with $w^+ = (\sum_{i=1}^b \frac{\beta_i}{\beta_1} w_i^\xi)^{1/\xi}$. Total OpVAR is asymptotically given by

$$\begin{aligned} \mathrm{VAR}^+_{\|t}(\kappa) &\sim u_1 + \frac{\beta_1}{\xi}\left[\left(\frac{w^+\lambda t}{1-\kappa}\right)^\xi - 1\right] \\ &\sim \sum_{i=1}^b \frac{\beta_i}{\xi}\left(\frac{w_i\lambda\, t}{1-\kappa}\right)^\xi \sim \sum_{i=1}^b \mathrm{VAR}^i_t(\kappa)\,, \qquad \kappa\uparrow 1\,. \end{aligned} \tag{2.57}$$

□

Finally, note that in Theorem 2.3.5 the assumption with regard to the severity is less strict than in Theorem 2.3.2 about the dominating cell OpVAR, which requires a regularly varying severity distribution for the dominant cell. We conclude this session with another example of Theorem 2.3.5, which moreover shows that Theorem 2.3.2 does not hold for any general dominating tail.

Example 2.3.10. [Weibull severities]
Consider a bivariate compound Poisson model with completely dependent cells and assume that the cells' severities are Weibull distributed according to

$$\overline{F}_1(x) = \exp(-\sqrt{x/2}) \qquad \text{and} \qquad \overline{F}_2(x) = \exp(-\sqrt{x})\,, \quad x>0\,. \tag{2.58}$$

Note that $\overline{F}_{1,2} \in \mathcal{S}\cap\mathcal{R}_\infty$. Equation (2.58) immediately implies that $\overline{F}_2(x) = o(\overline{F}_1(x))$. We find that $H(x_1) = \frac{3}{2}\,x_1$ implying that $\overline{F}^+ \in \mathcal{S}\cap\mathcal{R}_\infty$, since

$$\overline{F}^+(z) = \exp(-\sqrt{z/3})\,, \quad z>0\,. \tag{2.59}$$

It is remarkable that in this example the total severity (2.59) is heavier tailed than the stand alone severities (2.58), i.e. $F_{1,2}(x) = o(F^+(x))$ as $x\to\infty$. However, from

$$\mathrm{VAR}^1_t(\kappa) \sim 2\left[\ln\left(\frac{1-\kappa}{\lambda\, t}\right)\right]^2 \quad \text{and} \quad \mathrm{VAR}^2_t(\kappa) \sim \left[\ln\left(\frac{1-\kappa}{\lambda\, t}\right)\right]^2, \qquad \kappa\uparrow 1\,,$$

we find that the stand alone VARs are of the same order of magnitude:

$$\lim_{\kappa\uparrow 1}\frac{\mathrm{VAR}^2_t(\kappa)}{\mathrm{VAR}^1_t(\kappa)} = \frac{1}{2}\,.$$

Nevertheless, equation (2.54) of Theorem 2.3.5 still holds,

$$\mathrm{VAR}^+_t(\kappa) \sim 3\left[\ln\left(\frac{1-\kappa}{\lambda\, t}\right)\right]^2 = \mathrm{VAR}^1_t(\kappa) + \mathrm{VAR}^2_t(\kappa)\,, \qquad \kappa\uparrow 1\,.$$

□

2.3.4 Multivariate Compound Poisson Model with Independent Cells

Let us now turn to a multivariate compound Poisson model where the cell processes $S_i,\ i = 1, \dots, d$, are independent and so never jump together. Therefore, we may write the tail integral of S^+ as

$$\overline{\Pi}^+(z) = \Pi([z,\infty) \times \{0\} \times \cdots \times \{0\}) + \cdots + \Pi(\{0\} \times \cdots \times \{0\} \times [z,\infty))\,, \quad z \geq 0\,.$$

Recall from Example 2.2.5 that in the case of independence all mass of the Lévy measure Π is concentrated on the axes. Hence,

$$\begin{aligned}
\Pi([z,\infty) \times \{0\} \times \cdots \times \{0\}) &= \Pi([z,\infty) \times [0,\infty) \times \cdots \times [0,\infty)),\\
\Pi(\{0\} \times [z,\infty) \times \cdots \times \{0\}) &= \Pi([0,\infty) \times [z,\infty) \times \cdots \times [0,\infty)),\\
&\vdots \qquad\qquad \vdots\\
\Pi(\{0\} \times \{0\} \times \cdots \times [z,\infty)) &= \Pi([0,\infty) \times [0,\infty) \times \cdots \times [z,\infty)),
\end{aligned}$$

and we obtain

$$\begin{aligned}
\overline{\Pi}^+(z) &= \Pi([z,\infty) \times [0,\infty) \times \cdots \times [0,\infty)) + \cdots + \Pi([0,\infty) \times \cdots \times [0,\infty) \times [z,\infty))\\
&= \overline{\Pi}_1(z) + \cdots + \overline{\Pi}_d(z)\,. \qquad (2.60)
\end{aligned}$$

Now we are in the position to derive an asymptotic expression for total OpVAR in the case of independent cells.

Theorem 2.3.11. [OpVAR for the independent compound Poisson model] *Consider a multivariate compound Poisson model with independent cell processes $S_1, \dots, S_d$. Then S^+ defines a one-dimensional compound Poisson model with parameters*

$$\lambda^+ = \lambda_1 + \cdots + \lambda_d \quad \textit{and} \quad \overline{F}^+(z) = \frac{1}{\lambda^+}\left[\lambda_1 \overline{F}_1(z) + \cdots + \lambda_d \overline{F}_d(z)\right]\,, \quad z \geq 0\,. \qquad (2.61)$$

If $\overline{F}_1 \in \mathcal{S} \cap (\mathcal{R} \cup \mathcal{R}_\infty)$ and for all $i = 2, \dots, d$ there exist $c_i \in [0,\infty)$ such that

$$\lim_{x\to\infty} \frac{\overline{F}_i(x)}{\overline{F}_1(x)} = c_i\,, \qquad (2.62)$$

then, setting $C_\lambda = \lambda_1 + c_2\lambda_2 + \cdots + c_d\lambda_d$, total OpVAR can be approximated by

$$\mathrm{VAR}_t^+(\kappa) \sim F_1^{\leftarrow}\left(1 - \frac{1-\kappa}{C_\lambda\, t}\right)\,, \qquad \kappa \uparrow 1\,. \qquad (2.63)$$

Proof. From Proposition 2.2.7 we know that S^+ is a compound Poisson process with parameters λ^+ (here following from (2.60)) and F^+ as in (2.61) from which we conclude

$$\lim_{z\to\infty}\frac{\overline{F}^+(z)}{\overline{F}_1(z)} \;=\; \frac{1}{\lambda^+}[\lambda_1 + c_2\,\lambda_2 + \cdots + c_d\,\lambda_d] = \frac{C_\lambda}{\lambda^+} \in (0,\infty)\,,$$

i.e.

$$\overline{F}^+(z) \;\sim\; \frac{C_\lambda}{\lambda^+}\overline{F}_1(z)\,, \quad z\to\infty\,. \tag{2.64}$$

In particular, $\overline{F}^+ \in \mathcal{S}\cap(\mathcal{R}\cup\mathcal{R}_\infty)$ and S^+ defines a one-dimensional compound Poisson model. From (2.9) and (2.64) total OpVAR follows as

$$\mathrm{VAR}_t^+(\kappa) \sim F^{+\leftarrow}\left(1-\frac{1-\kappa}{\lambda^+\,t}\right) \sim F_1^{\leftarrow}\left(1-\frac{1-\kappa}{C_\lambda\,t}\right)\,, \qquad \kappa\uparrow 1\,. \qquad \square$$

We close this Section with some examples illustrating the Theorem above.

Example 2.3.12. [Dominating cells in the case of independence]
Assume that $c_i = 0$ for all $i \geq 2$; i.e. $\overline{F}_i(x) = o(\overline{F}_1(x))$, $i = 2,\ldots,d$. We then have $C_\lambda = \lambda_1$ and it follows from (2.63) that independent total OpVAR asymptotically equals the stand alone OpVAR of the first cell. In contrast to the completely dependent case (confer Proposition 2.3.7 and Example 2.3.10), this holds for the class $\mathcal{S}\cap(\mathcal{R}\cup\mathcal{R}_\infty)$ and not only for $F_1 \in \mathcal{R}$. $\square$

Example 2.3.13. [Pareto distributed severities, independence]
Consider a multivariate compound Poisson model with independent cells and Pareto distributed severities so that the constants c_i of Theorem 2.3.11 are given by

$$\lim_{x\to\infty}\frac{\overline{F}_i(x)}{\overline{F}_1(x)} = \left(\frac{\theta_i}{\theta_1}\right)^\alpha\,, \quad i=1,\ldots,b\,, \quad \text{and} \quad \lim_{x\to\infty}\frac{\overline{F}_i(x)}{\overline{F}_1(x)} = 0\,, \quad i=b+1,\ldots,d\,,$$

for some $b \geq 1$. Then

$$C_\lambda = \sum_{i=1}^b \left(\frac{\theta_i}{\theta_1}\right)^\alpha \lambda_i$$

and the distribution tail $\overline{F}^+$ satisfies

$$\overline{F}^+(z) \;=\; \frac{1}{\lambda^+}\sum_{i=1}^b \lambda_i \left(1+\frac{z}{\theta_i}\right)^{-\alpha} \;\sim\; \frac{1}{\lambda^+}\sum_{i=1}^b \lambda_i\,\theta_i^\alpha\,z^{-\alpha}\,, \qquad z\to\infty\,.$$

It follows that

$$\mathrm{VAR}_t^+(\kappa) \sim \left(\frac{t\,\sum_{i=1}^b \lambda_i\,\theta_i^\alpha}{1-\kappa}\right)^{1/\alpha} = \left(\sum_{i=1}^b \left(\mathrm{VAR}_t^i(\kappa)\right)^\alpha\right)^{1/\alpha}\,, \qquad \kappa\uparrow 1\,,$$

where $\mathrm{VAR}_t^i(\kappa)$ denotes the stand alone OpVAR of cell i according to (2.20). For identical cell frequencies $\lambda := \lambda_1 = \cdots = \lambda_b$ this further simplifies to

$$\mathrm{VAR}_t^+(\kappa) \sim \left(\frac{\lambda\, t}{1-\kappa}\right)^{1/\alpha} \left(\sum_{i=1}^{b} \theta_i^\alpha\right)^{1/\alpha}, \quad \kappa \uparrow 1\,.$$

□

Example 2.3.14. [GPD distributed severities, independence]
Consider a multivariate compound Poisson model of Definition 2.2.6 with independent cells and assume that within each cell large losses above some high threshold $u_i > 0$ have a GPD tail given by (2.19). We again have

$$\lim_{x\to\infty} \frac{\overline{F}_i(x)}{\overline{F}_1(x)} = \frac{w_i}{w_1}\left(\frac{\beta_i}{\beta_1}\right)^{1/\xi},\ i = 1,\ldots,b\,, \qquad \lim_{x\to\infty} \frac{\overline{F}_i(x)}{\overline{F}_1(x)} = 0\,,\ i = b+1,\ldots,d\,,$$

for $b \geq 1$. The distribution tail can then be calculated as

$$\overline{F}^+(z) \sim w^+ \left(1 + \xi\frac{z-u_1}{\beta_1}\right)^{-1/\xi} \sim \frac{1}{\lambda^+}\sum_{i=1}^{b} w_i\,\lambda_i \left(\frac{\beta_i}{\xi}\right)^{1/\xi} z^{-1/\xi}\,, \quad z\to\infty\,,$$

where $w^+ = \frac{1}{\lambda^+}\sum_{i=1}^{b}(\frac{\beta_i}{\beta_1})^{1/\xi}\lambda_i w_i$, and total OpVAR is asymptotically given by

$$\begin{aligned}\mathrm{VAR}_{\perp t}^+(\kappa) &\sim u_1 + \frac{\beta_1}{\xi}\left[\left(\frac{w^+\lambda t}{1-\kappa}\right)^{\xi} - 1\right] \\ &\sim \left(\sum_{i=1}^{b}\left(\frac{\beta_i}{\xi}\right)^{1/\xi}\frac{w_i\lambda_i\, t}{1-\kappa}\right)^{\xi} \sim \left(\sum_{i=1}^{b}\left(\mathrm{VAR}_t^i(\kappa)\right)^{1/\xi}\right)^{\xi}, \qquad \kappa\uparrow 1 \quad (2.65)\end{aligned}$$

□

Example 2.3.15. [Continuation of Example 2.3.10]
Consider a bivariate compound Poisson model with independent cells and Weibull distributed severities according to (2.58). Accoding to Theorem 2.3.11 we have $C_\lambda = \lambda_1$ and independent total OpVAR is asymptotically given by

$$\mathrm{VAR}_t^+(\kappa) \sim \mathrm{VAR}_t^1(\kappa) \sim 2\left[\ln\left(\frac{1-\kappa}{\lambda\, t}\right)\right]^2, \qquad \kappa\uparrow 1\,.$$

□

Let us briefly compare the cases of complete dependence and independence by using the important example of the GPD model. On one hand, Theorem 2.3.5 and Example 2.3.9 state that for complete dependence total asymptotic OpVAR is simply

the sum of the dominating cell's asymptotic stand-alone OpVARs. On the other hand, for independent cell severities, total OpVAR can be expressed in terms of a generalised mean M_p by

$$M_p(a_1,\dots,a_n) := \left(\frac{1}{n}\sum_{k=1}^{n} a_k^p\right)^{1/p}, \quad a_k \geq 0, \quad p \neq 0\,,$$

and (2.65) can be written for $b \leq d$ as

$$\mathrm{VAR}^+_{\perp t}(\kappa) \;\sim\; b^{\xi}\, M_{1/\xi}(\mathrm{VAR}^1_t(\kappa),\dots,\mathrm{VAR}^b_t(\kappa))\,, \qquad \kappa \uparrow 1\,.$$

Formally, the complete dependent case (2.57) can also be expressed by M_p, namely

$$\mathrm{VAR}^+_{\parallel t}(\kappa) \;\sim\; b\, M_1(\mathrm{VAR}^1_t(\kappa),\dots,\mathrm{VAR}^b_t(\kappa))\,, \qquad \kappa \uparrow 1\,.$$

A fundamental difference between both extreme dependence models is that, due to the dynamical dependence concept of Lévy copulas, the completely dependent model implies identical frequency λ for all cells, whereas the independent model allows for different cell frequencies. However, if high-severity losses mainly occur in one, say the first cell, both models yield the same asymptotic total OpVAR, namely the stand-alone VAR of the first cell; see Example 2.3.4.

Recall that simple-sum OpVAR (2.57) is often suggested as an upper bound for total OpVAR. This is also the basis for the new proposals of Basel II, where the standard procedure for calculating capital charges for operational risk is just the simple-sum OpVAR. Hence, our calculation has shown that regulators implicitly assume complete dependence between different cells as worst case scenario, meaning that losses within different business lines or risk categories always happen at the same instants of time. Moreover, they assume completely dependent loss severities.

This viewpoint is in the heavy-tailed case grossly misleading. To see this, assume the same frequency λ in all cells, also for the independent model, and denote by $\mathrm{VAR}^+_{\parallel}(\kappa)$ and $\mathrm{VAR}^+_{\perp}(\kappa)$ completely dependent and independent total OpVAR, respectively. Then, from (2.57) and (2.65), as a consequence of convexity ($0 < \xi < 1$) and concavity ($\xi > 1$) of the function $x \mapsto x^{1/\xi}$, we obtain

$$\frac{\mathrm{VAR}^+_{\perp}(\kappa)}{\mathrm{VAR}^+_{\parallel}(\kappa)} \;\sim\; \frac{\left(\sum_{i=1}^{b} w_i\, \beta_i^{1/\xi}\right)^{\xi}}{\sum_{i=1}^{b} w_i^{\xi}\, \beta_i} \quad \begin{cases} < 1\,, & 0 < \xi < 1\,, \\ = 1\,, & \xi = 1\,, \\ > 1\,, & \xi > 1\,. \end{cases} \tag{2.66}$$

This result says that for heavy-tailed severity data with GPD tail given by (2.19) subadditivity of OpVAR is violated because the sum of stand-alone OpVARs is smaller

$1/\xi$	$\mathrm{VAR}_{\parallel}^{+}$	$\mathrm{VAR}_{\perp}^{+}$
1.2		178.2
1.1		187.8
1.0	200.0	200.0
0.9		216.0
0.8		237.8
0.7		269.2

Table 2.3.16. *Comparison of total* OpVAR *for two operational risk cells (each with stand-alone* VAR *of EUR 100 million) in the case of complete dependence* ($\parallel$) *and independence* ($\perp$) *for different values of the tail parameter* ξ *in the relevant area* (*cf.* (2.66)).

than independent total OpVAR. This is a direct consequence of the Pareto-like tail, which we assumed for the loss severity distribution and is well-known in the financial literature; cf. Rootzén & Klüppelberg [62]. Nevertheless, to give an example for operational risk, consider two cells with constant stand-alone OpVAR of EUR 100 million, each calculated from a GPD model with fixed parameters $\beta_1 = \beta_2 = 1$, $w_1 = w_2 = 1$, and common tail parameter $\xi = \xi_1 = \xi_2$. Table 2.3.16 compares, for a realistic range of ξ-values (cf. Moscadelli [54])), total OpVAR both for completely dependent and independent data. Obviously, for $\xi > 1$, total OpVAR increases superlinearly, when taking on two independent risks, for example by opening two new subsidiaries in different parts of the world.

Even if we assume $0 < \xi < 1$ for all operational risk cells and thus $\mathrm{VAR}_{\parallel}^{+}(\kappa) > \mathrm{VAR}_{\perp}^{+}(\kappa)$, we obtain an interesting result concerning the relative "diversification benefit" in operational risk defined as $(\mathrm{VAR}_{\parallel}^{+} - \mathrm{VAR}_{\perp}^{+})/\mathrm{VAR}_{\parallel}^{+}$. Often diversification is understood to be directly linked to the notion of correlation – and particularly in the context of operational risk – to the loss-number correlation $\rho(N_1(t), N_2(t))$. For heavy-tailed data, however, it is well-known that correlation (even if it exists) is a misleading concept to describe diversification within a portfolio. Consider e.g. Figure 2.3.17, where the relative diversification benefit for two operational risk cells with $\theta_1 = \theta_2$ is plotted as a function of the tail parameter ξ. Obviously, relative diversification is very sensitive with regards to the value of ξ. In contrast to that, the loss-number correlation ρ of both models is constant, and it follows from (2.31) that $\rho_{\parallel} = 1$ and $\rho_{\perp} = 0$. Over and above, we know that dependence models, with regards to the frequency correlation have asymptotically undistinguishable OpVARs. As a consequence thereof, instead of trying to estimate precise frequency correlations between different cells, all effort should be directed into a more accurate modelling of the loss severity distribution.

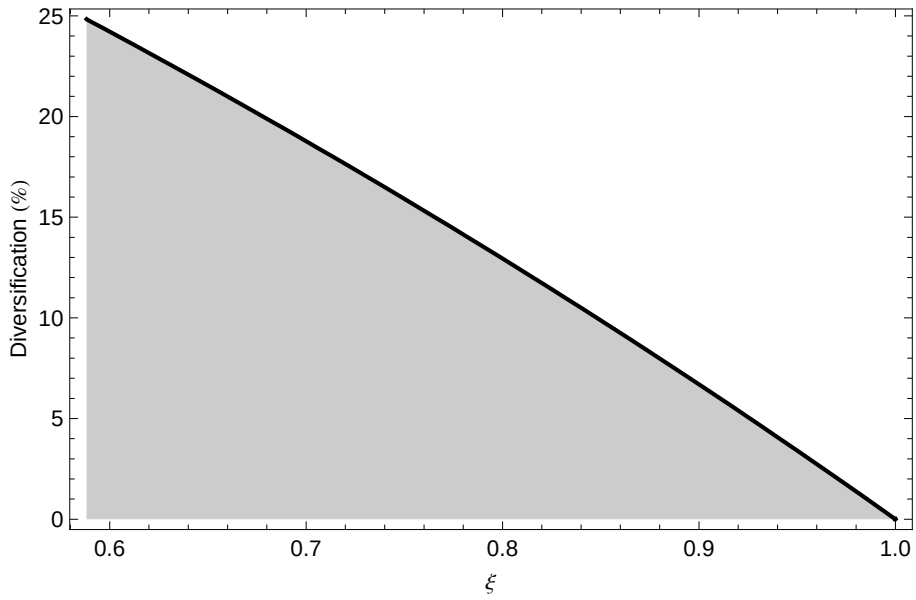

Figure 2.3.17. *Plot of the relative diversification benefit* $\frac{\mathrm{VAR}_{\|}^{+}-\mathrm{VAR}_{\perp}^{+}}{\mathrm{VAR}_{\|}^{+}} = 1 - 2^{\xi-1}$ *as given by* (2.66) *for two operational risk cells as a function of the tail parameter* ξ*.*

A final word of warning. It is beyond all dispute that operational risk is very material in most financial institutions. However, risk severities can be extreme by their very nature, recall for instance prominent examples such as Barings Bank (loss \$1.3 billion) or Sumitomo Corp. (loss \$2.6 billion). Moreover, our analysis shows that multivariate high-confidence OpVAR is very sensible to the parametrization of the severity distribution, an issue, which has already been pointed out by Mignola & Ugoccioni [51] for the univariate case. Altogether, this confirms the view that capital charges are not always the best way to deal with operational risk, and that risk measurement has always to be complemented by sound risk management and control processes.

2.3.5 Multivariate Compound Poisson Models of Regular Variation

Multivariate regular variation is an appropriate mathematical tool for discussing heavy tail phenomena as they occur for instance in operational risk. Moreover, as the notion of regular variation has proved useful for one-dimensional cell severity distributions, it seems natural to exploit the corresponding concept for the multivariate model. Readable overviews about multivariate regular variation are the books of Resnick [60, 61].

To simplify notation we denote by $\mathbb{E} := [\mathbf{0}, \boldsymbol{\infty}] \setminus \{\mathbf{0}\}$ where $\mathbf{0}$ an $\boldsymbol{\infty}$ are the zero and infinity vectors in $\mathbb{R}^d$, respectively. If for a general Borel set $A \subset \mathbb{E}$ its complement

in $\mathbb{E}$ is denoted by A^c, we introduce for $\boldsymbol{x} \in \mathbb{E}$ the complement

$$[\mathbf{0}, \boldsymbol{x}]^c := \mathbb{E} \setminus [\mathbf{0}, \boldsymbol{x}] = \{\boldsymbol{y} \in \mathbb{E} : \max_{1 \le i \le d} \frac{y_i}{x_i} > 1\} .$$

We also recall that a Radon measure is a measure, which is finite on all compacts. Finally, henceforth all operations and order relations of vectors are taken componentwise.

As already mentioned in Remark 2.1.7 (e), multivariate regular variation is best formulated in terms of vague convergence of measures. Moreover, from Lemma 6.1 of Resnick [61], p. 174, however, it suffices to consider regions $[\mathbf{0}, \boldsymbol{x}]^c$ for $\boldsymbol{x} \in \mathbb{E}$ which determine the convergence, and this is how we formulate our results.

We begin with regular variation of random vectors or, equivalently, of multivariate distribution functions. The idea is to have regular variation not only in some (or all) marginals, but along every ray starting in 0 and going through the positive cone to infinity. Clearly, this limits the set of possible dependence structures between the marginals, however, such models are still flexible enough to be broadly applied to various fields such as telecommunication, insurance, and last but not least VAR analysis in the banking industry. Furthermore, many of the dependence models implying multivariate regular variation can still be solved and analysed analytically.

Let us consider a random variable $\boldsymbol{X}$ with distribution function F that is—as our Lévy measure Π—concentrated on $\mathbb{E}$.

Assume there exists a Radon measure ν on $\mathbb{E}$ (i.e. a Borel measure that is finite on compact sets) such that

$$\lim_{t \to \infty} \frac{1 - F(t\boldsymbol{x})}{1 - F(t\mathbf{1})} = \lim_{t \to \infty} \frac{P(t^{-1}\boldsymbol{X} \in [\mathbf{0}, \boldsymbol{x}]^c)}{P(t^{-1}\boldsymbol{X} \in [\mathbf{0}, \mathbf{1}]^c)} = \nu([\mathbf{0}, \boldsymbol{x}]^c) \tag{2.67}$$

holds for all $\boldsymbol{x} \in \mathbb{E}$ which are continuity points of the function $\nu([\mathbf{0}, \cdot]^c)$. One can show that the above definition (2.67) implies that ν has a homogeneity property, i.e. there exists some $\alpha > 0$ such that

$$\nu([\mathbf{0}, s\boldsymbol{x}]^c) = s^{-\alpha} \nu([\mathbf{0}, \boldsymbol{x}]^c) , \quad s > 0 , \tag{2.68}$$

and we say that F has a *multivariate regularly varying tail with index* $-\alpha$ ($F \in \mathcal{R}_{-\alpha}$). Condition (2.67) also says that $\overline{F}(t\mathbf{1}) = 1 - F(t\mathbf{1})$ as a function of t is in $\mathcal{R}_{-\alpha}$. Define now $b(t)$ to satisfy $\overline{F}(b(t)\mathbf{1}) \sim t^{-1}$ as $t \to \infty$. Then, replacing t by $b(t)$ in (2.67) yields

$$\lim_{t \to \infty} tP\Big(\frac{\boldsymbol{X}}{b(t)} \in [\mathbf{0}, \boldsymbol{x}]^c\Big) = \nu([\mathbf{0}, \boldsymbol{x}]^c) . \tag{2.69}$$

In (2.69) the random variable $\boldsymbol{X}$ is normalised by the function $b(\cdot)$. As explained in Resnick (2007), Section 6.5.6, normalisation of all components by the same function

$b(\cdot)$ implies that the marginal tails of $\boldsymbol{X}$ satisfy for $i, j \in \{1, \dots, d\}$

$$\lim_{x\to\infty} \frac{\overline{F}_i(x)}{\overline{F}_j(x)} = \frac{c_i}{c_j}\,,$$

where $c_i, c_j \in [0, \infty)$. Let us assume that $c_1 > 0$, which means that F_1 is either heavier-tailed or tail equivalent to the other distributions $F_i, i = 2, \dots, d$. Then, we can always define $c_i, i = 2, \dots, d$ such that $c_1 = 1$, and we can choose $b(t)$ such that for $t \to \infty$

$$\overline{F}_1(b(t)) \sim t^{-1} \quad \Leftrightarrow \quad b(t) \sim \left(\frac{1}{\overline{F}_1}\right)^{\leftarrow}(t) \quad \Leftrightarrow \quad b(t) \sim \overline{F}_1{}^{\leftarrow}\left(\frac{1}{t}\right), \tag{2.70}$$

and by substituting in (2.69) we obtain a limit on the left-hand side of (2.69) with the same scaling structure as before.

To formulate analogous definitions for Lévy measures, note first that we can rewrite (2.67) by means of the distribution of $\boldsymbol{X}$ as

$$\lim_{t\to\infty} \frac{P_{\boldsymbol{X}}(t[\mathbf{0}, \boldsymbol{x}]^c)}{P_{\boldsymbol{X}}(t[\mathbf{0}, \mathbf{1}]^c)} = \nu([\mathbf{0}, \boldsymbol{x}]^c),$$

and similarly (2.69) as

$$\lim_{t\to\infty} tP_{\boldsymbol{X}}(b(t)[\mathbf{0}, \boldsymbol{x}]^c) = \lim_{t\to\infty} tP_{\boldsymbol{X}}([\mathbf{0}, b(t)\boldsymbol{x}]^c) = \nu([\mathbf{0}, \boldsymbol{x}]^c)\,. \tag{2.71}$$

Then, the analogue expression to (2.67) for a Lévy measure Π is simply

$$\lim_{t\to\infty} \frac{\Pi(t[\mathbf{0}, \boldsymbol{x}]^c)}{\Pi(t[\mathbf{0}, \mathbf{1}]^c)} = \lim_{t\to\infty} \frac{\Pi(\{\boldsymbol{y} \in \mathbb{E} : y_1 > tx_1 \text{ or } \cdots \text{ or } y_d > tx_d)}{\Pi(\{\boldsymbol{y} \in \mathbb{E} : y_1 > t \text{ or } \cdots \text{ or } y_d > t)} = \nu([\mathbf{0}, \boldsymbol{x}]^c)\,, \tag{2.72}$$

for all $x \in \mathbb{E}$ which are continuity points of the function $\nu([\mathbf{0}, \cdot]^c)$. Summarising what we have so far yields the following definition for multivariate regular variation for Lévy measures, now formulated in analogy to (2.69) or (2.71), respectively:

Definition 2.3.18. [Multivariate regular variation for spectrally positive Lévy processes]
(a) *Let Π be a Lévy measure of a spectrally positive Lévy process on $\mathbb{E}$. Assume that there exists a function $b : (0, \infty) \to (0, \infty)$ satisfying $b(t) \to \infty$ as $t \to \infty$ and a Radon measure ν on $\mathbb{E}$, called the limit measure, such that*

$$\lim_{t\to\infty} t\,\Pi([\mathbf{0}, b(t)\,\boldsymbol{x}]^c) = \nu([\mathbf{0}, \boldsymbol{x}]^c)) \tag{2.73}$$

for all $\boldsymbol{x} \in \mathbb{E}$ which are continuity points of the function $\nu([\mathbf{0}, \cdot]^c$. Then we call $\overline{\Pi}$ multivariate regularly varying.

(b) *The measure ν has a scaling property: there exists some $\alpha > 0$ such that for every $s > 0$*

$$\nu([\mathbf{0}, s\boldsymbol{x}]^c) = s^{-\alpha}\nu([\mathbf{0}, \boldsymbol{x}]^c)\,, \quad x \in \mathbb{E}\,, \tag{2.74}$$

i.e. $\nu([0,\cdot]^c)$ is homogeneous *of order $-\alpha$, and $\overline{\Pi}$ is called* multivariate regularly varying with index $-\alpha$ ($\Pi \in \mathcal{R}_{-\alpha}$).

Remark 2.3.19. (a) In (2.73), the scaling of all components of the tail integral by the same function $b(\cdot)$ implies

$$\lim_{x\to\infty} \frac{\overline{\Pi}_i(x)}{\overline{\Pi}_j(x)} = \tilde{c}_{ij} \;\in\; [0,\infty]\,, \tag{2.75}$$

for $1 \le i, j \le d$. We now focus on the case that all $\overline{\Pi}_i$ are tail-equivalent, i.e. $\tilde{c}_{ij} > 0$ for some i, j. In particular, we then have marginal regular variation $\overline{\Pi}_i \in \mathcal{R}_{-\alpha}$ with the same tail index, and thus for all $i = 1, \dots, d$

$$\begin{aligned} \lim_{t\to\infty} t\,\overline{\Pi}_i(b(t)\,x) &= \nu([0,\infty] \times \cdots \times (x,\infty] \times [0,\infty] \times \cdots \times [0,\infty]) \\ &= \nu_i(x,\infty] = \tilde{c}_i\, x^{-\alpha}\,, \qquad x > 0\,, \end{aligned} \tag{2.76}$$

for some $\tilde{c}_i > 0$. Specifically, we can always set $\tilde{c}_1 = 1$ and thus

$$\overline{\Pi}_1(b(t)) \sim t^{-1} \quad \Leftrightarrow \quad b(t) \sim \left(\frac{1}{\overline{\Pi}_1}\right)^{\leftarrow}(t)\,, \qquad t \to \infty\,. \tag{2.77}$$

(b) If $\overline{\Pi}_1$ is *standard regularly varying* (i.e. with index $\alpha = 1$ and slowly varying function $L \equiv 1$), we can take $b(t) = t$, which immediately follows from (2.77).
(c) There exists also a broader definition of multivariate regular variation which allows for different α_i in each marginal; see Theorem 6.5 of Resnick [61], p. 204. However, we have already dealt with the situation of dominant marginals and, hence, the above definition is the relevant one for us. □

From the point of view of dependence structure modeling, multivariate regular variation is basically a special form of multivariate dependence. Hence, a natural question in this context is how multivariate regular variation is linked to the dependence concept of a Lévy copula. The answer gives the following Theorem.

Theorem 2.3.20. [Lévy copulas and multivariate regular variation] *Let $\overline{\Pi}$ be a multivariate tail integral of a spectrally positive Lévy process in $\mathbb{E}$. Assume that the marginal tail integrals $\overline{\Pi}_i$ are regularly varying with index $-\alpha$. Then the following assertions hold.*
(1) If the Lévy copula $\widehat{C}$ is a homogeneous function of order 1, then $\overline{\Pi}$ is multivariate

regularly varying with index $-\alpha$.
(2) The tail integral $\overline{\Pi}$ is multivariate regularly varying with index $-\alpha$ if and only if the Lévy copula $\widehat{C}$ is regularly varying with index 1; i.e.

$$\lim_{t\to\infty} \frac{\widehat{C}(t\,(x_1,\dots,x_d))}{\widehat{C}(t\,(1,\dots,1))} = g(x_1,\dots,x_d)\,, \quad (x_1,\dots,x_d) \in [\mathbf{0},\boldsymbol{\infty})\,, \tag{2.78}$$

and $g(s\,x) = s\,g(x)$ for $x \in [\mathbf{0},\boldsymbol{\infty})$.

Proof. (1) For any Lévy copula $\widehat{C}$, we can write the Lévy measure $\Pi([\mathbf{0},\boldsymbol{x}]^c)$ for $\boldsymbol{x} \in \mathbb{E}$ as

$$\begin{aligned}
\Pi([\mathbf{0},\boldsymbol{x}]^c) &= \Pi\{\boldsymbol{y} \in \mathbb{E} : y_1 > x_1 \text{ or } \cdots \text{ or } y_d > x_d\} \\
&= \sum_{i=1}^{d} \overline{\Pi}_i(x_i) - \sum_{\substack{i_1,i_2=1\\ i_1<i_2}}^{d} \widehat{C}(\overline{\Pi}_{i_1}(x_{i_1}), \overline{\Pi}_{i_2}(x_{i_2})) \\
&\quad + \sum_{\substack{i_1,i_2,i_3=1\\ i_1<i_2<i_3}}^{d} \widehat{C}(\overline{\Pi}_{i_1}(x_{i_1}), \overline{\Pi}_{i_2}(x_{i_2}), \overline{\Pi}_{i_3}(x_{i_3})) \\
&\quad + \cdots + (-1)^{d-1} \widehat{C}(\overline{\Pi}_{i_1}(x_{i_1}), \dots, \overline{\Pi}_{i_d}(x_{i_d}))\,.
\end{aligned}$$

The homogeneity allows interchange of the factor t with $\widehat{C}$, which, together with marginal regular variation as formulated in (2.76), yields the limit as in (2.73):

$$\begin{aligned}
\lim_{t\to\infty} t\,\Pi([\mathbf{0},b(t)\,\boldsymbol{x}]^c) &= \sum_{i=1}^{d} \nu_i(x_i,\infty] - \sum_{\substack{i_1,i_2=1\\ i_1<i_2}}^{d} \widehat{C}(\nu_{i_1}(x_{i_1},\infty], \nu_{i_2}(x_{i_2},\infty]) \\
&\quad + \cdots + (-1)^{d-1}\, \widehat{C}(\nu_{i_1}(x_{i_1},\infty], \dots, \nu_{i_d}(x_{i_d},\infty]) \\
&= \nu\{\boldsymbol{y} \in \mathbb{E} : y_1 > x_1 \text{ or } \cdots \text{ or } y_d > x_d\} \\
&= \nu([\mathbf{0},\boldsymbol{x}]^c)\,, \qquad \boldsymbol{x} \in \mathbb{E}\,.
\end{aligned} \tag{2.79}$$

(2) This follows from the same calculation as in the proof of (1) by observing that asymptotic interchange of the factor t with $\widehat{C}$ is possible if and only if (2.78) holds. □

Remark 2.3.21. For this definition of multivariate regular variation of arbitrary functions we refer to Bingham et al. [7], Appendix 1.4. □

Excursus: From Lévy Copulas to Pareto Lévy Copulas

The general concept of multivariate regular variation mentioned in Remark 2.3.19 with possibly different marginals requires different normalizing functions $b_1(\cdot),\dots,b_d(\cdot)$ in

(2.73). In that case marginals are usually transformed to standard regular variation with $\alpha = 1$ and $L \equiv 1$. In this case the scaling property (2.74) in the limit measure ν always scales with $\alpha = 1$. This is equivalent to all marginal Lévy processes being one-stable. Such standardised Lévy process lead to the concept of the *Pareto Lévy copula*, which in the context of multivariate regular variation seems to be more natural than the classical Lévy copula with Lebesgue marginals, see Klüppelberg and Resnick [40] as well as in Böcker and Klüppelberg [18].

Definition 2.3.22. *Let $(\boldsymbol{X}_t)_{t\geq 0}$ be a Lévy process in $\mathbb{R}^d$ with a Lévy measure Γ that has standard 1-stable one-dimensional marginals. Then we call Γ a* Pareto Lévy measure *and the associated tail integral*

$$\overline{\Gamma}(\boldsymbol{x}) = \Gamma([x_1, \infty) \times \cdots \times [x_d, \infty)) =: \widetilde{C}(x_1, \ldots, x_d)\,, \quad \boldsymbol{x} \in \mathbb{E}\,,$$

is referred to as Pareto Lévy copula $\widetilde{C}$.

For every Lévy process we now can transform its marginal Lévy measures by

$$\overline{\Pi}_i(x) \to \overline{\Gamma}_i(x) = \overline{\Pi}_i \circ \Big(\frac{1}{\overline{\Pi}_i}\Big)^{\leftarrow}(x) = \frac{1}{x}, \quad x \in [0, \infty)\,, \tag{2.80}$$

yielding standard 1-stable marginal Lévy processes with Lévy measures $\overline{\Pi}_i(x) = x^{-1}$ for $x > 0$, i.e. the corresponding Pareto Lévy copula. Note that if we apply such transformation to the multivariate compound Poisson model, then the transformed 1-stable Lévy processes are not compound Poisson anymore. Instead they are of infinite variation and have an infinite number of small jumps per unit time expressed by $\lim_{x\downarrow 0} \overline{\Pi}_i(x) = \infty$. For definitions and references of stable Lévy processes in general see Cont and Tankov [24].

Lemma 2.3.23. *Let $(\boldsymbol{X}_t)_{t\geq 0}$ be a spectrally positive Lévy process (i.e. a Lévy process admitting only positive jumps) with Lévy measure Π on $\mathbb{E}$ and continuous marginal tail measures $\overline{\Pi}_1, \ldots, \overline{\Pi}_d$. Then*

$$\overline{\Pi}(\boldsymbol{x}) = \Pi([x_1, \infty] \times \cdots \times [x_d, \infty]) = \widetilde{C}\Big(\frac{1}{\overline{\Pi}_1(x_1)}, \ldots, \frac{1}{\overline{\Pi}_d(x_d)}\Big)\,, \quad \boldsymbol{x} \in \mathbb{E}\,,$$

and $\widetilde{C}$ is a Pareto Lévy copula.

Proof. Note that for all $\boldsymbol{x} \in \mathbb{E}$,

$$\widetilde{C}(x_1, \ldots, x_d) \;=\; \overline{\Pi}\Big(\Big(\frac{1}{\overline{\Pi}_1}\Big)^{\leftarrow}(x_1), \ldots, \Big(\frac{1}{\overline{\Pi}_d}\Big)^{\leftarrow}(x_d)\Big).$$

This implies for the one-dimensional marginal tail measures

$$\widetilde{C}(0, \ldots, x, \ldots, 0) = \overline{\Pi}_i \circ \Big(\frac{1}{\overline{\Pi}_i}\Big)^{\leftarrow}(x) = \frac{1}{x}, \quad x \in [0, \infty).$$

□

From the construction above it is also clear that if $\widehat{C}(x_1, \ldots, x_d)$ is a Lévy copula, then the associated Pareto Lévy copula $\widetilde{C}$ can be constructed by $\widetilde{C}(x_1, \ldots, x_d) = \widehat{C}(1/x_1, \ldots, 1/x_d)$. Hence, Theorem 2.2.3 can easily be restated for Pareto Lévy copulas. Furthermore, the following examples follow immediately from those given in Section 2.2.2.

Example 2.3.24. [Independence Pareto Lévy-copula]
The independence Pareto Lévy copula of is given by

$$\widetilde{C}_\perp(\boldsymbol{x}) = x_1^{-1} 1_{\{x_2=\ldots=x_d=0\}} + \cdots + x_d^{-1} 1_{\{x_1=\ldots=x_{d-1}=0\}}.$$

The resulting Lévy process with Pareto Lévy copula $\widetilde{C}_\perp$ is a standard 1-stable process with independent components. □

Example 2.3.25. [Complete (positive) dependence Pareto Lévy copula]
The complete (positive) dependence Pareto Lévy copula is given by

$$\widetilde{C}_\parallel(\boldsymbol{x}) = \min(x_1^{-1}, \ldots, x_d^{-1}).$$

□

Example 2.3.26. [Archimedian Pareto Lévy copula]
Let $\phi : [0, \infty] \to [0, \infty]$ be strictly decreasing with $\phi(0) = \infty$ and $\phi(\infty) = 0$. Assume that $\phi^{\leftarrow}$ has derivatives up to order d with $(-1)^k \frac{d^k \phi_t^{\leftarrow}}{dt^k} > 0$ for $k = 1, \ldots, d$. Then the following is a Pareto Lévy copula

$$\widetilde{C}(\boldsymbol{x}) = \phi^{\leftarrow}(\phi(x_1^{-1}) + \cdots + \phi(x_d^{-1})).$$

□

Example 2.3.27. [Clayton Pareto Lévy copula]
Take $\phi_t = t^{-\delta}$ for $\delta > 0$. Then the Archimedian Pareto Lévy copula

$$\widetilde{C}_\delta(\boldsymbol{x}) = (x_1^\delta + \cdots + x_d^\delta)^{-1/\delta}$$

is called Clayton Pareto Lévy copula. □

The Clayton Pareto Lévy copula $\widetilde{C}(x_1, x_2) = (x_1^\delta + x_2^\delta)^{-1/\delta}$ for $x_1, x_2 > 0$ is homogenous of order -1, and, consequently, the associated Clayton Lévy copula is homogenous of order 1, see also Example 2.2.10. From Theorem 2.3.20 we can conclude that, if the marginal Lévy tail measures $\overline{\Pi}_1$ and $\overline{\Pi}_2$ (i.e. before standardising the marginals) were

regularly varying with some index $-\alpha$, then the Lévy measure Π had bivariate regularly varying tail with index $-\alpha$.

The -1-homogeneity of the Clayton Pareto Lévy copula can be used to visualise its regularly varying dependence structure. In doing so, note also that

$$\Gamma([\mathbf{0}, \boldsymbol{x}]^c) = \overline{\Gamma}_1(x_1) + \overline{\Gamma}_2(x_2) - \widetilde{C}(x_1, x_2) = \frac{1}{x_1} + \frac{1}{x_2} - \left(\left(\frac{1}{x_1}\right)^{-\delta} + \left(\frac{1}{x_2}\right)^{-\delta}\right)^{-1/\delta}.$$

We now transform to polar coordinates by setting $x_1 = r\cos\varphi$ and $x_2 = r\sin\varphi$, and from the homogeneity property it follows

$$\Gamma([\mathbf{0}, \boldsymbol{x}]^c) = r^{-1}\Gamma([\mathbf{0}, (\cos\varphi, \sin\varphi)^\top]^c) =: \Gamma(r, \varphi).$$

This is depicted in Figure 2.3.28 where $\Gamma(r, \varphi)$ is plotted for $r = 1$ as a function of φ, and thus the Clayton dependence structure is plotted as a measure on the quatercircle.

OpVAR for Regularly Varying Dependence Structures

We now want to apply our results about multivariate regular variation of Lévy measures (which hold true for general spectrally positive Lévy processes) to the problem of calculating total OpVAR. Hence, we turn back to a multivariate compound Poisson process whose Lévy measure Π is multivariate regularly varying according to (2.73). In particular, this implies tail equivalence of the marginal Lévy measures and we can write (2.75) with some $c_i \in (0, \infty)$ as

$$c_i := \lim_{x\to\infty} \frac{\overline{\Pi}_i(x)}{\overline{\Pi}_1(x)} = \frac{\lambda_i \overline{F}_i(x)}{\lambda_1 \overline{F}_1(x)} =: \frac{\lambda_i}{\lambda_1}\tilde{c}_i\,, \tag{2.81}$$

i.e. $\lim_{x\to\infty} \overline{F}_i(x)/\overline{F}_1(x) = \tilde{c}_i$. We avoid situations where for some i we have $c_i = 0$, corresponding to cases in which for $x \to \infty$ the tail measure $\overline{\Pi}_i(x)$ decays faster than $x^{-\alpha}$, i.e. in (2.81) we only consider the heaviest tail measures, all of tail index $-\alpha$. This makes sense because we know from Theorem 2.3.2 that only the heaviest-tailed risk cells contribute to total OpVAR.

Theorem 2.3.29. [OpVAR for the compound Poisson model with multivariate regular variation] *Consider a multivariate compound Poisson model with multivariate regularly varying cell processes $S_1, \ldots, S_d$ with index $-\alpha$ and limit measure ν in* (2.73)*. Assume further that the severity distributions F_i for $i = 1, \ldots, d$ are strictly increasing and continuous. Then, S^+ defines a one-dimensional compound Poisson model with parameters satisfying for $z \to \infty$*

$$\lambda^+ \overline{F}^+(z) \;\sim\; \lambda_1\, \nu^+(1, \infty]\, \overline{F}_1(z) \in \mathcal{R}_{-\alpha}\,, \quad z \to \infty\,, \tag{2.82}$$

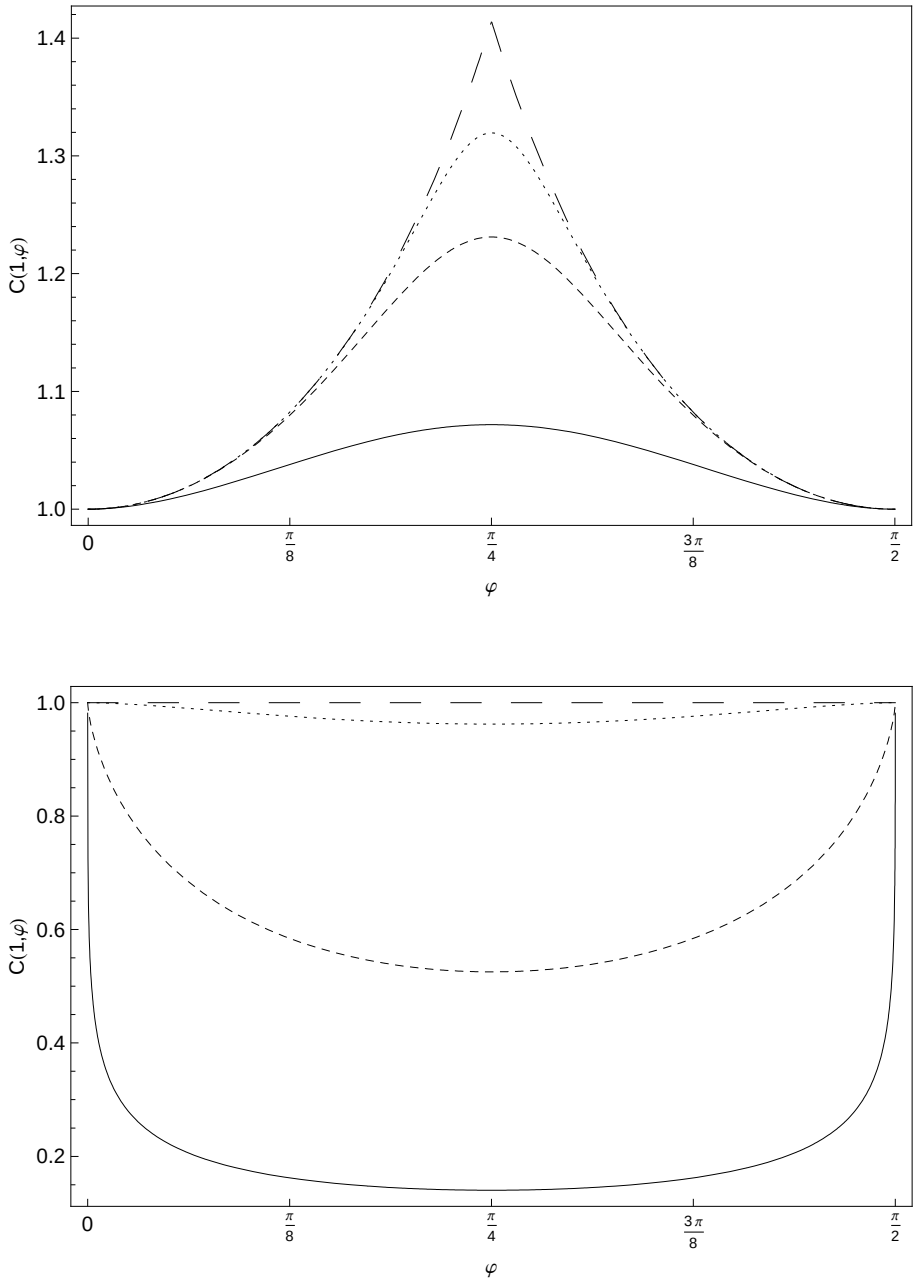

Figure 2.3.28. *The Pareto Lévy copula in polar coordinates* $\widetilde{C}(r,\varphi) = \overline{\Gamma}(r,\varphi)$ *as a function of the angle* $\varphi \in (0,\pi/2)$ *for* $r = 1$ *and different values of the dependence parameter.*
Top panel: $\theta = 1.8$ *(dotted line),* $\theta = 0.7$ *(dashed line),* $\theta = 0.3$ *(solid line).*
Bottom panel: $\theta = 2.5$ *(solid line),* $\theta = 5$ *(dashed line),* $\theta = 10$ *(dotted line),* $\theta = \infty$ *(complete positive dependence, long-dashed line).*

where $\nu^+(z,\infty] = \nu\{\boldsymbol{x} \in \mathbb{E} : \sum_{i=1}^d x_i > z\}$ *for* $z > 0$. *Furthermore, total OpVAR is asymptotically given by*

$$\mathrm{VAR}_t(\kappa) \sim F_1^{\leftarrow}\left(1 - \frac{1-\kappa}{t\,\lambda_1\,\nu^+(1,\infty]}\right), \quad \kappa \uparrow 1\,. \tag{2.83}$$

Proof. First recall that multivariate regular variation of $\overline{\Pi}$ implies regular variation of the marginal tail integrals, i.e. $\overline{\Pi}_i \in \mathcal{R}_{-\alpha}$ for all $i = 1, \ldots, d$. Analogously to Resnick [61], Proposition 7.3, p. 227, the tail measure $\overline{\Pi}^+$ is also regularly varying with index $-\alpha$, more precisely we have that

$$\begin{aligned} \lim_{t\to\infty} t\,\overline{\Pi}^+(b(t)\,z) &= \nu^+(z,\infty] \\ = \nu\{\boldsymbol{x} \in \mathbb{E} : \sum_{i=1}^d x_i > z\} &= z^{-\alpha}\nu\{\boldsymbol{x} \in \mathbb{E} : \sum_{i=1}^d x_i > 1\}\,. \end{aligned}$$

Now we can choose $b(t)$ so that $\lim_{t\to\infty} t\overline{\Pi}_1(b(t)) = 1$ and thus

$$\lim_{z\to\infty} \frac{\overline{\Pi}^+(z)}{\overline{\Pi}_1(z)} = \lim_{t\to\infty} \frac{t\overline{\Pi}^+(b(t))}{t\overline{\Pi}_1(b(t))} = \nu^+(1,\infty]\,.$$

Relation (2.82) follows immediately, and (2.83) by Theorem 2.1.15. □

According to the Theorem 2.3.29, for the wide class of regularly varying distributions, total OpVAR can effectively be written in terms of the severity distribution of the first cell. Specifically, the right-hand side of (2.83) can be understood as the stand-alone, asymptotic OpVAR of the first cell with an adjusted frequency parameter, namely $\lambda_1\nu^+(1,\infty]$. What remains is to find examples where $\nu^+(1,\infty]$ can be calculated analytically or numerically to understand better the influence of certain dependence parameters.

Example 2.3.30. [Revisiting the case of independent operational risk cells]
Before we present some explicit results for the Clayton Lévy copula below, let us consider again the particularly easy case with independent cells. Since then all mass is concentrated on the positive axes, we obtain

$$\nu^+(1,\infty] = \nu_1(1,\infty] + \cdots + \nu_d(1,\infty]\,. \tag{2.84}$$

From $\overline{\Pi}_1(b(t)) \sim t^{-1}$ it follows for the tail measure of the first cell (see also (2.76))

$$\lim_{t\to\infty} t\overline{\Pi}_1(b(t)z) = z^{-\alpha} = \nu_1(z,\infty]\,. \tag{2.85}$$

For $i = 2, \ldots, d$ we obtain by using (2.81),

$$\lim_{t\to\infty} t\overline{\Pi}_i(b(t)z) = \lim_{t\to\infty} \frac{\overline{\Pi}_i(b(t)z)}{\overline{\Pi}_1(b(t))} = \lim_{u\to\infty} \frac{\overline{\Pi}_i(uz)}{\overline{\Pi}_i(u)}\frac{\overline{\Pi}_i(u)}{\overline{\Pi}_1(u)} = \tilde{c}_i\,z^{-\alpha} = \nu_i(z,\infty]\,, \tag{2.86}$$

and therefore altogether $\nu^+(1,\infty] = 1 + \sum_{i=2}^d \tilde{c}_i$. By (2.82) together with $\lambda_1 \tilde{c}_i = \lambda_i c_i$ we finally recover the result of Theorem 2.3.11. $\square$

Let us consider a bivariate example where the marginal Lévy measures are not independent, and thus the limit measure $\nu^+(z,\infty]$ is not just the sum of the marginal limit measures as in (2.84). Instead, $\nu^+(z,\infty]$ has to be calculated by taking also mass between the positive axes into account (i.e. the grey shaded area in Figure 2.3.1), which can be done by representing $\nu^+(z,\infty]$ as an integral over a density.

Example 2.3.31. [Clayton Lévy copula]
The Clayton Lévy copula is for $\delta > 0$ defined as

$$\widehat{C}(u_1,\dots,u_d) = (u_1^{-\delta} + \cdots + u_d^{-\delta})^{-1/\delta}, \quad u_1,\dots,u_d \in (0,\infty),$$

see also Example 2.2.10. In Figures 2.3.33-2.3.38 we show sample paths of two dependent compound Poisson processes, where the dependence is modelled via a Clayton Lévy copula for different parameter values. With increasing dependence parameter δ we see more joint jumps.

Note that $\widehat{C}$ is homogenous of order 1. Hence, from Theorem 2.3.20, if $\overline{\Pi}_i \in \mathcal{R}_{-\alpha}$ for some $\alpha > 0$, the Lévy measure is multivariate regularly varying with index $-\alpha$. For $d = 2$, we obtain from (2.79) together with (2.85) and (2.86), and by setting $c := \tilde{c}_2$,

$$\nu([0,(x_1,x_2)]^c) = x_1^{-\alpha} + c\, x_2^{-\alpha} - \left[x_1^{\alpha\delta} + c^{-\delta}\, x_2^{\alpha\delta}\right]^{-1/\delta}, \quad x_1 > 0,\, x_2 > 0\,.$$

By differentiating we obtain the density ν' for $0 < \delta < \infty$ (the completely positive dependent case ($\delta \to \infty$) and the independent case ($\delta \to 0$) are not covered by the following calculation) as

$$\nu'(x_1,x_2) = c^{-\delta}\,\alpha^2(1+\delta)\, x_1^{-\alpha(1+\delta)-1} x_2^{\alpha\delta-1} \left(1 + c^{-\delta}\big(\frac{x_2}{x_1}\big)^{\alpha\delta}\right)^{-1/\delta-2}, \; x_1 > 0,\, x_2 > 0\,.$$

We then can write

$$\begin{aligned}
\nu^+(1,\infty] &= \nu\{(x_1,x_2) \in \mathbb{E} : x_1 + x_2 > 1\} \\
&= \nu((1,\infty] \times [0,\infty]) + \int_0^1 \int_{1-x_1}^\infty \nu'(x_1,x_2)\, dx_2\, dx_1 \\
&= \nu_1(1,\infty] + \int_0^1 \int_{1-x_1}^\infty \nu'(x_1,x_2)\, dx_2\, dx_1 \\
&= 1 + \alpha \int_0^1 \left(1 + c^{-\delta}\big(\frac{1}{x_1} - 1\big)^{\alpha\delta}\right)^{-1/\delta-1} x_1^{-1-\alpha} dx_1\,,
\end{aligned}$$

and substituting $v = \frac{1}{x_1} - 1$ we obtain

$$\begin{aligned} \nu^+(1,\infty] &= 1 + \alpha \int_0^\infty \left(1 + c^{-\delta} v^{\alpha\delta}\right)^{-1/\delta - 1} (1+v)^{\alpha-1} dv \\ &= 1 + c^{1/\alpha} \int_0^\infty \left(1 + s^\delta\right)^{-1/\delta - 1} (c^{1/\alpha} + s^{-1/\alpha})^{\alpha-1} ds\,. \end{aligned} \tag{2.87}$$

Since $g(y) := (1+y^\delta)^{-1/\delta-1}, y > 0$, is the density of a positive random variable Y_δ, we finally arrive at

$$\begin{aligned} \nu^+(1,\infty] &= 1 + c^{1/\alpha}\, E[(c^{1/\alpha} + Y_\delta^{-1/\alpha})^{\alpha-1}] \\ &=: 1 + c^{1/\alpha}\, C(\alpha,\delta)\,. \end{aligned} \tag{2.88}$$

Then, an analytical approximation for OpVAR follows together with expression (2.83),

$$\mathrm{VAR}_t^+(\kappa) \sim F_1^{\leftarrow}\left(1 - \frac{1-\kappa}{\lambda_1(1 + c^{1/\alpha}\, C(\alpha,\delta))\, t}\right), \quad \kappa \uparrow 1\,. \tag{2.89}$$

Note that $\mathrm{VAR}_t^+(\kappa)$ increases with $C(\alpha,\delta)$. For $\alpha = 1$ the constant $C(1,\delta) = 1$ implies that total OpVAR for all Clayton parameters in the range $0 < \delta < \infty$ is given by

$$\mathrm{VAR}_t^+(\kappa) \sim F_1^{\leftarrow}\left(1 - \frac{1-\kappa}{\lambda_1(1+c)\, t}\right) = F_1^{\leftarrow}\left(1 - \frac{1-\kappa}{(\lambda_1 + c_2\lambda_2)\, t}\right), \quad \kappa \uparrow 1\,,$$

which is (independent of the dependence parameter δ) equal to the independent OpVAR of Theorem 2.3.11. Note also the relation $c = \lambda_2/\lambda_1 c_2$ between the different constants in (2.81) Furthermore, $\nu^+(1,\infty]$ is greater or less than 2, according as α is greater or less than 1, respectively. If $\alpha\,\delta = 1$ we can solve the integral in (2.87) (similarly to Example 3.8 of Bregman and Klüppelberg [21]) and obtain

$$\nu^+(1,\infty] = \frac{c^{1+1/\alpha} - 1}{c^{1/\alpha} - 1}\,.$$

Figure 2.3.32 illustrates the tail measure $\nu^+(1,\infty]$ given in (2.88) for different values of θ and α. Note that according to (2.83), OpVAR increases with growing values for $\nu^+(1,\infty]$. Hence, Figure 2.3.32 shows us that in the case of $\alpha > 1$, a higher positive dependence structure leads to a higher OpVAR, whereas for $\alpha < 1$, it is the other way around: the lower the dependence structure is (i.e. losses are not tending to occur together), the higher OpVAR will be. This again shows how things may go awry, if one applies the well-behaving behaviour of normal distributions to extremely heavy-tailed distributions, and diversification may behave completely counterintuitive. Finally, note that for $\delta \to \infty$, independence occurs and $\nu^+(1,\infty] = 1 + c$ is constant as indicated by the horizontal long-dashed line in Figure 2.3.32. □

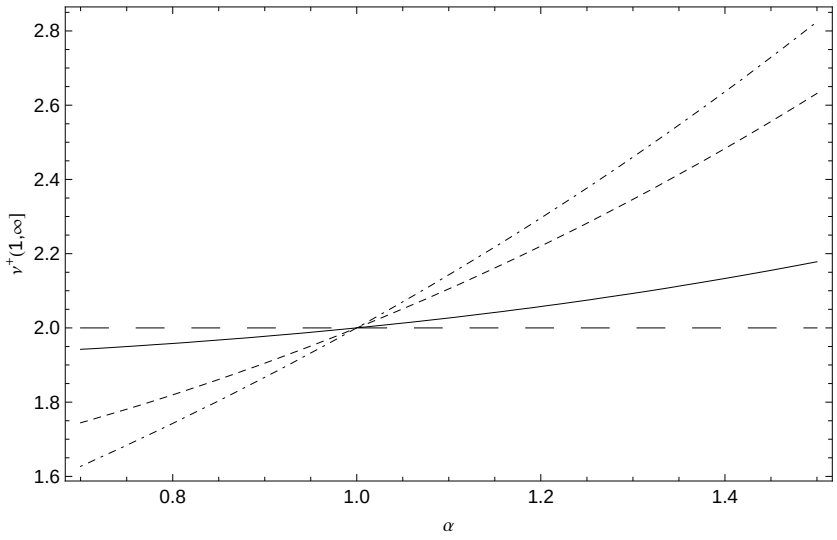

Figure 2.3.32. *Illustration of the tail measure* $\nu^+(1,\infty]$ *as given in* (2.88) *as a function of* α *for different values of the Clayton Lévy copula parameter* δ*. We have chosen 0.3 (light dependence, solid line), 1 (medium dependence, dashed line), and 10 (high dependence, dotted-dashed line). The long-dashed line corresponds to the independent case. Moreover, we have used* $c = 1$ *in all plots.*

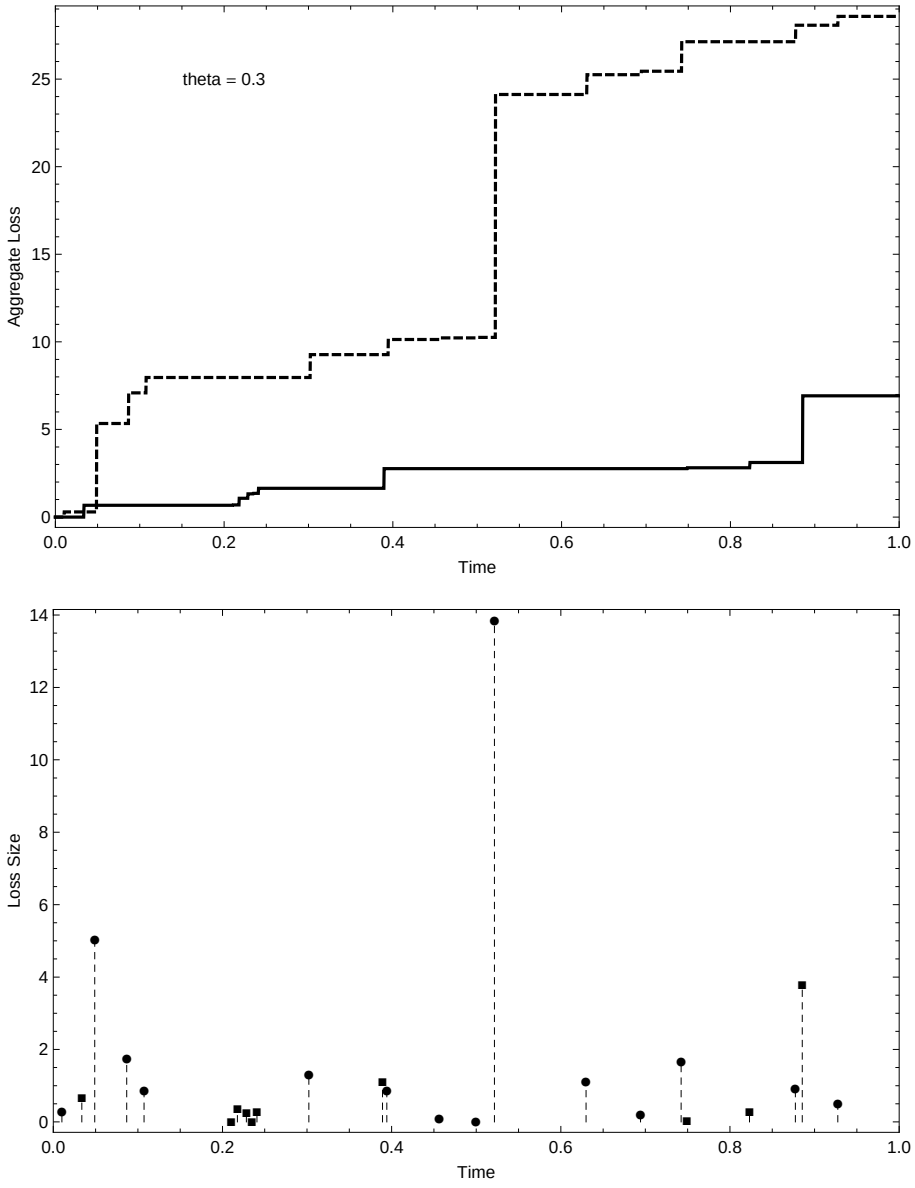

Figure 2.3.33. Simulation of the bivariate compound Poisson model as of Definition 2.2.6 with Clayton Lévy copula with parameter $\vartheta = 0.3$ (light dependence). Top panel: sample paths of the aggregate loss processes. Bottom panel: severity and occurrence times of losses. The univariate compound Poisson processes have frequencies of $\lambda_1 = \lambda_2 = 10$ and Pareto distributed severities with parameters $\alpha_1 = 1/\xi = 1.2$ and $\alpha_1 = 1/\xi = 2$.

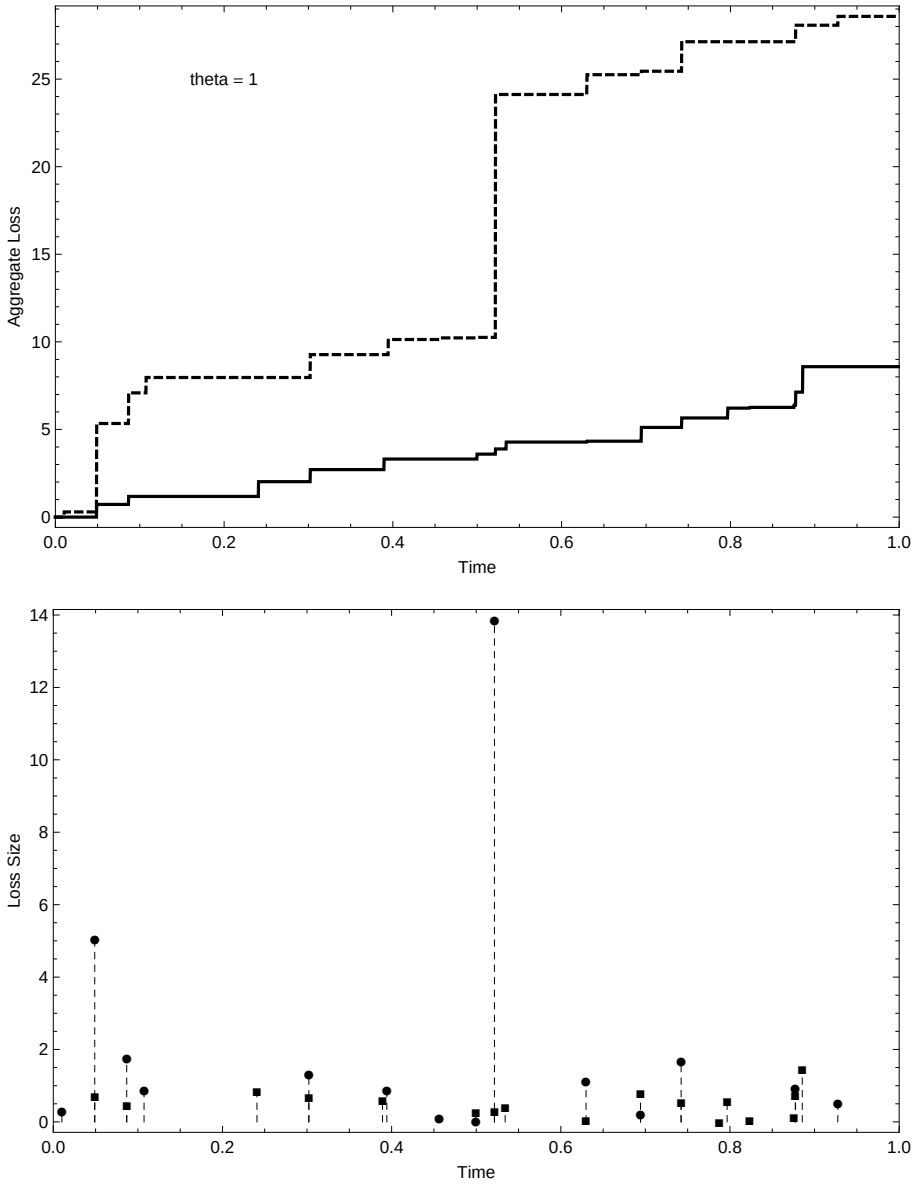

Figure 2.3.34. Simulation of the bivariate compound Poisson model as of Definition 2.2.6 with Clayton Lévy copula with parameter $\vartheta = 1$ (medium dependence). Top panel: sample paths of the aggregate loss processes. Bottom panel: severity and occurrence times of losses. The univariate compound Poisson processes have frequencies of $\lambda_1 = \lambda_2 = 10$ and Pareto distributed severities with parameters $\alpha_1 = 1/\xi = 1.2$ and $\alpha_1 = 1/\xi = 2$.

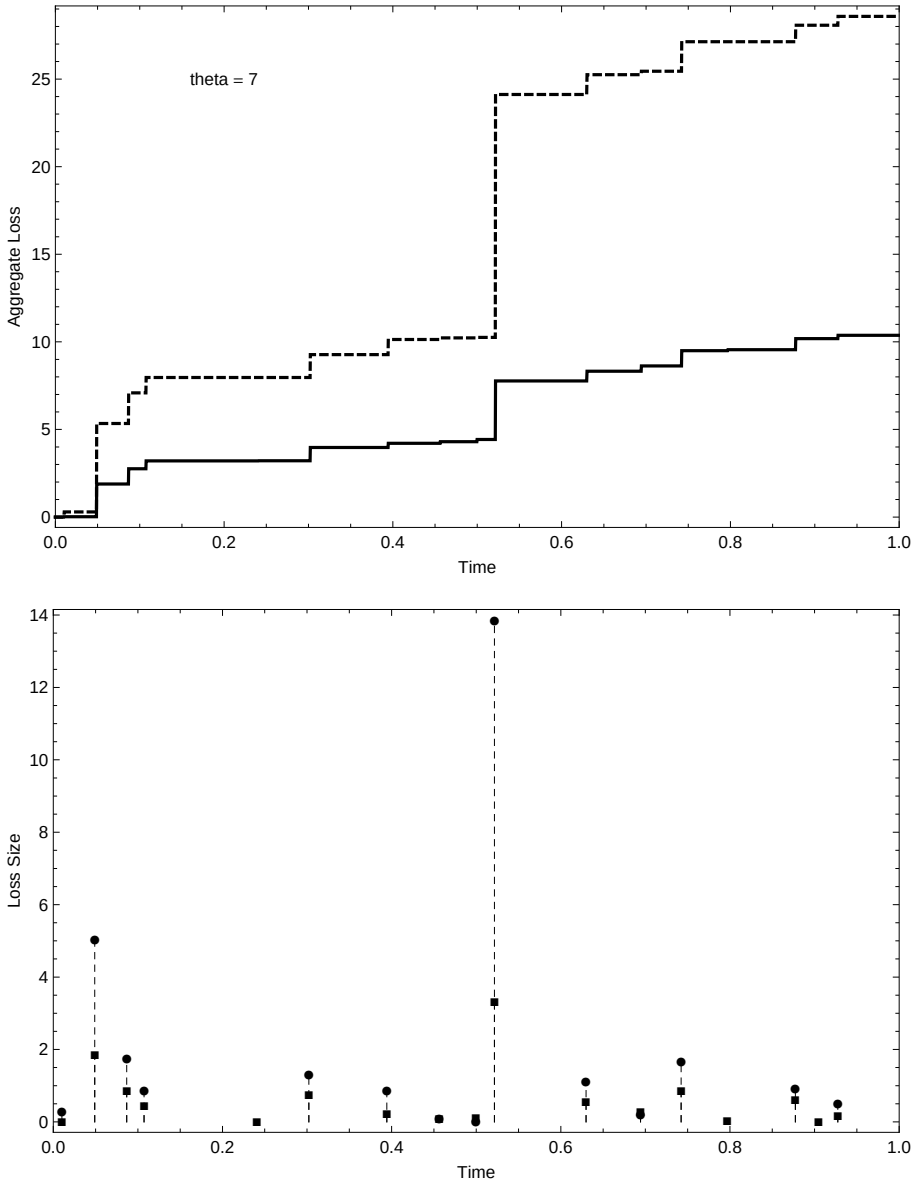

Figure 2.3.35. Simulation of the bivariate compound Poisson model as of Definition 2.2.6 with Clayton Lévy copula with parameter $\vartheta = 7$ (strong dependence). Top panel: sample paths of the aggregate loss processes. Bottom panel: severity and occurrence times of losses. The univariate compound Poisson processes have frequencies of $\lambda_1 = \lambda_2 = 10$ and Pareto distributed severities with parameters $\alpha_1 = 1/\xi = 1.2$ and $\alpha_1 = 1/\xi = 2$.

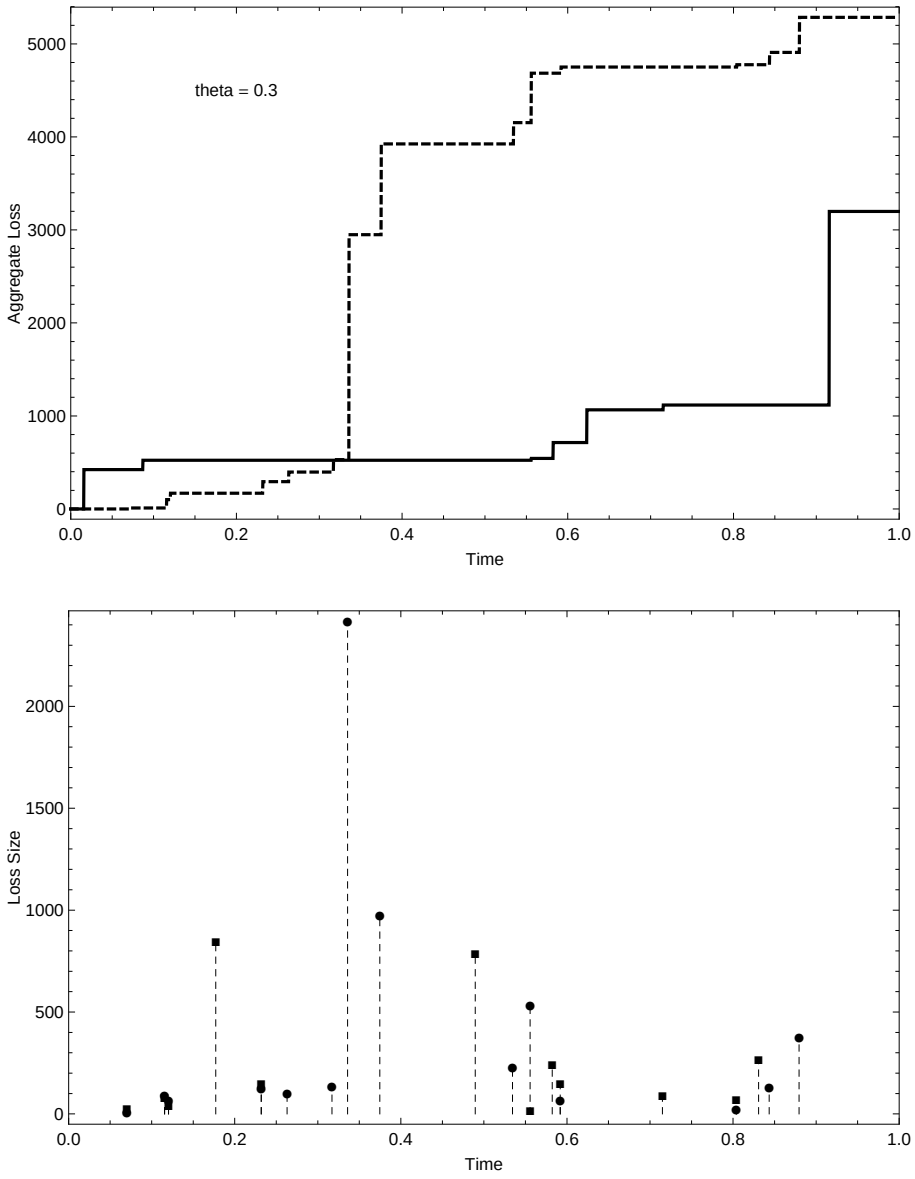

Figure 2.3.36. Simulation of the bivariate compound Poisson model as of Definition 2.2.6 with Clayton Lévy copula with parameter $\vartheta = 0.3$ (light dependence). Top panel: sample paths of the aggregate loss processes. Bottom panel: severity and occurrence times of losses. The univariate compound Poisson processes have frequencies of $\lambda_1 = \lambda_2 = 10$ and lognormal distributed severities with parameters $\mu_1 = \mu_2 = 5$ and $\sigma_1 = \sigma_2 = 1.5$.

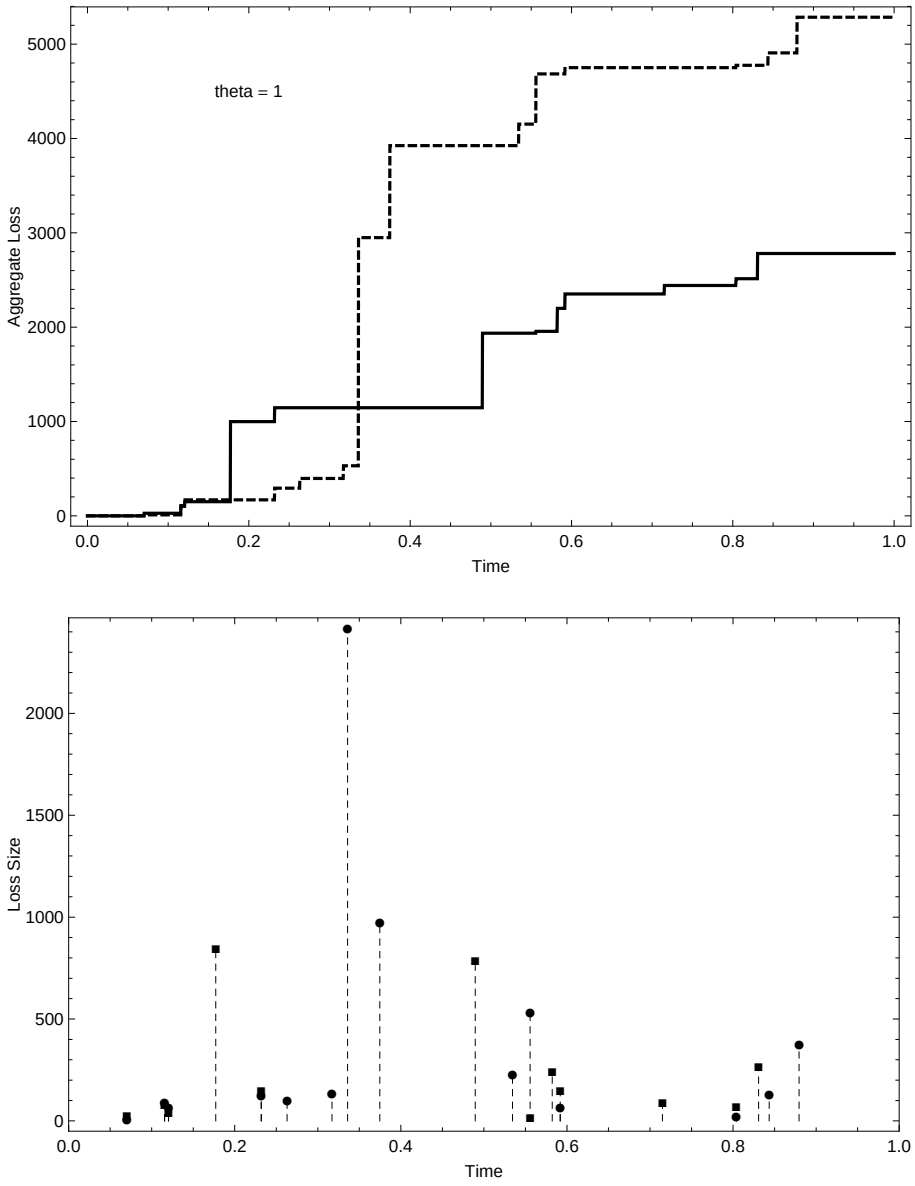

Figure 2.3.37. Simulation of the bivariate compound Poisson model as of Definition 2.2.6 with Clayton Lévy copula with parameter $\vartheta = 1$ (medium dependence). Top panel: sample paths of the aggregate loss processes. Bottom panel: severity and occurrence times of losses. The univariate compound Poisson processes have frequencies of $\lambda_1 = \lambda_2 = 10$ and lognormal distributed severities with parameters $\mu_1 = \mu_2 = 5$ and $\sigma_1 = \sigma_2 = 1.5$.

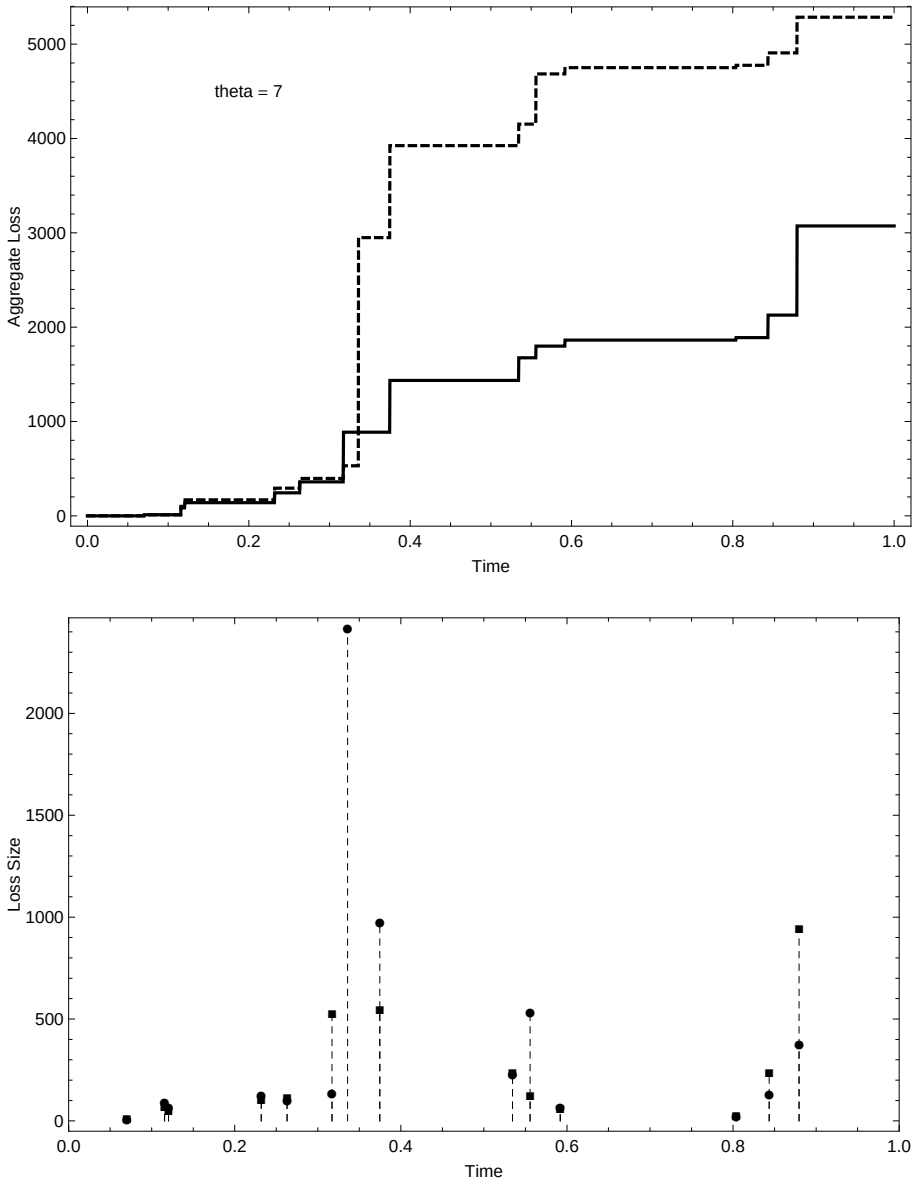

Figure 2.3.38. Simulation of the bivariate compound Poisson model as of Definition 2.2.6 with Clayton Lévy copula with parameter $\vartheta = 7$ (strong dependence). Top panel: sample paths of the aggregate loss processes. Bottom panel: severity and occurrence times of losses. The univariate compound Poisson processes have frequencies of $\lambda_1 = \lambda_2 = 10$ and lognormal distributed severities with parameters $\mu_1 = \mu_2 = 5$ and $\sigma_1 = \sigma_2 = 1.5$.

Chapter 3

A Continuous-Time Model for Business Risk

In this Chapter we suggest a continuous-time model for business risk calculation, which is based on a stochastic version of the discounted-cash-flow method. In particular, using different Gauss processes, we examine the long-term impact of the bank's future earnings fluctuations on the present value, leading to the so-called Capital-at-Risk.

3.1 Introduction

As can be seen e.g. by the cutting edge Section of RISK Magazine, research papers mainly focus on market risk, credit risk, and—with a little less attention—operational risk. Although these risk types are very important for financial institutions, the true landscape of risk is much more complex and far from being well explored and understood. There is a variety of "other" risks looming on the horizon, which seriously threaten a bank's profitability or which can disrupt or even destroy its business completely. Moreover, such risks often reflect an under-researched area of financial risk management, and established and ready-to-use measurement techniques are rarely available. Also banking supervisors demand that more attention is being paid to such "hard-to-measure" risks as the following Pillar-2 passages of the new international regulatory framework of Basel II [5] show:

- 731: "...Sound capital assessment include...policies and procedures designed to ensure that the bank identifies, *measures*, and reports *all material risks*."
- 742: "Although the Committee recognises that other risks [...] are not easily measurable, it expects industry to further develop techniques for managing *all aspects* of these risks."

This view has also been confirmed by different European supervisors, confer e.g. The Committee of European Banking Supervisors [25].

Capturing all material risks of a financial institution requires a broad risk self-assessment to find out which are the most relevant risk drivers for the bank. One of the most obvious variables to be monitored in this context are earning themselves. However, none of the Pillar 1 risks take earnings volatility *directly* as a primary driver into account, instead, they usually focus on aspects of the business environment that only indirectly affect the institution' earnings by virtue of e.g. failed processes, credit defaults, drop in share prices, or interest rate changes.

For an all-encompassing risk assessment it is therefore necessary to introduce an additional kind of risk that is directly linked to the uncertainty of specific earnings components not yet associated to other risk types. Usually, such an earnings-related potential loss, which can also threaten a bank's market capitalisation, is referred to as *business risk*.

Evidence for the growing importance of business risk was recently also given in a survey undertaken by the IFRI/CRO Forum about economic capital practices in leading financial institutions [36] where 85 % of the participants stated to include business risk in their aggregated economic capital assessment. Yet surprisingly, there is no common agreement on a precise definition, specific risk drivers, and measurement methodology for business risk, even though its absolute size in term of economic capital is comparable to that of operational risk, see again [36]. With this regard, we also performed a benchmark exercise on a sample of 15 international banks by analysing their risk management practice as disclosed in their official financial annual reports from 2004 to 2006. Again, we found that an increasing number of institutions are trying to quantify business risk in some way, even if different definitions and assumptions are adopted. Broadly speaking, approaches for business risk quantification can be divided into two main categories; top-down and bottom-up. Top-down techniques are linked to the general trend of the business environment and benchmark analysis based on external data is used for approximating business risk. In contrast to that, bottom-up approaches try to explicitly determine the volatility of particular, bank-internal economic time series (such as volumes, earnings, revenues, and expenses) at a more granular level, which is then transformed into a measure of business risk.

Here we we propose a bottom-up approach for modelling and measuring business risk where the dynamic of the underlying earnings is described in a continuous-time model.

3.2 Modelling Business CAR: a Simple Approach

3.2.1 Setting the Scene

Overlap with other risk types. Of course, the concept of "revenues" and "expenses" as we used so far is too general for measuring business risk. In particular, in order to avoid double counting and risk overlap, revenue and cost components that enter the business risk model must not directly or indirectly be used for the quantification of other risk types. To give an example, as potentially relevant revenues one may consider customer related provisions and net interest rate income, while on the cost side administrative expenses and depreciations may be included into business risk quantification. On the other hand, earnings related to trading activities would clearly cause an overlap with market risk, and should therefore not be included. Something similar holds for loan loss provisions, when they are captured within the bank's credit portfolio model.

However, the question which revenue and cost components are really relevant for modelling a particular firm's business risk, and which parts have to be excluded, is not an easy one. The answer crucially depends on the firm's definition of others risk types and its economic capital framework in general, and therefore setting up a business risk model should always be an integral part of the bank's overall risk-defining and assessment process. Moreover, one has to be aware that the availability and granularity of revenue and cost data may also depend on the firm's accounting rules, controlling standards, and IT infrastructure. As a consequence, the quality of data may differ from one legal entity to the other, and in order to achieve reliable results at aggregated level, great attention should be paid with regard to data selection and preparation. Hereafter, when we talk about earnings, we actually always mean non-credit and non-market earnings so that there is no double counting with other risk types that are already measured within a bank's economic capital model.

EAR versus CAR. Business risk can be defined as the potential loss in the company's earnings due to adverse, unexpected changes in business volume, margins, or both. Such losses can result above all from a serious deterioration of the market environment, customer shift, changes in the competitive situation, or internal restructuring. On one hand, these effects can lead to a drop in earnings in the short-term, e.g. within the next budget year, and are often measured in terms of earnings volatility are more general by EAR. On the other hand, volume or margin shrinking probably leads to a longer-lasting weakening of the earnings situation, thereby seriously diminishing the company's market capitalisation, and this risk is often referred to as CAR. As pointed out by Saita [64, 65], the recognition of such negative long-term effects on

earnings and the resulting impact on the market capitalisation is particular important for the shareholder perspective on capital and should also be used in the context of risk-adjusted performance measurement, e.g. by means of RAROC, EVA, or related concepts.

A convincing analysis proving this link between earnings' related risk and a company's loss in market value is given in Morrison, Quella and Slywotzky [53]. They found out that during a period of five years, 10 % of Fortune 1,000 companies lost (at least once) 25 % of their shareholder value within a one-month period, and that nearly all of these stock drops were a result of reduced quarterly earnings or reduced expected future earnings. Moreover, the majority of these earnings-shortfalls (about 58 %) were not owing to classical financial risks or operational losses but rather to what Quella et al. refer to as strategic risk factors, such as raising costs and margin squeeze, emerging global competitors, and customer priority shift etc.

If one considers business risk as a matter of market capitalisation and therefore measures it by CAR, one has to take the uncertainty of (all) future earnings into account. As mentioned above, such earnings fluctuations, i.e. the deviations of the realised earnings from the planned earnings trajectory, maybe the result of many different factors. However, for the model we suggest here, it is not necessary to explicitly link all these risk factors to future earnings. Instead we suppose that all risk factors together constitute some random "noise" effect, mixing with the expected earnings path; i.e. for $t \geq 0$ future cumulated earnings $E(t)$ can be written as

$$E(t) = f(t) + \text{“noise”}(t)\,, \qquad t \geq 0\,,$$

where f is a nonrandom function describing the expected earnings trajectory. Consequently, in Section 3.2.2 we model future earnings as a stochastic process $(E(t))_{t\geq 0}$.

Discounted-cash-flow method. Before we continue, however, it is worthwhile to recall some basic facts about company valuation, especially about the discounted-cash-flow method where expected future earnings are discounted to obtain the company's market value (see e.g. Goedhart, Koller and Wessels [33] or Pratt, Reilly and Schweihs [58] for a more detailed description). Denote by $\Delta E_{t_i} \in \mathbb{R}$ the company's earnings that are planned to be realized in $\Delta t_i = t_i - t_{i-1}$, i.e., between the future time periods t_{i-1} and t_i, defined for all $0 = t_0 < t_1 < \cdots < t_T$. One then usually defines the *present value* or *market value* of the company as

$$V(T) = \sum_{i=1}^{T} \frac{\Delta E_{t_i}}{(1+r_i)^{t_i}}\,, \tag{3.1}$$

where $r_i \in \mathbb{R}_+$ is a risk adjusted discount rate. Expression (3.1) simply says that a company's market value is just the sum of its discounted expected future earnings.

For our purposes, however, a continuous-time setting of the present value (3.1) is more feasible.

Definition 3.2.1. [Present value in continuous time] *Let $r(t)$ for $t \geq 0$ be a nonrandom positive function, representing the short-term discount rate with continuous compounding. Then, the* present value *of all earnings cumulated until the time horizon t is given by*

$$P(t) = P(0) + \int_0^t e^{-R(s)}\, dE(s)\,, \qquad t \geq 0\,,$$

with discount rate $R(t) := \int_0^t r(\tau)d\tau$ for $t \geq 0$, and

$$E(t) = E(0) + \int_0^t dE(s)\,, \qquad t \geq 0\,,$$

are the cumulated future earnings *as they are expected to be realised up to a time horizon t. Furthermore, we define*

$$P(\infty) := \lim_{t\to\infty} P(t)\,,$$

provided that the limit exists.

3.2.2 Model Definition and First Results

We begin our analysis of business CAR with a simple model based on Brownian motion. Such a model allows for closed-form solutions for business CAR and is therefore particularly useful to understand the nature and general properties if this important risk type.

Stochastic modelling of future cash flows. We define our model in a multivariate fashion in that it takes the dependence between different *cells* into account. Each cell could simply reflect a legal entity, business division, geographical region, or a combination of them. If one explicitly splits up revenues and costs instead of considering earnings directly, revenues and costs are represented by different cells, and for each cell we can define a cash-flow process representing the stochastic evolution of revenues, costs, or earnings. In the following we treat revenues as positive and costs as negative variables.

Definition 3.2.2. [Brownian motion cash flow (BMC) model] *Consider a d-dimensional standard Brownian motion $(W_1(t), \ldots, W_d(t))_{t\geq 0}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, the BMC model consists of:*

(1) Cash-flow processes.
For each business risk cell, indexed by $i = 1, \dots, m$, cumulated future cash flows $X_i(t)$ for $t \geq 0$ are described by a cash-flow process*, which is the strong continuous solution to the Itô-stochastic-differential equation*

$$dX_i(t) = \alpha_i(t)\, dt + \sum_{j=1}^{d} \sigma_{ij}(t)\, dW_j(t)\,, \qquad t \geq 0\,. \tag{3.2}$$

The bank's total aggregated cash flow is given by the aggregate cash-flow process

$$X(t) = \sum_{i=1}^{m} X_i(t) = \sum_{i=1}^{m} X_i(0) + \sum_{i=1}^{m} \int_0^t dX_i(s)\,.$$

Here, $\alpha_i(\cdot) > 0,\ i = 1, \dots, m,$ and $\sigma_{ij}(\cdot),\ i = 1, \dots, m; j = 1, \dots, d,$ are nonrandom functions of time, satisfying the integrability conditions $\int_0^t |\alpha_i(s)| ds < \infty$ and $\int_0^t \sigma_{ij}^2(s) ds < \infty$.

(2) Value process.
Let $R(\cdot) > 0$ be a nonrandom discount rate (see Definition 3.2.1) so that $\int_0^t \left(|\alpha_i(s)| e^{-R(s)} + \sigma_{ij}^2(s) e^{-2R(s)}\right) ds < \infty$. Then, the aggregate value process $(P(t))_{t \geq 0}$ *is defined by (setting $P(0) = 0$)*

$$\begin{aligned} P(t) &= \sum_{i=1}^{m} \int_0^t e^{-R(s)}\, dX_i(s)\,, \\ &= \sum_{i=1}^{m} \int_0^t \alpha_i(s)\, e^{-R(s)}\, ds + \sum_{j=1}^{d} \sum_{i=1}^{m} \int_0^t \sigma_{ij}(s)\, e^{-R(s)}\, dW_j(s)\,, \quad t \geq 0\,. \end{aligned} \tag{3.3}$$

Remark 3.2.3. (a) For every $t \geq 0$, the BMC model describes multivariate normally distributed cumulated cash flows $X_i(\cdot),\ i = 1, \dots, m$, with expectation

$$\mathbb{E} X_i(t) = X_i(0) + \int_0^t \alpha_i(s)\, ds\,, \qquad t \geq 0\,, \tag{3.4}$$

and cross-covariance function of $X_i(\cdot)$ and $X_k(\cdot),\ i, k = 1, \dots, m$, given by

$$\operatorname{cov}(X_i(t), X_k(t)) = \int_0^t \sum_{j=1}^{d} \sigma_{ij}(s)\, \sigma_{kj}(s)\, ds =: \int_0^t \mathbf{\Sigma}_{ik}(s)\, ds\,, \qquad t \geq 0\,.$$

Here, $\mathbf{\Sigma} := (\mathbf{\Sigma}_{ik}(t))_{ik}$ is called instantaneous cash flow covariance matrix, which is assumed to be positive definite for all $t \geq 0$.
(b) The variance of future cash flows $X_i(\cdot),\ i = 1, \dots d$, can be written as

$$\operatorname{var}(X_i(t)) = \int_0^t \mathbf{\Sigma}_{ii}(s)\, ds =: \int_0^t \sigma_i^2(s)\, ds\,, \qquad t \geq 0\,, \tag{3.5}$$

where $\sigma_i(\cdot),\ i = 1, \dots m$, are referred to as the instantaneous cash-flow volatilities.
(c) For nonzero $\sigma_i(\cdot)$, the cross-correlation function between $X_i(\cdot)$ and $X_k(\cdot)$ is

$$\operatorname{corr}(X_i(t), X_k(t)) = \frac{\operatorname{cov}(X_i(t), X_k(t))}{\sqrt{\operatorname{var}(X_i(t))}\,\sqrt{\operatorname{var}(X_k(t))}}\,, \qquad t \geq 0\,. \tag{3.6}$$

Informally, we denote the instantaneous correlation between $dX_i(\cdot)$ and $dX_k(\cdot)$ as

$$\operatorname{corr}(dX_i(t), dX_k(t)) = \frac{\mathbf{\Sigma}_{ik}(t)}{\sigma_i(t)\,\sigma_k(t)} =: \rho_{ik}(t)\,, \qquad t \geq 0\,.$$

(d) The value of the aggregate cash-flow process $X(t)$ at time $t \geq 0$ gives the total earnings of the bank that have been realised between 0 and t (cumulated earning). Its variance is given by

$$\begin{aligned}
\operatorname{var}(X(t)) &= \int_0^t \sum_{j=1}^{d} \left(\sum_{i=1}^{m} \sigma_{ij}(s) \right)^2 ds \\
&= \int_0^t \sum_{j=1}^{d} \sum_{i=1}^{m} \sum_{k=1}^{m} \sigma_{ij}(s)\,\sigma_{kj}(s)\,ds \\
&= \int_0^t \sum_{i=1}^{m} \sum_{k=1}^{m} \mathbf{\Sigma}_{ik}(s)\,ds =: \int_0^t \sigma^2(s)\,ds\,, \qquad t \geq 0\,,
\end{aligned} \tag{3.7}$$

where we call $\sigma(\cdot)$ the instantaneous aggregate cash-flow volatility.
(e) Note that the number d of independent Brownian motions $W_i(\cdot)$ can be different from the number m of business risk cells. Therefore, our model also allows for such realistic scenarios where the number of risk factors (represented by different Brownian motions) is greater than the number of clusters; think e.g. of a bank with two legal entities that are exposed to three risk factors, which affect their non-credit and non-market earnings and thus business risk. □

Example 3.2.4. [Bivariate BMC model with constant parameters]
Consider a simple bivariate BMC model with constant drift and diffusion parameters where the cash-flow processes are given by

$$\begin{aligned}
dX_1(t) &= \alpha_1\,dt + \sigma_1\,dW_1(t) \\
dX_2(t) &= \alpha_2\,dt + \sigma_2\big(\rho\,dW_1(t) + \sqrt{1-\rho^2}\,dW_2(t)\big)\,, \qquad t \geq 0\,.
\end{aligned} \tag{3.8}$$

From Remark 3.2.3 it follows that

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

implying for the variance of the i-th cumulated future cash flow

$$\mathrm{var}(X_i(t)) = \sigma_i^2 t\,, \qquad t \geq 0\,, \qquad i = 1, 2\,. \tag{3.9}$$

Moreover, the correlation between $X_1(\cdot)$ and $X_2(\cdot)$ and the instantaneous correlation are given for all $t \geq 0$ by

$$\mathrm{corr}(X_1(t), X_2(t)) = \mathrm{corr}(dX_1(t), dX_2(t)) = \rho\,.$$

Since all parameters are time-independent, this model can be calibrated quite easily. After discretisation of (3.8) by using the Euler method, σ_1 and σ_2 can be calculated directly from the standard deviations of the discrete increments $\Delta X_1(\cdot)$ and $\Delta X_2(\cdot)$. Then, according to (3.9), volatilities at a larger time-scale t can be derived by using the $\sqrt{t}$-scaling law. Finally, α_1 and α_2 can be estimated through the sample means of the discretisised incremental cash flows (3.8), or, alternatively, they can be obtained from the cumulated cash flows

$$\mathbb{E}X_i(t) = \alpha_i\, t + \text{const.}, \qquad t \geq 0\,,$$

for $i = 1, 2$ by regression analysis. □

Example 3.2.5. [Bivariate BMC model with time-dependent parameters]
We use a similar set up as in (3.8) but with time-dependent parameters for $t \geq 0$ of $\alpha_i(t) = \alpha_i t^{a_i}$ and $\sigma_i(t) = \sigma_i t^{b_i}$ for $a_i, b_i \geq 0$; $i = 1, 2$:

$$\begin{aligned} dX_1(t) &= \alpha_1\, t^{a_1}\, dt + \sigma_1\, t^{b_1}\, dW_1(t) \\ dX_2(t) &= \alpha_2\, t^{a_2}\, dt + \sigma_2\, t^{b_2} \big(\rho\, dW_1(t) + \sqrt{1-\rho^2}\, dW_2(t)\big)\,, \qquad t \geq 0\,. \end{aligned} \tag{3.10}$$

The expectations and variances of future cash flows can be calculated from (3.4) and (3.5); for $i = 1, 2$ one obtains

$$\mathbb{E}X_i(t) = \frac{\alpha_i}{1 + a_i}\, t^{1+a_i}\,, \qquad t \geq 0\,,$$

and

$$\mathrm{var}(X_i(t)) = \frac{\sigma_i^2}{2b_i + 1}\, t^{2b_i+1}\,, \qquad t \geq 0\,.$$

The instantaneous correlation in this model is still ρ, however, using (3.6), we have

$$\mathrm{corr}(X_1(t), X_2(t)) = \frac{\sqrt{2b_1 + 1}\,\sqrt{2b_2 + 1}}{1 + b_1 + b_2}\,\rho\,, \qquad t \geq 0\,.$$

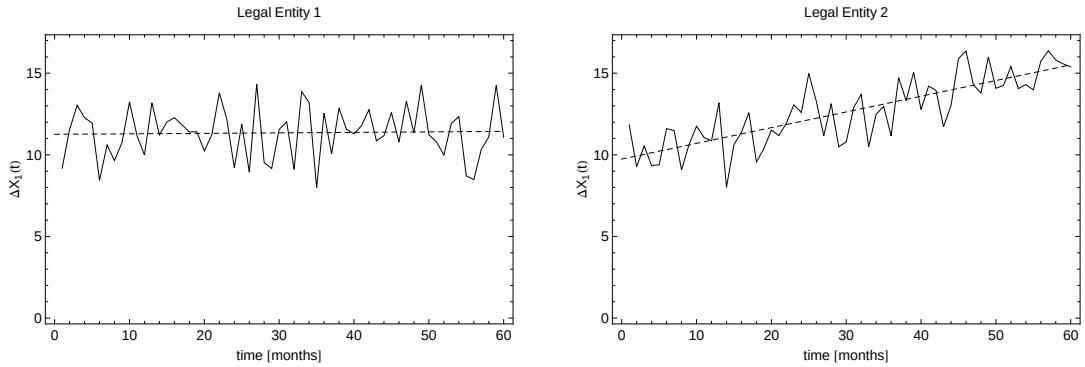

Figure 3.2.6. *Illustration of monthly earnings of two different legal entities as described in Example 3.2.5. In contrast to legal entity 1, the monthly earnings of legal entity 2 seems to have a positive linear drift which e.g. could be determined by linear regression (showed as dashed lines). Hence, in* (3.10) *one could set* $a_1 = 0$ *and* $a_2 = 1$*. Regarding the volatilities one could either use* $b_1 = b_2 = 0$ *(constant absolute volatilities) or* $b_1 = 0.5$ *and* $b_2 = 1.5$ *(constant Sharpe ratios).*

We now define the *signal-to-noise ratio*, also referred to as *Sharpe ratio*, for each cash-flow process as the ratio of its expected growths to the fluctuations, i.e. for $i = 1, 2$

$$\eta_i(t) = \frac{\mathbb{E}X_i(t) - X_i(0)}{\sqrt{\mathrm{var}(X_i(t))}}\,, \qquad t \geq 0\,. \tag{3.11}$$

Instead of constant volatilities $\sigma_i(\cdot) = \sigma_i$ as in the Example 3.2.4, we are now asking for constant Sharpe ratios $\eta_i(\cdot) = \eta_i$. Basically, this means that if we expect a certain relative growth (e.g. per year), the volatility is increasing with by same percentage, which makes sense from an economic perspective. Obviously, the Sharpe ratios of $X_1(\cdot)$ and $X_2(\cdot)$ are here constant for $b_i = a_i + \frac{1}{2}$. A typical situation as it may occour in practice is depicted in Figure 3.2.6, which shows the monthly earnings over 5 years for two hypothetical legal entities. □

Calculating business CAR. Let us now define business CAR, which depends on the distributional properties of the value process $(P(t))_{t\geq 0}$.

Definition 3.2.7. [Business CAR] *Consider different business risk cells with cash-flow processes* $X_i(\cdot)$*,* $i =, 1, \ldots, m$*, that are not attributable to other risk types, and define their corresponding market value process* $P(\cdot)$ *according to* (3.3)*. For* $t > 0$*, suppose that* F_t *is the distribution function of the value* $P(t)$ *with mean value* $\mathbb{E}P(t) < \infty$*. Then,* business CAR *at time horizon* t *and confidence level* $\kappa \in (0, 1)$ *is given by*

$$CAR_\kappa(t) = \mathbb{E}P(t) - F_t^{\leftarrow}(1 - \kappa)\,, \qquad t \geq 0\,, \tag{3.12}$$

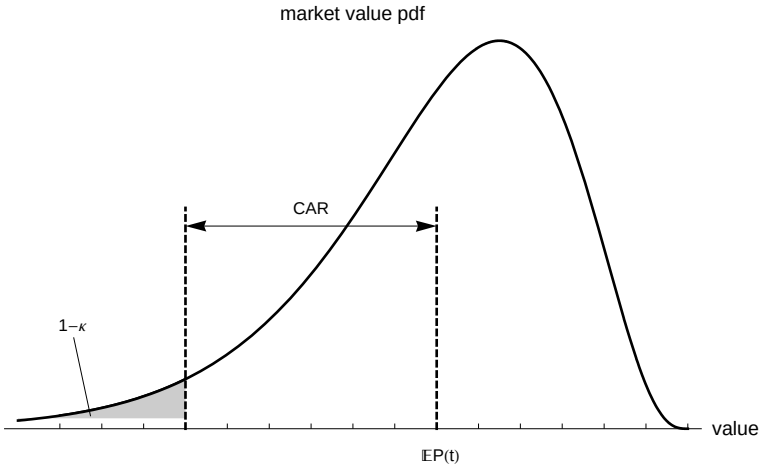

Figure 3.2.8. *Business CAR is defined as the difference between the expected market value* $\mathbb{E}P(t)$ *of the bank's aggregate value process and a very low quantile of its market value distribution* F_t*.*

where $F_t^{\leftarrow}(\kappa) = \inf\{x \in \mathbb{R} : \ F_t(x) \geq \kappa\}$, $0 < \kappa < 1$*, is the generalized inverse of* F_t*. If* F_t *is strictly increasing and continuous, we may write* $F_t^{\leftarrow}(\cdot) = F_t^{-1}(\cdot)$*.*

In the context of economic capital calculations, the confidence level κ is a number close to 1, e.g. $\kappa = 0.999$. In the case that the probability density function (pdf) of F_t exists, the definition of business CAR is illustrated in Figure 3.2.8.

In general, F_t and thus business CAR cannot be calculated analytically. For the BMC model, however, a closed-form expression for business CAR is available. This is a consequence of the following well-known result, which describes the distributional properties of the BMC model.

Proposition 3.2.9. *Consider the BMC model of Definition 3.2.2 with value process* $(P(t))_{t\geq 0}$*. Then, for every* $t > 0$*, the value* $P(t)$ *is normally distributed with expected value*

$$\mathbb{E}P(t) = \sum_{i=1}^{m} \int_0^t \alpha_i(s)\, e^{-R(s)}\, ds\,, \qquad t \geq 0\,, \tag{3.13}$$

and variance

$$\operatorname{var}(P(t)) = \int_0^t \sigma^2(s)\, e^{-2\,R(s)}\, ds\,, \qquad t \geq 0\,,$$

where $\sigma(\cdot)$ *is the instantaneous aggregate cash-flow volatility defined in* (3.7) *and* $R(\cdot)$ *the discount rate of Definition 3.2.1.*

Proof. Note that on the right-hand side of equation (3.3) the integrals in the first term are of the standard Riemann type, whereas the integrals in the second term given by

$I(t) = \int_0^t \sigma_{ij}(s)\, e^{-R(s)} dW_j(s)$ are Itô integrals with deterministic integrands. Then $I(t)$ is normally distributed with $\mathbb{E}I(t) = 0$ and $\mathrm{var}(I(t)) = \mathbb{E}\,(I(t)^2)$, see e.g. Shreve [66], Theorem 4.4.9. Using Itô's isometry we further obtain

$$\begin{aligned}
\mathrm{var}(P(t)) &= \sum_{j=1}^{d} \mathrm{var}\left(\sum_{i=1}^{m} \int_0^t \sigma_{ij}(s)\, e^{-R(s)}\, dW_j(s)\right) \\
&= \sum_{j=1}^{d} \mathbb{E}\left(\sum_{i=1}^{m} \int_0^t \sigma_{ij}(s)\, e^{-R(s)}\, dW_j(s)\right)^2 \\
&= \sum_{j=1}^{d} \int_0^t \left(\sum_{i=1}^{m} \sigma_{ij}(s)\, e^{-R(s)}\right)^2 ds \\
&= \int_0^t \sigma^2(s)\, e^{-2\,R(s)}\, ds\,, \qquad t \geq 0\,.
\end{aligned}$$

□

Since we now know that in the BMC model the bank's total market value $P(\cdot)$ is normally distributed, it is straightforward to calculate business CAR.

Theorem 3.2.10. [Business CAR for the BMC model] *Assume that future cash flows are described by a BMC model with instantaneous aggregate cash-flow volatility (see equation (3.7))*

$$\sigma(t) = \sum_{i=1}^{m} \sum_{j=1}^{d} \sum_{k=1}^{m} \sigma_{ij}(t)\, \sigma_{kj}(t)\,, \qquad t \geq 0\,, \tag{3.14}$$

and nonrandom discount rate $R(\cdot)$. Then, business CAR at time-horizon t and confidence level $\kappa \in (0,1)$ is given by

$$CAR_\kappa(t) = \Phi^{-1}(\kappa) \sqrt{\int_0^t \sigma^2(s)\, e^{-2\,R(s)}\, ds}\,, \qquad t \geq 0\,, \tag{3.15}$$

where Φ is the standard normal distribution function.

Proof. The assertion follows directly from Proposition 3.2.9 together with the definition of business CAR (3.12). □

This analytic expression for business CAR is mainly a consequence of using nonrandom interest rates. If one allows $R(\cdot)$ to be some continuous adapted interest rate process, the distribution function $F_t,\ t \geq 0$, of the aggregate value process $(P(t))_{t\geq 0}$ is

in general not normal anymore, and the result for business CAR will rarely be available in closed form.

Note that (3.15) only depends on the cash flows' covariance matrix and not on other model parameters such as drift parameters $\alpha_i(\cdot)$ or the initial values $X_i(0)$. As a consequence thereof, the BMC model can be calibrated easily, confer also Examples 3.2.4 and 3.2.5.

3.2.3 The Relationship Between EAR and CAR

Earnings-at-Risk. In the light of Definition 3.2.2 we now can qualify our notion of EAR, which we have already introduced in Section 3.2.1. From Remark 3.2.3 (d) we know that $\mathrm{var}(X(t))$ is the volatility of the bank's total aggregate cash flows accumulated between time 0 and t. It is also well-known that $X(t),\ t \geq 0$, is normally distributed, see e.g. Shreve [66], Theorem 4.4.9, so that EAR of the cumulated earnings $X(t)$ at $t \geq 0$ and confidence level $\kappa \in (0,1)$ is simply given by

$$\begin{aligned} \mathrm{EAR}_\kappa(t) &= \Phi^{-1}(\kappa)\,\sqrt{\mathrm{var}(X(t))} \\ &= \Phi^{-1}(\kappa)\,\sqrt{\int_0^t \sigma^2(s)\,ds}\,, \qquad t \geq 0\,, \end{aligned} \tag{3.16}$$

where Φ is the standard normal distribution function. In contrast to (3.15), we see that in (3.16) the volatility is not discounted. Moreover, it should be mentioned that according to what we have said in Section 3.2.1, the time parameter t in (3.16) should be chosen in a way that it reflects a *short-term horizon* so that discounting effects can be neglected. Finally, we define the instantaneous EAR as

$$\mathrm{ear}_\kappa(\cdot) = \Phi^{-1}(\kappa)\,\sigma(\cdot)\,. \tag{3.17}$$

EAR-CAR-transformation factors. An interesting question concerns the relation between EAR and CAR. Intuitively, CAR should be higher than EAR since CAR takes—in contrast to EAR—the *long-term uncertainty* of future cash flows into account. It has been suggested e.g. by Matten [49] or Saita [64, 65] that CAR is a constant multiplier of EAR, and that the multiplication factor depends only on a (risk-adjusted) discount factor and the time horizon t. Saita based his analysis of the EAR-CAR relationship on a discrete-time cash-flow model similar to (3.1) where EAR reflects the uncertainty of the ΔE_{t_i}, and [65], Section 5.8, gives a very readable overview about this topic.

In the case of the BMC model, we see by comparing (3.15) with (3.17) that such a proportionality between $\mathrm{EAR}(t)$ and $\mathrm{CAR}(t)$ does not hold for all $t \geq 0$ because of the time dependence of the instantaneous aggregate-cash-flow volatility $\sigma(\cdot)$. However,

the mean value theorem ensures that one can always chose a $\xi \in (0,t)$ so that (3.15) can be written as

$$\begin{aligned} \mathrm{CAR}_\kappa(t) &= \Phi^{-1}(\kappa)\,\sigma(\xi)\,\sqrt{\int_0^t e^{-2\,R(s)}\,ds} \\ &= \mathrm{ear}_\kappa(\xi)\,\sqrt{\int_0^t e^{-2\,R(s)}\,ds}\,, \qquad t \geq 0\,, \end{aligned}$$

and we can think of $\mathrm{ear}_\kappa(\xi)$ as an average EAR of the time interval $[0,t]$. The following two examples illustrate the relationship between EAR and CAR, in particular showing that—even in the quite simple framework of BMC models—EAR-CAR-transformation crucially depends on the specifications of the cash-flow processes $X_i(t),\ i =, 1, \ldots, d$.

Example 3.2.11. [BMC model with constant diffusion parameters]
Consider a BMC model with $\sigma_{ij}(\cdot) = \sigma_{ij} = \text{const.}$ for $i = 1, \ldots m; j = 1, \ldots d$. Then also the aggregate-cash-flow volatility (3.14) is constant and we obtain

$$\mathrm{CAR}_\kappa(t) = k_1(R,t)\,\Phi^{-1}(\kappa)\,\sigma = k_1(R,t)\,\mathrm{ear}_\kappa\,, \qquad t \geq 0\,,$$

with EAR-CAR transformation factor

$$k_1(R,t) = \int_0^t e^{-2\,R(s)}\,ds\,,$$

where we indicated the explicit dependence of k_1 on the discount rate $R(\cdot)$ by R. Hence, in this special case, business CAR is proportional to the (constant) instantaneous EAR. If furthermore the short-term discount rate is constant, $r(\cdot) = r$, we have that

$$R(t) = \int_0^t r\,d\tau = rt,$$

and we arrive at

$$k_1(r,t) = \sqrt{\frac{1 - \exp(-2\,r\,t)}{2\,r}}\,, \tag{3.18}$$

which is a simple function of the time horizon t and the short-term discount rate r. The higher r is, the smaller k_1 will be because future cash flows (and so their fluctuations) tend to have lower impact on the market value (and so on its uncertainty). Similarly, longer time horizons t lead to a growing k_1 because more uncertain future cash-flows are taken into account. In the limit $t \to \infty$, expression (3.18) simplifies to

$$\lim_{t\to\infty} k_1(r,t) = \frac{1}{\sqrt{2\,r}}\,.\,, \tag{3.19}$$

which, simply speaking, represents the maximum excess of CAR relative to the short-term measure EAR when all earning fluctuations up to infinity time horizon are taken into account. □

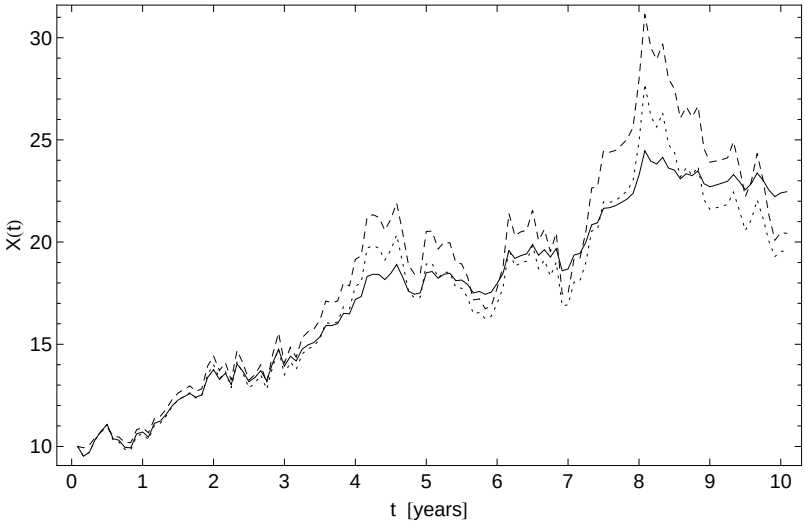

Figure 3.2.13. *Typical cash-flow paths $X_i(t)$ for BMC models with constant absolute cash-flow volatility (solid line) and constant Sharpe ratio (dashed line) referring to Examples 3.2.11 and 3.2.12, respectively. The parameters are $X_i(0) = 10$, $\alpha_i = 1.2$ and $\sigma_i = 1.4$ with a time horizon of 10 years and monthly increments. Moreover, the dotted path is obtained by the advanced model discussed in Section 3.3, especially Example 3.3.2, and we set $\tilde{\sigma}_i = \sigma_i / X_i(0) = 0.14$. We used the same seeds of normal random variables for three paths.*

Example 3.2.12. [BMC model with constant Sharpe ratio]
Consider a BMC model with cumulated cash-flows $X_i(\cdot)$ for $i = 1, \ldots, m$. As in Example 3.2.5 we consider the Sharpe ratios $\eta_i(\cdot)$ given by (3.11). Adopting a linear-growth model with constant drift parameters $\alpha_i(\cdot) = \alpha_i > 0$, $i = 1, \ldots, m$, we know from Example 3.2.5 that constant Sharpe ratios require square-root-of-time scalings of the instantaneous cash-flow volatilities, i.e. $\sigma_i(t) = c_i \sqrt{t}$ for some constants c_i, $i = 1, \ldots, m$. This implies that the aggregate-cash-flow volatility (3.14) for $t \geq 0$ can be written as $\sigma(t) = \sigma \sqrt{t}$, resulting in sample paths of $X_i(\cdot)$ that are in general more noisy than the one obtained in Example 3.2.11, a fact that is illustrated in the top panel of Figure 3.2.13.

Finally, using (3.15), we can calculate business CAR in the case of a constant short-term discount rate r to

$$\mathrm{CAR}_\kappa(t) = k_2(r,t)\, \Phi^{-1}(\kappa)\, \sigma\,, \qquad t \geq 0\,,$$

with EAR-CAR transformation factor

$$k_2(r,t) = \sqrt{\int_0^t s\, e^{-2rs}\, ds} = \frac{1}{2r} \sqrt{1 - (1+2rt)\exp(-2\,r\,t)} \tag{3.20}$$

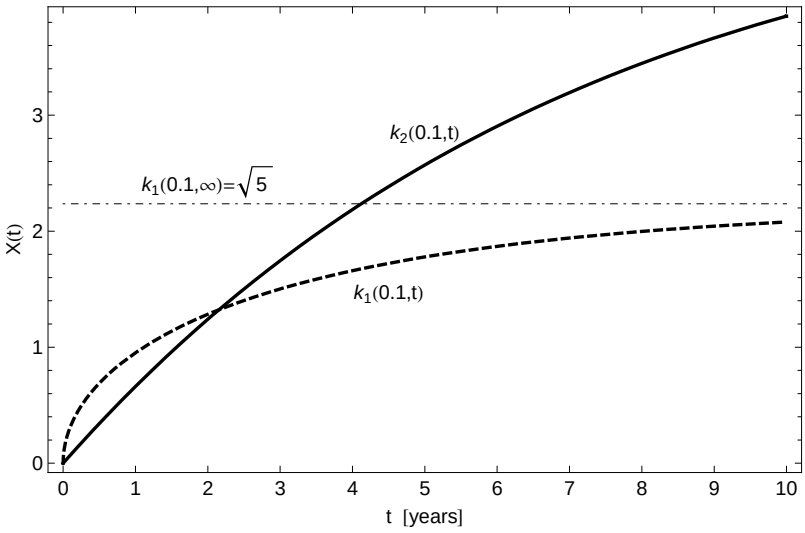

Figure 3.2.14. *EAR-CAR-transformation factors* $k_1(r,t)$ *and* $k_2(r,t)$ *according to* (3.18) *and* (3.20)*, respectively, as a function of time for a discount rate* $r = 0.1$*. The dotted-dashed line indicates the asymptote of* $k_1(r,t)$ *when* $t \to \infty$*.*

and

$$\lim_{t\to\infty} k_2(r,t) = \frac{1}{2\,r}\,. \tag{3.21}$$

□

Comparing the different BMC-model specifications of Example 3.2.11 and 3.2.12 with constant short-term discount rates r, one could expect that k_2 is greater than k_1 because in the previous example the Sharpe ratio was actually increasing with $\sqrt{t}$ (implying a decrease of future cash-flow fluctuation) whereas here it is constant by construction. Hence, k_2 accumulates more future uncertainty than k_1. As we see in Figure 3.2.14, this is indeed the case if t exceeds a certain threshold, i.e. $t > t_0(r)$, which for $r = 0.1$ is approximately given by $t_0(0.1) = 2.2$.

3.3 A Model with Level-Adjusted Volatility

In the BMC model the *absolute* changes of future cash flows $X_i(\cdot)$ are directly modelled by a Brownian motion, see equation (3.2), which in particular means that the uncertainty of a business cell's cash flow is independent from its absolute level. There is, however, no rational for this behaviour and intuitively one rather associates higher total earnings with higher earnings volatilities. As a possible remedy, one could describe future cash flows by a *geometric* Brownian motion as it is used e.g. for stock

prices in the famous famous Black-Scholes-Merton setting. Then, for $t \geq 0$ cumulated cash-flows $X_i(\cdot)$, $i = 1, \ldots, m$, are given by

$$X_i(t) = X_i(0) \exp\left[\int_0^t \left(\alpha_i(s) - \frac{1}{2}\sum_{j=1}^d \sigma_{ij}^2(s)\right) ds + \sum_{j=1}^d \sigma_{ij}(s)\, dW_j(s)\right],$$

implying an expected exponential cash-flow growth of

$$\mathbb{E}X_i(t) = X_i(0) \exp\left(\int_0^t \alpha_i(s) ds\right),$$

which, however, might be considered as too optimistic for most businesses. Alternatively, we suggest a model with still moderate growth but a kind of "cash-flow level adjusted" volatility. More precisely, we have the following definition.

Definition 3.3.1. [Level-adjusted BMC model] *Consider a d-dimensional standard Brownian motion* $(W_1(t), \ldots, W_d(t))_{t\geq 0}$ *on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$. *Then, the level-adjusted BMC model model consists of:*

(1) Cash-flow processes.
For each business risk cell, indexed by $i = 1, \ldots, m$, *cumulated future cash flows* $X_i(t)$ *for* $t \geq 0$ *are described by a* cash-flow process, *which is the strong continuous solution to the Itô-stochastic-differential equation*

$$dX_i(t) = \alpha_i(t)\, dt + X_i(t) \sum_{j=1}^d \tilde{\sigma}_{ij}(t)\, dW_j(t)\,, \qquad t \geq 0\,, \tag{3.22}$$

with aggregate cash-flow process

$$X(t) = \sum_{i=1}^m X_i(t) = \sum_{i=1}^m X_i(0) + \sum_{i=1}^m \int_0^t dX_i(s)\,.$$

Here, $\alpha_i(\cdot) > 0$, $i = 1, \ldots, m$, *and* $\tilde{\sigma}_{ij}(\cdot)$, $i = 1, \ldots, m; j = 1, \ldots, d$, *are nonrandom functions of time, satisfying the integrability conditions* $\int_0^t |\alpha_i(s)| ds < \infty$ *and* $\int_0^t \tilde{\sigma}_{ij}^2(s) ds < \infty$. *The matrix* $(\tilde{\sigma}_{ij}(t))_{ij}$ *is assumed to be positive definite for all* $t \geq 0$.

(2) Value process.
Let $R(\cdot) > 0$ *be a nonrandom discount rate (see Definition 3.2.1) so that* $\int_0^t \left(|\alpha_i(s)| e^{-R(s)} + \tilde{\sigma}_{ij}^2(s) e^{-2R(s)}\right) ds < \infty$. *Then, the* aggregate value process $(P(t))_{t\geq 0}$ *is defined by (setting* $P(0) = 0$*)*

$$P(t) \;=\; \sum_{i=1}^m \int_0^t e^{-R(s)}\, dX_i(s)\,, \qquad t \geq 0\,.$$

Let us first consider the cash-flow process and compare (3.22) with (3.2) of the BMC model. For the latter, the diffusion parameters σ_{ij} play the role of an *absolute* measure of uncertainty for the increments of $X_i(\cdot)$, whereas in (3.22) the $\tilde{\sigma}_{ij}$ describe the increments' fluctuations *relative* to the level of $X_i(\cdot)$. Furthermore, instead of (3.22) we may write

$$X_i(t) = X_i(0) + \int_0^t \alpha_i(s)\, ds + \sum_{j=1}^{d} \int_0^t X_i(s)\, \tilde{\sigma}_{ij}(s)\, dW_j(s)\,, \qquad t \geq 0\,, \tag{3.23}$$

and from the martingale property of the Itô integral it immediately follows that the expectation of $X_i(\cdot)$ is given by

$$\mathbb{E}X_i(t) = X_i(0) + \int_0^t \alpha_i(s)\, ds\,, \qquad t \geq 0\,,$$

i.e. it is the same as for the BMC model, and, in particular, the model does not exhibit exponential growth as it would be the case when geometric Brownian motion is used. We close this Section with an extended example that illustrates some properties of the level-adjusted volatility model.

Example 3.3.2. [Constant drift and diffusion parameters]
For the sake of simplicity we focus on the case of constant parameters $\alpha_i(\cdot) = \alpha_i$ and $\tilde{\sigma}_{ij}(\cdot) = \tilde{\sigma}_{ij}$. Note however, that the diffusion parameter of the process (3.22) is random and given by $X_i(\cdot)\,\tilde{\sigma}_{ij}$. In order to find a solution for (3.23) we define the function

$$F(t, W(t)) := \exp\big(-\sum_{j=1}^{d} \tilde{\sigma}_{ij} W_j(t) + \frac{1}{2}\sum_{j=1}^{d} \tilde{\sigma}_{ij}^2\, t\big)\,, \qquad t \geq 0\,.$$

Then, by using Itô's formula the differential of the product FX_i can be calculated as

$$d\big(F(t, W(t))\, X_i(t)\big) = \alpha_i\, F(t, W(t))\, dt\,,$$

which after integration finally yields (setting $W_j(0) = 0$ for $j = 1, \ldots, d$),

$$X_i(t) = F(-t, -W(t)) \left(X_i(0) + \alpha_i \int_0^t F(s, W(s))\, ds \right)\,, \qquad t \geq 0\,. \tag{3.24}$$

According to (3.24), the cumulated cash-flows $X_i(t)$ at time $t \geq 0$ are not normally distributed as they are in the BMC model. A one-dimensional example for a typical path of $X_i(\cdot)$ according to (3.24) is plotted as a dotted line in Figure 3.2.13.

The Itô representation of the value process $(P(t))_{t\geq 0}$ is given by

$$P(t) = \sum_{i=1}^{m} \alpha_i \int_0^t e^{-R(s)}\, ds + \sum_{j=1}^{d} \sum_{i=1}^{m} \tilde{\sigma}_{ij} \int_0^t X_i(s)\, e^{-R(s)}\, dW_j(s)\,, \quad t \geq 0\,, \tag{3.25}$$

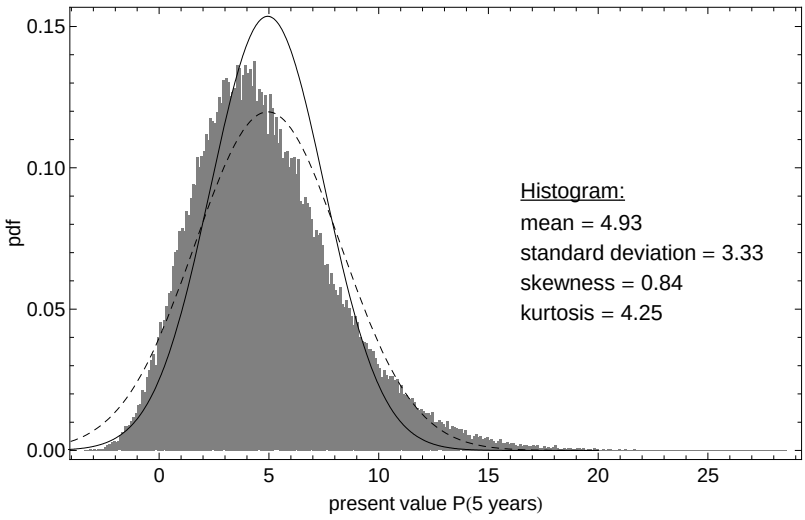

Figure 3.3.3. *The histogram shows the simulated present-value distribution for $t = 5$ years of the level-adjusted volatility model as discussed in Example 3.3.2 (with discount rate $R = rt$) as well as the mean, standard deviation, skewness, and kurtosis parameters of the simulated data. This is compared to a normal distribution with the same mean and standard deviation as the simulated data, plotted as a dashed line. The solid curve represents the normally distributed present value as obtained from the BMC model of Example 3.2.11. We set $X_i(0) = 10$ and use the yearly parameters $\alpha_i = 1.2$, $\sigma_i = 1.4$, $\tilde{\sigma}_i = \sigma_i / X_i(0) = 0.14$, and a discount rate of $r = 0.08$.*

which cannot be calculated in closed form. Note, however, that the expectation of $P(\cdot)$ is again given by (3.13) and therefore is the same as for the BMC model. This can also be seen in Figure 3.3.3 where we compare the distribution of the present value as obtained by (3.25) firstly to that of a normal distribution with the same mean and variance (dashed curve), and, secondly, to the normally distributed present value calculated from the BMC model of Example 3.2.11 (solid curve). One can see that (3.25) leads to a distribution that is more skewed to the right (positive skewness) and is more peaked and heavier-tailed than a normal distribution with the same variance (kurtosis larger than 3). □

EAR-CAR-transformation revisited. We have seen in Section 3.2.2 that for the BMC model both the cumulated cash-flows $X_i(\cdot)$ and the present value $P(\cdot)$ are normally distributed. This has two important consequences. First, business CAR (and thus the EAR-CAR-transformation factor) of the BMC model is independent of the expected growth and thus of $\alpha_i(\cdot)$, $i = 1, \ldots, m$. Second, the transformation factor is invariant under changes of the confidence level and equals the ratio of the present value volatility and the earnings volatility (see Examples 3.2.11 and 3.2.12). This is

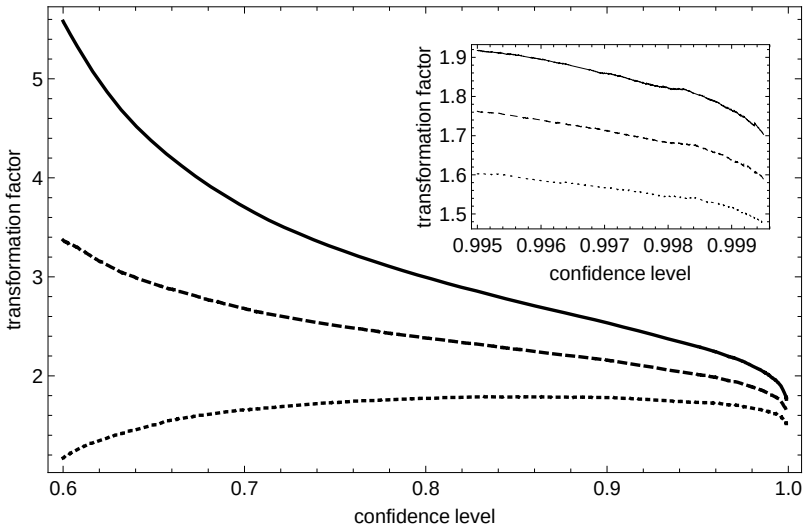

Figure 3.3.4. *EAR-CAR-transformation factor for the level-adjusted growth model of Example 3.3.2 for $t = 5$ years as a function of the confidence level κ and different values $\alpha_1 = 1.0$ (solid line), $\alpha_2 = 1.2$ (dashed line), and $\alpha_3 = 1.4$ (dotted line) of the growth parameter. We used $X_i(0) = 10$, $\tilde{\sigma}_i = 0.14$, and a constant discount rate of $r = 0.08$.*

a consequence of the well-known property of elliptically distributed random variables, which says that their quantiles can always be expressed in terms of their standard deviation.

However, such a behaviour cannot be expected for the level-adjusted BMC model because $P(\cdot)$ is not elliptically distributed. Figure 3.3.4 serves to illustrate this for the model of Example 3.3.2. It shows the results of a simulation study where the transformation factor between CAR and EAR (calculated at $t = 0$ where EAR is normally distributed with standard deviation $\sigma_i = \tilde{\sigma}_i X_i(0)$) is plotted as a function of their confidence level. The growth parameter is set to be $\alpha_i = 1.0, 1.2$, and 1.4. Note that the higher the growth rate α_i is, the lower the transformation factor and therefore the ratio between CAR and EAR will be. In contrast, if we compare the ratios of the volatilities of the present value and EAR, i.e. the volatility of the simulated histogram data to the initial absolute cash-flow volatility $\sigma_i = \tilde{\sigma}_i X_i(0) = 1.4$, we obtain 2.29, 2.38, and 2.46 for $\alpha_i = 1.0, 1.2$, and 1.4, respectively. Moreover, we see that an increasing confidence level leads to a decreasing transformation factor.

Summing up we can conclude that in general the question of how a EAR can be converted into an CAR is not straightforward to answer. While for the BMC model this seems to be easier and intuitively easier to grasp (since independent of the confidence level and the expected growth rate) it becomes rather involved for more general models like the one discussed in this Section.

3.4 Conclusion and Outlook

In this Chapter we suggested a multivariate continuous-time setting for assessing business risk using stochastic models for the future cash-flows of (non-credit, non-market, etc.) earnings. In contrast to scenario-based methods for estimating business risk, our model has the advantage that it results in a time-dependent probability distribution of future earnings, which allows for an accurate definition of VAR-like risk measures at any confidence level and time horizon.

We also investigated the relationship between EAR and CAR, which for general cash-flow processes turns out to be not straightforward, and, in particular, a constant multiplier converting EAR into CAR is usually not available. However, a simple EAR-CAR-transformation factor only depending on the time horizon and the discount rate can analytically be derived for the simple BMC model. Such a result may be useful when a fast and approximative VAR estimation based on some EAR figure is needed.

Since in our model the dependence structure between different legal entities or business units is reflected by their correlation matrix, risk-reducing strategies such as known from stock portfolio analysis can be straightforwardly applied.

One could think of several enhancements of the BMC model. A particular interesting possibility would be to introduce jumps in the earnings processes representing sudden and sharp falls in a company's earnings caused by e.g. a change in the competitor environment or a customer shift. A particular important class of jump process are the *Lévy processes*, which have become quite popular also in financial modelling, see e.g. Cont & Tankov [24]. Then, in addition to the covariance structure of the multivariate Brownian motion we already discussed, jump dependence between future earnings could be modelled by so-called Lévy copulas, see e.g. Böcker & Klüppelberg [17] for an application of this technique to operational risk. However, since every model should be seen also in the light of the data needed for its calibration, it may be wise to start with a well-known and established Brownian approach.

In our opinion, the development of advanced business risk models will be an important task in quantitative risk management, despite the difficulties and complexity discussed above. Equally important is to work on a harmonised definition of this material risk type, which clearly requires a closer collaboration between practitioners and academics.

Chapter 4

An Analysis of Inter-Risk Correlation and Risk Aggregation with Focus on Market and Credit Risk

Here we consider in detail the aggregation-by-risk-type approach, which can be used as an approximation to the total risk of a financial firm. In particular, we show how expert knowledge can be successfully exploited to estimate inter-risk correlations, and, moreover, we compare different aggregation methods based on copulas and linear correlation. We then consider the interaction between credit risk and market risk by combining a Merton-like factor model for credit risk with linear factor model for market risk. We analytically calculate their inter-risk correlation and show how inter-risk correlation bounds can be derived. Furthermore, we suggest a new estimator for the Gaussian copula parameter that can be applied to almost arbitrary credit portfolios.

4.1 Introduction

A core element of modern risk management and control is analysing the *capital adequacy* of a financial institution, which is concerned with the assessment of the firm's required capital to cover the risks it takes. To this end, financial firms seek to quantify their overall risk exposure by aggregating all individual losses associated with different risk types or business units, and then compare this figure with a so-called *risk taking capacity*, defined as the total amount of capital serving as a buffer against potential losses.

Until now no standard procedure for risk aggregation has emerged. According to

an industry survey of The Joint Forum [2], a widespread approach in the banking industry is *aggregation across risk types* where the marginal loss distributions of all relevant risk types are independently modelled from their dependence structure. This approach splits up into three steps:

- First, assign every individual risk position to a certain risk type.
- Second, calculate an aggregated measure (such as VaR) for every risk type by using separate, risk-type specific techniques and methodologies.
- Third, integrate all pre-aggregated risk figures of different risk types to obtain the overall capital number, henceforth called *aggregated risk capital*.

The easiest solution for the last step is simply to add up all pre-aggregated risk figures; this however, is only a rough estimate of the bank-wide total risk exposure. Moreover, banks usually try to reduce their overall risk by accounting for diversification between different risk types because it allows them either to reduce their capital buffer (and thus expensive equity capital) or to increase their business volume (and thus to generate additional earnings).

In practice, a widespread approach for aggregating different risk types is the so-called *square-root-formula approach* or *variance-covariance approach*. Though mathematically justified only in the case of elliptically distributed risk types (with the multivariate normal or t-distribution as prominent examples), this approach is very often used as a first approximation because then aggregated economic capital (EC) can be determined without simulation. In doing so, let us describe pre-aggregated losses associated with the i-th risk type by a random variable $X_i,\ i = 1, \ldots, m$, and define the associated EC at confidence level $\alpha \in (0, 1)$ by (for the sake of clarity, we drop the time-horizon index t)

$$\mathrm{EC}_{X_i}(\alpha) = F_i^{-1}(\alpha) - \mathbb{E}X_i \,,$$

where $\mathbb{E}X_i$ is the expected loss of risk type i (which we assume exists), and $F_i^{-1}(\cdot)$ is the inverse of the distribution function $F_i(\cdot) = P(X_i \leq \cdot)$, which here and in the sequel is always assumed to be strictly increasing and continuous. If moreover the linear correlations between risk types are denoted by $\rho_{ij} = \mathrm{corr}(X_i, X_j)$ for $i, j = 1, \ldots, m$, then aggregated economic capital $\mathrm{EC}^+(\alpha) := \mathrm{EC}_{X_1+\cdots+X_m}(\alpha)$ for elliptically distributed $(X_1, \ldots, X_m)$ is given by

$$\mathrm{EC}^+(\alpha) = \sqrt{\sum_{i,j=1}^{m} \mathrm{EC}_{X_i}(\alpha)\, \rho_{ij}\, \mathrm{EC}_{X_j}(\alpha)} \,. \tag{4.1}$$

Hence, a typical problem of risk aggregation is the calculation of the *inter-risk correlation matrix* $\mathbf{C} := (\rho_{ij}),\ i, j = 1, \ldots, m$, and its estimation and calibration is a core problem for implementing aggregation by risk type in practice.

A technique that is more advanced compared to the variance-covariance approach is based on copulas. In the last years, copulas have become very popular in basically all areas of financial modelling. Applied to bank-wide risk aggregation, copula techniques are discussed e.g. in Dimakos & Aas [28], Rosenberg & Schuermann [63], or Ward & Lee [70]. Here, however, we want to move in a slightly different direction and put particular emphasis on the fact that, in the context of risk aggregation, reliable empirical loss data are rare, and, consequently, sound statistical estimates of the inter-risk correlation are difficult to obtain. Therefore, we stress the usage of expert knowledge and show new ways how this could be done in a proper way. Moreover, we illustrate how the condition of positive semi-definiteness of the inter-risk-correlation matrix can be used to set bounds on unknown matrix elements.

4.2 Aggregation by Risk Type

4.2.1 Inter-Risk Correlation

As already said, the variance-covariance method is currently the dominant approach for risk aggregation and, consequently, dependence between different risk types is described in terms of *linear (Pearson) correlation*, which for two random variables X and Y is defined as

$$\rho = \operatorname{corr}(X, Y) = \frac{\operatorname{cov}(X, Y)}{\operatorname{var}(X) \operatorname{var}(X)},$$

where $\operatorname{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}X\,\mathbb{E}Y$ is the covariance between X and Y and $\operatorname{var}(X)$, $\operatorname{var}(Y)$ the variance of X and Y.

It is important to understand how inter-risk correlation is usually determined in practice. First, it can be derived from a purely statistical analysis based on empirical data, using e.g. Pearson product-moment correlation. For sample data $(x_i, y_i),\ i = 1, \ldots, n$, it is given by

$$\hat{\rho} = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2} \sqrt{\sum_i^n (y_i - \bar{y})^2}}, \tag{4.2}$$

where $\bar{x}$ and $\bar{y}$ are the sample means of X and Y, respectively. There are many other estimators for linear correlation available that are often superior to (4.2), see e.g. Devlin, Gnanadesikan & Kettering [27], or Lindskog [45] for overviews about statistical properties of different estimators.

However, since data in sufficient volume and quality are often hard to come by, a second approach relies on expert judgements, and in our estimation expert-based inter-risk correlations are widely used when risk aggregation is performed in practice. Therefore, evaluation of expert knowledge merits closer examination and this will be done in Section 4.2.4.

4.2.2 Copulas and Tail Dependence

Besides the variance-covariance approach where inter-risk dependence is completely captured by linear correlation, different risk figures can alternatively be aggregated by means of a *distributional copula*, which has become a standard concept for characterising the dependence structure between different distribution functions. General readable introductions about this subject with focus on financial applications are Cherubini, Luciano & Vecchiato [23] or Embrechts, Lindskog & McNeil [30].

Distributional copulas are multivariate distribution functions with uniform marginals. They are used for dependence modelling within the context of Sklar's theorem, which says that any multivariate distribution with continuous marginals can be transformed into a multivariate distribution with uniform marginals, called the copula. Copulas have the important property that they are invariant under strictly increasing transformations. As a consequence thereof, for a multivariate distribution function $F(x_1, \ldots, x_m)$, the copula is that part that does not depend on its marginal distributions $F_1(x_1), \ldots, F_m(x_m)$.

Theorem 4.2.1. [Sklar's Theorem, see e.g. Nelsen [52], Joe [38]] *Let F be a d-dimensional distribution function with marginals $F_1, \ldots, F_d$. Then there exists a copula $C : [0,1]^d \to [0,1]$ such that for all $x_1, \ldots, x_d \in [0, \infty]$*

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)). \tag{4.3}$$

If the marginal distribution functions $F_1, \ldots, F_d$ are continuous, then this copula is unique. Otherwise, it is unique on $\mathrm{Ran}F_1 \times \cdots \times \mathrm{Ran}F_d$.
Conversely, if C is a copula and $F_1, \ldots, F_d$ are marginal distribution functions, then function F in (4.3) *defines a d-dimensional distribution function with marginal distribution functions $F_1, \ldots, F_d$.*

Among all copulas discussed in the literature, maybe those most frequently used for risk aggregation are the *Gaussian copula* and the *t copula*, and both allow for natural extensions of the variance-covariance approach. Recall that equation (4.1) holds only for multivariate elliptically distributed risk types, i.e. having elliptical marginals *and* an elliptical dependence structure. Now, because both the Gaussian copula and the

t copula describe elliptical dependence, they allow to model the same dependence structure as in (4.1), however, in combination with more general, non-elliptical marginal distributions for each risk type.

For basically all practically relevant cases, the Gaussian copula and t copula can be parameterised by means of a correlation matrix $\mathbf{R}$, and, in the case of the t copula, by an additional degree-of-freedom parameter $\nu > 0$, which controls the degree of extreme dependence between the *tails* of the marginal distribution functions. More generally, one can define the important concept of (upper) *tail dependence*, which for two random variables X and Y with distribution functions F_X and F_Y is given by

$$\lambda := \lim_{u \to 1} P\big(Y > F_Y^{-1}(u) \,|\, X > F_X^{-1}(u)\big)\,, \tag{4.4}$$

where $F_X^{-1}(\cdot)$ and $F_Y^{-1}(\cdot)$ denote the quantiles of the distribution function of X and Y, respectively. Hence, focusing on risk aggregation, tail dependence λ between two risk types is the conditional probability that an (asymptotically) high loss in risk type Y occurs, given that another (asymptotically) high loss in risk type X has already been observed. If $\lambda = 0$ then X and Y are called tail-independent. Moreover, it can be shown (Embrechts et al. [30]) that, in contrast to linear correlation, λ is solely a property of the underlying copula, i.e. tail dependence does not depend on the specific marginals F_X and F_Y. For instance, for a bivariate t copula with correlation parameter R and degree-of-freedom parameter ν one has (see again Embrechts et al. [30])

$$\lambda_t(R, \nu) = 2 - 2F_{\nu+1}\big(\sqrt{\nu+1}\sqrt{1-R}/\sqrt{1+R}\big)\,, \tag{4.5}$$

where F_ν is the Student-t distribution function with ν degrees of freedom. Note that the tail dependence (4.5) is not vanishing even for $R = 0$. As a matter of fact, bivariate t distributed random variables are still dependent even if they are uncorrelated.

The different dependence structures of a bivariate Gaussian copula on one hand and a bivariate t copula with $\nu = 3$ on the other hand are depicted in Figure 4.2.2. For each scatter plot of joint losses (defined as positive values), we assume a Student-t distribution for the first risk type (say market risk), and a Vasicek distribution for the second risk type, therefore reflecting credit loss of a loan portfolio (see e.g. Vasicek [69], and Table 4.2.6 for details about the parametrisation we are using here). Furthermore, we used a correlation parameter of $R = 0.66$ for both copulas. While the bodies of both plots (i.e. the area of small or medium-sized losses of market and credit risk) are quite similar, we see that with increasing loss sizes the t copula causes less scattering and more points that are concentrated around the 45-degree line compared to the Gaussian copula. Consequently, losses in market and credit risk are more likely to occur simultaneously if their dependence is given by a t copula instead of a Gaussian

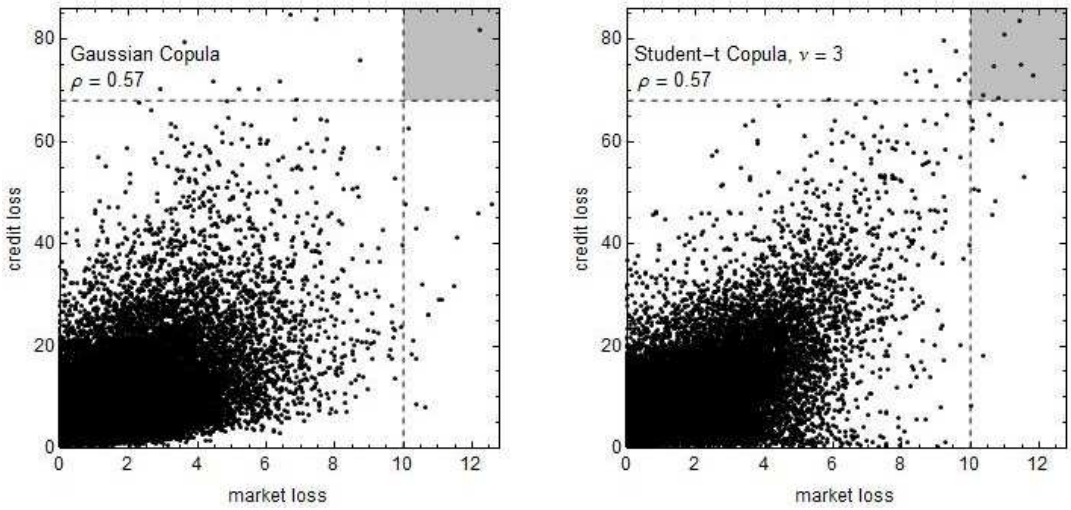

Figure 4.2.2. *Scatter plot of a simulation of joint losses in market and credit risk. Market risk losses are described by a Student-t distribution with degrees of freedom 10 and a 99.95 % quantile of 10 (indicated by the vertical dashed line). Credit risk follows a Vasicek distribution with average default probability $p = 0.3$ %, uniform asset correlation $\varrho = 8$ %, and a 99.95 % quantile of about 68 (indicated by the horizontal dashed line). The correlation parameter of both copulas is set to $R = 0.66$ and for the t copula we further have $\nu = 3$. Note that the linear correlation between both risk types is different from R and for both copulas estimated to be about $\rho = 0.57$.*

one. Especially, there is a significant amount of extreme joint losses above each risk type's 99.95 % quantile (indicated by the gray-shaded region in the upper right edge of the panel) in the case of the t copula, whereas for the Gaussian copula only very few extreme joint losses can be observed. This observation is not just a random effect but can be explained systematically. One can show that a Gaussian copula with correlation parameter $R < 1$ does not have tail dependence, i.e. $\lambda_{\text{Gauss}} = 0$, and therefore very extreme losses do not tend to occur simultaneously. In contrast, the t copula has a non-zero tail dependence (even for $R = 0$), and λ_t is increasing in the correlation parameter R and decreasing in ν. Furthermore, with growing ν, the t copula fades into the Gaussian copula so that in the limit $\nu \to \infty$ we obtain $\lambda_t = 0$. This property is useful for practical applications such as stress testing of inter-risk aggregation because one can use the t copula to explore the impact of different degrees of tail dependence on the aggregated EC.

It is worthwhile mentioning that the correlation matrix $\mathbf{R}$ of the Gaussian or t copula does not equal the usual inter-risk correlation matrix $\mathbf{C}$ calculated from empirical

loss data. In particular, for pairwise correlations $R_{ij}, i, j = 1, \ldots, m$, in principle all values of the interval $[0, 1]$ can be attained, which we know is in general not true for $\rho_{ij}, i, j = 1, \ldots, m$, and comonotonicity (countermonotonicity) between risk types X_i and X_j is reached for $R_{ij} = 1$ ($R_{ij} = -1$). However, as we will see in the next Section, there are algebraic boundary conditions for the whole matrix $\mathbf{R}$, which again may restrict some R_{ij} to a subinterval of $[0, 1]$. As an example, consider again Figure 4.2.2 where the copula parameter between market and credit risk is set to $R = 0.66$. The corresponding linear correlation can be estimated by applying (4.2) to the sample data shown in the scatter plots, yielding for both copulas a value of about $\rho = 0.57$.

A useful estimator for $\mathbf{R}$ is based on Kendall's tau rank correlation τ, which is defined as the probability of concordance minus the probability of discordance, i.e.

$$\tau(X, Y) = P[(X - X')(Y - Y') > 0] - P[(X - X')(Y - Y') < 0]\,,$$

where (X', Y') is an independent copy of (X, Y). If $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ are two samples of X and Y, then the standard estimator for Kendall's tau is given by

$$\hat{\tau} = \frac{n_c - n_d}{\sqrt{(n_c + n_d + n_x)(n_c + n_d + n_y)}}\,.$$

Here, n_c is the number of concordant pairs of observations and n_d is the number of discordant pairs, where a concordant pair (x_i, y_i) and (x_j, y_j) is one such that both $x_i > x_j$ and $y_i > y_j$, or both $x_i < x_j$ and $y_i < y_j$. A discordant pair is one such that $x_i > x_j$ and $y_i < y_j$, or $x_i < x_j$ and $y_i > y_j$. Moreover, n_x denotes the number of so-called ties where $x_i = x_j$, and similarly, n_y counts the ties in y in which $y_i = y_j$.

It was pointed out by Embrechts et al. [30], and Lindskog, McNeil & Schmock [47] that Kendall's tau can be used to construct a robust and efficient estimator for the correlation parameters R_{ij} and thus the matrix $\mathbf{R}$ of a Gaussian or t copula. In doing so, we note that for elliptical distributions without atoms, Kendalls tau is linked to the linear correlation coefficient by

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

Now, since copulas are invariant under strictly increasing transformations, relationship (4.6) also holds true for basically all marginal risk type distributions and can therefore be used as an estimator for R in the case of elliptical dependence structures.

4.2.3 Inter-Risk Correlation Matrix

We now assume that, using such methods as described in the previous Section, for given risk types $X_i,\ i = 1, \ldots, m$, we have estimated their linear correlations ρ_{ij}, or,

alternatively, the correlation parameters R_{ij} of a Gaussian or t copula. By definition, in order to be a proper correlation matrix, the pairwise correlation estimates must satisfy several conditions: The matrices $\mathbf{C}$ and $\mathbf{R}$ must be symmetric and all their diagonal elements must be equal to 1, moreover, they have to be *positive semi-definite*. The first two conditions can be checked easily, and for the latter there exist several mathematically equivalent ways of testing it. Using modern computer algebra systems, a particular convenient method is to calculate the eigenvalues of the matrix; then, if all eigenvalues are non-negative, the matrix is positive semi-definite and hence a proper correlation matrix. If some eigenvalues are negative, a positive semi-definite matrix has to be generated from the raw before risk aggregation can be performed. Algorithms for doing this are discussed e.g. in Higham [35], Rebonato & Jäckel [59], or Lindskog [45], who came to the conclusion that the best technique is the *eigenvalue method* where negative eigenvalues of the raw matrix are replaced by a small positive constant $\delta > 0$, and then the new matrix is calculated invoking standard results from linear algebra.

In order to illustrate the potential impact of the positive semi-definitess requirement on calculating an inter-risk correlation matrix, let us consider the following four-dimensional symmetric matrix based on average industry data according the survey of the IFRI/CRO Forum [36] (see Section 4.2.5 for further details):

$$\mathbf{R} = \left(\begin{array}{c|cccc} & \text{market} & \text{credit} & \text{oprisk} & \text{business} \\ \hline \text{market} & 1 & 0.66 & 0.30 & R \\ \text{credit} & 0.66 & 1 & 0.30 & R \\ \text{oprisk} & 0.30 & 0.30 & 1 & R \\ \text{business} & R & R & R & 1 \end{array}\right), \tag{4.7}$$

where we assumed that business risk has an unknown but uniform correlation parameter R with other risks. What can we say about the possible values of R, for instance, would it be possible to adopt the conservative assumption of $R = 1$? A look at Figure 4.2.3 reveals that $\mathbf{R}$ is a proper correlation matrix only if R is smaller than about 0.78, which therefore is the correct upper bound (instead of $R = 1$) to be used for a conservative calculation of aggregated EC. Though the property of positive semi-definiteness is in general difficult to grasp, consider the following intuitive argument: Assume that R in (4.7) is set to the extreme value of $R = 1$, implying that business risk comonotonously depends on market risk, credit risk, and operational risk, i.e. their loss variables always move in lock steps. This, however, implies at least a medium degree of comovement also between e.g. operational risk and market risk, or between operational risk and credit risk, since their random variables are coupled by virtue of the business risk random variable. This is, however, in contradiction to their relatively low correlation of 0.3 in (4.7). Consequently, the coupling of business risk to the other risk types must be less

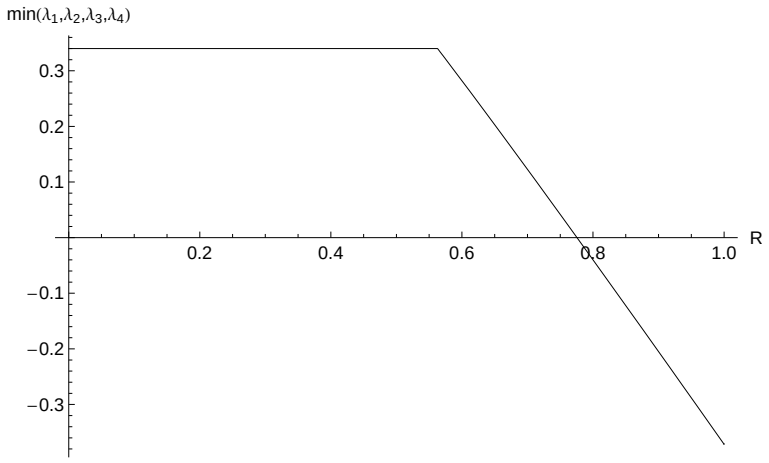

Figure 4.2.3. *Plot of* $\min(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ *of all eigenvalues* λ_i *of the symmetric matrix* (4.7) *as a function of the unknown correlation parameter* R *of business risk. For* $R < 0.78$ *all eigenvalues are positive and* $\boldsymbol{R}$ *becomes a well-defined correlation matrix.*

strong, implying an inter-risk correlation that is somewhat smaller than 1.

4.2.4 Elicitation of Expert Knowledge

Introduction

Because of the often scarce empirical data it is usually not guaranteed that statistical estimators like (4.2) yield reliable and robust results for the inter-risk correlation, and in this case it is necessary to draw on expert opinions. This has also been acknowledged by the Committee of European Banking Supervisors in their recent consultation paper [26], where they explicitly distinguish statistical techniques versus expert judgements. However, because elicitation of expert knowledge is not an easy task, supervisors are usually concerned about how the subjectivity of expert estimates can be addressed and controlled.

Therefore, in this Section we try to shed some light on this subject matter and review some basic concepts and key results about expert judgements. A reliable elicitation of expert probabilities requires knowledge both in statistics and in psychology of judgements under uncertainty. Excellent overviews about the methodological framework of elicitation of probabilities are O'Hagan et al. [56] or Jenkinson [37], and they were the source of inspiration for the following brief presentation on this intriguing subject.

To understand how experts construct probability judgements it is important to dis-

tinguish *frequency probability* versus personal or *subjective probability*. Simply stated, frequency probability of an *event* is the relative proportion of its occurrence in an, ideally infinitive, series of repetitions, think e.g. of the relative number of "heads" in a (obviously repeatable) coin-tossing game. Indeed, the *strong law of large numbers* ensures that the sample average converges in probability towards the expected value, i.e. for the coin-tossing game we have that $P(\lim_{n\to\infty} \frac{H_n}{n} = \frac{1}{2}) = 1$, where H_n denotes the number of "heads" in a series of n trials. Hence, frequency probability can be used to describe the statistical uncertainty (also referred to as *aleatory* uncertainty) in one ore more instances of a random process. On the other hand, if there is also systematic uncertainty (also referred to as *epistemic* uncertainty) due to imperfect knowledge, or if the "game" cannot be repeated, the definition of frequency probability does not suffices. Put another way, one then has to associate the notion of probability with the uncertainty of knowledge rather than with the results of repeatable experiments. This leads to the definition of subjective probability as a person's *degree of belief* in an uncertain proposition.

Given the frequency interpretation of probability, it is clear that also the personal probability should not contradict some formal laws of probability, such as the multiplication law $P(A \cap B) = P(A)\,P(B)$ for independent events A and B. However, empirical research revealed that, in practice, people's judgement of probabilities is often biased and non-coherent, i.e. violating formal axioms of probability. The reasons for that are manifold, e.g. the expert's statement of probability crucially depends on the facilitator's questions but also on the response scale on which the expert is asked to express his personal probability. Another reason is that the information on which the respondent's judgement is built may be based only on a sub-sample or subjective selection of what is potentially relevant. A ground-breaking work regarding the psychological foundation that governs the heuristics that are employed in making judgements of probabilities was the *heuristics and biases* research programme by Tversky & Kahneman [67][1]. According to [67], people rely their judgements of probability on some heuristic principles instead of a more complex assessment that requires more effort. In general, such approximative heuristics work sufficiently well "but sometimes they lead to severe and systematic errors." ([67], p. 1124). By way of illustration, let us consider the following frequently cited experiment, see e.g. Tversky & Kahneman [68]:

> Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student she was deeply concerned with issues of discrim-

[1]Kahneman was awarded the 2002 Nobel Prize in Economic Sciences for this and related work, and Tversky, who died in 1996, was acknowledged in the announcement because the Royal Swedish Academy of Sciences does not award prizes posthumously.

ination and social justice and also participated in antinuclear demonstrations.

Respondents were asked to tick the most likely alternative between

- Linda is a bank teller,
- Linda is a bank teller and is active in the feminist movement.

The majority of people (typically more than 80 % percent) rate the second opinion as more likely than the first one; a result, which is obviously in contradiction to the laws of probability because the conjunction of two events (Linda is a bank teller *and* an active feminist) can never be greater than either of the stand-alone events (e.g. Linda is a bank teller).

Eliciting Inter-Risk Dependence Measures

What are the implications of this research for estimating inter-risk correlation based on experts' opinion? First to mention is that judging correlations between two random variables contains the complexity of multivariate elicitation. Hence, one can expect to encounter similar or even bigger problems compared to those found when assessing means, medians or variances of univariate distributions. It has therefore been suggested that in order to elicit association (such as linear correlation) in a joint distribution, it is more reliable to elicit joint probabilities or conditional probabilities rather than the association measure directly. With this respect, O'Hagan et al. [56], p. 110, suggest a method for estimating the correlation between two random variables X and Y by means of their positive quadrant probability, which is defined as $P(X > x_{\text{med}}, Y > y_{\text{med}})$ where x_{med} and y_{med} are the medians of X and Y, respectively. For example, if X and Y are bivariate normally distributed with correlation ρ and standard deviations σ_x and σ_y, then their positive quadrant probability can be analytically calculated as (confer Lindskog [45], Example 2.2)

$$\begin{aligned} P(X > x_{\text{med}}, Y > y_{\text{med}}) &= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \int_0^\infty\int_0^\infty \exp\big[-\frac{1}{2(1-\rho^2)}\big(\frac{x^2}{\sigma_x^2} - \frac{2\rho xy}{\sigma_x\sigma_y} + \frac{y^2}{\sigma_y^2}\big)\big] \\ &= \frac{1}{4} + \frac{\arcsin(\rho)}{2\pi}\,. \end{aligned} \tag{4.8}$$

Solving (4.8) for ρ, an elicited positive quadrant probability can be transformed into a measure for the linear correlation coefficient.

Naturally, in the context of risk aggregation, two risk types are seldom bivariate normally distributed and (4.8) can not be used directly. However, the basic idea can be easily generalised to elicit the parameters of a given bivariate copula. It is worth

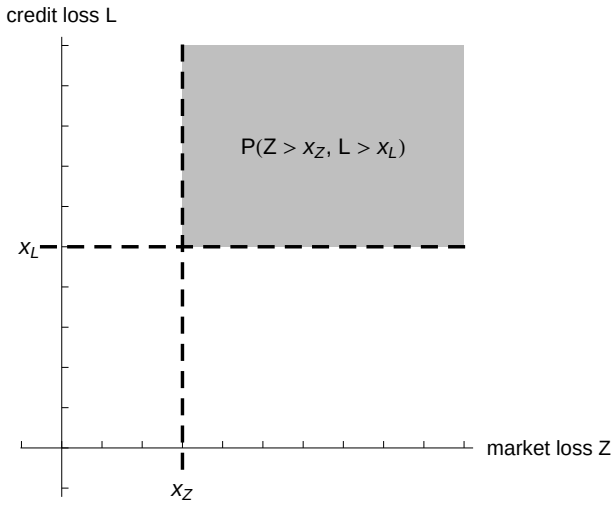

Figure 4.2.4. *Joint probability $P(Z > x_Z, L > x_L)$ that credit loss is above x_L and market loss is above x_Z. As described in the text, this probability can be used to estimate the correlation parameter R of a bivariate Gaussian or t copula.*

mentioning that in order to extract the parameters of a bivariate copula by estimating a joint probability such as (4.8), one needs to have already solved any assumptions about the marginal distribution functions as well as the copula. We now exemplify this technique for the Gaussian and t copula.

We now already adopt the notation we use in Section 4.3 and denote the credit risk and market risk random variables by L and Z, respectively. Let us consider the joint probability that market risk loss and credit risk loss exceed certain thresholds of x_L and x_Z, i.e. $P(Z > x_Z, L > x_L)$. From Figure 4.2.4 we immediately obtain

$$\begin{aligned} P(Z > x_Z, L > x_L) &= 1 - P(Z \leq x_Z) - P(L \leq x_L) + P(Z \leq x_Z, L \leq x_L) \\ &= 1 - F_Z(x_Z) - F_L(x_L) + C_R(F_Z(x_Z), F_L(x_L)) \,, \end{aligned} \tag{4.9}$$

where F_L and F_Z are the marginal distribution functions of pre-aggregated credit and market risk, and $C_R(\cdot, \cdot)$ is a bivariate Gaussian or t copula, depending on the unknown parameter R to be estimated. Expression (4.9), which can be understood as a generalisation of (4.8) to arbitrary marginals and bivariate copulas, can usually not be solved analytically for the copula parameter, here R. However, numerical techniques or a graphical analysis may help to infer R from an elicited value for $P(Z > x_Z, L > x_L)$.

It has been argued by several authors such as O'Hagan et al. [56] or Gokhale & Press [34] that people usually have more quantitative feeling about one-dimensional probabilities instead of (bivariate) joint probabilities. Hence, to diminish the bias of an expert estimate, they suggest to reduce a joint probability to a conditional probability,

which is a statement regarding only one random variable. In our example of market and credit risk, we therefore could do better in eliciting

$$P(L > x_L | Z > x_Z) = \frac{P(Z > x_Z, L > x_L)}{P(Z > x_Z)}\,, \tag{4.10}$$

i.e. the probability that credit losses will be above x_L given that a market risk loss greater than x_Z has already occurred. In order to illustrate this with a numerical example, let us first specify x_L and x_Z by the expectations $\mathbb{E}L$ and $\mathbb{E}Z$ of the marginal credit risk and market risk distributions. We use the same Student-t and Vasicek distribution for marginal market and credit risk as in Figure 4.2.2, and moreover we set without loss of generality $\mathbb{E}Z = 0$, whereas expected credit loss is roughly about 7.0 as described in Section 4.2.5. From (4.9) and (4.10) we then obtain

$$\begin{aligned} P(L > \mathbb{E}L | Z > 0) &= 2\,P(Z > 0, L > \mathbb{E}L) \\ &= 1 - 2\,F_L(\mathbb{E}L) + 2\,C_R\big(\frac{1}{2}, F_L(\mathbb{E}L)\big)\,. \end{aligned} \tag{4.11}$$

The left panel in Figure 4.2.5 shows the conditional probability $P(L > \mathbb{E}L | Z > 0)$ as a function of the copula parameter R both for a Gaussian and a t copula with 3 degrees of freedom. One can see that there is essentially no difference between both copulas, implying that extreme dependence between two risk types cannot be assessed by means of (4.11). This is mainly a consequence of the fact that we have chosen such moderate values for the thresholds x_L and x_Z, namely the expected losses of L and Z, respectively. If we set both x_L and x_Z in (4.9) to more extreme values, e.g. to high quantiles such as 99 %, one can graphically differentiate both cpoulas as it is shown in the right panel of Figure 4.2.5. The conditional probabilities derived from a Gaussian or a t copula are now different because, as we explained in Section 4.2.2, both copulas differ in the way they are reflecting extreme dependence. In particular, recall that $\lambda_{\text{Gauss}} = 0$, and therefore, in order to achieve a certain level of extreme joint loss probability $P(L > x_L | Z > x_Z)$, the correlation parameter of the Gaussian copula must be higher compared to the one of a t copula. Since in the right panel of Figure 4.2.5 the thresholds x_L and x_Z are chosen as high quantiles (99 %) of the marginal loss distributions, the plot effectively is a numerical approximation of the tail dependence (4.4) for different values of R. The exact tail dependence parameter for a t copula was given in (4.5) and is illustrated in Figure 4.2.5 as a dotted line, specifically, for $R = 0$ and $\nu = 3$ we have $\lambda_t = 0.12$.

Based on several expert estimates for different thresholds x_L and x_Z, one could in principle use the described method to determine both R and the degree-of-freedom parameter ν of a t copula. However, because such procedure would require expert

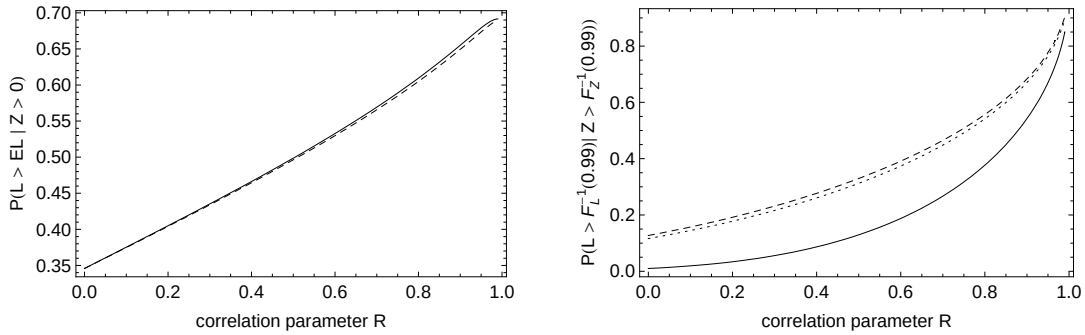

Figure 4.2.5. *Plot of the conditional probability $P(L > x_L|Z > x_Z)$ as a function of the parameter R of a Gaussian copula (solid line) and a t copula with 3 degrees of freedom (dashed line) according to* (4.10)*. In the case of the Gaussian copula, independence between market and credit risk is reached for $R = 0$, and the joint probability is then equal to the probability that credit losses are above x_L. Moreover, marginal market and credit risk follow the same Student-t and Vasicek distributions as used in Figure 4.2.2. Left panel: Here $x_L = \mathbb{E}L$ and $x_Z = \mathbb{E}Z = 0$. Right panel: $x_L = F_L^{-1}(0.99)$ and $x_Z = F_Z^{-1}(0.99)$. As explained in the text, this plot can therefore be interpreted as a numerical approximation of the tail dependence λ, which for the t copula is additionally shown as a dotted line.*

knowledge about high-quantile losses such as used in the right panel of Figure 4.2.5, one might expect a high degree of uncertainty when they are assessed by expert judgement.

Let us again consider the left panel of Figure 4.2.5. Statistically it is clear from (4.10) that the conditional probability $P(L > \mathbb{E}L|Z > 0)$ always takes values in a subinterval of $[0, 1]$, in particular, we have in the case of independent market and credit risk $P(L > \mathbb{E}L|Z > 0) = P(L > \mathbb{E}L) > 0$. Assuming a Gaussian copula, this is equivalent to $R = 0$. However, since humans cannot be guaranteed to act according to the rules of probability theory, we cannot take it for granted that this is also clear for the expert. The facilitator should therefore carefully explain the underlying statistical concept to the expert at the beginning of the elicitation exercise and maybe use the lower boundary $P(L > \mathbb{E}L)$ as an anchor that helps to calibrate the expert's estimate of $P(L > \mathbb{E}L|Z > 0)$. Finally, a possible question to the expert could be this one:

> Q: Suppose you were told that the actual yearly loss in market risk is above its expected value. Can you now determine the probability that also yearly credit loss is above the expected one?

The expert's estimate for $P(L > \mathbb{E}L|Z > 0)$ can then be transformed into an estimate for R either graphically by Figure 4.2.5 or numerically by solving (4.11). If one moreover is interested in the linear correlation coefficient ρ between both risk types, one can calculate it from simulated random variates obtained from the copula C_R together with the assumed marginal distributions, just as we did in Figure 4.2.2.

Needless to say, the described elicitation approach can be used to assess associations between almost every kind of risk-type marginals and, in principle, also in combination with other copulas. However, as we have already seen by comparing the Gaussian copula and the t copula in Figure 4.2.5, this method cannot straightforwardly be used to assess a possible extreme dependence between two risk types without touching the thorny issue of extreme value estimation based on expert opinions.

4.2.5 An Aggregation Exercise

After the inter-risk correlation matrix $\mathbf{C}$ or the matrix $\mathbf{R}$ of an elliptical copula is determined, one can begin with the aggregation of the different risk types. This Section is devoted to illustrate this aggregation procedure by considering a bank-typical example. We assume that pre-aggregated EC for each risk type has already been determined at confidence level of 99.95 % and time horizon of one year, and that the portfolio composition based on these pre-aggregated risk-type figures is given by 10 % market risk, 61 % credit risk, 14 % operational risk, and 15 % business risk. These values represent an industry average obtained from large banks that were participating the survey of the

Risk	Distribution function	Parameters
MR	$F(x) = F_\nu\left(\frac{x-\mu}{\sigma}\right), \quad x \in \mathbb{R}$	$\mu = 0,\ \sigma = 2.18, \nu = 10$
CR	$F(x) = \Phi\big[\frac{1}{\sqrt{\varrho}}\big(\sqrt{1-\varrho}\,\Phi^{-1}(\frac{x}{X}) - \Phi^{-1}(p)\big)\big], \quad x > 0$	$X = 2338.64,$ $p = 0.3\%, \varrho = 8\%$
OR	$F(x) = \Phi\left[\frac{\ln x-\mu}{\sigma}\right], \quad x > 0$	$\mu = -0.893, \sigma = 1.089$
BR	$F(x) = \Phi\left[\frac{x-\mu}{\sigma}\right], \quad x \in \mathbb{R}$	$\mu = 0,\ \sigma = 4.56$

Table 4.2.6. *Marginal distributions for market risk (MR), credit risk (CR), operational risk (OR), and business risk (BR), where F_ν is the Student-t distribution function with ν degrees of freedom and Φ is the standard normal distribution function. MR follows a scaled Student-t and CR is described by a Vasicek distribution with total exposure X, uniform asset correlation ϱ, and average default probability p. Operational risk (OR) is assumed to be lognormally distributed and business risk (BR) is modelled by a normal distribution. The parameters for the distribution functions are chosen so that MR, CR, OR, and BR absorb 10, 61, 14, and 15 units of EC at 99.95 % confidence level. Finally, only credit risk and operational risk have non-zero expected losses of about 7.0 and 0.7, respectively.*

IFRI/CRO Forum [36], Figure 16. For the different risk types we assume distributions and parameterisations as widely used in practice, see Table 4.2.6. In particular, the Vasicek distribution for credit risk will be further explained in Section 4.3.4. Also regarding the specification of the inter-risk correlation matrix we stick to the IFRI/CRO survey and use average benchmark results reported in [36], Figure 10. This is, however, not so straightforward since the linear correlation not only describes the dependence between different risk types but also is impacted by the marginal risk distributions. This is simply a consequence of the fact that linear correlation is (in contrast to copulas) not invariant under strictly increasing transformations. Hence, we simply interpret these benchmark correlations as the correlation matrix $\mathbf{R}$ of a Gaussian or t copula as described in Section 4.2.2, and thus independent of the marginals. We write

$$\mathbf{R} = \left(\begin{array}{c|cccc} & \text{market} & \text{credit} & \text{oprisk} & \text{business} \\ \hline \text{market} & 1 & 0.66 & 0.30 & 0.58 \\ \text{credit} & 0.66 & 1 & 0.30 & 0.67 \\ \text{oprisk} & 0.30 & 0.30 & 1 & 0.60 \\ \text{business} & 0.58 & 0.67 & 0.60 & 1 \end{array}\right). \tag{4.12}$$

Note that this matrix is positive semi-definite and thus a proper correlation matrix. Before we go on with our aggregation exercise, let us have a look at the corresponding correlation matrix $\mathbf{C}$ of linear correlation coefficients $\rho_{ij},\ i,j = 1,2,3,4$. As linear

	Aggregated Economic Capital	
Sum	Gaussian copula	t copula ($\nu = 5$)
100	79.57	85.95
	(82.39)	(83.10)

Table 4.2.7. *Aggregated EC at confidence level of 99.95 % and time horizon of one year for the portfolio consisting of market, credit, operational, and business risk as specified in Table 4.2.6. The correlation matrix* $\boldsymbol{R}$ *used for the Gaussian and t copula is given by* (4.12)*. The values in brackets are obtained by the variance-covariance formula* (4.1) *where we have used the simulated inter-risk correlations* (4.13) *and* (4.14)*, respectively.*

correlation depends on the dependence structure, we get different results for different copulas. Using the marginal distributions of Table 4.2.6 together with a Gaussian copula, a simulation yields

$$\mathbf{C}_{\text{Gauss}} = \left(\begin{array}{c|cccc} & \text{market} & \text{credit} & \text{oprisk} & \text{business} \\ \hline \text{market} & 1 & 0.57 & 0.30 & 0.42 \\ \text{credit} & 0.57 & 1 & 0.26 & 0.55 \\ \text{oprisk} & 0.30 & 0.26 & 1 & 0.43 \\ \text{business} & 0.42 & 0.55 & 0.43 & 1 \end{array} \right), \tag{4.13}$$

whereas in the case of a t copula with 5 degrees of freedom we obtain

$$\mathbf{C}_{\text{Student}} = \left(\begin{array}{c|cccc} & \text{market} & \text{credit} & \text{oprisk} & \text{business} \\ \hline \text{market} & 1 & 0.58 & 0.30 & 0.44 \\ \text{credit} & 0.58 & 1 & 0.26 & 0.61 \\ \text{oprisk} & 0.30 & 0.26 & 1 & 0.44 \\ \text{business} & 0.44 & 0.61 & 0.44 & 1 \end{array} \right). \tag{4.14}$$

The results of the aggregation analysis are shown in Table 4.2.7. One can see that total aggregated EC in the case of a Gaussian dependence structure is slightly below the result obtained from the variance covariance approach. However, because these results depend on the relative importance of the different risk types as well as the specific forms of the marginal distributions, one should not generalise this observation without investigation.

4.3 Correlation Between Credit and Market Risk

Here we combine a Merton-like factor model for credit risk with a linear factor model for another risk type—henceforth referred to as market risk—and investigate their correlation and the resulting aggregate risk. Both models are driven by a set of (macroeconomic) factors $Y = (Y_1, \ldots, Y_K)$ where the factor weights are allowed to be zero so that a risk type may only depend on a subset of Y. As a result, we obtain closed-form expressions for their inter-risk correlation.

4.3.1 Factor Models for Credit Risk

To describe credit portfolio loss, we choose a classical structural model as it can be found e.g. in Bluhm, Overbeck and Wagner [8]. Within these models, a borrower's credit quality (and so his default behaviour) is driven by its *asset-value process*, or, more generally and in the case of unlisted customers, by a so-called "ability-to-pay" process. Consider a portfolio of n loans. Then, default of an individual obligor i is described by a Bernoulli random variable $L_i, i = 1, \ldots, n$, with $\mathbb{P}(L_i = 1) = p_i$ and $\mathbb{P}(L_i = 0) = 1 - p_i$ where p_i is the obligor's *probability of default* within time period $[0, T]$ for $T > 0$. Following Merton's idea, counterparty i defaults if its asset value log-return A_i falls below some threshold D_i, sometimes referred to as *default point*, i.e.

$$L_i = \mathbb{1}_{\{A_i < D_i\}}\,, \qquad i = 1, \ldots, n\,.$$

If we denote the exposure at default net recovery of an individual obligor by e_i, portfolio loss is finally given by

$$L^{(n)} = \sum_{i=1}^{n} e_i\, L_i\,. \tag{4.15}$$

For a credit portfolio of n obligors, credit portfolio loss $L^{(n)}$ at time horizon T is driven by n realizations of the asset values A_i, which usually are assumed to depend on factors $(Y_1, \ldots, Y_K)$. The following factor model is widely spread in financial firms and similar versions are implemented in various software packages for credit risk.

Definition 4.3.1. [Normal factor model for credit risk] *Let $Y = (Y_1, \ldots, Y_K)$ be a K-dimensional random vector of (macroeconomic) factors with multivariate standard normal distribution. Then, in the normal factor model, each of the standardized asset value log-returns A_i, $i = 1, \ldots, n$, depends linearly on Y and a standard normally distributed idiosyncratic factor (or noise term) ε_i, independent of all Y_k, i.e.*

$$A_i = \sum_{k=1}^{K} \beta_{ik} Y_k + \sqrt{1 - \sum_{k=1}^{K} \beta_{ik}^2}\; \varepsilon_i\,, \qquad i = 1, \ldots, n\,, \tag{4.16}$$

with factor loadings β_{ik} *satisfying* $R_i^2 := \sum_{k=1}^K \beta_{ik}^2 \in [0,1]$ *describing the variance of* A_i *that can be explained by the systematic factors* Y*.*

For later usage we recall some properties of the normal factor model as can be found e.g. in Bluhm et al. [8] or McNeil, Frey and Embrechts [50], Chapter 8.

Remark 4.3.2. (a) The log-returns $A_1, \ldots, A_n$ are standard normally distributed and dependent with correlations

$$\rho_{ij} := \operatorname{corr}(A_i, A_j) = \sum_{k=1}^K \beta_{ik}\beta_{jk}\,, \qquad i,j = 1,\ldots,n\,, \tag{4.17}$$

the so-called *asset correlations* between A_i and A_j.
(b) The default point D_i of every obligor is related to its default probability p_i by

$$D_i = \Phi^{-1}(p_i)\,, \qquad i = 1,\ldots,n\,, \tag{4.18}$$

where Φ is the standard normal distribution function.
(c) The joint default probability of two obligors is given by

$$p_{ij} := \mathbb{P}(L_i = 1, L_j = 1) = \mathbb{P}(A_i \le D_i, A_j \le D_j) = \begin{cases} \Phi_{\rho_{ij}}(D_i, D_j)\,, & i \neq j\,, \\ p_i\,, & i = j\,, \end{cases} \tag{4.19}$$

where $\Phi_{\rho_{ij}}$ denotes the bivariate normal distribution function with standard marginals and correlation ρ_{ij} given by (4.17). Moreover, the *default correlation* between two different obligors is given by

$$\operatorname{corr}(L_i, L_j) = \frac{p_{ij} - p_i\,p_j}{\sqrt{p_i(1-p_i)\,p_j(1-p_j)}}\,, \qquad i,j = 1,\ldots,n\,. \tag{4.20}$$

(d) Conditional on a realization $y = (y_1, \ldots, y_K)$ of the factors Y, defaults of different obligors are independent. Moreover, the *conditional default probability* is given by

$$\begin{aligned} p_i(y) &= \mathbb{P}(L_i = 1 \,|\, Y = y) \\ &= \mathbb{P}\left(\sum_{k=1}^K \beta_{ik}\,y_k + \sqrt{1 - \sum_{k=1}^K \beta_{ik}^2}\;\varepsilon_i \le D_i\right) \\ &= \Phi\left(\frac{D_i - \sum_{k=1}^K \beta_{ik}\,y_k}{\sqrt{1 - \sum_{k=1}^K \beta_{ik}^2}}\right). \end{aligned}$$

□

A strong assumption of the model above is the multivariate normal distribution of the factor variables $Y = (Y_1, \ldots, Y_K)$, and thus of the asset value log-returns A_i. It is well known that the normal distribution has very light tails and therefore may seriously underestimate large fluctuations of the (macroeconomic) factors, eventually leading to model risk of the normal factor model for credit risk.

A generalisation allowing for heavier tails as well as a stronger dependence between different counterparties is the class of normal variance mixture distributions, where the covariance structure of the A_i is disturbed by means of a positive mixing variable W_L, see e.g. McNeil et al. [50], Section 3.2. A particular interesting model is the following one, confer also Kostadinov [44]:

Definition 4.3.3. [Shock model for credit risk] *Let $Y = (Y_1, \ldots, Y_K)$ be a K-dimensional random vector of (macroeconomic) factors with multivariate standard normal distribution. Then, in the shock model, each of the standardized asset value log-returns $\widehat{A}_i$, $i = 1, \ldots, n$, can be written as*

$$\widehat{A}_i = W_L \cdot \sum_{k=1}^{K} \beta_{ik} Y_k + W_L \cdot \sqrt{1 - \sum_{k=1}^{K} \beta_{ik}^2}\, \varepsilon_i \,, \qquad i = 1, \ldots, n \,, \tag{4.21}$$

where $W_L = \sqrt{\nu_L / S_{\nu_L}}$ and S_{ν_L} is a $\chi^2_{\nu_L}$ distributed random variable with ν_L degrees of freedom, independent of Y and the idiosyncratic factor ε_i.

The mixing variable W_L can be interpreted as a "global shock" driving the variance of all factors. Such an overarching shock may occur from political distress, severe economic recession or some natural disaster.

We conclude this Section with some general remarks about the shock model for credit risk (see again Bluhm et al. [8] and McNeil et al. [50] as well as references therein).

Remark 4.3.4. (a) In general, let $X = (X_1, \ldots, X_n)$ be a standardised multinormal vector with covariance matrix R and S_{ν_L} is a chi-square variable with ν_L degrees of freedom. Then $(X_1, \ldots, X_n)/\sqrt{S_{\nu_L}/\nu_L}$ has a multivariate t-distribution with correlation matrix R and ν_L degrees of freedom. Hence, from (4.16) and (4.21) it follows for the shock model for credit risk that the vector of standardised asset value log-returns $(\widehat{A}_1, \ldots, \widehat{A}_n)$ is t-distributed with ν_L degrees of freedom, in particular, it has the same asset correlation ρ_{ij} as the normal factor model given by equation (4.17).
(b) The default point $\widehat{D}_i$ of the shock model is linked to the obligor's default probability by

$$\widehat{D}_i = t_{\nu_L}^{-1}(p_i) \,, \qquad i = 1, \ldots, n \,, \tag{4.22}$$

where t_{ν_L} is the Student t distribution function with ν_L degrees of freedom.
(c) The joint default probability $\widehat{p}_{ij}$ for two obligors can be written as

$$\widehat{p}_{ij} = t_{\nu_L;\rho_{ij}}(\widehat{D}_i, \widehat{D}_j)\,, \qquad i \neq j\,, \tag{4.23}$$

where $t_{\nu_L;\rho_{ij}}$ denotes the standard bivariate Student t distribution function with correlation ρ_{ij} given by (4.17) and degree of freedom parameter ν_L. □

4.3.2 Joint Factor Models for Credit and Market Risk

Market risk is related to a bank's potential loss associated with its trading activities. We assume that it is already pre-aggregated so that losses can be approximately described by a one-dimensional random variable Z (or $\widehat{Z}$, see below), which can be thought of as the bank-wide, aggregated profit and loss (P/L) distribution due to changes in some market variables, such as interest rates, foreign exchange rates, equity prices or the value of commodities.

Similarly as for credit risk, we explain fluctuations of the random variable Z by means of (macroeconomic) factors $(Y_1, \ldots, Y_K)$. We use the same macroeconomic factors for credit and market risk, where independence of risk from such a factor is indicated by a loading factor 0. If the pre-aggregated P/L can be described by a normal distribution, the following factor model is a sensible choice and can be used for risk aggregation. Even if this assumption does not hold exactly, it can be used as an important approximation for investigating inter-risk dependencies (we use the convention that losses correspond to positive values of Z.)

Definition 4.3.5. [Normal factor model for market risk] *Let $Y = (Y_1, \ldots, Y_K)$ be a random vector of (macroeconomic) factors with multivariate standard normal distribution. Then, the normal factor model for the pre-aggregated market risk P/L is given by*

$$Z = -\sigma \left(\sum_{k=1}^{K} \gamma_k Y_k + \sqrt{1 - \sum_{k=1}^{K} \gamma_k^2}\, \eta \right), \tag{4.24}$$

with factor loadings γ_k satisfying $\sum_{k=1}^{K} \gamma_k^2 \in [0,1]$, which is that part of the variance of Z which can be explained by the systematic factor Y. Furthermore, η is a standard normally distributed idiosyncratic factor, independent of Y, and σ is the standard deviation of Z.

Clearly, for an actively managed market portfolio the idiosyncratic factor η is more important than for an unmanaged portfolio (e.g. an index of stocks). As a matter

of fact, portfolio managers are paid owing to their skills to achieve the best possible portfolio performance that is independent of some macroeconomic indicators.

Note that both in Definition 4.3.1 of the credit factor model as well as above, the factor loadings β_{ik} and γ_k, respectively, are allowed to be zero. For instance, Y_k can be relevant for credit but not for market by setting $\gamma_k = 0$.

Definition 4.3.6. [Joint normal factor model for credit and market risk] *Let $Y = (Y_1, \ldots, Y_K)$ be a random vector of (macroeconomic) factors with multivariate standard normal distribution. Let the credit portfolio loss $L^{(n)}$ be given by* (4.15)*, and the asset value log-returns A_i for $i = 1, \ldots, n$ be modelled by the normal factor model* (4.16)*. Let Z be the pre-aggregated market risk P/L modelled by the normal factor model* (4.24)*. When the idiosyncratic factors ε_i for $i = 1, \ldots, n$ of the credit model are independent of η, then we call $(L^{(n)}, Z)$ the* joint normal factor model for credit and market risk*.*

In order to account for possible heavy tails in the market risk P/L, we again rely on the global shock approach already used for credit risk.

Definition 4.3.7. [Shock model for market risk] *Let $Y = (Y_1, \ldots, Y_K)$ be a random vector of (macroeconomic) factors with multivariate standard normal distribution. Then the shock model for the pre-aggregated market risk P/L is given by*

$$\widehat{Z} = -\sigma \left(W_Z \cdot \sum_{k=1}^{K} \gamma_k Y_k + W_Z \cdot \sqrt{1 - \sum_{k=1}^{K} \gamma_k^2}\, \eta \right), \tag{4.25}$$

where σ is a scaling factor, $W_Z = \sqrt{\nu_Z / S_{\nu_Z}}$, and S_{ν_Z} is a $\chi^2_{\nu_Z}$ distributed random variable with ν_Z degrees of freedom, independent of Y and the idiosyncratic factor η.

Definition 4.3.8. [Joint shock model for credit and market risk] *Let $Y = (Y_1, \ldots, Y_K)$ be a random vector of (macroeconomic) factors with multivariate standard normal distribution. Let the credit portfolio loss be given by* (4.15)*, now denoted as $\widehat{L}^{(n)}$, and the asset value log-returns $\widehat{A}_i$ for $i = 1, \ldots, n$ be modelled by the shock model* (4.21) *with shock variable W_L. Let $\widehat{Z}$ be the pre-aggregated market risk P/L modelled by the shock model* (4.25) *with shock variable W_Z.*

(1) (Independent shock model for credit and market risk). *If the credit model's idiosyncratic factors ε_i for $i = 1, \ldots, n$ are independent of η, and furthermore W_L is independent from W_Z, then we call $(\widehat{L}^{(n)}, \widehat{Z})$ the* independent shock model for credit and market risk*.*

(2) (Common shock model for credit and market risk). *If the credit model's idiosyncratic factors ε_i for $i = 1, \ldots, n$ are independent of η, and furthermore if we set*

$$W_L \equiv W_Z =: W\,,$$

then we call $(\widehat{L}^{(n)}, \widehat{Z})$ the common shock model for credit and market risk*.*

4.3.3 Inter-Risk Correlation

We now calculate the linear correlation between the credit risk portfolio loss and market risk. We begin with the normal model.

Normal Factor Model Approach

The proposed models shall now be used to investigate the dependence between credit risk $L^{(n)}$ and market risk Z, introduced by the factors Y. Let us start with the linear correlation, which is defined as

$$\operatorname{corr}(L^{(n)}, Z) = \frac{\operatorname{cov}(L^{(n)}, Z)}{\sqrt{\operatorname{var}(L^{(n)})}\sqrt{\operatorname{var}(Z)}}\,. \tag{4.26}$$

Although linear correlation only describes linear dependence between different random variables, it is a very popular and important concept in finance, frequently used both by practitioners and academics. Moreover, since we calculate expressions for linear correlation in closed form, we are able to analytically investigate the linear dependence structure between market and credit risk. Note also that the correlation completely describes the dependence in the joint normal factor model.

We begin with the joint normal factor model for credit and market risk.

Theorem 4.3.9. [Inter-risk correlation for the normal factor model] *Suppose that credit portfolio loss $L^{(n)}$ and market risk Z are described by the joint normal factor models of Definition 4.3.6. Then, linear correlation between $L^{(n)}$ and Z is given by*

$$\operatorname{corr}(L^{(n)}, Z) = \frac{\sum_{i=1}^{n} r_i\, e_i \exp\left(-\frac{1}{2}D_i^2\right)}{\sqrt{2\pi\,\operatorname{var}(L^{(n)})}}\,, \tag{4.27}$$

where D_i is the default point (4.18)*,*

$$r_i := \operatorname{corr}(A_i, Z) = \sum_{k=1}^{K} \beta_{ik}\gamma_k\,, \qquad i = 1, \ldots, n,$$

and

$$\operatorname{var}(L^{(n)}) = \sum_{i,j=1}^{n} e_i\, e_j\, (p_{ij} - p_i\, p_j)$$

with joint default probability p_{ij} given by (4.19)*.*

Proof. First we calculate the covariance between $L^{(n)}$ and Z. Using $\mathbb{E}(Z) = 0$, expression (4.15), and the fact that η in (4.24) is independent from Y (and thus from L_i), we can write

$$\operatorname{cov}(L^{(n)}, Z) = \mathbb{E}(ZL^{(n)}) = -\sigma \sum_{i=1}^{n} e_i \sum_{k=1}^{K} \gamma_k\, \mathbb{E}(Y_k L_i)\,. \tag{4.28}$$

To evaluate the expectation, conditioning with respect to $Y_k = y_k$ and using the law of iterated expectation yield

$$\begin{aligned}
\mathbb{E}(Y_k\, L_i) &= \mathbb{E}(Y_k\, L_i(Y_1,\ldots,Y_K)) \\
&= \mathbb{E}\big(\mathbb{E}(Y_k\, L_i(Y_1,\ldots,Y_K)\,|\,Y_k)\big) \\
&= \int_{-\infty}^{\infty} \mathbb{E}(Y_k\, L_i(Y_1,\ldots,Y_K)\,|\,Y_k = y_k)\; d\Phi(y_k) \\
&= \int_{-\infty}^{\infty} \mathbb{E}(y_k\, L_i(Y_1,\ldots,y_k,\ldots,Y_K))\; d\Phi(y_k) \\
&= \int_{-\infty}^{\infty} y_k\, \mathbb{E}(L_i(Y_1,\ldots,y_k,\ldots,Y_K))\; d\Phi(y_k)
\end{aligned}$$

where Φ is the standard normal distribution function. Using $\mathbb{E}(L_i) = \mathbb{P}(L_i = 1)$, we have

$$\begin{aligned}
\mathbb{E}(Y_k\, L_i) &= \int_{-\infty}^{\infty} y_k\, \mathbb{P}(L_i(Y_1,\ldots,y_k,\ldots,Y_K) = 1)\; d\Phi(y_k) \\
&= \int_{-\infty}^{\infty} y_k\, \mathbb{P}\left(\sum_{\substack{l=1\\l\neq k}}^{K} \beta_{il}\, Y_l + \beta_{ik}\, y_k + \sqrt{1-\sum_{j=1}^{K}\beta_{ij}^2}\;\varepsilon_i \le D_i\right) d\Phi(y_k) \\
&= \int_{-\infty}^{\infty} y_k\, \mathbb{P}\left(\sum_{\substack{l=1\\l\neq k}}^{K} \beta_{il}\, Y_l + \sqrt{1-\sum_{j=1}^{K}\beta_{ij}^2}\;\varepsilon_i \le D_i - \beta_{ik}\, y_k\right) d\Phi(y_k) \\
&=: \int_{-\infty}^{\infty} y_k\, \mathbb{P}\left(X \le D_i - \beta_{ik}\, y_k\right) d\Phi(y_k)
\end{aligned}$$

where X is normally distributed with variance $\mathrm{var}(X) = 1 - \beta_{ik}^2$. Hence, we obtain

$$\mathbb{E}(Y_k\, L_i) = \int_{-\infty}^{\infty} y_k\, \Phi\left(\frac{D_i - \beta_{ik}\, y_k}{\sqrt{1-\beta_{ik}^2}}\right) d\Phi(y_k)\,.$$

Since the derivative of the density φ of the standard normal distribution is given by $\varphi'(y) = y\,\varphi(y)$, it follows by partial integration that

$$\mathbb{E}(Y_k\, L_i) = -\frac{\beta_{ik}}{\sqrt{1-\beta_{ik}^2}} \int_{-\infty}^{\infty} y_k\, \varphi\left(\frac{D_i - \beta_{ik}\, y_k}{\sqrt{1-\beta_{ik}^2}}\right) \varphi(y_k)\, d\varphi(y_k)\,,$$

where the right-hand side is just $-\beta_{ik}$ times the density of a random variable $\Psi = \sqrt{1-\beta_{ik}^2}X + \beta_{ik}Y_k$ for standard normal iid X, Y_k at point D_i. Since Ψ is then again standard normal, we obtain

$$\mathbb{E}(Y_k\, L_i) = -\beta_{ik}\varphi(D_i) = -\frac{\beta_{ik}}{\sqrt{2\pi}}\, e^{-\frac{D_i^2}{2}}\,, \tag{4.29}$$

which together with (4.28) yields

$$\begin{aligned}\operatorname{cov}(L^{(n)}, Z) &= \frac{\sigma}{\sqrt{2\pi}} \sum_{i=1}^{n} \sum_{k=1}^{K} e_i \, \gamma_k \, \beta_{ik} \, e^{-\frac{D_i^2}{2}} \\ &= \frac{\sigma}{\sqrt{2\pi}} \sum_{i=1}^{n} e_i \, r_i \, e^{-\frac{D_i^2}{2}}\end{aligned}$$

where we have introduced $r_i := \operatorname{corr}(A_i, Z) = \sum_{k=1}^{K} \beta_{ik}\gamma_k$. With $\sqrt{\operatorname{var}(Z)} = \sigma$ and

$$\begin{aligned}\operatorname{var}(L^{(n)}) &= \sum_{i,j=1}^{n} e_i \, e_j \operatorname{cov}(L_i, L_j) \\ &= \sum_{i,j=1}^{n} e_i \, e_j \left(\mathbb{E}(L_i L_j) - \mathbb{E}(L_i)\mathbb{E}(L_j) \right) \\ &= \sum_{i,j=1}^{n} e_i \, e_j \left(p_{ij} - p_i \, p_j \right), \end{aligned} \tag{4.30}$$

where p_{ij} is the joint default probability (4.19), the assertion follows. □

Note that r_i may become negative if (some) factor loadings β_{ik} and γ_k have different signs. Therefore, in principle, also negative inter-risk correlations can be achieved in our model. Moreover, in (4.27) the term $e_i \, e^{-D_i^2/2}$ can be interpreted as a kind of rating-adjusted exposure. For instance, a relatively low default probability of debtor i corresponds to a relatively small value of $e^{-D_i^2/2}$. As a consequence thereof, for two obligors with equal exposure size e_i, the one with the better rating has less impact on inter-risk correlation as the low-rated creditor.

The fact that $\operatorname{corr}(L^{(n)}, Z)$ linearly depends on the correlation r_i and thus on the factor loadings γ_k, implies the following Proposition.

Proposition 4.3.10. [Inter-risk correlation bound for the joint normal factor model] *Suppose that credit portfolio loss $L^{(n)}$ is described by the normal factor model of Definition 4.3.1, and market risk Z by the normal factor model of Definition 4.3.5, however, with unknown factor loadings γ_k, $k = 1, \ldots, K$. Then, inter-risk correlation is bounded by*

$$|\operatorname{corr}(L^{(n)}, Z)| \leq \frac{\sum_{i=1}^{n} e_i \, R_i \exp\left(-\frac{1}{2} D_i^2\right)}{\sqrt{2\pi \operatorname{var}(L^{(n)})}} \tag{4.31}$$

with $R_i = \sqrt{\sum_{k=1}^{K} \beta_{ik}^2}$.

Proof. Since the obligor's exposures are assumed to be positive, $e_i \geq 0$, it follows from (4.27) that

$$|\mathrm{corr}(L^{(n)}, Z)| \leq \frac{\sum_i e_i \, |r_i| \exp\left(-\frac{1}{2} D_i^2\right)}{\sqrt{2\pi \sum_{ij} e_i \, e_j \, (p_{ij} - p_i \, p_j)}} \, .$$

From the Cauchy-Schwarz inequality together with $\sum_{k=1}^{K} \gamma_k^2 \leq 1$, it follows that

$$|r_i| = \left| \sum_{k=1}^{K} \beta_{ik} \gamma_k \right| \leq \left(\sum_{k=1}^{K} \beta_{ik}^2 \right)^{1/2} \left(\sum_{k=1}^{K} \gamma_k^2 \right)^{1/2} \leq \left(\sum_{k=1}^{K} \beta_{ik}^2 \right)^{1/2} \leq 1 \, ,$$

which completes the proof. □

Note that (4.31) does not depend on any specific market risk parameter. Therefore, solely based on the parametrisation of the normal credit factor model, a bound for inter-risk correlation can be derived. This bound then holds for all market risk portfolios described by Definition 4.3.5. Furthermore, as R_i^2 is that part of the variance of A_i which can be explained by the factor Y, it follows from (4.31) that the inter-risk correlation bound is affine linearly increasing with R_i. This is also intuitively clear because with increasing R_i^2, credit portfolio loss is more and more dominated by the systematic factor Y, which by construction drives the inter-risk dependence with market risk.

Shock Model Approach

We now investigate how the existence of global shocks affects inter-risk correlation. We consider both kinds of joint shock models for credit and market risk given by Definition 4.3.8 and calculate inter-risk correlation similarly to Theorem 4.3.9.

Theorem 4.3.11. [Inter-risk correlation for the joint shock model] *Suppose that credit portfolio loss $\widehat{L}^{(n)}$ and market risk $\widehat{Z}$ are described by the joint shock factor model of Definition 4.3.8.*
(1) (Independent shock model, Definition 4.3.8 (1)). *If shocks in credit and market risk are driven by different independent shock variables W_L and W_Z with degrees of freedom $\nu_L > 0$ and $\nu_Z > 2$, respectively, linear correlation between $\widehat{L}^{(n)}$ and $\widehat{Z}$ is given by*

$$\mathrm{corr}(\widehat{L}^{(n)}, \widehat{Z}) = \sqrt{\frac{\nu_Z - 2}{2}} \, \frac{\Gamma\left(\frac{\nu_Z - 1}{2}\right)}{\Gamma\left(\frac{\nu_Z}{2}\right)} \, \frac{\sum_{i=1}^{n} e_i \, r_i \left(1 + \frac{\widehat{D}_i^2}{\nu_L}\right)^{-\frac{\nu_L}{2}}}{\sqrt{2 \, \pi \, \mathrm{var}(\widehat{L}^{(n)})}} \, . \tag{4.32}$$

(2) (Common shock model, Definition 4.3.8 (2)). *If shocks in credit and market risk are*

driven by the same shock variable W with $\nu > 1$ degrees of freedom, linear correlation between $\widehat{L}^{(n)}$ and $\widehat{Z}$ is given by

$$\operatorname{corr}(\widehat{L}^{(n)}, \widehat{Z}) = \sqrt{\frac{\nu-2}{2}}\,\frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\,\frac{\sum_{i=1}^n e_i\, r_i \left(1+\frac{\widehat{D}_i^2}{\nu}\right)^{\frac{1-\nu}{2}}}{\sqrt{2\,\pi\,\operatorname{var}(\widehat{L}^{(n)})}}\,. \tag{4.33}$$

In both cases,

$$r_i := \operatorname{corr}(\widehat{A}_i, \widehat{Z}) = \sum_{k=1}^K \beta_{ik}\gamma_k\,, \qquad i = 1, \ldots, n,$$

and

$$\operatorname{var}(\widehat{L}^{(n)}) = \sum_{i,j=1}^n e_i\, e_j\,(p_{ij} - p_i\, p_j)\,.$$

Furthermore, $\widehat{D}_i$ and $\widehat{p}_{ij}$ are given by (4.22) *and* (4.23)*, respectively, with degree of freedom ν_L for the independent shock model (1) and ν for the common shock model (2).*

Proof. (1) Using (4.15) and the law of iterated expectation, we obtain

$$\operatorname{cov}(\widehat{L}^{(n)}, \widehat{Z}) = \sum_{i=1}^n e_i\,\mathbb{E}(\widehat{Z}\,\widehat{L}_i) = \sum_{i=1}^n e_i\,\mathbb{E}\big(\mathbb{E}(\widehat{Z}\,\widehat{L}_i \,|\, W_L, W_Z)\big). \tag{4.34}$$

Now, in the credit shock model of Definition 4.3.3 we have that for each loss variable $\widehat{L}_i$

$$\begin{aligned}
\mathbb{P}(\widehat{L}_i = 1) &= \mathbb{P}\left(W_L \sum_{k=1}^K \beta_{ik}\, Y_k + W_L \sqrt{1 - \sum_{j=1}^K \beta_{ij}^2}\;\varepsilon_i \le \widehat{D}_i\right) \\
&= \mathbb{P}\left(\sum_{k=1}^K \beta_{ik}\, Y_k + \sqrt{1 - \sum_{j=1}^K \beta_{ij}^2}\;\varepsilon_i \le \frac{\widehat{D}_i}{W_L}\right).
\end{aligned}$$

Hence, the shock factor model conditional on the shock variable W_L is equivalent to a normal factor model with adjusted default points $D_i^* := \widehat{D}_i / W_L$. Therefore, we obtain from (4.29) without any further calculation

$$\begin{aligned}
\mathbb{E}(\widehat{Z}\,\widehat{L}_i \,|\, W_L, W_Z) &= -\sigma \sum_{k=1}^K \gamma_k\, W_Z\, \mathbb{E}(Y_k\,\widehat{L}_i \,|\, W_L) \\
&= -\sigma \sum_{k=1}^K \gamma_k\, W_Z \left(-\frac{\beta_{ik}}{\sqrt{2\pi}}\, e^{-\frac{D_i^{*2}}{2}}\right) \\
&= \frac{\sigma\, r_i}{\sqrt{2\pi}}\, W_Z\, e^{-\frac{D_i^{*2}}{2}},
\end{aligned}$$

where $r_i = \sum_{k=1}^{K} \beta_{ik}\gamma_k$ for $i = 1, \ldots, n$. Integration over W_L and W_Z yields

$$\begin{aligned} \mathbb{E}\big(\mathbb{E}(\widehat{Z}\,\widehat{L}_i \,|\, W_L, W_Z)\big) &= \frac{\sigma\, r_i}{\sqrt{2\pi}} \int_0^\infty W_Z \, dF_{\nu_Z}(s) \int_0^\infty e^{-\frac{D_i^{*2}}{2}} \, dF_{\nu_L}(s) \qquad (4.35) \\ &= \frac{\sigma\, r_i}{\sqrt{2\pi}}\, \mathbb{E}(W_Z) \int_0^\infty e^{-\frac{\widehat{D}_i^2}{2\nu_L} s}\, f_{\nu_L}(s)\, ds \end{aligned}$$

where F_ν is the distribution function of a χ^2_ν distributed random variable with density $f_\nu(s)$. By substitution, we can perform the integration for $\nu_L > 0$,

$$\begin{aligned} \int_0^\infty e^{-\frac{\widehat{D}_i^2}{2\nu_L} s}\, f_{\nu_L}(s)\, ds &= \int_0^\infty \frac{2^{-\nu/2}\, s^{\nu/2-1}}{\Gamma\big(\frac{\nu}{2}\big)} \exp\left[-\left(1 + \frac{\widehat{D}_i^2}{\nu_L}\right)\frac{s}{2}\right] ds \\ &= \left(1 + \frac{\widehat{D}_i^2}{\nu_L}\right)^{-\frac{\nu_L}{2}} \int_0^\infty \frac{2^{-\nu/2}\, e^{-s/2}\, s^{\nu/2-1}}{\Gamma\big(\frac{\nu}{2}\big)}\, ds \\ &= \left(1 + \frac{\widehat{D}_i^2}{\nu_L}\right)^{-\frac{\nu_L}{2}} . \end{aligned}$$

Together with

$$\mathbb{E}(W_Z) = \sqrt{\frac{\nu_Z}{2}}\, \frac{\Gamma\left(\frac{\nu_Z - 1}{2}\right)}{\Gamma\big(\frac{\nu_Z}{2}\big)}$$

for $\nu_Z > 1$ we obtain

$$\mathbb{E}\big(\mathbb{E}(\widehat{Z}\,\widehat{L}_i \,|\, W_L, W_Z)\big) = \frac{\sigma\, r_i}{\sqrt{2\pi}} \sqrt{\frac{\nu_Z}{2}} \left(1 + \frac{\widehat{D}_i^2}{\nu_L}\right)^{\frac{-\nu_L}{2}} \frac{\Gamma\left(\frac{\nu_Z - 1}{2}\right)}{\Gamma\big(\frac{\nu_Z}{2}\big)}, \qquad (4.36)$$

Now, plugging (4.36) into (4.34), and using (4.30) together with

$$\operatorname{var}(\widehat{Z}) = \left(\frac{\nu_Z}{\nu_Z - 2}\right)\sigma^2, \quad \nu_Z > 2,$$

finally leads to

$$\operatorname{corr}(\widehat{L}^{(n)}, \widehat{Z}) = \sqrt{\frac{\nu_Z - 2}{2}}\, \frac{\Gamma\left(\frac{\nu_Z - 1}{2}\right)}{\Gamma\big(\frac{\nu_Z}{2}\big)}\, \frac{\sum_{i=1}^n e_i\, r_i \left(1 + \frac{\widehat{D}_i^2}{\nu_L}\right)^{\frac{-\nu_L}{2}}}{\sqrt{2\,\pi\, \operatorname{var}(\widehat{L}^{(n)})}} .$$

(2) Similarly to the case of independent shock variables, we are now conditioning on

the single shock variable W. Instead of (4.35), we obtain for $\nu > 1$ by substitution

$$\begin{aligned}
\mathbb{E}\big(\mathbb{E}(\widehat{Z}\,\widehat{L}_i \,|\, W)\big) &= \frac{\sigma\, r_i}{\sqrt{2\pi}} \int_0^\infty W\, e^{-\frac{D_i^{*2}}{2}}\, dF_\nu(s) \\
&= \frac{\sigma\, r_i}{\sqrt{2\pi}} \int_0^\infty \sqrt{\frac{\nu}{s}}\, e^{-\frac{\widehat{D}_i^2}{2\nu} s}\, f_\nu(s)\, ds \\
&= \frac{\sigma\, r_i}{\sqrt{2\pi}} \int_0^\infty \frac{2^{-\nu/2}\, \nu^{1/2}}{\Gamma(\nu/2)} s^{\frac{\nu-1}{2}-1} \exp\left[-\left(\frac{\widehat{D}_i^2}{\nu}+1\right)\frac{s}{2}\right] ds \\
&= \frac{\sigma\, r_i}{\sqrt{2\pi}} \int_0^\infty \frac{2^{-\nu/2}\, \nu^{1/2}}{\Gamma(\nu/2)} \left(\frac{\widehat{D}_i^2}{\nu}+1\right)^{-\frac{\nu-1}{2}} s^{\frac{\nu-1}{2}-1} e^{-s/2} ds \\
&= \frac{\sigma\, r_i}{\sqrt{2\pi}} \frac{\nu^{1/2}\, 2^{-1/2}}{\Gamma(\nu/2)} \left(\frac{\widehat{D}_i^2}{\nu}+1\right)^{\frac{1-\nu}{2}} \Gamma\left(\frac{\nu-1}{2}\right) \int_0^\infty \frac{2^{-\frac{\nu-1}{2}}}{\Gamma\left(\frac{\nu-1}{2}\right)} s^{\frac{\nu-1}{2}-1} e^{-s/2} ds \\
&= \frac{\sigma\, r_i}{\sqrt{2\pi}} \frac{\nu^{\frac{\nu}{2}} \left(\widehat{D}_i^2+\nu\right)^{\frac{1-\nu}{2}} \Gamma\left(\frac{\nu-1}{2}\right)}{\sqrt{2}\,\Gamma\left(\frac{\nu}{2}\right)}, \qquad (4.37)
\end{aligned}$$

which finally completes the proof. □

Analogously to the normal factor model, inter-correlation bounds can be derived.

Proposition 4.3.12. [Inter-risk correlation bounds for the joint shock model] *Suppose that credit portfolio loss $\widehat{L}^{(n)}$ is described by the shock model of Definition 4.3.3 and market risk $\widehat{Z}$ by the shock model of Definition 4.3.7, however, with unknown factor loadings $\gamma_k,\ k = 1, \ldots, K$.*

(1) *For the independent shock model, inter-risk correlation is bounded by*

$$|\mathrm{corr}(\widehat{L}^{(n)}, \widehat{Z})| \;\le\; \sqrt{\frac{\nu_Z - 2}{2}}\, \frac{\Gamma\left(\frac{\nu_Z-1}{2}\right)}{\Gamma\left(\frac{\nu_Z}{2}\right)}\, \frac{\sum_{i=1}^n e_i\, R_i \left(1+\frac{\widehat{D}_i^2}{\nu_L}\right)^{-\frac{\nu_L}{2}}}{\sqrt{2\,\pi\,\mathrm{var}(\widehat{L}^{(n)})}}. \qquad (4.38)$$

(2) *For the common shock model, inter-risk correlation is bounded by*

$$|\mathrm{corr}(\widehat{L}^{(n)}, \widehat{Z})| \;\le\; \sqrt{\frac{\nu - 2}{2}}\, \frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\, \frac{\sum_{i=1}^n e_i\, R_i \left(1+\frac{\widehat{D}_i^2}{\nu}\right)^{\frac{1-\nu}{2}}}{\sqrt{2\,\pi\,\mathrm{var}(\widehat{L}^{(n)})}}. \qquad (4.39)$$

In both cases (4.38) *and* (4.39) *is* $R_i = \sqrt{\sum_{k=1}^K \beta_{ik}^2}$.

Proof. This is analogously to the proof of Proposition 4.3.10 □

.

For practical purposes very relevant is the situation where credit risk quantification is based on a normal factor model, whereas heavy tails are assumed for the market risk, which therefore shall be described by the shock model approach. This can be referred to as a *hybrid factor model*, which is a special case of the joint shock model of Definition 4.3.3 with $\nu_L \to \infty$. We formulate our results as a Corollary.

Corollary 4.3.13. [Inter-risk correlation for the hybrid factor model] *Suppose that credit portfolio loss $L^{(n)}$ is described by the normal factor model of Definition 4.3.1, and market risk $\widehat{Z}$ by the shock model of Definition 4.3.7. Assume that the credit model's idiosyncratic factors ε_i for $i = 1, \ldots, n$ are independent of η, then we call $(L^{(n)}, \widehat{Z})$ the* hybrid factor model.
(1) *Inter-risk correlation is given by*

$$\operatorname{corr}(L, \widehat{Z}) = \sqrt{\frac{\nu_Z - 2}{2}} \frac{\Gamma\left(\frac{\nu_Z - 1}{2}\right)}{\Gamma\left(\frac{\nu_Z}{2}\right)} \operatorname{corr}(L, Z)$$

with $\operatorname{corr}(L, Z)$ *as in* (4.27).
(2) *If the factor loadings γ_k, $k = 1, \ldots, K$ of market risk are unknown, the inter-risk correlation bound is given by*

$$|\operatorname{corr}(L, \widehat{Z})| \leq \sqrt{\frac{\nu_Z - 2}{2}} \frac{\Gamma\left(\frac{\nu_Z - 1}{2}\right)}{\Gamma\left(\frac{\nu_Z}{2}\right)} |\operatorname{corr}(L, Z)|$$

with $|\operatorname{corr}(L, Z)|$ *as in* (4.31).

4.3.4 An Application to One-Factor Models

The results so far can be used to estimate inter-risk correlation between pre-aggregated market risk and credit risk. For the purpose of illustration and to gain deeper insights into the interaction between $L^{(n)}$ and Z, it is helpful to consider the one-factor versions of (4.16) and (4.25) as well as for the corresponding inter-risk correlation.

Joint One-Factor Models for Credit And Market Risk

Instructive examples regarding inter-risk correlation and its bounds can be obtained for one-factor models and they are useful to explain general characteristics and systematic behaviour of inter-risk correlation. In the context of credit risk, one-factor models can quite naturally be obtained by considering the special case of a large homogenous portfolio (LHP).

Let us start with a homogenous portfolio for which we define that $e_i = e$, $p_i = p$, $\beta_{ik} = \beta_k$ for $i = 1, \ldots, n$, and $k = 1, \ldots, K$. Then, by setting $\rho := \sum_{k=1}^{K} \beta_k^2$ and

$$\widetilde{Y} := \left(\sum_{k=1}^{K} \beta_k Y_k \right) / \sqrt{\rho}\,, \tag{4.40}$$

expression (4.16) for the general factor model can be transformed into a one-factor model

$$A_i = \sqrt{\rho}\,\widetilde{Y} + \sqrt{1-\rho}\,\varepsilon_i\,, \tag{4.41}$$

where $\widetilde{Y}$ is standard normally distributed and independent of ε_i, and ρ is the *uniform asset correlation* within the credit portfolio. If we now additionally increase the number of counterparties in the portfolio by $n \to \infty$, then the relative portfolio loss satisfies[2]

$$\frac{L^{(n)}}{n\,e} \overset{\text{a.s.}}{\to} \Phi\left(\frac{D - \sqrt{\rho}\,\widetilde{Y}}{\sqrt{1-\rho}} \right) =: L\,, \qquad n \to \infty\,, \tag{4.42}$$

where $D = \Phi^{-1}(p)$ and $n\,e$ is the total exposure of the credit portfolio. Often L is used as an approximative loss variable for large and almost homogeneous portfolios. For later usage recall that the variance of L is given by $\mathrm{var}(L) = p_{12} - p^2$ with $p_{12} = \Phi_\rho(D, D)$. Similarly, in the case of the shock model, the LHP approximation reads

$$\frac{\widehat{L}^{(n)}}{n\,e} \overset{\text{a.s.}}{\to} \Phi\left(\frac{\widehat{D}/W_L - \sqrt{\rho}\,\widetilde{Y}}{\sqrt{1-\rho}} \right) =: \widehat{L}\,, \qquad n \to \infty\,,$$

where now $\widehat{D} = t_{\nu_L}^{-1}(p)$. The variance is given by $\mathrm{var}(\widehat{L}) = \widehat{p}_{12} - p^2$ with $\widehat{p}_{12} = t_{\nu_L;\rho}(\widehat{D}, \widehat{D})$.

We now apply the one-factor approach to market risk so that both market and credit risk are systematically described by one and the same single factor $\widetilde{Y}$. To achieve this, we use (4.40) and

$$\tilde{\eta} := \frac{1}{\sqrt{1-\tilde{\gamma}^2}} \left[\sum_{k=1}^{K} \left(\gamma_k - \frac{\tilde{\gamma}}{\sqrt{\rho}}\,\beta_k \right) Y_k + \sqrt{1 - \sum_{k=1}^{K} \gamma_k^2}\,\eta \right]$$

with

$$\tilde{\gamma} := \frac{1}{\sqrt{\rho}} \sum_{k=1}^{K} \beta_k \gamma_k\,. \tag{4.43}$$

[2]Actually, there are less restrictive conditions for the exposures e_i and the individual default variables L_i under which the LHP approximation still holds, see e.g. in Bluhm et al. [8], Section 2.5.1.

Then, we obtain the formal identities

$$Z = -\sigma \left(\tilde{\gamma}\, \widetilde{Y} + \sqrt{1-\tilde{\gamma}^2}\, \tilde{\eta} \right) \tag{4.44}$$

and

$$\widehat{Z} = -\sigma\, W_Z \left(\tilde{\gamma}\, \widetilde{Y} + \sqrt{1-\tilde{\gamma}^2}\, \tilde{\eta} \right) \tag{4.45}$$

for the normal factor model (4.24) and for the shock model (4.25), respectively. In both cases, $\tilde{\eta}$ is standard normally distributed and independent of $\widetilde{Y}$. Moreover, Z in (4.44) is normally distributed with zero mean and variance $\mathrm{var}(Z) = \sigma^2$, whereas $\widehat{Z}$ in (4.45) follows a t-distribution with ν_Z degrees of freedom.

While the one-factor weight $\sqrt{\rho}$ for the credit portfolio depends only on the β_k, the one-factor weight $\tilde{\gamma}$ for market risk given by (4.43) is a function of $\beta_k \gamma_k$. In particular, in order to obtain non-vanishing systematic market risk within the one-factor model, both risk types have to share at least one common factor.

One-Factor Inter-Risk Correlation

The calculations of Section 4.3.3 easily apply to the case of the joint one-factor model of credit and market risk and the results regarding inter-risk correlation simplify considerably. We start with the normal one factor model.

Normal Factor Model Approach. Instead of (4.27) we now obtain

$$\mathrm{corr}(L^{(n)}_{\mathrm{hom}}, Z) = \frac{\sqrt{n}\, r\, e^{-D^2/2}}{\sqrt{2\pi}\sqrt{p_{12}(n-1) + p(1-np)}},$$

where $D = \Phi^{-1}(p)$ is the default point, $p_{12} = \Phi_\rho(D, D)$ is the joint default probability within the homogenous portfolio, and $r = \mathrm{corr}(Z, A_i) = \sqrt{\rho}\,\tilde{\gamma} = \sum_{k=1}^K \beta_k \gamma_k$. If the credit portfolio is not only homogenous but also very large, we arrive at the following LHP approximation for the inter-risk correlation between the credit portfolio loss (4.42) and market risk P/L (4.44) in the limit $n \to \infty$:

$$\mathrm{corr}(L, Z) = \frac{r\, e^{-D^2/2}}{\sqrt{2\pi(p_{12} - p^2)}}\,. \tag{4.46}$$

Given the uniform asset correlation $\rho = \sum_{k=1}^K \beta_k^2$ of a homogenous credit portfolio, it follows from (4.43) that $|\tilde{\gamma}| \le 1$, and thus $|r| \le \sqrt{\rho}$, implying the bounds

$$|\mathrm{corr}(L, Z)| \le \frac{\sqrt{\rho}\, e^{-D^2/2}}{\sqrt{2\pi(p_{12} - p^2)}} =: \psi(p, \rho)\,. \tag{4.47}$$

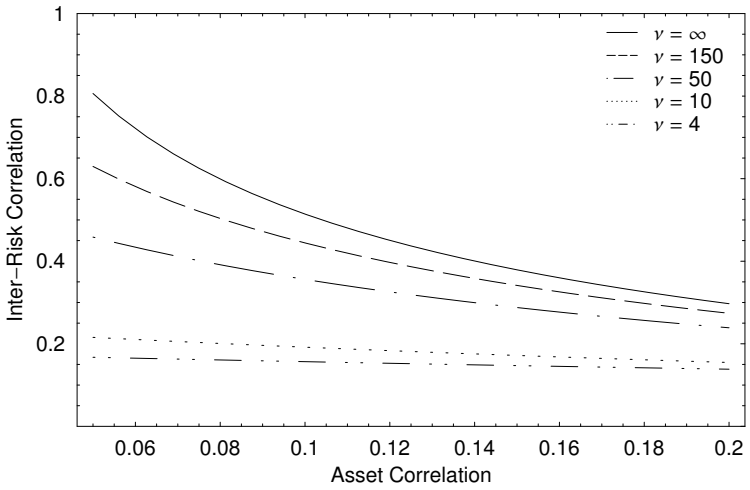

Figure 4.3.14. *LHP approximations of inter-risk correlation for $r = 0.2$ as a function of the asset correlation ρ according to the normal factor model (equation* (4.46)*) and the common shock model (equation* (4.49)*). The average default probability is assumed to be $p = 0.002$.*

Shock Model Approach. The LHP approximation for inter-risk correlation in the case of the independent shock model of Definition 4.3.8 (1) yields

$$\operatorname{corr}(\widehat{L}, \widehat{Z}) = \sqrt{\frac{\nu_Z - 2}{2}} \, \frac{\Gamma\left(\frac{\nu_Z - 1}{2}\right)}{\Gamma\left(\frac{\nu_Z}{2}\right)} \, \frac{r \left(1 + \frac{\widehat{D}^2}{\nu_L}\right)^{-\frac{\nu_L}{2}}}{\sqrt{2\,\pi\,(\widehat{p}_{12} - p^2)}}\,, \tag{4.48}$$

whereas for the common shock model of Definition 4.3.8 (2) we obtain

$$\operatorname{corr}(\widehat{L}, \widehat{Z}) = \sqrt{\frac{\nu - 2}{2}} \, \frac{\Gamma\left(\frac{\nu - 1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \, \frac{r \left(1 + \frac{\widehat{D}^2}{\nu}\right)^{\frac{1-\nu}{2}}}{\sqrt{2\,\pi\,(\widehat{p}_{12} - p^2)}}\,, \tag{4.49}$$

where $\widehat{D} = t_\nu^{-1}(p)$ is the default point, $\widehat{p}_{12} = t_{\nu;\rho}(\widehat{D}, \widehat{D})$ is the joint default probability within the homogenous portfolio, and $r = \operatorname{corr}(\widehat{Z}, \widehat{A}_i) = \sqrt{\rho}\,\tilde{\gamma}$. Similarly as for the LHP approximation of the normal factor model, bounds for the inter-risk correlation can be obtained from (4.48) and (4.49) together with $|r| \leq \sqrt{\rho}$. In the special case that $\nu_L = \nu_Z = \nu$ it follows by comparison of (4.48) and (4.49) that the assumption of one single common shock increases inter-risk correlation by a factor of $\sqrt{1 + \frac{\widehat{D}^2}{\nu}}$ compared to the independent shock model. For typical values of p this factor lies in a range of about 1.0–2.0.

To find out how global (macroeconomic) shocks affect inter-risk correlation, let us contrast the LHP approximations (4.48) and (4.49) of the shock models with that of the normal factor model (4.46). For this purpose, Table 4.3.16 as well as Figures 4.3.14

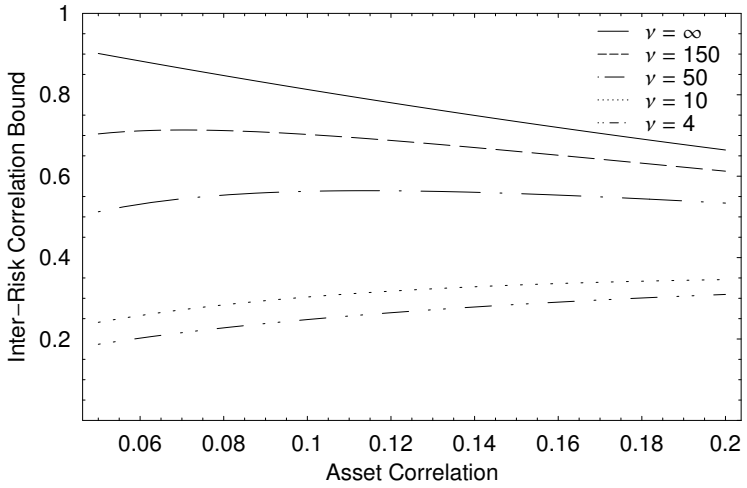

Figure 4.3.15. *LHP approximations of the inter-risk correlation bound as a function of the asset correlation ρ according to the normal factor model (equation* (4.47)*) and the common shock model (equation* (4.49) *with $r = \sqrt{\rho}$). The average default probability is assumed to be $p = 0.002$.*

and 4.3.15 compare the inter-risk correlation and its upper bound for the common shock model with the outcome of the normal factor model. One can see that the common shock model yields—particularly for small asset correlations—lower inter-risk correlations and bounds than the normal factor model. In the case of the independent shock model, the spread between the normal inter-risk correlation and the shocked inter-risk correlation would be even higher.

Needless to say, the one-factor asset correlation ρ is a popular parameter in the context of credit portfolio modelling. It is often used as a "single-number measure" to evaluate the average dependence structures between different counterparties of a credit portfolio. Moreover, it plays an important role also in the calculation formula for regulatory capital charges according to the *internal-ratings-based* (IRB) approach of Basel II [5]. Equations (4.47)-(4.49) now show that ρ is a very important parameter also beyond credit risk itself as it determines its maximum possible inter-risk correlation with another risk type, here market risk; see again Figure 4.3.15 where inter-risk correlation bounds are plotted as a function of ρ[3].

Similarly, for a fixed uniform asset correlation ρ, inter-risk correlation and its bound depend on the average default probability p and thus on the average rating of the credit portfolio. This is depicted in Figure 4.3.17 where LHP inter-risk correlation as well

[3]Note that ρ enters $|\mathrm{corr}(L, Z)|$ not only directly by $\sqrt{\rho}$ but also indirectly via the joint default probability $p_{12} = \Phi_\rho(\Phi^{-1}(p), \Phi^{-1}(p))$. This implies that $|\mathrm{corr}(L, Z)| \neq 0$ for $\rho \to 0$.

ρ	Normal Model	Common Shock Model		
	$\nu = \infty$	$\nu = 4$	$\nu = 10$	$\nu = 50$
		$p = 0.002$		
5 %	0.81 (0.90)	0.17 (0.19)	0.22 (0.24)	0.46 (0.51)
10 %	0.51 (0.81)	0.16 (0.25)	0.19 (0.30)	0.36 (0.56)
15 %	0.38 (0.73)	0.15 (0.28)	0.17 (0.33)	0.29 (0.56)
20 %	0.30 (0.66)	0.14 (0.31)	0.15 (0.35)	0.24 (0.53)
		$p = 0.02$		
5 %	0.85 (0.95)	0.27 (0.31)	0.37 (0.42)	0.62 (0.70)
10 %	0.57 (0.90)	0.25 (0.40)	0.33 (0.52)	0.48 (0.76)
15 %	0.44 (0.86)	0.24 (0.46)	0.29 (0.57)	0.39 (0.76)
20 %	0.37 (0.82)	0.22 (0.50)	0.27 (0.59)	0.33 (0.75)

Table 4.3.16. *LHP approximation for inter-risk correlation for the normal factor model* (4.46) *and the common shock model* (4.49) *using* $r = 0.2$ *but different values for* p *and average asset correlation* ρ*. The values in brackets indicate upper inter-risk correlation bounds for which* $r = \sqrt{\rho}$*.*

as the corresponding upper bounds are plotted as a function of the average portfolio rating. One can see that an improvement of the average portfolio rating structure decreases inter-risk correlation (as well as its bounds) and thus tends to result in a lower volatility of the total portfolio of market and credit.

A Moment Estimator for the Inter-Risk Correlation Bound. Even if the actual credit portfolio is not homogenous, the derived LHP approximation provides us with a useful estimator for approximating the upper inter-risk correlation bound.

Let us consider the normal factor model and so expression (4.47). For any credit loss distribution—obtained for instance by Monte Carlo simulation— estimators $\hat{p}$ and $\hat{\rho}$ for p and ρ, respectively, can be (numerically) obtained via moment matching. In doing so, we compare the empirical expected loss μ and the empirical variance ς^2 derived from the simulated credit portfolio with the corresponding figures of an LHP approximation, i.e. average default probability p and variance $\mathrm{var}(L)$. Thus we require that

$$\mu \;=\; e_{\mathrm{tot}}\,\hat{p} \tag{4.50}$$

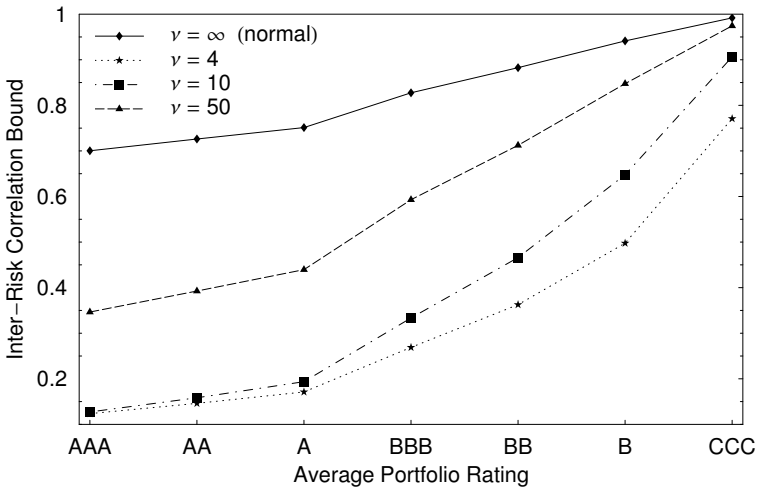

Figure 4.3.17. *LHP approximations of inter-risk correlation bound as a function of the average portfolio rating according to the normal factor model (equation* (4.47)*) and the common shock model (equation* (4.49) *with* $r = \sqrt{\rho}$*). The average asset correlation is assumed to be* $\rho = 10$ %*.*

and

$$\begin{aligned} \varsigma^2 &= e_{\text{tot}}^2 \, (p_{12} - \hat{p}^2) \\ &= e_{\text{tot}}^2 \left[\Phi_{\hat{\rho}}(\Phi^{-1}(\hat{p}), \Phi^{-1}(\hat{p})) - \hat{p}^2 \right], \end{aligned} \tag{4.51}$$

where e_{tot} denotes the total credit exposure. From (4.47) we then obtain the following moment estimator for the upper inter-risk correlation bound:

$$\hat{\psi}(\hat{p}, \hat{\rho}) = \frac{e_{\text{tot}}}{\varsigma} \, \frac{\sqrt{\hat{\rho}} \, \exp\left[-\frac{1}{2} \big(\Phi^{-1}(\hat{p}) \big)^2 \right]}{\sqrt{2\pi}} \, . \tag{4.52}$$

4.3.5 Estimation of the Gaussian Copula Parameter

We now concentrate on the joint normal factor model of Definition 4.3.6 and show how the inter-risk correlation between both risk types is related to the copula parameter of the Gaussian copula.

A general problem with copula techniques is which copula to choose and how to calibrate the it. A remarkable feature of the joint normal factor model for credit and market risk is that it can easily be interpreted as a Gaussian coupling model between market and credit risk.

Proposition 4.3.18. [Normal one-factor model and Gaussian copula] *Consider the joint normal one-factor model for credit and market risk, i.e. we consider* (4.41) *and*

$\tilde{\gamma}$	0.0	0.2	0.4	0.6	0.8	1.0
$\mathrm{corr}(L, Z)$	0.0	0.15	0.29	0.44	0.59	0.73

Table 4.3.19. *Relation between inter-risk correlation* $\mathrm{corr}(L, Z)$ *and Gaussian copula parameter* $\tilde{\gamma}$ *according to* (4.53) *for* $p = 0.002$ *and* $\rho = 15$ %.

(4.44) *where all idiosyncratic factors* ε_i *of the credit model are independent the idiosyncratic factors* $\tilde{\eta}$ *of market risk. Then,*
(1) *both risk types are coupled by a Gaussian copula with parameter* $\tilde{\gamma}$ *given by* (4.43)*;*
(2) *the copula parameter* $\tilde{\gamma}$ *and the inter-risk correlation* $\mathrm{corr}(L, Z)$ *are related by*

$$\tilde{\gamma} = \frac{\mathrm{corr}(L, Z)}{\psi} \tag{4.53}$$

where ψ *is the LHP approximation* (4.47) *for the inter-risk correlation bound.*

Proof. (1) An important characteristic of copulas is their invariance under monotonously increasing transformations. Since the portfolio loss L in the LHP approximation as given by (4.42) is a monotonously increasing function of $-\widetilde{Y}$, it follows that L and Z have the same copula as $-\widetilde{Y}$ and Z. For the latter we know from the one-factor representation of market risk (4.44) that they are bivariate normally distributed with correlation

$$\mathrm{corr}(-\widetilde{Y}, Z) = \tilde{\gamma}\,.$$

Hence, also L and Z are linked by a Gaussian copula with correlation parameter $\tilde{\gamma}$.
(2) This follows directly from (4.46) and (4.47) together with $r = \sqrt{\rho}\,\tilde{\gamma}$. □

It follows from (4.53) that the absolute value of the inter correlation between credit and market risk is always below the absolute value of the copula parameter $\tilde{\gamma}$. Furthermore, maximum inter-risk correlation corresponds to $\tilde{\gamma} = 1$ for which market risk is completely determined by one single risk factor without having any idiosyncratic component, cf. equation (4.44). A numerical example for (4.53) is given in Table 4.3.19.

Particularly important for practical applications is the question of how the Gaussian copula parameter can be estimated for general credit portfolios. Note that in this case Proposition 4.3.18 (1) is not directly applicable because β_k in (4.43) is only defined for a homogenous portfolio. However, we can extend the LHP approximation for a credit portfolio, which we have used to construct the estimator $\hat{\psi}$ for the inter-risk correlation bound given by (4.52), to the joint one-factor risk model of credit *and* market risk by matching the inter-risk correlations. If market and credit risk are described by the

joint normal factor model of Definition 4.3.6, we can calculate inter-risk correlation by Theorem 4.3.9 and compare it to the result in the case of the LHP approximation, i.e. expression 4.46. Then, using Proposition 4.3.18 (2) we arrive at the following general estimator for the copula parameter $\tilde{\gamma}$,

$$\widehat{\tilde{\gamma}}_1 = \frac{\widehat{\mathrm{corr}}(L, Z)}{\hat{\psi}} \tag{4.54}$$

where

$$\widehat{\mathrm{corr}}(L, Z) = \mathrm{corr}(L^{(n)}, Z)$$

with $\mathrm{corr}(L^{(n)}, Z)$ calculated as in (4.27).

An alternative estimator for $\tilde{\gamma}$ can be constructed by applying the right-hand side of (4.53) directly to a non-homogenous portfolio without introducing a one-factor approximation before. In this case it follows together with (4.27) and (4.31) that

$$\widehat{\tilde{\gamma}}_2 = \frac{\sum_{i=1}^{n} \sum_{k=1}^{K} \beta_{ik}\, \gamma_k\, e_i \exp\left(-\frac{1}{2} D_i^2\right)}{\sum_{i=1}^{n} e_i \sqrt{\sum_{k=1}^{K} \beta_{ik}^2}\, \exp\left(-\frac{1}{2} D_i^2\right)}\,. \tag{4.55}$$

Bibliography

[1] Aue, F., and Kalkbrenner, M. 2006. LDA at work: Deutsche Banks approach to quantifying operational risk. *Journal of Operational Risk* 1(**4**): 49–93.

[2] BCBS, IOSCO, IAIS (The Joint Forum). 2003. Trends in Risk Integration and Aggregation.

[3] Bee, M. 2005. Copula-based multivariate models with applications to risk management and insurance. Working Paper, University of Trento. Available at: www.gloriamundi.org.

[4] Barndorff-Nielsen, O., and Lindner, A. 2006. Lévy copulas: dynamics and transforms of Upsilon type. *Scand. J. Statistics* **34**: 298–316.

[5] Basel Committee on Banking Supervision. 2004. International Convergence of Capital Measurement and Capital Standards. Basel.

[6] Basel Committee on Banking Supervision. 2006. Observed range of practice in key elements of Advanced Measurement Approaches (AMA). Basel.

[7] Bingham, N.H., Goldie, C.M., and Teugels, J.L. 1987. *Regular Variation.* Cambridge: Cambridge University Press.

[8] Bluhm, C., Overbeck, L., and Wagner, C. 2003. *An Introduction to Credit Risk Modelling.* Boca Raton, Fla.: Chapman & Hall/CRC.

[9] Böcker, K. 2006. Operational Risk: Analytical results when high severity losses follow a generalized Pareto distribution (GPD)–A Note–. *Journal of Risk* **8**(4): 117–120.

[10] Böcker, K. 2008. Aggregation by Risk Type and Inter Risk Correlation. In: Resti, A. (ed.) *Pillar II in the New Basel Accord: the Challenge of the Internal Capital Adequacy Assessment Process.* London: Risk Books. In press.

[11] Böcker, K. 2008. Modelling and Measuring Business Risk. In: Resti, A. (ed.) *Pillar II in the New Basel Accord: the Challenge of the Internal Capital Adequacy Assessment Process.* London: Risk Books. In press.

[12] Böcker, K., and Hillebrand, M. 2007. Interaction of market and credit risk: an analysis of inter-risk correlation and risk aggregation. *PRMIA "2008 Enterprise Risk Management Symposium Award for Best Paper Challenging new Frontiers in Risk Management."* Submitted for publication. Available at: www.ma.tum.de/stat/.

[13] Böcker, K., and Klüppelberg, C. 2005. Operational VaR: a closed-form approximation. *RISK Magazine*: 90–93.

[14] Böcker, K., and Klüppelberg, C. 2007. Multivariate models for operational risk. *PRMIA "2007 Enterprise Risk Management Symposium Award for Best Paper Challenging new Frontiers in Risk Management."* Submitted for publication. Available at: www.ma.tum.de/stat/.

[15] Böcker, K., and Klüppelberg, C. 2007. Multivariate operational risk: dependence modelling with Lévy copulas. *2007 ERM Symposium Online Monograph, Society of Actuaries.* Available at: www.soa.org, and *Joint Risk Management Section Newsletter* **11**: 26–31.

[16] Böcker, K., and Klüppelberg, C. 2008. Economic Capital Modelling and Basel II Compliance in the Banking Industry. In: Jäger, W., Krebs, H.-J. (ed.) *Mathematics – Key Technology for the Future.* Berlin: Springer.

[17] Böcker, K., and Klüppelberg, C. 2008. Modelling and measuring multivariate operational risk with Lévy copulas. *Journal of Operational Risk* 3(**2**): 3–27.

[18] Böcker, K., and Klüppelberg, C. 2009. First Order Approximations to Operational Risk – Dependence and Consequences. In: Gregoriou, G.N. (ed.). *Operational Risk Towards Basel III, Best Practices and Issues in Modeling, Management and Regulation.* New York: Wiley. In press.

[19] Böcker, K., and Spielberg, H. 2005. Risikoaggregation mit Kopulas. *Die Bank* **8**: 56–59.

[20] Böcker, K., and Sprittulla, J. 2006. Operational VAR: meaningful means. *RISK Magazine*: 96–98.

[21] Bregman, Y., and Klüppelberg, C. 2005. Ruin estimation in multivariate models with Clayton dependence structure. *Scand. Act. J.* **6**: 462–480.

[22] Chavez-Demoulin, V., Embrechts, P., and Nešlehová, J. 2005. Quantitative models for operational risk: extremes, dependence and aggregation. *Journal of Banking and Finance* **30**: 2635–2658.

[23] Cherubini, U., Luciano, E., and Vecchiato, W. .2004. *Copula methods in finance.* Chichester: Wiley.

[24] Cont, R., and Tankov, P. 2004. *Financial Modelling With Jump Processes.* Boca Raton, Fla.: Chapman & Hall/CRC.

[25] Committee Of European Banking Supervisors (CEBS). 2004. The Application Of The Supervisory Review Process Under Pillar 2. London.

[26] Committee of European Banking Supervisors (CEBS). 2008. Consultation paper on technical aspects of diversification under Pillar 2. London.

[27] Devlin, S.J., Gnanadesikan, R., and Kettering, J.R. 1981. Robust estimation of dispersion matrices and principal components. *J. Amer. Statist. Assoc.* **76**: 354–362.

[28] Dimakos, X. K., and Aas, K. 2004. Integrated Risk Modelling. *Statistical Modelling* **4**: 265–277.

[29] Embrechts, P., Klüppelberg, C., and Mikosch, T. 1997. *Modelling Extremal Events for Insurance and Finance.* Berlin: Springer.

[30] Embrechts, P., Lindskog, F., and McNeil, A. 2001. Modelling Dependence with Copulas and Applications to Risk Management. In: Rachev, S. (ed.) *Handbook of Heavy Tailed Distributions in Finance.* Amsterdam: Elsevier: 329–384.

[31] Fahrmeir, L., Hamerle, A., and Tutz, G. 1996. *Multivariate statistische Verfahren.* Berlin: De Gruyter.

[32] Frachot, A., Roncalli, T., and Salomon, E. 2004. The correlation problem in operational risk. Working Paper, Credit Agricole. Available at: www.gloriamundi.org.

[33] Goedhart, M., Koller, T., and Wessels, D. 2005. *Valuation: Measuring and Managing the Value of Companies.* Hoboken NJ: Wiley.

[34] Gokhale, D.V., and James Press, S. 1982. Assessment of a Prior Distribution for the Correlation Coefficient in a Bivariate Normal Distribution. *Journal of the Royal Statistical Society A.* 145(**2**): 237–249.

[35] Higham, N.J. 2002. Computing the Nearest Correlation Matrix – a Problem From Finance. *IMA Journal of Numerical Analysis.* 22(**3**) 329–343.

[36] The Institute of the Chief Risk Officers (CROs)and Chief Risk Officers Forum (CRO Forum). 2007. Insights from the joint IFRI/CRO Forum survey on Economic Capital Practice and application.

[37] Jenkinson, D. 2005. The Elicitation of Probabilities–A Review of the Statistical Literature. Working Paper, Bayesian Elicitation of Experts' Probabilities (BEEP), University of Sheffield. Available at: www.shef.ac.uk/beep/publications.html.

[38] Joe, H. 1997. *Multivariate Models and Dependence Concepts.* London: Chapman & Hall.

[39] Kallsen, J., and Tankov, P. 2006. Characterization of dependence of multivariate Lévy processes using Lévy copulas. *J. Multiv. Anal.* **97**: 1551–1572.

[40] Klüppelberg, C., and Resnick, R.I. 2007. The Pareto copula, aggregation of risks, and the emperor's socks. *J. Appl. Probab.* 45(**1**): 67–84.

[41] Klugman, S., Panjer, H., and Willmot, G. 2004. *Loss Models - From Data to Decisions.* Hoboken NJ: Wiley.

[42] Klüppelberg, C., Lindner, A., and Maller, R. 2005. Continuous time volatility modelling: COGARCH versus Ornstein-Uhlenbeck models. In: Kabanov, Y., Lipster, R., and Stoyanov, J. (ed.) *From Stochastic Calculus to Mathematical Finance. The Shiryaev Festschrift.* Berlin: Springer: 393–419.

[43] Klüppelberg, C., and Mikosch, T. 1997. Large deviations of heavy-tailed random sums with applications in insurance and finance. *J. Appl. Prob.* **34**: 293–308.

[44] Kostadinov, K. 2006. Tail approximation for credit risk portfolios with heavy-tailed risk factors. *Journal of Risk* 8(**2**): 81–107.

[45] Lindskog, F. 2000. Linear correlation estimation. Research Paper, ETH Zurich. Available at: www.risklab.ch/Papers.html

[46] Lindskog, F., and McNeil, A. 2003. Common Poisson shock models: application to insurance and credit risk modelling. *ASTIN Bulletin* **33**: 209–238.

[47] Lindskog, F., McNeil, A., and Schmock, U. 2003 Kendall's Tau for Elliptical Distributions. In: Bol, G. et al. (ed.) *Credit Risk-measurement, evaluation and management.* Heidelberg: Physica-Verlag: 149–156.

[48] Loéve, M. 1978. *Probability Theory, Vol. I.* Berlin: Springer.

[49] Matten, C. 1996. *Managing Bank Capital, Capital Allocation and Performance Measurement.* Chichester: Wiley.

[50] McNeil, A.J., Frey, R., and Embrechts, P. 2005. *Quantitative Risk Management– Concepts, Techniques and Tools.* Princeton NJ: Princeton University Press.

[51] Mignola, G., and Ugoccioni, R. 2006. Sources of uncertainty in modelling operational risk losses. *Journal of Operational Risk* 1(**2**): 33–50.

[52] Nelsen, R. 1999. *An Introduction to Copulas*, New York: Springer.

[53] Morrison, D.J., Quella, J.A., and Slywotzky, A.J. 1999. Countering strategic risk with pattern thinking. *Mercer Management Journal* **11**. Available at: http://mysite.verizon.net/clarkm59/MMJ111CounteringRisk.pdf.

[54] Moscadelli, M. 2004. The modelling of operational risk: experience with the analysis of the data collected by the Basel Committee. Banca D'Italia, Termini di discussione No. 517.

[55] Nguyen, M.-T., and Ottmann, M. 2006. The fat tail. *OpRisk & Compliance.* March: 42–45.

[56] O'Hagan, A. et al. 2006. *Uncertain Judgements: Eliciting Experts' Probabilities.* Chichester: Wiley.

[57] Powosjowski, M.R., Reynolds, D., and Tuenter, J.H. 2002. Dependent events and operational risk. *Algo Research Quaterly* 5(**2**): 65–73.

[58] Pratt, S.P., Reilly, R.F., and Schweihs, R.P. 2000. *Valuing a Business: The Analysis and Appraisal of Closely Held Companies.* New York: McGraw-Hill.

[59] Rebonato, R., and Jäckel, P. 1999. The most general methodology to create a valid correlation matrix for risk management and option pricing purposes. *Journal of Risk* 2(**2**): 17–24.

[60] Resnick, S.I. 1987. *Extreme Values, Regular Variation, and Point Processes.* New York: Springer.

[61] Resnick, S.I. .2006. *Heavy-Tail Phenomena. Probabilistic and Statistical Modeling.* New York: Springer.

[62] Rootzén, H., and Klüppelberg, C. 1999. A single number can't hedge against economic catastrophes. *Ambio* 28(**6**): 550–555.

[63] Rosenberg, J.V., and Schuermann, T. 2004. A General Approach to Integrated Risk Management with Skewed, Fat-Tailed Risk. Working Paper, Federal Reserve Bank of New York Staff Reports, **185**.

[64] Saita, F. 2004. Risk Capital Aggregation: The Risk Manager's Perspective. EFMA Basel Meetings Paper. Available at: SSRN: http://ssrn.com/abstract=496684.

[65] Saita, F. 2007. *Value at Risk and Bank Capital Management.* San Diego: Elsevier.

[66] Shreve, E.S. 2004. *Stochastic Calculus for Finance II: Continuous-Time Models.* New York: Springer.

[67] Tversky, A., and Kahneman, D. 1974. Judgment under Uncertainty: Heuristics and Biases. *Science* (**185**): 1124–1131.

[68] Tversky, A., and Kahneman, D. 1983. Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* (**90**): 293-315.

[69] Vasicek, O. 2002. Loan portfolio value. *RISK Magazine.* December: 160–162.

[70] Ward, L., and Lee, D. 2002. Practical application of risk-adjusted return on capital framework, Dynamic Financial Analysis Discussion Paper. CAS Forum.